



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO.

Halisson Alberdan Cavalcanti Cardoso

Modelagem e Geometrização em PGNA

Recife

2021

Halisson Alberdan Cavalcanti Cardoso

Modelagem e Geometrização em PGNAA

Trabalho apresentado ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Ciência da Computação.

Área de Concentração: Mídia e Interação.

Orientador: Silvio de Barros Melo

Coorientador: Ricardo Martins de Abreu e Silva

Recife

2021

Catálogo na fonte
Bibliotecário Cristiano Cosme S. dos Anjos, CRB4-2290

C268m Cardoso, Halisson Alberdan Cavalcanti
 Modelagem e Geometrização em PGNAA/ Halisson Alberdan Cavalcanti
 Cardoso. – 2021.
 306 f.: il., fig., tab.

 Orientador: Silvio de Barros Melo.
 Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da
 Computação, Recife, 2021.
 Inclui referências e apêndices.

 1. Mídia e Interação. 2. GRASP. 3. PGNAA. 4.MCLLS. I. Melo, Stênio Silvio
 de Barros (orientador). II. Título

 006.7 CDD (23. ed.) UFPE - CCEN 2021 – 143

Halisson Alberdan Cavalcanti Cardoso

“Modelagem e Geometrização em PGNA”

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação.

Aprovado em: 22/04/2021.

Orientador: Prof. Dr. Silvio de Barros Melo

BANCA EXAMINADORA

Prof. Dr. Nivan Roberto Ferreira Júnior
Centro de Informática / UFPE

Prof. Dr. Carlos Costa Dantas
Departamento de Energia Nuclear / UFPE

Prof. Dr. Emerson Alexandre de Oliveira Lima
Escola Politécnica de Pernambuco / UPE

Prof. Dr. Bruno Gomes da Costa
Instituto Federal do Sertão Pernambucano / Campus Petrolina

Prof. Dr. Marco José da Silva
Universidade Tecnológica Federal do Paraná / Campus Curitiba

Dedico esse trabalho a todos que dele participaram, em especial ao seu Severino Cardoso (In Memoriam) e dona Maria da Conceição (In Memoriam)

AGRADECIMENTOS

As pessoas a quem devo o esforço dispendido ao longo de todo o trabalho e em nenhum momento desistiram de mim, e sempre acreditaram que eu pudesse chegar aonde eu quisesse, desde que, para isso, eu não desistisse dos meus objetivos. Em memória aos meus pais, dedico todas as honras e os agradecimentos por ter conseguido chegar a esta etapa na qual me encontro. Como filho, não tenho como descrever com palavras os meus sentimentos por tudo o que fizeram. Tudo o que possa fazer será pouco frente à dedicação e empenho que dedicaram na minha formação. Meu muito obrigado, quero estar sempre com eles!. Também gostaria de deixar meus agradecimentos àquela a quem eu dispendo todos os meus esforços, minha filha, Doralice Prado Cardoso. Ela foi a motivação que fez levantar e sacudir a poeira em todos os momentos em que pensei em desistir.

Não se pode esquecer de mencionar aqueles a quem eu devo efetivamente agradecer pela operacionalização do presente trabalho: aos meus professores, Sílvia de Barros Melo e Ricardo Martins de Abreu e Silva. Essas foram as pessoas que acreditaram em mim e depositaram seu voto de confiança para que pudesse dar origem a este trabalho. A eles dedico cada quantum de energia dispensado na elaboração deste documento. Espero, um dia, fazer por alguém o que os meus professores/orientadores fizeram pela minha pessoa. Se não tivesse sido pelo suporte e conhecimento deles, este trabalho não teria tido esse momento de desfecho. Meu muito obrigado!

Agradecimentos especiais a Katianne Cristina da Silva Veríssimo, pelo suporte fornecido ao longo dos anos vinculados ao CIn, como estudante de pós graduação.

Agradecimentos especiais para Ilker Meric, professor da Universidade de Bergen, e para seu aluno, Ícaro Valgueiro Malta Moreira, pela cessão dos dados utilizados no presente trabalho.

Gostaria de deixar meus agradecimentos aos integrantes da banca. Alguns deles acompanham o trabalho desde a etapa de qualificação: professores Nivan Júnior, Carlos Dantas e Emerson Lima. Agradeço pela força na viabilização e engrandecimento do nosso trabalho. Não poderia esquecer os agradecimentos a Bruno Gomes da Costa e ao professor Marco José da Silva.

Gostaria, também, de agradecer a todos aqueles que forneceram indiretamente condições indispensáveis à conclusão do nosso trabalho, e que se tornaram amigos durante o período de vínculo junto ao CIn: Seu Roberto, Suilan, Seu Santos, Carlos, Tadeu, Marcia, Julio, Luiz Melo, Lília, Daniel, Socorro, Amanda, Gabriela, Passira, Clécio, Dijaílson, Gabriela, Amanda, Inaldo...

a todos que por aqui passaram e exerceram suas atividades funcionais com profissionalismo difícil de ser encontrado. Fica, também, minhas desculpas se esqueci de enumerar algumas pessoas dentre os nomes listados acima.

Gostaria de deixar meus agradecimentos aos professores que passaram pelas coordenações do curso durante meu período de estadia e que demonstraram toda presteza e paciência com minha pessoa durante esse período: professores Edna e Aluizio e, em especial, professor Ricardo Prudêncio. Meu muito obrigado.

Não esquecendo, claro, dos meus amigos da APG que tanto me ajudaram nessa jornada até a presente data: Wanessa Veridiana, Elisa, Paulo Henrique, Bruno Maciel, Silas Garrido, Jailton Cardoso, Giorbis, Késia, Sidartha Carvalho, entre outros que talvez eu não tenha aqui mencionado por ter, injustamente, esquecido no momento dos agradecimentos.

“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê” (SCHOPENHAUER, 2003, p. 9).

RESUMO

O *Prompt Gamma Neutron Analysis Activation* (PGNAA) é uma das técnicas utilizadas atualmente na análise de materiais. Essa técnica define gráficos de intensidade espectral como função de energia (canais) para os elementos dentro de uma amostra. Esses gráficos (bibliotecas) permitem a determinação de frações de contagem percentual, ou frações em peso, de cada um dos elementos. O *Monte Carlo Library Least Squares* (MCLLS), principal abordagem ao PGNAA, possibilita tratamento linear para um problema inerentemente não linear. O Monte Carlo (MC), módulo do MCLLS, simula o setup experimental, gerando as bibliotecas, enquanto que o *Library Least Square* - Biblioteca Mínimos Quadrados (LLS) determina as frações em peso dos elementos do objeto estudado. As principais dificuldades enfrentadas no MCLLS são: (1) instabilidades numéricas advindas da etapa do LLS; (2) sobredeterminação de sistemas de equações; (3) dependência linear entre as bibliotecas; (4) radiações provenientes do espalhamento da radiação gama (background); (5) elevados custos computacionais. Esta tese propõe otimização do LLS no intuito de enfrentar as questões descritas acima e, simultaneamente, determinar as frações em peso dos elementos de uma amostra, fazendo uso dos algoritmos heurísticos *Greedy Randomized Adaptive Search Procedure* (GRASP) e *Continuous Greedy Randomized Adaptive Search Procedure* (CGRASP). A abordagem é baseada no particionamento dos espectros, determinando regiões (partições) que contenham os máximos de contagem espectral das bibliotecas (picos). Essas partições são utilizadas para otimização do GRASP dos canais inseridos na solução do LLS, como também auxiliam na estimativa e otimização pelo CGRASP do constituinte potencialmente desconhecido pertencente ao espectro da amostra. Os resultados evidenciam: (1) particionamento eficiente das bibliotecas, determinando os picos; (2) evidências de boa precisão das frações em peso dos constituintes, em comparação com trabalhos publicados (precisões médias da ordem de 3,16%, contra 8.80% de outros métodos); (3) sucesso na aproximação e estimação do constituinte desconhecido (precisão da ordem de 4,25%) inserido no espectro total. O nosso método mostrou-se promissor nas melhorias do método dos mínimos quadrados para as frações de contagem percentual dos elementos inseridos em uma amostra desconhecida, bem como evidenciou as potencialidades do particionamento dos dados, em contraposição às abordagens de particionamento das bibliotecas amplamente divulgadas no estado da arte.

Palavras-chaves: GRASP; PGNAA; MCLLS; mínimo quadrados; condicionamento.

ABSTRACT

Prompt Gamma Neutron Analysis Activation (PGNAA) is one of the main techniques currently used in materials analysis. This technique defines spectral intensity graphs as a function of energy (channels) for the elements within a sample. These graphs (libraries) allow the determination of fractions of percentage counting, or fractions by weight, of each element. The Monte Carlo Library Least Squares (MCLLS), the main approach to PGNAA, provides linear treatment for an inherently non-linear problem. Monte Carlo (MC), MCLLS module, simulates the experimental setup, generating the libraries while The Library Least Square module - Least Square Library (LLS) determines the weight fractions of the elements of the studied object. The main difficulties faced in MCLLS are (1) numerical instabilities arising from the LLS stage; (2) overdetermination of systems of equations; (3) dependency linear between libraries; (4) radiation from gamma radiation scattering (background); (5) high computational costs. This thesis proposes an optimization of LLS to face the main issues described above and at the same time properly determine the weight fractions of the elements included in the sample, making use of the heuristic algorithms Greedy Randomized Adaptive Search Procedure (GRASP) and Continuous Greedy Randomized Adaptive Search Procedure (CGRASP). The approach is based on the partitioning of the spectra, determining regions (partitions) that contain the maximum spectral counts of the libraries (peaks). These partitions are used to optimize the GRASP of the channels inserted in the LLS solution, as well as helping to estimate and optimize by CGRASP potentially unknown constituents belonging to the sample spectrum. The results show (1) efficient partitioning of the libraries, determining the peaks; (2) evidence of good precision of the constituent weight fractions, compared to published works (mean precision of the order of 3.16%, against 8.80% for other methods); (3) success in the approximation and estimation of the unknown constituent (accuracy of the order of 4.25% in the estimation of fractions of percentage counts of the omitted element) inserted in the total spectrum but estimated in the implemented algorithm. Our method showed promise in improving the least-squares method for fractions of percentage counts of elements inserted in an unknown sample. It evidenced the potential of data partitioning, in contrast to the partitioning approaches of libraries widely publicized in state of the art.

Keywords: GRASP; PGNAA; MCLLS; least-squares; conditioning.

LISTA DE FIGURAS

Figura 1 – Ilustração do decaimento. Absorção do nêutron induz à emissão de raios gama. Raios gama atrasados podem ser emitidos, antes de o núcleo retornar à estabilidade. (SÖLLRADL, 2014)	53
Figura 2 – Distribuição amostral das soluções nas iterações do GRASP	73
Figura 3 – Espectros da Base 1	92
Figura 4 – Sumarização da modelagem do Algoritmo GRASP ao domínio do problema. Escolhido um n -ésimo isótopo arbitrário, representado por um vetor de números. Esse vetor é omitido antes de os dados serem fornecidos como entrada ao algoritmo descrito na presente figura. São descritos como passos essenciais do método: (1) determinação dos picos e das partições aos quais pertencem, para fins de representação das bibliotecas conhecidas e fornecidas inicialmente (Seção 4.1); (2) determinação da partição que contém o maior resíduo, para fins de representação do elemento omitido, mas inserido no espectro total (Seção 4.2); (3) Construção de uma solução (Seção 4.4), definindo os parâmetros de interesse (<i>parâmetros de construção</i>) das partições e o intervalo de valores em torno do qual irá flutuar as estimativas do elemento omitido (estimativa pelo Algoritmo de Geometrização, Seção 4.3). A partir dessa etapa, destaca-se que há o início da modelagem do problema via Algoritmo GRASP; (4) Ajuste do elemento omitido, via Algoritmo CGRASP, nos compartimentos das bibliotecas fornecidas inicialmente (Seção 4.5); (5) Refinamento da solução construída, pela execução do passo de busca na partição de maior resíduo (Seção 4.5.2); (6) segue-se a mais uma execução do CGRASP; (7) execução da linkagem entre soluções, na partição de maior resíduo (Seção 4.7); (8) finalizando com mais um ajuste do CGRASP.	93
Figura 5 – Construção de vetores de compostos	96
Figura 6 – Sumarização da determinação dos picos e da partição de maior resíduo. Sumarização das etapas de determinação dos picos (Seção 4.1) das bibliotecas fornecidas inicialmente, e da partição de maior resíduo.	98
Figura 7 – Método de estimativa da altura nas partições.	100
Figura 8 – Sumarização da etapa construtiva.	102

Figura 9 – Sumariação do Algoritmo CGRASP	104
Figura 10 – Sumarização dos procedimentos tomados durante a execução do algoritmo de construção do CGRASP	106
Figura 11 – Processo construtivo do Algoritmo CGRASP, no ajuste da biblioteca omitida nas partições representativas das bibliotecas fornecidas inicialmente.	107
Figura 12 – Processo de busca do Algoritmo CGRASP, no ajuste da biblioteca omitida nas partições representativas das bibliotecas fornecidas inicialmente.	109
Figura 13 – Sumarização dos procedimentos tomados durante a execução do Algoritmo de busca do CGRASP	110
Figura 14 – Variação nos valores de altura para o composto omitido na partição de maior resíduo, definindo vizinhos da solução corrente.	112
Figura 15 – Busca por soluções vizinhas à solução atual.	113
Figura 16 – Sumarização do processo de busca por soluções melhores que a solução corrente.	114
Figura 17 – Sumarização da linkagem entre duas soluções.	116
Figura 18 – Procedimentos executados na linkagem de canais entre as soluções <i>start</i> e <i>target</i>	118
Figura 19 – Variação nos valores de altura para o composto extraído do espectro desconhecido.	120
Figura 20 – Gráfico apresentando as componentes dos vetores de compostos, resíduos e componentes do vetor espectral total, como função da iteração. São apresentados os resultados para a instância 1 (Figura 20a) e instância 2 (Figura 20b). Os componentes do vetor espectral total são apresentados (variável <i>tot</i> – 5), além dos valores de resíduos nas partições (variável <i>res</i> – 1).124	124
Figura 21 – Frações de Contagem percentuais no particionamento dos compostos, instância 1. As taxas de contagem espectral para o <i>Am</i> são as taxas reais. Valores em preto (○), vermelho(+), azul (✱) e verde (✱) indicam, respectivamente, as quantidades de <i>Am</i> , <i>Co</i> , <i>Cs</i> e <i>Na</i> . Os valores nominais de frações dos elementos estão descritos nas legendas do gráfico. Os valores dos resíduos são descritos nos pontos (✱), e os pontos (□) indicam o condicionamento normalizado das matrizes de covariância do particionamento126	126

- Figura 22 – Plots demonstrando o intervalo de valores para o resíduo durante 30 execuções do algoritmo proposto no particionamento dos dados. Na ausência do Am , 50% dos valores de resíduo encontram-se compreendidos entre 2×10^{-15} e 4×10^{-16} . A mesma interpretação pode ser estendida para os demais elementos por meio dos seus respectivos boxes. Os valores de resíduos são positivamente assimétricos para o Co , Cs e Na e aproximadamente simétrico no caso do Am 127
- Figura 23 – Plots demonstrando os valores de condicionamento. Na ausência do Am , os valores de resíduo estão compreendidos entre 3 e 4, com os dados distribuídos aproximadamente simetricamente. O Co apresenta esses valores compreendidos no intervalo 2 e 2.5; o condicionamento para a instância do Cs está no intervalo de 3.5 e 5 e, por fim, o Na tem seu intervalo definido entre os valores 3 e 4. 128
- Figura 24 – Frações de contagens percentuais, instância 1, omissão e aproximação do Am . Pontos (○), vermelho (+), azul (*) e verde (☆) indicam, respectivamente, as frações em peso para o Am , Co , Cs e Na . Os valores dos resíduos são descritos nos pontos ✕, e os pontos □ indicam o condicionamento normalizado das matrizes de covariância no particionamento. Percebe-se que à medida que o resíduo é reduzido, as frações de contagem percentuais convergem aos valores nominais. 130
- Figura 25 – Plot para os intervalos de resíduo após 30. Para a aproximação do Am , a mediana dos valores é aproximadamente 5×10^{-3} , com 50% dos valores encontrando-se compreendidos aproximadamente entre 1×10^{-3} e 10^{-2} . Na ausência do Co , a mediana é de aproximadamente 1.2×10^{-2} e 50% dos valores compreendidos entre 7×10^{-3} e 1.7×10^{-2} . O intervalo de valores para ausência do Cs são da ordem de 5×10^{-3} e 2.25×10^{-2} . Os valores de resíduo para a ausência do Na é de 2.5×10^{-3} e 7.5×10^{-3} . A mediana para a ausência do Am , C e Na são positivamente assimétricas; o plot para a ausência do Co mostra os dados aproximadamente simétricos. . . . 131

- Figura 26 – Plots para o intervalo de valores para o condicionamento após 30 execuções do Algoritmo GRASP. Na ausência do Am , 50% desses valores encontram-se compreendidos entre 2.8 e 3.8. Para os demais compostos, respectivamente, esses valores encontram-se entre: Co , 3 e 6; Cs , 3.8 e 5; Na , 3 e 4.6. A mediana para a ausência do Am é negativamente assimétrica; o plot para a ausência do Co e Na mostra os dados aproximadamente simétricos e positivamente assimétricos para a ausência do Co 132
- Figura 27 – Plots contendo os intervalos de frações de contagens percentuais para cada uma das instâncias. A figura 27a possui o plot com a mediana em verde para indicar que o Am está sendo aproximado. A mesma interpretação pode ser estendida para as Figuras 27b, 27c e 27d, na aproximação dos seus respectivos compostos (Co , Cs e Na). As frações de contagens percentuais na Figura 27a refletem aproximadamente os valores nominais. Na figura 27b existe uma grande variabilidade na aproximação do compostos aproximado, o Co , mas ainda englobando o valor nominal desse composto no intervalo citado. As interpretação da figura 27a pode ser estendida para as figuras 27c e 27d. 133
- Figura 28 – Estimativas da biblioteca omitida, da instância 1. Os pontos $\bigcirc$ são valores de menor contagem estimada pelo algoritmo da Seção 4.3, enquanto que representados por $+$ indicam a menor contagem total no intervalo de canais que definem o compartimento da partição. Pontos $\star$ são os valores de maior contagem estimada para a biblioteca omitida e os pontos $\star$ indicam o maior valor total para essa mesma biblioteca. Valores de estimativas são adequados, já que delimitam os valores de contagem espectral total dentro dos compartimentos. 135
- Figura 29 – Valores de estimativa de altura para as bibliotecas omitidas Am , (Figura 29a, instância 1), Co , (Figura 29b, instância 2), Cs (Figura 29c, instância 3) e Na (Figura 29d, instância 4) nas suas partições de maior resíduo. . . 136
- Figura 30 – Comportamento do condicionamento na minimização da métrica. Na Figura 30, são descritos os resultados para a instância 1. Os valores tendem a se estabilizar em um intervalo, evidenciando que a minimização dos resíduos podem ser associados a um decrescimento do condicionamento, mesmo que esse decrescimento não seja monotônico. 137

Figura 31 – Boxplot demonstrando a distribuição de valores do condicionamento das matrizes de covariância na determinação das frações de contagem percentual na omissão de uma das bibliotecas da solução inicial. O primeiro plot é considerado para a instância 1, e os demais plots são descritos para as demais instâncias.	138
Figura 32 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para a instância 1, no estágio da construção. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância 1.	141
Figura 33 – O gráfico 33a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 33b.	141
Figura 34 – O gráfico 34a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 34b.	142
Figura 35 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, instância 1 (Figura 32a) no estágio da busca. A figura 32b traz a distribuição dos canais que definem o compartimento da partição de maior resíduo na ausência do Am	143
Figura 36 – O gráfico 36a apresenta os valores das estimativas para a biblioteca omitida, no passo de busca do GRASP, instância um. A distribuição para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 36b.	144
Figura 37 – O gráfico 34a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância um. A distribuição para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 34b.	145
Figura 38 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do Am , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Am	146

Figura 39 – O gráfico 39a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 39b.	147
Figura 40 – O gráfico 40a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 40b.	147
Figura 41 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para a instância 1, no estágio da segunda busca. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do <i>Am</i>	148
Figura 42 – O gráfico 42a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 42b.	149
Figura 43 – O gráfico 43a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 43b.	149
Figura 44 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas figuras 44a e 44b são apresentadas as frações percentuais para a instância 1. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. Os pontos ○ indicam as aproximações para o <i>Am</i> ; os pontos + são as aproximações para <i>Co</i> ; as aproximações para o <i>Cs</i> são determinados por pontos *; pontos ☆ indicam as aproximações para o <i>Na</i> ; Os valores para os resíduos são pontos ✕ e os pontos □ indicam os valores para condicionamento.	151
Figura 45 – Valores de frações de contagem percentuais nas iterações do CGRASP. Na Figura 45a estão as distribuições para fração de contagem percentual dos compostos para a instância um, omissão do <i>Am</i> . Figura 45b, instância dois, omissão do <i>Co</i> , Figura 45c, instância três, omissão do <i>Cs</i> e, por fim, 45d, instância quatro, omissão do <i>Na</i>	154

- Figura 46 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 46a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (o Am é representado por pontos $\bigcirc$; o Co tem representação nos pontos $+$; o Cs é indicado por pontos $*$ e o Na é representado por pontos $\star$). A Figura 46b indica a área aproximada (pontos representados por $\bigcirc$), área real (pontos representados por $+$) e resíduo pontos representados por $*$), demonstrando a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 46c tem-se os valores de estimativa do composto omitido (Am) e a aproximação dos demais compostos. 156
- Figura 47 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 47a, têm-se os valores quantitativos dos isótopos que constituem o espectro total (pontos $\bigcirc$ representam o Am ; pontos $+$ representam o Co ; pontos $*$ representam o Cs ; pontos $\star$ representam o Na). Na Figura 47b tem-se a seguinte determinação: área calculada são os pontos $\bigcirc$; área real são os pontos $+$; resíduo são os pontos $*$, demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 47c tem-se a distribuição das frações de contagem percentual, instância 1, e a aproximação dos demais compostos. 158
- Figura 48 – Componentes dos vetores de compostos (Ar, Água e Óleo), resíduo e componentes do vetor espectral total, como função de partição. Os componentes do vetor espectral total são apresentados (variável $tot - 5$), além dos valores de resíduos nas partições (variável $res - 1$). 161
- Figura 49 – Frações percentuais dos compostos após particionamento dos dados. As partições de referências são anexadas à partição de maior resíduo no intuito definir a matriz de covariância. Pontos em preto ($\bigcirc$, Ar), vermelho($+$, Água) e azul ($*$, Óleo) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais de frações de cada composto encontram-se na legenda do gráfico. Os valores dos resíduos são descritos nos pontos em verde ($\star$), e o rosa ($\times$) indica o condicionamento normalizado das matrizes de covariância do particionamento. 162

- Figura 50 – Plot comparando as distribuições nos valores de condicionamento para as instâncias 5, 6 e 7. Na ausência do Ar, os valores de condicionamento estão amplamente distribuídos no intervalo de 0 e 0.35 com os dados distribuídos assimetricamente positivo. Para a instância da Água, os valores estão distribuídos no intervalo entre 0 e 0.07 163
- Figura 51 – Frações percentuais dos compostos após supressão e estimação do Ar no espectro total, instância 5. Valores em preto (○), vermelho(+) e azul(★) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais estão descritos na legenda. Os valores dos resíduos são descritos nos pontos ☆, e o ponto ✕ indica o condicionamento normalizado (entre 0 e 1) das matrizes de covariância na aproximação das bibliotecas. 166
- Figura 52 – Plots contendo os intervalos de frações de contagens percentuais após a omissão das bibliotecas de energia. As medianas em verde indicam a biblioteca que está sendo estimada. 50% dos valores, para as frações de contagens percentuais, nas Figuras 52a e 52b, estão contidos em um intervalo de pequena amplitude. Existe uma maior variabilidade desses valores na Figura 52c. A convergência dos valores preditos, aos valores nominais das frações de contagem percentual, é melhor nas Figuras 52a e 52b do que na Figura 52c. 167
- Figura 53 – Intervalo de valores para o resíduo durante 30 execuções do algoritmo. Na ausência e aproximação do Ar, a ordem de grandeza da mediana dos valores é de aproximadamente 1.8×10^{-1} , sendo que 50% dos valores de resíduo encontram-se compreendidos no intervalo 10^{-1} e 2×10^{-1} . Na ausência da Água, a mediana dos valores é de aproximadamente 6×10^{-1} e 50% dos valores de resíduos estão compreendidos entre 6×10^{-1} e 5×10^{-1} , e, para o Óleo, o intervalo é de 1×10^{-1} e 5×10^{-1} . A amplitude da distribuição do resíduo é maior na instância da ausência do Óleo, do que as demais bibliotecas. 168
- Figura 54 – Plots demonstrando o intervalo de valores para o condicionamento durante 30 execuções do GRASP para a Segunda base de dados. Na ausência do Ar, 50% desses valores encontram-se compreendidos entre 6 e 9. A mediana para a ausência do Ar é positivamente assimétricas. O plot para a ausência do Água e do Óleo mostram dados negativamente assimétricos. 169

- Figura 55 – Estimativas do Ar, instância 5. Os pontos ○ indicam os valores de menor contagem estimada pelo algoritmo da Seção 4.3 enquanto pontos representados por + indicam a menor contagem espectral total real no intervalo de canais que definem a partição. Pontos * são os valores de maior contagem estimada para a biblioteca omitida e os pontos ☆ indicam o maior valor real para essa mesma biblioteca. 170
- Figura 56 – Valores de estimativa de altura para o elemento omitido nas instâncias da Base de Dados 2. 171
- Figura 57 – Condicionamento na minimização da métrica. Na Figura 57, têm-se os valores de condicionamento, na minimização do resíduo, para a ausência e aproximação do Ar, instância 5. Os valores tendem a se estabilizar em um intervalo, evidenciando que a minimização dos resíduos podem ser associados a um decrescimento do condicionamento, mesmo que esse decrescimento não seja monotônico 172
- Figura 58 – Boxplot demonstrando a distribuição de valores do condicionamento das matrizes de covariância na determinação das frações de contagem percentual na omissão de uma das bibliotecas da solução inicial. O primeiro plot é considerado para a ausência do Ar, e os demais plots seguem a mesma descrição. 173
- Figura 59 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 59a e 59b são representados duas execuções arbitrárias para a instância 5. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. Os pontos ○ indicam as aproximações para o Ar; + são pontos para as aproximações para a Água; os pontos * indicam as aproximações para o Óleo; os valores de resíduo são representados por pontos ☆ e o símbolo ✖ indica os valores para condicionamento. 174
- Figura 60 – Valores de fração de contagem percentual nas iterações do CGRASP, Base de Dados 2. Na Figura 60a estão as distribuições para fração de contagem percentual dos compostos na omissão do Ar, instância cinco. Figura 60b, omissão da Água, instância seis, e Figura 60c, omissão do Óleo, instância sete. 177

- Figura 61 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, têm-se os valores quantitativos dos isótopos que constituem o espectro total. Na Figura 157b, tem-se a seguinte descrição: Área estimada representada pelos pontos($\bigcirc$); área real representada pelos pontos ($+$); resíduo representados pelos pontos ($*$). A evolução demonstra a métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c têm-se os valores de estimativa da biblioteca omitida (instância 5, omissa do Ar) e a aproximação dos demais compostos. 179
- Figura 62 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, têm-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar $\bigcirc$);(Água $+$);(Óleo $*$); na Figura 157btem-se a seguinte descrição: Área estimada ($\bigcirc$); área real ($+$); resíduo ($*$) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c têm-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos. 181
- Figura 63 – Componentes dos vetores de compostos como função de partição. Os componentes do vetor total são apresentados (variável *tot* – 5), além dos valores de resíduos (variável *res* – 1). 221
- Figura 64 – Resultados para determinação dos picos, Base de Dados 2. 222
- Figura 65 – Determinação quantitativa dos compostos por particionamento dos dados. Pontos ($\bigcirc$), vermelho($+$) , azul ($*$) e verde ($\star$) indicam, respectivamente, as quantidades de *Am*, *Co*, *Cs* e *Na*. Os valores nominais são descritos nas legendas dos Gráficos 65a, 65b e 65c. Os valores dos resíduos são descritos nos pontos ($\times$), e o amarelo ($\square$)indica o condicionamento normalizado das matrizes de covariância do particionamento. 223
- Figura 66 – Determinação quantitativa dos compostos por particionamento dos dados. Pontos ($\bigcirc$, Ar), vermelho($+$, Água) e azul ($*$, Óleo) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais estão descritos nas legendas dos Gráficos 66a e 66b. Os valores dos resíduos são descritos por pontos ($\star$), e o rosa ($\times$) indica o condicionamento normalizado das matrizes de covariância do particionamento. 224

- Figura 67 – Gráficos contendo as frações de contagens percentuais dos compostos após a omissão, e estimação, de uma biblioteca do espectro total. Pontos (○), vermelho (+), azul (*) e verde (☆) indicam, respectivamente, as quantidades de *Am*, *Co*, *Cs* e *Na*. Os valores dos resíduos são descritos nos pontos ✕, e □ indica o condicionamento normalizado das matrizes de covariância na aproximação. 225
- Figura 68 – Gráficos contendo a determinação de fração de contagens percentuais dos compostos. Pontos (○), vermelho(+) e azul(*) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais estão descritos nas legendas dos Gráficos 68a e 68b. Os valores dos resíduos são descritos nos pontos ☆, e o ponto ✕ indica o condicionamento das matrizes de covariância na aproximação das bibliotecas. 226
- Figura 69 – Estimativas da biblioteca omitida. Os pontos ○ e * são os valores de mínima e máxima, respectivamente, contagem espectral estimada, e os pontos representados por + e ☆ indicam o mínima e o máxima, respectivamente, contagem espectral real no intervalo do compartimento. As figuras 69a, 69b e 69c são resultados para as instâncias das bibliotecas, respectivamente, do *Co*, *Cs* e *Na*. 227
- Figura 70 – Gráficos estimativa de altura nas partições, Base de Dados 2. Estimativa da biblioteca omitida. Os pontos ○ são os valores de mínima contagem espectral estimado. Os pontos por + representam a mínima contagem espectral real total nos compartimentos. Os pontos * representam os valores de máxima contagem espectral estimada. Os pontos ☆ indicam o valor de máximo espectral real total. As figuras 70a e 70b são representativas para a omissão das bibliotecas, respectivamente da Água e do Óleo. 228
- Figura 71 – Comportamento do número de condicionamento na minimização do resíduo. A Figura 71a traz os resultados para a instância dois; a Figura 71b traz os resultados para a instância 3 e, por fim, a Figura 71c traz os resultados para a instância quatro. 229
- Figura 72 – Condicionamento na determinação das frações de contagem percentual, na omissão de uma das bibliotecas, nas ocorrências de minimização do resíduo. Na Figura 72a, resultados para a instância seis; Figura 72b, valores de condicionamento para a instância sete. 230

Figura 73 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 73a, para omissão do C_o , no estágio da construção. A figura 73b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do C_o	231
Figura 74 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 74a, para omissão do C_s , no estágio da construção. A figura 74b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do C_s	232
Figura 75 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 75a, para omissão do N_a , no estágio da construção. A figura 75b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do N_a	232
Figura 76 – O gráfico 76a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 76b.	233
Figura 77 – O gráfico 77a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 77b.	233
Figura 78 – O gráfico 78a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 78b.	234
Figura 79 – O gráfico 79a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 79b.	234
Figura 80 – O gráfico 80a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 80b.	235

Figura 81 – O gráfico 81a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 81b.	235
Figura 82 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 82a, para omissão do Co , no estágio da busca. A figura 82b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância dois.	236
Figura 83 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 83a, para omissão do Cs , no estágio da busca. A figura 83b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância três.	236
Figura 84 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 84a, para omissão do Na , no estágio da busca. A figura 84b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância três.	237
Figura 85 – O gráfico 85a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 85b.	237
Figura 86 – O gráfico 86a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 86b.	238
Figura 87 – O gráfico 87a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 87b.	238
Figura 88 – O gráfico 88a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 88b.	239

Figura 89 – O gráfico 89a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 89b.	239
Figura 90 – O gráfico 90a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 90b.	240
Figura 91 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do C_o , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do C_o	240
Figura 92 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do C_s , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do C_s	241
Figura 93 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do N_a , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do N_a	241
Figura 94 – O gráfico 94a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 94b.	242
Figura 95 – O gráfico 95a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 95b.	242
Figura 96 – O gráfico 96a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 96b.	243

Figura 97 – O gráfico 97a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 97b.	243
Figura 98 – O gráfico 98a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 98b.	244
Figura 99 – O gráfico 99a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 99b.	244
Figura 100–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 100a, para omissão do Co , no estágio da segunda busca. A figura 100b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Am	245
Figura 101–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 101a, para omissão do Cs , no estágio da segunda busca. A figura 101b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Am	245
Figura 102–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 102a, para omissão do Na , no estágio da segunda busca. A figura 102b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Na	246
Figura 103–O gráfico 103a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 103b.	246
Figura 104–O gráfico 104a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 104b.	247

Figura 105–O gráfico 105a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 105b.	247
Figura 106–O gráfico 106a apresenta os valores de constantes que definem a biblioteca omitida, no passo de segunda busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 106b.	248
Figura 107–O gráfico 107a apresenta os valores de constantes que definem a biblioteca omitida, no passo de segunda busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 107b.	248
Figura 108–O gráfico 108a apresenta os valores de constantes que definem a biblioteca omitida, no passo de segunda busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 108b.	249
Figura 109–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 109a, para omissão do Ar, no estágio da construção. A figura 109b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.	250
Figura 110–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 110a, para omissão da Água, no estágio da construção. A figura 110b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.	250
Figura 111–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 111a, para omissão do Óleo, no estágio da construção. A figura 111b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.	251
Figura 112–O gráfico 112a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 112b.	251

Figura 113–O gráfico 113a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 113b.	252
Figura 114–O gráfico 114a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 114b.	252
Figura 115–O gráfico 115a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 115b.	253
Figura 116–O gráfico 116a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 116b.	253
Figura 117–O gráfico 117a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 117b.	254
Figura 118–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 118a, para omissão do Ar, no estágio da busca. A figura 118b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.	254
Figura 119–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 119a, para omissão da Água, no estágio da busca. A figura 119b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.	255
Figura 120–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 120a, para omissão da Água, no estágio da busca. A figura 120b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.	255

Figura 121–O gráfico 121a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 121b.	256
Figura 122–O gráfico 122a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 122b.	256
Figura 123–O gráfico 123a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 123b.	257
Figura 124–O gráfico 124a apresenta os valores de constantes que definem a biblioteca omitida, no passo da primeira busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 124b.	257
Figura 125–O gráfico 125a apresenta os valores de constantes que definem a biblioteca omitida, no passo da primeira busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 125b.	258
Figura 126–O gráfico 126a apresenta os valores de constantes que definem a biblioteca omitida, no passo da primeira busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 126b.	258
Figura 127–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 127a, para omissão do Ar, no estágio da linkagem. A figura 127b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.	259
Figura 128–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 128a, para omissão da Água, no estágio da linkagem. A figura 128b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.	259

Figura 129–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 129a, para omissão do Óleo, no estágio da linkagem. A figura 129b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.	260
Figura 130–O gráfico 130a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 130b.	260
Figura 131–O gráfico 131a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 131b.	261
Figura 132–O gráfico 132a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 132b.	261
Figura 133–O gráfico 133a apresenta os valores de constantes que definem a biblioteca omitida, no passo da linkagem do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 133b.	262
Figura 134–O gráfico 134a apresenta os valores de constantes que definem a biblioteca omitida, no passo da linkagem do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 134b.	262
Figura 135–O gráfico 135a apresenta os valores de constantes que definem a biblioteca omitida, no passo da linkagem do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 135b.	263
Figura 136–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 136a, para omissão do Ar, no estágio da Segunda Busca. A figura 136b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.	263

Figura 137–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 137a, para omissão da Água, no estágio da Segunda Busca. A figura 137b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.	264
Figura 138–Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 138a, para omissão do Óleo, no estágio da Segunda Busca. A figura 138b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.	264
Figura 139–O gráfico 139a apresenta os valores de estimativas para a biblioteca omitida, no passo de Segunda Busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 139b.	265
Figura 140–O gráfico 140a apresenta os valores de estimativas para a biblioteca omitida, no passo de Segunda Busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 140b.	265
Figura 141–O gráfico 141a apresenta os valores de estimativas para a biblioteca omitida, no passo de Segunda Busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 141b.	266
Figura 142–O gráfico 142a apresenta os valores de constantes que definem a biblioteca omitida, no passo da Segunda Busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 142b.	266
Figura 143–O gráfico 143a apresenta os valores de constantes que definem a biblioteca omitida, no passo da Segunda Busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 143b.	267
Figura 144–O gráfico 144a apresenta os valores de constantes que definem a biblioteca omitida, no passo da Segunda Busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 144b.	267

- Figura 145–Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 145a e 145b os valores de frações de contagem percentual, resíduo e condicionamento, como função da iteração, para omissão do *Co*. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. O ponto $\bigcirc$ indica as aproximações para o *Am*; $+$ as aproximações para *Co*; $*$ as aproximações para o *Cs*; $\star$ as aproximações para o *Na*; $\times$ os valores para o resíduo e $\square$ os valores para número de condicionamento. 268
- Figura 146–Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 146a e 146b os valores de frações de contagem percentual, resíduo e condicionamento, como função da iteração, para omissão do *Cs*. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. O ponto $\bigcirc$ indica as aproximações para o *Am*; $+$ as aproximações para *Co*; $*$ as aproximações para o *Cs*; $\star$ as aproximações para o *Na*; $\times$ os valores para o resíduo e $\square$ os valores para número de condicionamento. 269
- Figura 147–Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 147a e 147b os valores de frações de contagem percentual, resíduo e condicionamento, como função da iteração, para omissão do *Na*. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. O ponto $\bigcirc$ indica as aproximações para o *Am*; $+$ as aproximações para *Co*; $*$ as aproximações para o *Cs*; $\star$ as aproximações para o *Na*; $\times$ os valores para o resíduo e $\square$ os valores para número de condicionamento. 269
- Figura 148–Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 148a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (*Am* $\bigcirc$); *Co* $+$; *Cs* $*$); na Figura 148b($\bigcirc$); área real $+$; resíduo $*$) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 148c tem-se os valores de estimativa do composto omitido (*Co*) e a aproximação dos demais compostos. 270

- Figura 149 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 149a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((*Am* ○); *Co* +; *Cs* *); na Figura 149b(○); área real +; resíduo *) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 149c tem-se os valores de estimativa do composto omitido (*Cs*) e a aproximação dos demais compostos. 271
- Figura 150 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 150a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((*Am* ○); *Co* +; *Cs* *); na Figura 150b(○); área real +; resíduo *) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 150c tem-se os valores de estimativa do composto omitido (*Na*) e a aproximação dos demais compostos. 272
- Figura 151 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 151a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((*Am* ○); *Co* +; *Cs* *); na Figura 151b são demonstrados: área calculada(○); área real +; resíduo *), demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 151c tem-se a distribuição das frações de contagem percentual, omitido *Co* (mediana em verde), e a aproximação dos demais compostos. 273
- Figura 152 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 152a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((*Am* ○); *Co* +; *Cs* *); na Figura 152b são demonstrados: área calculada(○); área real +; resíduo *), demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 152c tem-se a distribuição das frações de contagem percentual, omitido *Cs* (mediana em verde), e a aproximação dos demais compostos. 274
- Figura 153 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 153a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((*Am* ○); *Co* +; *Cs* *); na Figura 153b são demonstrados: área calculada(○); área real +; resíduo *), demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 153c tem-se a distribuição das frações de contagem percentual, omitido *Na* (mediana em verde), e a aproximação dos demais compostos. 275

- Figura 154–Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas figuras 154a e 154b valores referentes à omissão da Água. Os percentuais aproximam-e dos valores nominais à medida que os resíduos reduzem. ○ indicam as aproximações para o Ar; + as aproximações para a Água; * as aproximações para o Óleo; ☆ os valores para o resíduo ✕ os valores para número de condicionamento. 276
- Figura 155–Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas figuras 155a e 155b valores referentes à omissão do Óleo. Os percentuais aproximam-e dos valores nominais à medida que os resíduos reduzem. ○ indicam as aproximações para o Ar; + as aproximações para a Água; * as aproximações para o Óleo; ☆ os valores para o resíduo ✕ os valores para número de condicionamento. 277
- Figura 156–Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○); (Água +); (Óleo *); na Figura 157btem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Água) e a aproximação dos demais compostos. 278
- Figura 157–Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○);(Água +);(Óleo *); na Figura 157b tem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos. 279
- Figura 158–Resultados acumulados para a etapa de busca do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○);(Água +);(Óleo *); na Figura 157b tem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) demonstram a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos. 280

Figura 159 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○);(Água +);(Óleo *); na Figura 157b tem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) demonstram a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos. 281

LISTA DE TABELAS

Tabela 1 – Mapeamento Ternary' ($n = 2$)	77
Tabela 2 – Resumo dos trabalhos em PGNA e metodologias no ajuste do espectro.	87
Tabela 3 – Descrição das bases de dados estudadas e das instâncias criadas a partir das bases.	89
Tabela 4 – Descrição dos cinco primeiros canais das bibliotecas do Amerício, Cobalto, Césio e Sódio GRASP. Os dados de entrada para a primeira base possuem uma quantidade de 20148 canais.	122
Tabela 5 – Regiões para os picos dos elementos fornecidos GRASP	122
Tabela 6 – Partições contendo os picos dos constituintes. Algoritmo GRASP.	123
Tabela 7 – Valores médios de frações de contagens percentuais dos compostos inseridos no espectro total, após 30 execuções do algoritmo. A linha 1 contém os valores de frações percentuais dos compostos para a instância 1. As instâncias 2, 3 e 4 podem ser observadas a partir das linhas seguintes. As duas últimas colunas contêm os valores médios de resíduo e condicionamento para trinta execuções do algoritmo.	128
Tabela 8 – Valores aproximados das frações em peso para os menores valores do resíduo. A linha 1 da tabela determina os valores de frações dos elementos inseridos na amostra, estimando a biblioteca do <i>Am</i> e o aproximando-o por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas: linha 2 é referente à instância 2, omissão do <i>Co</i> ; linha 3 é referente à instância 3, omissão do <i>Cs</i> e a última linha refere-se à instância 4, omissão do <i>Na</i>	129
Tabela 9 – Valores de frações de contagem percentual aproximados considerando o menor condicionamento, nas ocorrências de minimização do resíduo, como métrica. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o <i>Am</i> da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas.	129

- Tabela 10 – Frações de contagem percentual médios para as execuções do CGRASP. A linha 1 determina as frações em peso para as bibliotecas, na ausência do *Am*, na solução inicial. A linha 2 apresenta os resultados para a ausência do *Co*. Linha 3 e 4, respectivamente, seguem a mesma descrição, para a ausência *Cs* e *Na*. Os valores das precisões estão descritos pelas colunas que se iniciam com o prefixo 'pre': *PreAm*, precisão para o *Am*; *PreCo*, precisão para o *Co*, e assim sucessivamente. 152
- Tabela 11 – Frações em peso aproximadas na determinação quantitativa de compostos inseridos no espectro total, tomando como métrica a aproximação do resíduo. A linha 1 da tabela determina os valores de frações percentuais omitindo o *Am* da geração do espectro total. A mesma descrição é válida para as demais linhas: linha 2, omite-se o *Co*; linha 3, omite-se o *Cs* e na última linha é omitido o *Na*. As colunas *PreAm*, *PreCo*, *PreCs* e *PreNa* indicam as precisões na aproximação das frações de contagem percentuais para o *Am*, *Co*, *Cs* e *Na*. A última coluna traz o menor valor de resíduo. . . . 153
- Tabela 12 – Valores de percentuais aproximados na determinação quantitativa de compostos inseridos em um espectro total. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o *Am* da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas. . . 153
- Tabela 13 – Resultados na determinação dos picos para a segunda base de dados. . . . 159
- Tabela 14 – Valores médios de frações de contagens percentuais na determinação quantitativa de compostos inseridos no espectro total. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o *Ar* da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas. 164
- Tabela 15 – Valores de frações de contagens percentuais na determinação quantitativa de compostos inseridos em um espectro total, tomando por métrica a minimização resíduo. A linha 1 exibe o *Ar* na descrição do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas: linha 2, omite-se o a Água; linha 3, omite-se o Óleo. 165

Tabela 16 – Valores de percentuais aproximados na determinação quantitativa de compostos inseridos em um espectro total no menor condicionamento nas ocorrências de minimização do resíduo. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o Ar da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas.	165
Tabela 17 – Valores médios de frações de contagens percentuais dos compostos inseridos no espectro total. A linha 1 contém os valores de frações percentuais dos compostos, na instância 5. Os resultados para as instâncias 6 e 7 são descritos nas demais linhas.	175
Tabela 18 – Valores de frações em peso na minimização do resíduo. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos para a instância 5. A mesma descrição é válida para as demais linhas: linha 2, omite-se a Água; linha 3, omite-se o Óleo.	176
Tabela 19 – Valores de percentuais aproximados na para as frações de contagem percentuais tomando por métrica a minimização do condicionamento. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, para a instância 5. A mesma descrição é válida para as demais linhas: linha 2, instância 6; linha 3, instância 7.	176
Tabela 20 – Comparativo com trabalhos do estado da arte. A tabela adota a notação Isótopo - Participação percentual para cada Dataset. Os datasets 1, 2, 3 e 4 apresentam a omissão, respectivamente, do <i>Am</i> , <i>Cs</i> , <i>Co</i> e <i>Na</i>	185
Tabela 21 – Tabela demonstrando o comparativo entre as frações percentuais antes e depois da execução da otimização por parte do algoritmo.	187

LISTA DE ABREVIATURAS E SIGLAS

AMPSO	<i>Adaptative Mutation based PSO</i>
CGRASP	<i>Continuos Greedy Randomized Adaptative Search Procedure</i>
FA	<i>Fireworks Algorithm</i>
FSA	<i>Full Spectrum Analysis</i> - Análise de espectro total
GAs	<i>Genetic Algorithms</i> - Algoritmos Genéticos
GM	<i>Geiger-Muller</i>
GRASP	<i>Greedy Randomized Adaptative Search Procedure</i>
HFG	<i>Hybrid Fuzzy Genetic</i>
LLS	<i>Library Least Square</i> - Biblioteca Mínimos Quadrados
MAE	<i>Mean Absolut Error</i>
MC	Monte Carlo
MCLS	<i>Monte Carlo Library Least Squares</i>
MCLLSX	<i>Monte Carlo Library Least Square Extended</i>
MCNP	<i>Monte Carlo N-Particle</i>
MLR	<i>Multiple Linear Regression</i>
MMQ	Método dos Mínimos Quadrados
NNLS	<i>Non Negative Least Squares</i> - Mínimos Quadrados Não Negativos
PGNAA	<i>Prompt Gamma Neutron Analysis Activation</i>
PSO	<i>Particle Swarm Optimization</i> - Otimização por enxame de partículas
RCL	Lista Restrita de Candidatos
RMSE	<i>Root Mean Square Error</i>
SGSD	<i>Sequential Gamma-ray Spectrum Deconvolution</i>
SOA	<i>Significance Of Abundance</i>
SVD	<i>Singular Value Decomposition</i> - Decomposição em Valores Singulares
VRP	<i>Vehicle Routing Problem</i> - Problema de roteamento de veículos

WSA

Whole Spectrum analysis - Análise de Espectro Total

LISTA DE SÍMBOLOS

γ Letra grega Gama

$\in$ Pertence

δ Delta

θ Teta

σ Sigma

μ Mi

Am Amerício

Co Cobalto

Na Sódio

Cs Césio

SUMÁRIO

1	INTRODUÇÃO	43
1.1	MOTIVAÇÃO E JUSTIFICATIVA	46
1.2	ESTADO DO PROBLEMA	47
1.3	OBJETIVOS	49
1.4	CONTRIBUIÇÕES	49
1.5	ESTRUTURA DA PROPOSTA	50
2	FUNDAMENTAÇÃO	52
2.1	PGNAA	52
2.2	CONDICIONAMENTO DE UM SISTEMA	56
2.3	DECOMPOSIÇÃO EM VALORES SINGULARES (SVD)	57
2.4	CONDICIONAMENTO E SVD	58
2.5	MÍNIMOS QUADRADOS (MMQ)	63
2.6	MCLLS	64
2.6.1	Definições	64
2.6.2	Fundamentação	65
2.6.3	Mau Condicionamento	67
2.7	GREEDY RANDOMIZED ADAPTATIVE SEARCH PROCEDURE	69
2.8	CONTINUOS GREEDY RANDOMIZED ADAPTATIVE SEARCH PROCEDURE	73
3	ANÁLISE ESPECTRAL E O ESTADO DA ARTE	79
4	METODOLOGIA	88
4.1	DETERMINAÇÃO DOS PICOS DAS BIBLIOTECAS FORNECIDAS	94
4.2	DETERMINAÇÃO DA PARTIÇÃO DE MAIOR RESÍDUO	95
4.3	DESCRIÇÃO DA ETAPA DA GEOMETRIZAÇÃO	98
4.4	MODELAGEM DO PASSO CONSTRUTIVO ALGORITMO GRASP	101
4.5	AJUSTE DO ELEMENTO OMITIDO VIA MODELAGEM DO CGRASP	102
4.5.1	Modelagem do Ajuste via CGRASP - Passo construtivo	103
4.5.2	Modelagem do Ajuste via CGRASP - Passo de Busca	106
4.6	MODELAGEM DO PASSO DE BUSCA DO ALGORITMO GRASP	110
4.7	MODELAGEM DO PASSO DE LINKAGEM ALGORITMO GRASP	114

4.7.1	Linkagem Grasp - Canais	117
4.7.2	Linkagem Grasp - Altura	119
5	RESULTADOS	121
5.1	RESULTADOS PARA OS ELEMENTOS	121
5.1.1	Determinação dos picos	122
5.1.2	Determinação dos Resíduos das Partições	123
5.1.3	Determinação de Frações Percentuais - Particionamento	125
5.1.4	Determinação da Fração Percentual dos Compostos por Aproximações	127
5.1.5	Aproximação dos Compostos - Estimativa das alturas nas partições	134
5.1.6	Condicionamento na Minimização da Métrica na aproximação dos compostos.	137
5.1.7	Resultados dos passos básicos do GRASP	139
5.1.7.1	<i>Resultados do Passo construtivo</i>	<i>139</i>
5.1.7.2	<i>Resultados do Passo de Busca</i>	<i>143</i>
5.1.7.3	<i>Resultados do Passo de Linkagem</i>	<i>145</i>
5.1.7.4	<i>Resultados do Passo da Segunda Busca</i>	<i>148</i>
5.1.8	Resultados do CGRASP	150
5.1.8.1	<i>Resultados do CGRASP - Execução Única.</i>	<i>150</i>
5.1.8.2	<i>Resultados do CGRASP - Execuções Acumuladas / Valores Métricas</i>	<i>152</i>
5.1.8.3	<i>Resultados do CGrasp - Etapa Construtiva.</i>	<i>155</i>
5.1.8.4	<i>Resultados do CGrasp - Etapa Busca.</i>	<i>157</i>
5.2	RESULTADOS PARA OS COMPOSTOS.	159
5.2.1	Determinação dos picos	159
5.2.2	Determinação dos Resíduos das Partições	160
5.2.3	Determinação de Frações Percentuais - Particionamento	161
5.2.4	Determinação da Fração Percentual dos Compostos por Aproximações	164
5.2.5	Aproximação dos Compostos - Estimativa das alturas nas partições	169
5.2.6	Condicionamento na Minimização da Métrica na aproximação dos compostos.	171
5.2.7	Resultados dos passos básicos do GRASP	173
5.2.8	Resultados do CGRASP	173

5.2.8.1	<i>Resultados do CGRASP - Execução Única.</i>	174
5.2.8.2	<i>Resultados do CGRASP - Execuções Acumuladas / Valores Métricas</i>	175
5.2.8.3	<i>Resultados do CGrasp - Execuções Acumuladas - Etapa Construtiva.</i>	178
5.2.8.4	<i>Resultados do CGRASP - Etapa de Busca.</i>	180
5.3	COMPARAÇÃO COM TRABALHOS DO ESTADO DA ARTE	182
6	CONCLUSÕES E TRABALHOS FUTUROS	188
6.1	TRABALHOS FUTUROS	189
	REFERÊNCIAS	191
	APÊNDICE A – NÚMERO DE CONDICIONAMENTO	199
	APÊNDICE B – MMQ	203
	APÊNDICE C – METODOLOGIA - PSEUDOCÓDIGOS	205
	APÊNDICE D – DETERMINAÇÃO DOS RESÍDUOS DAS PARTIÇÕES	221
	APÊNDICE E – DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - PARTICIONAMENTO	223
	APÊNDICE F – DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - APROXIMAÇÃO	225
	APÊNDICE G – APROXIMAÇÃO DOS COMPOSTOS - ESTIMATIVA DA ALTURA NAS PARTIÇÕES	227
	APÊNDICE H – CONDICIONAMENTO NA MINIMIZAÇÃO DA MÉTRICA	229
	APÊNDICE I – RESULTADOS ETAPAS BÁSICAS DO GRASP .	231
	APÊNDICE J – RESULTADOS CGRASP	268
	APÊNDICE K – PUBLICAÇÕES	282

1 INTRODUÇÃO

A técnica do *Prompt Gamma Neutron Analysis Activation* (PGNAA) é amplamente utilizada para análise e determinação dos elementos que formam uma amostra. Os principais mecanismos físicos aplicados no PGNAA para determinação da composição nuclear são espalhamento inelástico e captura de nêutrons (ZHANG; TUO, 2013) (ZHANG et al., 2015). Os espectros gama (gráficos de intensidade espectral como função da energia, também conhecidos por bibliotecas) de uma amostra são gerados, definindo seus elementos, seguindo-se à etapa da determinação das frações em peso desses elementos que estão inseridos no objeto em estudo. O PGNAA é não destrutivo, amplamente utilizado para análise elementar de amostras, tais como: (1) análise de carvão e minerais (CHANG et al., 2019); (2) exploração de poços de petróleo (MERIÇ, 2012; WANG; LI; GARDNER, 2008); (3) medição natural de radioatividade (BALDONCINI et al., 2018; CACIOLLI et al., 2012; MORSY; EL-WAHAB; EL-FARAMAWY, 2012); (4) medições em ambientes marinhos (ELEFThERIOU et al., 2013). Outra aplicação de PGNAA é sua utilização para detecção e identificação de explosivos e drogas ilícitas (KANAWATI et al., 2010; KANAWATI et al., 2011).

As determinações das frações em peso que constituem uma amostra em estudo foram abordadas com sucesso pelo método do *Monte Carlo Library Least Squares* (MCLLS) (GARDNER; XU, 2009) e pela técnica de análise das regiões de máxima intensidade espectral (picos) (GARDNER et al., 1997b). O MCLLS é composto pelo módulo do *Monte Carlo* (MC), responsável por simular a configuração experimental e gerar as bibliotecas dos elementos da amostra, e pelo sistema dos Mínimos Quadrados (*Library Least Squares*, LLS), encarregado por realizar a combinação linear das bibliotecas na tentativa de representar a amostra em estudo e gerar as frações em peso dos elementos dessa amostra. Conforme estudado por Gardner e Xu (2009), há evidências que a abordagem do MCLLS tem uma vantagem significativa sobre a técnica de análise de pico, já que o MCLLS faz uso de todo o espectro de raios gama em vez de focar apenas nas particularidades dos picos, fornecendo uma análise quantitativa mais precisa (GARDNER et al., 1997a).

O MCLLS propôs uma correção iterativa para um problema inerentemente não linear, como o PGNAA (GARDNER; XU, 2009). A fundamentação principal da abordagem do MCLLS é que a amostra em estudo pode ser definida como uma combinação linear das bibliotecas que compõem essa amostra, com as frações em peso realizando o papel dos coeficientes da combinação. Essa abordagem foi utilizada para muitos fins, incluindo fluxo multifásico de

hidrocarbonetos (WANG; LI; GARDNER, 2008);(MERIÇ, 2012). Maiores detalhes sobre o MCLLS e sua implementação pode ser encontrada nos seguintes trabalhos: Han e Gardner (2006), Han, Gardner e Metwally (2007), Shyu, Gardner e Verghese (1992), Meriç (2012).

Os principais desafios na análise de espectro por MCLLS são: (1) a sobredeterminação do problema, uma vez que o número de frequências de energia, ou canais, é maior que os constituintes inseridos na amostra (MERIÇ, 2012); (2) radiações provenientes do espalhamento da radiação gama (background) (MERIÇ, 2012); (3) elevado custo computacional para as simulações de MC (MAUCEC; RIGOLLET, 2004); (4) instabilidades numéricas no ajuste dos mínimos quadrados devido à inversão de matrizes no sistema do LLS.

Vários métodos foram propostos para superar essas dificuldades: (1) minimizar o número condicionamento das matrizes de covariância (MERIÇ, 2012); (2) aplicação do método dos mínimos quadrados não negativos a espectros de raios gama (BALDONCINI et al., 2018; BALDONCINI, 2010); (3) utilização de algumas janelas de energia das bibliotecas (ANDROULAKAKI et al., 2016), em contraposição ao uso de todos os canais de energia das bibliotecas. Alguns outros métodos que também merecem destaque são: (1) Algoritmos fireworks (ALAMANIOTIS; TSOUKALAS, 2018); (2) otimização por enxame de partículas, ou *Particle Swarm Optimization* - Otimização por enxame de partículas (PSO) (SHAHABINEJAD; VOSOUGHI, 2018); (3) técnicas baseadas em algoritmos genéticos (ALAMANIOTIS; MATTINGLY; TSOUKALAS, 2013); (4) abordagens por redes neurais (YOSHIDA et al., 2002); (5) Sistemas baseados em lógica Fuzzy (ALAMANIOTIS; CHOI; TSOUKALAS, 2015). Esses métodos, essencialmente, tinham por finalidade identificar as bibliotecas de uma amostra e realizar a devida análise quantitativa, ou determinar as frações em peso.

A inversão de matrizes utilizadas pelo sistema do LLS ocasiona a inserção de instabilidades numéricas devido ao mau condicionamento, implicando em soluções sensíveis a pequenas perturbações no sistema de equações lineares, introduzindo incertezas não desprezíveis nos resultados numéricos (ASTER; BORCHERS; THURBER, 2018; MERIÇ, 2012). Na abordagem do MCLLS, as instabilidades ocorrem principalmente por:

- Duas ou mais bibliotecas que são linearmente correlacionadas, implicando na dependência linear das colunas da matriz;
- Um espectro de biblioteca que tem uma contribuição irrelevante para o espectro total de raios gama. As taxas de contagem de uma biblioteca são muitas ordens de magnitude menor que os valores de contagem de outra biblioteca.

Este trabalho descreve uma abordagem de otimização do passo do LLS, sub-etapa do MCLLS, no intuito de enfrentar a sobre-determinação do sistema de equações, e ao mesmo tempo, determinar adequadamente as frações em peso dos elementos inseridos em uma amostra. A otimização será realizada pelo uso dos algoritmos Greedy Randomized Adaptative Procedure (GRASP) e Continuous Greedy Randomized Adaptative Search Procedure (CGRASP). A otimização levará em conta uma otimizada escolha dos dados de entrada, como também partirá da premissa de que alguma biblioteca não foi devidamente informada ao sistema do LLS pelo módulo do MC. Esse elemento será estimado ao longo do processo de otimização por uma nova abordagem que envolve uma proposta geométrica. A escolha para o algoritmo deve-se ao sucesso do mesmo na minimização do condicionamento das matrizes de covariância, realizando uma devida modelagem do Algoritmo GRASP ao problema do PGNA, presente no trabalho de Cardoso et al. (2020).

A abordagem é baseada em particionamento dos espectros, determinando as regiões que contenham os picos das bibliotecas conhecidas. Após a determinação dessas regiões, também conhecida por partições, é realizada a seleção dos canais nas partições para compor uma nova solução, concomitantemente às estimativas das intensidades espectrais de uma das bibliotecas, omitida nos passos iniciais do GRASP. Assim, procura-se fazer com que o conjunto de bibliotecas consiga descrever, o mais fielmente possível, o espectro total original.

Os resultados alcançados pelo presente trabalho indicam ser possível particionar os dados devidamente, com a consequente obtenção das matrizes de covariância $n \times n$, onde n é o número de bibliotecas, com números de condicionamento muito próximos ao das matrizes de dimensões 2×2 amplamente divulgadas no estado da arte. Além disso, os estudos quantitativos determinam um erro médio 3.16% em todas as instâncias estudadas, enquanto os principais métodos aplicados no âmbito do LLS do MCLLS apresentaram um erro médio nas determinações quantitativas da ordem de 8.80%. As bibliotecas estimadas apresentaram um erro médio de 4.22% nas instâncias apresentadas, não encontrando paralelo para essa iniciativa, até o presente momento, no estado da arte.

Os métodos aqui apresentados mostraram-se capazes de enfrentar as principais dificuldades no contexto do MCLLS aplicado no âmbito do PGNA. Além disso, também foram capazes de modelar adequadamente os princípios do Algoritmo GRASP ao domínio do problema, mostrando-se como uma nova alternativa às várias abordagens apresentadas na literatura para otimização do LLS.

1.1 MOTIVAÇÃO E JUSTIFICATIVA

As simulações de MCLLS apresentam a vantagem de eliminar as necessidades de calibrações experimentais e uso de amostras padrões. Essas simulações possuem elevados custos computacionais (MAUCEC; RIGOLLET, 2004). Esses custos podem ser minimizados implementando otimizações em nível do LLS do MCLLS. Melhorias na abordagem do condicionamento das matrizes de covariância presentes no trabalho de Meriç (2012) podem otimizar o processo, assim como propostas no particionamento das bibliotecas que descrevem o espectro total podem contribuir em melhorias ao processo do MCLLS.

Muitos métodos já foram efetivamente utilizados no intuito de enfrentar os problemas do MCLLS. No nosso trabalho, uma das motivações foram os resultados alcançados em Cardoso et al. (2020), em que a otimização das frações de contagem percentual alcançaram resultados satisfatórios, a ponto de estender a abordagem e passar a estimar, também, um biblioteca que não tenha sido fornecida inicialmente pelo módulo do MC ao sistema do LLS.

No trabalho citado de Cardoso et al. (2020), foi possível obter matrizes de covariância de dimensão 2×2 , com valores de condicionamento extremamente baixo, comparados com os trabalhos de Meriç (2012). Essa matrizes foram capazes de prover qualidade na determinação das frações em peso dos elementos inseridos em uma amostra em estudo. Essa abordagem levou em conta particionamento dos dados de entrada, obedecendo a critérios de independência linear, para posterior escolha de canais dentro dessas partições para construção das matrizes de covariância.

Alguns outros trabalhos que também contribuíram nessa proposta de extensão do trabalho de Cardoso et al. (2020) foram: (1) estudo quantitativo de frações em peso, no âmbito do LLS, a partir de matrizes de dimensão 2×2 (MERIÇ, 2012); (2) Extensão do trabalhos apresentado por Meriç (2012) para abordagem de bibliotecas que apresentassem intensidades espectrais residuais quando comparado as demais bibliotecas (CHANG et al., 2019); (3) particionamento de um conjunto de bibliotecas de entrada, na descrição de um espectro total, levando em conta o uso de conjuntos ativos e passivos e de métrica de adequação do espectro estimado ao espectro total (SHAHABINEJAD; VOSOUGHI, 2019); (4) métodos de particionamento de bibliotecas e de canais, na obtenção dos picos das bibliotecas (ANDROULAKAKI et al., 2016); (5) métodos de otimização dos LLS, do MCLLS, por meio de abordagens de enxame de partículas (SHAHABINEJAD; VOSOUGHI, 2018).

As novas abordagens a serem investigadas, relativamente ao trabalho de Cardoso et al.

(2020), passam pelos seguintes passos: (1) particionamento dos dados pela determinação das regiões em que existem os picos das bibliotecas, e não mais por determinação de regiões que apresentem independência linear. Essa abordagem permite que se escolha uma região representativa para cada uma das bibliotecas; (2) estimativas de uma das bibliotecas contidas no espectro total, mas não fornecida inicialmente pelo MC, nas regiões determinadas no passo um, por nova abordagem geométrica aplicada ao LLS. Essa biblioteca estimada será associada a uma partição, determinada pela região dos dados que apresenta a maior diferença entre o espectro total real e o espectro definido pelas bibliotecas conhecidas; (3) continuidade dos princípios de otimização dos canais inseridos nas partições, com os passos de busca e linkagem moldados de maneira similar ao de Cardoso et al. (2020), acrescidos da modelagem para otimização da biblioteca estimada nas partições; (4) estudo das matrizes de covariância $n \times n$, em que n é o número de bibliotecas estudadas no problema.

1.2 ESTADO DO PROBLEMA

A despeito da premissa de linearidade no MCLLS, PGNA é um problema inerentemente não linear (MERIÇ, 2012). Essa não linearidade deve-se, principalmente, ao espalhamento de nêutrons e atenuação de raios gama, levando a erros graves nas análises quantitativas (MERIÇ, 2012). Isso significa que as contagens imediatas de raios gama de um único constituinte dependerão não apenas da quantidade desse constituinte, mas também das quantidades de outros elementos devidos a tais efeitos.

As principais etapas do MCLLS envolvem (CHANG et al., 2019):

1. Designa-se um conjunto de constituintes iniciais assumidos como geradores de uma composição que se pense ser a mais próxima possível da composição real da amostra em estudo. Prepara-se uma representação MC para essa amostra aproximada, com as frações de pesos iguais para todos os constituintes.
2. Submete-se ao método MC a representação da amostra aproximada, no contexto da representação do arranjo experimental. O MC então retorna as bibliotecas de energia de cada constituinte.
3. As bibliotecas dos constituintes são submetidas a um sistema de equações lineares sobredeterminado (Mínimos Quadrados-LLS), tendo a biblioteca da amostra desconhecida como a coluna de termos independentes do sistema. Os coeficientes retornados pelo

sistema LLS são aproximações para as frações de peso dos constituintes na composição da amostra em estudo.

4. Um espectro aproximado para a amostra em estudo é construído tomando-se as bibliotecas dos constituintes e atenuando-as via multiplicação pelos respectivos coeficientes retornados pelo LLS, e somando-as. Se o espectro da amostra aproximada ainda divergir por larga margem do espectro da amostra em estudo, prepara-se a representação MC da amostra aproximada, utilizando-se os coeficientes do LLS como frações de peso na composição, e segue-se para o Passo 2. Senão, segue-se para o Passo 5.
5. Retornam-se as frações de peso determinados no Passo 3.

A abordagem descrita acima sugere um processo iterativo, no contexto do MCLLS, entre as etapas do MC e sistema do LLS. Na ocorrência de ineficiência por parte do sistema do LLS, uma nova chamada à etapa do MC é realizada para representação da amostra aproximada, implicando em uma carga experimental elevada (MAUCEC; RIGOLLET, 2004).

Muitas abordagens foram propostas ao sistema do LLS no intuito de prover uma maior qualidade às aproximações do MCLLS. Uma delas é minimizar as dimensões das matrizes de covariância a fim de mitigar as instabilidades numéricas associadas ao sistema, por meio da soma de bibliotecas (MERIÇ, 2012). Porém, mesmo com a redução da dimensionalidade das matrizes de covariância, ainda acontece a questão da sobre-determinação dos sistema de equações associado ao sistema do LLS. Sobre-determinação de equações ainda tende a ser uma fonte de instabilidades numéricas associadas ao problema.

Como alternativa às técnicas que fazem combinação de bibliotecas, pode-se enumerar a aplicação de algoritmos evolucionário, ao sistema do LLS, bastante consolidados: aplicação de algoritmos de enxame de partículas (SHAHABINEJAD; VOSOUGHI, 2018), em que a população de soluções é avaliadas, e essa população é modificada, segundo uma velocidade e uma posição, mantendo-se as melhores soluções; aplicação de Lógica Fuzzy e de Algoritmos Genéticos, em que os indivíduos trocam informações no intuito de se conseguir descendentes de melhor qualidade (ALAMANOTIS; JEVREMOVIC, 2015); determinação de regiões mais relevantes do espectro, determinando picos, e realizando a descrição em função das bibliotecas que se encontram nessas regiões (ANDROULAKAKI et al., 2016).

Uma desvantagem desses métodos acima é que eles não consideram todas as bibliotecas que efetivamente foram utilizadas na geração do espectro da amostra em estudo. Esses métodos

avaliam a qualidade de uma iteração arbitrária por meio de um teste de adequação que pode ignorar bibliotecas na descrição do espectro total.

Este trabalho descreve uma abordagem de otimização do passo do LLS, sub-etapa do MCLLS, no intuito de enfrentar sistemas de equações que apresentam sobre-determinação e, ao mesmo tempo, determinar adequadamente as frações em peso de todos os elementos inseridos em uma amostra em estudo, fazendo uso de algoritmos heurísticos e métricas adequadas para esse fim, acarretando menos chamadas ao módulo do MC.

1.3 OBJETIVOS

O objetivo principal deste doutorado é otimizar o módulo dos Mínimos Quadrados do MCLLS, selecionando adequadamente seções dos dados de entrada e estimando dados não fornecidos inicialmente pelo módulo do MC, de modo a tornar o processo mais preciso nas determinações das frações de contagem percentual nas saídas dos Mínimos Quadrados.

O principal objetivo é alcançado por meio dos seguintes objetivos específicos:

- Adaptar os conceitos GRASP e CGRASP ao domínio do problema de modo que ocorra uma condução eficiente do algoritmo no espaço de soluções disponível.
- Estimar adequadamente uma substância presente no espectro total, mas inicialmente desconhecida, por meio de modelos geométricos apropriados e inovadores, nunca antes propostos por métodos que objetivam a melhoria do LLS no contexto do MCLLS.
- Particionar os dados, resgatando os conceitos de picos, objetivando trabalhar com seções do mesmo no intuito de manipular apenas as seções mais importantes, mitigando questões como as diferenças de magnitude nas taxas de contagem das bibliotecas individuais na pesquisa do LLS.
- Analisar matrizes de covariância de maior dimensão que as publicadas no estado da arte.

1.4 CONTRIBUIÇÕES

As contribuições realizadas durante o trabalho são:

- Atingir maior precisão nas saídas dos Mínimos Quadrados, ou seja, aproximar as saídas dos mínimos quadrados dos verdadeiros valores das frações em peso, comparável com

trabalhos do estado da arte. Essa contribuição será essencial no sentido de realizar menos chamadas ao módulo do MC pelo sistema do LLS.

- Modelar o domínio do problema ao contexto dos algoritmos GRASP e CGRASP para otimização das frações de contagem percentual das bibliotecas, assim como na estimativa e aproximação da biblioteca omitida ao sistema do LLS.
- Determinar a estimativa e realizar a otimização dessa estimativa de biblioteca inserida no espectro total, mas que não tenha sido mapeada pelo processo do MC, por meio de uma abordagem geométrica aplicada ao contexto do processo do MCLLS. Essa ideia destaca-se como uma contribuição nunca antes proposta por abordagens que visam melhorias no contexto do MCLLS, oferecendo novas frentes de pesquisa e desenvolvimento à técnica do MCLLS.
- Determinação de matrizes de covariância de maior dimensão, com o menor número de condicionamento possível, como uma abordagem alternativa à proposta por (MERIÇ, 2012).

1.5 ESTRUTURA DA PROPOSTA

Seção 02 - Fundamentação.

Nesse capítulo serão trazidos os principais conceitos que serão utilizados diretamente nesta tese, como:

1. Conceitos sobre PGNAA;
2. Definição e implicações do mau condicionamento das matrizes de covariância;
3. As definições iniciais sobre a heurística utilizada no presente trabalho;
4. Definições matemáticas: número de condicionamento; *Singular Value Decomposition* - Decomposição em Valores Singulares (SVD); Método dos Mínimos Quadrados (MMQ);
5. Conceitos sobre a abordagem do MCLLS ao PGNAA;
6. Fundamentação dos Algoritmos utilizados nesta obra, o GRASP e CGRASP.

Seção 03 - Análise Espectral e o Estado da Arte.

São analisadas as principais publicações da área, além de ser realizada a descrição das características de cada um dos trabalhos apresentados nessa seção. Também serão abordadas, quando possível, as principais contribuições ao campo do PGNA, como: velocidade computacional, precisão das frações em peso, problemas enfrentados e as principais limitações.

Seção 04 - Metodologia.

O método proposto é apresentado nesse capítulo. A descrição do método iniciará por uma visão mais geral, para posterior definição das etapas intermediárias que o compõem. O Algoritmo geométrico será explicado nessas subetapas em maiores detalhes, assim como as saídas dessa etapa serão exploradas pelas etapas básicas do GRASP e pelo ajuste do CGRASP da biblioteca omitida do espectro total. O algoritmo proposto no presente trabalho manuseia variáveis em um espaço discreto (canais de energia são indexados por variáveis inteiras) e grandezas definidas em espaços contínuos (as contagens espectrais da biblioteca omitida do espectro total).

Seção 05 - Resultados

Nesse capítulo, abordam-se os resultados aplicados a duas bases de dados provenientes das simulações de MC: uma delas formada por elementos de Am, Co, Cs e Na. A Base 2 é formada pelos compostos *Ar*, *Água* e *Óleo*. Os resultados são avaliados por meio das diferenças de áreas existente entre o espectro total fornecido e o espectro estimado pelo algoritmo do presente trabalho. O referido capítulo abordará a discussão dos resultados.

Seção 06 - Conclusões e Trabalhos Futuros.

Serão apresentadas as conclusões, trabalhos publicados durante a pesquisa e propostas de trabalhos futuros. Também são discutidos os pontos fortes da atual proposta para que possa ocorrer uma continuidade dos atuais trabalhos, como também as falhas para que sejam alvo de investigações futuras.

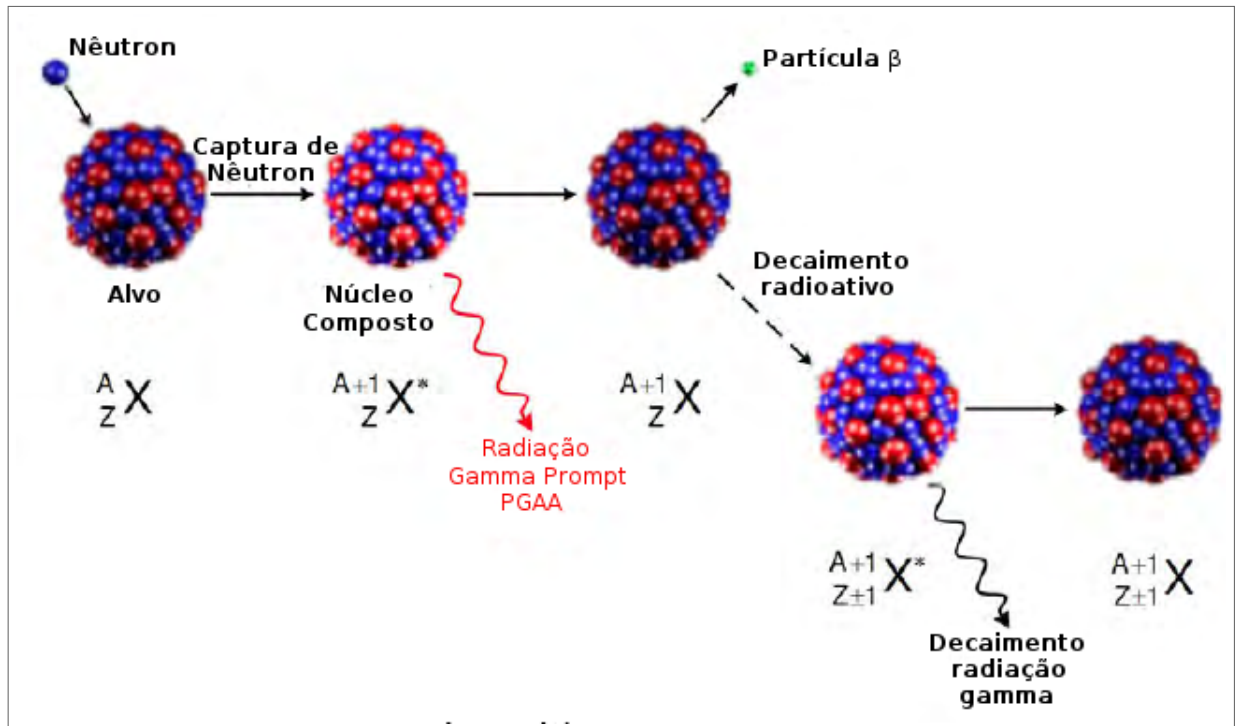
2 FUNDAMENTAÇÃO

O presente capítulo anuncia as fundamentações utilizadas no trabalho: (1) na Seção 2.1 são detalhados os mecanismos físicos e as principais abordagens metodológicas aplicadas ao problema do PGNA; (2) uma breve definição sobre condicionamento de um sistema de equação linear é fornecida na Seção 2.2; (3) a definição matemática de SVD é apresentada no capítulo 2.3; (4) a relação entre condicionamento e a decomposição em valores singulares aparece na Seção 2.4; (5) a fundamentação dos Mínimos Quadrados é determinada na Seção 2.5; (6) os detalhamentos do MCLS são descritos na Seção 2.6; (7) os algoritmos utilizados na tese são descritos na Seção 2.7 e 2.8.

2.1 PGNA

O uso de PGNA, baseada em feixe de nêutrons, é uma técnica rápida e não destrutiva, largamente utilizada para análise de amostras (BALDONCINI et al., 2018; CACIOLLI et al., 2012; MORSY; EL-WAHAB; EL-FARAMAWY, 2012). Pode-se sumarizar pelos seguintes passos: (1) a amostra que se deseja analisar é submetida a uma fonte de nêutrons; (2) essa amostra captura o nêutron emitido, tornando-se um núcleo emissor de radiação gama. Essa radiação emitida nessa etapa é o gama pronto; (3) a amostra, antes de retornar ao seu estado estável, ainda emite partículas β . Essa pode contribuir com ruído ao processo de análise do PGNA; (4) após a emissão de partículas β , nova radiação gama pode ser emitida; (5) a amostra em estudo retorna ao seu estado estável. Todo o processo pode ser sumarizado na Figura 1. Os principais mecanismos físicos aplicados no PGNA para determinação da composição nuclear são espalhamento inelástico e captura de nêutrons (ZHANG; TUO, 2013; ZHANG et al., 2015).

Figura 1 – Ilustração do decaimento. Absorção do nêutron induz à emissão de raios gama. Raios gama atrasados podem ser emitidos, antes de o núcleo retornar à estabilidade. (SÖLLRADL, 2014)



Fonte: (SÖLLRADL, 2014)

Alguns exemplos que podem ser citados a respeito da aplicação da técnica de análise por PGNA: (1) determinação de concentrações de elementos em metais de amostras geotérmica de fundo de oceano (WILDE, 2002) e também em amostras biológicas (ROSSBACH; HIEP, 1992); (2) uso em Spsychala, Michaelis e Fanger (1987) para determinar concentrações de elementos Na, Al, Si, K, Ca, Ti, Mn, Fe, Zn, C e S em amostras sedimentares, assim como os constituintes de Cr, Ni, Cu, Cd e Hg; (3) em Becker et al. (1994), determinam-se os dados de fração de massa para H, B, N, K, Ca, S, Cl, Cd, Sm e Gd na certificação de folhas de tomate; (4) no trabalho de Yonezawa, Matsue e Hoshi (1997), investigam-se as sensibilidades de 21 elementos em diversos materiais, como rochas, amostras biológicas, sedimentos, carvão, cabelo humanos e outras amostras ambientais no reator de PGNA; (5) identificação de concentrações de B por meio de análise de ativação gama de captura de nêutrons e materiais de referência biológicas pode ser conferido em Anderson, Cunningham e Mackey (1990); (6) por fim, a utilização de PGNA na determinação *on-line* de *calcário*, *sílica*, *alumínio* e *óxido de ferro* na fabricação de cimentos.

O espalhamento não resulta apenas nas interações de fótons com elétrons livres, mas

também com elétrons atômicos ligados. O pico pode ser considerado como a região do espectro em que os fótons interagem com os elétrons livres. A radiação espalhada na amostra, pode, portanto, ser considerada uma nova fonte de informações sobre uma propriedade da própria amostra (BODE et al., 2008), além da interação de fótons com elétrons livres. O Contínuo de Compton é constituído pela radiação espalhada por meio desses dois mecanismos.

Duas abordagens podem ser usadas para análise de raios gama na determinação das frações em peso dos elementos de uma amostra: (1) o método de análise de picos (GARDNER et al., 1997b) e (2) análise de espectro total (*Whole Spectrum analysis* - Análise de Espectro Total (WSA) ou *Full Spectrum Analysis* - Análise de espectro total (FSA)) (GARDNER; XU, 2009).

No caso da análise por picos (GARDNER et al., 1997b), a análise quantitativa, ou as frações em peso dos constituintes, são calculadas com base no número de contagens nos picos mais intensos dos espectros. Esse método possui a desvantagem de necessitar de um detector com alta resolução, além de não considerar a perda de informação do Contínuo de Compton. Na presença de diversas substâncias absorvendo nêutrons e emitindo radiação, simultaneamente, o sinal adquirido, no âmbito da análise por picos, exibe um alto grau de complexidade porque consiste em uma combinação de contagens provenientes diversos radionuclídeos. A medição de um espectro pode estar contaminado com radiações de fundo provenientes de raios cósmicos (SHULTIS; FAW, 2016).

No âmbito do método da análise de espectro inteiro (WSA), o estudo da radiação é realizado tomando em conta os picos do espectro e o Contínuo de Compton, ou seja, todo o espectro é utilizado na análise da radiação. Esse método municia vantagens caso a qualidade do detector seja um fator limitante (KNOLL, 2000). O ajuste do espectro medido é definido como uma combinação linear, para os diferentes isótopos (SALMON, 1961; SHAHABINEJAD; VOSOUGHI, 2018), com as frações em peso realizado o papel dos coeficientes na combinação linear. A fundamentação principal para a abordagem do PGNAa no âmbito do WSA será fornecida na Seção 2.6.2.

Grande parte dos materiais radioativos emitem radiação gama em uma janela específica de intensidades de energia. A coleção desses raios gama permite a definição de espectros característicos de um nuclídeo, possibilitando, assim, a definição da “impressão digital” do isótopo no domínio da radiação gama (TSOULFANIDIS, 2010). A análise espectroscópica trata da extração de informações dos sinais de raios gama, usando reconhecimento em algoritmos, e fazendo uma inferência sobre a identidade e propriedades dos isótopos radioativos (ALAMANIOTIS; JEVREMOVIC, 2015). Considerando que a seção transversal para espalhamento inelástico é

pequena, a análise por PGNAA está concentrada essencialmente na captura de nêutrons térmicos (ZHANG; TUO, 2013). Nêutrons de baixa energia são absorvidos pelo núcleo alvo, formando o núcleo composto, que tem energia igual à energia cinética dos nêutrons mais a energia de ligação de captura de nêutrons. Núcleos compostos emitem γ fóton para atingirem o estado fundamental em $10^{-3}s$.

Muitas metodologias aplicadas na análise por PGNAA adotam uma variedade de ferramentas computacionais para analisar espectros de raios gama obtidos por detectores. Em Salmon (1961), *Multiple Linear Regression* (MLR) foi utilizada para resolver o sistema sobredeterminado a partir da aproximação FSA, apesar de ter gerado valores negativos de frações percentuais, matematicamente inadmissível. A dificuldade em estimar as verdadeiras frações percentuais ocorrem devido ao fato de o problema inverso ser mal condicionado (MERIÇ, 2012). A superação dessa dificuldade tem sido proposta por vários métodos: minimizar o condicionamento da matriz de covariância (MERIÇ, 2012); aplicação do *Non Negative Least Squares* - Mínimos Quadrados Não Negativos (NNLS) (BALDONCINI, 2010; BALDONCINI et al., 2018; CACIOLLI et al., 2012); execução do FSA em várias diferentes janelas de energia (ANDROULAKAKI et al., 2016) e controle do processo de ajuste dos raios gama por um módulo de comparação (ALAMANIOTIS; CHOI; TSOUKALAS, 2015; SHAHABINEJAD; VOSOUGHI, 2018).

Alguns algoritmos heurísticos e evolucionários foram aplicados ao problema do PGNAA, dentre eles pode-se citar: *Genetic Algorithms* - Algoritmos Genéticos (GAs) (ALAMANIOTIS; MATTINGLY; TSOUKALAS, 2013; CARLEVARO; WILKINSON; BARRIOS, 2008); *Hybrid Fuzzy Genetic* (HFG) (ALAMANIOTIS; CHOI; TSOUKALAS, 2015); *Fireworks Algorithm* (FA) (ALAMANIOTIS; CHOI; TSOUKALAS, 2015). Todas as técnicas são aplicadas no intuito de superar as dificuldades do clássico MLR na análise de espectros de raios gama.

No contexto do uso de técnicas de inteligência artificial, redes neurais foram utilizadas na análise de espectros de raios gama (KANGAS et al., 2008; YOSHIDA et al., 2002), a despeito da dificuldade em processar entradas longas. Pode-se citar também a aplicação de algoritmos genéticos na análise desses espectros como descrito em Alamaniotis, Mattingly e Tsoukalas (2013), Carlevaro, Wilkinson e Barrios (2008), Garcia-Talavera e Ulicny (2003). Os algoritmos baseados em redes neurais montam modelos de nuclídeos, enquanto que os baseados em algoritmo genético focam no encapsulamento de radionuclídeos como descrito em Diver e Ireland (1997).

Metodologias para adaptação de picos, integração de mínimos quadrados e algoritmos de buscas aleatórias para avaliação dos parâmetros de montagem são encontradas em Kitamura et

al. (1976), enquanto uma aproximação apropriada adotando Levenberg-Marquardt é apresentada em Fatah e Ahmed (2011). As funções de base *B-spline* são derivadas para ajuste do espectro para reduzir flutuações estatísticas em Meng-Hua et al. (2009), fornecendo resultados precisos.

Metodologias desenvolvidas para análise de radiação inversa são descritas em Favorite (2004), Mattingly e Mitchell (2010). Adicionalmente, análise baseada em wavelet aplicada a espectros raios gama é apresentada em Sullivan, Martinez e Garner (2006), enquanto que uma metodologia de ajuste de fotopico é introduzido em Chen e Nanping (2013). Descrição de regressão linear múltipla pode ser encontrada em Burr e Hamada (2009), com aplicação na espectroscopia de raios gama em Pracilio et al. (2006).

Métodos estatísticos são apresentados em Meng e Ramsden (2000). Algoritmos de identificação foram desenvolvidos incluindo estatística Bayesiana em Stinnett e Sullivan (2013); testes estatísticos de razão de verossimilhança (FORSBERG et al., 2009); Fireworks Algorithm (ALAMANIOTIS; CHOI; TSOUKALAS, 2015); algoritmos de deconvolução (SHAHABINEJAD; VO-SOUGHI, 2019) e, por fim, algoritmos baseados em integração de coeficientes de correlação com avaliação baseada em lógica fuzzy, como em Alamaniotis et al. (2009). Esses métodos são capazes de encontrar as boas soluções em um grande espaço de solução multidimensional sem ficar preso em ótimos locais (MUKHERJEE, 2002; SHAHABINEJAD; HOSSEINI; SOHRABPOUR, 2016; SHAHABINEJAD; SOHRABPOUR, 2017).

2.2 CONDICIONAMENTO DE UM SISTEMA

O condicionamento de um problema reflete a sensibilidade da solução exata ser transformada após uma perturbação nos dados de entrada. Para quantificar esta ideia, assuma que o problema é definido por um conjunto de dados d . Denote $s(d)$ a solução exata do problema para os dados de entrada d . Se quaisquer pequenas transformações em d conduzem a pequenas transformações em $s(d)$, pode-se concluir que o problema é bem condicionado. De outro modo, o problema será considerado mal condicionado. Supondo que d_1 e d_2 são dois possíveis conjuntos de dados, o *Condicionamento* pode ser definido como (GILL; MURRAY; WRIGHT, 1991):

Definição 2.2.1 (Condicionamento). *O condicionamento é medido como o máximo valor da razão:*

$$\frac{\|s(d_1) - s(d_2)\|}{\|d_1 - d_2\|} \quad (2.1)$$

Um exemplo prático sobre condicionamento pode ser conferido no Apêndice A, Seção A.1. Uma consequência dessa definição será abordada pelo Teorema 2.2:

Teorema 2.2.1 (Condicionamento sistemas lineares). *Condicionamento para sistemas lineares pode ser definido por:*

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}^{-1}\| \|\mathbf{A}\| \quad (2.2)$$

A Equação 2.2 pode ser deduzida a partir da Definição 2.1. A prova matemática para o Teorema 2.2 encontra-se na Seção A.2.

Sabendo que a matriz identidade tem norma 1, então $\mathbf{I} = \mathbf{A}^{-1}\mathbf{A}$ e $\|\mathbf{A}^{-1}\mathbf{A}\| = 1 \leq \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$, conclui-se que $\text{cond}(\mathbf{A}) \geq 1$. Portanto, uma matriz bem condicionada tem condicionamento próximo de 1. Um problema mal condicionado, geralmente apresenta um condicionamento muito maior que 1.

2.3 DECOMPOSIÇÃO EM VALORES SINGULARES (SVD)

Decomposição em valores singulares (SVD) é de grande importância para cálculos numéricos. Qualquer matriz pode ser decomposta em valores singulares (ver Golub e Loan (2012)), tem grande aplicabilidade na solução de sistemas de equações sobredeterminados, principal característica da análise por radiação gama (HARTLEY; ZISSERMAN, 2004; LIMA, 2006; STRANG, 1993). O condicionamento da matriz, como será visto na seção 2.4, é fortemente relacionado aos valores singulares dessa matriz.

Dada uma matriz $\mathbf{A}$, o SVD é uma fatorização de $\mathbf{A}$ como $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$, onde $\mathbf{U}$ e $\mathbf{V}$ são matrizes ortogonais, e $\mathbf{D}$ é uma matriz diagonal com valores não negativos. A decomposição produz uma matriz diagonal $\mathbf{D}$ em que as entradas, convencionalmente, tem ordem decrescente.

A decomposição em valores singulares também pode ser realizada para matrizes não quadradas. Sendo $\mathbf{A}_{m \times n}$ matriz com dimensões $m \geq n$, $\mathbf{A}$ pode ser fatorada como $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ onde $\mathbf{U}$ é uma matriz $m \times m$, cujas colunas são ortogonais, $\mathbf{D}$ é uma matriz diagonal $m \times n$ e $\mathbf{V}$ é uma matriz ortogonal $n \times n$. Afirmar que $\mathbf{U}$ tem colunas ortogonais significa que $\mathbf{U}^T\mathbf{U} = \mathbf{I}_{n \times n}$.

As entradas da matriz diagonal $\mathbf{D}$ não são iguais aos seus autovalores. Para realizar essa

conexão entre valores singulares e autovalores, considere:

$$\begin{aligned}
 \mathbf{A} &= \mathbf{U}\mathbf{D}\mathbf{V}^T \\
 \mathbf{A}^T\mathbf{A} &= [\mathbf{V}\mathbf{D}\mathbf{U}^T][\mathbf{U}\mathbf{D}\mathbf{V}^T] \\
 &= \mathbf{V}\mathbf{D}^2\mathbf{V}^T, \quad \text{sendo } \mathbf{U}^T\mathbf{U} = \mathbf{I} \\
 \mathbf{A}^T\mathbf{A} &= \mathbf{V}\mathbf{D}^2\mathbf{V}^{-1}.
 \end{aligned} \tag{2.3}$$

então, a Equação 2.3 para a definição dos autovalores, indicando que as entradas de $\mathbf{D}^2$ são os autovalores de $\mathbf{A}^T\mathbf{A}$ e as colunas de $\mathbf{V}$ são os autovetores de $\mathbf{A}^T\mathbf{A}$. Em suma, os valores singulares de $\mathbf{A}$ são as raízes quadradas dos autovalores de $\mathbf{A}^T\mathbf{A}$.

Como $\mathbf{A}^T\mathbf{A}$ é simétrica e definida semi-positiva, os autovalores são reais e não negativos.

2.4 CONDICIONAMENTO E SVD

A decomposição em valores singulares fornece meios que permitem analisar o condicionamento de uma matriz. Essa seção abordará uma visão a respeito da relação que o condicionamento de uma matriz tem com o seus valores singulares.

Teorema 2.4.1 (Norma Matricial). *A norma de uma matriz pode ser definida por:*

$$\|\mathbf{A}\| = \max \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}, \quad \mathbf{x} \neq \vec{0} \tag{2.4}$$

Considere a seguinte razão:

$$\mu(\mathbf{A}, \mathbf{x}) = \frac{\|\mathbf{A}\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = \frac{(\mathbf{A}\mathbf{x})^T \mathbf{A}\mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \tag{2.5}$$

Devido ao fato que $\mu(\mathbf{A}, \gamma\mathbf{x}) = \mu(\mathbf{A}, \mathbf{x})$ para qualquer escalar não nulo, só a direção de $\mathbf{x}$ é que afetará μ .

Qualquer vetor que maximiza $\mu(\mathbf{A}, \mathbf{x})$ também deve dar o máximo valor para $\frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$ e para $\|\mathbf{A}\|_2$. O subscrito 2 está sendo aqui colocado para definir a norma l_2 .

Será mostrado que o SVD determina o vetor $\mathbf{x}$ para o qual o valor máximo de $\mu(\mathbf{A}, \mathbf{x})$ é atingido. Considere $\mathbf{v}_1, \dots, \mathbf{v}_n$ representando as n colunas de $\mathbf{V}$, que são ortonormais e geram $\mathbb{R}^n$. Qualquer vetor $\mathbf{x}$ pode unicamente ser representado como uma combinação linear dessas colunas:

$$\mathbf{x} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_n \mathbf{v}_n = \mathbf{V}[\mathbf{x}]_\beta \tag{2.6}$$

onde $[\mathbf{x}]_\beta = (\alpha_1, \dots, \alpha_n)^T$, vetor de coordenadas de $\mathbf{x}$ relativo à base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Será desenvolvida uma expressão para $\mu(\mathbf{A}, \mathbf{x})$ em termos de $[\mathbf{x}]_\beta$, utilizando para isso as propriedades de matrizes e SVD.

Usando Equação 2.6 e a ortogonalidade de $\mathbf{V}$, tem-se

$$\mathbf{x}^T \mathbf{x} = [\mathbf{x}]_\beta^T \mathbf{V}^T \mathbf{V} [\mathbf{x}]_\beta = [\mathbf{x}]_\beta^T [\mathbf{x}]_\beta = \sum_{i=1}^n \alpha_i^2 \quad (2.7)$$

Desde que $\mathbf{V}$ seja ortogonal e $\mathbf{S}$ seja diagonal, o vetor $\mathbf{Ax}$ pode ser expresso como uma combinação linear das colunas de $\mathbf{U}$, utilizando-se para isso da definição de SVD presente na Seção 2.3 ($\mathbf{A} = \mathbf{USV}^T$):

$$\begin{aligned} \mathbf{Ax} &= \mathbf{USV}^T (\mathbf{V}[\mathbf{x}]_\beta) \\ &= \mathbf{US}[\mathbf{x}]_\beta \\ &= \underbrace{\begin{pmatrix} u_{11} & u_{12} & \dots & u_{1k} \\ u_{21} & u_{22} & \dots & u_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nk} \end{pmatrix}}_{\mathbf{U}} \underbrace{\begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \end{pmatrix}}_{\mathbf{S}} \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}}_{[\mathbf{x}]_\beta} \\ &= \begin{pmatrix} u_{11}\sigma_1 & u_{12}\sigma_2 & \dots & u_{1k}\sigma_n \\ u_{21}\sigma_1 & u_{22}\sigma_2 & \dots & u_{2k}\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1}\sigma_1 & u_{n2}\sigma_2 & \dots & u_{nk}\sigma_n \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \\ &= \begin{pmatrix} u_{11}\sigma_1\alpha_1 + u_{12}\sigma_2\alpha_2 + \dots + u_{1k}\sigma_n\alpha_n \\ u_{21}\sigma_1\alpha_1 + u_{22}\sigma_2\alpha_2 + \dots + u_{2k}\sigma_n\alpha_n \\ \vdots \\ u_{n1}\sigma_1\alpha_1 + u_{n2}\sigma_2\alpha_2 + \dots + u_{nk}\sigma_n\alpha_n \end{pmatrix} \\ &= \sigma_1\alpha_1[\mathbf{u}_1] + \sigma_2\alpha_2[\mathbf{u}_2] + \dots + \sigma_n\alpha_n[\mathbf{u}_n] \\ &= \sum_{i=1}^n (\sigma_i\alpha_i)\mathbf{u}_i \end{aligned} \quad (2.8)$$

onde $\mathbf{u}_i$ denota a i -ésima coluna de $\mathbf{U}$.

Quando $\mathbf{x} = \mathbf{v}_i$, o vetor $[\alpha]$ de 2.6 é $\mathbf{e}_i$, o i -ésimo vetor da base canônica, isto é,

$\begin{bmatrix} 0 & 0 & \dots & 1 & \dots & 0 & 0 \end{bmatrix}$. A Equação 2.8, então, é simplificada para:

$$\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i \quad (2.9)$$

A partir de 2.8, é possível escrever:

$$\begin{aligned} \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} &= [\mathbf{x}]_\beta^T \mathbf{S}^T \mathbf{U}^T \mathbf{U} \mathbf{S} [\mathbf{x}]_\beta \\ &= [\mathbf{x}]_\beta^T \mathbf{S}^2 [\mathbf{x}]_\beta \\ &= \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix} \\ &= \sum_{i=1}^n \alpha_i^2 \sigma_i^2 \end{aligned} \quad (2.10)$$

Substituindo 2.10 e 2.7 na 2.5, obtém-se a seguinte equação:

$$\mu(\mathbf{A}, \mathbf{x}) = \frac{\sum_{i=1}^n \alpha_i^2 \sigma_i^2}{\sum_{i=1}^n \alpha_i^2} \quad (2.11)$$

Esta razão é maximizada quando $\alpha_1 \neq 0$ e $\alpha_2 = \dots = \alpha_n = 0$, já que σ_1 é o maior valor singular, o que significa que $\mathbf{x}$ é um múltiplo não nulo de $\mathbf{v}_1$ (a primeira coluna de $\mathbf{V}$). Para $\mathbf{x} = \gamma \mathbf{v}_1$ com $\gamma \neq 0$, a partir de 2.7 $\mathbf{x}^T \mathbf{x} = \gamma^2$, e a partir da 2.10 $\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = \gamma^2 \sigma_1^2$. Segue, então, que 2.11 $\mu(\mathbf{A}, \mathbf{x}) = \sigma_1^2$. Tomando a raiz quadrada, obtém-se:

$$\max_{\mathbf{x} \neq 0} \sqrt{\mu(\mathbf{A}, \mathbf{x})} = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \sigma_1 = \|\mathbf{A}\| \quad (2.12)$$

e o máximo é atingido quando $\mathbf{x} = \gamma \mathbf{v}_1$. Logo, sabendo que $\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|$, o valor máximo para $\|\mathbf{A}\mathbf{x}\|$ acontecerá quando:

$$\|\mathbf{A}\mathbf{x}\| = \|\mathbf{A}\| \|\mathbf{x}\| \quad \text{quando} \quad \mathbf{x} = \gamma \mathbf{v}_1 \quad (2.13)$$

O maior valor singular de $\mathbf{A}$ fornece o limite superior para o incremento no comprimento euclidiano de qualquer vetor transformado por $\mathbf{A}$. Pode-se observar também que a razão $\|\mathbf{A}\mathbf{x}\| / \|\mathbf{x}\|$ é um limite inferior do maior valor singular, já que $\frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \sigma_1$.

Por outro lado, o menor valor possível da razão $\mu(\mathbf{A}, \mathbf{x})$ ocorre quando $\mathbf{x}$ é um múltiplo não nulo de $\mathbf{v}_n$ (a última coluna de $\mathbf{V}$), isto é:

$$\min_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|_2}{\|\mathbf{x}\|_2} = \sigma_n \quad \text{quando} \quad \mathbf{x} = \gamma \mathbf{v}_n, \gamma \neq 0 \quad (2.14)$$

Se os k menores valores singulares de $\mathbf{A}$ são iguais, então qualquer combinação linear de k colunas de $\mathbf{V}$ também irá produzir uma razão mínima em 2.14. Para uma matriz não singular $\mathbf{A}$, tem-se um limite inferior para a redução relativa no comprimento euclidiano de qualquer vetor não nulo transformado por $\mathbf{A}$. Assim como no máximo na 2.12, a propriedade da 2.14 implica que a razão $\frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}$ para qualquer vetor $\mathbf{x}$ não nulo é um limite superior sobre o menor valor singular de A .

A análise anterior mostra que as colunas de $\mathbf{V}$ a partir do SVD de $\mathbf{A}$ fornece um conjunto de vetores ordenados:

$$\sigma_1 = \frac{\|\mathbf{Av}_1\|}{\|\mathbf{v}_1\|} \geq \frac{\|\mathbf{Av}_2\|}{\|\mathbf{v}_2\|} \geq \dots \geq \frac{\|\mathbf{Av}_n\|}{\|\mathbf{v}_n\|} = \sigma_n \quad (2.15)$$

A primeira e última colunas de $\mathbf{V}$ evidenciam as transformações extremas quando um vetor $\mathbf{x}$ é transformado por $\mathbf{A}$. O SVD também determina a decomposição da matriz inversa. Sendo $\mathbf{U}$ e $\mathbf{V}$ matrizes ortogonais e $\mathbf{S}$ diagonal, tem-se

$$\mathbf{A}^{-1} = \mathbf{VS}^{-1}\mathbf{U}^T \quad (2.16)$$

Os valores singulares de $\mathbf{A}^{-1}$ são os recíprocos aos valores singulares de $\mathbf{A}$. Devido ao fato de os valores singulares serem positivos, pode-se verificar que o maior valor singular de $\mathbf{A}^{-1}$ é $1/\sigma_n$, e o menor é $1/\sigma_1$, e as regras de ortogonalidade de $\mathbf{U}$ e $\mathbf{V}$ são inversas.

Tomando por base 2.6, é possível escrever:

$$\begin{aligned}
\mathbf{Ax} &= \mathbf{VS}^{-1}\mathbf{U}^T(\mathbf{U}([\mathbf{x}]_\beta)) \\
&= \mathbf{VS}^{-1}[\mathbf{x}]_\beta \\
&= \underbrace{\begin{pmatrix} v_{11} & v_{12} & \dots & v_{1k} \\ v_{21} & v_{22} & \dots & v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \dots & v_{nk} \end{pmatrix}}_{\mathbf{v}} \underbrace{\begin{pmatrix} 1/\sigma_n & 0 & \dots & 0 \\ 0 & 1/\sigma_{n-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/\sigma_1 \end{pmatrix}}_{\mathbf{S}^{-1}} \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}}_{[\mathbf{x}_\beta]} \\
&= \begin{pmatrix} v_{11}(1/\sigma_n) & v_{12}(1/\sigma_{n-1}) & \dots & v_{1k}(1/\sigma_1) \\ v_{21}(1/\sigma_n) & v_{22}(1/\sigma_{n-1}) & \dots & v_{2k}(1/\sigma_1) \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1}(1/\sigma_n) & v_{n2}(1/\sigma_{n-1}) & \dots & v_{nk}(1/\sigma_1) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \\
&= \begin{pmatrix} v_{11}(1/\sigma_n)\alpha_1 + v_{12}(1/\sigma_{n-1})\alpha_2 + \dots + v_{1k}(1/\sigma_1)\alpha_n \\ v_{21}(1/\sigma_n)\alpha_1 + v_{22}(1/\sigma_{n-1})\alpha_2 + \dots + v_{2k}(1/\sigma_1)\alpha_n \\ \vdots \\ v_{n1}(1/\sigma_n)\alpha_1 + v_{n2}(1/\sigma_{n-1})\alpha_2 + \dots + v_{nk}(1/\sigma_1)\alpha_n \end{pmatrix} \\
&= (1/\sigma_n)\alpha_1[\mathbf{v}_1] + (1/\sigma_{n-1})\alpha_2[\mathbf{v}_2] + \dots + (1/\sigma_1)\alpha_n[\mathbf{v}_n] \\
&= \sum_{i=1}^n ((1/\sigma_{n+1-i})\alpha_i)\mathbf{v}_i
\end{aligned} \tag{2.17}$$

Se $\mathbf{x} = \mathbf{v}_i$, é possível escrever, lembrando que nesse caso o vetor $[\alpha]$ é o i -ésimo $\mathbf{e}_i$ vetor unitário:

$$\mathbf{A}^{-1}\mathbf{v}_i = \frac{1}{\sigma_i}\mathbf{v}_i \tag{2.18}$$

Conhecendo, a partir de 2.10, que $\mathbf{x}^T\mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{x} = [\mathbf{x}_\beta]^T\mathbf{S}^{-2}[\mathbf{x}_\beta] = \sum_{i=1}^n \alpha_i^2 (\frac{1}{\sigma_i})^2$ e executando o mesmo procedimento da 2.12 e da 2.13, tem-se:

$$\max_{x \neq 0} \frac{\|\mathbf{A}^{-1}\mathbf{x}\|}{\|\mathbf{x}\|} = \|\mathbf{A}^{-1}\| = \frac{1}{\sigma_n} \tag{2.19}$$

O valor máximo é atingido quando $\mathbf{x} = \gamma \mathbf{v}_n$. Combinando 2.19, 2.12 e ??, é possível obter:

$$\text{cond}(A) = \frac{\sigma_1}{\sigma_n} \quad (2.20)$$

ou seja, a razão entre o maior e o menor valor singular da matriz $\mathbf{A}$ determina o condicionamento de $\mathbf{A}$.

2.5 MÍNIMOS QUADRADOS (MMQ)

Um sistema linear $m \times n$, $\mathbf{Ax} = \mathbf{b}$ é dito inconsistente se não tiver solução. Sistemas de equações lineares inconsistentes aparecem em muitas situações e é importante determinar uma aproximação de solução para tais sistemas. A principal abordagem consiste em mudar o problema, de modo que a equação $\mathbf{Ax} = \mathbf{b}$ não precise ser refeita. Esse mesmo tratamento será dado ao problema enfrentado no presente trabalho. Um sistema linear $m \times n$, $\mathbf{Ax} = \mathbf{b}$, é consistente se, e somente se, $\mathbf{b}$ pertence ao subespaço gerado pelas colunas de A .

Procura-se um vetor $\hat{\mathbf{x}}$ em $\mathbb{R}^n$ tal que $\mathbf{A}\hat{\mathbf{x}}$ esteja mais próximo possível da solução $\mathbf{b}$. Sendo o espaço coluna de $\mathbf{A}$ dado por W , então:

Teorema 2.5.1. *(Vetor mais próximo de um subespaço)(ANTON; RORRES, 2001) Sejam W e $\mathbf{v}$, respectivamente, um subespaço e um vetor de $\mathbb{R}^n$. O vetor $\mathbf{w} \in W$ tal que $\|\mathbf{v} - \mathbf{w}\|$ é mínimo é a projeção de $\mathbf{v}$ sobre W :*

$$\mathbf{w} = (\mathbf{v} \cdot \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{v} \cdot \mathbf{w}_2)\mathbf{w}_2 + \dots + (\mathbf{v} \cdot \mathbf{w}_m)\mathbf{w}_m \quad (2.21)$$

Diz-se que $\mathbf{w}$, neste caso, é o “vetor de W mais próximo de $\mathbf{v}$ ”.

Demonstração. Ver Seção B.1.

□

O vetor em W mais próximo de $\mathbf{b}$ é $\text{proj}_W \mathbf{b}$, isto é, $\|\mathbf{b} - \mathbf{w}\|$, para $\mathbf{w}$ em W , é mínimo quando $\mathbf{w} = \text{proj}_W \mathbf{b}$. Então, ao encontrar um vetor $\hat{\mathbf{x}}$ tal que $\mathbf{A}\hat{\mathbf{x}} = \text{proj}_W \mathbf{b}$, tem-se a certeza de que $\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|$ é o menor possível. Sabendo que $\mathbf{b} - \text{proj}_W \mathbf{b} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$ é ortogonal a todos os vetores em W (B.2), tem-se, então, que $\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$ é ortogonal a cada coluna de $\mathbf{A}$. Em termos de equação matricial, tem-se:

$$\mathbf{A}^T(\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}) = \mathbf{0} \quad (2.22)$$

ou , equivalentemente:

$$\mathbf{A}^T \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}^T \mathbf{b} \quad (2.23)$$

Qualquer solução de 2.23 é chamada de **aproximação por mínimos quadrados** para o sistema linear $\mathbf{Ax} = \mathbf{b}$. Se $\mathbf{A}$ for invertível, uma aproximação por mínimos quadrados para $\mathbf{Ax} = \mathbf{b}$ é simplesmente a solução usual $\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{b}$, já que $\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \mathbf{A}^{-1} \mathbf{A}^{-T} \mathbf{A}^T \mathbf{b} = \mathbf{A}^{-1} \mathbf{b}$.

Teorema 2.5.2. (*Posto e Solução por Mínimos Quadrados*)(KOLMAN, 2006)) Se $\mathbf{A}_{m \times n}$ é uma matriz cujo posto é $r(\mathbf{A}) = n$, então $\mathbf{A}^T \mathbf{A}$ é invertível e o sistema linear $\mathbf{Ax} = \mathbf{b}$ tem uma única aproximação por mínimos quadrados dada por $\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$, isto é, o sistema normal de equações tem uma única solução.

Demonstração. Ver Seção B.2

□

2.6 MCLLS

2.6.1 Definições

A abordagem do MCLLS ao PGNAA fundamenta-se em um processo iterativo entre dois sistemas computacionais: (1) o sistema de MC, em que se designa um conjunto de constituintes iniciais assumidos como geradores da composição mais próxima da composição real da amostra em estudo. A representação da amostra é submetida à simulação do arranjo experimental, retornando as bibliotecas de energia de cada constituinte e (2) as bibliotecas são submetidas aos Mínimos Quadrados, de posse do espectro total como a coluna de termos independentes do sistema. Os coeficientes retornados são aproximações para as frações em peso dos constituintes da amostra.

As simulações por MCLLS fornecem muitas vantagens em relação às observações e medições realizadas em ambiente de laboratório: não necessitam de maiores recursos como os requeridos por laboratório, e fornecem um visão sobre os parâmetros críticos, com estimativa das quantidades físicas de interesse, dentre eles a eficiência de parada dos raios gama nos contadores *Geiger-Muller* (GM). Elas são ferramentas computacionais que podem ser usadas para fornecer um valor médio de uma quantidade de interesse por amostragem aleatória iterativa de distribuições apropriadas, que descrevem o comportamento médio dessa quantidade (MERIÇ, 2012). O transporte de radiação através da matéria se encaixa nessa descrição, já que o caminho

de uma partícula propagando através de um absorvente é decidido com base nos resultados de vários eventos estocásticos.

A simulação de MCLLS inicia com a simulação da configuração experimental do problema na qual os valores das quantidades desejadas serão calculados. Predefinição de energia, tipo, coordenadas e direção inicial das partículas precisam ser fornecidos antes da execução da simulação. Dependendo do pacote de simulação, os materiais relevantes serão atribuídos a diferentes volumes definidos na configuração experimental do problema. Os dados de entrada são usados em conjunto com um gerador de números aleatórios para selecionar o resultado de um evento de maneira aleatória. As simulações de MCLLS necessitam de contabilidade, em que os contadores de partículas apropriados são incrementados sempre que um evento de interesse ocorre, como quando um determinado tipo de partícula cruza uma superfície. Finalmente, as médias das quantidades de interesse são calculadas, para um determinado número de partículas, por toda a história, desde a geração da mesma, até a absorção dentro da configuração experimental do problema.

Existem vários softwares de simulação de transporte de radiação de MCLLS para uso geral que implementam diferentes modelos de interação e algoritmos de transporte de partículas. Podem ser citados como exemplos os trabalhos de Allison et al. (2006), Briesmeister et al. (2000), Cullen (1998), Ferrari et al. (2005), Nelson, Rogers e Hirayama (1985), Salvat, Fernández-Varea e Sempau (2006). A maioria desses pacotes de simulação de MCLLS está disponível publicamente e sua distribuição não é restrita. Portanto, o usuário tem liberdade de escolher um pacote que seja o mais adequado para resolver um problema. O MCNP5 e o PENÉLOPE são capazes de simular elétrons / fótons em geometrias arbitrárias (CARVAJAL et al., 2009; EAKINS, 2009; VILCHES et al., 2008; SEMPAU et al., 1997; ZAMBURLINI et al., 2007).

2.6.2 Fundamentação

O principal propósito do MCLLS é realizar o ajuste do espectro experimental de raios gama pronto, usando simulações para as curvas de raios gama dos constituintes.

Na aproximação pelo MCLLS, considere a seguinte expressão:

$$n(i) = \sum_{k=1}^m x_k s_k(i) \quad (2.24)$$

que é uma relação linear entre $n(i)$, representando a contagem espectral total n da amostra em algum canal (i) , e as variáveis $s_k(i)$, que representam as contagens espectrais dos k -ésimos

constituintes da amostra em diferentes canais i . x_k representa a fração percentual do k -ésimo elemento.

A aproximação por mínimos quadrados corresponde à minimização da seguinte grandeza (minimização do erro χ -quadrado, busca realizada pela descida de gradiente por pontos críticos):

$$\chi^2 = \sum_{i=1}^t \frac{1}{\sigma_i^2} \left\{ n(i) - \sum_{k=1}^m x_k s_k(i) \right\}^2 \quad (2.25)$$

considerando σ_k^2 como a diferença entre t e m , onde t é total de canais e m é o número de isótopos.

O mínimo de χ^2 é encontrado igualando a derivada parcial de χ^2 a zero, com respeito aos parâmetros que indicam a fração em peso, chegando-se às seguintes equações:

$$\frac{\partial \chi^2}{\partial x_l} = -2 \sum_{i=1}^t \frac{1}{\sigma_i^2} s_l(i) \left\{ n_i - \sum_{k=1}^m x_k s_k(i) \right\} \quad (2.26)$$

Igualando a zero e realizando algumas manipulações algébricas:

$$\sum_{k=1}^m x_k \sum_{i=1}^t \frac{1}{\sigma_i^2} s_k(i) s_l(i) = \sum_{i=1}^t \frac{1}{\sigma_i^2} s_l(i) n(i), l = 1, \dots, m \quad (2.27)$$

Equação 2.27 representa um conjunto de equações lineares que podem ser dados em notação matricial, resultando em:

$$\sum_{k=1}^m A_{l,k} x_k = N_l. \quad (2.28)$$

De maneira mais abreviada, escreve-se:

$$\mathbf{Ax} = \mathbf{N} \quad (2.29)$$

Onde os elementos de $\mathbf{A}$ e $\mathbf{N}$ são dados a seguir:

$$A_{l,k} = \sum_{i=1}^t \frac{1}{\sigma_i^2} s_k(i) s_l(i). \quad (2.30)$$

$$N_l = \sum_{i=1}^t \frac{1}{\sigma_i^2} s_l(i) n(i). \quad (2.31)$$

A solução para o sistema B.4 pode ser determinada por:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{N}. \quad (2.32)$$

O PGNAÁ é inerentemente uma técnica não linear. Isso ocorre por causa de efeitos diversos, como atenuação de raios gama, levando sérios erros em análises quantitativas. Enfatiza-se, também, que as contagens de raios gama de um único constituinte dependerá também das quantidades de outros constituintes. O MCLLS tem a vantagem de considerar esses efeitos automaticamente ao apresentar uma abordagem linear para esse problema do PGNAÁ que tem características inerentemente não lineares. Pode-se, ainda, citar como vantagens do MCLLS (MERIÇ, 2012)

- Todo o espectro é considerado, ou seja, as regiões de pico e as regiões de Contínuo Compton.
- Resíduos podem revelar se existem bibliotecas faltantes, ou seja, LLS pode revelar contribuições de resíduos muito pequenas nas aproximações. Nesse ponto, entenda-se como bibliotecas faltantes alguma biblioteca que não tenha sido fornecida ao sistema do LLS pelo módulo do MC.
- Desvios padrões de frações em massa são avaliados a partir das buscas do LLS.

Um grande problema reside, também, na não linearidade do detector nas simulações de MC. Esse fato é levado em conta nas aplicações de funções de resposta do detector por código para fins específicos (CHANG et al., 2019).

Ainda nos estágios encontrados na abordagem do MC, duas modificações foram feitas no intuito de se conseguir realizar cálculos satisfatórios com a finalidade de detectar impurezas. A primeira modificação foi a subtração da biblioteca iterativa (MILLER; HOLMES; GARDNER, 2015); a segunda foi a utilização de uma técnica de biblioteca combinada iterativa (MERIÇ, 2012).

A abordagem de subtração é um método iterativo (MILLER; HOLMES; GARDNER, 2015) baseado na técnica de “descascamento” usada para extrair as bibliotecas indesejadas do espectro desconhecido. As bibliotecas indesejadas surgem devido a componentes da configuração experimental, blindagem e diversos outros materiais, podendo ser produzidas utilizando o CEARCPG e aplicadas para retirar do espectro experimental de raios gama.

2.6.3 Mau Condicionamento

A solução para a Equação 2.32 pode ser aproximada pelos Mínimos Quadrados Clássico, Equação 2.33, fornecendo os multiplicadores de cada biblioteca no ajuste do LLS às bibliotecas

fornecidas pelo MC:

$$\mathbf{x} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{b} \quad (2.33)$$

Um grande problema inerente da área de PGNAA é que, mesmo nos casos em que a matriz A é invertível, essa inversão pode contribuir com soluções de baixa qualidade devido ao mau condicionamento da matriz A . A fonte desse mau condicionamento reside em bibliotecas que tem uma forma espectral semelhante, ou em bibliotecas que tenham uma contribuição negligenciável ao espectro total. As soluções dos mínimos quadrados são acompanhadas de grandes incertezas (MERIÇ, 2012), tornando-se difícil estimar as quantidades corretas de frações do isótopos (SHAHABINEJAD; VOSOUGHI, 2018). Tais instabilidades na abordagem do MCLS são, em certos casos, inevitáveis.

Essas fontes de incertezas são provenientes principalmente de:

- Os espectros de duas ou mais bibliotecas tem formatos que são semelhantes, implicando em dependência linear nas colunas da matriz A (CHANG et al., 2019).
- A biblioteca de um dos constituintes tem uma pequena contribuição para o espectro total. As taxas de contagem de uma determinada biblioteca são muitas ordens de magnitude menores que as taxas de contagem de qualquer outra biblioteca usada na contagem (CHANG et al., 2019).

Na tentativa de superar os desafios, em Chang et al. (2019), são utilizados raios gama de 2 a 8 MeV no intuito de se obter as bibliotecas de radiação de fundo (*Background*). Os picos de raios gama de compostos que participam no espectro total apenas de maneira residual podem se sobrepor aos picos da biblioteca de background, compreendendo mais uma dificuldade no que diz respeito a compostos que apresentam essa característica.

Métodos de separação espectral já foram propostos em situações anteriores, como em Furr, Robinson e Robins (1968). Esse Método começa no pico de energia mais alto no espectro desconhecido. A biblioteca correspondente é identificada e subtraída do desconhecido com seu correspondente pico. O próximo pico é identificado e o mesmo procedimento é realizado. Existem algumas desvantagens a serem consideradas nesse método: a subtração inicial no pico de energia mais alto significa que os cálculos são realizados na área com a pior incerteza estatística (MILLER; HOLMES; GARDNER, 2015). Como a remoção espectral remove apenas uma biblioteca por vez, qualquer incerteza na subtração de uma biblioteca será refletida em

todas as subtrações subsequentes. Outra aproximação a ser utilizada é realizar o ajuste dos mínimos quadrados para o espectro total com as bibliotecas disponíveis e estudar os resíduos resultantes (GARDNER et al., 1997b). As localizações dos picos dos espectros não conhecidos podem ser determinados por esse método, mas a forma do Contínuo do Compton é mais difícil de diferenciar. A forma espectral completa é necessária para executar um ajuste pelo LLS. Isso significa que a forma dos elementos identificados teria que ser determinada através de outros métodos.

O desenvolvimento e implementação de métodos simulatórios de MC podem ser encontrados em Maucec e Rigollet (2004) para estudo de desempenho de sistemas de detecção baseado na técnica de análise de nêutrons rápidos/térmicos. O objetivo é desenvolver e implementar métodos de simulação para apoiar e avançar as técnicas de análise de dados dos espectros de raios γ característicos, levando à caracterização elementar de objetos inócuos usando a abordagem de análise de espectro total (FSA). A aplicação dessa técnica permitiu o cálculo de espectros padrão de raios gama individuais em cerca de 7 horas de computação em um computador padrão. A exceção foi o cálculo de espectro padrão de captura de nêutrons para o oxigênio. Devido à seção transversal de captura de nêutrons extremamente baixa, o procedimento de amostragem *Monte Carlo N-Particle* (MCNP) requerido para alcançar valores satisfatórios é menos eficiente. Consequentemente, um tempo computacional total de cerca de 24 horas foi necessário.

2.7 GREEDY RANDOMIZED ADAPTATIVE SEARCH PROCEDURE

Problemas de otimização que envolvem um número grande, mas finito, de alternativas surgem frequentemente na indústria, governo e ciência (FEO; RESENDE, 1995). Pode-se citar como exemplo: projetar redes de telecomunicações, agendamento de operações em uma fábrica, localizar reservas estratégicas de energia, rotear veículos e projetar um grande experimento. Em todos esses casos é possível enumerar todas as combinações de soluções e avaliar cada uma delas em relação ao objetivo declarado. Os que fornecem os resultados mais favoráveis são considerados ótimos. No entanto, do ponto de vista prático, é inviável seguir tal estratégia de enumeração completa, porque o número de combinações cresce exponencialmente com o tamanho do problema.

Muitos trabalhos foram feitos no intuito de se conseguir desenvolver métodos de busca ideal que não exigem explicitamente um exame de cada alternativa. Esta pesquisa deu origem

ao campo de otimização combinatória (PAPADIMITRIOU; STEIGLITZ, 1998), e uma capacidade crescente de resolver problemas tornou-se real. Alguns casos de aplicações bem sucedidas foram relatados em programação linear, como encontrado em Karmarkar e Ramakrishnan (1991).

Diversos problemas encontrados na indústria e no governo são computacionalmente intratáveis por sua natureza, ou suficientemente grandes para impedir o uso de algoritmos exatos. Nesses casos, métodos heurísticos são geralmente empregados para encontrar boas soluções, mas não necessariamente ótimas. A eficácia desses métodos depende de estratégias importantes, como a sua capacidade de evitar situações de ficar preso em ótimos locais e explorar a estrutura básica do problema. Além disso, procedimentos de reinicialização, randomização controlada, estruturas de dados eficientes e pré-processamento são bastante benéficos. Com base nessas noções, várias técnicas de busca heurística foram desenvolvidas, que comprovadamente melhoraram a capacidade de obter boas soluções para problemas de otimização combinatória difíceis. As mais promissoras dessas técnicas incluem simulated annealing (KIRKPATRICK, 1984), tabu search (GLOVER, 1989; GLOVER, 1990), algoritmos genéticos (GOLDBERG; HOLLAND, 1988) e GRASP (RESENDE; FEO, 1996).

O GRASP é um processo iterativo de duas fases: uma de construção e uma fase de busca local. A melhor solução é mantida como resultado. O pseudo-código é fornecido no Algoritmo 1. A linha 1 corresponde à entrada do problema. As iterações ocorrem da linha 2 à linha 6 e terminam quando algum critério de finalização, como o número máximo de iterações ocorridas ou a solução procurada for encontrada, for satisfeito. A construção de uma solução é realizada na linha 3. Os procedimentos para encontrar uma solução nas vizinhanças da solução construída, por meio de uma busca, são realizados na linha 4. Se uma solução melhor for encontrada, a solução corrente é atualizada na linha 5.

Algoritmo 1 Algoritmo Grasp

```

1: inputSolution
2: for {critério de parada não for satisfeito} do
3:   cons ← construirSolucao(inputSolution)
4:   sear ← buscaSolucao(cons)
5:   solAtua ← melhorSolucao(inputSolution, sear)
6: end for

```

Na fase de construção, a cada iteração do GRASP, uma solução viável é construída por vez, observando-se uma quantidade mínima de diferenças em relação a todas as soluções já construídas até então.

Associado ao Algoritmo GRASP, tem-se a lista contendo as melhores soluções até então encontradas. Em cada iteração, a escolha do próximo elemento para ser adicionado à Lista Restrita de Candidatos (RCL) é determinado, ordenando os elementos presentes na lista de candidatos com respeito a uma função objetivo. Essa lista contém os melhores candidatos. A heurística é adaptativa porque os elementos são atualizados a cada iteração com respeito aos valores da função objetivo. O componente probabilístico do GRASP é caracterizado pela escolha aleatória de um dos melhores candidatos durante as iterações, mas não necessariamente o melhor deles. Esta técnica permite que diferentes soluções sejam obtidas em cada iteração GRASP, não comprometendo a capacidade do componente adaptativo.

O Algoritmo 2 exibe o código básico para a etapa construtiva do GRASP. A solução é inicializada na linha 1. Na linha 3, é criada a RCL. Um candidato é escolhido ao acaso na linha 4 e adicionado à solução na linha 5. Por fim, calcula-se o valor da função objetivo para o elemento selecionado (linha 6) (FEO; RESENDE, 1995).

Algoritmo 2 Algoritmo Build Grasp

```

1: Solution = {};
2: while {Solução não for construída} do
3:   RCL ← Fazer();
4:   s ← selecRandom(RCL);
5:   Solution ← solution ∪ {s};
6:   ObjFun(s);
7: end while

```

As garantias de o GRASP apresentar as melhores soluções localmente não existem, por apresentar características também probabilísticas. Por isso, é benéfico aplicar uma busca em torno da solução construída. Esse algoritmo de busca funciona de maneira iterativa, substituindo sucessivamente a solução atual por uma solução melhor nas vizinhanças. Esse processo termina quando nenhuma solução melhor que a atual é encontrada. A *estrutura de vizinhança* N para uma solução P relaciona a solução s do problema para um subconjunto de soluções $N(s)$. A estrutura de um algoritmo de busca tem a forma geral demonstrada no Algoritmo 3. A escolha adequada de uma estrutura de vizinhança, técnicas de pesquisa e a solução inicial são a chave para o sucesso de um algoritmo de busca local.

Algoritmo 3 Algoritmo Search Grasp

```

1: while {s não for ótimo local} do
2:   t ← Encontre uma solução melhor  $t \in N(s)$ 
3:   s ← t

```

4: **end while**

5: return s ;

Embora os procedimentos de otimização locais possam exigir tempo exponencial a partir de um ponto de partida arbitrário, sua eficiência melhora significativamente a partir da melhora da solução inicial. Uma fase de construção eficiente que produza soluções iniciais de qualidade contribuem para uma convergência mais rápida do GRASP. Muitas vezes, soluções GRASP são geradas no tempo necessário para o procedimento de otimização local convergir de uma inicialização aleatória.

É difícil formalizar a qualidade das soluções encontradas durante o GRASP. Cada iteração GRASP produz uma amostra de solução a partir de uma distribuição desconhecida de todos os resultados obtidos (FEO; RESENDE, 1995). A média e a variância são funções da RCL. Se o tamanho da RCL é um, uma única solução é gerada e a variância da distribuição é nula. Dada uma função gananciosa, ou seja, uma função que sempre procura a melhor solução, o valor médio da solução neste caso deve ser bom, mas provavelmente abaixo do ideal. Se um limite de cardinalidade menos restritivo for imposto, muitas soluções diferentes serão produzidas, implicando em uma variação maior. Uma vez que a função objetivo é gananciosa, a solução média deve degradar sua qualidade. Devido a estatísticas de ordem e pelo fato de as amostras serem produzidas aleatoriamente, o melhor valor encontrado deve superar o valor médio (RESENDE; FEO, 1996). Na figura 2, para diferentes valores de cardinalidades (tamanho da lista = 1, 2, ..., 256), a distribuição dos valores de solução observados (3116, 3117, ..., 3126) obtidos para cada iteração, para 100.000 execuções do GRASP, são demonstrados. A simulação usa o código *GRASP-B*, de Resende e Feo (1996) para resolver instâncias do problema de satisfatibilidade *ssa7552-160* de 2ndDIMACS (JOHNSON; TRICK, 1996). No problema, procura-se maximizar o número de cláusulas satisfeitas. A instância do problema tem 1391 variáveis, 3126 cláusulas e 7025 literais. O valor da solução ótima é 3126. A simulação mostra que a solução gulosa ($|RCL| = 1$) tem um alto valor médio (3124.00) e o menor valor de variância (zero). À medida que essa restrição é relaxada, os valores médios diminuem e as variâncias aumentam. Com o aumento da variância, o número de amostras retiradas do conjunto de soluções ótimas também aumenta. Nenhuma solução é retirada do conjunto para $|RCL| = 1, 2, 4$. Cinco são obtidas para $|RCL| = 8$ e 689 são obtidas para $|RCL| = 256$.

Figura 2 – Distribuição amostral das soluções nas iterações do GRASP .

size	solution values										
RCL	3116	3117	3118	3119	3120	3121	3122	3123	3124	3125	3126
1									100000		
2							151	6053	93796		
4						75	1676	17744	80503	2	
8				1	50	750	6566	31257	61336	35	5
16				16	282	2485	13274	38329	45547	42	25
32		1	3	72	635	4196	16455	37937	40479	164	58
64		4	18	177	1213	5933	19553	37666	34832	441	163
128		4	36	269	1716	7324	21140	37186	34832	679	281
256	1	5	35	304	1980	7867	21792	36725	29027	1575	689

Fonte: Trabalho de Feo e Resende (1995)

Em (LOMBARD; TAMAYO; FONTANE, 2018), é apresentada uma abordagem para modelar o sistema de roteamento de veículos como função do tempo usando informações de cartografia on-line. O modelo é baseado em uma matriz multicamadas composta por tempos de viagem, substituindo a matriz de distância tradicional. Tempos de viagem são obtidos para cada passo na discretização do tempo. O modelo integra o fato de que o tempo de viagem entre dois pontos é modificado durante o horizonte de intervalo de tempo. Este modelo é aplicado a um problema de tamanho médio na área urbana de Paris usando o GRASP

O *Vehicle Routing Problem* - Problema de roteamento de veículos (VRP) (DANTZIG; RAMSER, 1959) foi introduzido pela primeira vez em 1959 ao abordar a otimização da entrega de gasolina. Na solução do problema, o trabalho de Lombard, Tamayo e Fontane (2018) utiliza o Algoritmo GRASP, com o método fundamentado no trabalho de Reyes, Savelsbergh e Toriello (2017) que resolveu uma versão do VRP.

2.8 CONTINUOUS GREEDY RANDOMIZED ADAPTATIVE SEARCH PROCEDURE

Problemas de otimização global objetivam uma solução $x^* \in S \subseteq \mathbb{R}^n$ tal que $f(x^*) \leq f(x)$, $\forall x \in S$, onde S é alguma região de $\mathbb{R}^n$ e a função objetivo é definida por $f: S \rightarrow \mathbb{R}$. Tal solução x^* é a chamada de *mínimo global*. Uma solução x' é um mínimo local em uma vizinhança $S_0 \subset S$ se $f(x') \leq f(x)$, $\forall x \in S_0$. Problemas de otimização global aparecem em muitas áreas, incluindo ciências dos materiais (SCHOEN, 2005), biologia, química e genética (CHEN; ZHOU; TANG, 2004), (WALES; SCHERAGA, 1999), ciência militar (KROKHMAL et al., 2003; KROKHMAL et al., 2004), engenharia elétrica (GRANVILLIERS; BENHAMOU, 2001; RATSCHKE; ROKNE, 1993), robótica (KEARFOTT, 1987; TSAI; MORGAN, 1985) e ciência dos transportes (GENDREAU; POTVIN, 2005; ZABARANKIN; URYASEV; MURPHEY, 2006).

O CGRASP é uma meta-heurística, composta por construção e melhorias locais, utilizada para resolver problemas de otimização global contínua sujeito a restrições n -dimensionais. Sem perda de generalidade, toma-se o domínio como o hiper-retângulo $S = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : l \leq x \leq u\}$, onde $l \in \mathbb{R}^n$ e $u \in \mathbb{R}^n$ tais que $u_i \geq l_i$ para $i = 1, \dots, n$ (HIRSCH et al., 2007).

A fundamentação do CGRASP está demonstrada no Algoritmo 4 (HIRSCH et al., 2007).

Algoritmo 4 Algoritmo CGrasp

```

1: input {  $n, l, u, f(\cdot), maxIters, maxNumIterNoImpro, numTimesToRun, maxDirToTry, \alpha$  };
2:  $f^* \leftarrow \infty$  ;
3: for  $\{j = 1, \dots, numTimesToRun\}$  do
4:    $x \leftarrow \text{getRand}(l, u)$  ;
5:    $h \leftarrow 1$  ;
6:    $numNoImprov \leftarrow 0$  ;
7:   for  $\{\text{iter} = 1, \dots, maxIters\}$  do
8:      $x \leftarrow \text{construcGreeRando}(x, f(\cdot), n, h, l, u, \alpha)$  ;
9:      $x^* \leftarrow \text{localImpro}(x, f(\cdot), n, h, l, u, maxDirToTry)$  ;
10:    if  $\{f(x) < f^*\}$  then
11:       $x^* \leftarrow x$  ;
12:       $f^* \leftarrow f(x)$  ;
13:       $numNoImpro \leftarrow 0$  ;
14:    else
15:       $numNoImpro \leftarrow numNoImpro + 1$  ;
16:    end if
17:    if  $\{numNoImpro \geq maxNumIterNoImpro\}$  then
18:       $h \leftarrow \frac{h}{2}$  ;
19:       $numNoImpro \leftarrow 0$  ;
20:    end if
21:  end for
22: end for

```

O procedimento no Algoritmo 4 toma como entrada para o problema a dimensionalidade n , os valores de limites superior e inferior para uma dimensão arbitrária (l, u) , o valor da função objetivo $f(\cdot)$, o valor de número máximo de iterações $maxIters$, o valor de número máximo de iterações sem melhorias $maxNumIterNoImpro$, número máximo de direções a ser pesquisada na busca $maxDirToTry$, número de vezes a executar o CGRASP $numTimesToRun$.

Na linha 2, inicia-se a função objetivo com o valor ∞ . Da linha 3 a 21, o Algoritmo CGRASP será executado por uma quantidade de iterações determinada pela variável $numTimesToRun$,

caracterizando-o, assim, como um procedimento *multi-start*. Para cada iteração, uma solução inicial x é construída (linha 8) sobre um conjunto de pontos uniformemente distribuído sobre um *box* em $\mathbb{R}^n$ limitado por l e u . O parâmetro h controla a discretização do espaço de busca, sendo inicializado como 1 (linha 5). A construção e o processo de busca local são chamados *maxIters* vezes, nas linhas 8 e 9.

Uma nova solução após a busca local é comparada contra a solução corrente na linha 10. Se a nova solução tem um valor de função objetivo menor que o valor da função objetivo corrente, a nova solução corrente é atualizada nas linhas 11 e 12, e o valor de *numNoImprov*, que controla a densidade do grid na busca, é ajustado com o valor zero. Se assim não acontecer, na linha 15, incrementa-se o valor de *numNoImprov*. Se *numNoImprov* torna-se maior que *maxNumIterNoImprov*, a densidade do grid na busca é reduzida à metade, e *numNoImprov* é ajustado para zero. Isso permitirá o CGRASP iniciar com uma discretização suave e aumentar de maneira adaptativa a densidade conforme o necessário, intensificando a busca por meio de uma discretização mais intensa. A melhor solução encontrada ao fim de todos os *numTimesToRun* é retornada.

O algoritmo da fase construtiva está descrito no Algoritmo 5. Ao iniciar o algoritmo, permite-se que todas as coordenadas de x sejam transformadas. Uma pesquisa é feita em cada uma das coordenadas (linha 8) não fixas, mantendo-se constantes todas as demais coordenadas. O valor que minimiza z_i , minimizará a função objetivo $f(z_i)$. A coordenada i é salva, assim como também o valor da função objetivo para essa coordenada (linhas 8 e 10). Ao se executar a pesquisa em cada uma das direções, forma-se uma RCL que contém as coordenadas i cujos valores g_i são menores ou iguais a $\alpha \cdot \max + (1 - \alpha) \cdot \min$, onde $\max$ e $\min$ são os valores máximo e mínimo que g_i assumiu durante a pesquisa sobre todas as coordenadas não fixas de x . Os valores para a variável $\alpha \in [0, 1]$ é um parâmetro definido pelo usuário do código. Formada a RCL, escolhe-se uma coordenada ao acaso, e nova solução corrente passa a ser x_j , fixando-se coordenada j . Isso assegura a fase aleatória do processo construtivo. Continua-se o processo até que todas as n coordenadas tenham sido fixadas. Após isso, o valor de x é retornado do processo construtivo.

Algoritmo 5 Algoritmo build CGrasp

```

1: input  $\{x, f(\cdot), n, h, l, u, \alpha\}$ ;
2:  $S \leftarrow \{1, 2, \dots, n\}$ ;
3: while  $\{S \neq \emptyset\}$  do
4:    $\min \leftarrow +\infty$ ;
5:    $\max \leftarrow -\infty$ ;
6:   for  $\{i = 1, \dots, n\}$  do
7:     if  $\{i \in S\}$  then
8:        $z_i \leftarrow \text{LinSear}(x, h, i, n, f(\cdot), l, u)$ ;
9:        $g_i \leftarrow f(z_i)$ ;
10:      if  $\{\min > g_i\}$  then
11:         $\min \leftarrow g_i$ ;
12:      end if
13:      if  $\{\max < g_i\}$  then
14:         $\max \leftarrow g_i$ ;
15:      end if
16:    end if
17:  end for
18:   $RCL \leftarrow \emptyset$ ;
19:  for  $\{i = 1, \dots, n\}$  do
20:    if  $\{i \in S \ \&\& \ g_i \leq (1 - \alpha) * \min + \alpha * \max\}$  then
21:       $RCL \leftarrow RCL \cup \{i\}$ ;
22:    end if
23:  end for
24:   $j \leftarrow \text{RandSelec}(RCL)$ ;
25:   $x_j \leftarrow z_j$ ;
26:   $S \leftarrow S - \{j\}$ ;
27: end while
28: return(x);

```

A busca local inicia a partir de um ponto $x \in \mathbb{R}^n$, gerando um conjunto de direções. A partir de uma dessas direções, é determinada a direção em que a função objetivo melhora seu valor.

Para um problema bidimensional (facilmente generalizado para n dimensões), dado um ponto x , existe um conjunto de direções para serem analisados na busca local $\{(1, 0), (0, 1), (-1, 0), (0, -1), (-1, 1), (1, -1), (-1, -1)\}$. Em geral, em $\mathbb{R}^n$, cada direção irá ser um vetor de comprimento n , onde cada componente desse vetor pertence a $\{-1, 0, 1\}$. É possível verificar que serão possíveis $3^n - 1$ direções, excluindo o caso em que todos os elementos do vetor são iguais a

Tabela 1 – Mapeamento Ternary' ($n = 2$)

$i \rightarrow$	Ternary(i) $\rightarrow$	Ternary(i)' $\rightarrow$	$i \rightarrow$	Ternary(i) $\rightarrow$	Ternary(i)' $\rightarrow$
1 $\rightarrow$	01 $\rightarrow$	{0, 1}	2 $\rightarrow$	02 $\rightarrow$	{0, -1}
3 $\rightarrow$	10 $\rightarrow$	{1, 0}	4 $\rightarrow$	11 $\rightarrow$	{1, 1}
5 $\rightarrow$	12 $\rightarrow$	{1, -1}	6 $\rightarrow$	20 $\rightarrow$	{-1, 0}
7 $\rightarrow$	21 $\rightarrow$	{-1, 1}	8 $\rightarrow$	22 $\rightarrow$	{-1, -1}

0. Pode-se fazer um mapeamento entre cada número no conjunto $\{1, \dots, (3^n - 1)\}$ e cada uma das direções. A função Ternary mapeia cada $r \in \mathbb{N}$ para uma representação ternária (base 3). Por exemplo, $\text{Ternary}(15) = 120$. Dito isso, realiza-se uma modificação em Ternary (chamando-a por Ternary'), com $r \in \mathbb{N}$ mapeado dentro da representação ternária, mas com o '2' sendo substituído por '-1'. Então, a função Ternary' mapeia cada $r \in \mathbb{N}$ dentro de vetores de direção. A saída será a direção de um vetor (HIRSCH et al., 2007).

Algoritmo 6 Algoritmo Search CGrasp

```

1: Input  $x, f(\cdot), n, h, l, u, maxDirTry$ 
2:  $impro \leftarrow true$ ;
3:  $D \leftarrow \emptyset$ ;
4:  $x^* \leftarrow x$ ;
5:  $f^* \leftarrow f(x)$ ;
6:  $numTry \leftarrow \min\{(3^n - 1), maxDirTry\}$ ;
7: while  $\{impro\}$  do
8:    $impro \leftarrow false$ ;
9:   while  $\{|D| \leq numTry \ \&\& \ notImpro\}$  do
10:      $r \leftarrow \text{randomSelect}(1, (3^n - 1) \notin D)$ 
11:      $D \leftarrow D \cup \{r\}$ ;
12:      $d \leftarrow \text{Ternary}'(r)$ 
13:      $x \leftarrow x^* + h \cdot d$ ;
14:     if  $\{l \leq x \leq u\}$  then
15:       if  $\{f(x) < f^*\}$  then
16:          $x^* \leftarrow x$ ;
17:          $f^* \leftarrow f(x)$ ;
18:          $D \leftarrow \emptyset$ ;
19:          $impro \leftarrow true$ ;
20:       end if
21:     end if
22:   end while
23: end while

```

A busca local é iniciada a partir de uma solução $start\ x \in S \subseteq \mathbb{R}^n$. A melhor solução corrente é inicializada na linha 4. Se o número de direções a ser pesquisado for elevado, a linha 6 trata de pesquisar em uma quantidade de direções mínima entre $(3^n - 1)$ e um valor definido pelo usuário $maxDirTry$. Iniciando a partir de um ponto x^* , da linha 9 à linha 22, constroem-se diferentes vizinhos à solução atual, escolhida uma arbitrária direção. O número total de direções será representada pela variável $numTry$. Uma solução é definida na linha 12, escolhida uma direção arbitrária. Se a solução x construída na linha 13 for viável e melhor do que a solução corrente x^* , x^* é configurada como a solução corrente (linha 16). O conjunto de direções escolhidas podem mudar durante esse processo, assim como também a ordem em que essas direções são escolhidas.

3 ANÁLISE ESPECTRAL E O ESTADO DA ARTE

No trabalho de Meriç (2012), existe a preocupação quanto ao uso do MCLLS para analisadores PGNAAs na determinação de quantidades em um fluxo multifásico. Comprovou-se que o fato de a aproximação dos mínimos quadrados não ser suficientemente sensível em determinar as frações de contagem percentual do Óleo e gases deve-se, por sua vez, ao fato de a matriz de covariância a ser invertida no sistema de equação linear ser quase singular ou mal condicionada. Um novo método iterativo foi proposto para tratar os casos de mau condicionamento no problema de fluxo multifásico. Cálculos finais indicam que o método proposto pode ser capaz de efetivamente minimizar o mau condicionamento, apesar de utilizar todo o conjunto de dados do espectro de raios gama, acarretando a sobredeterminação do sistema de equações.

O trabalho de Meriç (2012) enfrenta os casos de mau condicionamento concatenando, por colunas, uma biblioteca com a soma de todas as outras bibliotecas, compondo, assim, as bibliotecas fornecidas ao sistema do LLS. Dentre as possíveis combinações, escolhe-se a instância cuja matriz apresenta o menor condicionamento na determinação das frações em peso. Esse método não implica em maiores custos computacionais, além de simplificar os modelos matemáticos para determinação das frações em peso determinando matrizes de covariância de dimensão dois. Também pode enumerar que contribuiu na resolução do PGNAAs, determinando com precisão as frações em peso, mas poderia aprimorar o enfrentamento das causas do mau condicionamento realizando uma seleção dos canais mais adequados para o cálculo das matrizes de covariância e consequente determinação das frações em peso.

No trabalho de Chang et al. (2019), é realizada a medição de impurezas de cloro em amostras de Óleo cru. Nesse trabalho, é utilizado o *Monte Carlo Library Least Square Extended* (MCLLSX) para determinar a quantidade de elementos com frações residuais, ou elementos que tenham baixa participação no espectro total. O MCLLSX faz uso da técnica iterativa de separação de espectros para remover bibliotecas conhecidas, no módulo do MC, sendo realizada em sequência a construção de uma biblioteca de radiação de fundo, para ser usada posteriormente na etapa dos mínimos quadrados. A incorporação dessa técnica ao módulo do MC traz uma carga computacional adicional para uma etapa que apresenta elevados custos computacionais. O referido trabalho também faz uso da implementação da técnica de combinações de bibliotecas para minimizar problemas relacionados às várias ordens de diferença

de magnitude nas taxas de contagem das bibliotecas individuais na pesquisa do LLS, como descrito em Meriç (2012). Logo, as observações aplicadas ao trabalho de Meriç (2012) podem ser aplicadas ao trabalho de Chang et al. (2019) no estágio da determinação das frações em peso.

Os resultados apresentados por Chang et al. (2019) indicam que o método PGNAA em conjunto com o MCLLSX pode ser considerado viável para medições de impurezas de cloro em amostras de petróleo bruto com boa precisão, com a desvantagem que o método proposto foi testado com uma única impureza na amostra de Óleo. Uma outra ausência sentida no trabalho de Chang et al. (2019) é que não há caracterização das fases no fluxo multifásico, sendo conhecidas previamente. Ele utiliza a combinação de bibliotecas, como em Meriç (2012), para o ajuste das bibliotecas ao espectro total, sempre na tentativa de minimizar o condicionamento.

Em Shahabinejad e Vosoughi (2019), um novo algoritmo, *Sequential Gamma-ray Spectrum Deconvolution* (SGSD), é apresentado para analisar espectros de raios gama. O algoritmo SGSD produz uma sequência de espectros convergindo para melhor estimativa do detector de raios gama. Em cada etapa, um isótopo de um espectro desconhecido é identificado e a respectiva resposta do detector é construída. O algoritmo SGSD explora as vantagens do método de conjunto ativo do NNLS (GILL; MURRAY; WRIGHT, 1981) e as técnicas de otimização evolutiva (SHAHABINEJAD; VOSOUGHI, 2018) para analisar um dado espectro de raios gama. Apesar de conseguir caracterizar corretamente os elementos envolvidos na descrição da amostra em estudo, essa abordagem usa diversos isótopos de uma fonte mista de raios gama no intuito de determinar as contagens de frações percentuais. Contudo, o método do SGSD necessita de um grande número de bibliotecas nas etapas iniciais para a devida identificação e descrição do espectro total, implicando em elevados custos computacionais na etapa do MC.

Muitas soluções com boas precisões podem ser encontradas pelo algoritmo SGSD presente em Shahabinejad e Vosoughi (2019), mas existem instâncias utilizadas em que o NNLS obtém resultados melhores que o método SGSD. O SGSD também apresenta sensibilidades a dados que apresentam elevados erros estatísticos, ocasionado pela baixo tempo de aquisição dos dados. Mas os resultados gerais apresentados pelo SGSD são mais precisos que os do algoritmo NNLS para identificar e determinar as frações de contagem percentual dos isótopos na análise de espectros de raios gama. O conjunto de soluções candidatas é movida através do espaço de busca seguindo uma dinâmica por um ótimo, para uma função de custo descrita da seguinte maneira:

$$\left\{ \begin{array}{l} \text{minimize } \sqrt{\frac{MAE}{max_{mae}} + \frac{RMSE}{max_{rmse}}} \\ \text{sujeito a} \\ MAE = \frac{1}{n} \| \mathbf{N} - \mathbf{N}_{est} \| \\ RMSE = \frac{1}{\sqrt{n}} \| \mathbf{N} - \mathbf{N}_{est} \| \end{array} \right.$$

O PSO de Shahabinejad e Vosoughi (2019) direciona as soluções candidatas para o ótimo, memorizando a melhor posição encontrada por todas as soluções candidatadas e as melhores posições encontradas por cada solução candidata no processo de busca. A grande motivação para a escolha do *Mean Absolut Error* (MAE) e *Root Mean Square Error* (RMSE) é explorar as características para encontrar uma boa solução em termos de duas qualidades: RMSE minimiza o viés na solução, enquanto o MAE mede quão perto os espectros estão dos medidos. Uma desvantagem de tal método é que a descrição de um espectro total pode não ter solução única, já que uma gama de bibliotecas é disponibilizada para que o espectro total seja descrito como função de algumas delas. As frações de contagem percentual são avaliados segundo um teste de significância *Significance Of Abundance* (SOA), e o espectro estimado é avaliado pelo Theil Teste (THEIL, 1967). Os passos do algoritmo podem ser descritos por:

- Execução PSO.
- Teste SOA.
- Atualização de bibliotecas que compõem o desconhecido.
- Theil test.

A proposta desenvolvida em Androulakaki et al. (2016) foi aplicada para obter estimativas de concentração de atividade para medições em ambiente marinho. As concentrações de atividades estimadas estavam de acordo com as concentrações determinadas experimentalmente, dentro das incertezas. O referido trabalho encontra-se no contexto do FSA, que oferece a possibilidade de quantificar isótopos, verificar as estimativas de frações em peso nas regiões de picos, além de abordar a detecção de anomalias (identificar presença de isótopos que não foram considerados inicialmente na análise). Os picos são selecionados em janelas selecionadas pelo usuário, sem critérios para sua detecção, e os valores teoricamente estimados são comparados com os resultados experimentais correspondentes, apesar de a precisão dos resultados depender da

qualidade dos espectros de MC. Essa precisão é função da energia de calibração e a seleção apropriada de várias janelas que serão utilizadas na minimização do χ^2 . O espectro total é descrito como função de janelas de canais que apresentam picos em seu interior.

As frações em peso encontradas em Androulakaki et al. (2016) são determinadas janela a janela, comparando as integrais do espectro estimado com o espectro real. A abordagem apresenta eficiência computacional, por seccionar devidamente os dados tomando por base os picos, mas não apresenta essas regiões como função da quantidade de bibliotecas estudadas, ou seja, não é possível associar os picos a todas as bibliotecas disponíveis. Esse método também apresenta a desvantagem de considerar picos espúrios, ou seja, picos que não digam respeito a nenhuma biblioteca em questão, sendo selecionado para descrição do espectro total. Esse fato também pode contribuir para instabilidades numéricas na aproximação dos LLS, afetando diretamente a qualidade das soluções. O método apresenta avanços no sentido de descrever sistemas com menor número de equações, mas depende fortemente da qualidade das bibliotecas simuladas reproduzirem com fidelidade a ambiente experimental.

O trabalho apresentado por Shahabinejad e Vosoughi (2018) é uma simplificação do trabalho de Shahabinejad e Vosoughi (2019). O algoritmo de análise de espectros de raios gama baseado em otimização de enxame de partículas (PSO) é desenvolvido para identificar diferentes isótopos de uma fonte mista de raios gama e identificar suas frações de contagem percentual, baseando-se em uma modificação do algoritmo presente em Mukherjee (2002). O algoritmo PSO é iterativo e imita comportamentos sociais observados na natureza para conduzir a otimização de problemas complexos. O método PSO é usado para ajuste à resposta de um detector de cintilação NaI de 3×3 polegadas e o processo de ajuste é controlado por um teste de significância de abundância e cálculo de coeficiente de Theil.

O algoritmo desenvolvido em Shahabinejad e Vosoughi (2018) foi testado para vários espectros de raios gama, incluindo combinações de Co, Cs, Na e Am. O método apresentou avanços no sentido de não apresentar valores de frações em peso negativos, diferentemente do método com que foi comparado, o MLR. Os resultados mostram uma boa correspondência com as frações de contagem percentual reais dos compostos, apesar de não evidenciar diferenças estatísticas significativas com algoritmos clássicos. As aproximações das frações em peso não são realizadas pelo método dos mínimos quadrados, mas é abordado segundo o coeficiente de Theil, como realizado no trabalho Shahabinejad e Vosoughi (2019), apesar de essa métrica não ter sido bem sucedida na definição do espectro estimado, ao ser comparado com o espectro real.

A abordagem descrita em Alamaniotis e Jevremovic (2015) é muito semelhante a trazida por Shahabinejad e Vosoughi (2018). As principais diferenças residem em: (1) no algoritmo aplicado para aproximação das frações em peso, enquanto o autor de Shahabinejad e Vosoughi (2018) utiliza o PSO, em Alamaniotis e Jevremovic (2015) fazem uso de algoritmos genéticos; (2) Alamaniotis e Jevremovic (2015) realizam uma identificação prévia das bibliotecas por meio de lógica Fuzzy (ALAMANOTIS et al., 2011; MURRAY; BUTTERFIELD; BAIRD, 2000). Os dois são utilizados simultaneamente para identificar isótopos e suas contribuições em um espectro arbitrário. O módulo de lógica Fuzzy se concentra na identificação de isótopos no espectro, enquanto que o algoritmo genético ajusta e calcula as frações de contagem percentual dos isótopos identificados.

O ajuste do espectro estudado em Alamaniotis e Jevremovic (2015) é controlado por um procedimento de avaliação baseado no teste de significância de abundância (SOA) dos coeficientes de Theil (ALAMANOTIS; JEVREMOVIC, 2015; THEIL, 1967). Esta combinação apresenta um novo mecanismo para seleção de isótopos no ajuste de espectro, eliminando ajuste manual baseado em intervenção de usuário. As frações em peso no ajuste são determinadas com uma boa precisão, mas o módulo de identificação de picos e, conseqüentemente, das bibliotecas não faz uso dessa informação no sentido de diminuir o tamanho do sistema de equações e das conseqüentes instabilidades numéricas associadas, preocupando-se em particionar as bibliotecas que descrevem o espectro total. Os tempos computacionais para execução do algoritmo não são elevados, mas as iterações são funções das bases de dados estudadas. Nesse método, existem evidências de melhorias entre o módulo de identificação das bibliotecas e a execução do GAs para determinação das frações em peso das bibliotecas.

O uso da meta-heurística GRASP na minimização do número de condicionamento da matrizes de covariância utilizada no MCLLS é apresentado em Cardoso, Melo e Silva (2018). Nesse trabalho, a abordagem é aplicada a um particionamento dos espectros de bibliotecas. Esse particionamento utiliza critérios para que as matrizes de covariância construídas ao longo do GRASP apresentem características de independência linear, utilizando-se do fato que a dimensão das bibliotecas de energia ser muito maior que o número de isótopos.

O algoritmo GRASP apresentado por Cardoso, Melo e Silva (2018) é composto pela etapa de construção de uma solução, seguida por uma busca nas vizinhanças da solução construída. Não são apresentados determinações das frações em peso, mas o particionamento por independência linear contribui significativamente para a redução no condicionamento das matrizes de covariância. Os tempos computacionais não são elevados, utilizando a mesma

abordagem apresentada por Meriç (2012). Apesar de o êxito não ter sido alcançado em duas das nove instâncias, os resultados indicam que a nova abordagem pode superar o mau condicionamento do problema, contribuindo para uma efetiva melhora na etapa dos mínimos quadrados do MCLLS.

O trabalho apresentado em Cardoso, Melo e Silva (2018) foi ampliado em Cardoso et al. (2020). Em Cardoso et al. (2020), foi adicionado ao GRASP um mecanismo que fará uso de memória das melhores soluções encontradas durante a execução do algoritmo, chamado *path-relinking*. Também destaca-se a avaliação das frações de contagem percentual presentes, utilizando as matrizes de covariâncias definidas pelo GRASP no passo dos Mínimos Quadrados. Os resultados foram comparados com os valores de frações em peso determinados pelo trabalho de Meriç (2012). A aplicação do GRASP no particionamento dos dados de bibliotecas de energia para posterior determinação das frações mostrou-se promissora ao realizar a comparação dos resultados com o que se encontra publicado no estado da arte.

Em Meric et al. (2014), o uso de ativação de nêutrons de raios gama em conjunto com a abordagem de mínimos quadrados foram considerados para medições de fração de massa de componentes em fluxo multifásico. Os resultados apresentados indicam que o PGNAA em conjunto com o MCLLS é viável para medição de fluxo multifásico, desde que o mau condicionamento nos mínimos quadrados possa ser tratado adequadamente. A primeira abordagem consiste em um setup físico experimental de combinar um analisador PGNAA empregando uma fonte de nêutron Cf-252 com um densímetro de Cs-137. A segunda abordagem foi iterativa, que se baseia na minimização do condicionamento, como presente em Meriç (2012), nos cálculos do LLS para compostos como Água, Óleo, Gás, Sal e Background. Os resultados indicam que ambos os métodos podem ser utilizados para superar a questão do mau condicionamento. Mas foi indicado que o método iterativo fornece melhores resultados que o analisador PGNAA combinado com o densímetro. Abordagem muito semelhante à de Meriç (2012), em que as modificações consistem na introdução de uma nova fonte de nêutrons. Pode-se, então, determinar que a descrição do trabalho de Meriç (2012) se estendem ao presente trabalho.

Em Miller, Holmes e Gardner (2015) é proposto um método para determinar a radiação de fundo presente nos espectros experimentais através de uma abordagem matemática. Bibliotecas de interesse são subtraídas de um espectro total usando multiplicadores para essa biblioteca e subtraindo os resultados do espectro total. Esse processo prossegue até que se considere que o pico dessa biblioteca considerada seja extraído. Após a remoção de todas as bibliotecas, o

restante do espectro experimental deve representar uma aproximação do background. Essa aproximação pode ser usada em conjunto com os mínimos quadrados para determinar as contribuições de radiação de fundo para o espectro total. A operação de subtrair o pico de mais alta energia muitas vezes significa que os cálculos estão sendo realizados na área de pior incerteza estatística. Como as bibliotecas são extraídas uma por vez, uma incerteza para uma biblioteca irá se refletir em todas as subtrações, ocasionando incertezas em cascatas para os passos subsequentes. Outra aproximação é executar o ajuste dos mínimos quadrados para o espectro total com as bibliotecas disponíveis e estudar os resíduos do ajuste resultante para obter as bibliotecas faltantes. O pico de bibliotecas faltantes pode ser determinado usando este método, mas o desenho do espectro total é difícil de ser definido. A forma espectral completa é necessária para executar o ajuste do LLS, implicando na necessidade em identificar a forma dos elementos faltantes de outra maneira. O trabalho não menciona frações em peso no ajuste dos mínimos quadrados, mas sim com o percentual que a biblioteca conhecida cobre do espectro total.

O presente trabalho aborda a questão do particionamento dos dados, determinando os picos das bibliotecas conhecidas. Essa iniciativa irá estender a abordagem presente em (CARDOSO et al., 2020), em que a secção dos dados é realizada por critérios de independência linear entre as bibliotecas fornecidas. Acredita-se que a determinação das regiões do picos contribuirá em formação de partições mais restritas e especializadas. As determinações quantitativas deste trabalho são realizadas pelos método dos Mínimos Quadrados, após a aplicação dos algoritmos heurísticos GRASP e CGRASP. Os algoritmos aqui aplicados serão os mesmos que os aplicados em (CARDOSO et al., 2020), mas com uma diferenciação na abordagem: a otimização agora será realizada nos canais e nas estimativas de uma biblioteca faltante. A cada nova estimativa para essa biblioteca faltante nas iterações do GRASP, é realizado o ajuste dos mínimos quadrados. Esse novo ajuste produz um resíduo entre o espectro total e o estimado, que é associado a novos valores de frações em peso. Nas ocorrências de minimização de resíduos, os valores de condicionamento serão observados. O ajuste dos mínimos quadrados será realizado para a ocorrência do menor valor de condicionamento, e os valores serão comparados com os determinados no ajuste dos mínimos quadrados na minimização do resíduo.

Os trabalhos acima citados são classificados segundo alguns critérios:

- Um dos critérios avaliados será se o trabalho em questão determina a detecção de picos.

A resposta para esse item será **Aplicável** ou **Não Aplicável**;

- Um outro item a ser avaliado será a análise quantitativa, ou a determinação das frações em peso. Mais uma vez, as respostas para esse item será **Aplicável** ou **Não Aplicável**;
- O condicionamento do sistema de equações será alvo de avaliação nesse item, classificando-o como **Aplicável** ou **Não Aplicável** ao trabalho descrito;
- Será verificado se o artigo apresenta abordagens para trabalhar com seções dos dados, classificando-o **Aplicável** ou **Não Aplicável**;
- Os algoritmos utilizados neste trabalho e as finalidades da aplicação desse algoritmo, seja para análise quantitativa ou para detecção de picos;
- Os principais modelos utilizados nas aproximações quantitativas ou na detecção de picos. Dentre esses modelos, pode-se citar: mínimos quadrados, coeficiente de Theil, Testes de significância, lógica Fuzzy para detecção de picos;
- Será verificado se o trabalho descrito aplica a parametrização das partições criadas. A resposta para esse item será **Aplicável** ou **Não Aplicável**.

Tabela 2 – Resumo dos trabalhos em PGNA e metodologias no ajuste do espectro.

Trabalho	Deteção Picos	Análise Quantit.	Nº Condi.	Partic.	Algoritmos	Modelos	Parâmetri.
(MERİÇ, 2012)	Não aplicável.	Aplicável.	Aplicável.	Não aplicável.	Processo iterativo: biblioteca contra a soma de todas as demais.	Frações em peso pelo mínimos quadrados; Não há enfrentamento de elemento faltante.	Não aplicável.
(CHANG et al., 2019)	Não aplicável.	Aplicável.	Aplicável.	Não aplicável.	Iterativo na determinação do <i>background</i> ; abordagem iterativa de Meriç (2012) na determinação quantitativa.	mínimos quadrados para as frações em peso; não há enfrentamento de biblioteca faltante.	Não aplicável.
(SHAHABINEJAD; VOSOUGHI, 2019)	Não aplicável.	Aplicável.	Não aplicável.	Não aplicável.	Algoritmo PSO.	Conjuntos Ativos e Passivos; Testes de significância e de adequação.	Não aplicável.
(ANDROULAKAKI et al., 2016)	Aplicável.	Aplicável.	Não aplicável.	Aplicável.	Algoritmo iterativo em diferentes janelas, comparando espectro real e estimado.	Processo de ajuste realizado pela minimização do χ^2 .	Não aplicável.
(SHAHABINEJAD; VOSOUGHI, 2018).	Não aplicável.	Aplicável.	Não aplicável.	Não aplicável.	Aplicação do <i>Adaptive Mutation based PSO</i> (AMPSO) a vetores de frações de contagem percentual que são modificados a cada iteração.	Função de custo (norma 2) aos vetores de frações de contagem percentual real e estimado na aplicação do PSO; coeficiente de Theil ao fim das iterações.	Não aplicável.
(ALAMANIOTIS; JEVREMOVIC, 2015)	Aplicável.	Aplicável;	Não aplicável.	Não Aplicável.	Lógica Fuzzy e GAs.	Teste de significância SOA e Teste de Theil de adequação.	Não Aplicável.
(CARDOSO; MELO; SILVA, 2018).	Não aplicável.	Não aplicável.	Aplicável.	Aplicável.	Utilização do GRASP na obtenção de matrizes de covariância: passo de construção e de buscas nos entornos da solução construída.	Utilização da metodologia presente em (MERİÇ, 2012) para uma seção dos dados selecionados pelo GRASP.	Aplicável (compartimentalização das partições).
(CARDOSO et al., 2020).	Não Aplicável.	Aplicável.	Aplicável	Divisão dos dados seguindo critérios para obtenção de valores L.I.	Algoritmo GRASP, com o passo do Path Relinking.	Mínimos quadrados na determinação da fração de contagem percentual.	Aplicável.
(MERİÇ et al., 2014).	Não Aplicável.	Aplicável	Aplicável.	Não Aplicável.	Algoritmo Iterativo.	Mínimos quadrados na determinação da fração de contagem percentual.	Não Aplicável.
(MILLER; HOLMES; GARDNER, 2015).	Aplicável.	Aplicável.	Não Aplicável.	Não aplicável.	Algoritmo iterativo para geração da biblioteca representativa do <i>background</i> .	Ajuste realizados pelos mínimos quadrados.	Não aplicável.
Presente Trabalho	Aplicável.	Aplicável.	Aplicável.	Aplicável.	Algoritmos heurísticos de otimização GRASP e CGRASP.	Ajuste realizados pelos mínimos quadrados.	Aplicável.

4 METODOLOGIA

A Seção da metodologia abordará uma descrição mais geral, detalhando gradualmente cada passo enumerado nessa descrição geral. Na Figura 4, é realizada a descrição dos passos aplicados aos dados de entrada, sempre destacando as entradas e saídas desses passos específicos. Cada um desses algoritmos será alvo de uma descrição, objetivando um melhor entendimento dos passos aplicados.

No intuito de apresentar fluidez à leitura do presente trabalho, e para fins de melhor entendimento da Figura 4, será utilizada a seguinte nomenclatura:

- A primeira base de dados será constituída pelos elementos Amerício (*Am*), Cobalto (*Co*), Césio (*Cs*) e Sódio (*Na*). A segunda base de dados será constituída pelos compostos Ar, Água e Óleo. Mais detalhes da configuração experimental para a geração dos dados da primeira base pode ser encontrada em Meric (2012). A segunda base é privada, cedida pelo aluno de Ilker Meric, Ícaro Valgueiro Malta Moreira. As referências para cada uma das instâncias é fornecida nas tabelas abaixo:

Tabela 3 – Descrição das bases de dados estudadas e das instâncias criadas a partir das bases.

Base	Instância	Elemento Omitido	Elemento Fornecido
Base 1	1	Amerício (<i>Am</i>)	Cobalto
			Césio
			Sódio
	2	Cobalto (<i>Co</i>)	Amerício
			Césio
			Sódio
	3	Césio (<i>Cs</i>)	Amerício
			Cobalto
			Sódio
	4	Sódio (<i>Na</i>)	Amerício
			Cobalto
			Césio
Base 2	5	Ar	Água
			Óleo
	6	Água	Ar
			Óleo
	7	Óleo	Ar
			Água

Fonte: Próprio autor

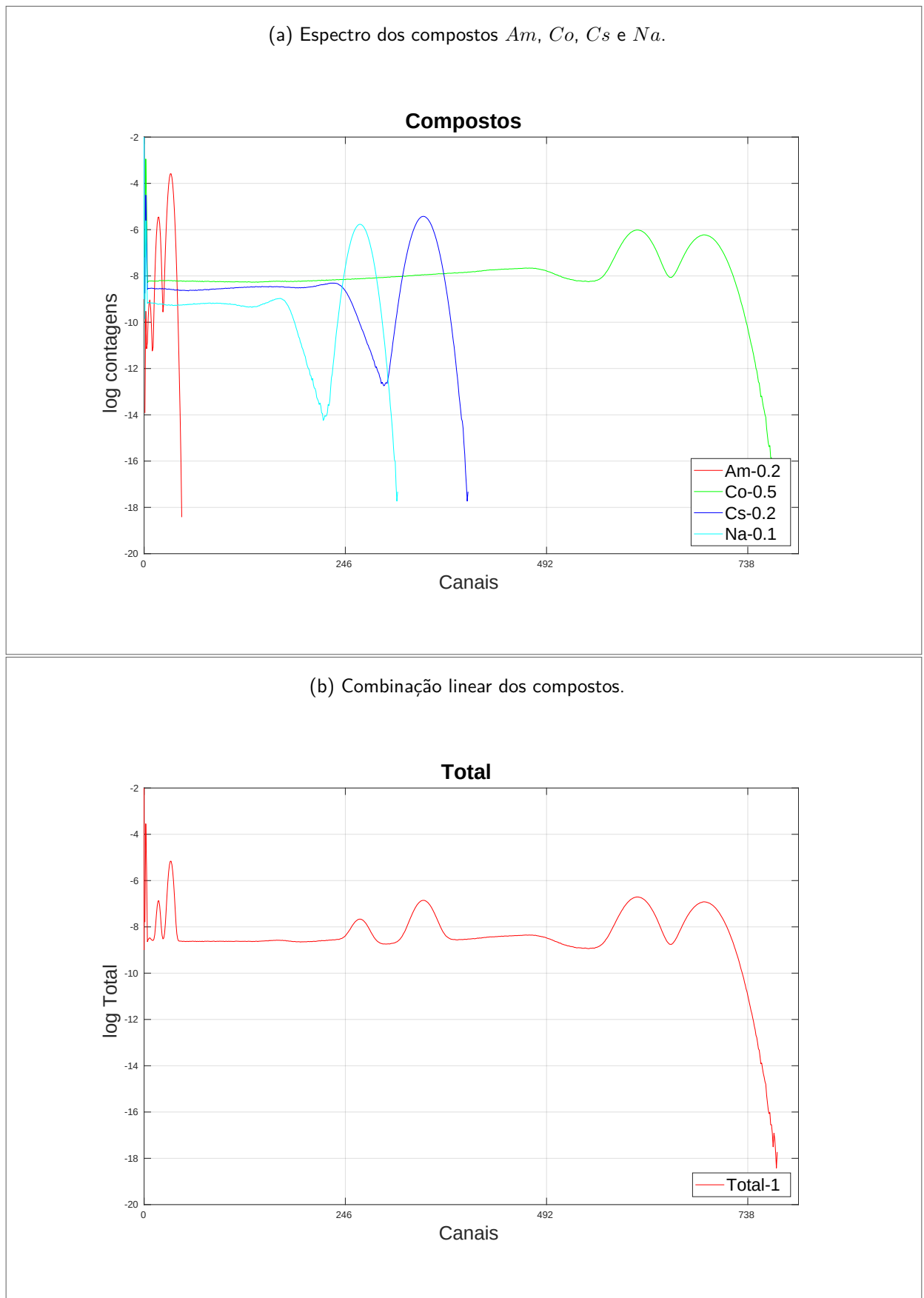
- O termo **Biblioteca, isótopo**, ou **compostos**, corresponderá a um vetor de números reais, que diz respeito às intensidades espectrais de um elemento. Esses termos serão utilizados no intuito de se evitar repetições, ou para utilização daquele que seja mais adequado ao local onde esteja sendo empregado. A linha do vetor onde se encontra uma contagem espectral será denominada **canal**. O termo **frequências de energia** também pode ser usado em substituição ao termo **canal**.
- Os dados de entrada são divididos em regiões de 60 canais, nas duas bases. Essa quantidade de canais, nas regiões, foi considerada como suficiente para determinar locais que representassem as bibliotecas fornecidas. Essas regiões são denominadas **partições**. O particionamento dos dados é importante, no intuito de se evitar a sobredeterminação do sistema de equações utilizado pelo LLS.
- Dentro de uma partição, não necessariamente todos os canais pertencentes às partições serão utilizados nas aproximações dos mínimos quadrados. Seguindo-se à divisão dos dados em partições, o algoritmo se encarregará de escolher, aleatoriamente, os canais

de cada partição que serão introduzidos na solução inicial. As regiões que irão conter os canais das partições que efetivamente foram utilizados, e modificados, ao longo de nosso algoritmo, serão chamadas de **compartimentos**. Define-se um compartimento para cada um dos n -ésimas bibliotecas utilizadas.

- Cada compartimento conterá a região de maior inflexão nas taxas de contagem espectral, para um elemento fornecido inicialmente, como pode ser observado na Figura 3a. Essa região é denominada por **pico**.
- A matriz de coeficientes utilizada na aproximação dos mínimos quadrados será nomeada por **A**.
- O termo 'frações de contagem percentual' terá o mesmo significado do termo 'frações em peso'. Esses termos farão referência às frações utilizadas nos LLS para definição do espectro total.
- O termo resíduo será utilizado para o valor de $\{\|\mathbf{N} - \mathbf{N}_{est}\|\}$ contido na Equação 4.1.
- A partir do passo construtivo do Algoritmo GRASP, alguns parâmetros serão atualizados ao longo de toda a execução do algoritmo implementado. Esses parâmetros são:
 1. Os canais inseridos nos compartimentos.
 2. Área estimada para a biblioteca omitida, nas partições definidas.
 3. Vetor de contagem espectral total nos compartimentos.
 4. Estimativa da contagem espectral da biblioteca omitida nos compartimentos. Essa estimativa é realizada de acordo com o intervalo de valores definido na Seção 4.3.
 5. Frações em peso para os $n - 1$ elementos/compostos definido inicialmente, assim como para o elemento omitido.
 6. Frações em peso para os n elementos definidos inicialmente, sem a omissão de nenhuma biblioteca. Esse resultado será importante, pois será comparado com os resultados em que $n - 1$ elementos são fornecidos inicialmente, e um deles é omitido.

Os parâmetros descritos acima são utilizados como entrada em diversas etapas do algoritmo descrito neste trabalho. Como esses parâmetros são alvo de atualizações frequentes, serão denominados **parâmetros de construção**.

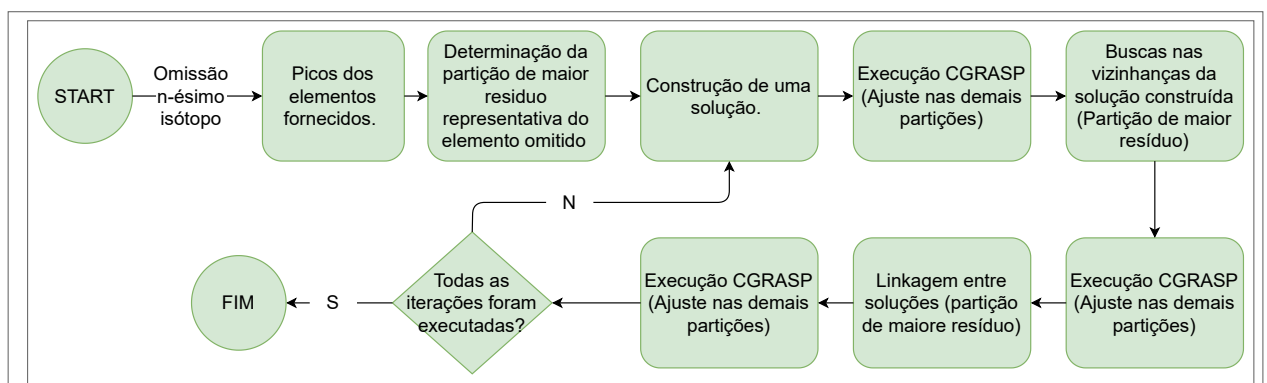
- Os desenhos espectrais para a primeira base de dados estão representada na Figura 3. Nesta Figura, são descritas as intensidades espectrais como função dos canais. É possível observar o espectro total como uma combinação linear das bibliotecas. As formas espectrais para a segunda base de dados serão omitidas, por questões de semelhança.

Figura 3 – Espectros da Base 1 para os elementos do *Am*, *Co*, *Cs* e *Na*.

Fonte: Próprio autor

- O algoritmo utilizado no presente trabalho é apresentado na Figura 4. Na Seção 4.1, serão descritos, brevemente, os procedimentos realizados para a determinação dos picos das bibliotecas fornecidas inicialmente. Na Seção 4.2, são descritos os detalhes para a determinação da partição representativa do elemento omitido e inserido no espectro total; a determinação para o intervalo de valores m na estimativa das contagens espectrais para o elemento omitido são descritos na Seção 4.3; a modelagem dos passos básicos do Algoritmo GRASP ao nosso problema estão presentes na Seções 4.4, 4.6 e 4.7. Por fim, na Seção 4.5, há a descrição para o ajuste do elemento omitido nas partições representativas das bibliotecas fornecidas. Importante enfatizar que a modelagem do Algoritmo GRASP ao domínio do problema inicia a partir da etapa nomeada por '**Construção de uma solução e estimativa do intervalo de valores para o elemento omitido**'.

Figura 4 – Sumarização da modelagem do Algoritmo GRASP ao domínio do problema. Escolhido um n -ésimo isótopo arbitrário, representado por um vetor de números. Esse vetor é omitido antes de os dados serem fornecidos como entrada ao algoritmo descrito na presente figura. São descritos como passos essenciais do método: (1) determinação dos picos e das partições aos quais pertencem, para fins de representação das bibliotecas conhecidas e fornecidas inicialmente (Seção 4.1); (2) determinação da partição que contém o maior resíduo, para fins de representação do elemento omitido, mas inserido no espectro total (Seção 4.2); (3) Construção de uma solução (Seção 4.4), definindo os parâmetros de interesse (*parâmetros de construção*) das partições e o intervalo de valores em torno do qual irá flutuar as estimativas do elemento omitido (estimativa pelo Algoritmo de Geometrização, Seção 4.3). A partir dessa etapa, destaca-se que há o início da modelagem do problema via Algoritmo GRASP; (4) Ajuste do elemento omitido, via Algoritmo CGRASP, nos compartimentos das bibliotecas fornecidas inicialmente (Seção 4.5); (5) Refinamento da solução construída, pela execução do passo de busca na partição de maior resíduo (Seção 4.5.2); (6) segue-se a mais uma execução do CGRASP; (7) execução da linkagem entre soluções, na partição de maior resíduo (Seção 4.7); (8) finalizando com mais um ajuste do CGRASP.



Fonte: Próprio autor

- As restrições matemáticas do presente trabalho na aplicação dos procedimentos descritos

na Figura 4 são:

$$\left\{ \begin{array}{l} \text{minimize} \{ \|\mathbf{N} - \mathbf{N}_{est}\| \} \\ \text{where } \mathbf{N}_{est} = \mathbf{A}\mathbf{x} + \mathbf{B} \\ \mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{N} \\ \mathbf{S}_n = \{s_{n1}, \dots, s_{nn}\}, l \leq s_{nk} \leq u \\ 0 \leq x_j \leq 1 \end{array} \right. \quad (4.1)$$

onde $\mathbf{N}$ é o vetor de contagem espectral total e $\mathbf{N}_{est}$ é o vetor de contagem espectral estimado pelos passos do presente algoritmo descrito na Figura 4. A matriz $\mathbf{A}$ contém os vetores de compostos. O vetor $\mathbf{x}$ contém as frações em peso para bibliotecas inseridas na amostra e $\mathbf{B}$ é o termo de erro. A terceira restrição indica as aproximações dos mínimos quadrados para as frações de contagem percentual dos elementos inseridos no espectro total. A quarta restrição indica a estimativa da n -ésima biblioteca, omitida do espectro total, estimada nas n regiões particionadas dos dados de entrada, onde cada componente s_{nk} tem que obedecer a um intervalo determinado, segundo o processo de geometrização (Seção 4.3). Maiores detalhes da construção dessa matriz podem ser encontradas na Seção 4.2. A quinta restrição indica que as frações de contagem percentual devem estar compreendidas entre os valores 0 e 1.

4.1 DETERMINAÇÃO DOS PICOS DAS BIBLIOTECAS FORNECIDAS

O processamento inicial dos dados divide os canais em conjuntos de 60. Cada um dos canais será percorrido, calculando-se os valores das derivadas primeira, segunda e o valor da seguinte Equação 4.2.

$$T = m + \delta \quad (4.2)$$

onde m é a média e δ é o desvio padrão nos cinco vizinhos mais próximos ao canal corrente. A vantagem em se escolher a Equação 4.2 para detecção de picos é que ela apresenta robustez na verificação de pequenas transformações nos N vizinhos mais próximos, onde N é um número pré-determinado. Para o presente trabalho escolheu-se o valor de N igual a 5 (ALAMANIS; JEVREMOVIC, 2015).

Não foram utilizados métodos de suavização para determinação das derivadas. As derivadas foram discretizadas, onde a diferença infinitesimal foi usada como a diferença entre dois pontos

contíguos. O espectro foi assumido como uma discretização de uma função em que é possível garantir a continuidade da derivada primeira e segunda, ou seja, uma discretização de uma função C_2

A saída para o algoritmo de determinação de picos será uma estrutura de dados, contendo o canal do pico e a biblioteca ao qual esse canal pertence.

O método de detecção de picos é uma adaptação de técnicas tradicionais para encontrar regiões de inflexão, como cálculo de derivadas, com procedimentos utilizados no trabalho de Alamaniotis e Jevremovic (2015) para detecção de picos por lógica Fuzzy.

4.2 DETERMINAÇÃO DA PARTIÇÃO DE MAIOR RESÍDUO

A partição de maior resíduo será definida como aquela que apresentar maior diferença entre o espectro estimado a partir dos $n - 1$ elementos fornecidos inicialmente com seus picos e o espectro total. Parte-se do pressuposto que essa partição contém informações relevantes sobre a biblioteca omitida, já que nessa região provavelmente localiza-se o pico desse elemento, devido ao elevado valor do resíduo. Essa partição é definida como a representativa para o pico da biblioteca omitida.

A definição da partição de maior resíduo segue-se à determinação dos picos. Nessa etapa, a restrição 1 da Equação 4.1 será aplicada a $n - 1$ bibliotecas, para n igual ao número de elementos inseridos no espectro total.

As entradas para essa parte do algoritmo serão:

- Os picos definidos na Seção 4.1;
- As partições às quais os picos das bibliotecas conhecidas pertencem;
- As $n - 1$ bibliotecas conhecidas;
- O vetor contendo as contagens do espectro total.

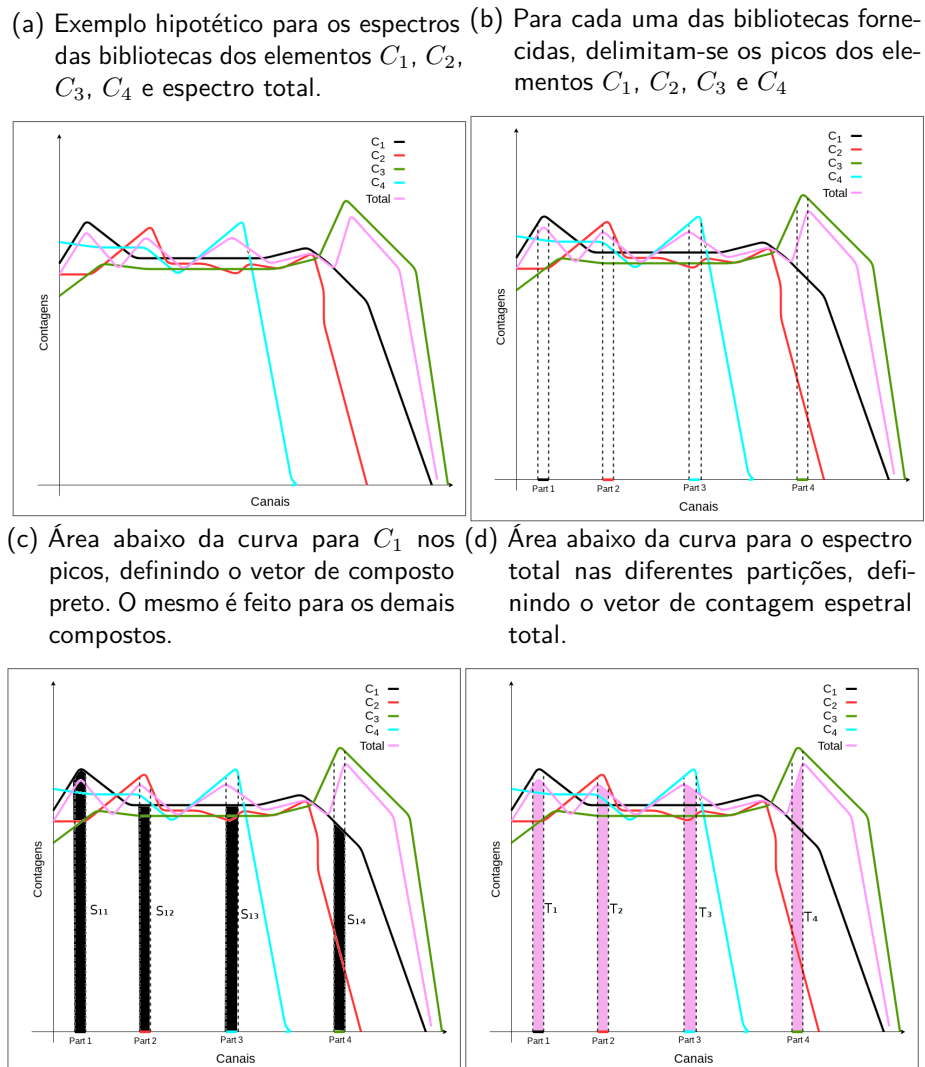
As saídas, após o processamento das entradas, serão:

- Os valores de resíduos em cada partição;
- Os vetores de compostos, as áreas abaixo da curva, para uma biblioteca arbitrária, ao longo das partições;
- Os vetores de contagem espectral total ao longo das partições;

- O máximo valor de resíduo encontrado e a partição em que esse valor de resíduo foi encontrado.

O passo inicial consiste na determinação dos vetores de compostos, e do vetor de contagem espectral total, em todas as partições dos dados, segundo descrito na Figura 5. Como o vetor de contagem espectral total é definido como combinação linear das bibliotecas, com as frações em peso desempenhando o papel dos coeficientes na combinação, o vetor de contagem espectral total tende a se encontrar definido na região compreendida pelas bibliotecas.

Figura 5 – Descrição dos procedimentos realizados para a construção dos vetores de compostos e do vetor de contagem espectral total.



Fonte: Próprio autor

A determinação dos vetores de compostos, para as bibliotecas fornecidas inicialmente na definição da contagem espectral total e do vetor de contagem espectral total, conduz à

definição do seguinte sistema de equações na determinação do N_{est} (vetor de contagem total estimado) (restrição 2, Equação 4.1) e posterior cálculo do resíduo (restrição 1, Equação 4.1):

$$\underbrace{\begin{bmatrix} s_{11} & s_{21} & s_{31} \\ s_{12} & s_{22} & s_{32} \\ s_{13} & s_{23} & s_{33} \\ s_{14} & s_{24} & s_{34} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} \end{bmatrix}}_A \cdot \underbrace{\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}}_x = \underbrace{\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \mathbf{1} \end{bmatrix}}_N \quad (4.3)$$

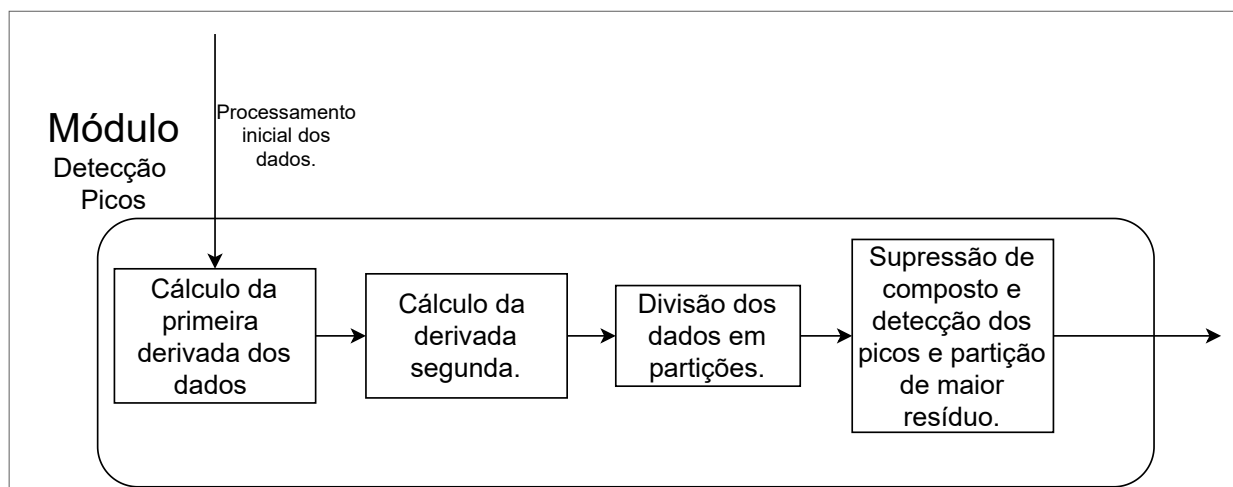
onde s_{ij} é a componente do vetor de compostos i na partição j . Cada componente é definida como a soma das contagens espectrais para o compostos i na partição j , de acordo com a Figura 5c; α , β e γ são os valores das frações percentuais dos compostos e t_j é a componente do vetor de contagens espectrais totais na partição j . Essa componente é a soma das contagens espectrais totais na partição j .

Os valores das frações α , β e γ são determinados pela seguinte aproximação dos mínimos quadrados:

$$\mathbf{x} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot (\mathbf{A})^T \cdot \mathbf{b} \quad (4.4)$$

Na Figura 6, é apresentado o sumário do processamento inicial dos dados presentes na Seção 4.1 e nesta seção. Maiores detalhes sobre o módulo de determinação dos picos e partições podem ser encontrados no Apêndice C, Seção C.2.

Figura 6 – Sumarização das etapas de determinação dos picos (Seção 4.1) das bibliotecas fornecidas inicialmente, e da partição de maior resíduo.



Fonte: Próprio autor

4.3 DESCRIÇÃO DA ETAPA DA GEOMETRIZAÇÃO

A modelagem da geometrização para estimativa de um elemento omitido do espectro total está presente na etapa denominada '**Construção de uma solução e estimativa do intervalo de valores para o elemento omitido.**' na Figura 4 (Seção 4.4). Logo, os procedimentos realizados para estimativa do intervalo de valores da biblioteca omitida são realizados no âmbito do passo construtivo do GRASP. A introdução de um algoritmo segundo uma modelagem geométrica é uma das contribuições inovadoras deste trabalho, já que esses mesmos procedimentos não são encontrados em trabalhos similares do estado da arte. Vale lembrar que o termo 'Geometrização' é utilizado para se referir à proposta da definição do intervalo de valores em torno do qual irá ser estabelecida as estimativas da biblioteca omitida. Essa estimativa será realizada por meio de contagens espectrais constantes ao longo dos compartimentos definidos dentro de cada uma das partições.

Parte-se do pressuposto que os compostos, representados como ênuplas de valores de contagem, em que cada dimensão é uma dada partição de canais, geram um espaço de compostos (Figura 7a). Nesta representação, cada composto corresponde a um ponto do espaço. Combinando compostos básicos, podem-se obter outros compostos através de médias ponderadas dos mesmos, ou seja, os pesos somam a 1 (já que são frações de peso). Três pontos não colineares, quando combinados dessa forma, usando todos os valores de pesos possíveis, geram um plano de compostos. A etapa de geometrização define um plano no espaço $n - 1$ dimensional para um conjunto de pontos n dimensionais, onde n é o número de bibliotecas.

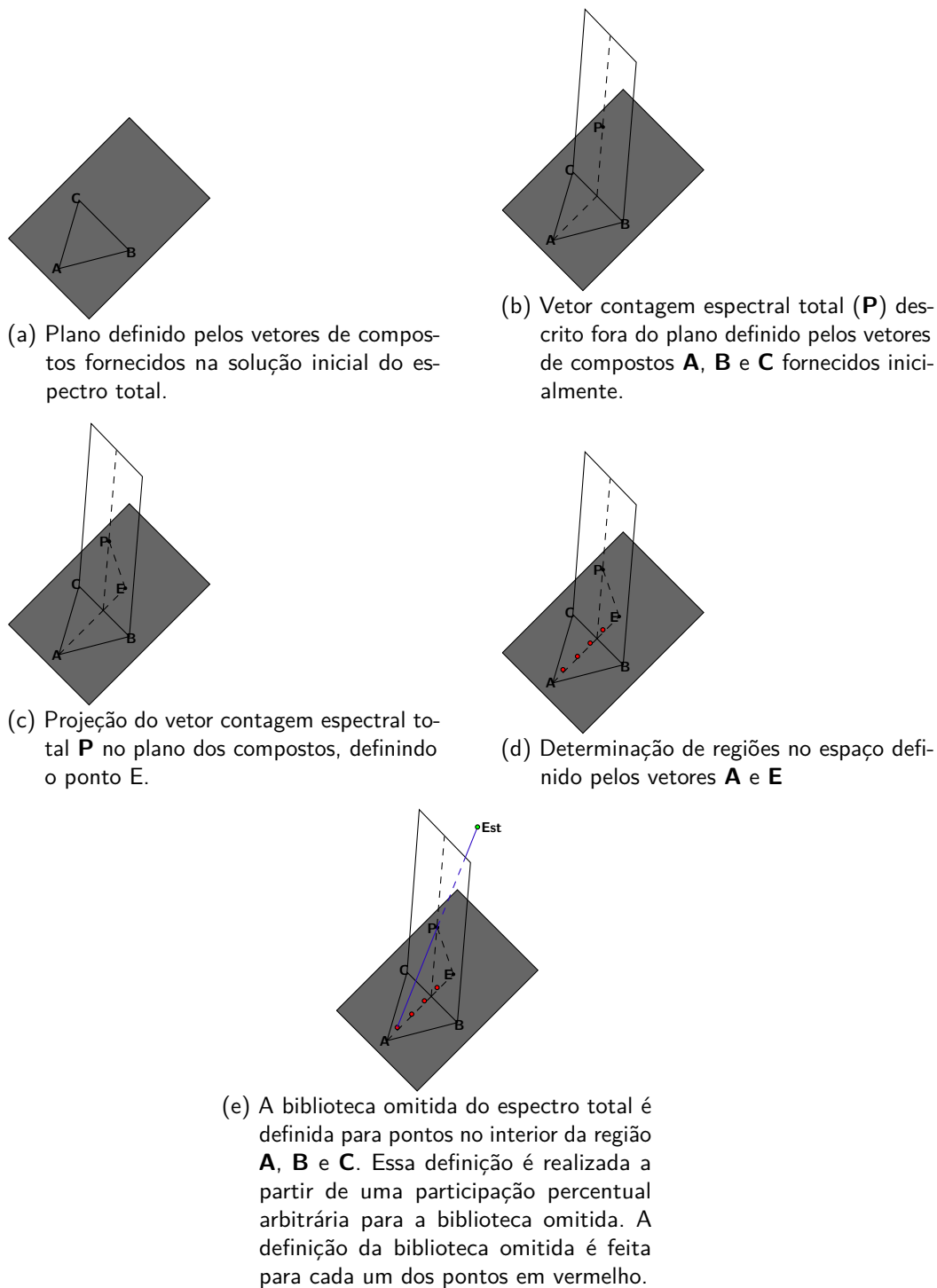
Considere que os três constituintes que geram o plano são a solução inicial para o problema do MCLLS. Se por acaso houver um constituinte a mais que não foi levado em conta na solução inicial e não foi considerado nessa construção geométrica, este constituinte extra é representado como um ponto fora do plano gerado pelos três constituintes. Quando se combinam esses quatro constituintes para produzir o espectro total, o que se obtém é um ponto também fora do plano gerado pelos três da solução inicial (ponto P na Figura 7b). Logo, se por acaso se ignora a existência do constituinte extra, e forma-se o sistema tendo apenas os três da solução inicial como geradores do composto total, obtém-se um sistema sem solução (na geometria: o ponto P na Figura 7b está fora do plano, e os três pontos do plano só podem gerar compostos cujos pontos estão no plano). Dessa forma, para resolver o sistema, utiliza-se o MMQ, que oferece uma solução dentro do plano mas que seja o mais próximo possível de P , ou seja, a solução MMQ é a projeção ortogonal de P sobre o plano gerado pelos constituintes da solução inicial (ponto E na Figura 7c).

O vetor de contagem espectral total não pertence ao plano gerado pelos vetores de constituintes básicos definidos na Figura 7b, levando ao pressuposto de um sistema descrito em função dos três vetores de constituintes contido no plano definido por eles, mais um quarto vetor não pertinente ao mesmo plano (vetor nomeado por *Est* na Figura 7e). Para isso, uma forma de reduzir o espaço de busca pelo constituinte omitido é:

- Faz-se a projeção ortogonal do vetor de contagem espectral total no plano definido pelos três vetores de constituintes básicos (Figura 7c), após a definição do plano definido pelos vetores de compostos (Figura 7a) e do vetor de contagem espectral total (Figura 7b);
- Em função desses três vetores de compostos básicos, verifica-se qual deles tem coordenada baricêntrica negativa (Figura 7d) na projeção do vetor de contagem total no plano de compostos. Nas experimentações, caso a projeção indique percentuais todos positivos de participação do vetores de constituintes, escolhe-se um ponto arbitrário para prosseguir na etapa descrita na Figura 7e;
- Utiliza-se a região do espaço determinada pela solução do MMQ e pelo vetor de compostos que tenha apresentado coordenada baricêntrica negativa após a projeção do MMQ (Segmento $\overline{AE}$ Figura 7d);
- Vários pontos são escolhidos na região definida no último item e, para cada um deles, faz-se a projeção da biblioteca omitida na solução inicial do espectro total (Figura 7e). A

pesquisa é realizada no triângulo invertido que completa o tetraedro gerado pelo ponto P e os constituintes básicos.

Figura 7 – Método de estimativa da altura nas partições.



Fonte: Próprio autor

O pseudocódigo para estimação das contagens espectrais da biblioteca omitida, encontra-se

no Apêndice C, Seção C.6.

São fornecidos como parâmetros de entrada para o Algoritmo 13:

1. O compartimento onde se deseja estimar o intervalo para a biblioteca omitida;
2. As bibliotecas de contagens espectrais fornecidas inicialmente ao espectro total;
3. A biblioteca de contagem espectral total;
4. As partições onde não esteja realizando a estimativa do elemento omitido;
5. Estimativa inicial de percentual para a biblioteca omitida da solução inicial.

As estimativa dentro da n -ésima partição, após a definição do compartimento, será função dos valores de contagem espectral dos elementos fornecidos inicialmente.

Como saída para o algoritmo, destaca-se:

1. Os valores de intervalo de contagem espectral para o elemento omitido, dentro do compartimento em questão.

4.4 MODELAGEM DO PASSO CONSTRUTIVO ALGORITMO GRASP

A partir da etapa descrita na presente seção, o passo construtivo, inicia-se a modelagem do problema via Algoritmo GRASP. O pseudocódigo para essa modelagem pode ser encontrado no Apêndice C.1, Seção C.3. O passo construtivo, ao realizar a aproximação das frações em peso para as bibliotecas, irá adicionar uma coluna à matriz de coeficientes contida na Equação 4.3. Essa coluna corresponde ao vetor de contagens espectrais estimados, determinando, assim, a primeira aproximação para os mínimos quadrados do algoritmo da Figura 4.

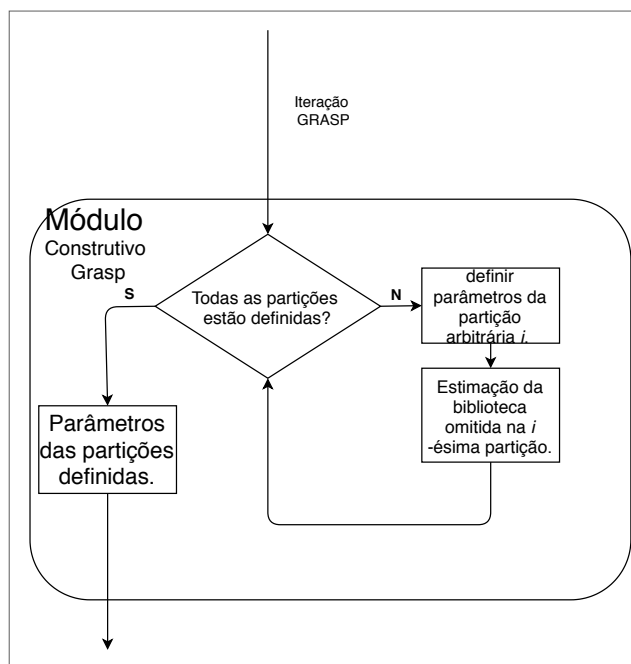
Pode-se enumerar como entradas para o passo construtivo do Algoritmo GRASP:

- Os parâmetros das partições, como: valores de resíduos, os vetores de compostos e o vetor de contagem espectral total das partições;
- As contagens espectrais totais e as bibliotecas individuais;
- Número mínimo de canais que podem ser escolhidos das partições para definir os compartimentos. No presente trabalho, convencionou-se que esse número mínimo sejam 10 canais;

- Número inteiro que determina que biblioteca será aproximada.

As saídas serão os *parâmetros de construção*. Esses parâmetros servirão como entrada para o algoritmo de ajuste das alturas no Algoritmo CGRASP, como também serão utilizados em outros módulos do algoritmo implementado na presente tese. O resumo do passo construtivo pode ser encontrado na Figura 8.

Figura 8 – Sumarização da etapa construtiva.



Fonte: Próprio autor

4.5 AJUSTE DO ELEMENTO OMITIDO VIA MODELAGEM DO CGRASP

O Algoritmo CGRASP encarrega-se de realizar a estimativa do elemento omitido nos compartimentos representativos das bibliotecas fornecidas inicialmente. O algoritmo é composto por dois sub-algoritmos: (1) algoritmo responsável pela construção de uma solução (Seção 4.5.1); (2) rotina responsável por realizar buscas nas vizinhanças da solução construída (Seção 4.5.2).

Na chamada do Algoritmo CGRASP, destacam-se os seguintes parâmetros de entrada:

- A melhor solução até então encontrada;
- Os parâmetros de construção, gerados como saída em algum dos passos básicos do GRASP;

- Dimensionalidade em torno da qual o algoritmo irá realizar a busca. Para o presente problema, a dimensionalidade é definida como $n = 2$;
- Número de vezes em que o CGRASP será executado na expectativa de alcançar uma melhoria;
- Ponteiro que indica o número máximo de iterações sem melhorias. Ao atingir esse número máximo, a discretização da dimensão pesquisada será aumentada, na tentativa de obter melhores soluções;
- Número máximo de direções em que o algoritmo irá fazer a busca;
- Constante α a ser utilizada no algoritmo de construção do CGRASP.

Como saída para a chamada do CGRASP, destaca-se:

- Os parâmetros de construção atualizados;
- O menor valor de resíduo, e a solução contendo as frações em peso atualizadas.

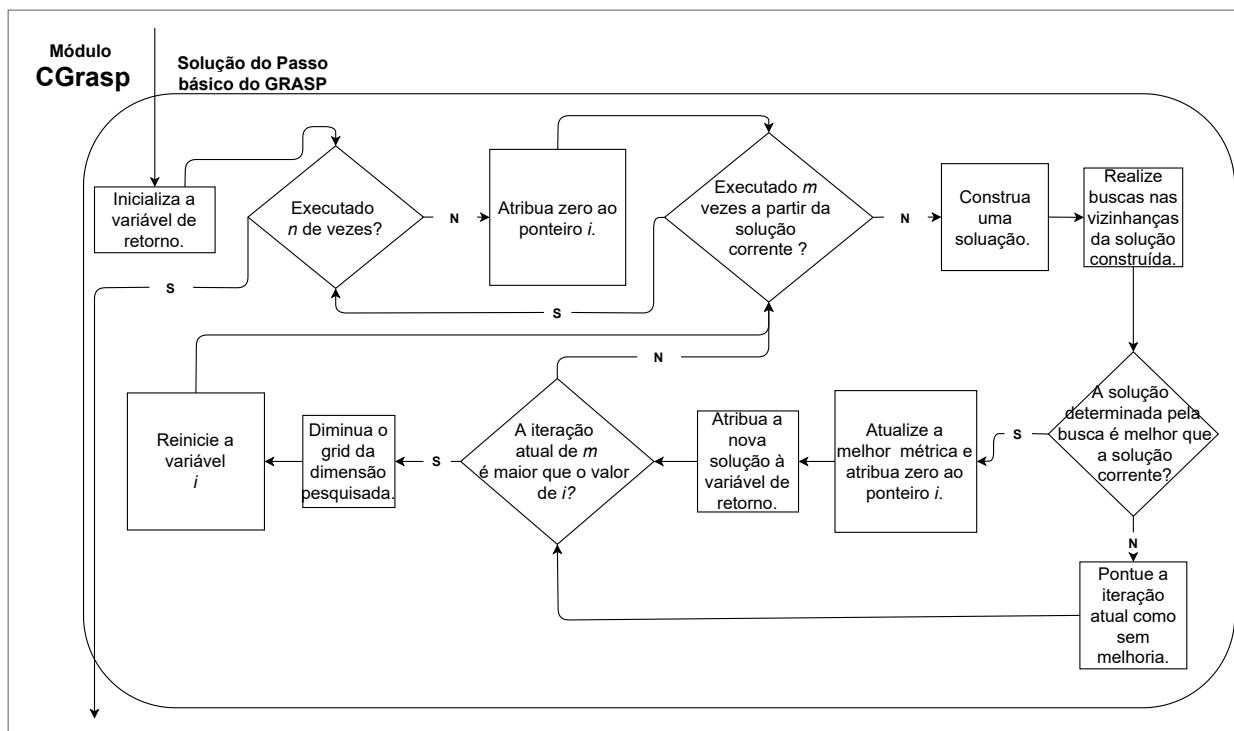
Os procedimentos realizados para determinar a estimativa do elemento omitido pela modelagem do CGRASP estão contidos na Figura 9. O Algoritmo CGRASP é executado n vezes. O ponteiro i é definido para verificar o número de iterações sem melhorias. Um novo processo iterativo é iniciado, sendo executados m vezes para uma solução corrente arbitrária. Constrói-se uma solução e buscas são realizadas nas vizinhanças dessa solução construída. A solução corrente é atualizada se os resultados da busca indicarem uma nova solução melhor que a solução atual. Se não for encontrada uma solução melhor que a corrente após o passo da busca, a iteração m atual é apontada como uma iteração sem melhorias. Verifica-se se o ponteiro i é menor que o valor de m . Se isso for verdade, diminui-se a discretização da dimensão contínua em torno da qual ocorre a pesquisa do algoritmo. No presente trabalho, essa dimensão é definida como a estimativa das contagens espectrais do elemento omitido, limitado pelos valores permitidos para essas estimativas no processo de geometrização.

O pseudocódigo do CGRASP modelado para o problema em questão pode ser encontrado no Apêndice C.7.

4.5.1 Modelagem do Ajuste via CGRASP - Passo construtivo

A presente seção descreverá a construção de uma solução. Antes de realizar as considerações do algoritmo, será descrito como foi realizada a modelagem de dimensão no contexto dos

Figura 9 – Sumarização do Algoritmo CGRASP.



Fonte: Próprio autor

algoritmos do CGRASP. A dimensão enumerada na presente seção é modelada como o par $\#partição, estimativa$, onde $\#partição$ é a partição escolhida e $estimativa$ é o valor da contagem espectral da biblioteca omitida. Para uma base de dados compostos por quatro bibliotecas, e levando em conta que as estimativas para a biblioteca omitida podem ser incrementadas ou decrementadas, o número de dimensões possíveis passa a ser $4 \times 2 = 8$ dimensões.

O passo construtivo possui um mecanismo de pesquisa, ao visitar todas as dimensões, e determinar se a solução associada à dimensão está inserida em um intervalo, com limites definidos pelo próprio algoritmo de construção. A solução associada a essa dimensão é salva, dentro de uma lista, em uma primeira etapa, caso essa solução esteja no interior desse intervalo. Por fim, seleciona-se aleatoriamente uma das soluções dentro da lista para ser a nova solução corrente. O pseudocódigo que descreve o algoritmo de construção do CGRASP pode ser encontrado no Apêndice C.8, Algoritmo 15.

Na chamada da função de construção do CGRASP, podem-se destacar os seguintes parâmetros:

1. Os *parâmetros de construção* atualizados por um dos passos básicos do GRASP;

2. Métrica associada à solução corrente;
3. O número de direções em torno do qual a construção realizará a pesquisa para configuração do intervalo utilizado na avaliação das soluções;
4. A densidade inicial do grid que determina como será percorrida a contagem espectral do elemento omitido;
5. Os valores de máximo e mínimo a serem respeitados para os valores de estimativa da biblioteca não fornecida;
6. Parâmetro arbitrário compreendido entre $[0, 1]$, utilizado na avaliação das soluções associadas às diferentes soluções.

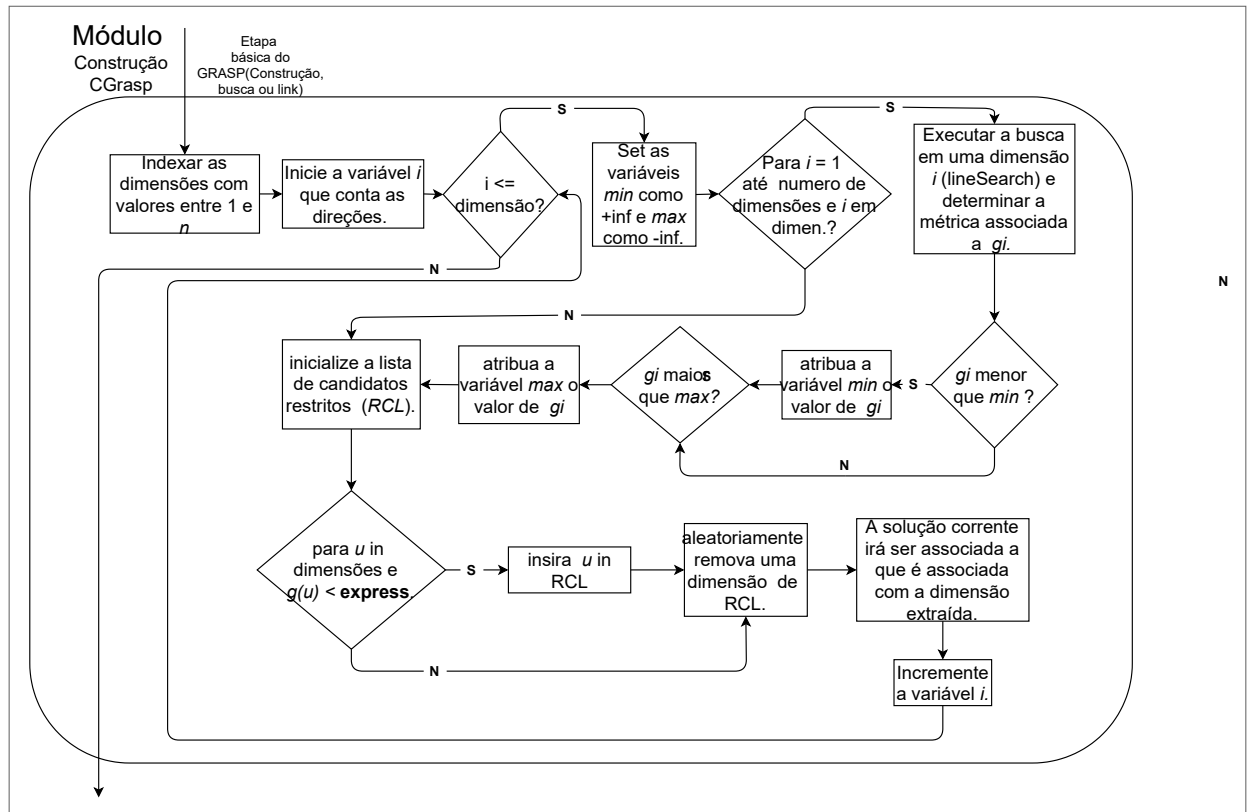
Enumeram-se como saídas para o algoritmo desta seção:

1. Os *parâmetros de construção* atualizados após o passo construtivo;
2. Menor valor de métrica após o passo construtivo.

A Figura 10 apresenta um resumo do passo construtivo no ajuste das alturas nas partições para a biblioteca omitida do espectro total. As prováveis dimensões utilizadas na construção de uma solução são indexadas pelo intervalo $[1, \dots, n]$, onde n é o número de dimensões possíveis. Os ponteiros *max* e *min* são definidos e inicializados, seguindo-se a uma visita a cada uma das dimensões. As visitas às dimensões irão atualizar os valores das variáveis *max* e *min*. Se uma das dimensões atende ao critério que o valor funcional é menor que uma expressão pré-definida, a dimensão é escolhida para ser inserida na lista associada ao processo construtivo. Ao fim da visita a todas as dimensões, uma dimensão é escolhida arbitrariamente da lista e a solução associada a essa dimensão passa a ser a solução da etapa de construção. Uma nova iteração é iniciada, mas a dimensão escolhida na iteração anterior é descartada em uma nova pesquisa. Esse processo é denominado *fixar a dimensão*. O processo prossegue até que todas as direções tenham sido fixas.

A Figura 11 traz os procedimentos fundamentais na construção de uma solução do Algoritmo CGRASP.

Figura 10 – Sumarização dos procedimentos tomados durante a execução do algoritmo de construção do CGRASP



Fonte: Próprio autor

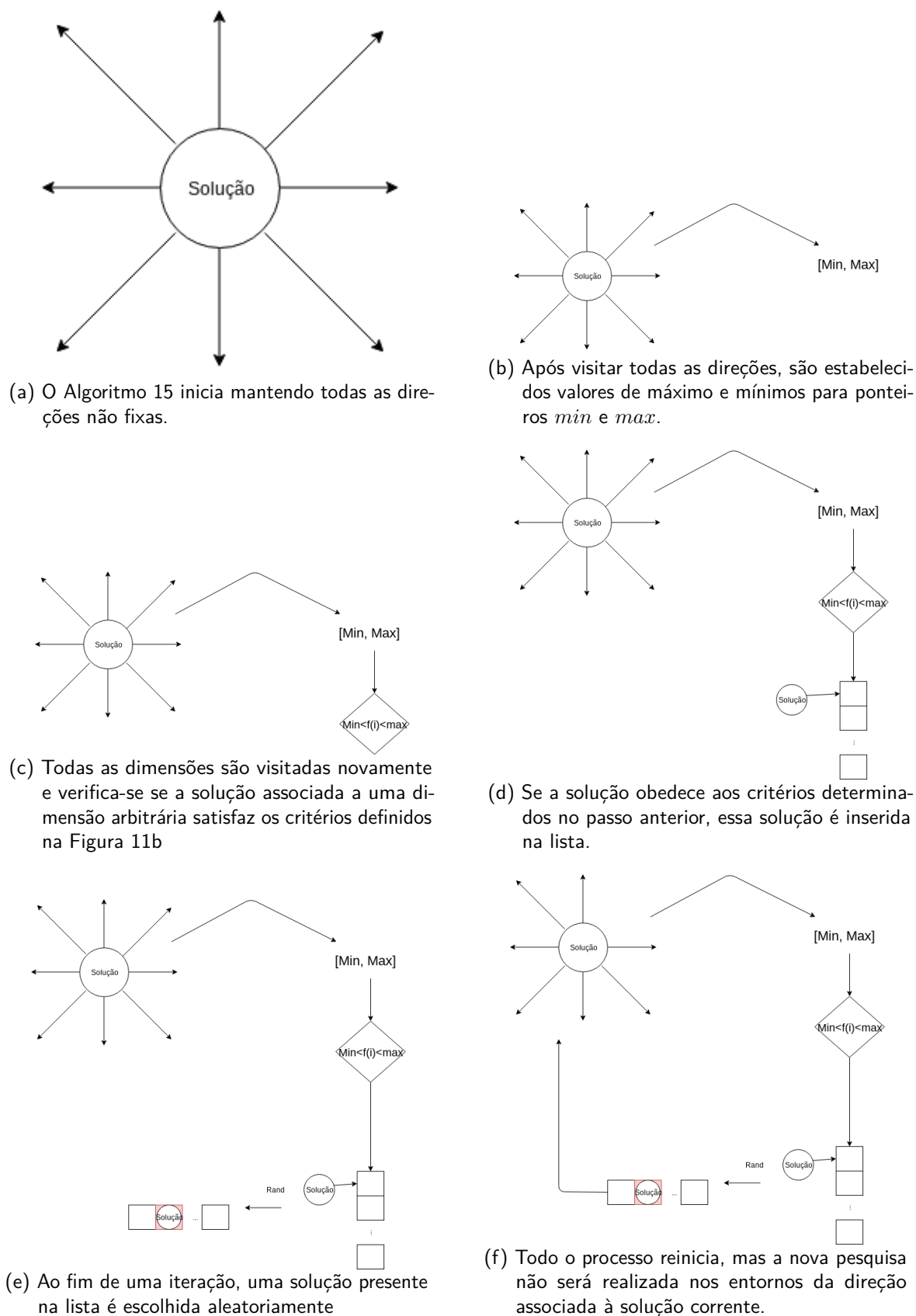
4.5.2 Modelagem do Ajuste via CGRASP - Passo de Busca

As saídas do passo construtivo do Algoritmo CGRASP, Seção 4.5.1, serão fornecidas como entrada para o algoritmo descrito na presente seção. A descrição do pseudocódigo referente a essa etapa pode ser encontrada no Apêndice C.9, Algoritmo 16. A modelagem de dimensões para este algoritmo é a mesma modelagem realizada na seção do passo construtivo.

Pode-se enumerar como entradas para o algoritmo de Busca do CGRASP:

1. Parâmetros de construção determinados no passo construtivo;
2. Valor da função objetivo a ser minimizada;
3. Número de dimensões em torno do qual a busca será realizada;
4. Intervalo $[l, u]$ que deve conter os valores de estimativa da biblioteca omitida;
5. Número máximo de tentativas a ser realizado na busca por soluções melhores que a atual;

Figura 11 – Processo construtivo do Algoritmo CGRASP, no ajuste da biblioteca omitida nas partições representativas das bibliotecas fornecidas inicialmente.



Fonte: Próprio autor

6. Densidade do grid que determina a discretização da contagem espectral (variável contínua) do elemento omitido nas partições que representam as bibliotecas fornecidas.

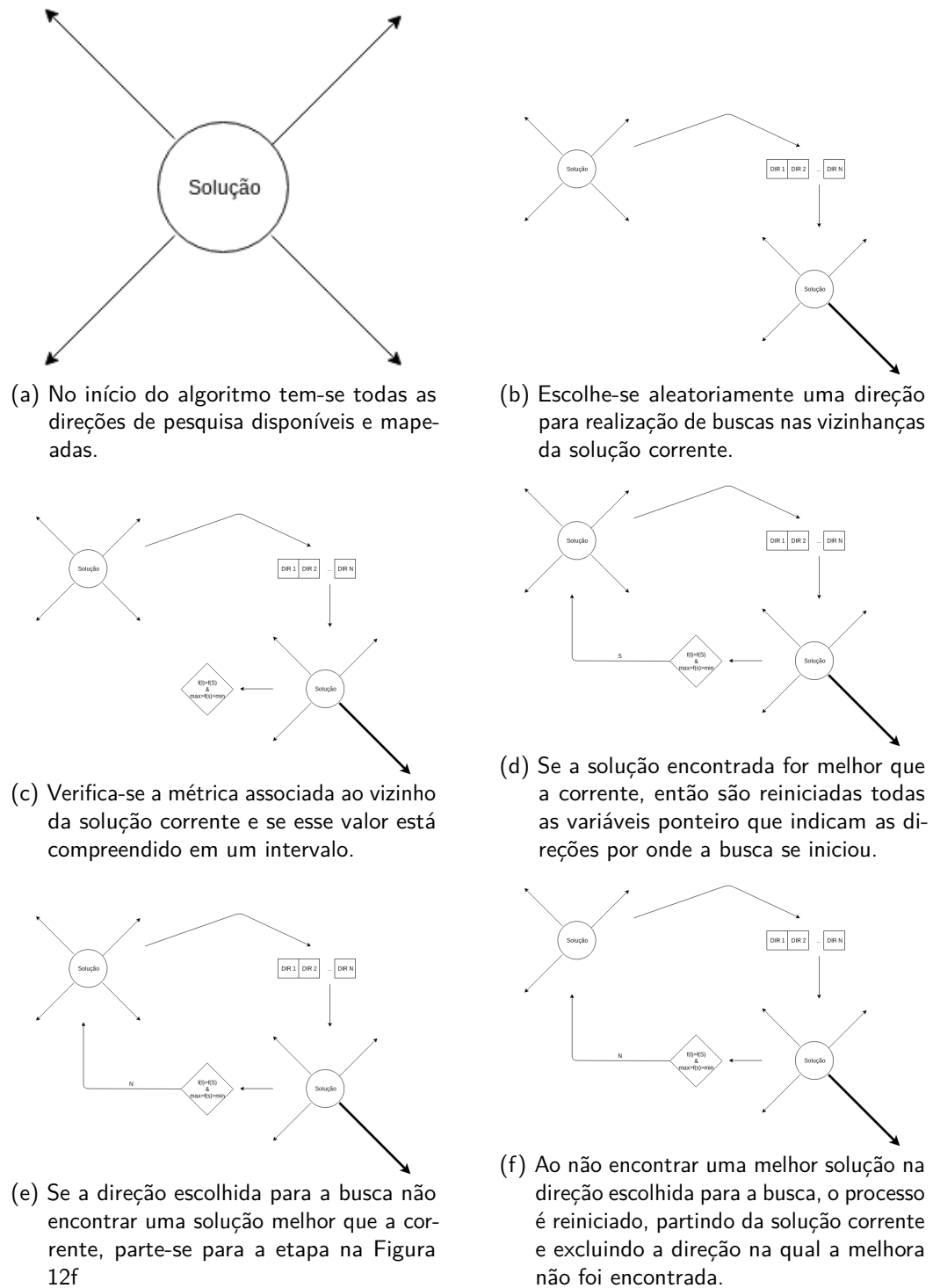
As seguintes saídas são estabelecidas:

1. Os *parâmetros de construção* atualizados após a etapa de busca do CGRASP;
2. O menor valor da métrica atualizado, após o processo da busca do CGRASP.

A Figura 12 demonstra a dinâmica envolvida na busca por soluções melhores que as definidas no passo construtivo do CGRASP.

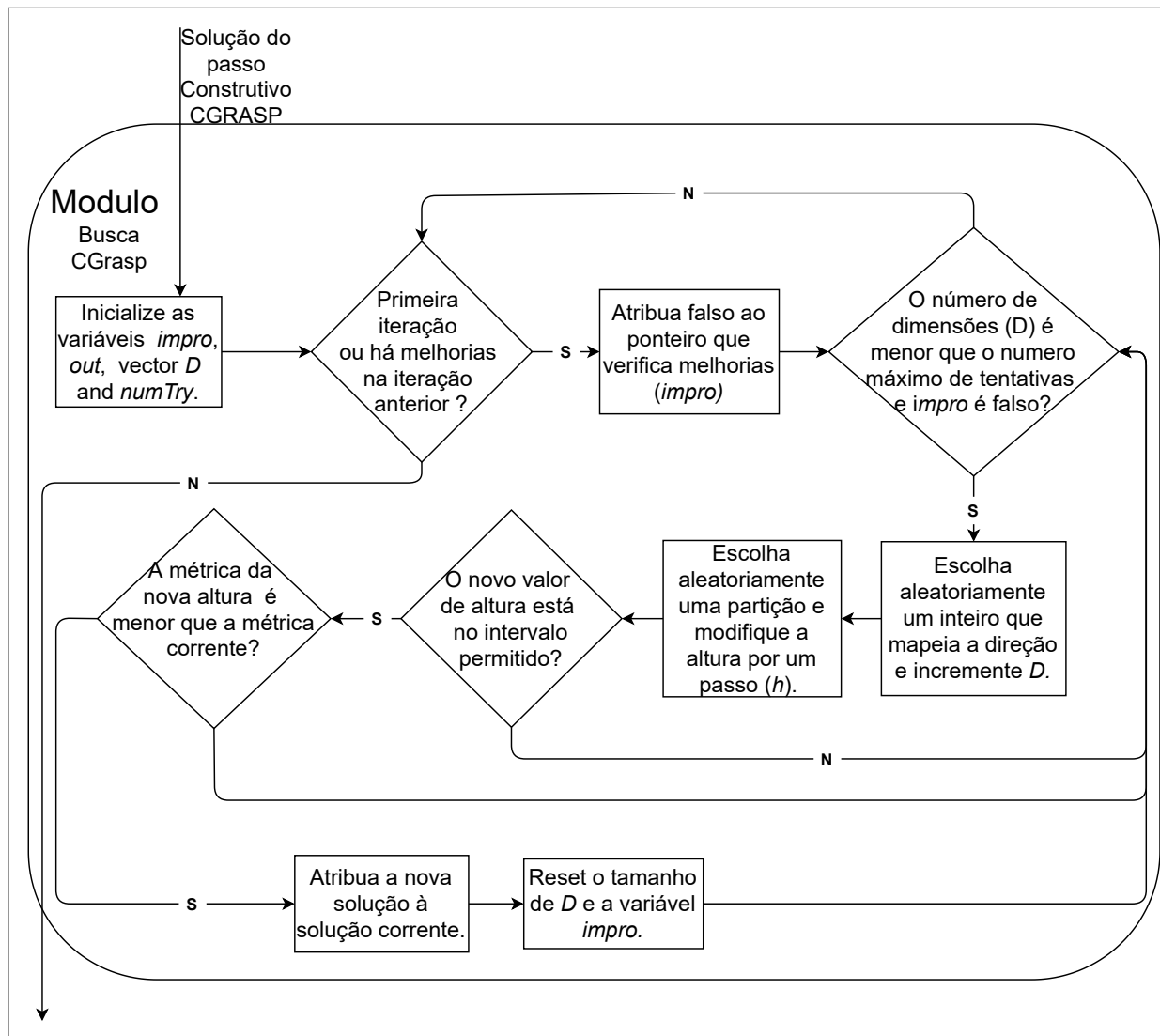
O sumário do processo de busca do CGRASP está descrito na Figura 13. A solução vinda do passo construtivo é fornecida como entrada para a etapa de busca. Após a inicialização das variáveis *impro* (verifica melhorias no algoritmo), (*Vector D*) (vetor de dimensões) e da variável *numTry* (parâmetro a ser utilizado caso não se queira pesquisar todas as dimensões), verifica-se se a variável *impro* é verdadeira e se o número de iterações permitidos não foi ultrapassado. Verificado se o tamanho de dimensões é menor que o número máximo de tentativas a serem realizadas, e se nenhuma melhoria até então foi encontrada, escolhe-se um inteiro que define a direção a ser mapeada. A direção, nesse caso, é um valor inteiro que mapeia a contagem espectral corrente para cima ou para baixo. Após escolher aleatoriamente a direção de modificação da contagem espectral, a partição que receberá a modificação é escolhida aleatoriamente. Se a nova altura estiver no intervalo permitido para a partição, determina-se se a solução é melhor que a solução corrente, após a atualização dos parâmetros de construção. Sendo verdadeiro, a nova solução será a solução corrente, e as variáveis *tamanho de D* (para que todas as direções voltem a ser alvo de pesquisa) e *impro* serão reinicializadas.

Figura 12 – Processo de busca do Algoritmo CGRASP, no ajuste da biblioteca omitida nas partições representativas das bibliotecas fornecidas inicialmente.



Fonte: Próprio autor

Figura 13 – Sumarização dos procedimentos tomados durante a execução do Algoritmo de busca do CGRASP



Fonte: Próprio autor

4.6 MODELAGEM DO PASSO DE BUSCA DO ALGORITMO GRASP

O Algoritmo de busca receberá os parâmetros descritos no processo construtivo do GRASP, Seção 4.4. O Algoritmo de busca será o responsável por definir e modelar vizinhos a partir da solução vinda do passo construtivo. Realizada a modelagem, a cada novo vizinho definido os parâmetros associados a esse vizinho serão calculados. Se o novo vizinho for melhor que a solução corrente, esse vizinho será definido, então, como a nova solução corrente. O pseudocódigo e detalhamento dos passos executados podem ser encontrados no Algoritmo 10, Apêndice C, Seção C.4. Dentre os parâmetros de entrada para o Algoritmo de Busca, pode-se citar:

1. Os elementos associados aos *parâmetros de construção* definidos no passo construtivo do Algoritmo GRASP;
2. O tamanho da lista do passo de busca;
3. O número máximo de iterações que será executado na pesquisa por melhores soluções;
4. As partições dos dados;
5. O valor mínimo de resíduo até então encontrado.

Pode-se enumerar como saídas para o Algoritmo de Busca:

1. Os *parâmetros da construção* de uma solução, atualizados;
2. O novo valor de mínimo de métrica a ser objetivada nas iterações seguintes.

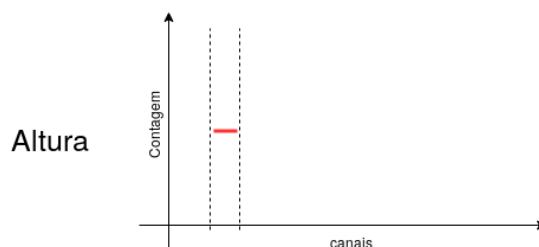
A definição de vizinhos no Algoritmo 10 será dada por um par de valores $(alt, larg)$, onde alt é o valor de contagem espectral do elemento omitido e $larg$ é a janela de canais utilizada na definição biblioteca omitida do espectro total, ambos na partição de maior resíduo. A busca definirá vizinhos pela modificação de uma única dessas variáveis.

A figura 14 descreve a definição de vizinhos, a partir da solução corrente, pela modificação nos valores de contagem espectral do elemento omitido na partição de maior resíduo. As contagens espectrais nessa partição vão ser modificadas segundo um constante aleatória, mas que definem a nova estimativa da biblioteca omitida dentro do intervalo definido no processo de geometrização.

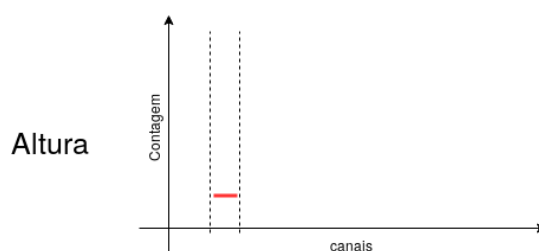
A definição de vizinhos, a partir da modificação dos canais da solução corrente, é descrita na Figura 15. Os canais inseridos, ou retirados, foram configurados de modo que não necessitassem ser contíguos. É possível adicionar ao fim de um compartimento um canal que não é o canal seguinte ao último canal do compartimento, definindo, assim, “descontinuidades” nos canais inseridos nos compartimentos. Da mesma maneira, também foi possível configurar a busca, de modo que fosse possível retirar canais que se encontrassem no meio do compartimento, definindo também “descontinuidades” no compartimento pela retirada de canais.

O resumo da etapa de busca do GRASP está descrito na Figura 16. Como etapa inicial, modifica-se a contagem espectral da biblioteca omitida, na partição de maior resíduo, com consequente recálculo dos *parâmetros de construção* e da métrica. A lista do algoritmo é preenchida se a nova solução for melhor que a solução corrente. Se não for melhor que a

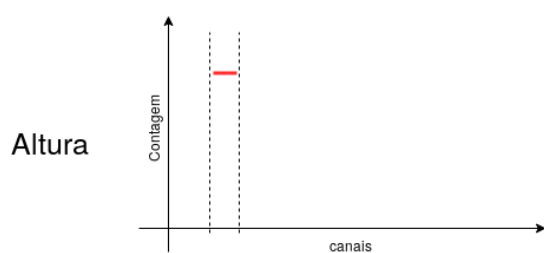
Figura 14 – Variação nos valores de altura para o composto omitido na partição de maior resíduo, definindo vizinhos da solução corrente.



- (a) O valor de contagem espectral para a biblioteca omitida no espectro total é feito por meio de um número aleatório que determinará essa contagem dentro do intervalo estimado permitido, de acordo com o descrito na Seção 4.3.



- (b) O valor da contagem espectral modificado por meio de um novo valor aleatório que determinará uma contagem espectral menor que a contagem corrente. Esse valor que diminuirá a contagem espectral estará necessariamente compreendido no intervalo $(0, 1)$. Esse novo valor está compreendido no intervalo definido no processo de geometrização.



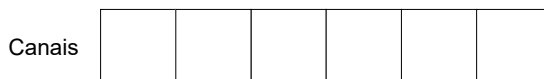
- (c) O valor da contagem espectral modificado por meio de um novo valor aleatório que determinará uma contagem espectral maior que a contagem corrente. Esse valor que aumentará a contagem espectral será necessariamente maior que 1.

Fonte: próprio autor

Figura 15 – Busca por soluções vizinhas à solução atual



- (a) A solução inicial na partição de maior resíduo pode ter, por exemplo, 8 canais.



- (b) Na busca por vizinhos à solução apresentada em 15a são retirados dois canais na partição de maior resíduo.

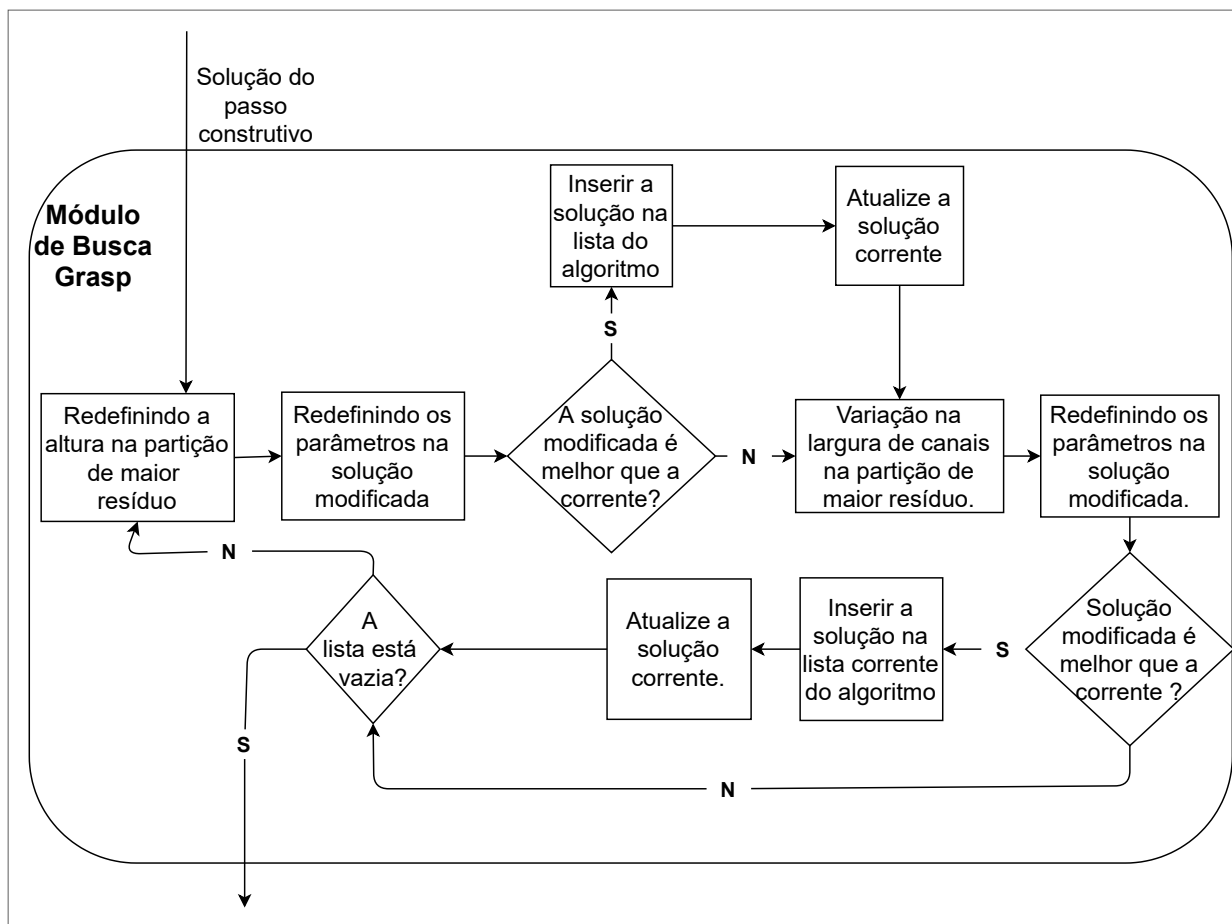


- (c) Os vizinhos da solução apresentada em 15a também são buscados por meio da adição de canais à solução corrente.

Fonte: próprio autor

solução corrente, procede-se à modificação dos canais inseridos na partição de maior resíduo. Na etapa seguinte, após a inserção de canais na partição de maior resíduo, procede-se ao recálculo dos parâmetros de construção. A nova solução sendo melhor que a solução corrente, esta última é substituída por essa nova solução. O processo de busca encerra-se se a lista a ele associada estiver vazia ao fim de uma iteração.

Figura 16 – Sumarização do processo de busca por soluções melhores que a solução corrente.



Fonte: próprio autor

4.7 MODELAGEM DO PASSO DE LINKAGEM ALGORITMO GRASP

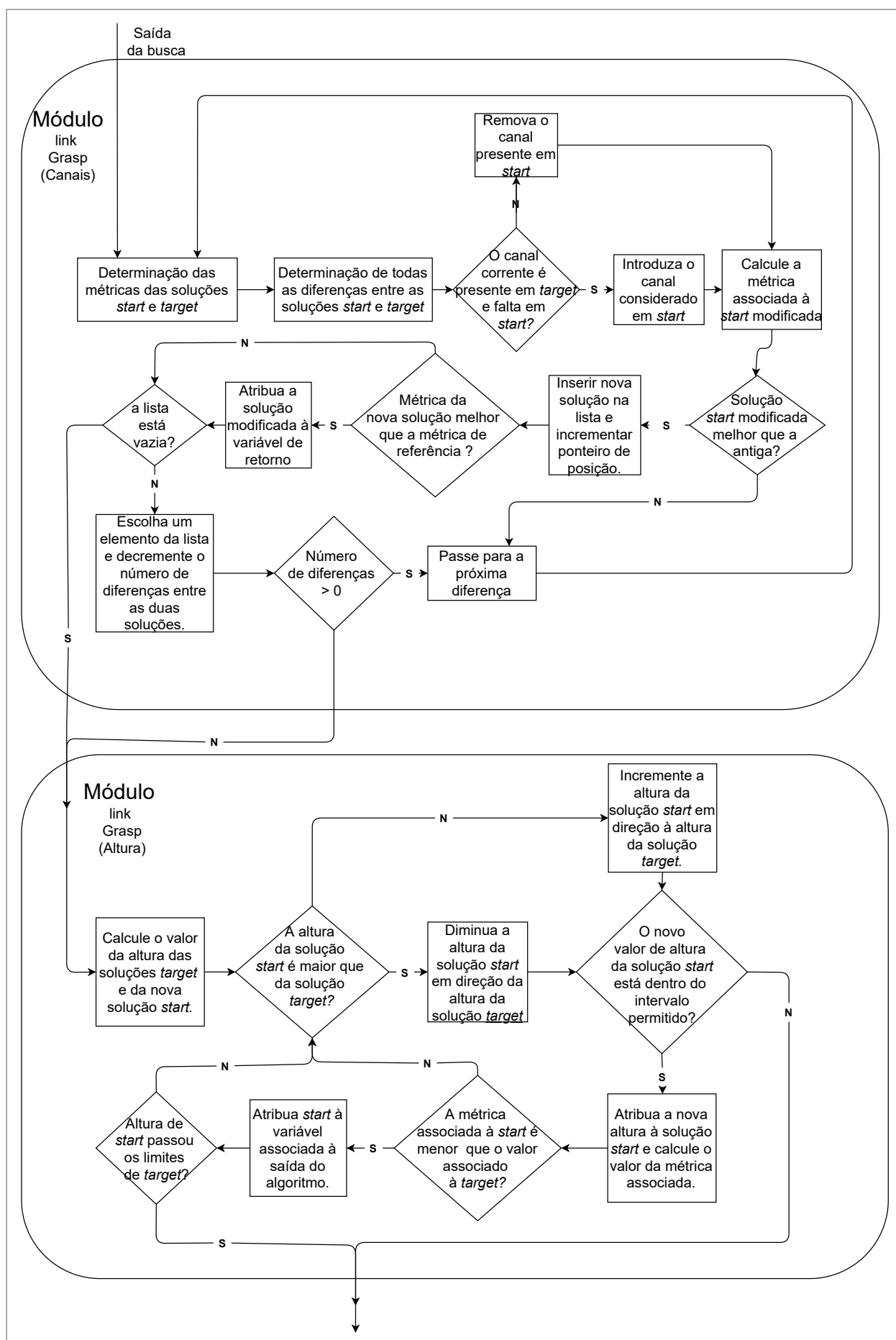
Na Figura 17, é demonstrada a sumarização do processo de linkagem entre duas soluções. Uma das soluções, nomeada por *start*, tem o valor de métrica de pior qualidade que a segunda solução, nomeada por *target*. As diferenças entre essas duas soluções são definidas a partir de duas variáveis: (1) a primeira delas são os canais inseridos na partição de maior resíduo das soluções. Os canais inseridos nas partições das duas soluções são comparados, um a um, podendo ocorrer o processo de subtração ou de adição de canais na solução *start*, de acordo com a comparação a ser feita; (2) a segunda variável leva em conta a definição de contagem espectral da biblioteca omitida na partições de maior resíduo das soluções *start* e *target*. Essa contagem espectral pode ser majorada, ou minorada, na solução *start*, de acordo com a comparação a ser feita na solução *target*, após um processo de discretização da diferença de contagem espectral entre essas duas soluções. A contagem espectral da biblioteca omitida, na partição de maior resíduo, da solução *start*, é guiada pelo valor definido na solução *target*.

O pseudocódigo referente ao processo de linkagem é apresentado no Apêndice C, Seção C.5. A linkagem de canais está apresentada na Seção C.5.1, e linkagem na contagem espectral de duas soluções encontra-se na Seção C.5.2.

A sumarização do processo de linkagem geral está descrita na Figura 17. O passo inicial é determinar as métricas e as diferenças entre as soluções *start* e *target*. São comparados os canais entre as soluções, verificando os canais que estão inseridos na solução *target* e ausentes na solução *start*, adicionando um a um em *start* e recalculando as métricas no intuito de verificar se a nova solução *start* é melhor que a solução *start* corrente. As verificações também ocorrem pela verificação dos canais presentes em *start* e ausentes em *target*. Esses canais são subtraídos da solução *start*, um por um, e a cada subtração, todas as métricas são recalculadas para verificar se a nova solução *start* é melhor que solução *start* corrente.

O segundo passo da linkagem verifica as definições de contagem espectral para a biblioteca omitida, na partição de maior resíduo, para as duas soluções. A diferença de contagem espectral entre as duas soluções é discretizada, e o valor da contagem da solução *start* é modificado, a partir da diferença mínima definida na discretização, em direção à contagem da solução *target*. A cada modificação, os parâmetros são recalculados, para verificar se a nova solução *start* é melhor que a solução corrente.

Figura 17 – Sumarização da linkagem entre duas soluções.



4.7.1 Linkagem Grasp - Canais

A linkagem remove os canais da solução *start* que estão ausentes na solução *target*. Essa subtração é feita a partir da premissa de que esses canais são os responsáveis por determinar a menor qualidade da solução inicial relativamente à solução alvo. Os canais presentes na solução *target* e ausentes na solução *start* também podem indicar o fato de que esses canais da solução *target* são os responsáveis pela menor qualidade da solução *start*. Nesse caso, os canais serão inseridos a partir da solução *target* em *start*. O algoritmo para linkagem dos canais pode ser conferido no Apêndice C Seção C.5.1.

Tem-se como parâmetros de entrada para o algoritmo de linkagem:

1. As soluções *start* e *target* e os seus respectivos *parâmetros de construção*;
2. Número mínimo de canais que os compartimentos das partições devem conter;
3. As partições de referência, contendo os picos dos elementos fornecidos, e a partição de maior resíduo para o elemento omitido do espectro total;
4. Número inteiro indicando a biblioteca a ser omitida;
5. Menor valor de resíduo até então encontrado.

Como saída para o Algoritmo 11, pode-se enumerar:

1. Os *parâmetros de construção* atualizados pela linkagem;
2. Novo valor de métrica após o processo de linkagem.

A Figura 18a mostra uma comparação entre os canais de duas soluções *start* e *target*, e como os procedimentos de linkagem são realizados após a comparação entre duas soluções.

Figura 18 – Procedimentos executados na linkagem de canais entre as soluções *start* e *target*.

Partições Arbitrárias

Solução
Inicial

30	31	32	33	34	35	36	37

Solução
Referência

33	34	35	36	37	38	39	40

- (a) Canais inseridos nas soluções *start* (Solução Inicial) e *target* (Solução Referência). Existem canais comuns às duas soluções, e canais que pertencem exclusivamente a cada uma delas. Canais presentes em *start* e ausentes em *target* contribuem para a pior qualidade de *start*. Canais presentes em *target* e ausentes em *start* contribuem para a melhor qualidade de *target*.

Partições Arbitrárias

Solução
Inicial

~~30~~	31	32	33	34	35	36	37

Solução
Referência

33	34	35	36	37	38	39	40

- (b) Exclui-se um canal na solução inicial, e verifica-se a melhora na qualidade da métrica associada a essa solução *start* modificada. Esse procedimento é feito com os canais 30, 31 e 32.

Partições Arbitrárias

Solução
Inicial

30	31	32	33	34	35	36	37	38

Solução
Referência

33	34	35	36	37	38	39	40

- (c) Posteriormente, um canal da solução de referência é adicionado ao conjunto de canais da solução inicial, verificando-se se ocorre a melhora na qualidade da solução *start*, realizando-se um novo cálculo de métrica.

Fonte: Próprio autor

4.7.2 Linkagem Grasp - Altura

A linkagem também é definida na aproximação da biblioteca omitida da solução *start* em direção à solução *target*, na partição de maior resíduo. O Algoritmo 12 descreve essa linkagem, sendo encontrado no Apêndice C, Seção C.5.2.

Como parâmetros de entrada para o Algoritmo 12, enumeram-se:

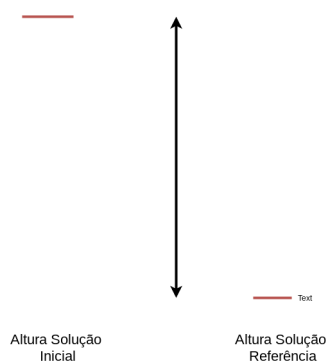
1. *Parâmetros de construção* resultantes do processo de linkagem na Seção 4.7.1, para as soluções *start* e *target*;
2. Número de divisões definidos a partir da diferença das contagens espectrais das soluções *start* e *target*, na partição de maior resíduo;
3. Número inteiro que indica a biblioteca omitida do espectro total;
4. Mínimo valor de resíduo até então encontrado ao longo de todo o Algoritmo GRASP.

Os parâmetros de saída podem ser enumerados por:

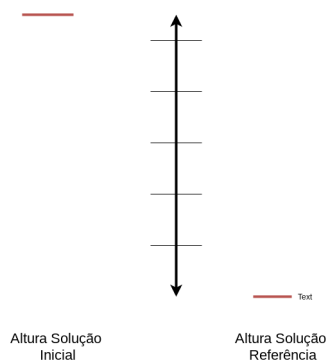
1. Atualização dos *parâmetros de construção* para uma nova solução melhor que *start*.
2. Novo valor de menor resíduo (*minRes*), caso seja encontrada uma solução melhor que a solução de referência.

A Figura 19a demonstra o processo de linkagem dos valores de contagem espectral para duas soluções *start* e *target* arbitrárias.

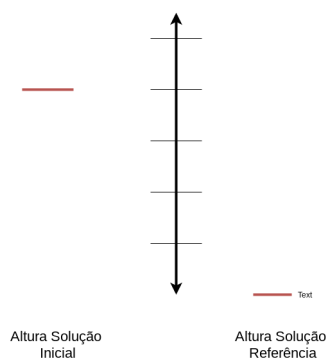
Figura 19 – Variação nos valores de altura para o composto extraído do espectro desconhecido.



- (a) Diferença de altura entre as soluções *start* (solução inicial) e a solução *target*.



- (b) Definida a diferença de altura na figura 19a, discretiza-se essa diferença para que possa ser varrida as diversas diferenças entre as duas soluções.



- (c) Aqui no caso, a altura da solução inicial será levada em direção a altura da solução *target*, que, no caso, será diminuída.

Fonte: Próprio autor

5 RESULTADOS

Alguns resultados que foram relevantes para o trabalho são descritos nas seguintes seções: (1) na Seção 5.1.1, são determinados os picos dos elementos que são fornecidos inicialmente ao módulo do LLS; (2) em 5.1.2 são determinadas as partições que apresentam o maior resíduo, para cada uma das instâncias; (3) as frações de contagem percentual no particionamento dos dados são descritas na Seção 5.1.3, após determinação dos compartimentos de cada partição, com as contagens espectrais reais das bibliotecas. Essa seção tem a finalidade de avaliar as saídas das frações em peso, sem a necessidade da estimativa de uma biblioteca, na tentativa de parametrizar os resultados seguintes; (4) as frações em peso, nos compartimentos das partições, mediante a aproximação de uma biblioteca pela geometrização, estão descritas na Seção 5.1.4; (5) A avaliação dos resultados para as estimativas do elemento omitido, e a respectiva comparação com o espectro total, encontra-se na Seção 5.1.5; (6) os valores de condicionamento, na minimização dos resíduos, são descritos na Seção 5.1.6, a fim de realizar a comparação entre essas duas variáveis na condução do GRASP; (7) os resultados para as subetapas do GRASP têm descrição nas Seções 5.1.7.1, 5.1.7.2, 5.1.7.3 e 5.1.7.4; (8) por fim, são avaliados os resultados nos ajustes do CGRASP à estimativa da biblioteca omitida nas Seções 5.1.8.1 (execução única do GRASP), 5.1.8.2 (execuções acumuladas), 5.1.8.3 (etapa construtiva) e 5.1.8.4 (etapa de busca).

5.1 RESULTADOS PARA OS ELEMENTOS

Antes da apresentação dos resultados, realiza-se uma apresentação dos dados da Base 1. A Tabela 4 traz uma seção dos cinco primeiros canais dos dados de entrada para o algoritmo implementado. Cada coluna (Amerício (Am), Cobalto (Co), Césio (Cs) e Sódio (Na)) representa as intensidades da contagem espectral para um elemento (compostos, no caso da Base de Dados 2), como função dos canais de energia (variáveis discretas), descritos na coluna um.

Tabela 4 – Descrição dos cinco primeiros canais das bibliotecas do Amerício, Cobalto, Césio e Sódio GRASP. Os dados de entrada para a primeira base possuem uma quantidade de 20148 canais.

Canais	Amerício	Cobalto	Césio	Sódio
1	9e-07	0.000749	0.000158	5.24e-05
2	7.28e-05	0.0525	0.011	0.00365
3	1.45e-05	0.00767	0.00174	0.000613
4	1.9e-05	0.000251	0.000187	0.000102
5	5.58e-05	0.000267	0.000196	0.000107

Fonte: Próprio autor

5.1.1 Determinação dos picos

A presente seção trará os resultados para a determinação dos picos dos elementos contidos na Base de Dados 1. São verificadas as saídas do algoritmo implementado para a detecção de picos.

Os resultados apresentados para os picos das bibliotecas da Base 1 são apresentados a seguir:

Tabela 5 – Regiões para os picos dos elementos fornecidos GRASP

Instância	Elemento Omitido	Elemento Fornecido	Canal do Pico
1	Amerício (<i>Am</i>)	<i>Co</i>	603
		<i>Cs</i>	340
		<i>Na</i>	263
2	Cobalto (<i>Co</i>)	<i>Am</i>	32
		<i>Cs</i>	340
		<i>Na</i>	263
3	Césio (<i>Cs</i>)	<i>Am</i>	32
		<i>Co</i>	603
		<i>Na</i>	263
4	Sódio (<i>Na</i>)	<i>Am</i>	32
		<i>Co</i>	603
		<i>Cs</i>	340

Fonte: Próprio autor

Os valores da Tabela 5 trazem os canais onde se encontram os picos das bibliotecas fornecidas inicialmente. Para cada biblioteca omitida, são demonstrados os canais onde são encontrados os picos de cada biblioteca fornecida, segundo descrito na Seção 5.1.1.

O pico para o *Co* foi encontrado no canal 603, para as instâncias 1, 3 e 4, de acordo com o descrito na Figura 3. O pico para o *Cs* nas instâncias 1, 2 e 4 está no canal 340, mais uma vez, em conformidade com a descrição de picos na Figura 3. A mesma discussão pode ser estendida para os resultados do *Am* e *Na*. Os resultados indicam que o método foi bem sucedido na determinação dos picos e consequente escolha das partições para os elementos fornecidos inicialmente ao espectro total. Os resultados da determinação dos picos podem ser comparados com a Figura 3. Os picos dos três elementos, na comparação com a figura citada, residem onde os algoritmos determinam os picos.

5.1.2 Determinação dos Resíduos das Partições

A presente seção revela os valores de resíduos, juntamente com as componentes dos vetores de compostos, em cada uma das partições. As Figuras 20a e 20b apresentam os resultados para os valores das componentes dos vetores de compostos e resíduos como função da iteração nas instâncias 1 e 2. Os resultados da Base 2, instâncias 5 e 6, estão na Seção 5.2.2. As demais instâncias estão descritas no Apêndice D: Seção D.1 para as instâncias 3 e 4.

A Tabela 6 apresenta o número das partições de maior resíduo para as instâncias 1, 2, 3 e 4. A partição 01 é a de maior resíduo para a omissão do *Am*. Comparando esse resultado com a Tabela 5, percebe-se que a primeira partição contém efetivamente o pico do *Am*. O mesmo pode ser generalizado para os demais compostos (a partição 6 também indica o maior resíduo para o *Cs*, cujo pico está no canal 350; O *Na* tem o maior resíduo na partição 5, com pico localizado no entorno do canal 260). A exceção reside na determinação da partição de maior resíduo na omissão do *Co*. Esse caso será analisado quando da apresentação dos gráficos dos vetores de compostos como função da iteração, na discussão que se segue.

Tabela 6 – Partições contendo os picos dos constituintes. Algoritmo GRASP

–	Composto 1 <i>Am</i>	Composto2 <i>Co</i>	Composto 3 <i>Cs</i>	Composto 4 <i>Na</i>
Partição	1	6	6	5

Fonte: Próprio autor

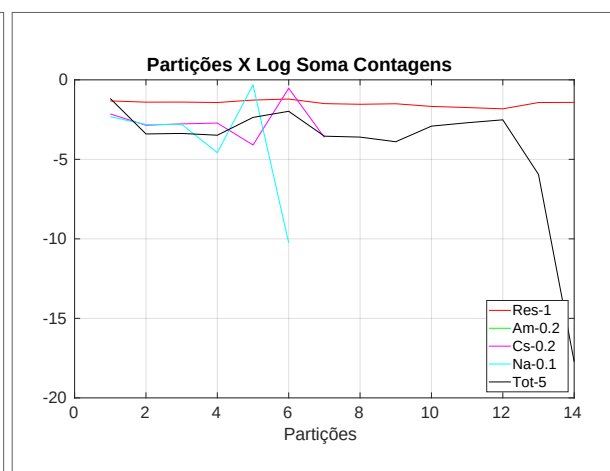
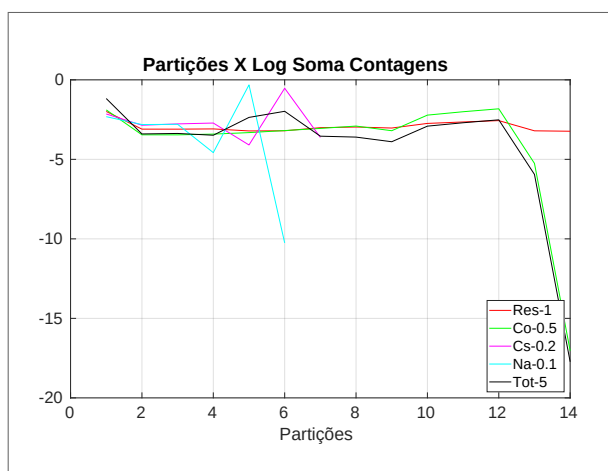
A Figura 20a, para a instância 1, e Figura 20b, para a instância 2, apresentam os seguintes valores: (1) valores das componentes dos vetores de compostos das bibliotecas fornecidas (gráficos verde, rosa e azul); (2) valores do resíduo nas partições, a partir das bibliotecas fornecidas (gráfico em vermelho); (3) as componentes do vetor espectro total (gráfico em preto) como função da partição em estudo. No caso da omissão do *Am*, o maior resíduo

está localizado na partição 1, local onde ocorre o pico desse elemento. Para a primeira base de dados, a exceção encontra-se na determinação do pico do *Co* (Figura 20b). A ocorrência do pico do *Co* é localizada em um região em que as bibliotecas fornecidas vão a zero, não sendo, portanto, preciso o método para determinação desse pico. Mas o método adota uma localização para descrição do *Co* que será bem sucedida na sua estimação e determinação da fração em peso, como será visto posteriormente na Seção 5.1.4.

Figura 20 – Gráfico apresentando as componentes dos vetores de compostos, resíduos e componentes do vetor espectral total, como função da iteração. São apresentados os resultados para a instância 1 (Figura 20a) e instância 2 (Figura 20b). Os componentes do vetor espectro total são apresentados (variável *tot* – 5), além dos valores de resíduos nas partições (variável *res* – 1).

(a) Componentes dos vetores de compostos *Co*, *Cs* e *Na* como função de partição. Os componentes do vetor espectro total são apresentados (variável *tot* – 5), além dos valores de resíduos nas partições (variável *res* – 1)

(b) Componentes dos vetores de compostos *Am*, *Cs* e *Na* como função de partição. Os componentes do vetor espectro total são apresentados (variável *tot* – 5), além dos valores de resíduos nas partições (variável *res* – 1).



Fonte: Próprio autor

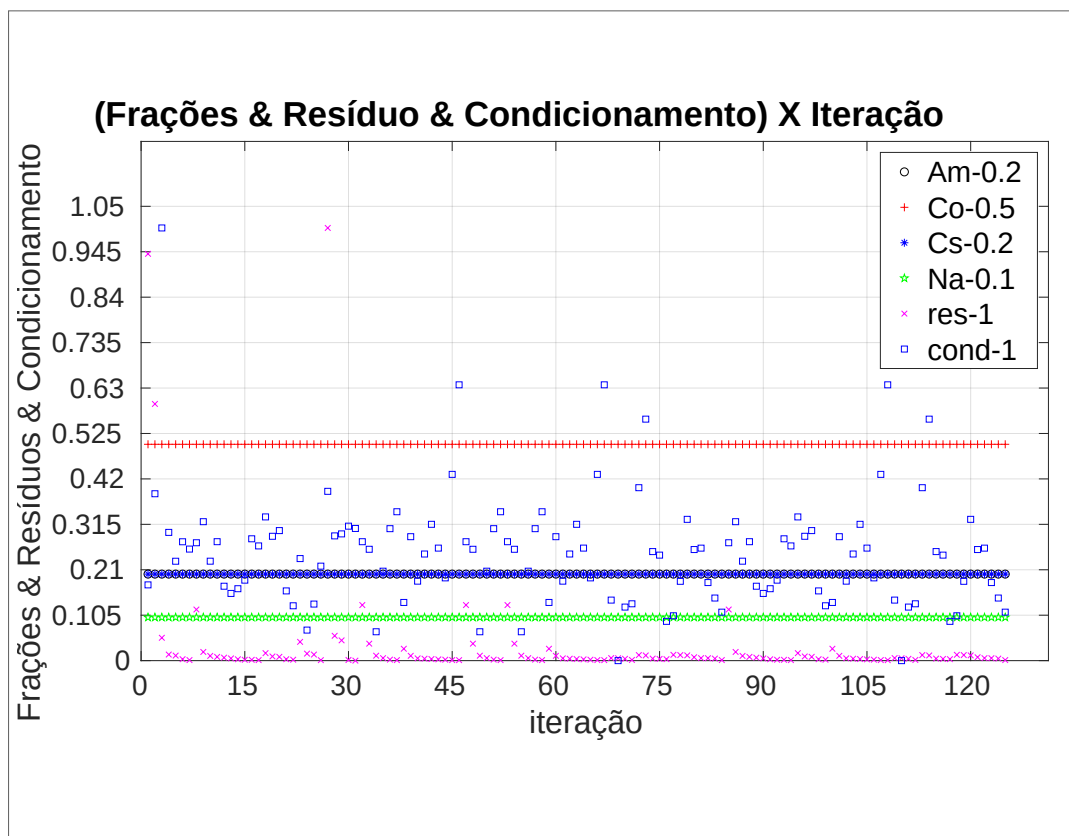
5.1.3 Determinação de Frações Percentuais - Particionamento

A atual seção define os valores de frações em peso para trinta execuções acumuladas do algoritmo para os elementos da Base de Dados 1, particionando os dados, sem realizar a estimativa de uma das bibliotecas. Os resultados aqui apresentados são importantes, no intuito de analisar as frações em peso utilizando uma seção dos dados. Essa abordagem evita a questão da sobre-determinação do sistema de equações. Os resultados aqui descritos também servirão de parâmetros para as aproximações das frações de contagem percentual dos elementos quando um deles estiver sendo aproximado, na Seção 5.1.4.

O resultado para a instância 1 será apresentado aqui. As demais instâncias da Base 1, por questões de similaridades, serão apresentadas no Apêndice E, Seção E.1.

A Figura 21 demonstra as frações em peso, valores de resíduos e do condicionamento como função da iteração. Aqui, foram destacadas algumas iterações, já que são muitas saídas como essa em trinta execuções do algoritmo. Importante destacar como as frações em peso, determinadas pela aproximação dos mínimos quadrados, descrevem com exatidão os valores nominais das frações de contagem percentual, como pode ser observado na legenda do Gráfico 21. Os valores de condicionamento oscilam ao longo das iterações, enquanto que os valores de resíduos são significativamente reduzidos, a partir de uma região do gráfico que pode ser identificada como uma execução do algoritmo.

Figura 21 – Frações de Contagem percentuais no particionamento dos compostos, instância 1. As taxas de contagem espectral para o *Am* são as taxas reais. Valores em preto (○), vermelho(+), azul (*) e verde (☆) indicam, respectivamente, as quantidades de *Am*, *Co*, *Cs* e *Na*. Os valores nominais de frações dos elementos estão descritos nas legendas do gráfico. Os valores dos resíduos são descritos nos pontos (x), e os pontos (□) indicam o condicionamento normalizado $[0, 1]$ das matrizes de covariância do particionamento

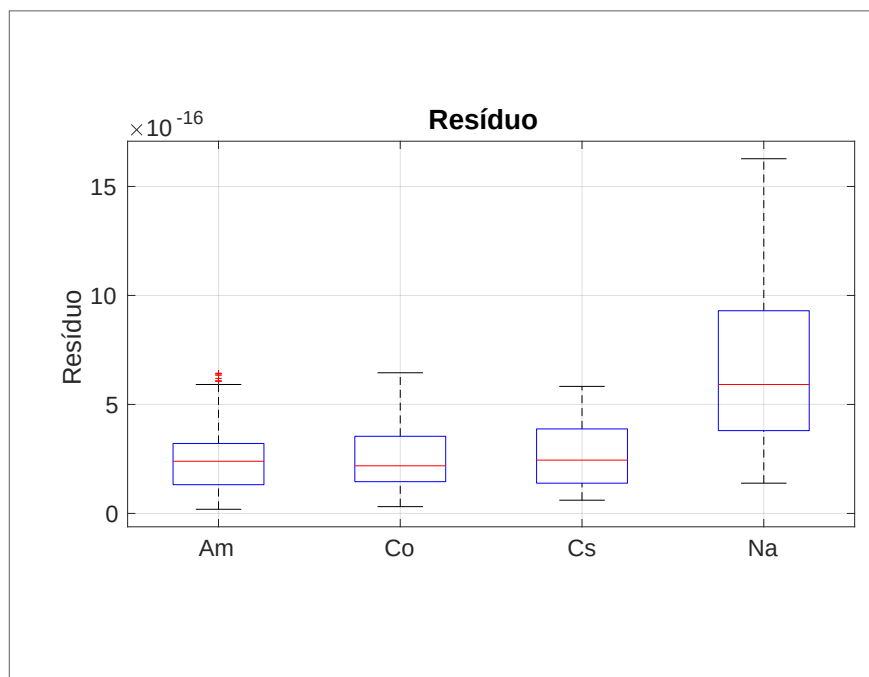


Fonte: Próprio autor

Na figura 22, são apresentadas as distribuições para os valores do resíduo para as instâncias 1, 2, 3 e 4. Existem evidências de que esses valores são aproximadamente nulos, dentre todos os casos estudados. Esses resultados são indicativos iniciais que a minimização do resíduo contribui na convergência das frações de contagem percentual aos valores nominais de frações de cada biblioteca. Essa evidência irá guiar as aproximações das frações quando uma das bibliotecas for aproximada.

A distribuição para os valores de condicionamento para as matrizes de covariância são demonstrados na Figura 23. Para a instância 1, esses valores são aproximadamente simétricos e compreendidos entre 3 e 4. Os valores são distribuídos negativamente simétricos no caso do *Cs*, e positivamente simétricos no caso do *Co* e *Na*. Esses valores de condicionamento são comparáveis aos valores atingidos nos trabalhos que utilizam matrizes de dimensão 2×2 (MERIÇ, 2012) como métrica na qualidade da aproximação para as frações em peso.

Figura 22 – Plots demonstrando o intervalo de valores para o resíduo durante 30 execuções do algoritmo proposto no particionamento dos dados. Na ausência do *Am*, 50% dos valores de resíduo encontram-se compreendidos entre 2×10^{-15} e 4×10^{-16} . A mesma interpretação pode ser estendida para os demais elementos por meio dos seus respectivos boxes. Os valores de resíduos são positivamente assimétricos para o *Co*, *Cs* e *Na* e aproximadamente simétrico no caso do *Am*.



Fonte: Próprio autor

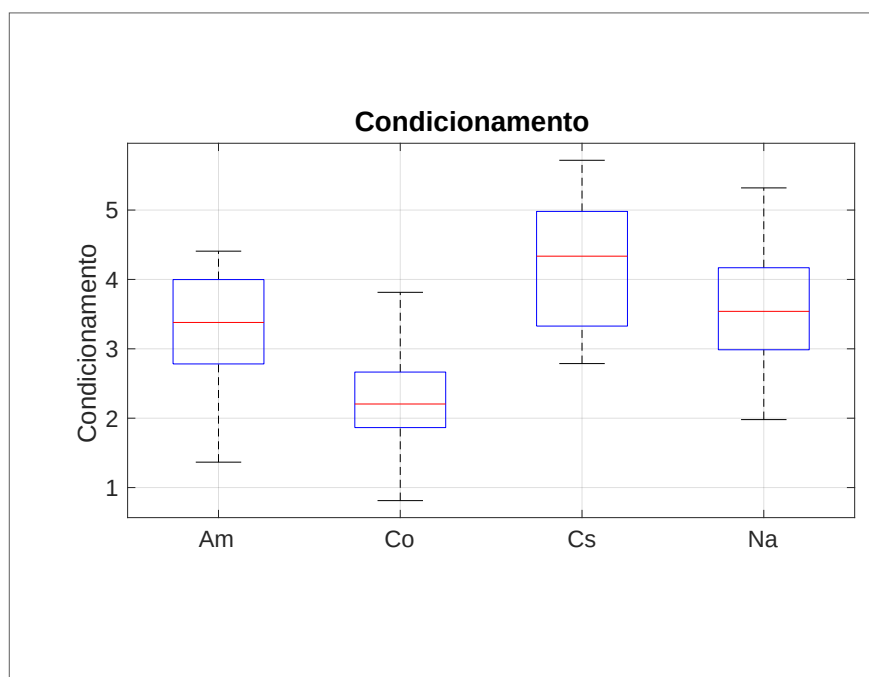
5.1.4 Determinação da Fração Percentual dos Compostos por Aproximações

Na presente seção, serão apresentados os resultados para as frações em peso dos compostos na omissão de uma biblioteca do espectro total. Os resultados farão referência à instância um. As demais instâncias da Base de dados um serão apresentadas no Apêndice F, Seção F.1, já que os resultados apresentam similaridades.

Os valores médios, de 30 execuções, das frações de contagens percentuais, de resíduo e de condicionamento, ao omitir uma biblioteca e aproximá-la, estão demonstrados na Tabela 7. A primeira linha contém as médias das frações de contagens percentuais (*Am*, *Co*, *Cs* e *Na*) de todos os elementos para a instância 1. As colunas *PreAm*, *PreCo*, *PreCs* e *PreNa* contêm as precisões para todas as aproximações. As duas últimas colunas contêm os valores médios de resíduos e de condicionamento das matrizes de covariância. As demais linhas são referentes às instâncias 2, 3 e 4, seguindo a mesma interpretação fornecida à linha 1.

Na Tabela 8, são determinados os valores de fração de contagem percentual na minimização do resíduo. Os valores de resíduos estão descrito na última coluna da Tabela 8. A primeira

Figura 23 – Plots demonstrando os valores de condicionamento. Na ausência do *Am*, os valores de resíduo estão compreendidos entre 3 e 4, com os dados distribuídos aproximadamente simetricamente. O *Co* apresenta esses valores compreendidos no intervalo 2 e 2.5; o condicionamento para a instância do *Cs* está no intervalo de 3.5 e 5 e, por fim, o *Na* tem seu intervalo definido entre 3 e 4.



Fonte: Próprio autor

Tabela 7 – Valores médios de frações de contagens percentuais dos compostos inseridos no espectro total, após 30 execuções do algoritmo. A linha 1 contém os valores de frações percentuais dos compostos para a instância 1. As instâncias 2, 3 e 4 podem ser observadas a partir das linhas seguintes. As duas últimas colunas contêm os valores médios de resíduo e condicionamento para trinta execuções do algoritmo.

	Am	PreAm	Co	PreCo	Cs	PreCs	Na	PreNa	Resíduo	Condic.
1	0.14557	0.27216	0.56182	0.12364	0.1764	0.11798	0.11562	0.15617	0.022774	3.8183
2	0.35066	0.75328	0.23122	0.53756	0.22463	0.12315	0.19194	0.91938	0.02494	7.6554
3	0.19661	0.016929	0.56321	0.12643	0.12175	0.39124	0.1181	0.18101	0.029824	5.7452
4	0.20567	0.028345	0.51298	0.025969	0.18701	0.064925	0.094139	0.058614	0.023908	6.3937

Fonte: Próprio autor

linha descreve os resultados para a instância 1, e as demais linhas apresentam os resultados das instâncias 2, 3 e 4.

Tabela 8 – Valores aproximados das frações em peso para os menores valores do resíduo. A linha 1 da tabela determina os valores de frações dos elementos inseridos na amostra, estimando a biblioteca do *Am* e o aproximando-o por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas: linha 2 é referente à instância 2, omissão do *Co*; linha 3 é referente à instância 3, omissão do *Cs* e a última linha refere-se à instância 4, omissão do *Na*.

	Am	PreAm	Co	PreCo	Cs	PreCs	Na	PreNa	Resíduo
1	0.19976	0.0011945	0.50027	0.0005372	0.20002	0.0001025	0.099925	0.0007476	6.4825e-05
2	0.20085	0.004256	0.49719	0.0056122	0.20064	0.003179	0.10132	0.01319	0.00041551
3	0.20004	0.0001915	0.50004	7.8e-05	0.20005	0.000232	0.099701	0.0029895	4.1566e-05
4	0.19146	0.042682	0.49955	0.0008966	0.19968	0.0016145	0.10931	0.093075	0.0003487

Fonte: Próprio autor

Na Tabela 9, são determinados os valores de fração de contagem percentual para o menor valor de condicionamento em todas as ocorrências de minimização do resíduo para uma instância: linha 1, instância 1 e, assim sucessivamente. O menor valor de condicionamento está descrito na última coluna da Tabela 9 para cada uma das instâncias.

Tabela 9 – Valores de frações de contagem percentual aproximados considerando o menor condicionamento, nas ocorrências de minimização do resíduo, como métrica. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o *Am* da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas.

	Am	PreAm	Co	PreCo	Cs	PreCs	Na	PreNa	Condic.
1	0.18757	0.062148	0.48407	0.031851	0.2033	0.016518	0.12337	0.23367	6.89
2	0.19999	4.35e-05	0.46719	0.065611	0.20847	0.04234	0.11967	0.19672	1.83564
3	0.17213	0.13937	0.60504	0.21009	0.12652	0.36742	0.092828	0.071723	4.64264
4	0.20931	0.046564	0.52946	0.058921	0.19105	0.044747	0.069961	0.30039	1.92046

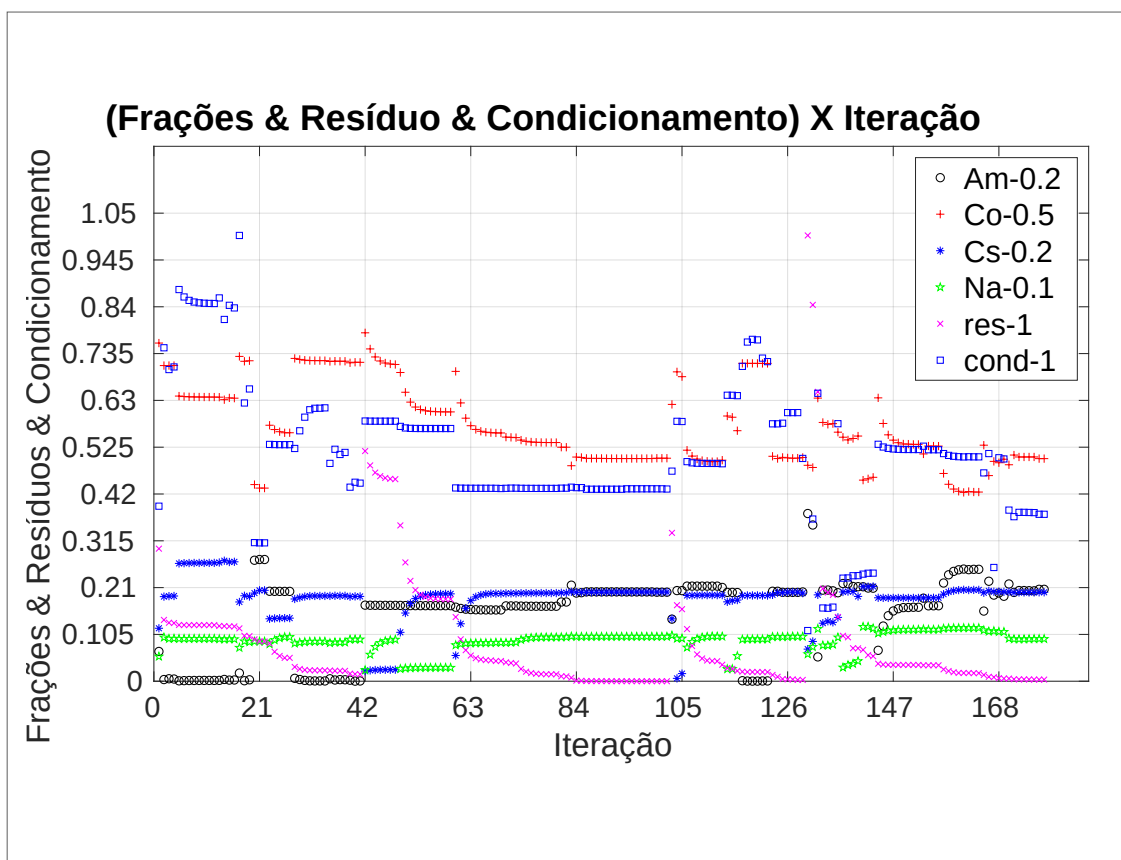
Fonte: Próprio autor

Das possíveis métricas apresentadas nas Tabelas 7 (frações em peso determinadas por valores médios de trinta execuções), 8 (frações em peso determinadas na minimização do resíduo) e 9 (frações em peso determinadas na minimização do condicionamento em todas as ocorrências de minimização do resíduo), a que apresenta evidências de ser eficiente na precisão das frações em peso é a minimização do resíduo. Os valores médios de trinta execuções não se mostraram tão eficientes quanto à minimização do resíduo. Apesar da minimização do condicionamento ser um preditor da diminuição das instabilidades numéricas inerentes a

um sistema de equações linear, a minimização dos seus valores não obteve êxito tão bom na determinação das frações em peso quanto a minimização dos resíduos. Mas é importante destacar que esses valores de condicionamento calculados, para matrizes de dimensão quatro, não são encontrados nos últimos trabalhos publicados do estado da arte.

Na Figura 24, são apresentados os resultados das frações de contagem percentual, resíduo e condicionamento, como função da iteração, em trinta execuções do algoritmo para a instância 1. A Figura 24 demonstra que os valores de fração de contagem percentual convergem para os valores nominais à medida que os valores de resíduo são reduzidos. As regiões em que os resíduos são monotonicamente reduzidos evidenciam uma execução arbitrária do nosso algoritmo, dentre as execuções acumuladas apresentadas na Figura 24. Esse fato pode ser observado nas iterações 100 e 120, por exemplo, que é possível identificar como uma execução particular do nosso algoritmo.

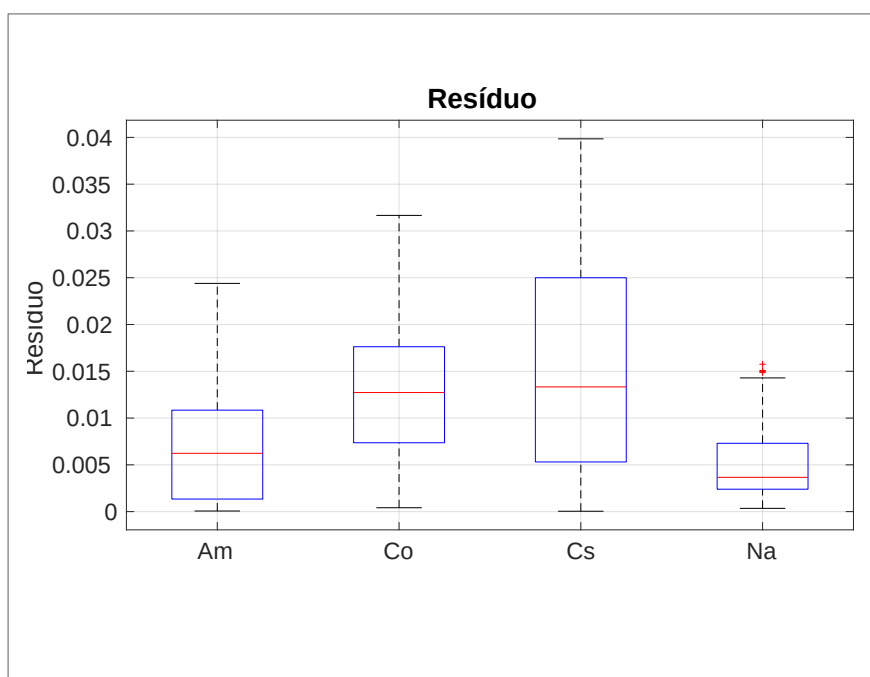
Figura 24 – Frações de contagens percentuais, instância 1, omissão e aproximação do *Am*. Pontos (○), vermelho (+), azul (*) e verde (☆) indicam, respectivamente, as frações em peso para o *Am*, *Co*, *Cs* e *Na*. Os valores dos resíduos são descritos nos pontos ×, e os pontos □ indicam o condicionamento normalizado ([0, 1]) das matrizes de covariância no particionamento. Percebe-se que à medida que o resíduo é reduzido, as frações de contagem percentuais convergem aos valores nominais.



Fonte: Próprio autor

As distribuições dos valores dos resíduos na aproximação dos elementos estão descritas na Figura 25. O primeiro box descreve as distribuições dos resíduos, em trinta execuções do algoritmo, ao realizar a determinação das frações de contagem percentual na ausência do *Am*. Os demais boxes seguem a mesma descrição realizada para o primeiro box para as instâncias do *Co*, *Cs* e *Na*. Ao realizar a comparação das distribuições dos resíduos, mediante a estimativa de uma das bibliotecas, com as distribuições de resíduos presente na Figura 22, percebe-se que as técnicas de estimação de bibliotecas geram a introdução de incertezas, mediante o aumento desses valores.

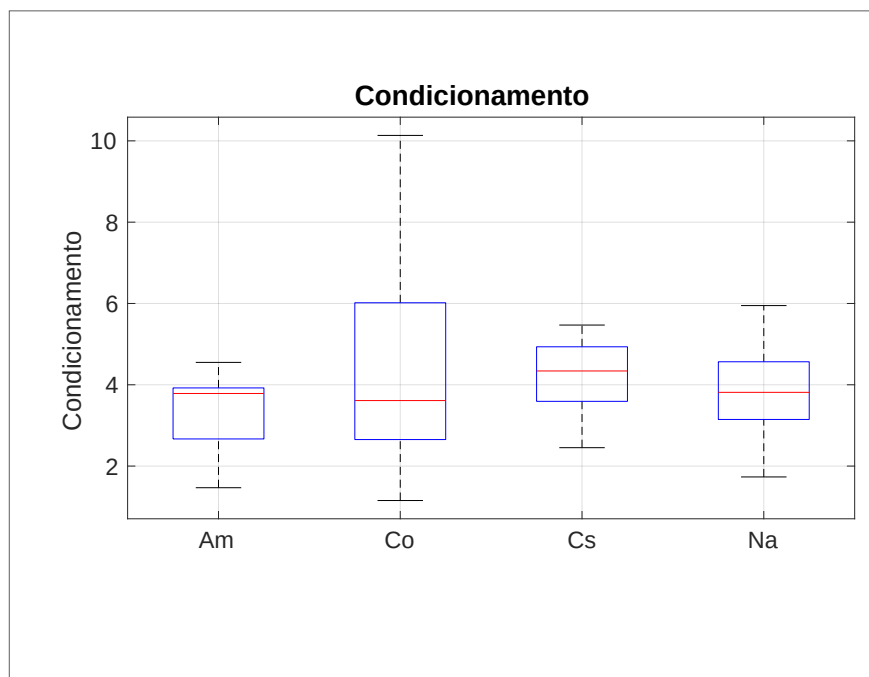
Figura 25 – Plot para os intervalos de resíduo após 30. Para a aproximação do *Am*, a mediana dos valores é aproximadamente 5×10^{-3} , com 50% dos valores encontrando-se compreendidos aproximadamente entre 1×10^{-3} e 10^{-2} . Na ausência do *Co*, a mediana é de aproximadamente 1.2×10^{-2} e 50% dos valores compreendidos entre 7×10^{-3} e 1.7×10^{-2} . O intervalo de valores para ausência do *Cs* são da ordem de 5×10^{-3} e 2.25×10^{-2} . Os valores de resíduo para a ausência do *Na* é de 2.5×10^{-3} e 7.5×10^{-3} . A mediana para a ausência do *Am*, *C* e *Na* são positivamente assimétricas; o plot para a ausência do *Co* mostra os dados aproximadamente simétricos.



Fonte: Próprio autor

As distribuições dos valores de condicionamento na aproximação dos elementos estão descritas na Figura 26. Ao realizar a comparação com os valores de condicionamento presentes na Figura 23, verifica-se que as técnicas de estimação de uma das bibliotecas agregaram um aumento médio nos valores de condicionamento das matrizes de covariância.

Figura 26 – Plots para o intervalo de valores para o condicionamento após 30 execuções do Algoritmo GRASP. Na ausência do *Am*, 50% desses valores encontram-se compreendidos entre 2.8 e 3.8. Para os demais compostos, respectivamente, esses valores encontram-se entre: *Co*, 3 e 6; *Cs*, 3.8 e 5; *Na*, 3 e 4.6. A mediana para a ausência do *Am* é negativamente assimétrica; o plot para a ausência do *Co* e *Na* mostra os dados aproximadamente simétricos e positivamente assimétricos para a ausência do *Co*.

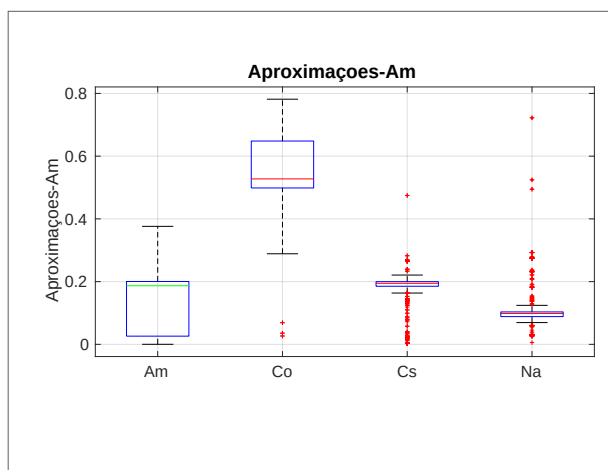


Fonte: Próprio autor

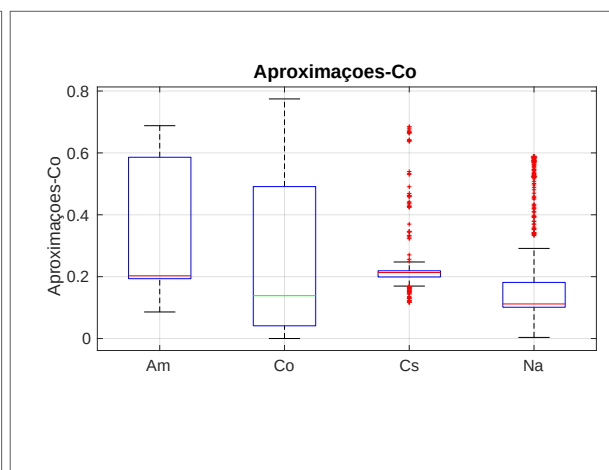
A Figura 27 apresenta as distribuições das frações de contagem percentual em 30 execuções do algoritmo desenvolvido no presente trabalho. A Figura 27a demonstra as frações de peso para a instância 1, com a mediana das distribuições em verde para indicar que esse elemento está sendo estimado para essa instância. As demais instâncias são descritas nas Figuras 27b (instância 2), 27c (instância 3) e 27d (instância 4). A amplitude para as frações em peso é menor para as instâncias 1, 3 e 4 do que para a instância 2, de acordo com o descrito pela Figura 27. A instância 2 apresentou maiores dificuldades na determinação da região de maior resíduo, muito provavelmente devido ao fato de a biblioteca do *Co* ocorrer isoladamente no espectro, gerando dificuldades na determinação do seu pico. Todas as demais bibliotecas vão a zero antes da ocorrência do pico do *Co*. Apesar disso, o nosso algoritmo foi capaz de determinar uma partição representativa para a biblioteca omitida do *Co*, e foi possível determinar, com boa precisão, as frações em peso de todas as bibliotecas inseridas no espectro total.

Figura 27 – Plots contendo os intervalos de frações de contagens percentuais para cada uma das instâncias. A figura 27a possui o plot com a mediana em verde para indicar que o *Am* está sendo aproximado. A mesma interpretação pode ser estendida para as Figuras 27b, 27c e 27d, na aproximação dos seus respectivos compostos (*Co*, *Cs* e *Na*). As frações de contagens percentuais na Figura 27a refletem aproximadamente os valores nominais. Na figura 27b existe uma grande variabilidade na aproximação do compostos aproximado, o *Co*, mas ainda englobando o valor nominal desse composto no intervalo citado. As interpretação da figura 27a pode ser estendida para as figuras 27c e 27d.

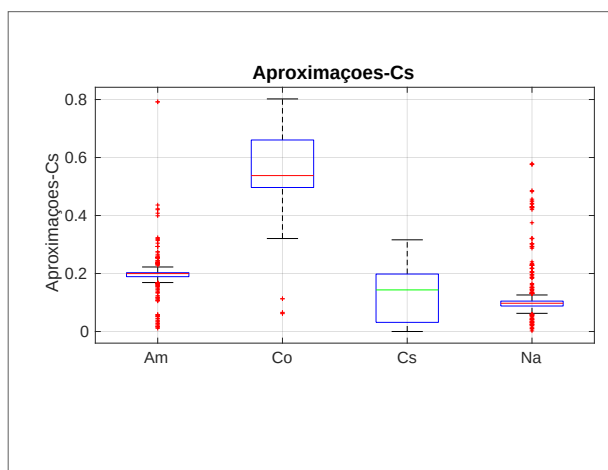
(a) Distribuição das frações de contagem percentuais após supressão do *Am*, instância 1, do espectro total.



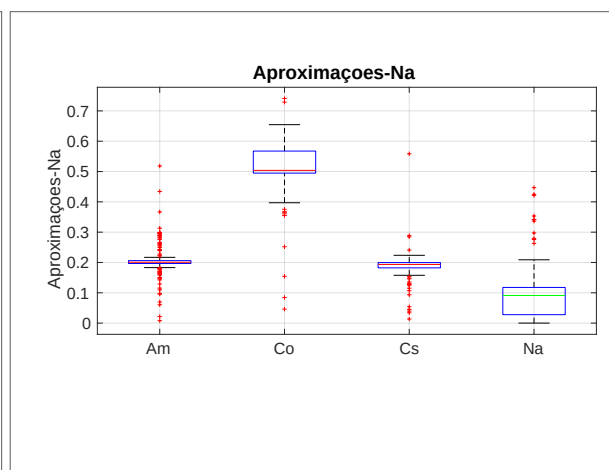
(b) Distribuição das frações de contagem percentuais após supressão do *Co*, instância 2, do espectro total.



(c) Distribuição das frações de contagem percentuais após supressão do *Cs*, instância 3, do espectro total.



(d) Distribuição das frações de contagem percentuais após supressão do *Na*, instância 4, do espectro total.



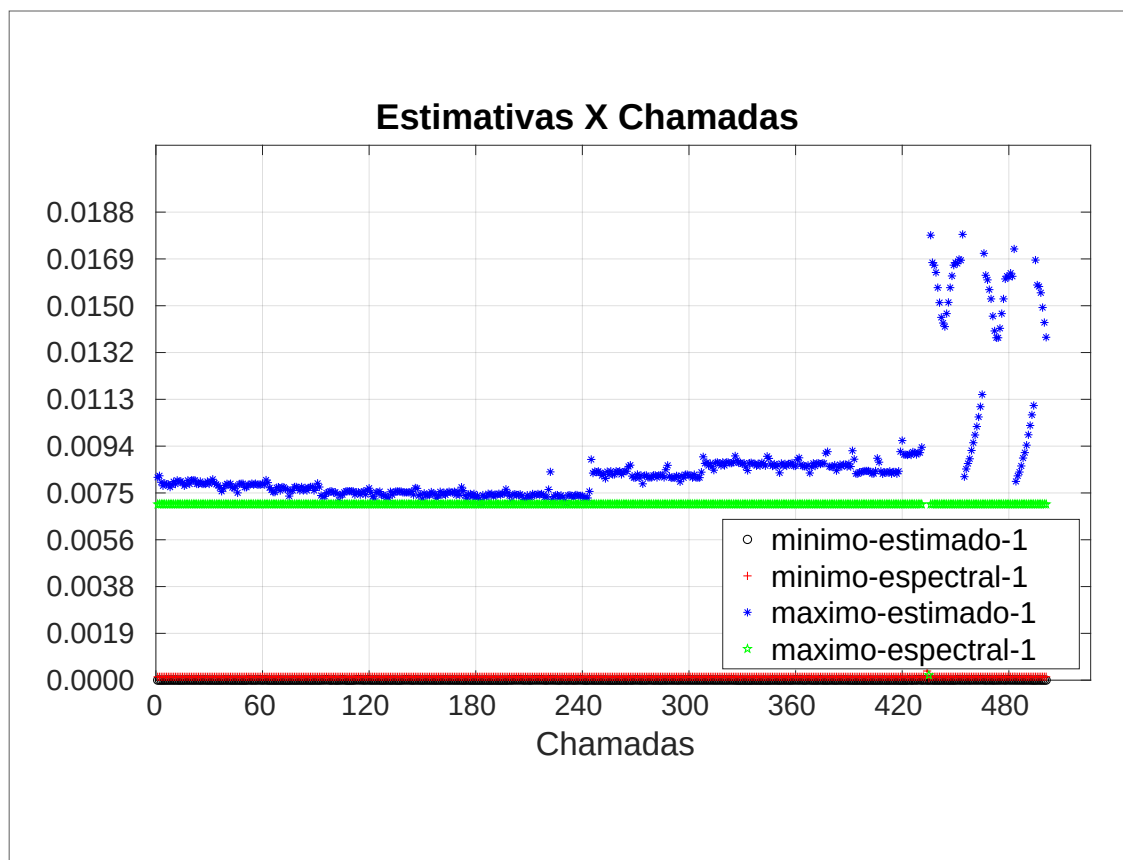
Fonte: Próprio autor

5.1.5 Aproximação dos Compostos - Estimativa das alturas nas partições

A presente seção analisará os resultados para as estimativas da biblioteca omitida, comparando-as aos valores de contagem espectral total. Como o espectro total é uma combinação linear das bibliotecas, é esperado que as estimativas mínimas e máximas sejam, respectivamente, menores e maiores que as contagens espectrais totais reais, aproximadamente, dentro dos compartimentos definidos. Esse fato irá, então, guiar a qualidade das aproximações realizadas na Seção 4.3 para a biblioteca omitida. Os resultados para a instância 1 são apresentados na Seção atual, enquanto que os resultados para as demais instâncias da Base 1 são apresentados no Apêndice G, Seção G.1.1, por questões de similaridades aos resultados apresentados na presente seção.

A Figura 28 apresenta os valores de estimativa de contagem espectral normalizados. Nessa Figura, é possível constatar que os valores estimados englobam os valores de contagens espectrais totais, como esperado. Então, há evidências de que os valores estimados podem ser ajustados no intuito de estimar devidamente as contagens espectrais da biblioteca omitida por meio dos procedimentos de geometrização.

Figura 28 – Estimativas da biblioteca omitida, da instância 1. Os pontos $\circ$ são valores de menor contagem estimada pelo algoritmo da Seção 4.3, enquanto que representados por $+$ indicam a menor contagem total no intervalo de canais que definem o compartimento da partição. Pontos $*$ são os valores de maior contagem estimada para a biblioteca omitida e os pontos $\star$ indicam o maior valor total para essa mesma biblioteca. Valores de estimativas são adequados, já que delimitam os valores de contagem espectral total dentro dos compartimentos.

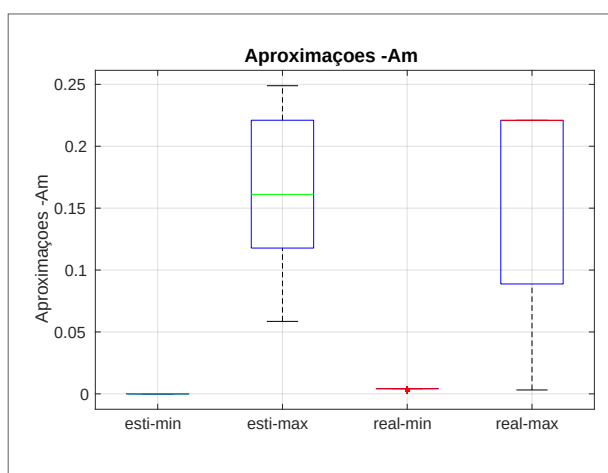


Fonte: Próprio autor

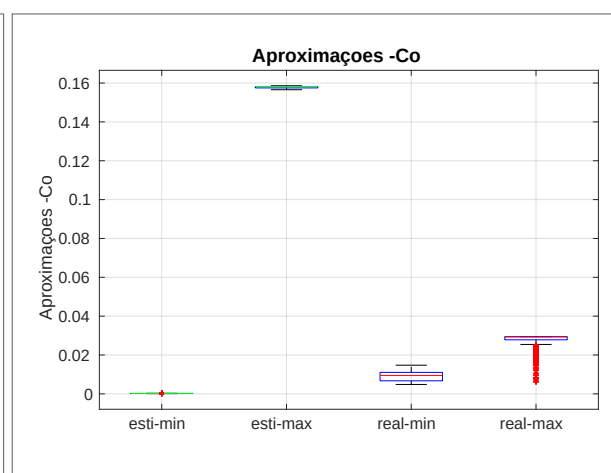
As distribuições apresentadas nas Figuras 29a, 29b, 29c e 29d ratificam que distribuição dos valores de contagem espectral estimada são limites para os valores de contagem espectral total nos compartimentos das partições. A exceção a essa regra é para a instância 1, em que os valores estimados não necessariamente delimitam as contagens espectrais totais. Esse fato pode ser devido às contagens reais para esse elemento irem a zero nos primeiros canais de energia, impactando a comparação com a contagem total em regiões em que não existem valores diferentes de zero para a instância 1. Esses resultados podem ser enumerados como uma das maiores contribuições do presente trabalho, não encontrando abordagem similar no estado da arte para emparelhamento do sistema do LLS no ajuste dos espectro fornecidos pelo MC.

Figura 29 – Valores de estimativa de altura para as bibliotecas omitidas *Am*, (Figura 29a, instância 1), *Co*, (Figura 29b, instância 2), *Cs* (Figura 29c, instância 3) e *Na* (Figura 29d, instância 4) nas suas partições de maior resíduo.

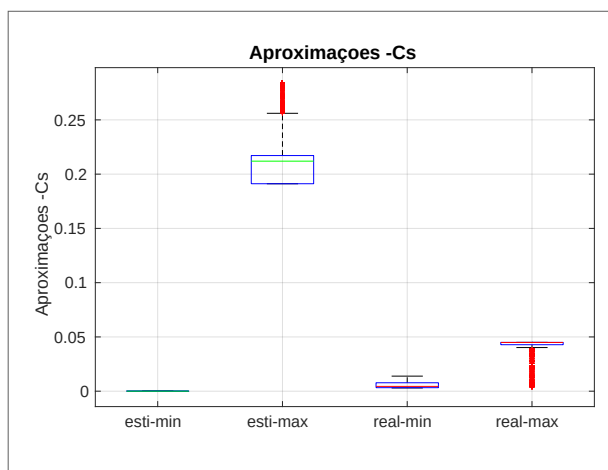
(a) Comparativo entre as estimativas do elemento omitido e as contagens espectrais totais, para a ausência do *Am*, instância 1.



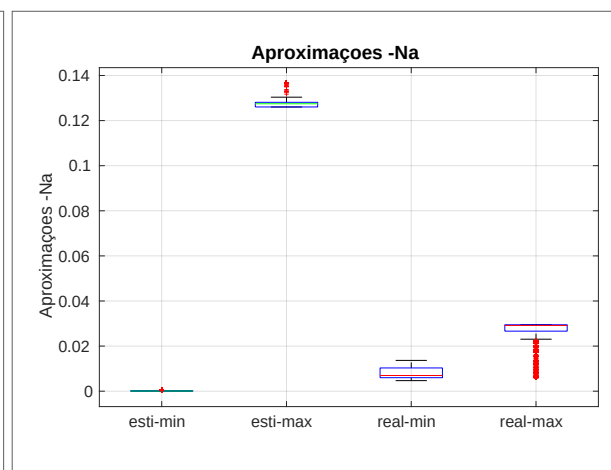
(b) Comparativo entre as estimativas do elemento omitido e as contagens espectrais totais, para a ausência do *Co*, instância 2.



(c) Comparativo entre as estimativas do elemento omitido e as contagens espectrais totais, para a ausência do *Cs*, instância 3.



(d) Comparativo entre as estimativas do elemento omitido e as contagens espectrais totais, para a ausência do *Na*, instância 4.



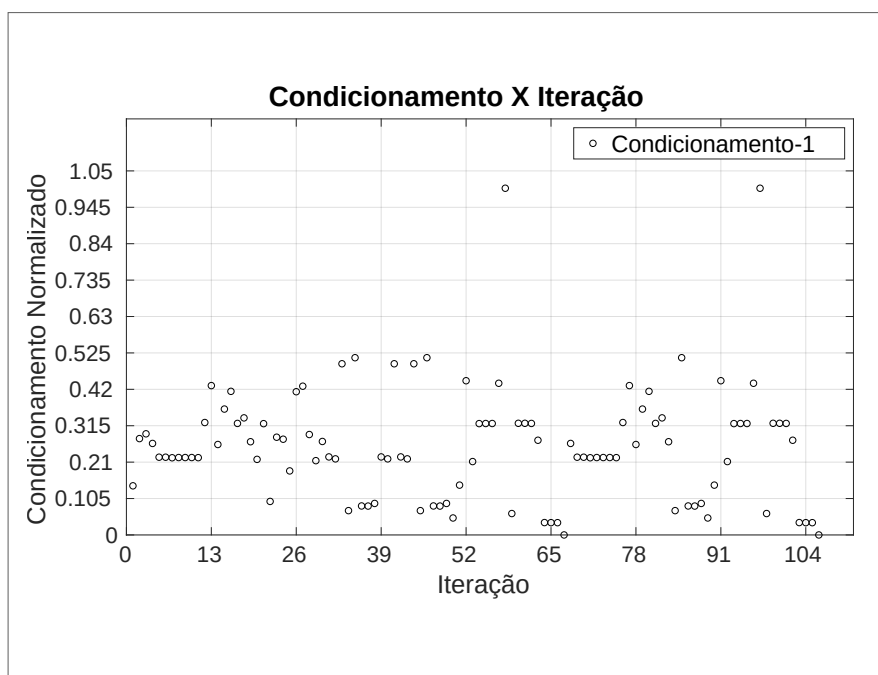
Fonte: Próprio autor

5.1.6 Condicionamento na Minimização da Métrica na aproximação dos compostos.

O comportamento do condicionamento das matrizes de covariância na minimização da métrica também foi alvo de investigação. Os resultados para a instância um são descritos na presente seção, nas ocorrências de minimização dos resíduos, como descritos na Seção 5.1.4, enquanto que as instâncias dois, três e quatro tem seus resultados descritos no Apêndice H, Seção H.1.

Os valores de condicionamento normalizados, em trinta execuções do algoritmo, apresentados na Figura 30, demonstram uma tendência de convergência na minimização da métrica, mas essa convergência não ocorre de maneira monotônica. Esse fato evidencia que os valores de condicionamento podem vir a ser utilizados como critérios para a qualidade da determinação das frações de contagem percentual, caso seja observada a diferença nesses valores entre iterações consecutivas. Esse fato também leva à conclusão de que a minimização dos resíduos também conduz a uma diminuição das instabilidades numéricas associadas ao problema, já que o condicionamento apresenta essa tendência a ser minimizado, mesmo que essa tendência não seja monotônica.

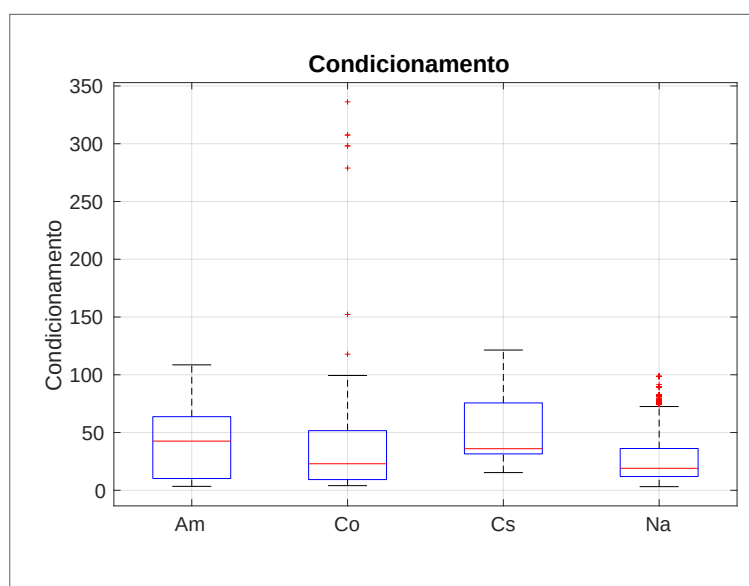
Figura 30 – Comportamento do condicionamento na minimização da métrica. Na Figura 30, são descritos os resultados para a instância 1. Os valores tendem a se estabilizar em um intervalo, evidenciando que a minimização dos resíduos podem ser associados a um decrescimento do condicionamento, mesmo que esse decrescimento não seja monotônico.



Fonte: Próprio autor

Na Figura 31 são apresentadas as distribuições dos verdadeiros valores de condicionamento na determinação das frações de contagem percentual. O intervalo de condicionamento para os compostos *Am* e *Cs* é de maior amplitude que os demonstrados para os compostos *Co* e *Na*. Os valores de condicionamento apresentados para as matrizes de dimensão 4 são próximos aos valores determinados pelo trabalho de Meriç (2012) para matrizes de dimensão 2. Esse fato é importante de ser destacado, já que se espera que as matrizes de dimensão 4 tenham condicionamento maiores que as matrizes de dimensão 2, o que não se verifica para o presente caso. Maiores valores de condicionamento implicam em maiores instabilidades numéricas ao problema, refletindo-se na qualidade da determinação das frações em peso.

Figura 31 – Boxplot demonstrando a distribuição de valores do condicionamento das matrizes de covariância na determinação das frações de contagem percentual na omissão de uma das bibliotecas da solução inicial. O primeiro plot é considerado para a instância 1, e os demais plots são descritos para as demais instâncias.



Fonte: Próprio autor

5.1.7 Resultados dos passos básicos do GRASP

As seções 5.1.7.1, 5.1.7.2, 5.1.7.3 e 5.1.7.4 apresentam os resultados para os quatro passos básicos do GRASP: (1) o passo construtivo (Seção 5.1.7.1); (2) o estágio da primeira busca (Seção 5.1.7.2); (3) a etapa de linkagem (Seção 5.1.7.3) e, por fim, (4) a etapa da segunda busca (Seção 5.1.7.4). A descrição dessas etapas serão feitas tomando por base a instância 1. As demais instâncias para a Base 1, instâncias 2, 3 e 4, por questões de similaridades nos resultados, são apresentados no Apêndice I: na Seção I.1.1 são descritos os resultados para o passo construtivo para essas instâncias; em I.1.2 são apresentados os resultados para o estágio de busca; os resultados da linkagem para as instâncias estão contemplados na Seção I.1.3 e, por fim, são apresentados os resultados para a segunda busca na Seção I.1.4.

Os seguintes parâmetros serão observados na apresentação dos resultados para as etapas básicas do GRASP:

- Canais na partição de maior resíduo;
- Distribuição de valores para os canais que definem a partição de maior resíduo;
- Gráfico contendo valores de integral abaixo da curva para a biblioteca omitida do espectro total;
- Distribuição de valores para a biblioteca omitida da solução inicial;
- Box e gráfico contendo os valores dos fatores multiplicativos que irão alterar a contagem espectral da biblioteca omitida.

5.1.7.1 Resultados do Passo construtivo

As informações trazidas pelo passo construtivo, descritas na abertura da Seção 5.1.7, são descritos na presente seção para a instância 1. Os resultados para o passo construtivo, observando-se cada um dos parâmetros descritos na presente seção, para as instâncias 2, 3 e 4, podem ser conferidos no Apêndice I.1, Seção I.1.1. Na presente seção, serão observadas as saídas para os intervalos que definem o compartimento da partição de maior resíduo, os valores de estimativa da biblioteca omitida e os valores das constantes que modificam as estimativas da biblioteca omitida durante o passo construtivo. Esse mesmo procedimento será adotado para as Seções 5.1.7.2, 5.1.7.3 e 5.1.7.4.

A Figura 32 apresenta os números inteiros que indexam os canais que definem o compartimento da partição de maior resíduo. O passo construtivo determina esses valores de início e fim do compartimento de maneira aleatória, obedecendo-se à restrição de que haja um mínimo de dez canais no interior desse compartimento. Os valores são números contíguos, sem interrupções. Esse fato era esperado do passo construtivo, já que o mesmo não realiza refinamentos em busca de melhores soluções. O gráfico 32a demonstra que o número de canais dentro do compartimento tem variações bruscas, desde o número mínimo de dez canais, até ocorrências em que o número de canais é de aproximadamente de quarenta. A questão de os compartimentos não apresentarem ‘vazios’ nos canais que os definem mostra-se de grande relevância, de modo que novos procedimentos podem ser adotados na modelagem do problema pelos algoritmos que abordam o PGNAA pela otimização do LLS.

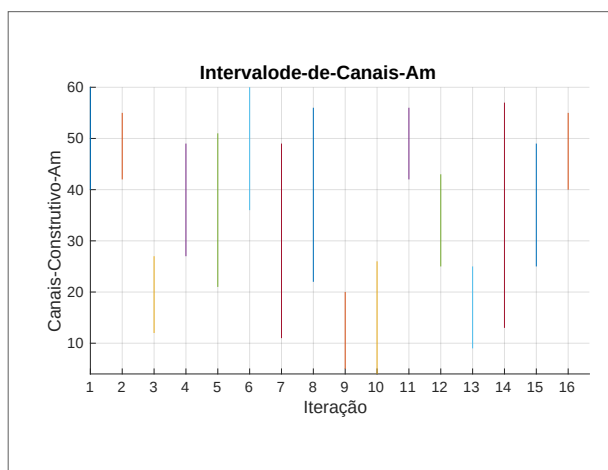
A distribuição desses canais, apresentada na Figura 32b, apresenta-se relativamente simétrica, com a distribuição que define o menor canal dentro do compartimento, mais ampla que a distribuição do maior canal. Como a amplitude da distribuição para o canal inicial foi maior, a liberdade para escolher o canal final do compartimento torna-se menor, demonstrando uma distribuição com menor amplitude. As duas distribuições são aproximadamente simétricas, fornecendo ao algoritmo a importante característica de aleatoriedade que auxilia no sentido de aumentar as probabilidades de encontrar uma solução que seja mais próxima possível de ótimas soluções.

As estimativas para a biblioteca omitida, na partição de maior resíduo, no passo construtivo do GRASP, estão apresentadas na Figura 33a (valores normalizados), com as distribuições dessas valores na Figura 33b. Essas estimativas apresentam uma tendência de definição inicial no intervalo pertencente ao primeiro quartil. Esse fato pode ser explicado a partir da questão que os valores estimados apresentaram-se bem próximos dos valores de contagem espectral total real, como pôde ser visto na Figura 28, restringindo, assim, as estimativas dessa biblioteca durante o estágio de construção de uma solução. Mas esse parâmetro evidencia que as estimativas são efetivamente aleatórias, contribuindo, assim, para fornecer as características necessárias de aleatoriedade ao algoritmo, possibilitando determinar soluções que sejam ótimas soluções locais.

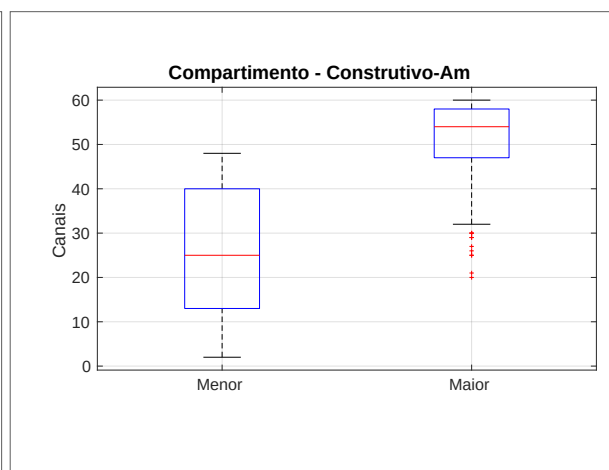
A figura 34a apresenta as distribuições do fator que modifica a contagem da biblioteca estimada, após o primeiro ajuste do CGRASP na solução vinda do passo construtivo do GRASP. Esses fatores podem decrementar os valores de contagem espectral estimada, caso estejam no intervalo compreendido entre $[0, 1]$, ou podem incrementar, caso seja escolhido para esse

Figura 32 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para a instância 1, no estágio da construção. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância 1.

(a) Intervalo de canais como função da iteração, instância 1, no estágio de construção de uma solução



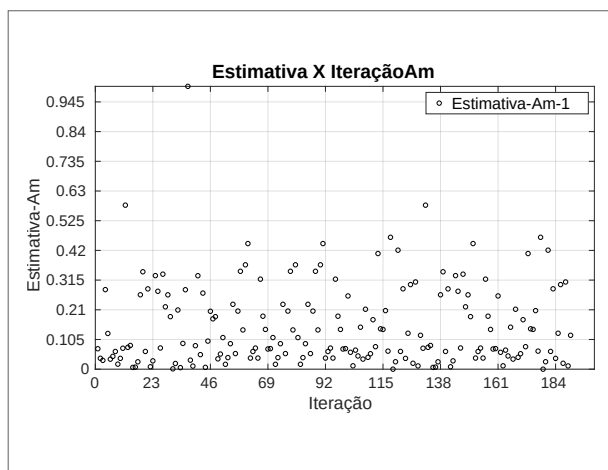
(b) Distribuição de canais que define o compartimento na partição de maior resíduo, no estágio de construção de uma solução.



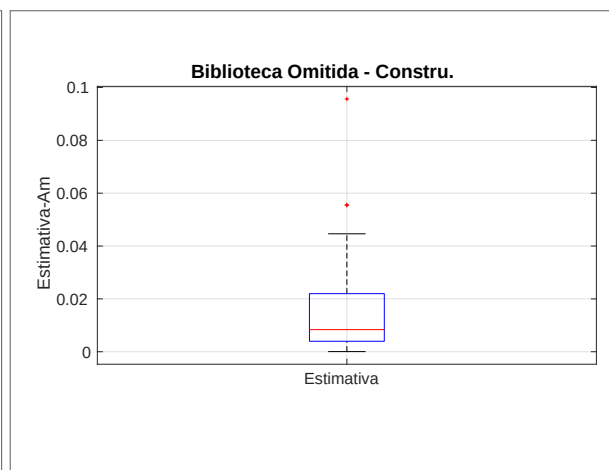
Fonte: Próprio autor

Figura 33 – O gráfico 33a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 33b.

(a) Estimativas para o elemento omitido durante o passo construtivo, instância um, ausência do *Am*.



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo construtivo, instância um.

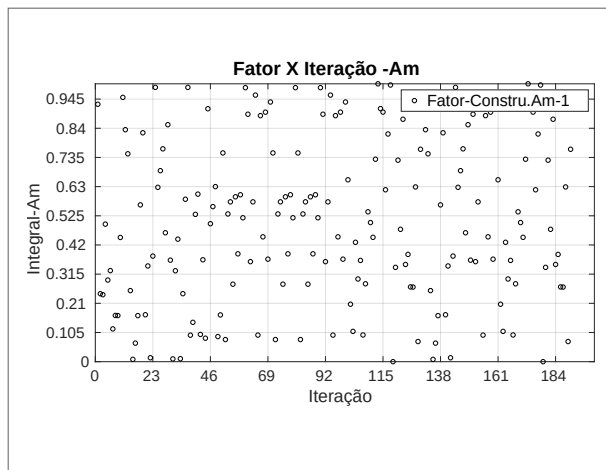


Fonte: Próprio autor

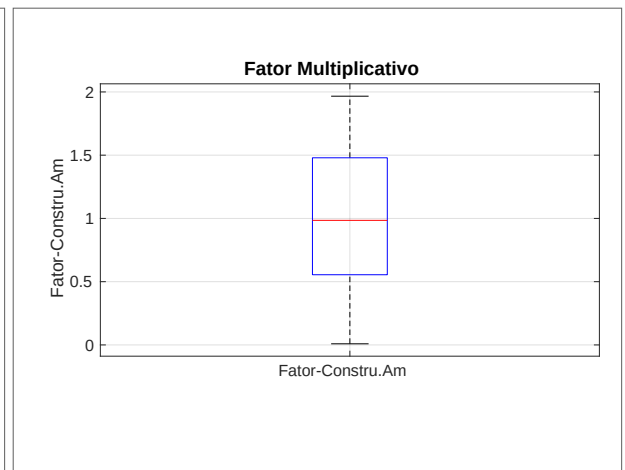
fator uma valor maior que um. O valor dessa constante é limitada superiormente pelo número que leva a estimativa ao limite superior do intervalo definido pela geometrização. Importante observar como esse fator é aleatório, ratificando mais uma vez a importante característica de aleatoriedade aplicada ao algoritmo, devidamente modelado e adaptado ao domínio do problema. Essa característica também pode ser visualizada pela Figura 34b, em que a distribuição desses valores são simétricos.

Figura 34 – O gráfico 34a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 34b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância um, no passo construtivo.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância um, passo construtivo.



Fonte: Próprio autor

5.1.7.2 Resultados do Passo de Busca

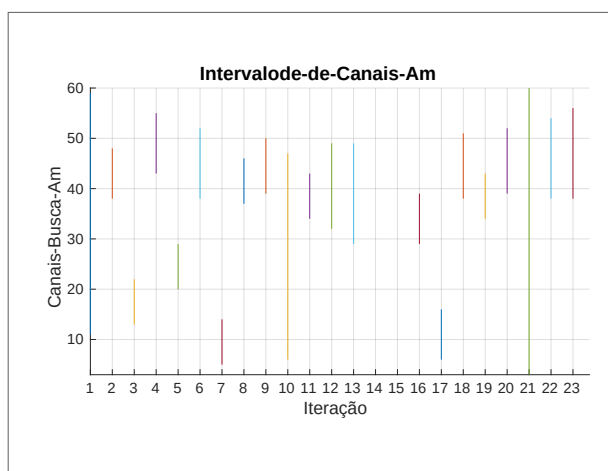
A presente seção trará uma discussão sobre os resultados dos principais parâmetros no passo de busca do GRASP, comparando-os com os parâmetros definidos no passo construtivo. As demais instâncias da Base 1 podem ser conferidas no Apêndice I.1, na Seção I.1.2.

A figura 35a apresenta a definição do compartimento na partição de maior resíduo para o estágio da primeira Busca. Comparativamente à Figura 32a, o passo de busca evidencia uma diminuição na amplitude dos intervalos que definem os compartimentos, indicando que ocorreram retiradas de canais no compartimento da partição de maior resíduo desde o passo construtivo até o fim do estágio da primeira busca.

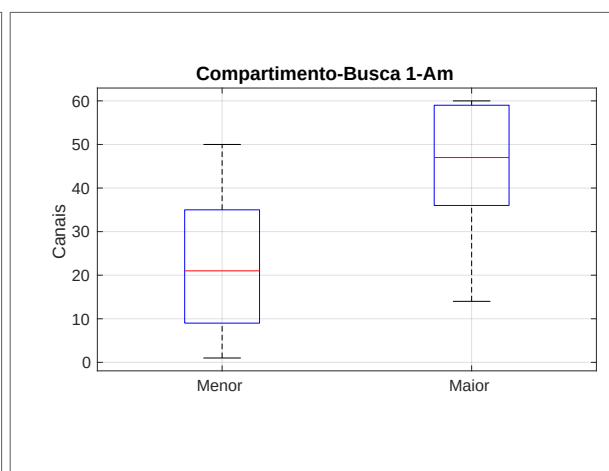
As distribuições apresentadas na Figura 35b determinam um diminuição dos compartimentos que iniciam no canal quarenta, ao realizar a comparação com a Figura 32b, e um aumento no número de compartimentos que encerram no canal quarenta. A premissa do passo de busca em retirar, e adicionar canais, tende a diminuir a amplitude dos compartimentos, como também faz com que a distribuição dos menores e maiores canais apresente-se mais simétrico, ao comparar a Figura 35b com a Figura 32b.

Figura 35 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, instância 1 (Figura 32a) no estágio da busca. A figura 32b traz a distribuição dos canais que definem o compartimento da partição de maior resíduo na ausência do *Am*.

(a) Intervalo de canais como função da iteração, omitindo *Am*, no estágio de busca de uma solução



(b) Distribuição de canais que define o compartimento na partição de maior resíduo, no estágio de busca de uma solução.



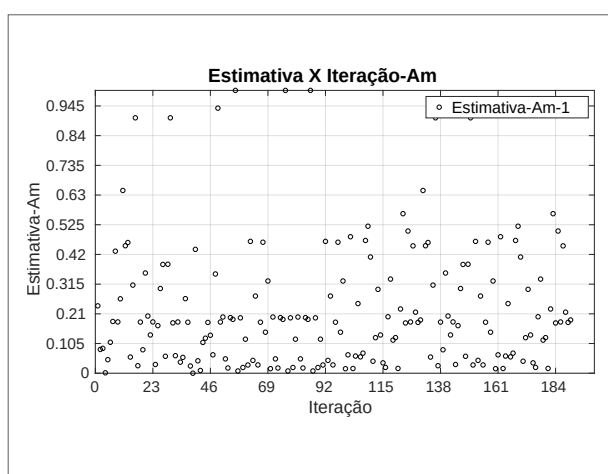
Fonte: Próprio autor

As estimativas para a biblioteca omitida, na partição de maior resíduo, no passo de busca, são apresentadas na Figura 36a. A diminuição da amplitude dos compartimentos, apresentada

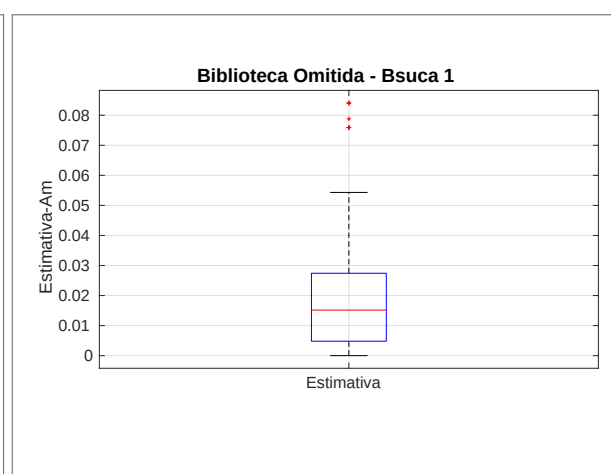
na Figura 35a, vem acompanhada por um aumento nos valores da estimativa da biblioteca omitida, conforme pode ser conferido comparando as Figuras 33a e 36a. A comparação entre as Figuras 36b e 33b demonstra que existe a tendência em se realizar uma majoração na contagem espectral do elemento omitido no passo de busca.

Figura 36 – O gráfico 36a apresenta os valores das estimativas para a biblioteca omitida, no passo de busca do GRASP, instância um. A distribuição para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 36b.

(a) Estimativas para o elemento omitido durante o passo de busca, instância um, ausência do *Am*.



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de busca, instância um.



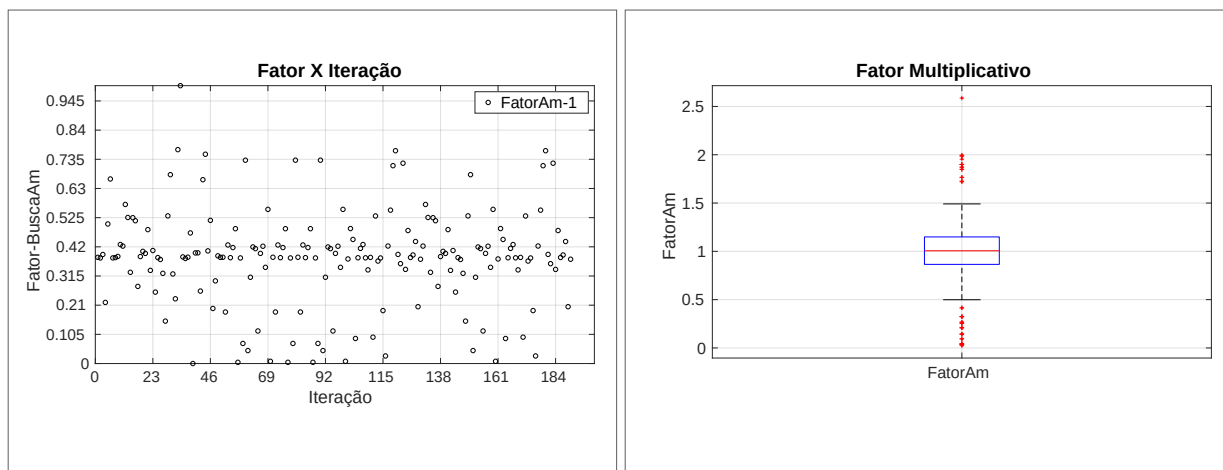
Fonte: Próprio autor

As Figuras 37a e 37b apresentam, respectivamente, o gráfico com as constantes que modificam a biblioteca omitida no passo de busca e as distribuições dessas constantes em trinta execuções do algoritmo. Os resultados apresentados pelo passo de busca para a constante multiplicativa apresentam a peculiaridade desses valores normalizados estarem compreendidos, boa parte deles, entre 0.4 e 0.42, diferentemente do que foi apresentado pelo Gráfico 34a em que os valores apresentavam-se melhor distribuídos. A distribuição da Figura 37b apresenta-se simétrica, mas de menor amplitude do que a apresentada na Figura 34b, ratificando que esse passo tende a restringir a distribuição desses valores. As constantes determinadas no passo de busca percorrem o espaço de soluções com uma eficiência menor do que as saídas determinadas na seção anterior.

Figura 37 – O gráfico 34a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância um. A distribuição para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 34b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância um, no passo da primeira busca.

(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância um, passo de busca.



Fonte: Próprio autor

5.1.7.3 Resultados do Passo de Linkagem

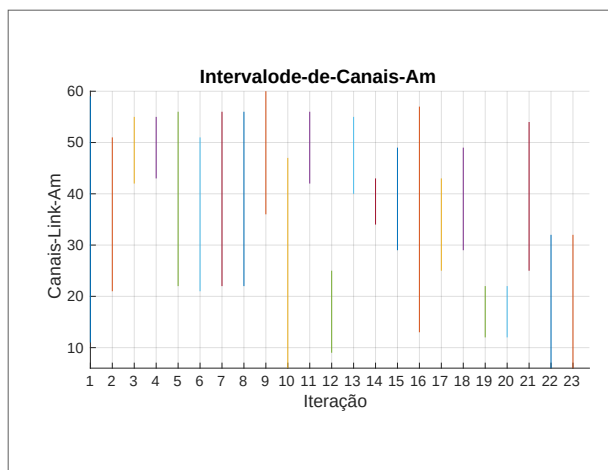
Os resultados da linkagem na modificação dos parâmetros dos compartimentos são descritos na presente seção. Os resultados serão comparados com as observações determinadas no passo construtivo, Seção 5.1.7.1, e no passo de busca, Seção 5.1.7.2. As descrições aqui presentes se aplicam à instância 1. As demais instâncias da Base 1 podem ser conferidas no Apêndice 1, Seção I.1.3.

Os canais inseridos no compartimento da partição de maior resíduo, na Figura 38a, seguem a mesma tendência observada no passo construtivo e no passo de busca: os valores dos inteiros que indexam os canais a serem inseridos são contíguos, mas com a observação de que a amplitude desses intervalos, mais uma vez, aumentam no passo de linkagem, comparativamente ao passo de busca. Esse fato é decorrente da premissa que a comparação entre as soluções *start* e *target* tende a aumentar o número de canais nos compartimentos, inserindo os canais da solução *target* na solução *start*. A distribuição desses canais, apresentada na Figura 38b, demonstra que os compartimentos apresentam maior tendência a serem definidos até o canal quarenta, diferentemente do passo de busca, em que ocorreu uma maior tendência dos compartimentos serem definidos até o canal trinta e cinco, aproximadamente, de acordo com a Figura 35b.

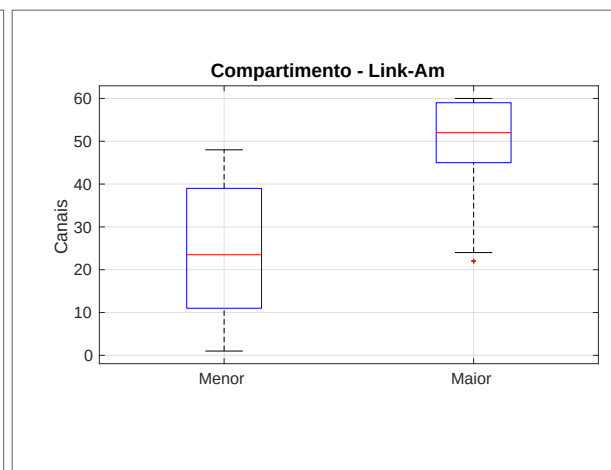
Comparando-se as Figuras 39a e 36a, percebe-se que não há diferenças significativas na estimativa da contagem espectral do elemento omitido na partição de maior resíduo entre os

Figura 38 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do Am , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Am .

(a) Intervalo de canais, como função da iteração, para a instância 1, no estágio de linkagem entre soluções.



(b) Distribuição de canais que define o compartimento na partição de maior resíduo, no estágio de linkagem de uma solução.



Fonte: Próprio autor

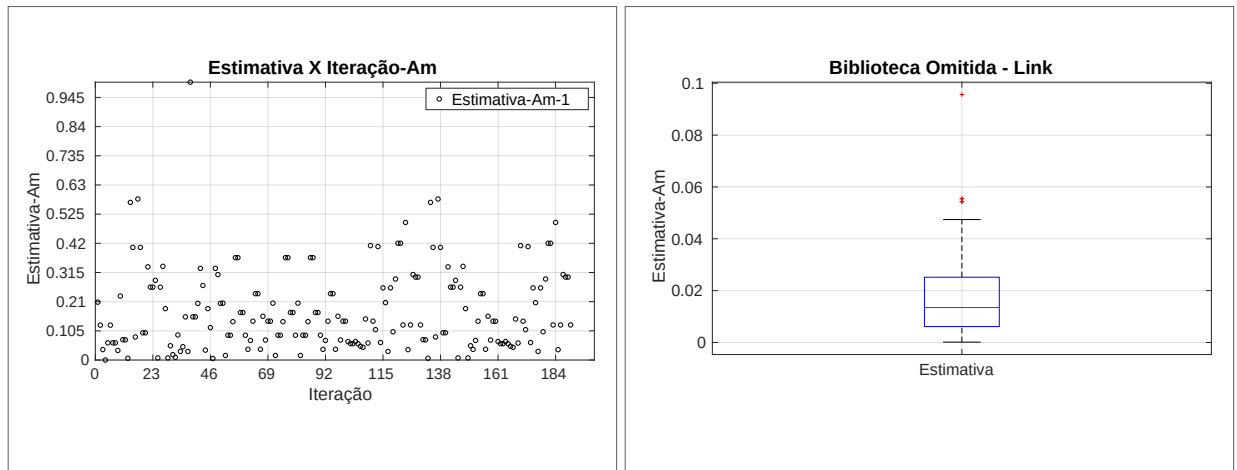
passo de busca e linkagem. O passo construtivo define de maneira aleatória esses valores, a partir da definição do intervalo pelo passo de geometrização, sendo que os passos de busca e linkagem tendem, aproximadamente, a incrementar essa estimativa. Esse fato pode ser ratificado a partir das distribuições presentes nas Figuras 33b, 36b e 39b.

As constantes que modificam os valores das estimativas da contagem espectral da biblioteca omitida, presente nas Figuras 34a, 37a e 40a, não apresentam diferenças significativas, mas ratificam, mais uma vez, que esses valores são igualmente distribuídos no intervalo definido para eles pela geometrização, contribuindo para a aleatoriedade do algoritmo e a qualidade esperada para as soluções. As distribuições presentes nas Figuras 34b, 37b e 40b descrevem que grande parte dos intervalos de valores, determinados pela geometrização para modificação da estimativa estão compreendidos entre 0.5 e 1.5.

Figura 39 – O gráfico 39a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 39b.

(a) Estimativas para o elemento omitido durante o passo de linkagem, instância um, ausência do Am .

(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de linkagem, instância um.

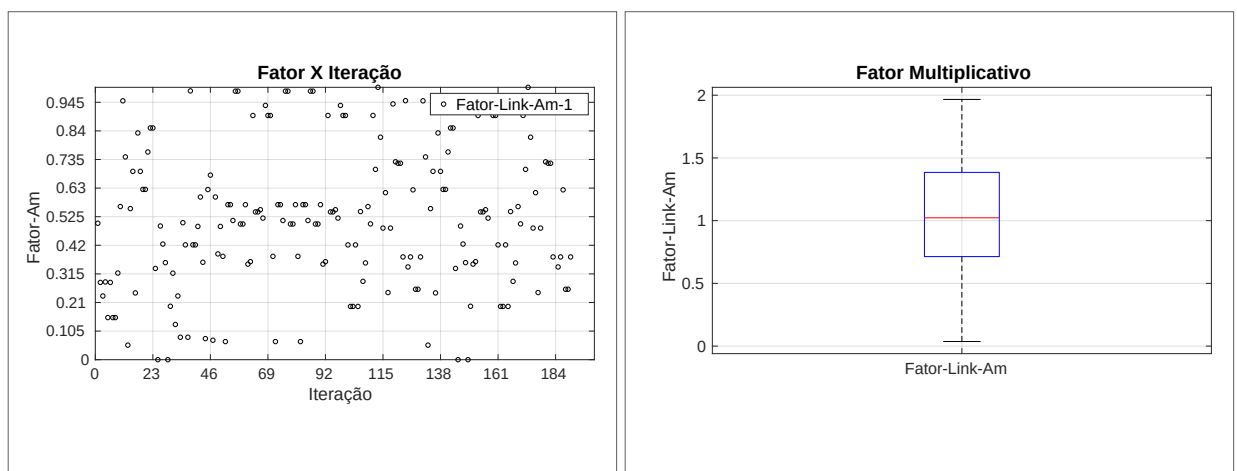


Fonte: Próprio autor

Figura 40 – O gráfico 40a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 40b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância um, no passo da linkagem.

(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância um, passo de linkagem.



Fonte: Próprio autor

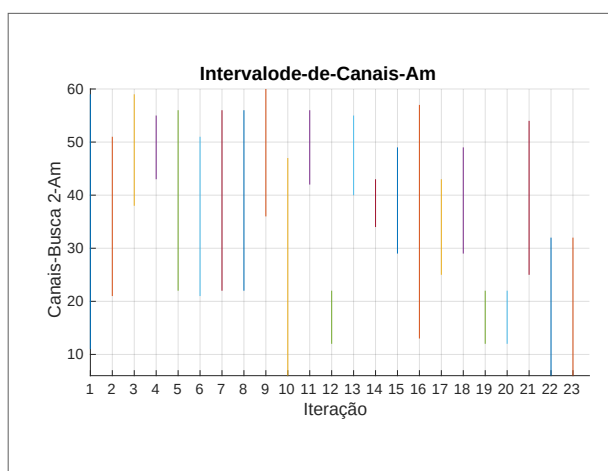
5.1.7.4 Resultados do Passo da Segunda Busca

Os resultados da presente seção são aplicados à instância 1. Os resultados aplicados as demais instâncias podem ser conferidos no Apêndice I.1, na Seção I.1.4.

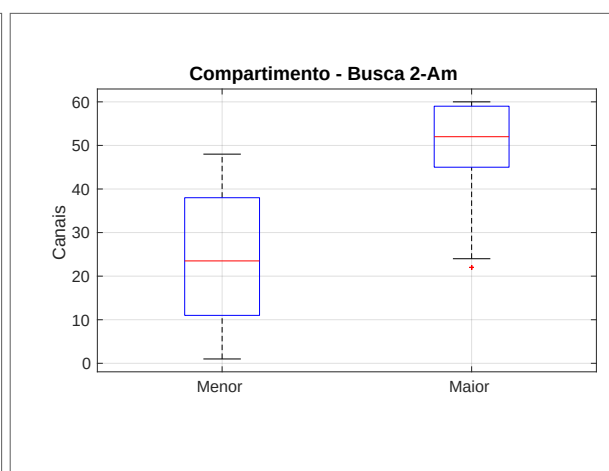
Os canais utilizados para definir os compartimentos, presentes na Figura 41a, não apresentam diferenças significativas, na amplitude dos intervalos, em relação ao passo descrito na Seção 5.1.7.3, Figura 38a. As distribuições apresentadas na Figura 41b são significativamente semelhantes às distribuições da Figura 38b, evidenciando que a segunda busca não deve realizar mudanças significativas nas soluções vindas a partir da linkagem. A linkagem, portanto, é o estágio do AlgoritmoGRASP que irá definir as principais características das saídas nas iterações do nosso algoritmo.

Figura 41 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para a instância 1, no estágio da segunda busca. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do *Am*.

(a) Intervalo de canais como função da iteração, instância 1, no estágio de segunda busca de uma solução



(b) Distribuição de canais que define o compartimento na partição de maior resíduo, no estágio de segunda busca de uma solução.



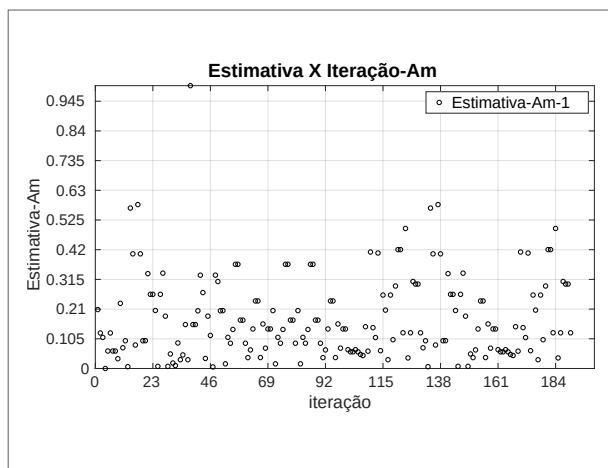
Fonte: Próprio autor

As estimativas para a biblioteca omitida, descritas na Figura 42a, também não escapam das observações realizadas na etapa de linkagem, apresentando evidências de similaridades em relação aos resultados da linkagem. Esse fato confirma, mais uma vez, que o passo da segunda busca não realiza modificações significativas nos parâmetros determinados no passo de linkagem, ratificada pelas distribuições desses valores na Figura 42b.

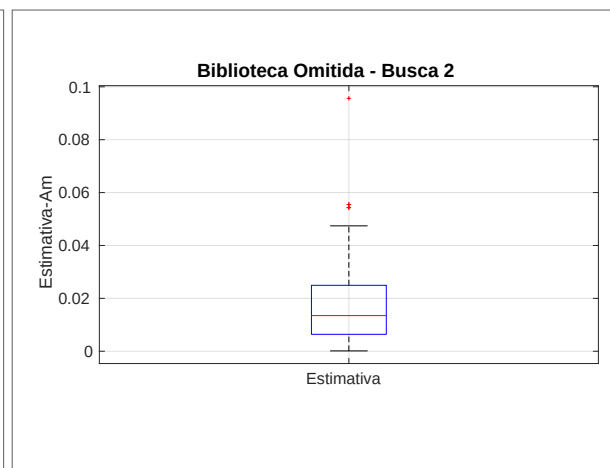
As constantes, mais uma vez, determinadas nas Figuras 43a e 43b, são significativamente aleatórias, com valores distribuídos no intervalo de 0.5 a 1.5. A definição das constantes,

Figura 42 – O gráfico 42a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 42b.

(a) Estimativas para o elemento omitido durante o da segunda busca, instância um, ausência do Am .



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de segunda busca, instância um.

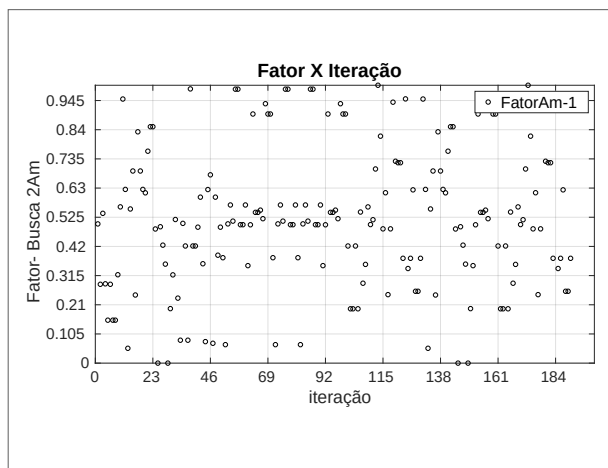


Fonte: Próprio autor

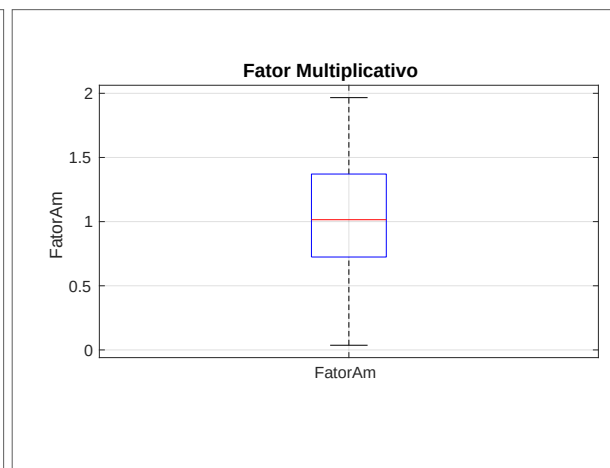
ao longo das quatro etapas básicas do GRASP, é evidenciada como aleatória, aplicando ao algoritmo trabalhado a aleatoriedade necessária no intuito de apresentar soluções com a maior diversidade possível, garantindo que as melhores soluções encontradas sejam melhores que as soluções médias.

Figura 43 – O gráfico 43a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 43b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância um, no passo da segunda busca.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância um, passo da segunda busca.



Fonte: Próprio autor

5.1.8 Resultados do CGRASP

O Algoritmo CGRASP será alvo de investigação da convergência das aproximações das frações em peso aos valores das frações nominais. As variáveis que serão alvo de investigação na execução do CGRASP, para a instância 1, são:

- Resultados de uma iteração arbitrária (Seção 5.1.8.1) para as frações de contagem percentual, resíduo e o condicionamento da matriz de covariância como função da iteração. Essa questão é importante no intuito de verificar a convergência do algoritmo proposto;
- Tabelas apresentando os valores de frações de contagem percentuais estimadas e os valores de métrica associados, na Seção 5.1.8.2 para a instância 1. Ainda na Seção 5.1.8.2, são apresentadas as distribuições das frações de contagem percentual nas melhorias do CGRASP (Figura 45);
- Gráficos, com as frações de contagem percentual como função da iteração e valores de métricas reais e estimadas como função da iteração para o acumulado das execuções para a instância 1, nas etapas de construção e busca local, nas Seções 5.1.8.3 e 5.1.8.4.

Os resultados para as demais instâncias, por similaridade, serão apresentadas no Apêndice J, Seção J.1: resultados para uma única execução estão apresentados na Seção J.1.1; etapa Construtiva e os parâmetros associados a essa etapa estão na Seção J.1.2; os resultados da busca estão na Seção J.1.3.

5.1.8.1 Resultados do CGRASP - Execução Única.

Os resultados para execução única do CGRASP são descritos na presente seção, para a instância 1. As aproximações das frações de contagem percentual, resíduo e condicionamento são descritos como função da iteração nos gráficos mostrados na presente seção.

As Figuras 44a e 44b apresentam os resultados de frações de contagem percentual para duas execuções arbitrárias para a instância 1. As demais instâncias da Base 1 estão descritas no Apêndice J.1, Seção J.1.1.

Os valores das frações de contagem percentual convergem aos valores nominais à medida que o valor do resíduo vai sendo reduzido. Importante destacar que essa convergência das frações também ocorre para a biblioteca estimada. Na figura 44a, os pontos ○ indicam as

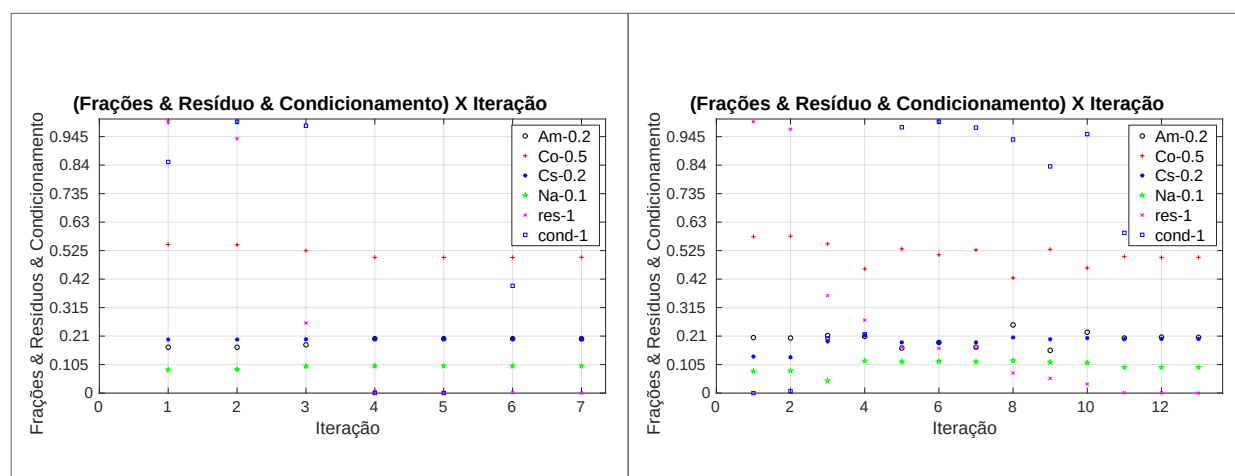
aproximações para o Am ; os pontos $+$ demonstram as aproximações para Co ; pontos $*$ indicam as aproximações para o Cs ; $\star$ são pontos para as aproximações para o Na ; os pontos $\times$ são os valores para o resíduo e o número de condicionamento é representados por pontos $\square$. Importante observar como a minimização do resíduo conduziu, de maneira efetiva, as frações de contagem percentual estimada pelos mínimos quadrados em direção aos verdadeiros valores de frações de contagem percentual. Esse fato revela a importância do ajuste do CGRASP nas saídas finais do algoritmo trabalhado na presente tese.

Os valores de condicionamento (pontos $\square$) não apresentaram uma minimização monotônica, como no caso do resíduo, mas percebe-se que existe uma tendência de delimitação desse número a um intervalo, à medida que as saídas das frações de contagem percentual estimadas vão se aproximando dos valores nominais. Esse fator apresenta um indicativo de que o número de condicionamento pode, futuramente, servir como parâmetro da qualidade das aproximações das frações em peso, ao determinar a diferença do condicionamento entre duas iterações consecutivas.

Figura 44 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas figuras 44a e 44b são apresentadas as frações percentuais para a instância 1. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. Os pontos $\circ$ indicam as aproximações para o Am ; os pontos $+$ são as aproximações para Co ; as aproximações para o Cs são determinados por pontos $*$; pontos $\star$ indicam as aproximações para o Na ; Os valores para os resíduos são pontos $\times$ e os pontos $\square$ indicam os valores para condicionamento.

(a) Percentuais dos elementos e respectivas métricas. Instância 1

(b) Percentuais dos elementos e respectivas métricas. Instância 1



Fonte: Próprio autor

5.1.8.2 Resultados do CGRASP - Execuções Acumuladas / Valores Métricas

Os valores de resíduos e de condicionamento serão alvo de novas investigações na presente seção, na determinação da qualidade das frações de contagem percentual, agora no âmbito da execução do Algoritmo CGRASP.

A Tabela 10 contém os valores médios de frações de contagem percentual, de resíduos e de condicionamento, em todas as chamadas do CGRASP, em trinta execuções do algoritmo. Mais uma vez, assim como já descrito para a Tabela 7, os valores médios não podem ser definidos como parâmetros para a melhor saída do nosso algoritmo.

Tabela 10 – Frações de contagem percentual médios para as execuções do CGRASP. A linha 1 determina as frações em peso para as bibliotecas, na ausência do *Am*, na solução inicial. A linha 2 apresenta os resultados para a ausência do *Co*. Linha 3 e 4, respectivamente, seguem a mesma descrição, para a ausência *Cs* e *Na*. Os valores das precisões estão descritos pelas colunas que se iniciam com o prefixo 'pre': *PreAm*, precisão para o *Am*; *PreCo*, precisão para o *Co*, e assim sucessivamente.

	Am	PreAm	Co	PreCo	Cs	PreCs	Na	PreNa	Resíduo	Condicionamento
1	0.1577	0.21151	0.55059	0.10119	0.17347	0.13267	0.11778	0.17783	0.015658	3.7065
2	0.33878	0.69391	0.26875	0.46251	0.22267	0.11336	0.16822	0.68218	0.024844	8.5442
3	0.20576	0.028795	0.54163	0.083253	0.12726	0.36372	0.12516	0.25157	0.025542	6.6154
4	0.18509	0.074543	0.48633	0.027344	0.17287	0.13564	0.15532	0.55324	0.00596	6.0599

Fonte: Próprio autor

As frações de contagem percentual são melhores preditas na minimização dos resíduos, conforme pode ser conferido na Tabela 11, e também pôde ser conferido na Tabela 8. Uma dos grandes destaques do presente trabalho foi verificar como o resíduo consegue ser bem sucedido ao determinar os valores das frações em peso, comparativamente ao condicionamento (Tabela 12) e à determinação de valores médios de trinta execuções, descritos na Tabela 10.

Observando-se do ponto de vista dos condicionamentos, assim como na Tabela 9, no âmbito agora das saídas do CGRASP, as frações de contagem percentual descritas na Tabela 12 não apresentaram a mesma eficácia em suas predições que os valores encontrados na Tabela 11 para a minimização dos resíduos. Esse fato consolida os resíduos como a melhor métrica na predição das frações em peso.

As Figuras 45a, 45b, 45c e 45d trazem as distribuições dos valores de frações de contagem percentual para as execuções do CGRASP na Base de Dados 1. As medianas em verde, em cada uma das descrições das instâncias, indicam o elemento omitido e aproximado nas iterações do CGRASP. As instâncias um (Figura 45a), três (Figura 45c) e quatro (Figura 45d), apresentem distribuições das frações de contagem percentual relativamente melhor centradas em torno dos

Tabela 11 – Frações em peso aproximadas na determinação quantitativa de compostos inseridos no espectro total, tomando como métrica a aproximação do resíduo. A linha 1 da tabela determina os valores de frações percentuais omitindo o *Am* da geração do espectro total. A mesma descrição é válida para as demais linhas: linha 2, omite-se o *Co*; linha 3, omite-se o *Cs* e na última linha é omitido o *Na*. As colunas *PreAm*, *PreCo*, *PreCs* e *PreNa* indicam as precisões na aproximação das frações de contagem percentuais para o *Am*, *Co*, *Cs* e *Na*. A última coluna traz o menor valor de resíduo.

	Am	PreAm	Co	PreCo	Cs	PreCs	Na	PreNa	Resíduo
1	0.19976	0.0011945	0.50027	0.0005372	0.20002	0.0001025	0.099925	0.0007476	6.4825e-05
2	0.19996	0.0002195	0.49731	0.0053798	0.20209	0.010436	0.10083	0.008315	0.00042434
3	0.20004	0.0001915	0.50004	7.8e-05	0.20005	0.000232	0.099701	0.0029895	4.1566e-05
4	0.20236	0.011801	0.50860	0.01720	0.19182	0.040909	0.09721	0.02790	1.4888e-08

Fonte: Próprio autor

Tabela 12 – Valores de percentuais aproximados na determinação quantitativa de compostos inseridos em um espectro total. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o *Am* da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas.

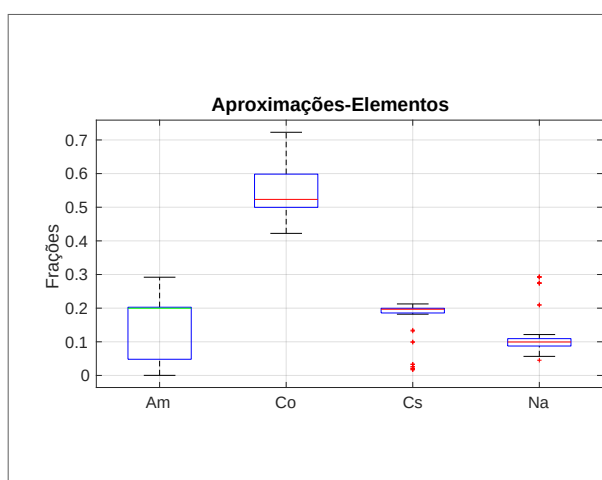
	Am	PreAm	Co	PreCo	Cs	PreCs	Na	PreNa	Nº Condicionamento
1	0.11129	0.44356	0.59841	0.19683	0.19639	0.018053	0.092084	0.07916	1.9964
2	0.16181	0.19095	0.30757	0.38487	0.42621	1.1311	0.084411	0.15589	1.2527
3	0.2089	0.044521	0.50443	0.0088532	0.1646	0.17701	0.12187	0.2187	3.0192
4	0.19573	0.021349	0.53077	0.061531	0.22013	0.10065	0.052174	0.47826	1.6995

Fonte: Próprio autor

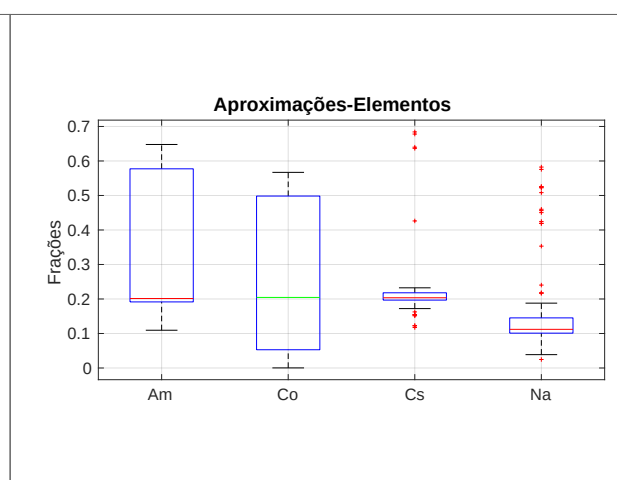
seus verdadeiros valores de frações que a instância dois. Apesar de a instância dois apresentar uma grande amplitude nas predições das frações nas saídas do CGRASP, ainda assim, foi possível ser bem sucedido na aproximação das frações, nessa instância, na minimização dos resíduos. A amplitude na determinação dos resultados para a instância 2 foi mais uma vez verificadas nas saídas do CGRASP, assim como foi observado nas saídas do GRASP, Seção 5.1.4. Mais uma vez associa-se esse fato à questão do pico do *Co* estar presente em uma região do espectro que ocorre após todas as outras contagens terem ido a zero.

Figura 45 – Valores de frações de contagem percentuais nas iterações do CGRASP. Na Figura 45a estão as distribuições para fração de contagem percentual dos compostos para a instância um, omissão do *Am*. Figura 45b, instância dois, omissão do *Co*, Figura 45c, instância três, omissão do *Cs* e, por fim, 45d, instância quatro, omissão do *Na*.

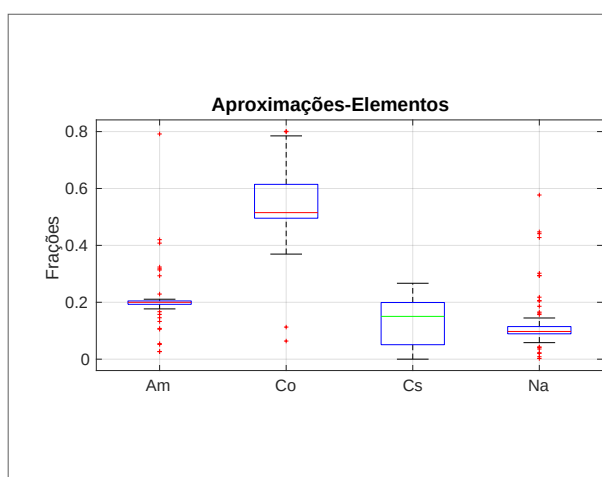
(a) Distribuição das Frações de contagem percentual nas iterações do CGRASP, instância um, mediante ausência e estimacão da biblioteca do *Am*.



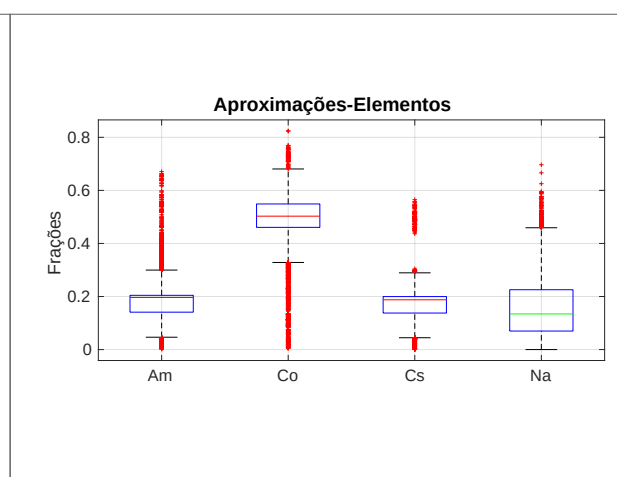
(b) Distribuição das Frações de contagem percentual nas iterações do CGRASP, instância dois, mediante ausência e estimacão da biblioteca do *Co*.



(c) Distribuição das Frações de contagem percentual nas iterações do CGRASP, instância três, mediante ausência e estimacão da biblioteca do *Cs*.



(d) Distribuição das Frações de contagem percentual nas iterações do CGRASP, instância quatro, mediante ausência e estimacão da biblioteca do *Na*.



Fonte: Próprio autor

5.1.8.3 Resultados do CGrasp - Etapa Construtiva.

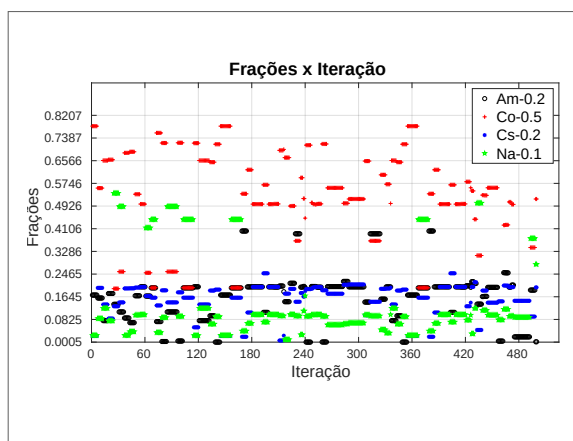
A qualidade das soluções na etapa construtiva do CGRASP é apresentada na presente seção. Os gráficos de frações em peso e de resíduo serão comparados tomando por base as mesmas iterações para avaliar a qualidade das saídas durante a etapa construtiva do CGRASP. O detalhamento dos resultados nessa seção serão para a instância 1, enquanto que os resultados para as demais instâncias da Base 1 podem ser conferidos no Apêndice J, Seção J.1.2, por questões de similaridades aos resultados da instância 1.

A Figura 46a descreve as frações de contagem percentual, para a primeira instância (omissão do Am), para os 4 elementos da Base de Dados 1 (pontos $\circ$ representam o Am , pontos $+$ representam o Co , pontos $*$ representam o Cs e pontos $\star$ representam o Na), como função da iteração. A Figura 46b exibe as correspondências de métrica real, métrica aproximada e o valor do resíduo como função das mesmas iterações presentes na Figura 46a. As convergências para as frações dos elementos não é evidente como apresentado na Figura 24, mas comparando-se as Figuras 46a e 46b, principalmente nas regiões em que os valores de resíduo são menores, é possível perceber as frações convergindo para os valores nominais. Esse fato pode ser constatado nas iterações 180 a 210, aproximadamente, assim como nas iterações 280 a 320. As iterações centradas por volta da iteração 420 também apresentam essa característica, demonstrando a capacidade da métrica em realizar a convergência das frações aos valores nominais.

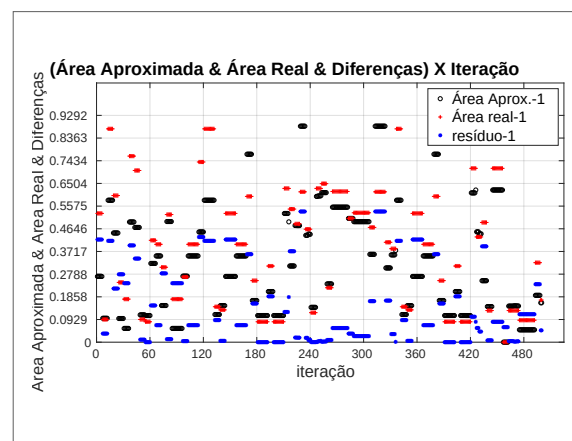
Os resultados apresentados para a etapa construtiva trazem uma boa perspectiva no sentido de convergência dos valores de fração de contagem percentual à medida que os valores de resíduos são reduzidos. As distribuições das frações de contagem percentual descritas na Figura 46c apresentam evidências de qualidade na precisão das frações de contagem percentuais dos elementos procurados. O elemento que apresenta maior amplitude nas distribuições de suas frações é o Co . Os demais elementos, inclusive o que está sendo estimado, apresenta uma amplitude nos valores menor que as apresentadas para o elemento do Co . Esse resultado do Co atribui-se às dificuldades no nosso algoritmo em localizar o pico dessa biblioteca, mais uma vez. Mas essa questão não foi proibitiva para uma boa aproximação das frações em peso para as bibliotecas da Base 1.

Figura 46 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 46a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (o *Am* é representado por pontos ○; O *Co* tem representação nos pontos +; o *Cs* é indicado por pontos * e o *Na* é representado por pontos ☆). A Figura 46b indica a área aproximada (pontos representados por ○), área real (pontos representados por +) e resíduo pontos representados por *), demonstrando a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 46c tem-se os valores de estimativa do composto omitido (*Am*) e a aproximação dos demais compostos.

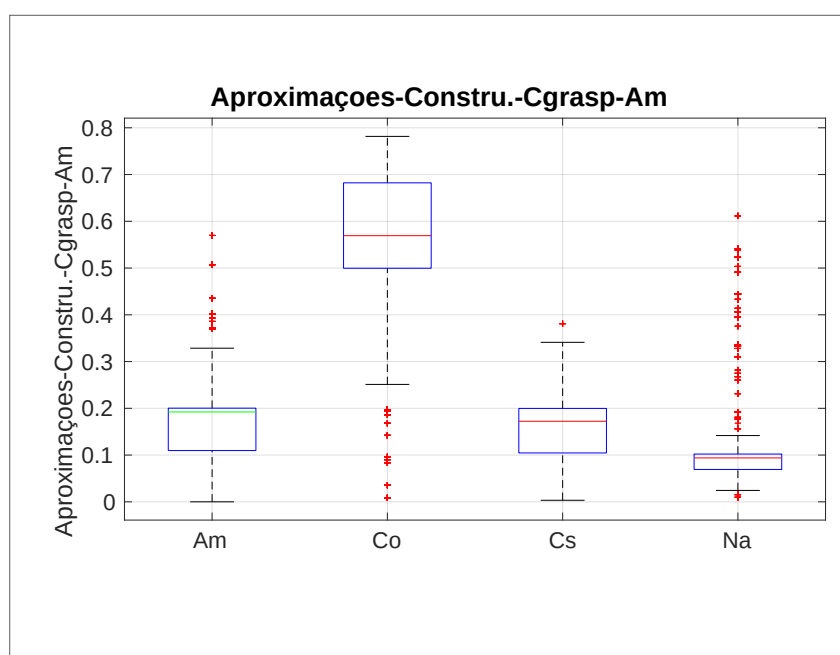
(a) Frações dos isótopos inseridos no espectro total, instância 1. O *Am* é representado por pontos ○; O *Co* é indicado por pontos +; O *Cs* tem representação nos pontos *; O *Na* é indicado por pontos ☆. Resultados para a etapa construtiva.



(b) Medidas da área abaixo da curva do espectro total. Área calculada é representada por pontos ○; área real é indicada por pontos + e o resíduo representada por pontos *. Resultados para a etapa construtiva.



(c) Box contendo as distribuições nas determinações percentuais das bibliotecas. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP.



Fonte: Próprio autor

5.1.8.4 Resultados do CGrasp - Etapa Busca.

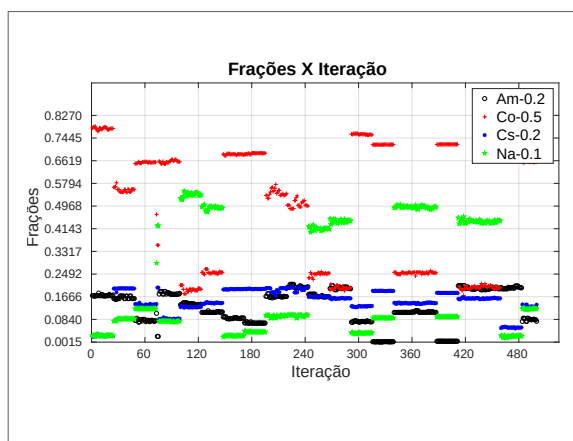
A Figura 47 apresenta as saídas do Algoritmo de Busca do CGRASP para a primeira base de dados. As demais instâncias são apresentadas no Apêndice J, Seção J.1.3. A interpretação a ser realizada para a Figura 47 é a mesma interpretação atribuída à Figura 46. Mais uma vez são comparadas as frações de contagem percentual com valores de área real, aproximada e resíduo. As distribuições das frações de contagem percentual são apresentadas à parte.

As evidências apresentadas na Seção 5.1.8.3 podem ser estendidas para os resultados apresentados na presente seção. A comparação entre os gráficos 47a e 47b evidenciam a convergência dos valores de frações de contagem percentual mediante a minimização da métrica utilizada. Esse fato pode ser evidenciado nas iterações 180 a 240, por exemplo. Esse é mais um indício de que a minimização do resíduo serviu como uma métrica de qualidade para boas aproximações das frações em peso. Os resultados também indicam que a modelagem da busca ao problema foi adequada, ao escolher, aleatoriamente, a partição para o ajuste da biblioteca omitida.

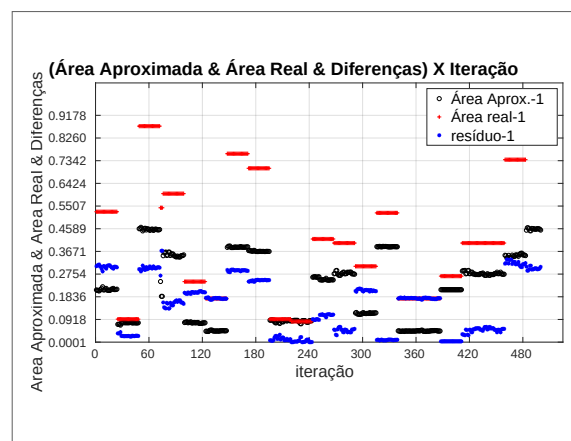
As distribuições das frações de contagem percentuais contidas na Figura 47c demonstram evidências de que as aproximações estão próximas dos valores nominais de frações para as bibliotecas que constituem o espectro total.

Figura 47 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 47a, têm-se os valores quantitativos dos isótopos que constituem o espectro total (pontos $\circ$ representam o *Am*; pontos $+$ representam o *Co*; pontos $*$ representam o *Cs*; pontos $\star$ representam o *Na*). Na Figura 47b tem-se a seguinte determinação: área calculada são os pontos $\circ$; área real são os pontos $+$; resíduo são os pontos $*$, demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 47c tem-se a distribuição das frações de contagem percentual, instância 1, e a aproximação dos demais compostos.

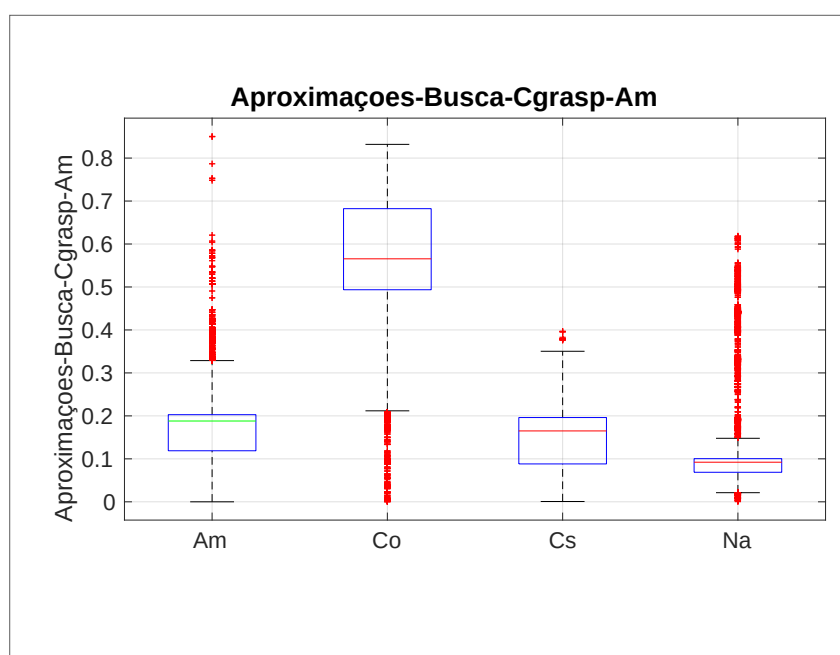
(a) Frações dos isótopos inseridos no espectro total, instância 1. O *Am* é representado por pontos $\circ$; O *Co* é indicado por pontos $+$; O *Cs* tem representação nos pontos $*$; O *Na* é indicado por pontos $\star$. Resultados para o passo de busca local.



(b) Medidas da área abaixo da curva do espectro total. Área calculada são representados por pontos $\circ$; área real são representados por pontos $+$ e o resíduo são representados por pontos $*$. Resultados para o passo de busca local.



(c) Distribuição das determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP. Passo de busca local.



Fonte: Próprio autor

5.2 RESULTADOS PARA OS COMPOSTOS.

As seções seguintes apresentarão os resultados para a Base de Dados 2. Essa base de dados, de acordo com a Tabela 3, tem suas instâncias numeradas pelos valores 5, 6 e 7. Os dados de entrada para a segunda Base de Dados apresentam as mesmas características que os dados de entrada para a Base de Dados 1: vetores representam as bibliotecas de entrada, e os canais são representados pela valor da linha considerada. Os resultados seguintes, em sua maioria, são para a instância 5. Por uma questão de semelhança nos resultados, as demais instâncias são apresentadas no apêndice. Como as duas bases de dados são diferentes na natureza de sua produção, ao menos a instância 5 será apresentada na seção principal dos resultados.

As descrições das seções que se seguem levam em conta a mesma descrição utilizada na abertura da Seção 5: Seção 5.2.1 traz os resultados para determinação dos picos; a determinação da partição de maior resíduo encontra-se na Seção 5.2.2; determinação das frações em peso no particionamento dos dados está presente na Seção 5.2.3; as frações de contagem percentuais na aproximação de uma das bibliotecas está presente na Seção 5.2.4; as estimativas para o elemento omitido está na Seção 5.2.5; condicionamento das matrizes de covariância da aproximação de uma biblioteca está na Seção 5.2.6; resultados dos passos básicos do GRASP está em 5.2.7; por fim, os resultados do CGRASP está em 5.2.8.

5.2.1 Determinação dos picos

Na análise dos dados para determinação dos picos para a Base 2 encontram-se os seguintes resultados:

Tabela 13 – Resultados na determinação dos picos para a segunda base de dados.

Instância	Elemento Omitido	Elemento Fornecido	Canal do Pico
5	Ar	Água	452
		Óleo	137
6	Água	Ar	34
		Óleo	452
7	Óleo	Ar	34
		Água	452

Fonte: Próprio Autor

Os valores da Tabela 13 determinam os picos das bibliotecas fornecidas inicialmente. Os

valores dos picos, devido à ruídosidade da base e consequente ausência do gráfico das bibliotecas, serão comparados com os valores das partições que encontram maior resíduo, na Seção 5.2.2. Importante notar como na ausência do Ar, o pico determinado para a Água e o Óleo foram os canais 452 e 137, respectivamente. Mas na ausência da Água, o Óleo foi definido tendo seu pico no canal 452. Esse fato ocorre porque a Água e o Óleo tem seu pico no mesmo canal, mas o algoritmo foi programado para não realizar essa determinação no intuito de não gerar dependência linear dos dados.

5.2.2 Determinação dos Resíduos das Partições

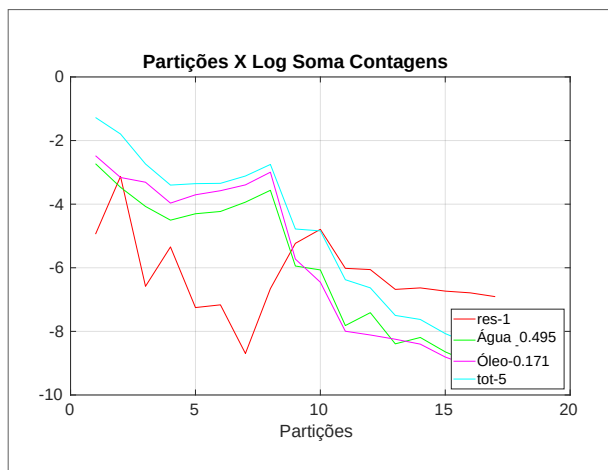
Na presente seção, são apresentados os componentes do vetor de compostos, resíduos e componentes do vetor espectral total como função das partições, para as instâncias 5 (Figura 48a) e 6 (Figura 48b). A instância 7 tem resultado descrito na Seção D.2.

A Figura 48a determina a partição 1 como a de maior resíduo, enquanto que as Figura 48b prediz a partição de maior resíduo aproximadamente compreendida entre as partições 7 e 8.

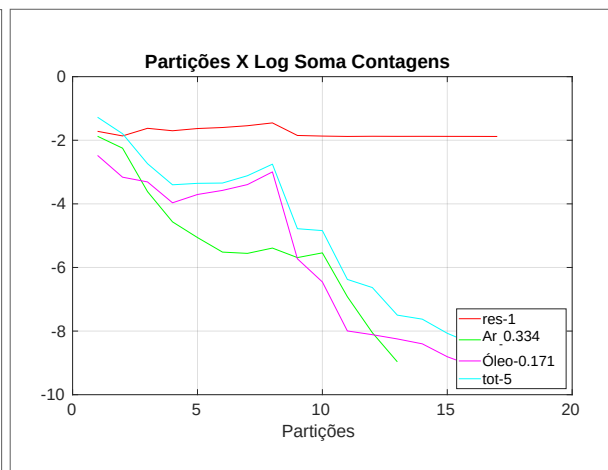
Ao comparar com os resultados da seção anterior, a partição de maior resíduo para o Ar está realmente no pico desse composto. O mesmo pode ser afirmado para a partição de maior resíduo da Água, partição essa compreendida na região do pico desse elemento, determinado no canal 452.

Figura 48 – Componentes dos vetores de compostos (Ar, Água e Óleo), resíduo e componentes do vetor espectral total, como função de partição. Os componentes do vetor espectral total são apresentados (variável $tot - 5$), além dos valores de resíduos nas partições (variável $res - 1$).

(a) Componentes dos vetores de compostos Água e Óleo como função da partição. Os componentes do vetor espectral total são apresentados (variável $tot-5$), além dos valores de resíduos nas partições (variável $res-1$)



(b) Componentes dos vetores de compostos Ar e Óleo como função da partição. Os componentes do vetor espectral total são apresentados (variável $tot-5$), além dos valores de resíduos nas partições (variável $res-1$)



Fonte: Próprio Autor

5.2.3 Determinação de Frações Percentuais - Particionamento

O objetivo da presente seção é o mesmo realizado para a Seção 5.1.3. O particionamento dos dados servirá como um guia para indicar valores de resíduo e condicionamento possíveis de serem atingidos na aproximação das frações em peso, quando uma das bibliotecas estiver sendo aproximada, na seção seguinte, Seção 5.2.4.

A Figura 49 traz frações de contagem percentual no particionamento da segunda base de dados. Os resultados levam em conta a mesma metodologia usada na Seção 5.1.3. Para esse caso, são utilizadas matrizes de covariância de ordem 3.

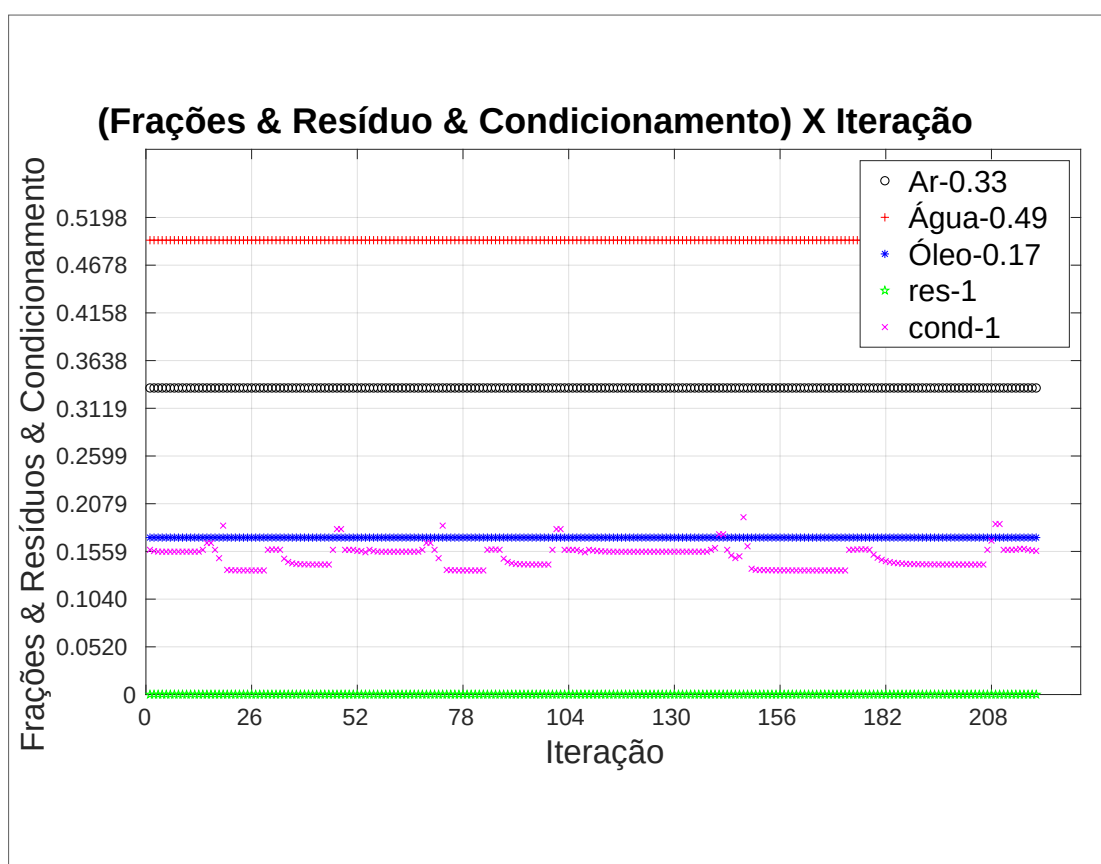
Por questões de simplificação, os resultados da presente seção são aplicados para a instância 5. As outras instâncias podem ser conferidas na Seção E.2, Apêndice E.

Como descrito na Seção 5.1.3, os valores de frações definidos pelos mínimos quadrados no particionamento dos dados descrevem com exatidão os valores nominais de frações. Essa evidência é importante no sentido de esses valores de frações serem acompanhados por resíduos mínimos, que irá servir de parâmetro quando uma das bibliotecas estiver sendo aproximada.

Fato a ser destacado também a partir do gráfico 49: os valores de condicionamento, a medida que as iterações avançam, tendem a reduzir, diferentemente do que foi observado

na Seção 5.1.3, Figura 21. As matrizes de covariância de dimensão três, logo, apresentam evidências de menor instabilidades numéricas que as matrizes de dimensão quatro. A redução no condicionamento, lembrando, é um indício de menores instabilidades numéricas associadas ao sistema de equações, implicando em aproximações para as frações em peso de melhor qualidade.

Figura 49 – Frações percentuais dos compostos após particionamento dos dados. As partições de referências são anexadas à partição de maior resíduo no intuito definir a matriz de covariância. Pontos em preto ($\circ$, Ar), vermelho($+$, Água) e azul ($*$, Óleo) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais de frações de cada composto encontram-se na legenda do gráfico. Os valores dos resíduos são descritos nos pontos em verde ($\star$), e o rosa ($\times$) indica o condicionamento normalizado das matrizes de covariância do particionamento.



Fonte: Próprio autor

Aqui não são trazidas as distribuições dos valores de resíduo na determinação quantitativa dos compostos por serem muito próximos a zero, em todos os possíveis casos a serem estudados.

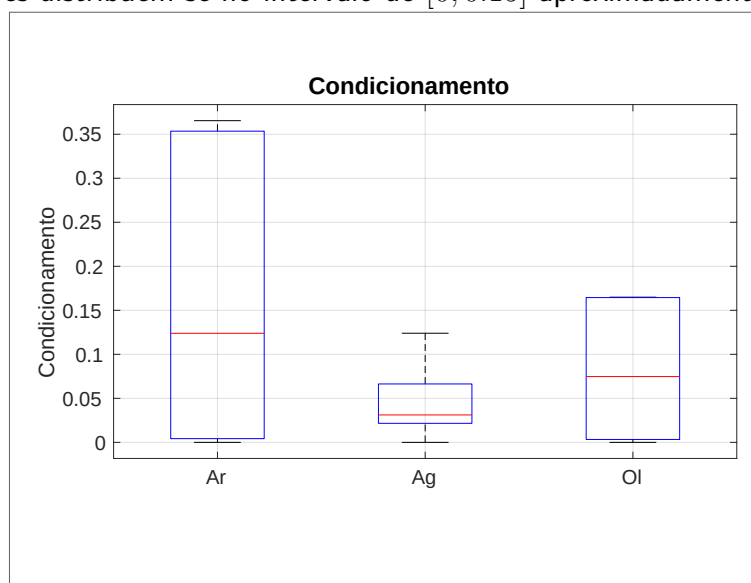
Os valores para condicionamento das matrizes de covariância estão demonstrados na Figura 50. As distribuições dos valores de condicionamento para o caso das matrizes de covariância de dimensão três são significativamente menores que os valores de condicionamento definidos para as matrizes de dimensão quatro. Isso conduz à constatação de que dimensionalidade do

problema também é um dos fatores responsáveis por introduzir instabilidades numéricas ao problema.

Figura 50 – ,

com os dados distribuídos positivamente assimétricos. No ultimo caso, os valores distribuem-se no intervalo de 0 e 0.15] aproximadamente de forma simétrica.]Plot comparando as distribuições nos valores de condicionamento para as instâncias 5, 6 e 7. Na ausência do Ar, os valores de condicionamento estão amplamente distribuídos no intervalo de $[0, 0.35]$ com os dados distribuídos assimetricamente positivo. Para a instância da Água, os valores estão distribuídos no intervalo de $[0, 0.07]$, com os dados distribuídos positivamente assimétricos. No ultimo caso, os valores distribuem-se no intervalo de $[0, 0.15]$ aproximadamente de forma

simétrica.



Fonte: Próprio autor

5.2.4 Determinação da Fração Percentual dos Compostos por Aproximações

Na presente seção, serão apresentados os resultados para as frações em peso dos compostos na omissão de uma biblioteca do espectro total. Os resultados farão referência à instância cinco. As demais instâncias da Base 2, por questões de similaridades com os resultados da instância 5, serão apresentadas no Apêndice F, Seção F.2

A interpretação para a Tabela 14 é a mesma contida na Tabela 7. São apresentados os valores médios de frações de contagem percentual para os compostos estudados, como também os valores médios de resíduos e condicionamento, em trinta execuções do algoritmo.

A Tabela 15 apresenta os valores de frações de contagem percentual, mediante a omissão de uma biblioteca, na minimização do resíduo. Já a Tabela 16 traz as mesmas saídas que a Tabela 15, mas descritas como função do menor valor de resíduo, nas ocorrências de minimização do resíduo, em trinta execuções do nosso algoritmo.

Possível observar, mais uma vez, como descrito na Seção 5.1.4, que a minimização dos valores de resíduos é uma métrica adequada para determinar, com boa precisão, os valores de frações de contagem percentual. Os valores médios das trinta execuções, apesar de, em algumas instâncias, ter encontrado frações que são apropriadas, e melhores que os valores descritos na mesma tabela da Base 1, ainda assim não são melhores que as frações encontradas na Tabela 15. A comparação entre as Tabelas 14 e a Tabela 7 leva-nos a inferir que as tendências de imprecisões aumentam, à medida que aumenta a dimensionalidade do problema. Essa mesma observação pode ser estendida para a Tabela 16, comparando-a com a Tabela 9. Os valores de condicionamento apresentam tendências a minimizarem, à medida que a Base 2 apresenta uma dimensionalidade menor que as matrizes previstas para a Base de Dados 1, comparando-se as Tabelas 16 e 9.

Tabela 14 – Valores médios de frações de contagens percentuais na determinação quantitativa de compostos inseridos no espectro total. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o Ar da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas.

	Ar	PreAr	Ag	PreAg	OI	PreOI	Resíduo	Condic.
1	0.32646	0.010725	0.48394	0.012377	0.18967	0.11569	0.0025337	5.1814
2	0.34321	0.040025	0.4192	0.1445	0.23746	0.39681	0.00049819	4.7115
3	0.32637	0.010985	0.59345	0.21112	0.080177	0.52837	0.00028644	5.9248

Fonte: Próprio Autor

Tabela 15 – Valores de frações de contagens percentuais na determinação quantitativa de compostos inseridos em um espectro total, tomando por métrica a minimização resíduo. A linha 1 exibe o Ar na descrição do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas: linha 2, omite-se o a Água; linha 3, omite-se o Óleo.

	Ar	PreAr	Ag	PreAg	OI	PreOI	Resíduo
1	0.33315	0.0095539	0.49512	0.010442	0.17173	0.010179	1.8451e-06
2	0.33367	0.011135	0.49606	0.01236	0.17009	0.00052588	8.7566e-05
3	0.3339	0.011828	0.49565	0.011529	0.17045	0.0026312	1.0784e-06

Fonte: Próprio Autor

Tabela 16 – Valores de percentuais aproximados na determinação quantitativa de compostos inseridos em um espectro total no menor condicionamento nas ocorrências de minimização do resíduo. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, omitindo o Ar da geração do espectro total e o aproximando por funções constantes nas diferentes partições. A mesma descrição é válida para as demais linhas.

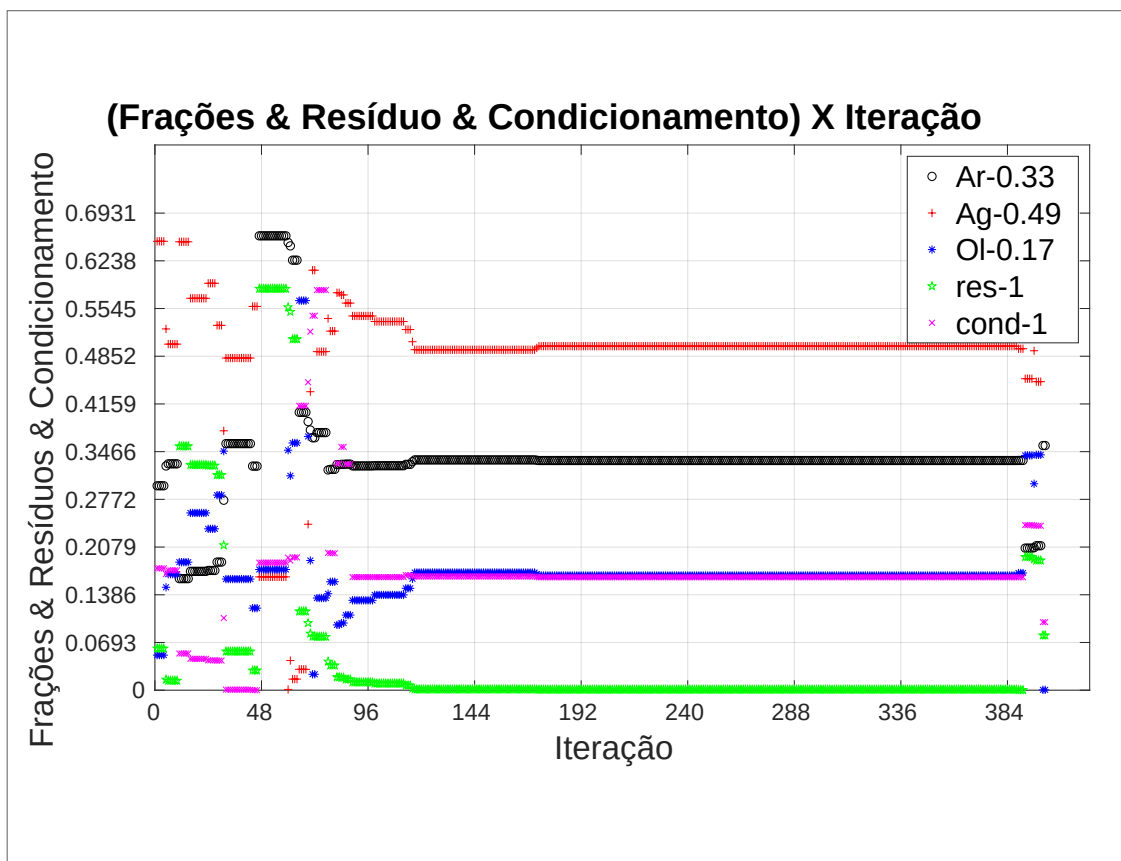
	Ar	PreAr	Ag	PreAg	OI	PreOI	Condic.
1	0.32553	0.013532	0.55754	0.13784	0.11921	0.29875	2.5671
2	0.42947	0.30142	0.0042574	0.99131	0.56624	2.3308	1.96334
3	0.30474	0.076545	0.69533	0.41904	7.216e-06	0.99996	2.54550

Fonte: Próprio Autor

Na Figura 51 são apresentados os resultados das contagens de frações percentuais para a omissão da biblioteca do Ar. Os resultados para a omissão das bibliotecas da Água (Figura 68a) e do Óleo (68b) estão descritos no Apêndice F, Seção F.2.

Importante considerar regiões dos Gráficos 51 em que, à medida que os valores de resíduo são reduzidos, os valores de frações de contagem percentual convergem para os valores nominais (região compreendida a partir da iteração 100 do Gráfico 51. Os resultados para a Base de Dados 2 apresentaram uma facilidade de convergência dos valores de frações, aos valores nominais, mais rápido que os valores descritos para a Base de Dados 1.

Figura 51 – Frações percentuais dos compostos após supressão e estimação do Ar no espectro total, instância 5. Valores em preto (○), vermelho(+) e azul(*) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais estão descritos na legenda. Os valores dos resíduos são descritos nos pontos ☆, e o ponto × indica o condicionamento normalizado $([0, 1])$ das matrizes de covariância na aproximação das bibliotecas.



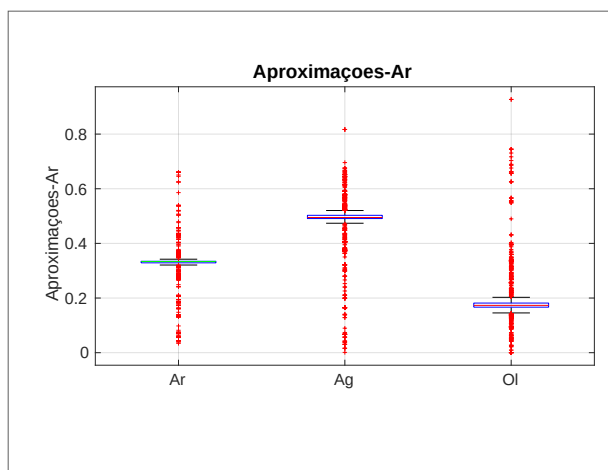
Fonte: Próprio autor

Na Figura 52 têm-se as distribuições das frações de contagem percentuais para as instâncias da Base 2. As amplitudes das distribuições mostram-se próximas dos valores nominais nos casos 52a e 52b, mas foi identificada uma amplitude maior no caso da apresentada na Figura 52c para a omissão do Óleo. Mas importante destacar, apesar de a instância sete ter apresentado uma variabilidade nas saídas maior que as instâncias cinco e seis, ainda assim, na minimização dos resíduos, o algoritmo apresentado foi capaz de prever, com boa qualidade de precisão, os valores de frações de contagem percentual das bibliotecas estudadas.

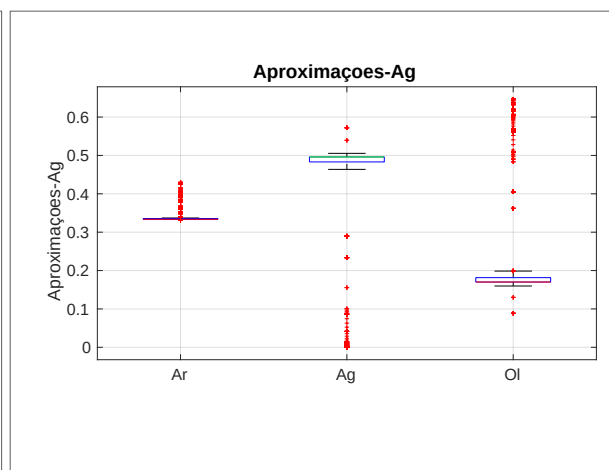
A distribuição dos valores dos resíduos, para cada uma das instâncias da Base de Dados 2, está descrita na Figura 53. Importante perceber como esses valores estão compreendidos em intervalos consideravelmente maiores que os descritos pela Figura 25, o que leva, mais uma vez, à indicação de que a dimensionalidade do problema é preponderante no sentido de aumentar as incertezas entre o espectro real e o espectro estimado na determinação das frações em peso

Figura 52 – Plots contendo os intervalos de frações de contagens percentuais após a omissão das bibliotecas de energia. As medianas em verde indicam a biblioteca que está sendo estimada. 50% dos valores, para as frações de contagens percentuais, nas Figuras 52a e 52b, estão contidos em um intervalo de pequena amplitude. Existe uma maior variabilidade desses valores na Figura 52c. A convergência dos valores preditos, aos valores nominais das frações de contagem percentual, é melhor nas Figuras 52a e 52b do que na Figura 52c.

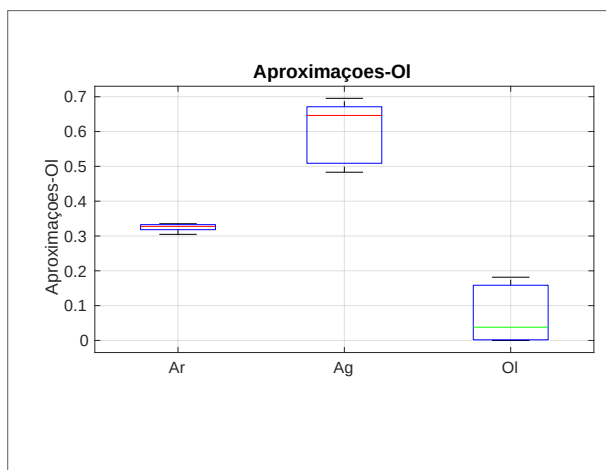
(a) Frações percentuais dos compostos após supressão do Ar do espectro total.



(b) Frações percentuais dos compostos após supressão da Água do espectro total.



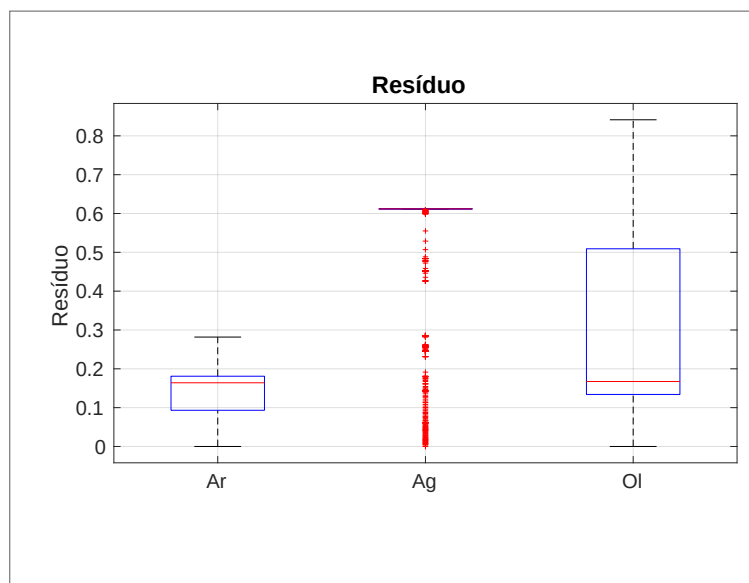
(c) Frações percentuais dos compostos após supressão do Óleo do espectro total.



Fonte: Próprio Autor

das bibliotecas.

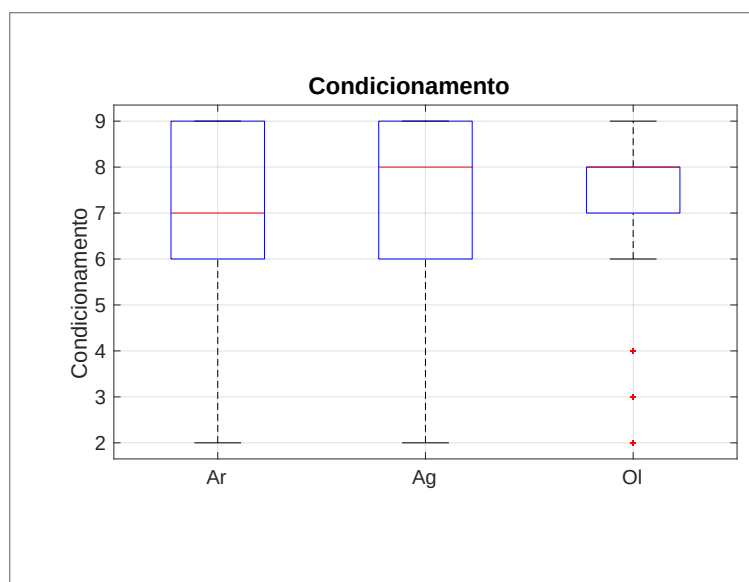
Figura 53 – Intervalo de valores para o resíduo durante 30 execuções do algoritmo. Na ausência e aproximação do Ar, a ordem de grandeza da mediana dos valores é de aproximadamente 1.8×10^{-1} , sendo que 50% dos valores de resíduo encontram-se compreendidos no intervalo 10^{-1} e 2×10^{-1} . Na ausência da Água, a mediana dos valores é de aproximadamente 6×10^{-1} e 50% dos valores de resíduos estão compreendidos entre 6×10^{-1} e 5×10^{-1} , e, para o Óleo, o intervalo é de 1×10^{-1} e 5×10^{-1} . A amplitude da distribuição do resíduo é maior na instância da ausência do Óleo, do que as demais bibliotecas.



Fonte: Próprio Autor

Importante observar que as distribuições dos números de condicionamento presentes na Figura 54, para a Base de Dados 2, apresentam valores médios maiores que os valores descritos na Figura 26 para a Base de Dados 1. Esse fato pode encontrar explicação na natureza da segunda base de dados, que é mais ruidosa, e menos suave, que a Base de Dados 1.

Figura 54 – Plots demonstrando o intervalo de valores para o condicionamento durante 30 execuções do GRASP para a Segunda base de dados. Na ausência do Ar, 50% desses valores encontram-se compreendidos entre 6 e 9. A mediana para a ausência do Ar é positivamente assimétrica. O plot para a ausência do Água e do Óleo mostram dados negativamente assimétricos.



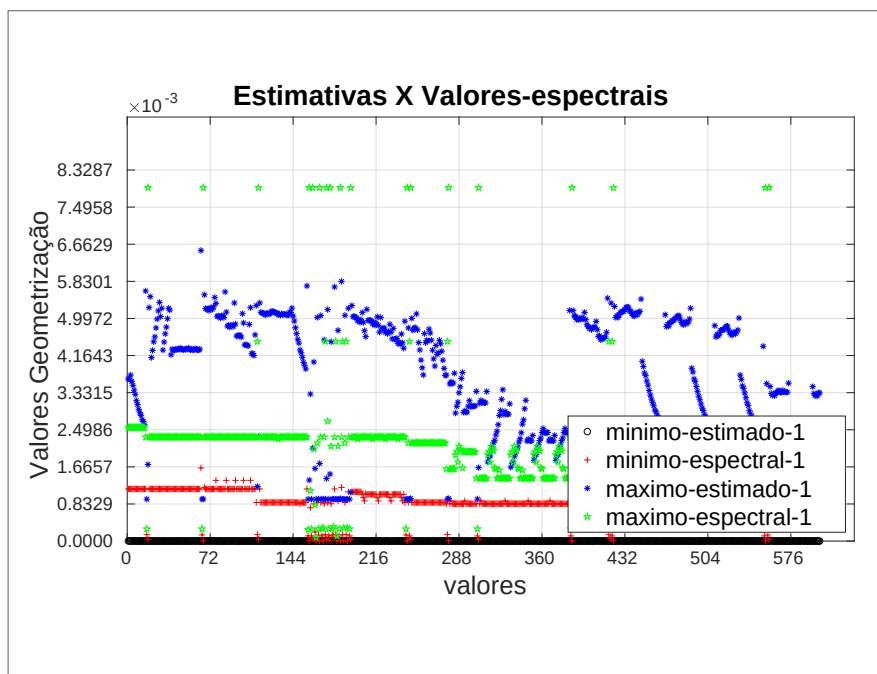
Fonte: Próprio Autor

5.2.5 Aproximação dos Compostos - Estimativa das alturas nas partições

As estimativas da biblioteca omitida e as contagens espectrais totais reais serão mais uma vez comparadas na presente seção, mas dessa vez para a Base 2. Os resultados para a instância cinco são descritos na presente seção. As instâncias seis e sete têm os resultados descritos no Apêndice G, Seção G.2.

Os resultados encontrados na Seção 5.1.5 são idênticos aos apresentados na presente seção para a segunda base de dados, instância cinco. Na Figura 55, para os resultados associados à instância cinco, determinam, mais uma vez, que as contagens estimadas são os limitantes inferior e superior para a contagem espectral total nos compartimentos das partições, como era de se esperar para o espectro total definido como combinação linear das bibliotecas. Interessante observar que os valores estimados máximos tendem a convergir em direção aos valores de contagem espectral total máxima, em algumas regiões do Gráfico 55. Esse fato encontra explicação dentro da natureza da Base de Dados 2, que apresenta características de ruídos mais evidente que a Base de Dados 1. Em algumas regiões dos canais, as contagens espectrais dos demais elementos são muito pequenas, fazendo com que o processo de otimização conduza os valores da biblioteca estimada em direção aos valores de contagem espectral total real.

Figura 55 – Estimativas do Ar, instância 5. Os pontos $\circ$ indicam os valores de menor contagem estimada pelo algoritmo da Seção 4.3 enquanto pontos representados por $+$ indicam a menor contagem espectral total real no intervalo de canais que definem a partição. Pontos $*$ são os valores de maior contagem estimada para a biblioteca omitida e os pontos $\star$ indicam o maior valor real para essa mesma biblioteca.

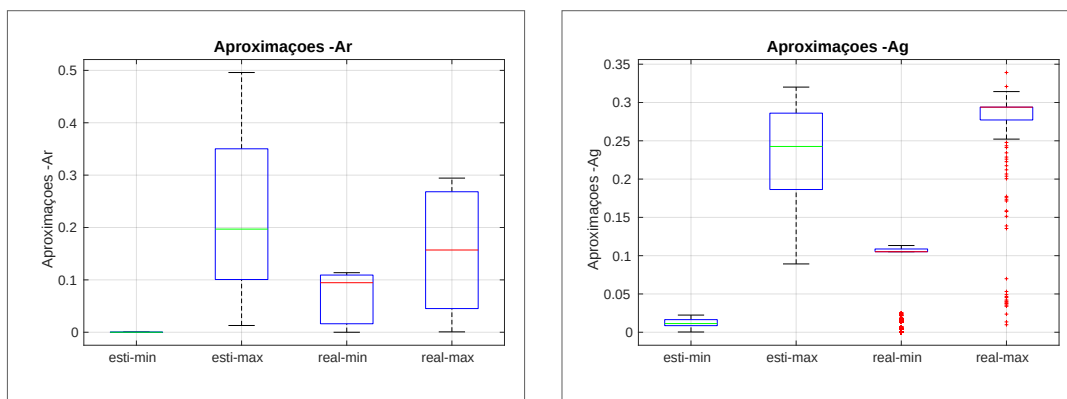


Fonte: Próprio autor

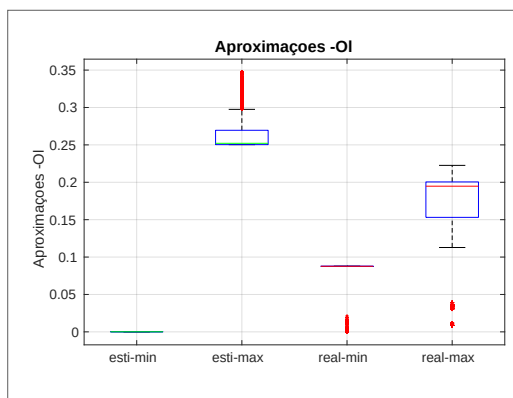
As Figuras 56a (instância 5), 56b (instância 6) e 56c (instância 7), contêm distribuição de valores da estimativa da contagem espectral omitida e compara com os valores de contagem espectral total. As distribuições apresentadas com valores em verde indicam as estimativas de contagem espectral. As distribuições em vermelho indicam as contagens espectrais totais para a instância estudada, dentro dos compartimentos das partições. Na Base de Dados 2, é possível identificar interseção entre as estimativas máximas e os valores de contagem espectral total máxima, contrariando a hipótese de que as estimativas englobariam os valores de contagens espectrais totais reais. Mas esse fato não foi proibitivo para a determinação das frações em peso ter sido bem sucedida. Esse fato pode encontrar justificativas na qualidade da Base 2, que é uma base mais ruidosa que a Base 1.

Figura 56 – Valores de estimativa de altura para o elemento omitido nas instâncias da Base de Dados 2.

- (a) Comparativo global entre as estimativas mínimas e máximas do elemento omitido do espectro e suas respectivas contagens espectrais totais reais mínimas e máximas, para a ausência do Ar.
- (b) Comparativo global entre as estimativas mínimas e máximas do elemento omitido do espectro e suas respectivas contagens espectrais totais reais mínimas e máximas, para a ausência da Água.



- (c) Comparativo global entre as estimativas mínimas e máximas do elemento omitido do espectro e suas respectivas contagens espectrais totais reais mínimas e máximas, para a ausência do Óleo.



Fonte: Próprio Autor

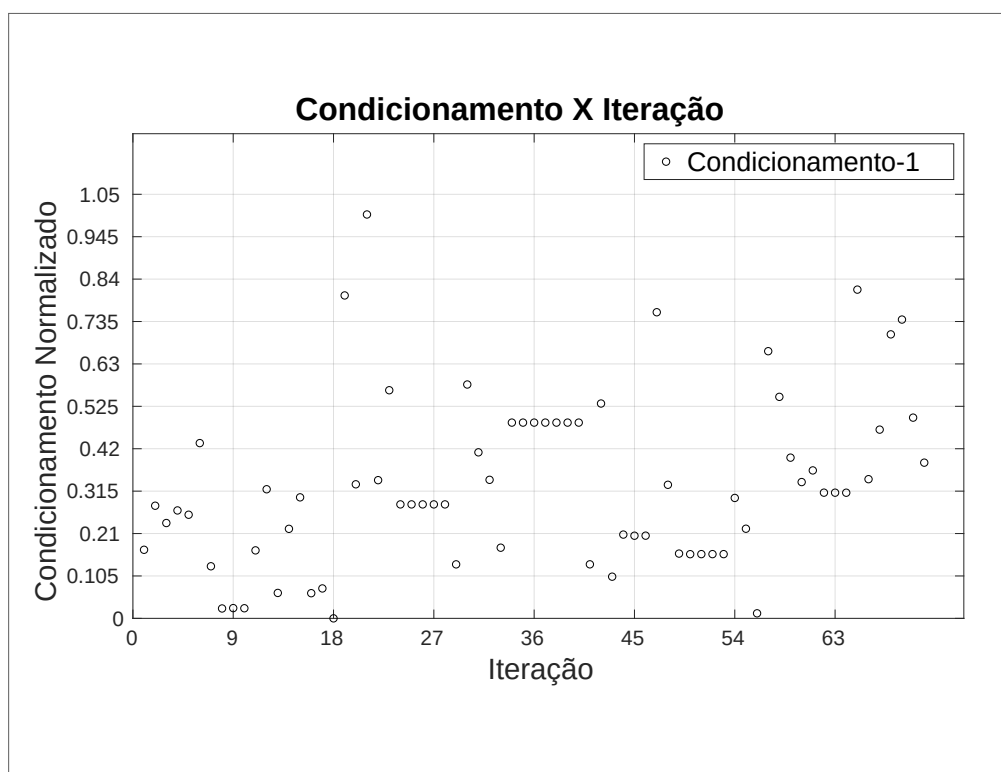
5.2.6 Condicionamento na Minimização da Métrica na aproximação dos compostos.

Na presente seção será, investigado o comportamento do condicionamento na minimização do resíduo. Serão apresentados os resultados para a instância cinco. As demais instâncias da Base 2, por questões de similaridades, serão apresentadas no apêndice H, Seção H.2.

Os Gráficos aqui demonstram um intervalo de valores normalizados, para o condicionamento das matrizes de covariância na minimização da métrica, muito semelhante aos apresentados na Seção 5.1.6. O Gráfico 57 apresenta evidências que os valores de condicionamento tende a se restringir a um intervalo específico, à medida que os valores de resíduo são minimizados.

Esse fato comprova-se pela convergência desses valores a um intervalo, à medida que as iterações avançam, ou seja, mais uma evidência de que a boa qualidade das frações em peso das bibliotecas em estudo é função da diminuição das instabilidades numéricas associada ao sistema de equações.

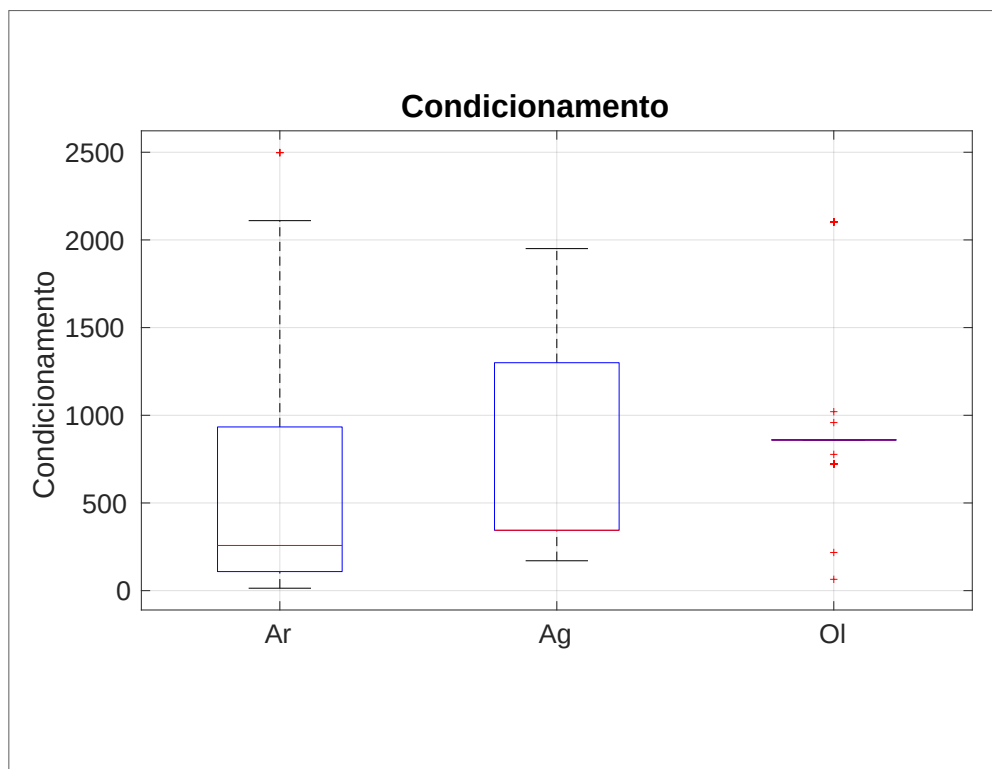
Figura 57 – Condicionamento na minimização da métrica. Na Figura 57, têm-se os valores de condicionamento, na minimização do resíduo, para a ausência e aproximação do Ar, instância 5. Os valores tendem a se estabilizar em um intervalo, evidenciando que a minimização dos resíduos podem ser associados a um decrescimento do condicionamento, mesmo que esse decrescimento não seja monotônico .



Fonte: Próprio autor

Na Figura 58, são demonstradas as distribuição dos verdadeiros valores de condicionamento na determinação das frações de contagem percentual. Os resultados da distribuição de valores para a instâncias 5 e 6 têm pequena amplitude, o que não pode ser conferido nos resultados presenciados para o Óleo. Comparativamente aos valores de condicionamento apresentados no trabalhos de (MERIÇ, 2012) (matrizes de dimensão 2×2), são valores de intervalos próximos, o que é bom, considerando que as matrizes de dimensão 3×3 apresentam tendências de maior condicionamento.

Figura 58 – Boxplot demonstrando a distribuição de valores do condicionamento das matrizes de covariância na determinação das frações de contagem percentual na omissão de uma das bibliotecas da solução inicial. O primeiro plot é considerado para a ausência do Ar, e os demais plots seguem a mesma descrição.



Fonte: Próprio Autor

5.2.7 Resultados dos passos básicos do GRASP

Os resultados dos passos básicos do GRASP para a Base 2, por uma questão de brevidade e semelhança aos resultados descritos para a Base 1 na Seção 5.1.7, estarão descritos no Apêndice I.2.1 para as instâncias 5, 6 e 7. A Seção I.2.1 apresenta os resultados para a etapa construtiva; os resultados para a etapa de busca são descritos na Seção I.2.2; na Seção I.2.3, são descritos os resultados para a etapa de linkagem e, por fim, os resultados para a segunda busca são enumerados na Seção I.2.4.

5.2.8 Resultados do CGRASP

A descrição para os resultados do CGRASP, presente nas Seções 5.2.8.1, 5.2.8.2, 5.2.8.3 e 5.2.8.4, leva em conta a mesma descrição realizada na Seção 5.1.8. Serão alvo de investigação na presente seção as frações de contagem percentual, resíduo e condicionamento, como função da iteração, no contexto de uma execução única do CGRASP; no acumulado das execuções do algoritmo, assim como em cada uma das subetapas do CGRASP.

5.2.8.1 Resultados do CGRASP - Execução Única.

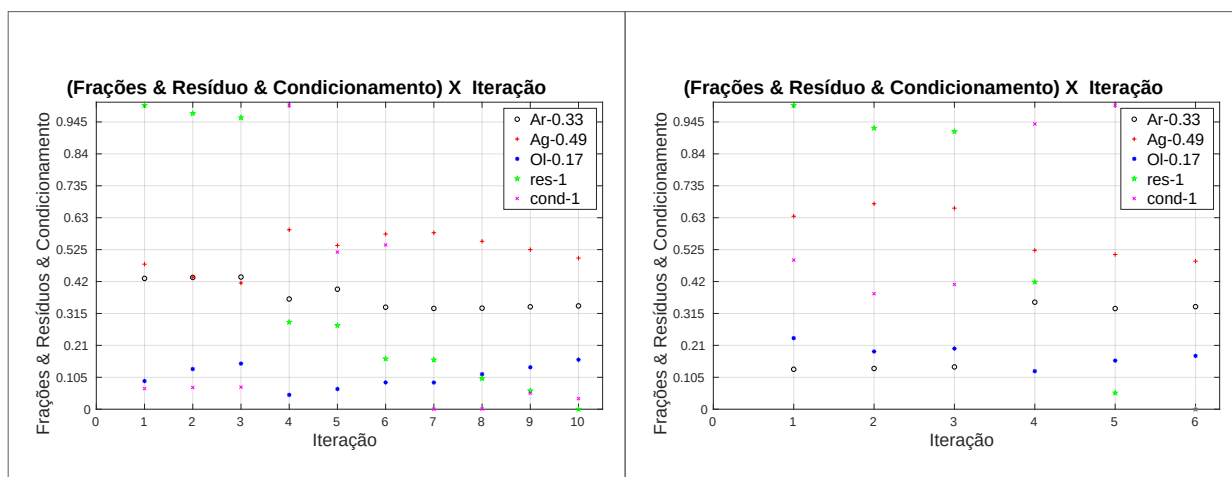
A presente seção terá por objetivo investigar as mesmas questões descritas na Seção 5.1.8.1. As frações de contagem percentual serão analisadas conjuntamente à métrica utilizada. Também será alvo de destaque os valores de condicionamento sendo confrontados com as frações e os valores de resíduo. Os resultados para a instância cinco serão aqui apresentados, enquanto as instâncias seis e sete estarão presentes no Apêndice J, Seção J.2.1.

As Figuras 59a, 59b trazem os resultados da fração de contagem percentual, resíduo e condicionamento normalizados, como função da iteração, para a instâncias 5. Os resultados observados são similares aos observados para a instância um, primeira base de dados, Seção 5.1.8.1. À medida que os valores de resíduos são reduzidos, os valores das frações de contagem percentual convergem aos valores nominais, descritos na legenda dos Gráficos 59a e 59b. Os valores de condicionamento, mais uma vez, não são minimizados monotonicamente, mas apresentam evidências do valor do condicionamento estar contido em um intervalo específico na aproximação das frações aos valores nominais. Os resultados para a Base 2, em consonância com os resultados da Base 1, ratificam a abordagem das aproximações das frações em peso pela minimização dos resíduos.

Figura 59 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 59a e 59b são representados duas execuções arbitrárias para a instância 5. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. Os pontos $\circ$ indicam as aproximações para o Ar; $+$ são pontos para as aproximações para a Água; os pontos $\ast$ indicam as aproximações para o Óleo; os valores de resíduo são representados por pontos $\star$ e o símbolo $\times$ indica os valores para condicionamento.

(a) Percentuais dos elementos e respectivas métricas. Instância 5.

(b) Percentuais dos elementos e respectivas métricas. Instância 5.



Fonte: Próprio Autor

5.2.8.2 Resultados do CGRASP - Execuções Acumuladas / Valores Métricas

Os resultados apresentados na presente seção, para a Base 2, obedecem à mesma sequência e descrição dos resultados apresentados na Seção 5.1.8.2 para a Base 1. As métricas de avaliação dos resultados serão mais uma vez comparadas, e serão observados as execuções acumuladas para o Algoritmo CGRASP. As distribuições dos valores de resíduo e condicionamento encerram o detalhamento dos resultados na seção atual.

A Tabela 17 contém os valores médios de frações de contagem percentual nas iterações do Algoritmo CGRASP. Mais uma vez, da mesma maneira que a apresentada para a Tabela 14, os valores médios em trinta execuções do algoritmo não apresentam resultados satisfatórios para as frações em peso dos constituintes da amostra, apesar de os resultados apresentados possuírem mais qualidade que os determinados para a primeira base de dados, apresentando evidências que a dimensionalidade do problema pode ser um fator limitante à abordagem apresentada.

Tabela 17 – Valores médios de frações de contagens percentuais dos compostos inseridos no espectro total. A linha 1 contém os valores de frações percentuais dos compostos, na instância 5. Os resultados para as instâncias 6 e 7 são descritos nas demais linhas.

	Ar	PreAg	Ag	PreAg	OI	PreOI	Resíduo	Condicionamento
1	0.30795	0.066827	0.48467	0.010879	0.20744	0.22024	0.0045606	5.2904
2	0.34157	0.035048	0.41232	0.15854	0.24588	0.44635	0.00054274	5.1936
3	0.32705	0.0089262	0.57238	0.16812	0.10652	0.37341	0.00021191	5.8356

Fonte: Próprio Autor

A Tabela 18 apresenta as frações de contagem percentual na minimização do resíduo, em trinta execuções do algoritmo. Fica evidente, mais uma vez, que as minimizações do resíduo conduzem as aproximações dos mínimos quadrados aos valores nominais de frações em peso dos compostos estudados. Os valores de precisão para as frações de contagem percentual, determinados pelas colunas *PreAr*, *PreAg* e *PreOI*, são muito próximos de zero, indicando, assim, a boa qualidade das aproximações.

Os menores valores de condicionamento, para cada uma das instâncias, nas ocorrências de minimização da métrica, e os correspondentes valores de frações de contagem percentual, são descritos na Tabela 19. As instâncias apresentarem valores de frações de contagem percentual de melhor qualidade do que os apresentados para a Base 1. Os valores, quando comparados com a Tabela 18, não apresentam melhor qualidade na minimização do condicionamento do que na minimização do resíduo.

Tabela 18 – Valores de frações em peso na minimização do resíduo. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos para a instância 5. A mesma descrição é válida para as demais linhas: linha 2, omite-se a Água; linha 3, omite-se o Óleo.

	Ar	PreAr	Ag	PreAg	OI	PreOI	Resíduo
1	0.33315	0.0095539	0.49512	0.010442	0.17173	0.010179	1.8451e-06
2	0.33317	0.0096194	0.49406	0.0082786	0.17309	0.018173	8.7366e-05
3	0.33052	0.0015606	0.49068	0.0013829	0.1706	0.0035553	2.5205e-05

Fonte: Próprio Autor

Tabela 19 – Valores de percentuais aproximados na para as frações de contagem percentuais tomando por métrica a minimização do condicionamento. A linha 1 da tabela determina os valores de frações percentuais dos diferentes compostos, para a instância 5. A mesma descrição é válida para as demais linhas: linha 2, instância 6; linha 3, instância 7.

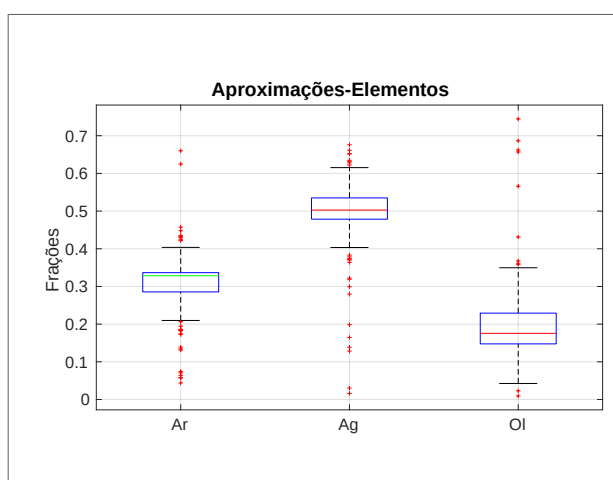
	Ar	PreAr	Ag	PreAg	OI	PreOI	Condicionamento
1	0.32553	0.013532	0.55754	0.13784	0.11921	0.29875	2.6552
2	0.38204	0.15768	0.00018298	0.99963	0.61731	2.6312	2.1581
3	0.31129	0.056694	0.68797	0.40401	0.074552	0.56146	3.0956

Fonte: Próprio Autor

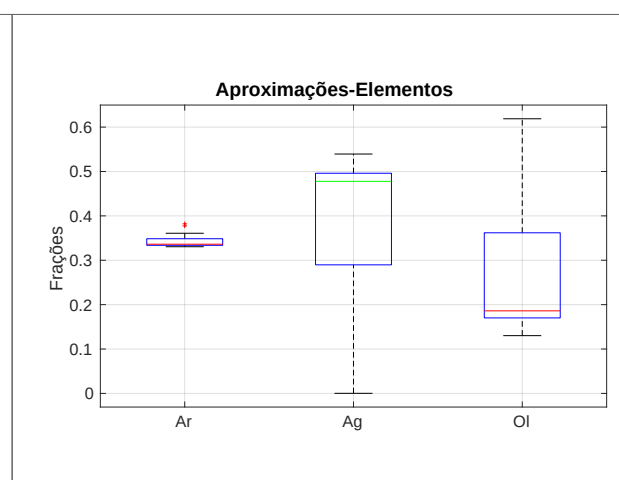
Na Figura 60, são apresentadas as distribuições de frações de contagem percentual para as possíveis instâncias da Base 2. A instância cinco apresentou as distribuições melhor centradas em torno dos valores de fração nominais, enquanto que as instâncias seis e sete mostraram maior amplitude comparativamente à instância cinco. Mas mesmo evidenciado-se uma maior amplitude para as instância seis e sete, a minimização do resíduo foi capaz de determinar uma boa qualidade nas frações de contagem percentual dos compostos inseridos na amostra.

Figura 60 – Valores de fração de contagem percentual nas iterações do CGRASP, Base de Dados 2. Na Figura 60a estão as distribuições para fração de contagem percentual dos compostos na omissão do Ar, instância cinco. Figura 60b, omissão da Água, instância seis, e Figura 60c, omissão do Óleo, instância sete.

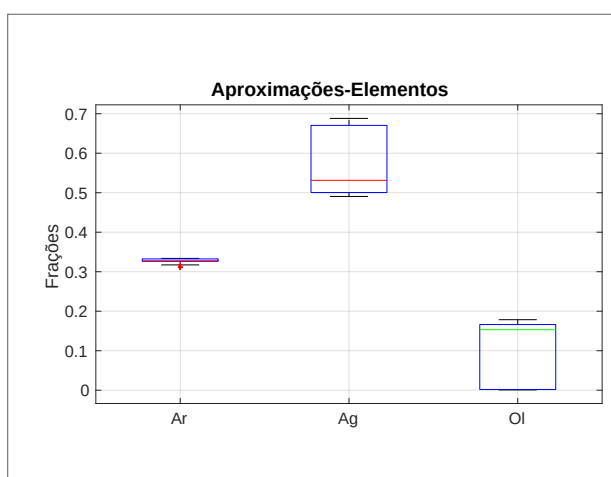
(a) Distribuição das Frações de contagem percentual nas iterações do CGRASP, instância cinco, mediante ausência e estimação da biblioteca do Ar.



(b) Distribuição das Frações de contagem percentual nas iterações do CGRASP, instância seis, mediante ausência e estimação da biblioteca da Água.



(c) Distribuição das Frações de contagem percentual nas iterações do CGRASP, instância sete, mediante ausência e estimação da biblioteca do Óleo.



Fonte: Próprio Autor

5.2.8.3 Resultados do CGrasp - Execuções Acumuladas - Etapa Construtiva.

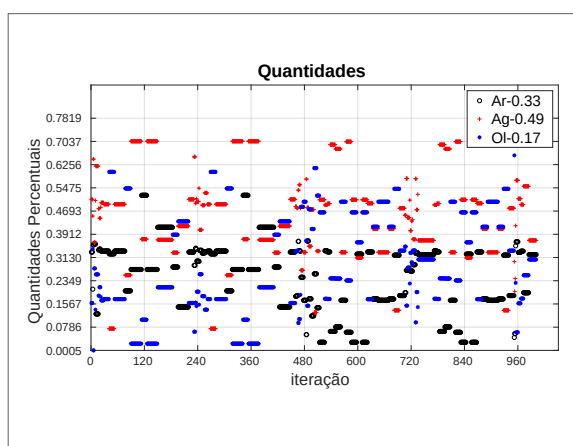
A presente seção realiza a mesma descrição que a apresentada na Seção 5.1.8.3. Serão apresentados os fatos observados para a instância cinco, sendo que as demais instâncias para a Base 2 estão apresentadas no Apêndice J, Seção J.2.2, por questões de similaridades nos resultados.

O passo construtivo para a instância cinco define regiões em que o resíduo é reduzido, como demonstra a Figura 61b, e que podem ser associados a regiões em que as frações de contagem percentual convergem aos valores nominais, como pode ser observado ao comparar a Figura 61b com a Figura 61a. A diferença para os resultados da presente seção, com os resultados da Seção 5.1.8.3, é que as aproximações para a segunda base de dados na fase construtiva do CGRASP são mais difíceis: é possível localizar regiões em que os resíduos são baixos mas, mesmo assim, não existe qualidade nas aproximações das frações em peso. A constatação desse fato nos resultados da Base 1 são baixos, comparados aos resultados demonstrados pela Base 2. Mas esse fato não foi impeditivo para uma boa qualidade global do algoritmo na determinação das frações em peso na minimização dos resíduos.

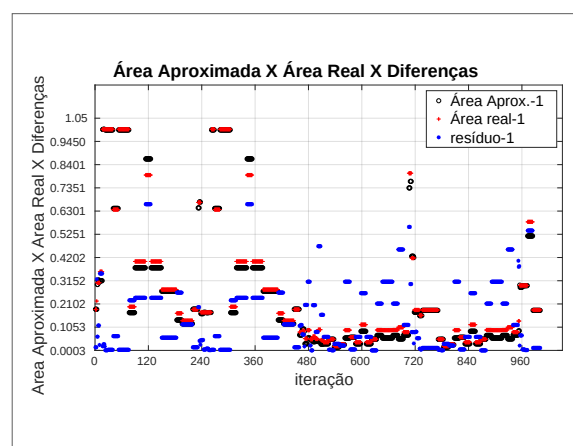
As distribuições das frações em peso para bibliotecas da Base 2, descritas na Figura 61c, apresentam-se centradas em torno dos valores das frações em peso nominais, demonstrando, assim, a capacidade em determinar uma boa qualidade para as frações ainda no passo construtivo. Essa capacidade em determinar boas qualidades de soluções no passo construtivo deve-se ao fato de esse estágio ser dotado de um mecanismo de pesquisa ao longo das direções possíveis. Esta pesquisa, ao longo dessa execução, pode encontrar soluções de boa qualidade.

Figura 61 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, têm-se os valores quantitativos dos isótopos que constituem o espectro total. Na Figura 157b, tem-se a seguinte descrição: Área estimada representada pelos pontos (○); área real representada pelos pontos (+); resíduo representados pelos pontos (*). A evolução demonstra a métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c têm-se os valores de estimativa da biblioteca omitida (instância 5, omissão do Ar) e a aproximação dos demais compostos.

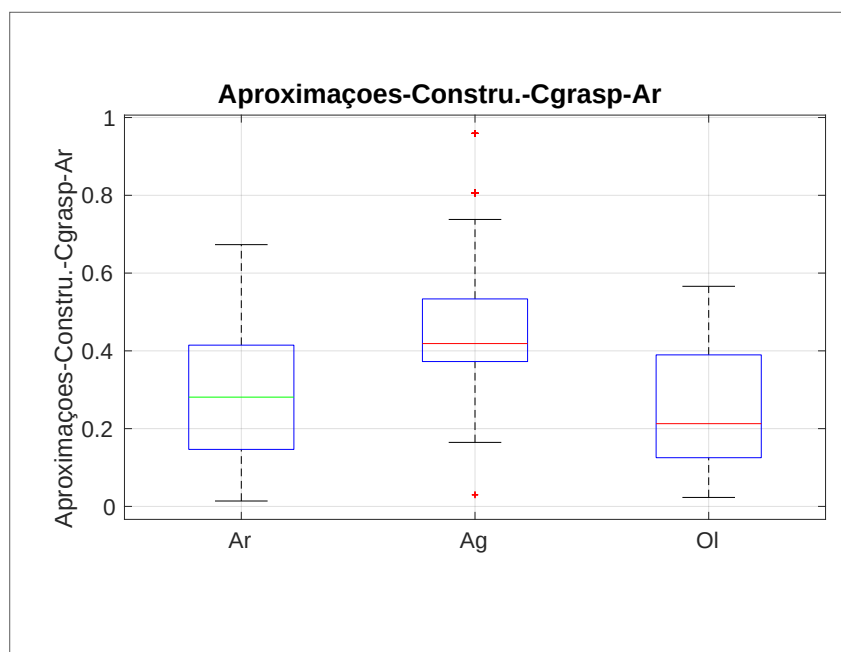
(a) Valores das quantidades dos isótopos inseridos no espectro total na instância 5 (pontos em ○ representam o Ar; pontos em + representam a Água; pontos em * representam o Óleo.)



(b) Medidas da área abaixo da curva do espectro total. Área estimada representada pelos pontos ○; área real representada pelos pontos +; os pontos * representam o resíduo.



(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP. instância 5.



Fonte: Próprio Autor

5.2.8.4 Resultados do CGRASP - Etapa de Busca.

A etapa de busca determina os resultados descritos nas Figuras 62. Por uma questão de simplificação, as Figuras 158 e 159 são apresentadas no Apêndice J, Seção J.2.3, instâncias Água e Óleo, respectivamente. A interpretação para os gráficos acima mencionados é a mesma que as realizadas na Seção 5.2.8.3 para a etapa construtiva.

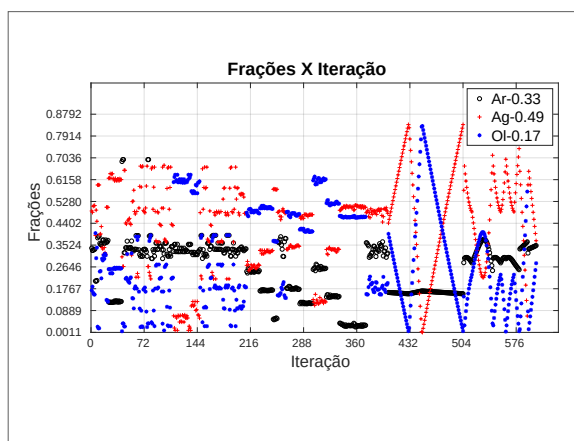
A comparação entre as Figuras 62a e 62b, para a omissão da biblioteca do Ar, trazem os mesmos resultados descritos nas Seções 5.1.8.4 e 5.2.8.3. Os valores de métrica, ao serem minimizados, implicam na possível na convergência das frações de contagem percentual aos valores nominais.

Existem regiões em que os valores de resíduos são mínimos, como pode ser percebido nas imediações posteriores a da iteração 360, em que os valores das saídas das frações de contagem percentual permanecem concentradas ao redor dose seus verdadeiros valores.

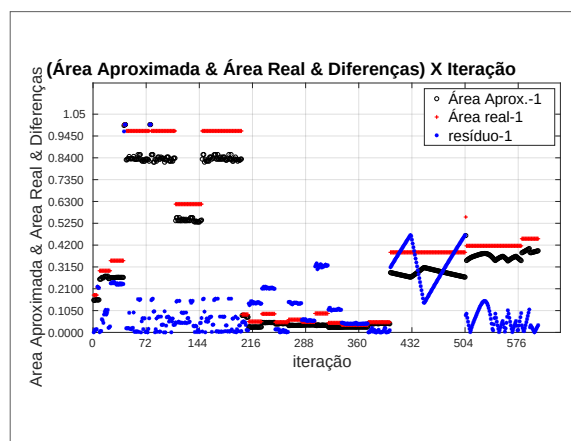
Para o caso da Base 2, evidenciou-se de uma maneira expressiva como o Algoritmo de Busca consegue ser eficaz no sentido determinar valores de frações de contagem percentual de boa qualidade a partir das soluções vindas do passo construtivo. Essas evidências podem ser conferidas a partir da comparação entre os gráficos 61c e 62c.

Figura 62 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, têm-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○);(Água +);(Óleo *); na Figura 157btem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c têm-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos.

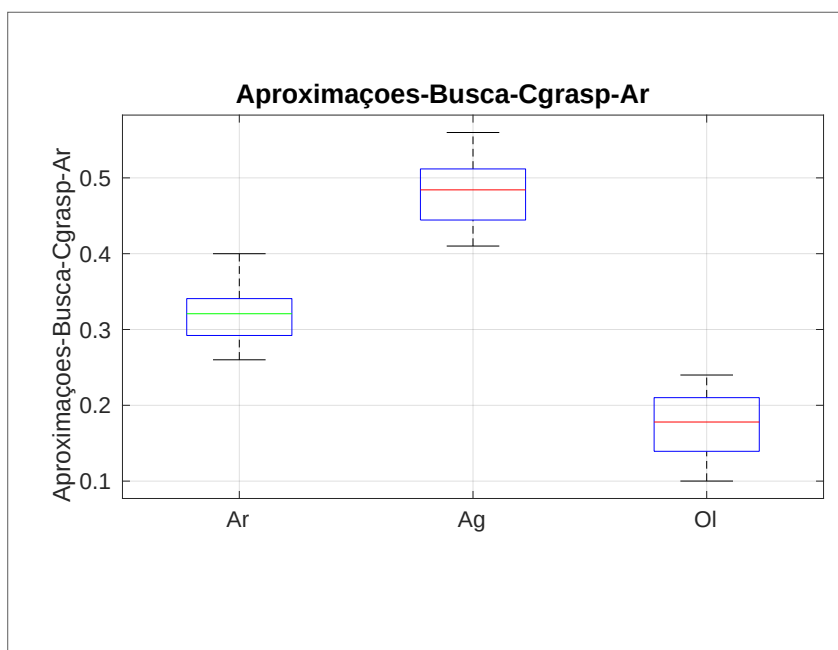
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão do Ar (Ar ○); Água +; Óleo *.



(b) Medidas da área abaixo da curva do espectro total. Área calculada (○); área real +; resíduo *.



(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP. Omissão Ar.



Fonte: Próprio Autor

5.3 COMPARAÇÃO COM TRABALHOS DO ESTADO DA ARTE

A comparação do método descrito na tese, com trabalhos publicados no estado da arte, será feita na presente seção. Essa comparação é realizada com métodos que trabalharam base de dados semelhantes à Base 1, apresentada no presente trabalho. Serão avaliadas as saídas dos métodos, e os erros cometidos na determinação das frações percentuais por eles. Nosso algoritmo será submetido a mesma avaliação para fins de comparação. Por fim, também são apresentadas as frações de contagem percentual determinadas pelo simples ajuste dos LLS, antes da execução do nosso algoritmo, e as frações de contagem percentuais obtidas após a execução dos procedimentos de otimização descritos na presente obra.

Na Tabela 20, são comparados os resultados de frações de contagens percentuais do presente trabalho com os resultados apresentados por trabalhos publicados nos últimos anos.

O nosso algoritmo é comparado com: (1) o Algoritmo SGSD (SHAHABINEJAD; VOSOUGHI, 2019); (2) o Algoritmo NNLS (SHAHABINEJAD; VOSOUGHI, 2019); (3) Algoritmo PSO (SHAHABINEJAD; VOSOUGHI, 2018); (4) Algoritmo MLR (SHAHABINEJAD; VOSOUGHI, 2018); (5) Algoritmo GAs (ALAMANIOTIS; JEVREMOVIC, 2015).

A primeira coluna da Tabela 20 leva em conta a notação **Isótopo - Fração Percentual** para cada um dos métodos. O dataset 1, na primeira linha e na primeira coluna, por exemplo, traz a notação $^{60}\text{Co} - 48\% | 48\% | 50\%$. Essa notação indica que o método SGSD levou em conta uma fração real de 48%; o NNLS, uma participação real de 48%, e o algoritmo adotado no presente trabalho adotou uma participação no espectro total de 50%. A mesma descrição é válida para os demais datasets.

Para cada Dataset, as colunas dois, três e quatro são subdivididas em colunas: **F.A** - *Fraction Abundance*; **S.D** - *Standard Deviation*; **ER** - *Error*. As colunas acima mencionadas podem ser preenchidas com o valor '-' caso o isótopo não se aplique ao dataset mencionado. Caso o método comparado identifique uma biblioteca que não faça parte do espectro, como no caso do Dataset 1 em que o *Na*, *Eu* e o *Am* não fazem parte das bibliotecas inseridas no método descrito para o NNLS, mas que foram identificadas como participantes no espectro total, será usada a abreviatura **N/A** para **Não Aplicável**.

Importante destacar que as instâncias em que a biblioteca esteja marcada com **Elemento*** indica que ocorre a omissão e estimativa dessa biblioteca em nosso método. Então, as quatro primeiras linhas da Tabela 20 (Dataset 1, Dataset2, ..., Dataset 4) apresentam, respectivamente, a omissão dos elementos *Am*, *Cs*, *Co* e *Na*.

Para o caso do Dataset 1, o método do SGSD identificou as bibliotecas do *Co* e do *Cs* na descrição do espectro total, como era de se esperar, já que o espectro total é composto pelos dois elementos. Para essas bibliotecas, houve uma determinação de 48.44% e 51.56%, respectivamente, apresentando um erro de 0.91% e 0.84%. O método do NNLS identificou esses mesmos elementos, com um erro de determinação de 0.38% e 1.07%, respectivamente. A observação a ser feita para o método do NNLS é que ele identificou as bibliotecas do *Na*, *Eu* e *Am*, mesmo esses elementos não tendo participação no espectro total. O método do presente trabalho atingiu para essas bibliotecas uma fração em peso de 49.98% e 19.97%, com erros de 0.04% e 0.15%, sendo o melhor comparativamente aos métodos descritos para o Dataset 1. A biblioteca estimada para o nosso método no caso do Dataset 1 apresentou um erro de 0.2%, o que é muito bom, considerando que as estimativas para essa biblioteca evidentemente introduzem incertezas e instabilidades ao conjunto de equações dos LLS. A observação a ser feita consiste na aproximação das frações em peso para o *Co*, que apresentou um desvio padrão de 5.21, maior que os demais métodos para o primeiro conjunto de dados. Mas esse fato, mesmo assim, não foi impeditivo da obtenção de resultados eficazes na determinação da sua fração em peso, que apresentou erro da ordem de 0,04%.

O método do SGSD, inserido na abordagem do FSA, apresentou valores de erros de frações percentuais que são bastante adequados quando comparados com os erros para cada uma das instâncias. Um dos fatores que pesa contra o SGSD é que ele apresenta custos computacionais elevados ao requerer um grande número de bibliotecas na validação do seu conjunto ativo. Também pode-se enumerar como desvantagens do método de Shahabinejad e Vosoughi (2019) é que ele apresentou resultados similares ao NNLS, além de apresentar sensibilidades a erros estatísticos. Por fim, as soluções apresentadas pelo SGSD tendem a não ser uma solução única, já que diversas combinações de bibliotecas podem ser realizadas na descrição do espectro total.

A proposta da presente tese apresenta a vantagem de não manipular o estágio do MC, como descrito em Shahabinejad e Vosoughi (2019), não apresentando dependência de uma grande quantidade de bibliotecas de entrada. Nosso método não manipula o universo total dos canais das bibliotecas consideradas, o que é um fator decisivo para a qualidade das soluções apresentadas. Pode-se enumerar também como potencialidade do nosso método a capacidade de analisar um grande conjunto de possíveis soluções ao problema proposto. O trabalho de Shahabinejad e Vosoughi (2019), de todos os conjuntos de possíveis soluções, enumera a primeira delas que apresentar adequação ao espectro da amostra em estudo, não necessariamente explorando o conjunto de possíveis soluções.

Para o caso do Dataset seis, por exemplo, o método do PSO identificou os elementos do *Am*, *Co* e *Na*, obtendo valores de fração e peso, respectivamente, de 51.85%, 41.59% e 5.50% e erros de 4.01%, 6.37% e 3.98%. O método do MLR atingiu as marcas para esses elementos de 52.14%, 41.37% e 5.47%, com erros de 4.58%, 6.86% e 4.55%, respectivamente. Para esse mesmo dataset, o método baseado em Lógica Fuzzy e GAs utilizou elementos diferentes que os métodos do PSO e MLR, identificando as bibliotecas do *Am*, *Co* e *Cs*. Nesse caso, para o *Am*, *Co* e *Cs* identificou frações em peso de, respectivamente, 1,5%, 67% e 31.5%, encontrando erros de 66.06%, 7.75% e 27.11% para cada um deles. Comparando-se os números encontrados pelo método do GAs, o nosso método, para qualquer uma das instâncias trabalhadas, encontrou erros para os valores das frações evidentemente melhores que os determinados pelo Algoritmo GAs.

Os algoritmos genéticos apresentados em Alamaniotis e Jevremovic (2015) são aplicados após a caracterização dos elementos inseridos na amostra por meio de Lógica Fuzzy. A desvantagem do presente trabalho, referencialmente aos procedimentos tomados nesta obra, é que a caracterização dos elementos não é utilizada em um momento posterior para determinação da seção de canais mais relevantes das bibliotecas de entrada. Utiliza os mesmos testes de abundância para as frações em peso e adequação para o espectro total descritos por Shahabinejad e Vosoughi (2019), fazendo uso, assim, da descrição do espectro por um particionamento de bibliotecas. Mais uma vez, incorre-se no problema da exploração das possíveis soluções, além de manipular sistemas de equações sobredeterminados.

Tomando por base o Data set três aplicado ao nosso algoritmo, encontrou-se um valor de fração em peso de 47.95% para a biblioteca estimada (*Co*), com um erro de 4.10%. A biblioteca do *Co*, e a instância em que o *Co* foi omitido, foram os casos mais difíceis na determinação das frações em peso para o algoritmo descrito na presente tese. Mais uma vez, destaca-se que essa dificuldade ocorre, presumivelmente, ao fato de o pico da biblioteca do *Co* ocorrer após todas as demais bibliotecas irem a zero. A interpretação para os demais resultados podem ser extrapolados de acordo com as descrições realizadas nos dois últimos parágrafos.

As conclusões que podem ser retiradas a partir dos valores descritos na Tabela 20 é que valores de erros das frações em peso apresentam-se em conformidade com os valores divulgados pelos últimos trabalhos publicados na área para otimização dos LLS no contexto do MCLLS.

Os resultados apresentados no contexto do WSA

A Tabela 21 apresenta os valores das frações em peso anterior ao processo de otimização, e as mesmas frações em peso posterior ao processo de otimização, para uma iteração arbitrária

Tabela 20 – Comparativo com trabalhos do estado da arte. A tabela adota a notação **Isótopo - Participação percentual** para cada Dataset. Os datasets 1, 2, 3 e 4 apresentam a omissão, respectivamente, do *Am*, *Cs*, *Co* e *Na*.

Dataset # Isótopo - Fração Percentual.	Análise dos Resultados								
	SGSD			NNLS			Nosso AI.		
	F.A (%)	S.D (%)	ER (%)	F.A (%)	S.D (%)	ER (%)	F.A (%)	S.D (%)	ER (%)
Dataset 1									
1) ⁶⁰ Co - 48% 48% 50%	1) 48.44	1.95	0.91	1) 47.82	1.97	0.38	1) 49.98	5.21	0.04
2) ¹³⁷ Cs - 52% 52% 20%	2) 51.56	3.11	0.84	2) 51.44	3.11	1.07	2) 19.97	2.08	0.15
3) ²² Na - 0% 0% 10%	3) -	-	-	3) 0.55	312	N/A	3) 9.93	1.03	0.70
4) ¹⁵² Eu - 0% 0% 0%	4) -	-	-	4) 0.03	8789	N/A	4) -	-	-
5) ²⁴¹ Am* - 0% 0% 20%	5) -	-	-	5) 0.16	2474	N/A	5) 20.04	2.09	0.2
Dataset 2									
1) ⁶⁰ Co - 0% 0% 50%	1) -	-	-	1) 1.43	46.39	N/A	1) 49.61	4.60	0.78
2) ¹³⁷ Cs* - 55.45% 55.45% 20%	2) 56.63	2.00	2.13	2) 55.37	2.04	0.15	2) 19.26	1.81	4.00
3) ²² Na - 6.61% 6.61% 10%	3) 8.38	14.37	26.85	3) 8.60	13.97	30.23	3) 11.36	1.12	13.6
4) ¹⁵² Eu - 37.94% 37.94% 0%	4) 34.99	5.13	7.78	4) 34.60	5.18	8.82	4) -	-	-
5) ²⁴¹ Am - 0% 0% 20%	5) -	-	-	5) -	-	-	5) 19.74	1.83	1.30
Dataset 3									
1) ⁶⁰ Co* - 100% 100% 50%	1) 72.55	1.94	27.45	1) 67.81	2.06	32.19	1) 47.95	2) 5.35	4.10
2) ¹³⁷ Cs - 0% 0% 20%	2) -	-	-	2) 3.67	64.53	N/A	2) 20.16	2.22	0.8
3) ²² Na - 0% 0% 10%	3) -	-	-	3) 3.81	66.21	N/A	3) 10.10	1.12	1.00
4) ¹⁵² Eu - 0% 0% 0%	4) 27.45	13.78	N/A	4) 24.71	15.20	N/A	4) -	-	-
5) ²⁴¹ Am - 0% 0% 20%	5) -	-	-	5) -	-	-	5) 21.76	2.49	8.80
Dataset 4									
1) ⁶⁰ Co - 0% 0% 50%	1) -	-	-	1) -	-	-	1) 50.29	3.01	1.18
2) ¹³⁷ Cs - 0% 0% 20%	2) -	-	-	2) -	-	-	2) 19.73	1.18	1.35
3) ²² Na* - 0% 0% 10%	3) -	-	-	3) -	-	-	3) 9.14	0.58	8.60
4) ¹⁵² Eu - 0% 0% 0%	4) -	-	-	4) -	-	-	4) -	-	-
5) ²⁴¹ Am - 0% 0% 20%	5) -	-	-	5) -	-	-	5) 20.82	1.25	4.10
Dataset # Isótopo - Fração Percentual.	PSO			MLR			Fuzzy - GA		
	F.A (%)	S.D (%)	ER (%)	F.A (%)	S.D (%)	ER (%)	F.A (%)	S.D (%)	ER (%)
Dataset 5									
1) ²⁴¹ Am - 0% 0% 0%	1) -	-	-	1) -1.56	422	5.59	1) 0	0	0
2) ⁶⁰ Co - 100% 100% 100%	2) 93.04	1.68	6.96	2) 94.41	1.66	5.59	2) 98.9	1	1.10
3) ¹³⁷ Cs - 0% 0% 0%	3) -	-	-	3) -	-	-	3) -	-	-
4) ²² Na - 0% 0% 0%	4) -	-	-	4) -	-	-	4) -	-	-
5) ¹⁵² Eu - 0% 0% 0%	5) 6.96	60.63	-	5) 7.15	58.93	-	5) -	-	-
6) ²³⁹ Pu - 0% 0% 0%	6) -	-	-	6) -	-	-	6) 1.1	2	N/A
Dataset 6									
1) ²⁴¹ Am - 49.85% 49.85% 4.42%	1) 51.85	1.37	4.01	1) 52.14	1.37	4.58	1) 1.5	1.5	66.06
2) ⁶⁰ Co - 44.42% 44.42% 72.63%	2) 41.59	0.41	6.37	2) 41.37	0.41	6.86	2) 67.0	20	7.75
3) ¹³⁷ Cs - 0% 0% 22.96%	3) -	-	-	3) 0	0	0	3) 31.5	19.5	27.11
4) ²² Na - 5.73% 5.73% 0%	4) 5.50	5.54	3.98	4) 5.47	5.57	4.55	4) -	-	-
5) ¹⁵² Eu - 0% 0% 0%	5) -	-	-	5) -	-	-	5) -	-	-
6) ²³⁹ Pu - 0% 0% 0%	6) -	-	-	6) -	-	-	6) -	-	-
Dataset 7									
1) ²⁴¹ Am - 0% 0% 0%	1) -	-	-	1) -	-	-	1) -	-	-
2) ⁶⁰ Co - 45.20% 45.20% 74.25%	2) 44.92	0.97	0.61	2) 44.92	0.97	0.61	2) 71.2	6	4.10
3) ¹³⁷ Cs - 48.97% 48.97% 25.75%	3) 48.85	1.52	0.25	3) 48.85	1.52	0.25	3) 28.8	6	11.84
4) ²² Na - 5.83% 5.83% 0%	4) 6.23	12.64	6.83	4) 6.23	12.64	6.83	4) -	-	-
5) ¹⁵² Eu - 0% 0% 0%	5) -	-	-	5) -	-	-	5) -	-	-
6) ²³⁹ Pu - 0% 0% 0%	6) -	-	-	6) -	-	-	6) -	-	-

Fonte: Próprio Autor

do algoritmo presente na tese.

A coluna 1 da Tabela 21 demonstra a base de dados trabalhada. Os procedimentos de ajuste aos LLS está presente na coluna 2. Esses procedimentos podem ser o ajuste tradicional dos LLS, descrito no campo 'Ajuste LLS', e o procedimento do ajuste do LLS realizado após o processo de otimização, no campo descrito por 'Ajuste otimizado LLS'. A descrição dos elementos das referidas bases encontra-se na coluna 3, com as bibliotecas estimadas com a notação elemento*. As saídas das frações em peso são determinadas na coluna 4, com os respectivos cálculos de erros na coluna 5.

No contexto do MCLLS, caso o setup experimental realizasse o ajuste pelos LLS sem a devida otimização, o módulo do MC seria inicializado novamente, com novas entradas para esse módulo determinadas pelas frações em peso saídas do sistema do LLS, para uma nova simulação experimental e geração de novas bibliotecas. Esse processo, como um todo, é custoso computacionalmente.

As saídas do LLS, otimizadas pelo presente algoritmo, implicariam em uma convergência mais rápida para o sistema do MCLLS, já que os erros para as frações em peso das bibliotecas são minimizados no processo de otimização. Essa maior precisão das frações em peso por parte do LLS implica em um número menor de chamadas ao módulo do MC.

Tabela 21 – Tabela demonstrando o comparativo entre as frações percentuais antes e depois da execução da otimização por parte do algoritmo.

Base	Procedimento	Elemento	Fração Calculada (%)	Erro (%)
Base 1	Ajuste LLS	<i>Am</i>	0.29	98.55
		<i>Co</i>	52.48	4.96
		<i>Cs</i>	47.14	135
		<i>Na</i>	0.065	35
	Ajuste otimizado LLS	<i>Am</i>	19.96	0.2
		<i>Co</i>	49.60	0.8
		<i>Cs</i> *	19.93	0.35
		<i>Na</i>	10.48	4.8
Base 2	Ajuste LLS	Ar	39.45	19.54
		Água	-0.14	-
		Óleo	60.71	257
	Ajuste otimizado LLS	Ar	33.46	1.39
		Água*	48.01	2.02
		Óleo	18.53	7.35

Fonte: Próprio Autor

6 CONCLUSÕES E TRABALHOS FUTUROS

Na tese, foram tratados dois problemas no enfrentamento do ajuste do LLS ao espectro de uma amostra em estudo. O primeiro reside em definir as frações em peso dos constituintes de uma amostra através da análise dos seus respectivos espectros. Já segundo problema consiste em determinar espectros de constituintes ausentes através da análise dos espectros de constituintes conhecidos.

A principal abordagem utilizada atualmente na técnica do PGNAA é o método do MCLLS. No método do MCLLS, o módulo do MC fornece os espectros, e o sistema do LLS determina as frações em peso.

O presente trabalho contribuiu na melhoria do MCLLS, fornecendo ao sistema do LLS um método de otimização aberta. Os métodos escolhidos, bem como os algoritmos heurísticos GRASP e CGRASP, foram devidamente modelados ao domínio do problema.

Na execução dos algoritmos, o processo de otimização foi aplicado a regiões dos dados de entrada, determinadas pelas contagens espectrais máximas (picos). O GRASP, como método de otimização genérico que pode ser aplicado a diversas situações, problemas e fenômenos distintos que exigem uma devida modelagem ao problema, exigiu que as etapas de construção, busca e linkagem fosse devidamente modeladas ao domínio do problema do PGNAA.

A metodologia aplicada apresentou avanços na modelagem do GRASP frente ao trabalho de Cardoso et al. (2020), ao otimizar ambos os problemas do PGNAA descritos no presente trabalho.

O modelo foi testado em duas bases de dados. Foram aplicadas às bases um inovador método geométrico para o segundo problema descrito nesse trabalho que reside na estimação do espectro ausente. Os resultados para a determinação desse espectro conduzem à conclusão que a abordagem é adequada frente aos métodos lineares para a determinação do espectro total, muito embora haja espaço para investigações adicionais na modelagem geométrica.

Os resultados precisos para as frações em peso nos bancos de dados utilizados permitem ratificar os procedimentos da seção dos dados e que são alternativos aos principais métodos de particionamento de espectros (SHAHABINEJAD; VOSOUGHI, 2019; ALAMANIOTIS; CHOI; TSOUKALAS, 2015; SHAHABINEJAD; VOSOUGHI, 2018). A obra atual determinou regiões do espectro baseado no particionamento pelos picos de contagem das bibliotecas; diferentemente do trabalho de Cardoso et al. (2020), que realizou o particionamento dos dados por critérios

de independência linear, apresentando mais uma alternativa aos poucos trabalhos do estado da arte que manipulam regiões específicas representativas das bibliotecas de energia.

O condicionamento da matriz de covariância apresentou relevância para a boa qualidade das frações em peso do problema posto pelo sistema do LLS, mas a minimização dessa grandeza não é garantia para que as melhores soluções possíveis sejam encontradas. Foi comprovado que é possível atingir valores de condicionamento extremamente baixos para matrizes de dimensão $n \times n$, comparando-os com as matrizes de dimensão 2×2 (MERIÇ, 2012). Mas os valores de condicionamento não se mostraram tão eficazes na avaliação das saídas quanto os valores dos resíduos.

O presente trabalho apresentou estimativas das frações em peso, na abordagem do MCLLS à técnica do PGNAA. As frações foram avaliadas a partir de uma otimização do LLS pelos algoritmos heurísticos GRASP e CGRASP, dotados de uma função objetivo única. Os resultados sugerem que o modelo foi capaz de melhorar a resolução do método MCLLS por meio da devida modelagem aos algoritmos aplicados e adequada geometrização do processo de análise do espectro gama na estimativa de biblioteca não fornecida ao sistema do LLS pelo módulo do MC. Esses conceitos culminaram em uma evidente adequação para as frações em peso, oferecendo maior precisão nas saídas dos LLS e ratificando um inovador modelo geométrico para estimativa do espectro ausente. Além disso, a abordagem geométrica pode ser implementada em qualquer aplicação do PGNAA, porque lida com aspectos fundamentais do método, apresentando potencial para novos desenvolvimentos que poderão ser essenciais na consolidação do MCLLS como principal técnica relacionada ao PGNAA.

6.1 TRABALHOS FUTUROS

Por fim, como trabalhos futuros, algumas direções podem ser seguidas, a exemplo de:

- Contextualizar as métricas apresentadas em outros trabalho, como (SHAHABINEJAD; VOSOUGHI, 2018), (SHAHABINEJAD; VOSOUGHI, 2019) para teste de significância e de adequação, ao trabalho apresentado e comparação com a métrica apresentada no presente trabalho;
- Analisar as possibilidades de particionamento dos dados, como realizado no presente trabalho, conjuntamente com a separação de bibliotecas, como descrito em Meriç (2012).

Há evidências de que nessa configuração seja possível minimizar os valores de número de condicionamento apresentados no presente trabalho;

- Implementar o algoritmo no sentido de determinar contribuições indesejadas presentes como uma biblioteca de background;
- Analisar diferenças no número de condicionamento entre iterações consecutivas como métrica na determinação das frações de contagem percentual e consequente ajuste ao espectro total;
- Propor uma métrica às estimativas de altura contidas na Seção 4.3 no intuito de viabilizar convergência em menor tempo dos valores de frações de contagem percentual;
- Analisar saídas realizadas por outros métodos de ajuste ao espectro total diferente do método dos mínimos quadrados utilizados neste trabalho. Trabalhos como (SHAHABINEJAD; VOSOUGHI, 2019), (SHAHABINEJAD; HOSSEINI; SOHRABPOUR, 2016) realizam o ajuste ao espectro total pelo solução analítica dos sistemas de equações;
- Investigar a relação entre a qualidade das soluções e as características de gulosidade e de aleatoriedade do algoritmo proposto, e sugerir uma maneira de maior efetividade do nosso algoritmo;
- Explorar outras regiões contidas na Figura 7 de modo a aperfeiçoar as estimativas da biblioteca faltante, realizando uma conexão entre a heurística proposta para definições de novas soluções e a geometria proposta na Figura 7;
- Reduzir as instabilidades numéricas introduzidas no processo de ajuste do espectro total devido as estimativas da biblioteca faltante;
- Investigar as possibilidades de diminuir a quantidade de canais que compõem inicialmente as partições;
- Determinar o impacto da escolha de outros valores de parâmetros α na expressão do passo construtivo do CGRASP;
- Explorar a efetividade do estágio de busca do CGRASP em aprimorar as saídas vindas do passo construtivo.

REFERÊNCIAS

- ALAMANOTIS, M.; CHOI, C. K.; TSOUKALAS, L. H. Application of fireworks algorithm in gamma-ray spectrum fitting for radioisotope identification. *International Journal of Swarm Intelligence Research (IJSIR)*, IGI Global, v. 6, n. 2, p. 102–125, 2015.
- ALAMANOTIS, M.; GAO, R.; TSOUKALAS, L. H.; JEVREMOVIC, T. Intelligent order-based method for synthesis of nrf spectra and detection of hazardous materials. In: IEEE. *2009 21st IEEE International Conference on Tools with Artificial Intelligence*. [S.l.], 2009. p. 658–665.
- ALAMANOTIS, M.; IKONOMOPOULOS, A.; JEVREMOVIC, T.; TSOUKALAS, L. H. Intelligent recognition of signature patterns in nrf spectra. *Nuclear technology*, Taylor & Francis, v. 175, n. 2, p. 480–497, 2011.
- ALAMANOTIS, M.; JEVREMOVIC, T. Hybrid fuzzy-genetic approach integrating peak identification and spectrum fitting for complex gamma-ray spectra analysis. *IEEE Transactions on Nuclear Science*, IEEE, v. 62, n. 3, p. 1262–1277, 2015.
- ALAMANOTIS, M.; MATTINGLY, J.; TSOUKALAS, L. H. Pareto-optimal gamma spectroscopic radionuclide identification using evolutionary computing. *IEEE Transactions on Nuclear Science*, IEEE, v. 60, n. 3, p. 2222–2231, 2013.
- ALAMANOTIS, M.; TSOUKALAS, L. H. Assessment of gamma-ray-spectra analysis method utilizing the fireworks algorithm for various error measures. In: *Critical Developments and Applications of Swarm Intelligence*. [S.l.]: IGI Global, 2018. p. 155–181.
- ALLISON, J.; AMAKO, K.; APOSTOLAKIS, J.; ARAUJO, H.; DUBOIS, P. A.; ASAI, M.; BARRAND, G.; CAPRA, R.; CHAUVIE, S.; CHYTRACEK, R. et al. Geant4 developments and applications. *IEEE Transactions on nuclear science*, IEEE, v. 53, n. 1, p. 270–278, 2006.
- ANDERSON, D.; CUNNINGHAM, W.; MACKEY, E. Determination of boron in food and biological reference materials by neutron capture prompt- γ activation. *Fresenius' Journal of Analytical Chemistry*, Springer, v. 338, n. 4, p. 554–558, 1990.
- ANDROULAKAKI, E.; KOKKORIS, M.; TSABARIS, C.; ELEFTHERIOU, G.; PATIRIS, D.; PAPPA, F.; VLASTOU, R. In situ γ -ray spectrometry in the marine environment using full spectrum analysis for natural radionuclides. *Applied Radiation and Isotopes*, Elsevier, v. 114, p. 76–86, 2016.
- ANTON, H.; RORRES, C. *Algebra Linear com Aplicações*. Bookman, 2001. ISBN 9788573078473. Disponível em: <<http://books.google.com.br/books?id=pOaaSKP9lcMC>>.
- ASTER, R. C.; BORCHERS, B.; THURBER, C. H. *Parameter estimation and inverse problems*. [S.l.]: Elsevier, 2018.
- BALDONCINI, M. Applicazione del metodo non negative least square alla full spectrum analysis nel processo di calibrazione di uno spettrometro di raggi gamma portatile. graduate thesis. *University of Ferrara*, 2010.
- BALDONCINI, M.; ALBÉRI, M.; BOTTARDI, C.; CHIARELLI, E.; RAPTIS, K. G. C.; STRATI, V.; MANTOVANI, F. Investigating the potentialities of monte carlo simulation for assessing soil water content via proximal gamma-ray spectroscopy. *Journal of environmental radioactivity*, Elsevier, v. 192, p. 105–116, 2018.

BECKER, D.; ANDERSON, D.; LINDSTROM, R.; GREENBERG, R.; GARRITY, K.; MACKEY, E. Use of inaa, pgaa, and rnaa to determine 30 elements for certification of an srm: Tomato leaves, 157a. *Journal of radioanalytical and nuclear chemistry*, Springer, v. 179, n. 1, p. 149–154, 1994.

BODE, P.; BUENO, M.; GORAIEB, K.; KOSTER-AMMERLAAN, M. The compton continuum: Nuisance or new information in naa? *Journal of radioanalytical and nuclear chemistry*, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV . . . , v. 278, n. 3, p. 727–732, 2008.

BRIESMEISTER, J. F. et al. Mcnptm-a general monte carlo n-particle transport code. *Version 4C, LA-13709-M, Los Alamos National Laboratory*, p. 2, 2000.

BURR, T.; HAMADA, M. Radio-isotope identification algorithms for nai γ spectra. *Algorithms*, Molecular Diversity Preservation International, v. 2, n. 1, p. 339–360, 2009.

CACIOLLI, A.; BALDONCINI, M.; BEZZON, G.; BROGGINI, C.; BUSO, G.; CALLEGARI, I.; COLONNA, T.; FIORENTINI, G.; GUASTALDI, E.; MANTOVANI, F. et al. A new fsa approach for in situ γ ray spectroscopy. *Science of the Total Environment*, Elsevier, v. 414, p. 639–645, 2012.

CARDOSO, H.; MELO, S. B.; SILVA, R. M. Optimizing the condition number in spectral analysis. In: *IEEE SMC 2018 IEEE International Conference On Systems, Man and Cybernetics*. [S.l.: s.n.], 2018.

CARDOSO, H. A. C.; CARVALHO, S. A. L. de; MELO, S. de B.; SILVA, R. M. de A.; SANTOS, S. G. T. de C.; DANTAS, C. C. Spectral analysis and optimization of the condition number problem. *Computer Physics Communications*, Elsevier, p. 107587, 2020.

CARLEVARO, C.; WILKINSON, M.; BARRIOS, L. A genetic algorithm approach to routine gamma spectra analysis. *Journal of Instrumentation*, IOP Publishing, v. 3, n. 01, p. P01001, 2008.

CARVAJAL, M.; GARCÍA-PAREJA, S.; GUIRADO, D.; VILCHES, M.; ANGUIANO, M.; PALMA, A.; LALLENA, A. Monte carlo simulation using the penelope code with an ant colony algorithm to study mosfet detectors. *Physics in Medicine & Biology*, IOP Publishing, v. 54, n. 20, p. 6263, 2009.

CHANG, H. P.; MERIC, I.; SUDAC, D.; NAđ, K.; OBHOđAŠ, J.; HOU, G.; ZHANG, Y.; GARDNER, R. P. Implementation of the monte carlo library least-squares (mcLLS) approach for quantification of the chlorine impurity in an on-line crude oil monitoring system. *Radiation Physics and Chemistry*, Elsevier, v. 155, p. 197–201, 2019.

CHEN, F.; NANPING, W. Gaussian fitting in gamma-ray spectral decomposition. *Nuclear Science and Techniques*, v. 21, n. 4, p. 214–214, 2013.

CHEN, L.; ZHOU, T.; TANG, Y. Protein structure alignment by deterministic annealing. *Bioinformatics*, Oxford University Press, v. 21, n. 1, p. 51–62, 2004.

CULLEN, D. E. *TART98 a coupled neutron-photon 3-D, combinatorial geometry time dependent monte carlo transport code*. [S.l.], 1998.

DANTZIG, G. B.; RAMSER, J. H. The truck dispatching problem. *Management science*, Informs, v. 6, n. 1, p. 80–91, 1959.

DIVER, D.; IRELAND, D. Spectral decomposition by genetic algorithm. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 399, n. 2-3, p. 414–420, 1997.

EAKINS, J. On the effect of updated mcnp photon cross section data on the simulated response of the hpa tld. *Radiation protection dosimetry*, Oxford University Press, v. 134, n. 1, p. 66–71, 2009.

ELEFThERIOU, G.; TSABARIS, C.; ANDROULAKAKI, E.; PATIRIS, D.; KOKKORIS, M.; KALFAS, C.; VLASTOU, R. Radioactivity measurements in the aquatic environment using in-situ and laboratory gamma-ray spectrometry. *Applied Radiation and Isotopes*, Elsevier, v. 82, p. 268–278, 2013.

FATAH, A.; AHMED, A. Analysis of gamma-ray spectra using levenberg-marquardt method. *spectrum*, v. 1, p. 3, 2011.

FAVORITE, J. A. Using the schwinger variational functional for the solution of inverse transport problems. *Nuclear science and engineering*, Taylor & Francis, v. 146, n. 1, p. 51–70, 2004.

FEO, T. A.; RESENDE, M. G. Greedy randomized adaptive search procedures. *Journal of global optimization*, Springer, v. 6, n. 2, p. 109–133, 1995.

FERRARI, A.; SALA, P. R.; FASSO, A.; RANFT, J. *FLUKA: A multi-particle transport code (Program version 2005)*. [S.I.], 2005.

FORSBERG, P.; AGARWAL, V.; PERRY, J.; GAO, R.; TSOUKALAS, L. H.; JEVREMOVIC, T. Peakseek: A statistical processing algorithm for radiation spectrum peak identification. In: IEEE. *2009 21st IEEE International Conference on Tools with Artificial Intelligence*. [S.I.], 2009. p. 674–678.

FURR, A.; ROBINSON, E.; ROBINS, C. A spectrum stripping technique for qualitative activation analysis using monoenergetic gamma spectra. *Nuclear Instruments and Methods*, Elsevier, v. 63, n. 2, p. 205–209, 1968.

GARCIA-TALAVERA, M.; ULICNY, B. A genetic algorithm approach for multiplet deconvolution in γ -ray spectra. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 512, n. 3, p. 585–594, 2003.

GARDNER, R.; GUO, P.; WANG, Y.; SOOD, A.; LEE, S.; DOBBS, C. Feasibility of neutron activation methods for measurement of sodium and aluminium in green liquor. *Applied radiation and isotopes*, Elsevier, v. 48, n. 10-12, p. 1355–1372, 1997.

GARDNER, R. P.; SOOD, A.; WANG, Y.; LIU, L.; GUO, P.; GEHRKE, R. Single peak versus library least-squares analysis methods for the pgnaa analysis of vitrified waste. *Applied radiation and isotopes*, Elsevier, v. 48, n. 10-12, p. 1331–1335, 1997.

GARDNER, R. P.; XU, L. Status of the monte carlo library least-squares (mcLLS) approach for non-linear radiation analyzer problems. *Radiation Physics and Chemistry*, Elsevier, v. 78, n. 10, p. 843–851, 2009.

GENDREAU, M.; POTVIN, J.-Y. Metaheuristics in combinatorial optimization. *Annals of Operations Research*, Springer, v. 140, n. 1, p. 189–213, 2005.

GILL, P.; MURRAY, W.; WRIGHT, M. *Numerical linear algebra and optimization*. Addison-Wesley Pub. Co., Advanced Book Program, 1991. (Numerical Linear Algebra and Optimization, v. 1). ISBN 9780201126495. Disponível em: <<http://books.google.com.br/books?id=ng3vAAAAMAAJ>>.

GILL, P. E.; MURRAY, W.; WRIGHT, M. H. *Practical optimization*. [S.l.]: Academic Press, 1981.

GLOVER, F. Tabu search—part i. *ORSA Journal on computing*, INFORMS, v. 1, n. 3, p. 190–206, 1989.

GLOVER, F. Tabu search—part ii. *ORSA Journal on computing*, INFORMS, v. 2, n. 1, p. 4–32, 1990.

GOLDBERG, D. E.; HOLLAND, J. H. Genetic algorithms and machine learning. *Machine learning*, Springer, v. 3, n. 2, p. 95–99, 1988.

GOLUB, G.; LOAN, C. V. *Matrix Computations*. Johns Hopkins University Press, 2012. (Johns Hopkins Studies in the Mathematical Sciences). ISBN 9781421407944. Disponível em: <<https://books.google.com.br/books?id=X5YfsuCWpxMC>>.

GRANVILLIERS, L.; BENHAMOU, F. Progress in the solving of a circuit design problem. *Journal of Global Optimization*, Springer, v. 20, n. 2, p. 155–168, 2001.

HAN, X.; GARDNER, R. P. Cearcpg: A monte carlo simulation code for normal and coincidence prompt gamma-ray neutron activation analysis (pgnaa), invited, mark mills award winner. *Transactions of the American Nuclear Society*, American Nuclear Society, v. 95, p. 521–522, 2006.

HAN, X.; GARDNER, R. P.; METWALLY, W. Cearcpg: a monte carlo simulation code for normal and coincidence prompt-gamma-ray neutron activation analysis. *Nuclear science and engineering*, Taylor & Francis, v. 155, n. 1, p. 143–153, 2007.

HARTLEY, R. I.; ZISSERMAN, A. *Multiple View Geometry in Computer Vision*. Second. [S.l.]: Cambridge University Press, ISBN: 0521540518, 2004.

HIRSCH, M. J.; MENESES, C. N.; PARDALOS, P. M.; RESENDE, M. G. C. Global optimization by continuous grasp. *Optimization Letters*, v. 1, n. 2, p. 201–212, Mar 2007. ISSN 1862-4480. Disponível em: <<https://doi.org/10.1007/s11590-006-0021-6>>.

JOHNSON, D. S.; TRICK, M. A. *Cliques, coloring, and satisfiability: second DIMACS implementation challenge, October 11-13, 1993*. [S.l.]: American Mathematical Soc., 1996. v. 26.

KANAWATI, W. E.; CARASCO, C.; PEROT, B.; MARIANI, A.; RAOUX, A.-C.; VALKOVIC, V.; SUDAC, D.; OBHODAS, J.; BARICEVIC, M. Gamma-ray signatures improvement of the euritrack tagged neutron inspection system database. *IEEE Transactions on Nuclear Science*, IEEE, v. 57, n. 5, p. 2879–2885, 2010.

KANAWATI, W. E.; PEROT, B.; CARASCO, C.; ELEON, C.; VALKOVIC, V.; SUDAC, D.; OBHODAS, J. Conversion factors from counts to chemical ratios for the euritrack tagged neutron inspection system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 654, n. 1, p. 621–629, 2011.

- KANGAS, L. J.; KELLER, P. E.; SICILIANO, E. R.; KOUZES, R. T.; ELY, J. H. The use of artificial neural networks in pvt-based radiation portal monitors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 587, n. 2-3, p. 398–412, 2008.
- KARMAKAR, N.; RAMAKRISHNAN, K. Computational results of an interior point algorithm for large scale linear programming. *Mathematical Programming*, Springer, v. 52, n. 1-3, p. 555–586, 1991.
- KEARFOTT, R. B. Some tests of generalized bisection. *ACM Transactions on Mathematical Software (TOMS)*, ACM, v. 13, n. 3, p. 197–220, 1987.
- KIRKPATRICK, S. Optimization by simulated annealing: Quantitative studies. *Journal of statistical physics*, Springer, v. 34, n. 5-6, p. 975–986, 1984.
- KITAMURA, M.; TAKEDA, Y.; KAWASE, K.; SUGIYAMA, K. A compact fitting algorithm for accurate analysis of gamma spectra. *Nuclear Instruments and Methods*, Elsevier, v. 136, n. 2, p. 363–367, 1976.
- KNOLL, G. F. Radiation detection and measurement John Wiley & sons. Inc., New York, p. 65, 2000.
- KOLMAN, B. *Introdução a algebra linear com aplicações*. LTC, 2006. ISBN 9788521614784. Disponível em: <<http://books.google.com.br/books?id=fJuKMwAACAAJ>>.
- KROKHMAL, P.; MURPHEY, R.; PARDALOS, P.; URYASEV, S.; ZRAZHEVSKI, G. Robust decision making: Addressing uncertainties in distributions. In: *Cooperative control: Models, applications and algorithms*. [S.l.]: Springer, 2003. p. 165–185.
- KROKHMAL, P.; MURPHEY, R.; PARDALOS, P.; URYASEV, S. Use of conditional value-at-risk in stochastic programs with poorly defined distributions. In: *Recent developments in cooperative control and optimization*. [S.l.]: Springer, 2004. p. 225–241.
- LIMA, E. L. *Álgebra linear*. [S.l.: s.n.], 2006.
- LOMBARD, A.; TAMAYO, S.; FONTANE, F. Modelling the time-dependent vrp through open data. *arXiv preprint arXiv:1804.07555*, 2018.
- MATTINGLY, J.; MITCHELL, D. J. A framework for the solution of inverse radiation transport problems. *IEEE Transactions on Nuclear Science*, IEEE, v. 57, n. 6, p. 3734–3743, 2010.
- MAUCEC, M.; RIGOLLET, C. Monte carlo simulations to advance characterisation of landmines by pulsed fast/thermal neutron analysis. *Applied radiation and isotopes*, Elsevier, v. 61, n. 1, p. 35–42, 2004.
- MENG-HUA, Z.; LIANG-GANG, L.; DONG-XU, Q.; ZHONG, Y.; AO-AO, X. Least square fitting of low resolution gamma ray spectra with cubic b-spline basis functions. *Chinese Physics C*, IOP Publishing, v. 33, n. 1, p. 24, 2009.
- MENG, L.-J.; RAMSDEN, D. An inter-comparison of three spectral-deconvolution algorithms for gamma-ray spectroscopy. *IEEE Transactions on nuclear science*, IEEE, v. 47, n. 4, p. 1329–1336, 2000.
- MERIÇ, I. Nuclear methods for subsea fluid characterization. The University of Bergen, 2012.

MERIC, I.; JOHANSEN, G. A.; MATTINGLY, J.; GARDNER, R. On the ill-conditioning of the multiphase flow measurement by prompt gamma-ray neutron activation analysis. *Radiation Physics and Chemistry*, Elsevier, v. 95, p. 401–404, 2014.

MILLER, I.; HOLMES, T. W.; GARDNER, R. P. An analytical approach for treating background in spectral analysis measurements. *Radiation Physics and Chemistry*, Elsevier, v. 116, p. 87–91, 2015.

MORSY, Z.; EL-WAHAB, M. A.; EL-FARAMAWY, N. Determination of natural radioactive elements in abo zaabal, egypt by means of gamma spectroscopy. *Annals of Nuclear Energy*, Elsevier, v. 44, p. 8–11, 2012.

MUKHERJEE, B. A high-resolution neutron spectra unfolding method using the genetic algorithm technique. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 476, n. 1-2, p. 247–251, 2002.

MURRAY, W.; BUTTERFIELD, K.; BAIRD, W. Automated radioisotope identification using fuzzy logic and portable czr detectors. In: IEEE. *2000 IEEE Nuclear Science Symposium. Conference Record (Cat. No. 00CH37149)*. [S.l.], 2000. v. 2, p. 9–129.

NELSON, W. R.; ROGERS, D. W.; HIRAYAMA, H. *The EGS4 code system*. [S.l.], 1985.

PAPADIMITRIOU, C. H.; STEIGLITZ, K. *Combinatorial optimization: algorithms and complexity*. [S.l.]: Courier Corporation, 1998.

PRACILIO, G.; ADAMS, M. L.; SMETTEM, K. R.; HARPER, R. J. Determination of spatial distribution patterns of clay and plant available potassium contents in surface soils at the farm scale using high resolution gamma ray spectrometry. *Plant and Soil*, Springer, v. 282, n. 1-2, p. 67–82, 2006.

RATSCHEK, H.; ROKNE, J. Experiments using interval analysis for solving a circuit design problem. *Journal of Global Optimization*, Springer, v. 3, n. 4, p. 501–518, 1993.

RESENDE, M. G.; FEO, T. A. A grasp for satisfiability. In: CITESEER. *CLIQUE, COLORING, AND SATISFIABILITY: THE SECOND DIMACS IMPLEMENTATION CHALLENGE, VOLUME 26 OF DIMACS SERIES ON DISCRETE MATHEMATICS AND THEORETICAL COMPUTER SCIENCE*. [S.l.], 1996.

REYES, D.; SAVELSBERGH, M.; TORIELLO, A. Vehicle routing with roaming delivery locations. *Transportation Research Part C: Emerging Technologies*, Elsevier, v. 80, p. 71–91, 2017.

ROSSBACH, M.; HIEP, N. T. Prompt gamma cold neutron activation analysis applied to biological materials. *Fresenius' journal of analytical chemistry*, Springer, v. 344, n. 1-2, p. 59–62, 1992.

SALMON, L. Analysis of gamma-ray scintillation spectra by the method of least squares. *Nuclear Instruments and Methods*, Elsevier, v. 14, p. 193–199, 1961.

SALVAT, F.; FERNÁNDEZ-VAREA, J. M.; SEMPAY, J. Penelope-2006: A code system for monte carlo simulation of electron and photon transport. In: NUCLEAR ENERGY AGENCY, ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT BARCELONA, SPAIN. *Workshop proceedings*. [S.l.], 2006. v. 4, n. 7.

SCHOEN, M. P. Intelligent optimization techniques, genetic algorithms, tabu search, simulated annealing, and neural networks, dt pham and d. karaboga, springer: Berlin, heidelberg, new york; springer london: London, 2000, 302pp, isbn 1-85233-028-7. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, Wiley Online Library, v. 15, n. 6, p. 287–289, 2005.

SCHOPENHAUER, A. *Fragmentos para a história da filosofia*. [S.l.]: Editora Iluminuras Ltda, 2003.

SEMPAU, J.; ACOSTA, E.; BARO, J.; FERNÁNDEZ-VAREA, J.; SALVAT, F. An algorithm for monte carlo simulation of coupled electron-photon transport. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, Elsevier, v. 132, n. 3, p. 377–390, 1997.

SHAHABINEJAD, H.; HOSSEINI, S.; SOHRABPOUR, M. A new neutron energy spectrum unfolding code using a two steps genetic algorithm. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 811, p. 82–93, 2016.

SHAHABINEJAD, H.; SOHRABPOUR, M. A novel neutron energy spectrum unfolding code using particle swarm optimization. *Radiation Physics and Chemistry*, Elsevier, v. 136, p. 9–16, 2017.

SHAHABINEJAD, H.; VOSOUGHI, N. Analysis of complex gamma-ray spectra using particle swarm optimization. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 911, p. 123–130, 2018.

SHAHABINEJAD, H.; VOSOUGHI, N. Sgsd: A novel sequential gamma-ray spectrum deconvolution algorithm. *Annals of Nuclear Energy*, Elsevier, v. 132, p. 369–380, 2019.

SHULTIS, J. K.; FAW, R. E. *Fundamentals of Nuclear Science and Engineering Third Edition*. [S.l.]: CRC press, 2016.

SHYU, C.; GARDNER, R.; VERGHESE, K. Development of the monte carlo library least-squares method of analysis for prompt gamma-ray analyzers. *Transactions of the American Nuclear Society*, v. 65, n. 1, p. 12–14, 1992.

SÖLLRADL, S. *Developments in prompt gamma-ray neutron activation analysis and cold neutron tomography and their application in non-destructive testing*. Tese (Doutorado) — Verlag nicht ermittelbar, 2014.

SPYCHALA, M.; MICHAELIS, W.; FANGER, H. Prompt gamma-ray neutron activation analysis for multi-element determination in sediment samples. *Journal of Radioanalytical and Nuclear Chemistry*, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV . . . , v. 112, n. 2, p. 331–339, 1987.

STINNETT, J.; SULLIVAN, C. J. An automated isotope identification algorithm using bayesian statistics. In: IEEE. *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*. [S.l.], 2013. p. 1–7.

STRANG, G. The fundamental theorem of linear algebra. *The American Mathematical Monthly*, Taylor & Francis, v. 100, n. 9, p. 848–855, 1993.

SULLIVAN, C. J.; MARTINEZ, M.; GARNER, S. Wavelet analysis of sodium iodide spectra. *IEEE Transactions on Nuclear Science*, IEEE, v. 53, n. 5, p. 2916–2922, 2006.

THEIL, H. *Economics and information theory*. [S.l.], 1967.

TSAI, L.-W.; MORGAN, A. P. Solving the kinematics of the most general six-and five-degree-of-freedom manipulators by continuation methods. *Journal of Mechanisms, Transmissions, and Automation in Design*, American Society of Mechanical Engineers, v. 107, n. 2, p. 189–200, 1985.

TSOULFANIDIS, N. *Measurement and detection of radiation*. [S.l.]: CRC press, 2010.

VILCHES, M.; GARCIA-PAREJA, S.; GUERRERO, R.; ANGUIANO, M.; LALLENA, A. Monte carlo simulation of the electron transport through air slabs: A comparative study of penelope, geant3, geant4 and egsnrc monte carlo codes. *IEEE Transactions on Nuclear Science*, IEEE, v. 55, n. 2, p. 710–716, 2008.

WALES, D. J.; SCHERAGA, H. A. Global optimization of clusters, crystals, and biomolecules. *Science*, American Association for the Advancement of Science, v. 285, n. 5432, p. 1368–1372, 1999.

WANG, J.; LI, F.; GARDNER, R. P. On the use of prompt gamma-ray neutron activation analysis for determining phase amounts in multiphase flow. *Measurement Science and Technology*, IOP Publishing, v. 19, n. 9, p. 094005, 2008.

WILDE, P. Neutron-induced prompt gamma activation analysis (pgaa) of metals and non-metals in ocean floor geothermal vent-generated samples. *J. Anal. At. Spectrom*, v. 17, p. 32–37, 2002.

YONEZAWA, C.; MATSUE, H.; HOSHI, M. Multi-element analysis of environmental samples by cold and thermal guided neutron induced prompt gamma-ray measurement. *Journal of radioanalytical and nuclear chemistry*, Springer, v. 215, n. 1, p. 81–85, 1997.

YOSHIDA, E.; SHIZUMA, K.; ENDO, S.; OKA, T. Application of neural networks for the analysis of gamma-ray spectra measured with a ge spectrometer. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 484, n. 1-3, p. 557–563, 2002.

ZABARANKIN, M.; URYASEV, S.; MURPHEY, R. Aircraft routing under the risk of detection. *Naval Research Logistics (NRL)*, Wiley Online Library, v. 53, n. 8, p. 728–747, 2006.

ZAMBURLINI, M.; BYUN, S.; PEJOVIĆ-MILIĆ, A.; PRESTWICH, W.; CHETTLE, D. Evaluation of mcnp5 and egs4 for the simulation of in vivo strontium xrf measurements. *X-Ray Spectrometry: An International Journal*, Wiley Online Library, v. 36, n. 2, p. 76–81, 2007.

ZHANG, J.; TUO, X. Pgnaa neutron source moderation setup optimization. *arXiv preprint arXiv:1309.1308*, 2013.

ZHANG, Z.; CHONG, Y.; CHEN, X.; JIN, C.; YANG, L.; LIU, T. Pgnaa system preliminary design and measurement of in-hospital neutron irradiator for boron concentration measurement. *Applied Radiation and Isotopes*, Elsevier, v. 106, p. 161–165, 2015.

APÊNDICE A – NÚMERO DE CONDICIONAMENTO

A.1 EXEMPLO DEFINIÇÃO NÚMERO DE CONDICIONAMENTO

Uma ilustração do número de condicionamento: Considere o problema de se determinar as raízes do polinômio:

$$p(x) = (x - 1)^4,$$

para o qual as quatro raízes são reais e iguais a 1. Suponha uma pequena perturbação, da ordem de 10^{-8} , de tal modo que a equação a ser resolvida será:

$$(x - 1)^4 = 10^{-8}$$

Uma das raízes é $1 + 10^{-2}$, ou seja, a raiz real foi transformada de 1 para $1 + 10^{-2}$. Isso faz com que o valor da Equação 2.1 seja da ordem de 10^6 , configurando-o como mau condicionamento devido ao fato de ser muito maior que 1.

A.2 NÚMERO DE CONDICIONAMENTO SISTEMAS LINEARES

Considere **A** uma matriz não singular (ou seja, determinante não nulo, o que implica a existência e unicidade da inversa) e o seguinte sistema linear:

$$\mathbf{Ax} = \mathbf{b} \tag{A.1}$$

cuja solução exata é $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Perturba-se, então, o lado direito do sistema, de **b** para $\mathbf{b} + \delta\mathbf{b}$, permanecendo **A** inalterado. Representando a solução para o problema perturbado por $(\mathbf{x} + \delta\mathbf{x}_b)$, tem-se que:

$$\mathbf{A}(\mathbf{x} + \delta\mathbf{x}_b) = \mathbf{b} + \delta\mathbf{b} \tag{A.2}$$

A ocorrência de um ‘ δ ’ precedendo a nomeação do vetor **x** significa a perturbação, de modo que $\delta\mathbf{b}$ é a transformação no vetor **b** e $\delta\mathbf{x}_b$ é a perturbação na solução do problema

Sabendo que:

$$\begin{aligned} \mathbf{A}(\mathbf{x} + \delta\mathbf{x}_b) &= \mathbf{b} + \delta\mathbf{b} \\ \mathbf{Ax} + \mathbf{A}\delta\mathbf{x}_b &= \mathbf{b} + \delta\mathbf{b} \end{aligned} \tag{A.3}$$

e que $\mathbf{Ax} = \mathbf{b}$, implica que $\mathbf{A} \delta \mathbf{x}_b = \delta \mathbf{b}$, e

$$\delta \mathbf{x}_b = \mathbf{A}^{-1} \delta \mathbf{b} \quad (\text{A.4})$$

Usando propriedades de normas, obtém-se um limite superior para $\|\delta \mathbf{x}_b\|$

$$\|\delta \mathbf{x}_b\| \leq \|\mathbf{A}^{-1}\| \|\delta \mathbf{b}\| \quad (\text{A.5})$$

onde a igualdade é verificada para um vetor $\delta \mathbf{b}$. A Equação A.5 mostra que o tamanho da transformação de $\mathbf{Ax} = \mathbf{b}$ resultante a partir da transformação de $\delta \mathbf{b}$ no lado direito pode ser tão grande quanto $\|\mathbf{A}^{-1}\|$ multiplicado por $\|\delta \mathbf{b}\|$.

Pela mesma propriedade que permitiu escrever a Equação A.5, também é possível escrever a seguinte expressão, a partir da Equação A.1:

$$\|\mathbf{b}\| \leq \|\mathbf{A}\| \|\mathbf{x}\| \quad (\text{A.6})$$

onde a igualdade se verifica para certo valor de $\|\mathbf{x}\|$.

Multiplicando a Equação A.5 e A.6, é possível obter:

$$\|\mathbf{b}\| \|\delta \mathbf{x}_b\| \leq \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \|\mathbf{x}\| \|\delta \mathbf{b}\| \quad (\text{A.7})$$

Multiplicando a Equação A.7 por $\frac{1}{\|\mathbf{x}\|}$ (a fim de obter a perturbação relativa), e rearrumando os termos, tem-se:

$$\frac{\|\delta \mathbf{x}_b\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}^{-1}\| \|\mathbf{A}\| \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|} \quad (\text{A.8})$$

Ou seja, a transformação relativa na exata solução (lado esquerdo da equação) tem um limite superior quando Equação A.8 obedecer a igualdade para algum vetor $\mathbf{b}$ e $\mathbf{x}$.

Perturbando-se a matriz $\mathbf{A}$, mantendo-se $\mathbf{b}$ fixo :

$$(\mathbf{A} + \delta \mathbf{A})(\mathbf{x} + \delta \mathbf{x}_a) = \mathbf{b} \quad (\text{A.9})$$

O subscrito 'a' na variável $\mathbf{x}$ indica a transformação na matriz $\mathbf{A}$. Assume-se que $\|\delta \mathbf{A}\|$ pequeno suficiente de modo a fazer com que $\mathbf{A}$ permaneça não singular. Um processo similar ao da Equação A.3 produz:

$$\begin{aligned}
\mathbf{b} &= \mathbf{A}\mathbf{x} + \mathbf{A}\delta\mathbf{x}_A + \delta\mathbf{A}\mathbf{x} + \delta\mathbf{A}\delta\mathbf{x}_A \\
\mathbf{A}\delta\mathbf{x}_A &= -\delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x}_A) \\
\delta\mathbf{x}_A &= \mathbf{A}^{-1}(-\delta\mathbf{A}(\mathbf{x} + \delta\mathbf{x}_A))
\end{aligned} \tag{A.10}$$

pela propriedade de normas, é possível escrever, a partir da Equação A.10:

$$\begin{aligned}
\|\delta\mathbf{x}_A\| &\leq \|\mathbf{A}^{-1}\| \|\delta\mathbf{A}\| \|\mathbf{x} + \delta\mathbf{x}_A\| \\
\frac{\|\delta\mathbf{x}_A\|}{\|\mathbf{x} + \delta\mathbf{x}_A\|} &\leq \|\mathbf{A}^{-1}\| \|\delta\mathbf{A}\| \\
\frac{\|\delta\mathbf{x}_A\|}{\|\mathbf{x} + \delta\mathbf{x}_A\|} &\leq \|\mathbf{A}^{-1}\| \|\mathbf{A}\| \frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}
\end{aligned} \tag{A.11}$$

O valor $\|\mathbf{A}^{-1}\| \|\mathbf{A}\|$ aparece nas Equações A.8 e A.11, e determina a máxima possível transformação relativa na solução exata do sistema linear por meio da transformação dos dados.

Rearrmando as Equações A.8 e A.11, obtém-se:

$$\text{cond}(\mathbf{A}) \geq \frac{\|\delta\mathbf{x}_b\| / \|\mathbf{x}\|}{\|\delta\mathbf{b}\| / \|\mathbf{b}\|} \quad e \quad \text{cond}(\mathbf{A}) \geq \frac{\|\delta\mathbf{x}_A\| / \|\mathbf{x} + \delta\mathbf{x}_A\|}{\|\delta\mathbf{A}\| / \|\mathbf{A}\|} \tag{A.12}$$

Equação A.12 mostra que a transformação relativa na solução exata é muito maior que a perturbação em $\mathbf{b}$ ou $\mathbf{A}$.

A.3 EXEMPLO DE MAU CONDICIONAMENTO EM SISTEMAS LINEARES

Considere o seguinte exemplo após as definições das expressões contidas na Equação A.12 (GILL; MURRAY; WRIGHT, 1991).

$$\mathbf{A} = \begin{pmatrix} 0.550 & 0.423 \\ 0.484 & 0.372 \end{pmatrix} \tag{A.13}$$

A matriz $\mathbf{A}$ é mau condicionada. Considere o vetor $\mathbf{b}$ e seu vetor perturbado $\mathbf{b} + \delta\mathbf{b}$

$$\tilde{\mathbf{b}} = \begin{pmatrix} 0.12707 \\ 0.1228 \end{pmatrix} = \begin{pmatrix} 0.127 \\ 0.112 \end{pmatrix} + \begin{pmatrix} 0.00007 \\ 0.00028 \end{pmatrix} = \mathbf{b} + \delta\mathbf{b} \tag{A.14}$$

A solução para $\mathbf{Ax} = \mathbf{b}$ é $\mathbf{x} = (1, -1)^T$, mas a solução exata de:

$$\mathbf{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}} \quad (\text{A.15})$$

é

$$\tilde{\mathbf{x}} = (1.7, -1.91)^T \quad (\text{A.16})$$

A relativa perturbação, medida em norma infinita é, sabendo que a solução exata de $\mathbf{Ax} = \mathbf{b}$ é $\mathbf{x} = (1, -1)^T$, e a solução exata de $\mathbf{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ é $\tilde{\mathbf{x}} = (1.7, -1.91)^T$, de tal modo que $\delta\mathbf{x}_b = (0.7, -0.91)^T$.

$$\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} = \frac{0.00028}{0.127} \approx 2.2 \times 10^{-3} \text{ and } \frac{\|\delta\mathbf{x}_b\|}{\|\mathbf{x}\|} = 0.91 \quad (\text{A.17})$$

Esse valor de perturbação ($\delta\mathbf{b}$) resulta em uma transformação de mais de 400 vezes na solução. Neste caso, o número de condicionamento tem um valor mínimo de 400.

Também é possível observar o mau condicionamento fazendo uma pequena perturbação no elemento (2,1) da matriz $\mathbf{A}$.

$$\mathbf{A} + \delta\mathbf{A} = \begin{pmatrix} 0.550 & 0.423 \\ 0.483 & 0.372 \end{pmatrix} \quad (\text{A.18})$$

A solução exata de $\hat{\mathbf{x}}$ é:

$$(\mathbf{A} + \delta\mathbf{A})\hat{\mathbf{x}} = \mathbf{b} \quad (\text{A.19})$$

$\hat{\mathbf{x}} = (-0.4536, 0.8900)^T$, onde:

$$\|\delta\mathbf{x}_a\| = \|\hat{\mathbf{x}} - \mathbf{x}\| = 1.89 \quad (\text{A.20})$$

A relativa transformação na solução exata é aproximadamente duas mil vezes a transformação na matriz $\mathbf{A}$.

APÊNDICE B – MMQ

B.1 VETOR PRÓXIMO DE UM SUBESPAÇO

seja $\mathbf{w}$ um vetor qualquer em W . Então:

$$\mathbf{v} - \mathbf{w} = ((\mathbf{v} - \text{proj}_W \mathbf{v}) + (\text{proj}_W \mathbf{v} - \mathbf{w})) \quad (\text{B.1})$$

estando $\mathbf{w}$ e $\text{proj}_W \mathbf{v}$ em W , $\text{proj}_W \mathbf{v} - \mathbf{w}$ está em W . Antes de dar continuidade a Equação B.1, será provado que $\mathbf{v} - \text{proj}_W \mathbf{v}$ é ortogonal a todos os vetores em W (Prova B.2, considerando que o vetor $\mathbf{w}$ pode ser gerado por uma base $S = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ e que $\mathbf{w}_i \bullet \mathbf{w}_j = 0$, $i \neq j$, e $\mathbf{w}_i \bullet \mathbf{w}_i = 1$, $1 \leq i \leq m$. Considera-se também que: $\text{proj}_W \mathbf{v} = (\mathbf{v} \bullet \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{v} \bullet \mathbf{w}_2)\mathbf{w}_2 + \dots + (\mathbf{v} \bullet \mathbf{w}_m)\mathbf{w}_m$, tem-se, para um vetor $\mathbf{v} \in \mathbb{R}^n$):

$$\begin{aligned} \mathbf{u} &= \mathbf{v} - \text{proj}_W \mathbf{v} \\ \mathbf{u} \bullet \mathbf{w}_i &= (\mathbf{v} - \text{proj}_W \mathbf{v}) \bullet \mathbf{w}_i \\ \mathbf{u} \bullet \mathbf{w}_i &= \mathbf{v} \bullet \mathbf{w}_i - \text{proj}_W \mathbf{v} \bullet \mathbf{w}_i \\ \mathbf{u} \bullet \mathbf{w}_i &= \mathbf{v} \bullet \mathbf{w}_i - [(\mathbf{v} \bullet \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{v} \bullet \mathbf{w}_2)\mathbf{w}_2 + \dots + (\mathbf{v} \bullet \mathbf{w}_m)\mathbf{w}_m] \bullet \mathbf{w}_i \\ \mathbf{u} \bullet \mathbf{w}_i &= \mathbf{v} \bullet \mathbf{w}_i - (\mathbf{v} \bullet \mathbf{w}_i)(\mathbf{w}_i \bullet \mathbf{w}_i) \\ \mathbf{u} \bullet \mathbf{w}_i &= 0. \end{aligned} \quad (\text{B.2})$$

comprovando que qualquer vetor $\mathbf{v} - \text{proj}_W \mathbf{v}$, $\mathbf{v} \in \mathbb{R}^n$, é ortogonal a todos os vetores em W . Lembrando que os vetores $(\mathbf{v} - \text{proj}_W \mathbf{v}) \bullet (\text{proj}_W \mathbf{v} - \mathbf{w})$ são ortogonais, é possível retornar à B.1

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 &= (\mathbf{v} - \mathbf{w}) \bullet (\mathbf{v} - \mathbf{w}) \\ &= ((\mathbf{v} - \text{proj}_W \mathbf{v}) + (\text{proj}_W \mathbf{v} - \mathbf{w})) \bullet ((\mathbf{v} - \text{proj}_W \mathbf{v}) + (\text{proj}_W \mathbf{v} - \mathbf{w})) \\ &= \|\mathbf{v} - \text{proj}_W \mathbf{v}\|^2 + \|\text{proj}_W \mathbf{v} - \mathbf{w}\|^2 \end{aligned} \quad (\text{B.3})$$

Se $\mathbf{w} \neq \text{proj}_W \mathbf{v}$, então $\|\text{proj}_W \mathbf{v} - \mathbf{w}\|^2$ é positivo e :

$$\|\mathbf{v} - \mathbf{w}\|^2 > \|\mathbf{v} - \text{proj}_W \mathbf{v}\|^2 \quad (\text{B.4})$$

tem-se, então, que $\text{proj}_W \mathbf{v}$ é o vetor em W que minimiza $\|\mathbf{v} - \mathbf{w}\|^2$. Logo, minimiza $\|\mathbf{v} - \mathbf{w}\|$

B.2 POSTO E SOLUÇÃO POR MÍNIMOS QUADRADOS

Considerando $r(\mathbf{A})$ o posto da matriz $\mathbf{A}$. Se $r(\mathbf{A}) = n$, então as colunas de $\mathbf{A}$ são linearmente independentes. A matriz $\mathbf{A}^T \mathbf{A}$ é invertível desde que o sistema linear $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{0}$ tenha apenas a solução trivial. Multiplicando ambos lados de $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{0}$ por $\mathbf{x}^T$ à esquerda, obtém-se:

$$\mathbf{0} = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = (\mathbf{A} \mathbf{x})^T (\mathbf{A} \mathbf{x}) = (\mathbf{A} \mathbf{x}) \bullet (\mathbf{A} \mathbf{x}) \quad (\text{B.5})$$

sabendo, a partir das propriedades do produto interno, que $\mathbf{u} \bullet \mathbf{u} > 0$ para $\mathbf{u} \neq \mathbf{0}$; $\mathbf{u} \bullet \mathbf{u} = 0$ se e somente se $\mathbf{u} = \mathbf{0}$, tem-se que $\mathbf{A} \mathbf{x} = \mathbf{0}$. Isso implica que há uma combinação linear nula das colunas linearmente independentes de $\mathbf{A}$, logo $\mathbf{x} = \mathbf{0}$, solução trivial. Portanto, $\mathbf{A}^T \mathbf{A}$ é invertível e a aproximação por mínimos quadrados tem solução única dada por $\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$

APÊNDICE C – METODOLOGIA - PSEUDOCÓDIGOS

C.1 ALGORITMO PRINCIPAL - GRASP

A presente seção define as linhas gerais para a definição do Algoritmo GRASP por meio do pseudocódigo 7. O início dos procedimentos descritos na presente seção são realizados após a definição dos picos dos elementos fornecidos, e da partição de maior resíduo para o elemento omitido.

Algoritmo 7 Algoritmo Grasp

```

1: input newTot, canais, esc, retorResi, comp, tamLis, iterMax, omt
2: output retor, part, aprox
3: metObj  $\leftarrow$  inf ;
4: metMax  $\leftarrow$  0 ;
5: li  $\leftarrow$  lista(empty) ;
6: iter  $\leftarrow$  1; tamLis  $\leftarrow$  size(li) ;
7: while {t  $\leq$  iterMax } do
8:   dif  $\leftarrow$  false ;
9:   ponMax = 0 ;
10:  while {dif = false} do
11:    xc  $\leftarrow$  const(inputCons) ;
12:    xc  $\leftarrow$  CGRASP(param xc) ;
13:    metXc  $\leftarrow$  metrica(xc) ;
14:    dif  $\leftarrow$  compare(li, xc) ;
15:    if {dif = true && j  $\leq$  tamLis} then
16:      li(j,1)  $\leftarrow$  xc ;
17:      if {metXc > metMax} then
18:        metMax  $\leftarrow$  metXc ;
19:        ponMax  $\leftarrow$  j ;
20:      end if
21:      if {metXc < metObj} then
22:        output  $\leftarrow$  xc
23:        metObj  $\leftarrow$  metXc ;
24:      end if
25:      j  $\leftarrow$  j + 1
26:    else
27:      if {metXc < metMax} then
28:        li(ponMax,1)  $\leftarrow$  xc ;
29:        [ponMax, metObj, metMax]  $\leftarrow$  update(li) ;

```

```

30:         end if
31:     end if
32: end while
33:  $ys \leftarrow \text{search}(xc, tamList, canPar, maxIter, ref)$ ;
34:  $ys \leftarrow \text{CGRASP}(\text{param } ys)$ ;
35:  $metYs \leftarrow \text{metrica}(ys)$ ;
36: if { $metYs < metMax$ } then
37:      $li(ponMax, 1) \leftarrow ys$ ;
38:      $[ponMax, metObj, meMax] \leftarrow \text{update}(li)$ ;
39: end if
40:  $ks \leftarrow \text{link}(ys, eleLi, canPar, minCha, pas, ref)$ ;
41:  $ks \leftarrow \text{CGRASP}(\text{param } ks)$ ;
42:  $metKs \leftarrow \text{metrica}(ks)$ ;
43: if { $metKs < metMax$ } then
44:      $li(ponMax, 1) \leftarrow ks$ ;
45:      $[ponMax, metObj, meMax] \leftarrow \text{val}(li)$ ;
46: end if
47:  $ws \leftarrow \text{search}(ks, tamList, canPar, maxIter, ref)$ ;
48:  $ws \leftarrow \text{CGRASP}(\text{param } ws)$ ;
49:  $metWs \leftarrow \text{metrica}(ws)$ ;
50: if { $metWs < metMax$ } then
51:      $li(ponMax, 1) \leftarrow ws$ ;
52:      $[ponMax, metObj, meMax] \leftarrow \text{val}(li)$ ;
53: end if
54: end while

```

São enumerados como *input* para o Algoritmo 7: (1) os valores de contagem espectral total contidos (variável *newTot*); (2) os canais de cada uma das partições (variável *canais*); (3) as partições de referência contendo os fotopicos dos compostos inseridos na estrutura de dados *esc*; (4) estrutura de dados contendo parâmetros das partições na variável (variável *retorResi*); (5) as bibliotecas de contagem espectral (variável *comp*); (6) definição do tamanho da lista para o GRASP (variável inteira *tamLis*); (7) número máximo de iterações (variável *iterMax*) para o algoritmo e (8) valor inteiro que indica que biblioteca deve ser omitida.

O Algoritmo 7 trará como saída:

1. Os parâmetros de cada uma das partições, agrupados na variável *retor*:
 - a) Intervalo de canais das partições.
 - b) Área aproximada para a partição em questão (variável *Aprox*).

- c) Matriz de compostos da partição, de acordo com a Equação 4.3 (variável *vetComp*).
 - d) Vetor das contagens espectrais totais (variável *vetTot*).
 - e) Valor de contagem do composto suprimido (variável *heighRet*).
 - f) Valores de máximo e mínimo da contagem total para o intervalo da partição. Esse valor é útil no sentido de determinar a qualidade da aproximação da biblioteca omitida.
2. Frações de contagem percentual dos compostos no particionamento dos dados (variável *part*).
 3. Frações de contagem percentual aproximadas pelos mínimos quadrados na omissão de uma biblioteca do espectro total (variável *aprox*).

Logo, pode-se concluir que o Algoritmo GRASP pode ser sumarizado por:

$$[retor, part, aprox] = f(newTot, canais, esc, retorResi, comp, tamLis, iterMax) \quad (C.1)$$

A variável que irá identificar a melhora durante a execução é iniciada na linha 3. Soluções com valores de métrica menor do que o valor inicializado serão retornados pelo algoritmo, com respectivos valores de percentuais. Na linha 4 será inicializada a variável responsável pela substituição de soluções na lista. Na linha 5 inicia-se a lista, vazia, com tamanho n . Linha 6 tem-se a variável que contabiliza as iterações do algoritmo, juntamente com a variável inteira que indica o tamanho da lista associada ao algoritmo.

Inicializa a variável *dif* (linha 8, Algoritmo 7) para verificar se uma nova solução construída é suficientemente diferente das soluções até então definidas. Na linha 9 tem-se início à variável *ponMax* que servirá como ponteiro para alocar na lista uma solução melhor do que a pior solução da lista.

Calcula-se a métrica associada à solução construída (linha 13, Algoritmo 7) e verifica-se se essa solução é suficientemente diferente das soluções da lista (linha 14). Sendo verdadeiro, insere-se a solução na lista (linha 16) e os ponteiros (*metMax*, *ponMax*, *metObj* e *j*) são atualizados (linhas 18, 19, 23 e 25).

Se a lista estiver preenchida, a execução do algoritmo ocorrerá a partir da linha 27. Da linha 27 à linha 30 ele irá inserir a nova solução construída na lista; caso ela seja melhor do que a pior solução até então no interior da lista, ela atualizará os valores de todos os ponteiros

(*ponMax*, *metObj*, *metMax*, linha 29), fornecendo a lista *li* como parâmetro de entrada para a função *update*.

Passado o processo de construção, tem-se início o processo de busca (linha 33). Para essa sub-rotina, listam-se os seguintes parâmetros de entrada: (1) a solução advinda da construção *xc*; (2) tamanho da lista associada à busca *tamList*; (3) os canais inseridos em cada uma das partições *canPar*; (4) o número máximo de execuções da busca (*maxIter*) e (5) as partições de referência (*ref*). Após a execução da busca, tem-se a execução do CGRASP (34). Das linhas 36 a 39 verifica-se se o resultado da busca é melhor do que a pior solução presente na lista, e atualizam-se os ponteiros (linha 38). A solução encontra nessa iteração substitui a pior solução da lista, se a nova solução for melhor.

O mesmo processo descrito para a primeira busca nas vizinhanças da solução construída pode ser enumerado para o processo de “linkagem”: tem-se como parâmetro de entrada: (1) a solução vinda da busca (*ys*); (2) as soluções inseridas na lista para que seja escolhida aleatoriamente uma delas e comparada com a solução da primeira busca; (3) os canais inseridos em cada uma das partições (variável *canPar*); (4) o número mínimo de canais que uma nova solução vinda da linkagem deve ter (variável *minCha*); (5) a quantidade de vezes a ser alterada a diferença de altura das soluções *start* e *target* na partição de maior resíduo (variável *passo*) e (6) as partições de referência (*ref*). Após isso atualizam-se os ponteiros (linha 45).

Na linha 47 inicia-se o segundo processo de busca e atualização das variáveis de ponteiro na linha 52.

C.2 DETECÇÃO DE FOTOPICOS

No Algoritmo 8 é apresentada a abordagem para detecção de picos e da partição de maior resíduo. Na linha 6 do Algoritmo 8 para detecção de picos e partição de maior resíduo, os dados são divididos em partições de 60 canais. A função que secciona os dados, são fornecidos: (1) o tamanho da matriz de bibliotecas de compostos contendo as contagens espectrais e (2) o número de canais que se deseja em cada partição. As partições que contiverem picos de compostos são atribuídas à variável *esc* (7). O maior resíduo e a respectiva partição são determinados na linha 8.

Algoritmo 8 Algoritmo Resíduos

- 1: **input** {*compostos*}
- 2: **output** {*est*}
- 3: $pd \leftarrow \text{deri}(\text{compostos})$

```

4: sd ← deri2(pd)
5: pi ← picos(pd, sd)
6: [mu, ca] ← mult(tamComp, canPart)
7: esc ← ref(canaís)
8: est ← resi(esc, canaís, compostos, total)

```

C.3 PASSO CONSTRUTIVO GRASP

Na presente seção há uma descrição detalhada do pseudocódigo da etapa construtiva do GRASP.

Algoritmo 9 Algoritmo Build Grasp

```

1: input partRef, estPart, comp, tot, can, numMin, col
2: output dat
3: for { i ∈ partRef } do
4:   chaStar ← randChanOfPart(partRef(i,1))
5:   maxChanPart ← greatChannelPart(partRef(i,1))
6:   j ← firtsChannelPart(partRef(i,1))
7:   while { j ∈ partRef(i,1) && j ≤ maxChanPart } do
8:     j ← randi(partRef(i,1))
9:     if { j ≤ terPart && (j − chaStar) ≥ numMin } then
10:      ter ← j
11:    end if
12:  end while
13:  temp ← defPart( chaStar, ter, tot, comp, can, partMaxRes )
14:  est(i,1) ← temp
15: end for
16: dat ← data(est)

```

Na linha 3, itera-se sobre cada uma das partições de referência, fornecendo como parâmetro de entrada a estrutura de dados contendo as informações delas. Na linha 4 escolhe-se aleatoriamente um canal para definir o início da partição considerada. A linha 5 define-se a variável que receberá o inteiro que indexa o maior canal da partição em consideração. Na linha 6, inicializa-se um ponteiro *j* que receberá o inteiro que indexa o início da partição em destaque. O valor do inteiro que definirá o último canal da partição está definido na linha (8). Escolhido esse valor de término *j* da janela, define-se a partição (14) criando um estrutura de dados contendo:

C.4 PASSO DE BUSCA GRASP

Algoritmo 10 Algoritmo Search Grasp

```

1: input cons, tamLis, maxIter, ref, col, minRes ;
2: output resSear ;
3: output  $\leftarrow$  cons ;
4: lis  $\leftarrow$  createLis(tamLis) ;
5: metMin  $\leftarrow$  minRes ;
6: while { notEmptyList } do
7:   lis  $\leftarrow$  createLis(tamLis) ;
8:   t  $\leftarrow$  1 ;
9:   iter  $\leftarrow$  1 ;
10:  while { t  $\leq$  tamLis && iter  $<$  maxIter } do
11:    copConsA1  $\leftarrow$  cons
12:    varPartI  $\leftarrow$  varAlt(copConsA1) ;
13:    metri  $\leftarrow$  metrica(varAltPart) ;
14:    if { metri  $<$  metMin && t  $\leq$  tamLis } then
15:      lista(t,1)  $\leftarrow$  copConsA1 ;
16:      metMin  $\leftarrow$  metri ;
17:      t  $\leftarrow$  t + 1 ;
18:    end if
19:    copConLar  $\leftarrow$  cons
20:    varLargPart  $\leftarrow$  enlarge(copConLar) ;
21:    metri  $\leftarrow$  metrica(varLargPart) ;
22:    if { metri  $<$  metMin && t  $\leq$  tamLis } then
23:      lista(t,1)  $\leftarrow$  copConLar ;
24:      metMin  $\leftarrow$  metri ;
25:      t  $\leftarrow$  t + 1 ;
26:    end if
27:    iter  $\leftarrow$  iter + 1 ;
28:  end while
29:  if { notEmptylis } then
30:    newSolu  $\leftarrow$  randi (lis) ;
31:    resSear  $\leftarrow$  newSolu ;
32:  end if
33: end while

```

Pode-se destacar como dados de entrada na rotina 10: (1) Os *parâmetros da construção* (variável *cons*); (2) Tamanho da lista associada à execução da busca local (variável *tamLis*); (3) número máximo de iterações para a busca local (variável *maxIter*); (4) matriz de canais indexando as partições que servirão de referência (variável *ref*); (5) variável inteira indexando

a biblioteca a ser omitida e posteriormente aproximada (variável *col*); (6) valor mínimo de métrica a ser perseguida pela busca (variável *minRes*).

Na linha 4 é inicializada a lista da busca, juntamente com a variável ponteiro que verificará as melhoras na execução do algoritmo (linha 5).

Da linha 6 a linha 32, serão executadas as iterações do algoritmo em busca de soluções melhores nas vizinhanças da solução atual. Essas linhas serão executadas enquanto houver elementos na lista melhor do que a solução corrente. Na linha 7, reinicializa-se a lista para a ocorrência de uma nova iteração. Linha 8 e 9, respectivamente, tem-se um ponteiro para verificar qual a última posição da lista preenchida e o número de ordem da iteração.

Na linha 11 faz-se uma cópia da solução corrente, para que as modificações da solução atual ocorram nessa cópia, a cada iteração do *loop* que se inicia na 10. Isso se faz necessário para que as modificações sejam únicas a cada iteração da iteração da linha 10.

A definição de vizinhos no Algoritmo 10 será dado por um par de valores (*alt*, *larg*), onde *alt* é o valor da altura e *larg* é a janela de canais utilizada na definição biblioteca omitida do espectro total, ambos na partição de maior resíduo. A definição de vizinhos da solução corrente estão definidos nas figuras 15 e 14.

Na linha 12 ocorre uma modificação na altura do composto suprimido na partição de maior resíduo, com novo cálculo de métrica na linha 13, após a modificação. Se essa modificação faz essa nova solução ser classificada como melhor que a solução corrente (linha 14), ela será inserida na lista de melhoras da busca, e os ponteiros de verificação serão atualizados (ponteiro de métrica mínima, linha 16 e ponteiro de tamanho da lista, linha 17). Na linha 20 faz-se uma modificação na largura da partição de maior resíduo, seguindo-se ao cálculo de nova métrica após modificação (linha 21). Nas linhas 22 a 26 faz-se o mesmo procedimento que o realizado para as linhas 14 a 18.

Chegando-se a linha 29, verifica-se se a lista não está vazia. Se essa condição for satisfeita, escolhe-se aleatoriamente um elemento da lista na linha 30 para ser a nova solução corrente. A partir daí, tem-se início ao processo de linkagem (Algoritmo 11)

C.5 PASSO DE LINKAGEM GRASP

C.5.1 Linkagem Canais GRASP

Algoritmo 11 Algoritmo Link Grasp

1: **input** *start*, *target*, *numMin*, *ref*, *col*, , *minMet* ;

```

2: output resul;
3: metStar  $\leftarrow$  metrica(start); metTar  $\leftarrow$  metrica(target); metMin  $\leftarrow$  metStar;
4: if {metStar < metTar} then
5:   temp  $\leftarrow$  target; target  $\leftarrow$  start; start  $\leftarrow$  target; metMin  $\leftarrow$  metStar
6: end if
7: channSt  $\leftarrow$  findChann(start); channTr  $\leftarrow$  findChann(target);
8: StarNTr  $\leftarrow$  dif(channSt, channTr); TrNStar  $\leftarrow$  dif(channTr, channSt);
9: difTo  $\leftarrow$  StarNTr + TrNStar
10: lista  $\leftarrow$  newList; positi  $\leftarrow$  1;
11: while { difTo > 0 } do
12:   copStar  $\leftarrow$  start;
13:   channSt  $\leftarrow$  findChann(start); channTr  $\leftarrow$  findChann(target);
14:   StarNTr  $\leftarrow$  dif(channSt, channTr); TrNStar  $\leftarrow$  dif(channTr, channSt);
15:   difTo  $\leftarrow$  StarNTr + TrNStar; list  $\leftarrow$  newList; positi  $\leftarrow$  1;
16:   for { i  $\in$  StarNtr == 0 } do
17:     newSolu  $\leftarrow$  retChan(start); metNewSol  $\leftarrow$  metrica (newSolu);
18:     if {metNewSol < metStar} then
19:       list(positi)  $\leftarrow$  newSolu; positi  $\leftarrow$  positi + 1;
20:       if { metNewSol < metRef } then
21:         resul  $\leftarrow$  newSolu; metRef  $\leftarrow$  metNewSol;
22:       end if
23:     end if
24:   end for
25:   for { i  $\in$  TrNStar == 0 } do
26:     newSolu  $\leftarrow$  insChan(start); metNewSol  $\leftarrow$  metrica (newSolu);
27:     if {metNewSol < metStar} then
28:       list(positi)  $\leftarrow$  newSolu; positi  $\leftarrow$  positi + 1;
29:       if { metNewSol < metRef } then
30:         resul  $\leftarrow$  newSolu; metRef  $\leftarrow$  metNewSol;
31:       end if
32:     end if
33:   end for
34:   if { lista (empty) } then
35:     start  $\leftarrow$  randiElement(lista);
36:   end if
37:   difTo  $\leftarrow$  difTo - 1;
38: end while

```

Na linha 3 tem-se o cálculo da métrica associada à solução *start*, à solução *target*, além da atribuição do valor da métrica da solução *start* à métrica mínima.

Da linha 4 à 6, tem-se a inversão das soluções, caso a solução *start* possua um valor de métrica menor que o valor de métrica da solução *target*, com consequente atualização da métrica mínima.

Na linha 7, tem-se a determinação dos canais que estão inseridos em cada uma das duas soluções. Na variável *StarNTr* (Algoritmo 11, linha 8), tem-se um vetor que conterá os valores

0 para um canal arbitrário que não pertença à solução *target* e pertença à solução *start*, e valor 1 caso contrário. Já a variável *TrNStar* informa os canais presentes na solução *target* e ausentes na solução *start*, inserindo 0 nas posições de um vetor que atendam a esse requisito, e 1 caso contrário.

Na linha 9, somam-se os valores presentes em 8, definindo-se, assim, a quantidade total de diferenças entre as duas soluções.

Na linha 10, inicia-se a lista associada ao algoritmo. Enumera-se, também, a declaração de um ponteiro que irá alocar soluções para as posições na lista.

Na linha 11, incia-se a iteração que será executado até esgotarem todas as diferenças entre as duas soluções. A linha 12 é executada toda vez que uma nova diferença entre as duas soluções é verificada. Uma diferença verificada é resetada ao executar a linha 12, verificando-se, assim, diferença por diferença.

As linhas 13, 14 e 15, semelhantemente às linhas 7, 8, 9 e 10, são reexecutadas para que uma nova solução *start* adotada ao fim de uma iteração tenha seus parâmetros redefinidos.

Da linha 16 à linha 24, é analisada uma diferença arbitrária presente na solução *start* e ausente na solução *target*. Nessa situação, pressupõe-se ser necessário retirar um canal presente em *start* na tentativa de aproximá-la de *target*. Esse retirada de canal é feita na linha 17, com consequente cálculo da métrica. Na linha 18 verifica-se se essa nova solução é melhor que a solução corrente, inserindo-se essa solução na lista (19), atualizando o ponteiro responsável pela introdução da solução. (19). Redefinem-se os ponteiros de verificação de melhor solução e de retorno da função (21).

Das linhas 25 à linha 33, os procedimentos são os mesmos que os tomados das linhas 16 a 24, com a diferença que os procedimentos agora são tomados no intuito de se adicionar à solução *start* os canais que estão presentes na solução *target*.

Escolhe-se uma nova solução para *start*, a partir da lista do algoritmo, nos comandos compreendidos entre as linhas 34 a 36. Feito isso, decrementa-se o número total de diferenças entre as duas soluções de entrada na linha 37.

C.5.2 Linkagem Altura GRASP

Algoritmo 12 Algoritmo LinkAlt Grasp

```

1: input start, target, passo, col, fun;
2: output resul;
3: resul  $\leftarrow$  target; hStart  $\leftarrow$  start.altura; hTarget  $\leftarrow$  target.altura;
4: disPasso  $\leftarrow$  (hStart - hTarget)/(passo); metStart  $\leftarrow$  metrica(start); metTarget  $\leftarrow$  metrica(target);

```

```

5: if {  $hStart > hTarget$  } then
6:   for {  $i = 1 : passo$  } do
7:      $h \leftarrow hStart - i * passo$ ;
8:     if {  $h > minPart \ \&\& \ h < maxPart$  } then
9:        $start.altura \leftarrow h$ ;  $newMet \leftarrow metrica(start)$ ;
10:      if {  $newMet < metTarget$  } then
11:         $resul \leftarrow start$ ;
12:      end if
13:    end if
14:  end for
15: else
16:   for {  $i = j : passo$  } do
17:      $h \leftarrow hStart + j * passo$ ;
18:     if {  $h > minPart \ \&\& \ h < maxPart$  } then
19:        $start.altura \leftarrow h$ ;  $newMet \leftarrow metrica(start)$ ;
20:       if {  $newMet < metTarget$  } then
21:          $resul \leftarrow start$ ;
22:       end if
23:     end if
24:   end for
25: end if

```

São definidos os seguintes valores na linha 3 do Algoritmo 12: a variável de saída é configurada como a solução *target*, são recuperados os valores de altura da solução *start* e *target*, respectivamente, nas variáveis *hStart* e *hTarget*.

Calculam-se os valores da distância de um passo (*disPasso*) e das métricas das duas soluções na linha 4.

Na linha 5, verifica-se se a solução *start* tem altura maior que a solução *target*. Essa condição sendo verificada, inicia-se na linha 6 o laço que irá modificar a altura da solução *start* por uma quantidade de *passo* vezes. A cada iteração desse laço, a altura será modificada em direção à altura da solução *target* por um valor de *h* (linha 7). A cada modificação de altura, estando essa nova altura dentro dos valores delimitados pela aproximação calculada na Seção 4.3, calcula-se o valor da métrica da solução *start* (9). Se o valor da nova altura fizer com que a nova solução seja a melhor, atribui-se essa solução ao resultado dessa subrotina (linha 11).

A descrição contida nas linhas 15 a 24 é idêntica à descrição das linhas 5 a 15, com a diferença que agora o procedimento tomará em conta a altura da solução *start* ser menor que a altura da solução *target*.

C.6 ALGORITMO DE ESTIMATIVA DAS CONTAGENS ESPECTRAIS

Algoritmo 13 Algoritmo estimativa suprimido

```

1: input star, end, comp, tot, estim, partRef
2: output estPart
3:  $mS \leftarrow \text{fazMat}(star, end, comp, esti, partRef)$ 
4:  $mT \leftarrow \text{fazMat}(star, end, tot, partRef)$ 
5:  $coorbar \leftarrow ((mS)^T \cdot (mS))^{-1} \cdot (mS)^T \cdot (mT)$ 
6: if  $\{anyCoorBar < 0\}$  then
7:   for  $\{u = 1 : tamCoorBar\}$  do
8:      $ponNeg \leftarrow mS(:,u)$ 
9:   end for
10: else
11:    $ponNeg \leftarrow mS(:,rand)$ 
12: end if
13:  $pont \leftarrow mS \cdot coorBar$ 
14:  $vet \leftarrow pont - ponNeg$ 
15:  $pass \leftarrow \text{const}$ 
16:  $stop \leftarrow \text{false}$ 
17:  $h \leftarrow 1$ 
18:  $listPont \leftarrow \text{newList}(\text{empty})$ 
19: while  $\{stop == false\}$  do
20:    $newPont \leftarrow ponNeg + passo \cdot h \cdot vet$ 
21:    $newCoorbari \leftarrow ((mS)^{-1} \cdot (mS)) \cdot (mS)^T \cdot newPont$ 
22:   if  $\{any(newCoorBari) < 0\}$  then
23:      $parou \leftarrow \text{true}$ 
24:   else
25:      $pontAr \leftarrow mS \cdot newCoorBari$ 
26:      $listPont(:,h) \leftarrow \text{ponto}$ 
27:   end if
28:    $h = h + 1$ 
29: end while
30:  $tamList = \text{size}(listPont)$ 
31:  $lisEst \leftarrow \text{newList}(\text{empty})$ 
32: for  $\{q = 1 : tamList\}$  do
33:    $desc \leftarrow ((mT - ((1 - percen) \cdot (listPont(q,1)))) / (percen))$ 
34:    $listEst(q,1) \leftarrow desc$ 
35: end for
36:  $tamListEst \leftarrow \text{size}(listEst)$ 
37:  $estPart \leftarrow \text{nemList}(\text{empty})$ 
38: for  $\{z = 1 : qtdPart\}$  do
39:    $min \leftarrow \min(listEst(:,w) / (tamPart(z)))$ 

```

```

40:   $max \leftarrow \max(\text{listEst}(:,w)/ (\text{tamPart}(z)))$ 
41:   $estPart \leftarrow [\min, \max]$ 
42:   $estPart(z) \leftarrow estPart$ 
43: end for

```

Na linha 3, é chamada uma função para construir a matriz de vetores de bibliotecas de contagens espectrais, passado a variável *starWide* de início e o de fim *endWide* da partição considerada. Esse método somará os valores de contagem espectral das bibliotecas fornecidas na soluções inicial do espectro total, nas partições de referência. Essa soma também será feita para a contagem espectral total e atribuída a variável mT (Algoritmo 13 linha 4).

Na linha 5 tem-se a aproximação dos mínimos quadrados para determinação das contagens de fração percentual no plano definido pelos vetores de compostos básicos.

Na linha 6, encontrado algum valor de fração percentual que seja negativo vindo a partir da aproximação dos mínimo quadrados, salva-se o vetor que está associado à coordenada baricêntrica negativa (linha 8); não acontecendo a última condição, é escolhido aleatoriamente um vetor de compostos para que se dê prosseguimento ao algoritmo (linha 11).

Na linha 13, calcula-se o ponto determinado pela projeção dos mínimos quadrados, no plano de bibliotecas fornecido, para posterior cálculo do vetor diretor (linha 14) que servirá de suporte para a definição dos pontos presentes na Figura 7d.

Na linha 15, define-se uma constante que será responsável por caminhar ao longo do segmento definido por *vet* (Algoritmo 13 linha 14). Na linha 16, irá inicializar uma variável que controlará o loop que se inicia na linha 19. No comando que consta na linha 17, tem-se uma variável inteira que determinará quantos passos *pass* serão dados a partir do vetor *vet* em busca de pontos que atendam os critérios contidos na linha 22. Na linha 18 é criada uma lista que armazene todos os valores de pontos que estarão compreendidos entre *pont* e *ponNeg*, iniciando a partir do ponto *ponNeg*.

Na linha 20, tem-se a determinação de um novo ponto a partir do *ponNeg*. Para esse novo ponto, determina-se os seus valores de coordenadas baricênticas (linha 21) em relação aos valores de vetores de compostos que estão no mesmo plano. Essa determinação se faz necessária para conhecer se esse novo ponto está compreendido na região determinado pelos vetores de compostos. Se esse novo ponto obedecer a esse critério, é armazenado na lista de pontos (linha 26).

Verifica-se o tamanho da lista na linha 30. Na linha 33, calcula-se o ponto desconhecido, tomando por base que esse ponto será dado por $desc \cdot (percen) + (1 - percen) * estim(q, 1) =$

mT . Essa estimativa será adicionada à lista contendo vários outros pontos estimados (linha 36).

Cada ponto estimado inserido em *listEst*, conterá valores de estimativa da partição considerada. Os valores de coordenadas de cada um dos pontos serão divididos pelos respectivos tamanhos da partição (linha 39 e 40), retornando os valores máximo e mínimo de estimativa para a partição após dividir pelo tamanho z da mesma. Os valores de máximo e mínimo para uma partição z serão armazenados *estPart*

C.7 ALGORITMO PRINCIPAL DO CGRASP

O algoritmo inicializa as variáveis ponteiros que identificarão os valores de melhor métrica (linha 3). Enumera-se também a estrutura de dados escolhida como variável de saída (*retur*, inicializada na linha 4). O algoritmo tem um valor limite de execuções, descrito na abertura do comando *for* na linha 6 (variável *numTimesToRun*). Inicia-se na linha 7 a variável ponteiro para o número de execuções sem nenhum aperfeiçoamento (*numIterNoImprov*). Das linhas 8 a 23 serão construídas soluções e realizadas buscas em torno dessas soluções por um número *maxIters* de vezes, variável de controle do comando *for* da linha 8. Se for encontrada uma nova solução que seja melhor que a solução corrente, nas linhas 13 a 15 trata-se de atribuir os resultados às variáveis ponteiros, e incrementar a variável *numIterNoImprov* na linha 17, caso não seja encontrada uma solução melhor do que a atual na presente iteração. Das linhas 19 a 22, diminui-se a densidade do grid de busca caso o número de tentativas em se melhorar a solução atual tenha sido superado. Recomenda-se que o número de *maxIters* seja superior ao número de *maxNumIterNoImpro*.

Algoritmo 14 Algoritmo CGrasp

```

1: input perce, n, fun, maxIters, maxNumIterNoImprov, numTimesToRun, maxDirToTry, alpha
2: output retur
3:  $\text{min} \leftarrow \text{inf}$ 
4: retur  $\leftarrow \text{outGrasp}$ 
5:  $h \leftarrow 1$ 
6: for  $\{i = 1 : \text{numTimesToRun}\}$  do
7:   numIterNoImprov  $\leftarrow 0$ 
8:   for  $\{i = 1 : \text{maxIters}\}$  do
9:      $xc \leftarrow \text{build}(x, \text{fun}, n, h, x.\text{max}, x.\text{min}, \text{alpha})$ 
10:     $yc \leftarrow \text{improve}(xc, \text{fun}, n, h, x.\text{min}, x.\text{max}, \text{maxDirToTry}, \text{qtdPart})$ 
11:     $\text{metYc} \leftarrow \text{metrica}(yc)$ 

```

```

12:     if { metYc < min } then
13:         output  $\leftarrow$  yc
14:         min  $\leftarrow$  metYc
15:         numIterNoImprov  $\leftarrow$  0
16:     else
17:         numIterNoImprov  $\leftarrow$  numIterNoImprov + 1
18:     end if
19:     if { numIterNoImprov > maxNumIterNoImprov } then
20:         h  $\leftarrow$  h/2
21:         numIterNoImprov  $\leftarrow$  0
22:     end if
23: end for
24: end for

```

C.8 ALGORITMO FASE CONSTRUTIVA CGRASP

Algoritmo 15 Algoritmo Build CGrasp

```

1: input  $x, fun, n, h, l, u, \alpha$ 
2: output  $x_j$ 
3: dimen  $\leftarrow$  [1 2 ... n]
4: i  $\leftarrow$  1
5: tamDimen  $\leftarrow$  size(dimen)
6: while {i <= tamDimen} do
7:     min  $\leftarrow$  +inf
8:     max  $\leftarrow$  - inf
9:     for {  $k == 1 : tamDimen$  } do
10:        if {  $i$  in dimen } then
11:            zi  $\leftarrow$  lineSearch (x, h, fun, i, n, l, u)
12:            gi  $\leftarrow$  f(zi)
13:            if { min > gi } then
14:                min  $\leftarrow$  gi
15:            end if
16:            if { max < gi } then
17:                max  $\leftarrow$  gi
18:            end if
19:        end if
20:    end for
21:    RCL  $\leftarrow$  0
22:    for {all  $i$  in dimen} do
23:        if {  $i$  in dimen && gi < (1- $\alpha$ ) · min +  $\alpha$  · max } then

```

```

24:         RCL  $\leftarrow$  RCL  $\cup$  { $i$ }
25:     end if
26: end for
27:      $j \leftarrow \text{randi}([1 \text{ tamRCL}])$ 
28:      $x_j \leftarrow z_i$ 
29:      $S \leftarrow S - \{j\}$ 
30: end while

```

Inicia-se permitindo que todas as dimensões da solução x sejam modificadas, definindo as dimensões inicialmente como não fixas. O processo construtivo irá operar sobre uma dimensão arbitrária, mantendo-se constante todas as outras. Na linha 4 inicia ponteiro i que indexa a dimensão pesquisada e na linha 5 é determinado o tamanho do conjunto de dimensões. A função *lineSearch* executará a modificação da contagem espectral da biblioteca não fornecida inicialmente ao espectro total nas partições consideradas e calculará o valor da métrica associada (linha 11). Forma-se uma lista de candidatos (RCL linhas 21 a 26) que contém as coordenadas não fixas i cujos valores z_i são menores que ou iguais a $\alpha \cdot \max + (1 - \alpha) \cdot \min$. Os valores de $\min$ e $\max$ são determinados como os valores, respectivamente, do menor e maior valor da métrica associado às dimensões do problema. No Algoritmo 15 essas definições encontram-se a partir da linha 13 à linha 19. O valor de α é arbitrariamente escolhido entre $\in [0, 1]$. Escolhe-se aleatoriamente uma coordenada a partir da RCL (linhas 27), atribui-se à saída x_j e fixa-se a coordenada j de x , assegurando a aleatoriedade na fase construtiva. O processo é continuado até que todas as coordenadas de x tenham sido fixadas.

C.9 ALGORITMO FASE BUSCA CGRASP

A melhor solução corrente é inicializada na linha 6. Na linha 7 faz-se a opção do número de direções como sendo o valor mínimo entre o número de direções total para uma dimensionalidade arbitrária n ($3^n - 1$) e um número passado como parâmetro na chamada da função (*maxTry*). Essa opção se faz necessária desde que o número de possíveis direções de busca pode ser altamente extenso. Na linha 8, inicializa-se o vetor D que armazenará inteiros que mapearão as diferentes direções da busca. O laço *while* que se inicia na linha 9 será executado enquanto houver melhoras na execução do algoritmo. A variável presente na linha 10 será reinicializada para que uma nova iteração ocorra em direção à solução melhor que a solução atual. O laço da linha 11 será executado enquanto o vetor ponteiro de direções for menor que a quantidade máxima de tentativas, ou enquanto não encontrar soluções de melhor métrica que a solução

corrente. Na linha 12 será escolhido, aleatoriamente, um número inteiro que não pertença ao vetor D , que será utilizado posteriormente no mapeamento de uma direção (linha 14). Na linha 16, após escolher-se uma partição na linha 15, altera-se a altura da partição escolhida, dependendo da direção. Sendo verdade que a nova altura da partição esteja compreendida entre valores máximo e mínimo (linha 17) e que o valor da função para a nova altura tenha métrica melhor que o valor da função para a solução corrente (linha 18), são executadas as linhas 19 a 21 em que: configura-se a variável de retorno *sai* como a partição com seu novo valor de altura (linha 19); o ponteiro que armazena as direção são setados todas as suas posições como zero (linha 20); a variável ponteiro *impro* é setada como *true* (linha 21).

Algoritmo 16 Algoritmo Improve CGrasp

```

1: input  $x, fun, n, h, l, u, maxTry$ 
2: output sai
3:  $min = inf$ 
4:  $impro \leftarrow true$ 
5:  $xb \leftarrow x$ 
6:  $sai \leftarrow f(xb)$ 
7:  $numTry \leftarrow \min\{3^n - 1, maxTry\}$ 
8:  $D \leftarrow \text{zeros}(numTry, 1)$ 
9: while  $\{impro\}$  do
10:    $impro \leftarrow false$ 
11:   while  $\{|D| < maxTry \ \&\& \ \text{not} impro\}$  do
12:      $r \leftarrow \text{randi}(1, 3^n - 1) \notin D$ 
13:      $D \leftarrow D \cup \{r\}$ 
14:      $dir \leftarrow \text{mapeia}(r)$ 
15:      $part \leftarrow \text{randi}(\text{partições})$ 
16:      $alt \leftarrow alt + h \cdot d$ 
17:     if  $\{l < alt < u\}$  then
18:       if  $\{fun(newAlt) < fun\}$  then
19:          $sai \leftarrow part(newAlt)$ 
20:          $D \leftarrow 0$ 
21:          $impro \leftarrow true$ 
22:       end if
23:     end if
24:   end while
25: end while

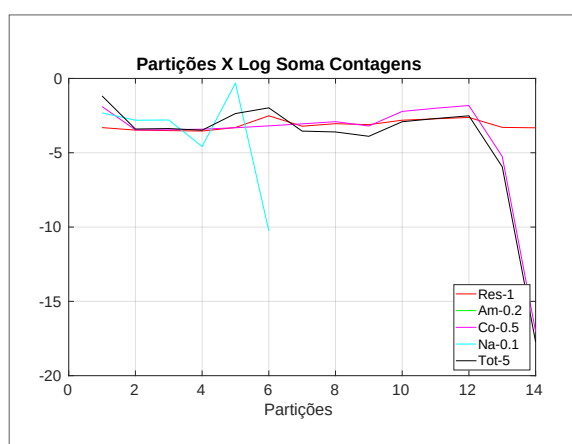
```

APÊNDICE D – DETERMINAÇÃO DOS RESÍDUOS DAS PARTIÇÕES

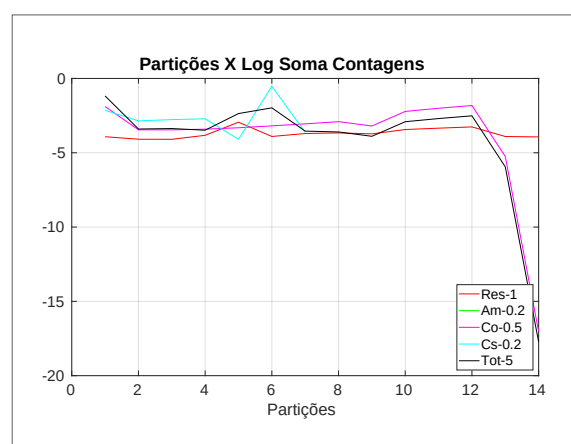
D.1 DETERMINAÇÃO DOS RESÍDUOS DAS PARTIÇÕES - ELEMENTOS

Figura 63 – Componentes dos vetores de compostos como função de partição. Os componentes do vetor total são apresentados (variável $tot - 5$), além dos valores de resíduos (variável $res - 1$).

(a) Componentes dos vetores de compostos (Am , Co e Na) $\times$ partição. Os componentes do vetor total são apresentados (variável $tot - 5$), e os valores dos resíduos (variável $res - 1$).



(b) Componentes dos vetores de compostos (Am , Co e Cs) $\times$ partição. Os componentes do vetor total são apresentados (variável $tot - 5$), além dos valores dos resíduos (variável $res - 1$).

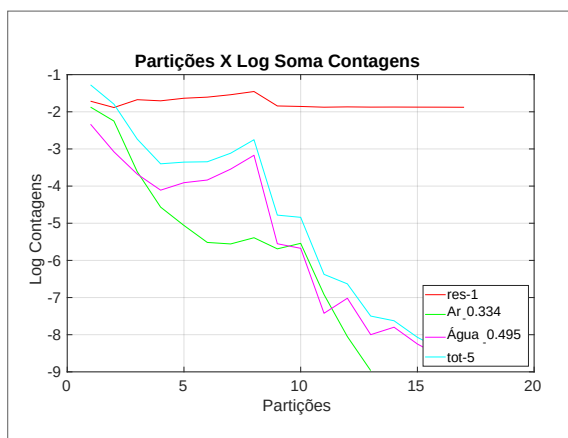


Fonte: Próprio Autor

D.2 DETERMINAÇÃO DOS RESÍDUOS DAS PARTIÇÕES - COMPOSTOS

Figura 64 – Resultados para determinação dos picos, Base de Dados 2.

- (a) Componentes dos vetores de compostos (Ar e Água) $\times$ partição. As componentes do vetor total são apresentados (variável *tot* – 5), e os valores dos resíduos (variável *res* – 1)



Fonte: Próprio Autor

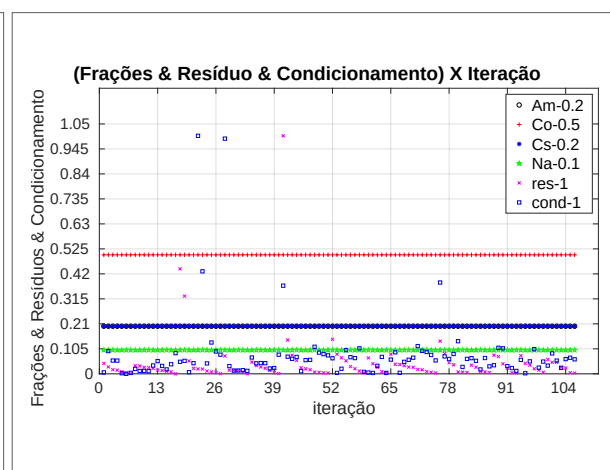
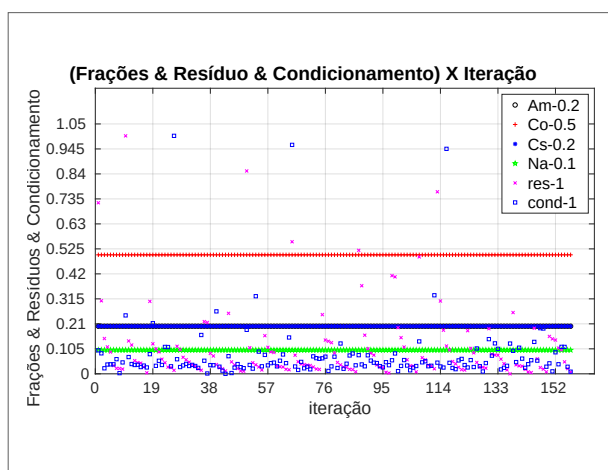
APÊNDICE E – DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - PARTICIONAMENTO

E.1 DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - PARTICIONAMENTO DOS ELEMENTOS

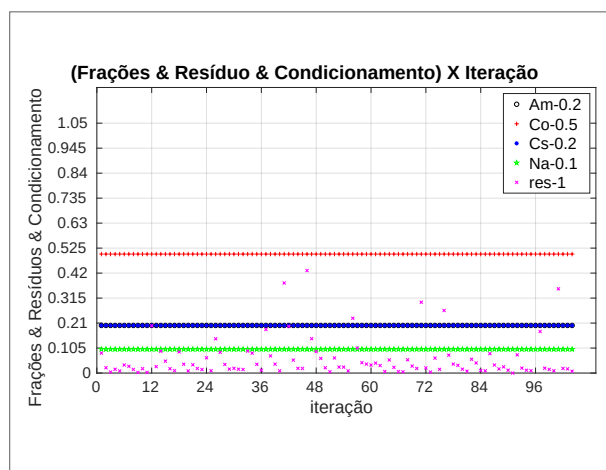
Figura 65 – Determinação quantitativa dos compostos por particionamento dos dados. Pontos (○), vermelho(+), azul (*) e verde (☆) indicam, respectivamente, as quantidades de *Am*, *Co*, *Cs* e *Na*. Os valores nominais são descritos nas legendas dos Gráficos 65a, 65b e 65c. Os valores dos resíduos são descritos nos pontos (✕), e o amarelo (□) indica o condicionamento normalizado das matrizes de covariância do particionamento.

(a) Frações de Contagem percentuais. As partições de referência juntamente com a partição de maior resíduo, instância 2. As contagens espectrais para o *Co* são reais.

(b) Frações de Contagem percentuais. As partições de referência juntamente com a partição de maior resíduo, instância 3. As contagens espectrais para o *Cs* são reais.



(c) Frações de Contagem percentuais. As partições de referência juntamente com a partição de maior resíduo, instância 4. As contagens espectrais para o *Na* são reais.

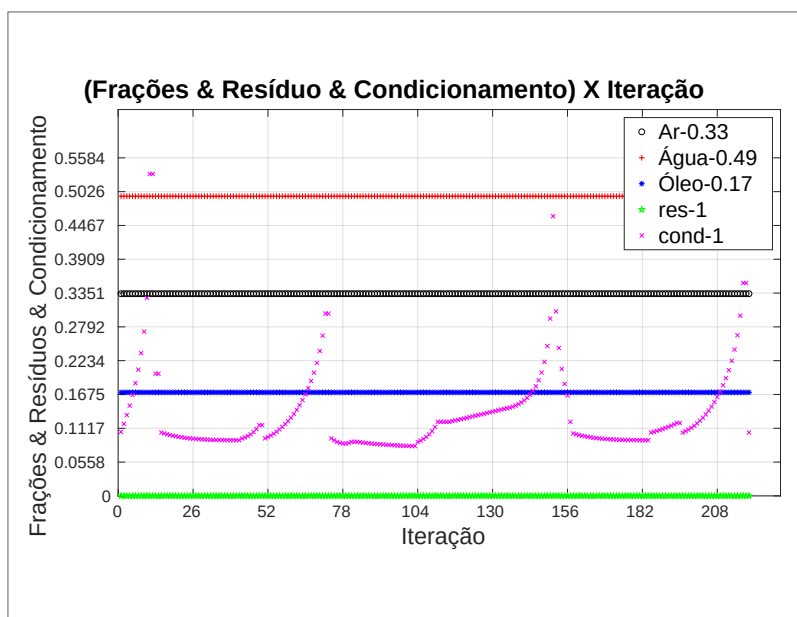


Fonte: Próprio Autor

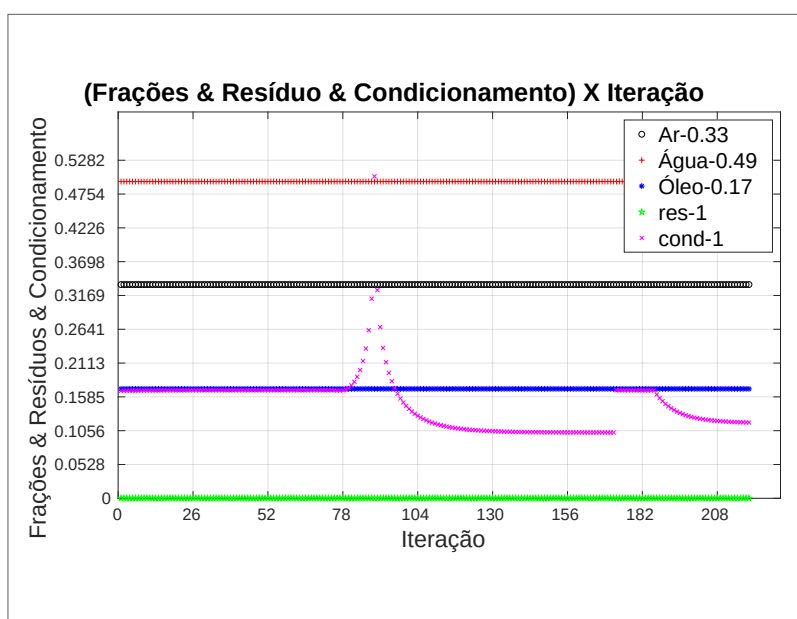
E.2 DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - PARTICIONAMENTO DOS COMPOSTOS

Figura 66 – Determinação quantitativa dos compostos por particionamento dos dados. Pontos (○, Ar), vermelho(+, Água) e azul (*, Óleo) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais estão descritos nas legendas dos Gráficos 66a e 66b. Os valores dos resíduos são descritos por pontos (☆), e o rosa (✕) indica o condicionamento normalizado das matrizes de covariância do particionamento.

- (a) Frações de Contagem percentuais. As partições de referência juntamente com a partição de maior resíduo, instância 6. As contagens espectrais para o Água são reais.



- (b) Frações de Contagem percentuais. As partições de referência juntamente com a partição de maior resíduo, instância 7. As contagens espectrais para o Óleo são reais.



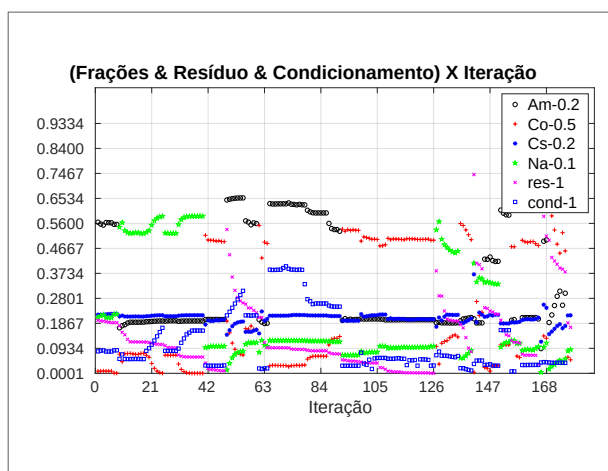
Fonte: Próprio Autor

APÊNDICE F – DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - APROXIMAÇÃO

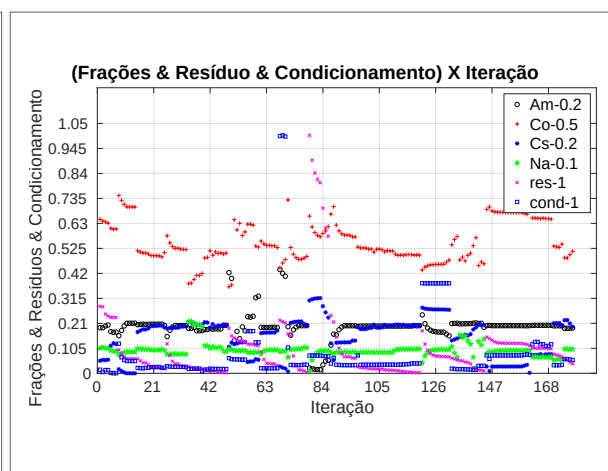
F.1 DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - APROXIMAÇÃO DOS ELEMENTOS

Figura 67 – Gráficos contendo as frações de contagens percentuais dos compostos após a omissão, e estimação, de uma biblioteca do espectro total. Pontos (○), vermelho (+), azul (*) e verde (☆) indicam, respectivamente, as quantidades de *Am*, *Co*, *Cs* e *Na*. Os valores dos resíduos são descritos nos pontos ×, e □ indica o condicionamento normalizado $[0, 1]$ das matrizes de covariância na aproximação.

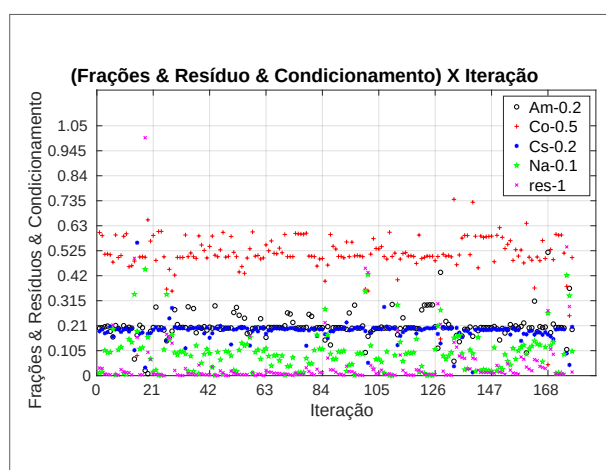
(a) Frações de contagens percentuais dos compostos, após supressão e estimativa do *Co* no espectro total.



(b) Frações de contagens percentuais dos compostos, após supressão e estimativa do *Cs* no espectro total.



(c) Frações de contagens percentuais dos compostos, após supressão e estimativa do *Na* no espectro total.

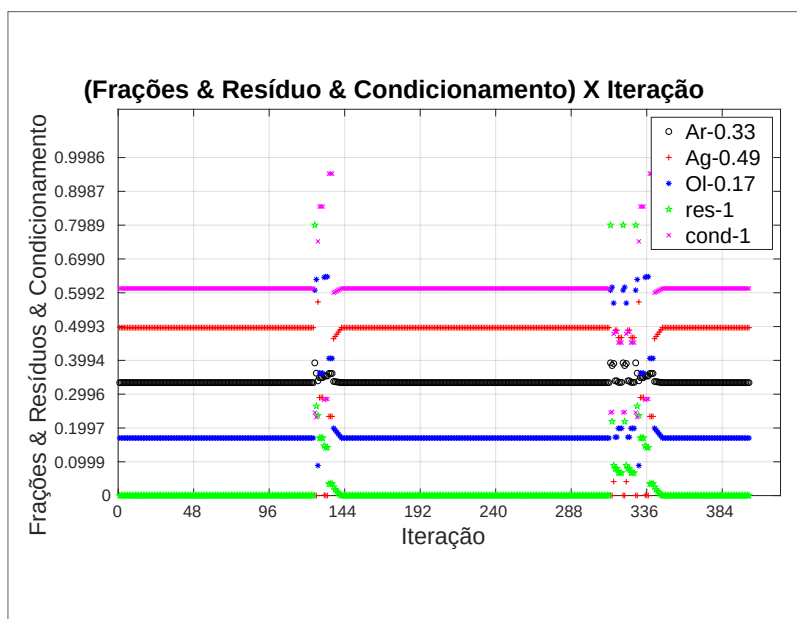


Fonte: Próprio Autor

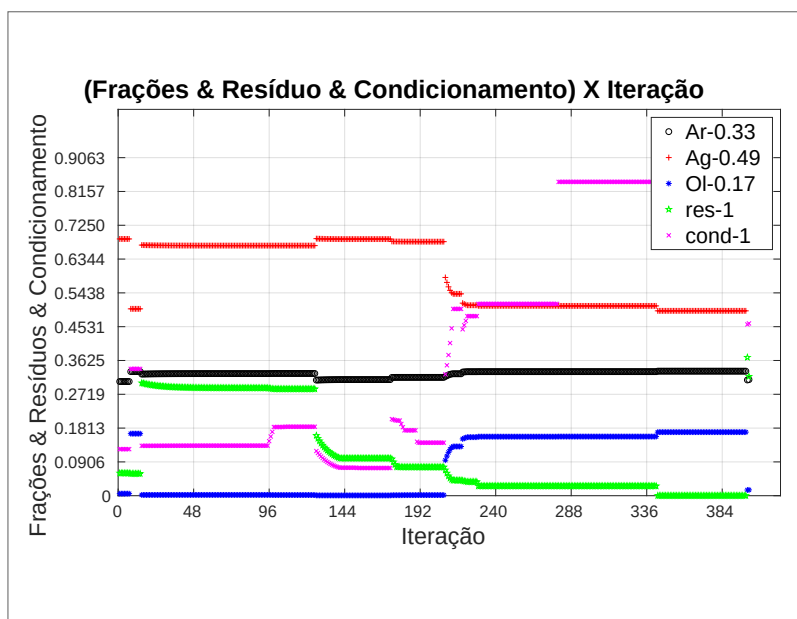
F.2 DETERMINAÇÃO DE FRAÇÕES PERCENTUAIS - APROXIMAÇÃO DOS COMPOSTOS

Figura 68 – Gráficos contendo a determinação de fração de contagens percentuais dos compostos. Pontos (○), vermelho(+) e azul(*) indicam, respectivamente, as quantidades de Ar, Água e Óleo. Os valores nominais estão descritos nas legendas dos Gráficos 68a e 68b. Os valores dos resíduos são descritos nos pontos ☆, e o ponto ✕ indica o condicionamento $([0, 1])$ das matrizes de covariância na aproximação das bibliotecas.

(a) Frações percentuais dos compostos após supressão da Água do espectro total.



(b) Frações percentuais dos compostos após supressão do Óleo do espectro total.



Fonte: Próprio Autor

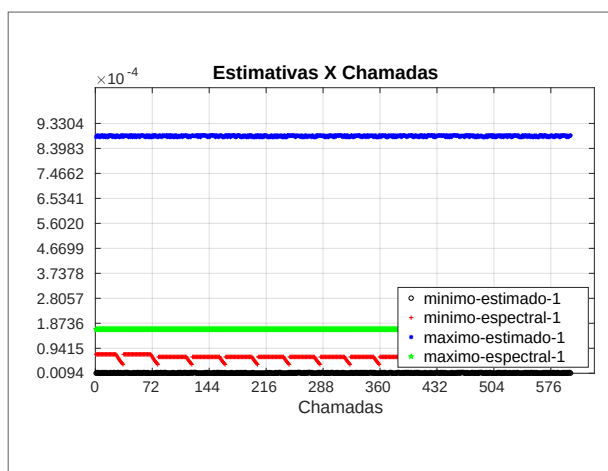
APÊNDICE G – APROXIMAÇÃO DOS COMPOSTOS - ESTIMATIVA DA ALTURA NAS PARTIÇÕES

G.1 APROXIMAÇÃO DOS ELEMENTOS - ESTIMATIVA DA ALTURA NAS PARTIÇÕES

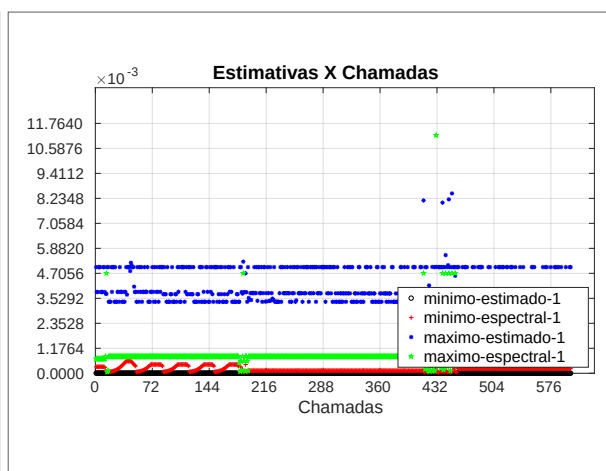
G.1.1 Gráficos para Estimativa da Altura

Figura 69 – Estimativas da biblioteca omitida. Os pontos $\circ$ e $\star$ são os valores de mínima e máxima, respectivamente, contagem espectral estimada, e os pontos representados por $+$ e $\star$ indicam o mínima e o máxima, respectivamente, contagem espectral real no intervalo do compartimento. As figuras 69a, 69b e 69c são resultados para as instâncias das bibliotecas, respectivamente, de Co , Cs e Na .

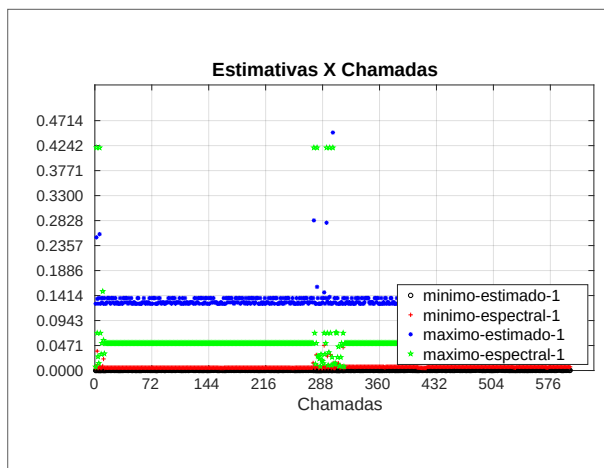
(a) Estimativas para a biblioteca do Co , instância 1. Valores estimados, com medianas em verde, são comparados com os valores de contagens espectrais totais.



(b) Estimativas para a biblioteca do Cs , instância 3. Valores estimados, com medianas em verde, são comparados com os valores de contagens espectrais totais.



(c) Estimativas para a biblioteca do Na , instância 4. Valores estimados, com medianas em verde, são comparados com os valores de contagens espectrais totais.



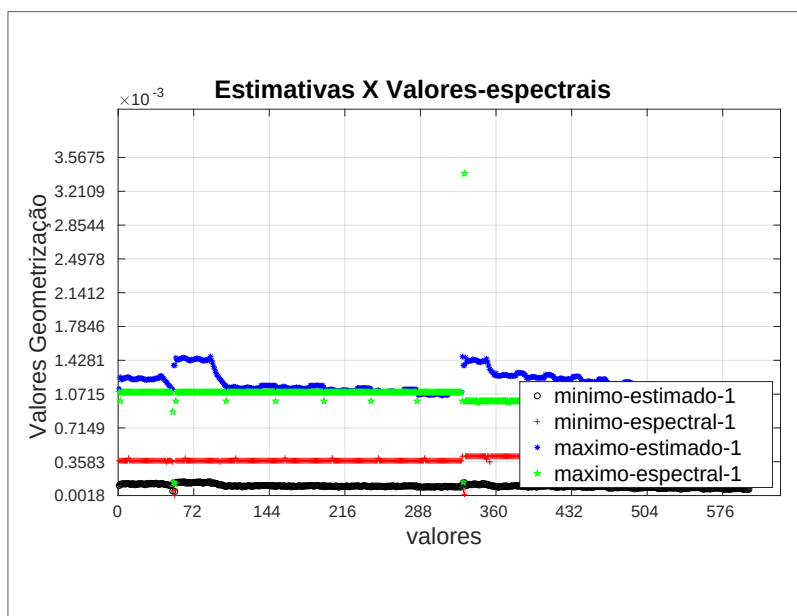
Fonte: Próprio Autor

G.2 APROXIMAÇÃO DOS COMPOSTOS - ESTIMATIVA DA ALTURA NAS PARTIÇÕES

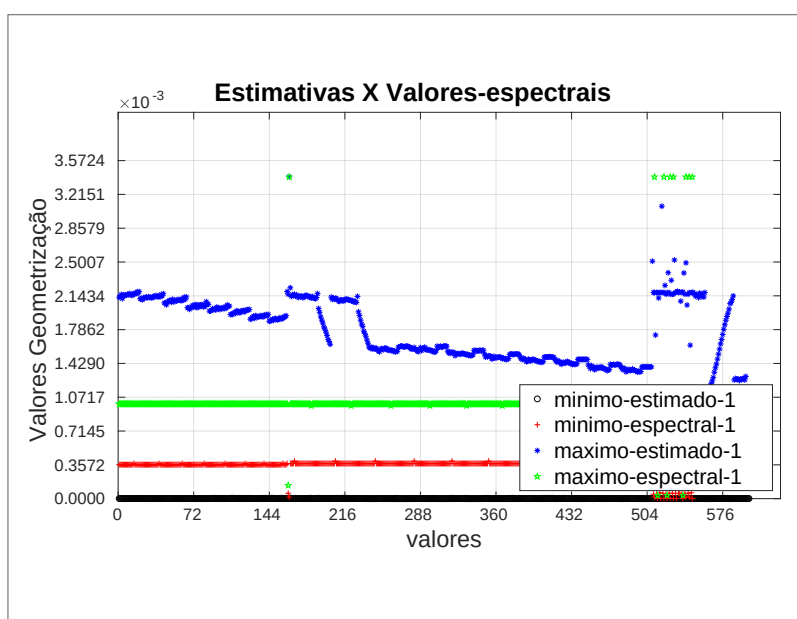
G.2.1 Gráficos para Estimativa da Altura

Figura 70 – Estimativa da biblioteca omitida. Os pontos $\circ$ são os valores de mínima contagem espectral estimado. Os pontos por $+$ representam a mínima contagem espectral real total nos compartimentos. Os pontos $\ast$ representam os valores de máxima contagem espectral estimada. Os pontos $\star$ indicam o valor de máximo espectral real total. As figuras 70a e 70b são representativas para a omissão das bibliotecas, respectivamente da Água e do Óleo.

(a) Estimativas para a biblioteca da Água suprimida, instância 6.



(b) Estimativas para a biblioteca do Óleo suprimido, instância 7.



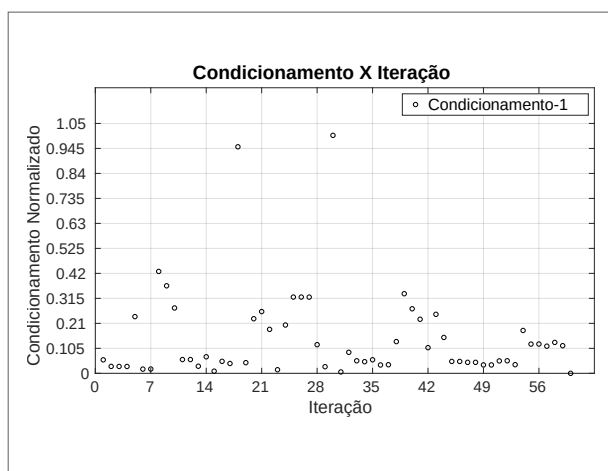
Fonte: Próprio Autor

APÊNDICE H – CONDICIONAMENTO NA MINIMIZAÇÃO DA MÉTRICA

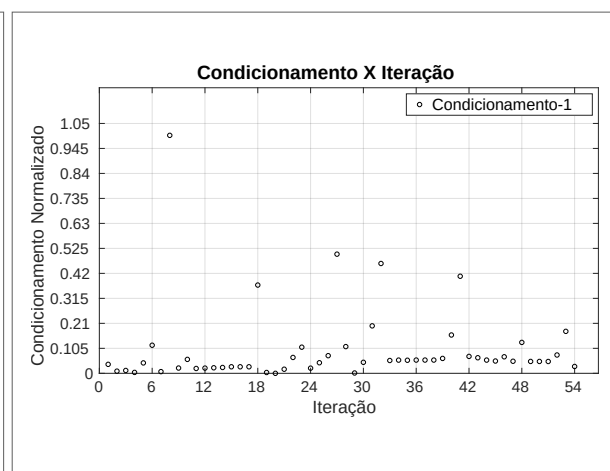
H.1 NÚMERO DE CONDICIONAMENTO NA MINIMIZAÇÃO DA MÉTRICA - ELEMENTOS

Figura 71 – Comportamento do número de condicionamento na minimização do resíduo. A Figura 71a traz os resultados para a instância dois; a Figura 71b traz os resultados para a instância 3 e, por fim, a Figura 71c traz os resultados para a instância quatro.

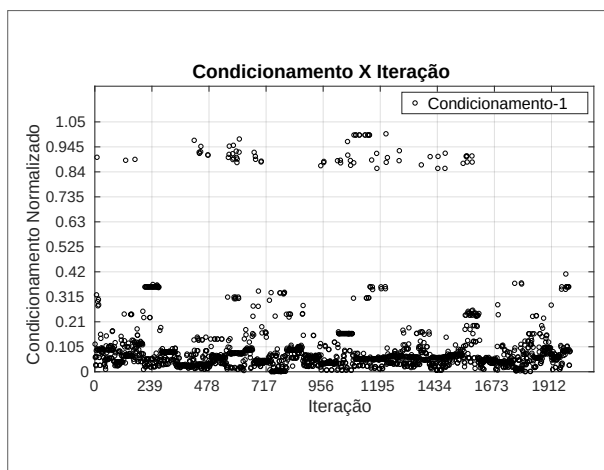
(a) Valores do condicionamento nas ocorrências da minimização do resíduo, instância dois.



(b) Valores do condicionamento nas ocorrências da minimização do resíduo, instância três.



(c) Valores do condicionamento nas ocorrências da minimização do resíduo, instância quatro.

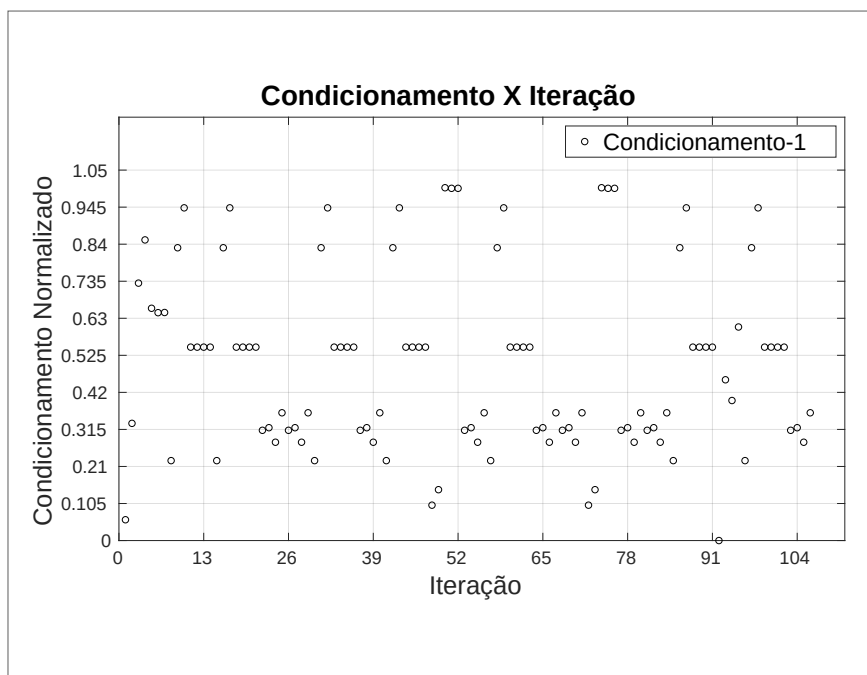


Fonte: Próprio Autor

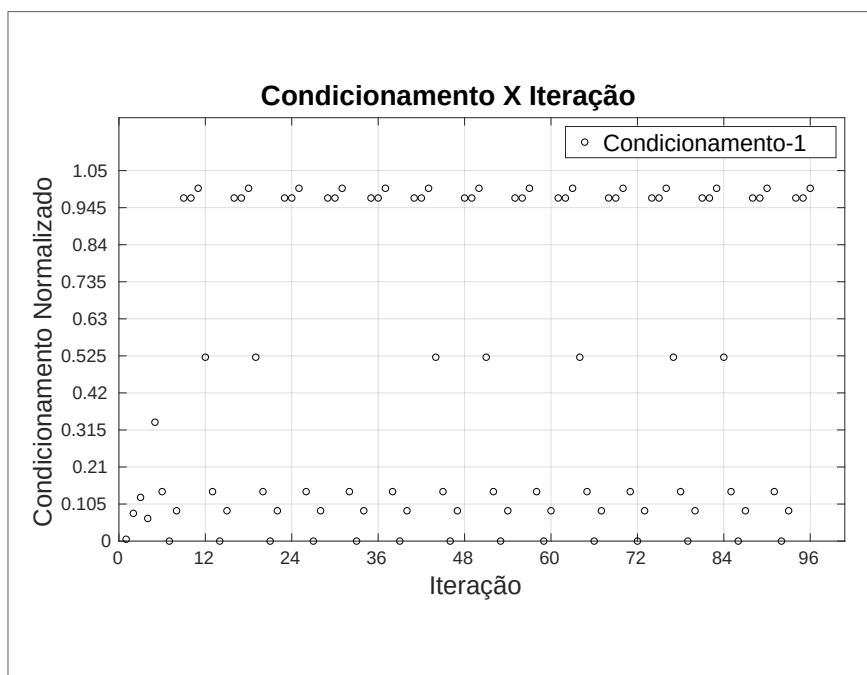
H.2 CONDICIONAMENTO NA MINIMIZAÇÃO DA MÉTRICA - COMPOSTOS

Figura 72 – Condicionamento na determinação das frações de contagem percentual, na omissão de uma das bibliotecas, nas ocorrências de minimização do resíduo. Na Figura 72a, resultados para a instância seis; Figura 72b, valores de condicionamento para a instância sete.

- (a) Valores de condicionamento nas ocorrências de minimização do resíduo para a instância seis.



- (b) Valores de condicionamento nas ocorrências de minimização do resíduo para a instância sete.



Fonte: Próprio Autor

APÊNDICE I – RESULTADOS ETAPAS BÁSICAS DO GRASP

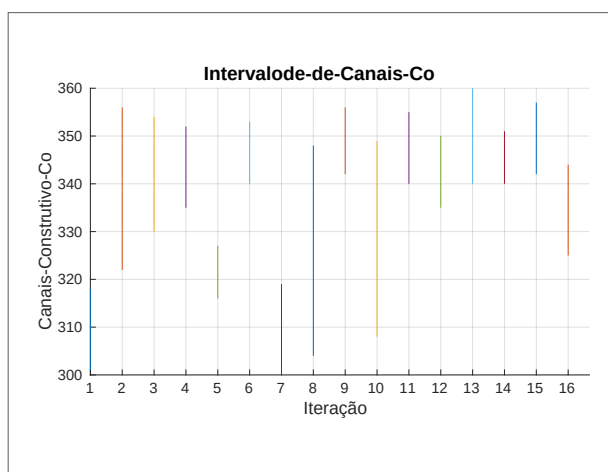
I.1 BASE DE DADOS 1

I.1.1 Resultados do Passo construtivo

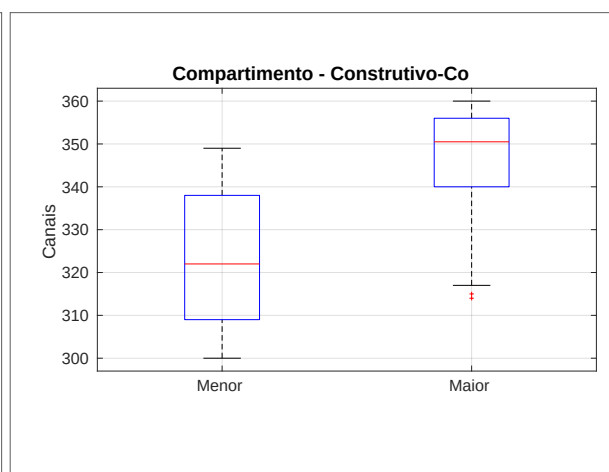
Escolha dos Canais nos Compartimentos

Figura 73 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 73a, para omissão do Co , no estágio da construção. A figura 73b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Co .

(a) Intervalo de canais como função da iteração, omitindo Co , no estágio de construção de uma solução.



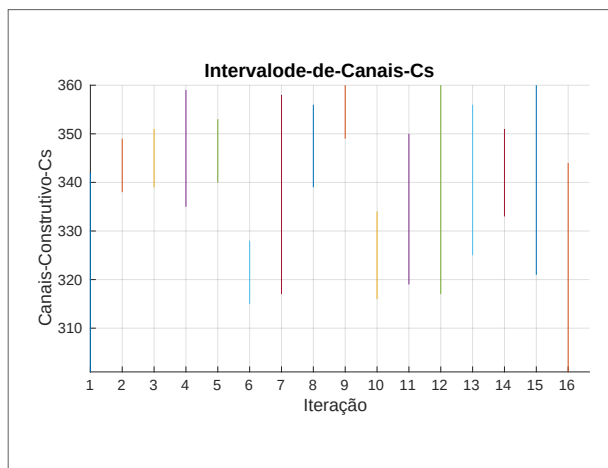
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de construção, instância dois.



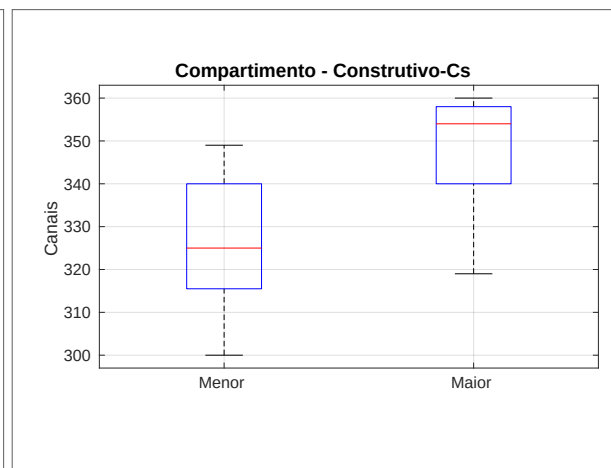
Fonte: Próprio Autor

Figura 74 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 74a, para omissão do Cs , no estágio da construção. A figura 74b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Cs .

(a) Intervalo de canais como função da iteração, omitindo Cs , no estágio de construção de uma solução.



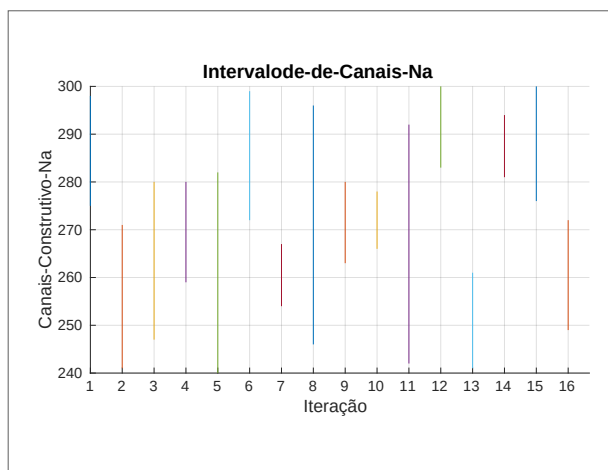
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de construção, instância três.



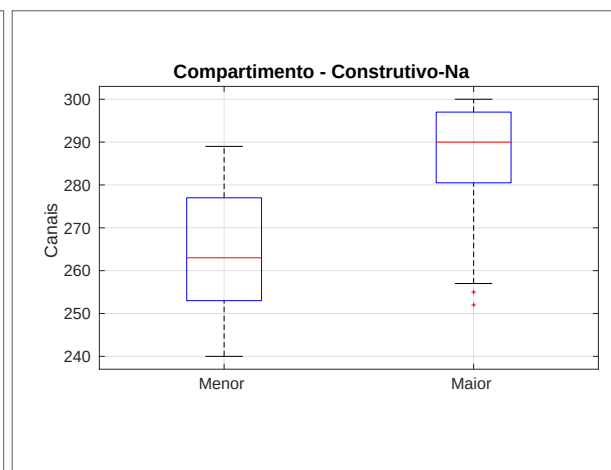
Fonte: Próprio Autor

Figura 75 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 75a, para omissão do Na , no estágio da construção. A figura 75b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Na .

(a) Intervalo de canais como função da iteração, omitindo Na , no estágio de construção de uma solução.



(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de construção, instância quatro.



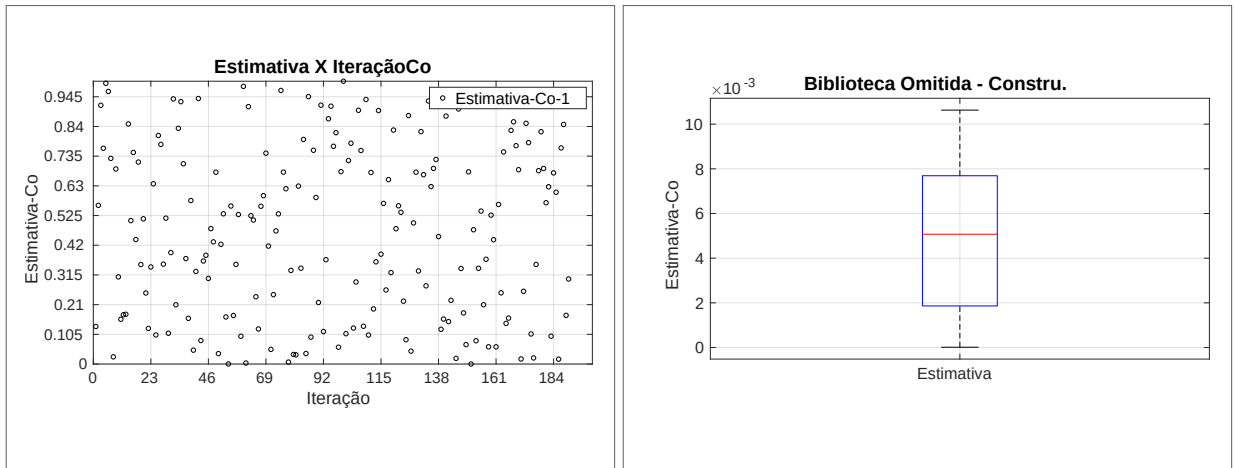
Fonte: Próprio Autor

Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 76 – O gráfico 76a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância um. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 76b.

(a) Estimativas para o elemento omitido durante o passo construtivo, instância dois, ausência do C_o .

(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo construtivo, instância dois.

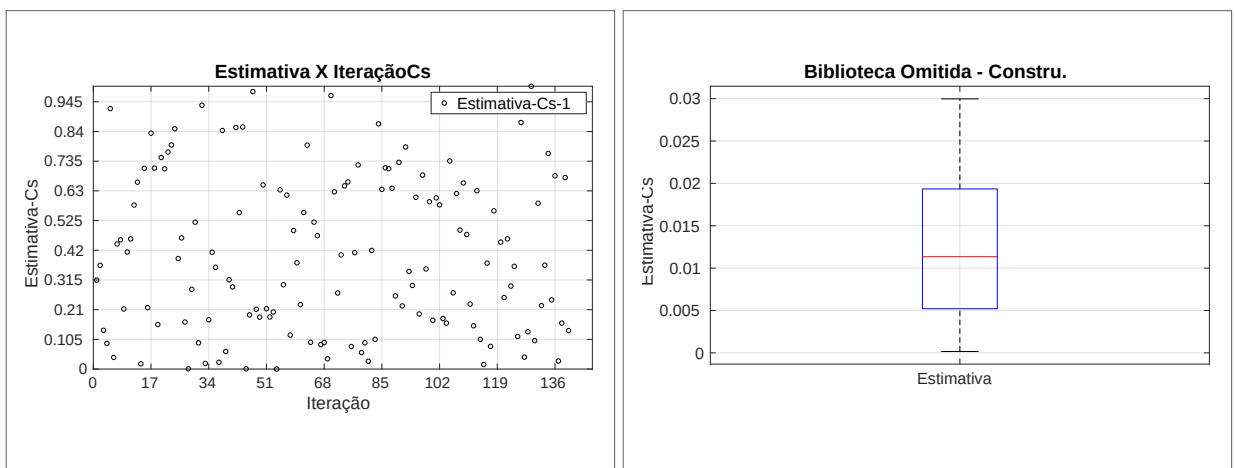


Fonte: Próprio Autor

Figura 77 – O gráfico 77a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 77b.

(a) Estimativas para o elemento omitido durante o passo construtivo, instância três, ausência do C_s .

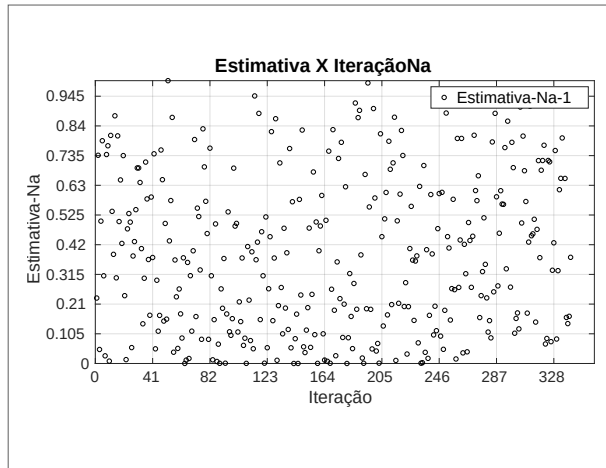
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo construtivo, instância três.



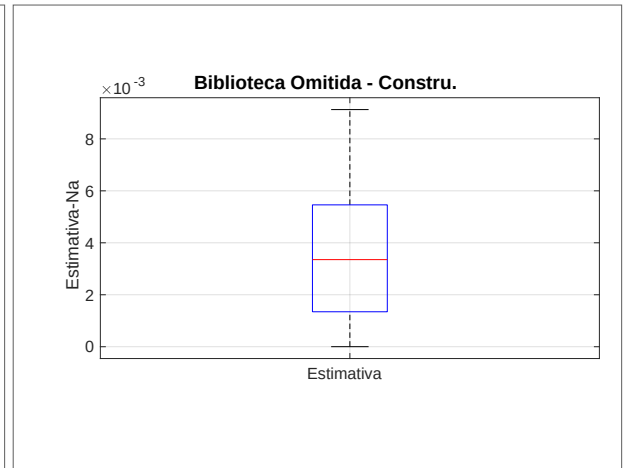
Fonte: Próprio Autor

Figura 78 – O gráfico 78a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 78b.

(a) Estimativas para o elemento omitido durante o passo construtivo, instância quatro, ausência do Na .



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo construtivo, instância quatro.

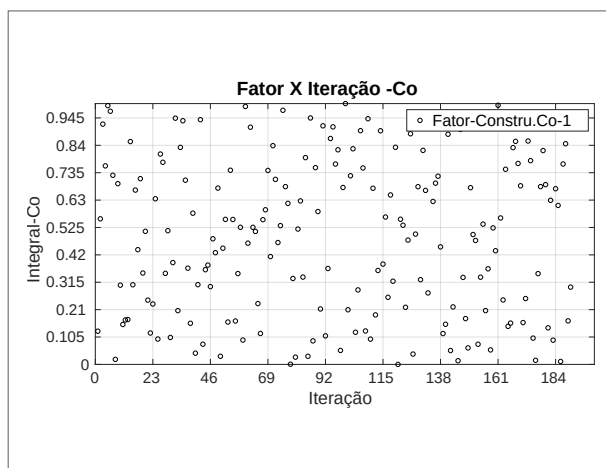


Fonte: Próprio Autor

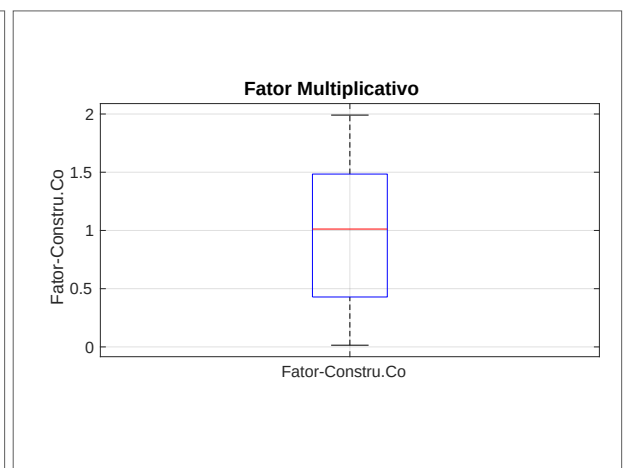
Constante Modificadora da Biblioteca Omitida

Figura 79 – O gráfico 79a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 79b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância dois, no passo construtivo.



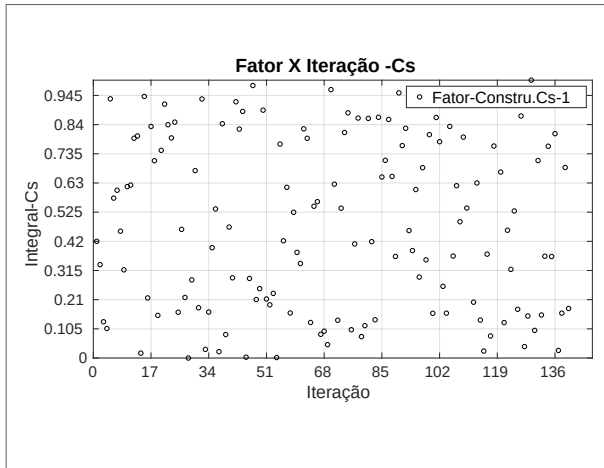
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância dois, passo construtivo.



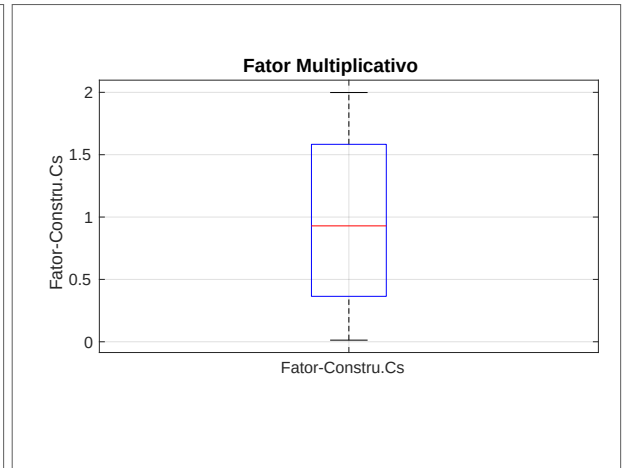
Fonte: Próprio Autor

Figura 80 – O gráfico 80a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 80b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância três, no passo construtivo.



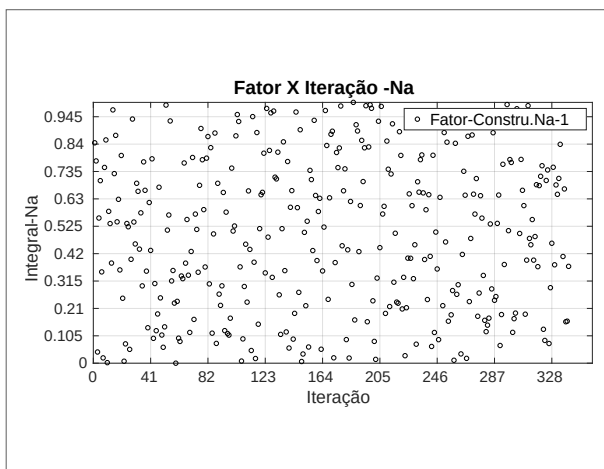
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância três, passo construtivo.



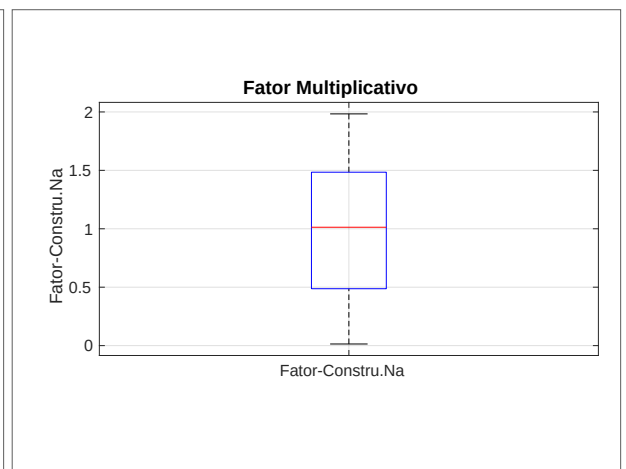
Fonte: Próprio Autor

Figura 81 – O gráfico 81a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 81b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância quatro, no passo construtivo.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância quatro, passo construtivo.



Fonte: Próprio Autor

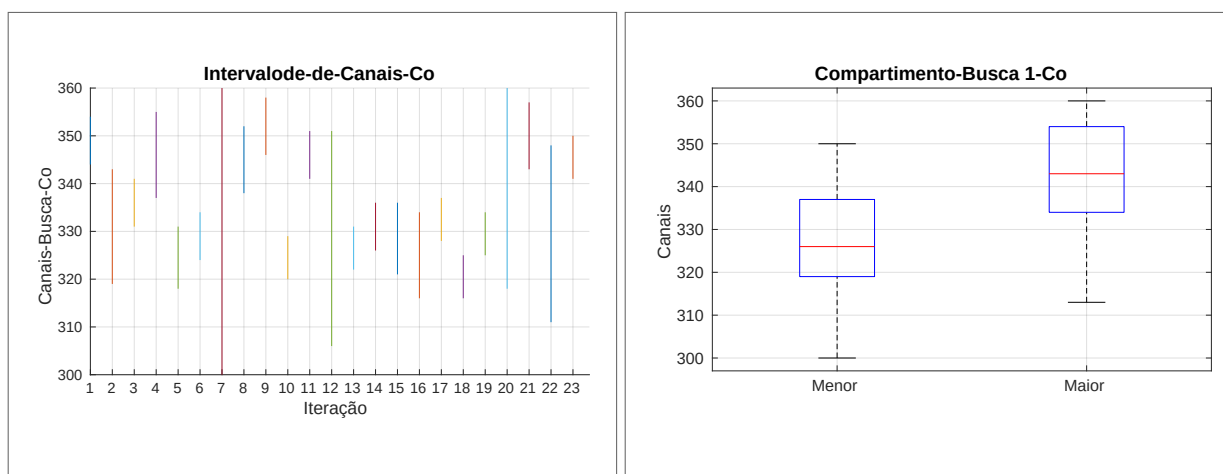
I.1.2 Resultados do Passo de Busca

Escolha dos Canais nos Compartimentos

Figura 82 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 82a, para omissão do Co , no estágio da busca. A figura 82b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância dois.

(a) Intervalo de canais, como função da iteração, omitindo Co , no estágio de busca de uma solução.

(b) Distribuição dos canais que definem o compartimento na partição de maior resíduo, estágio de busca de uma solução, instância dois.

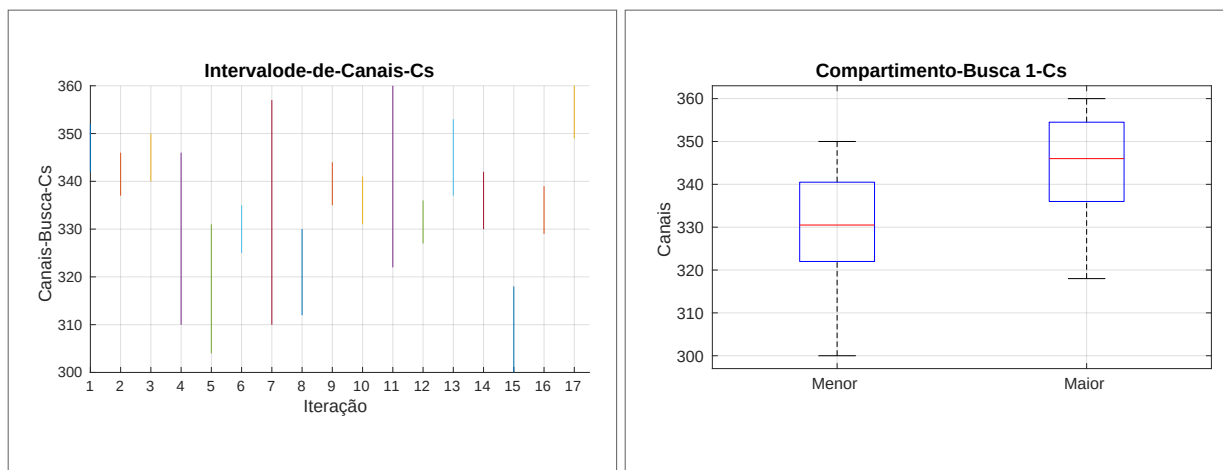


Fonte: Próprio Autor

Figura 83 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 83a, para omissão do Cs , no estágio da busca. A figura 83b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância três.

(a) Intervalo de canais como função da iteração, omitindo Cs , no estágio de busca de uma solução.

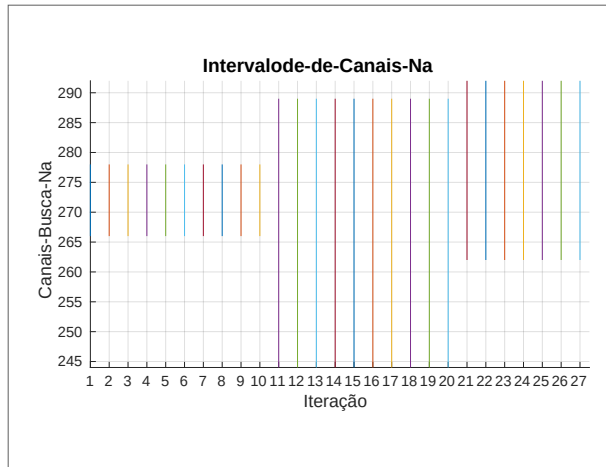
(b) Distribuição dos canais que definem o compartimento na partição de maior resíduo, estágio de busca de uma solução, instância três.



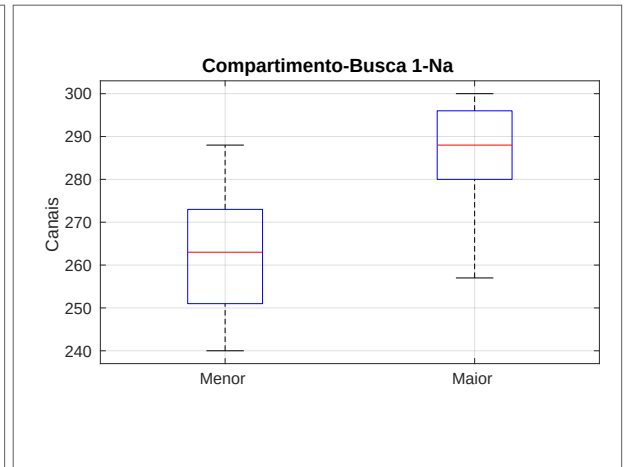
Fonte: Próprio Autor

Figura 84 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 84a, para omissão do Na , no estágio da busca. A figura 84b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo para a instância três.

(a) Intervalo de canais como função da iteração, omitindo Na , no estágio de busca de uma solução.



(b) Distribuição dos canais que definem o compartimento na partição de maior resíduo, estágio de busca de uma solução, instância quatro.

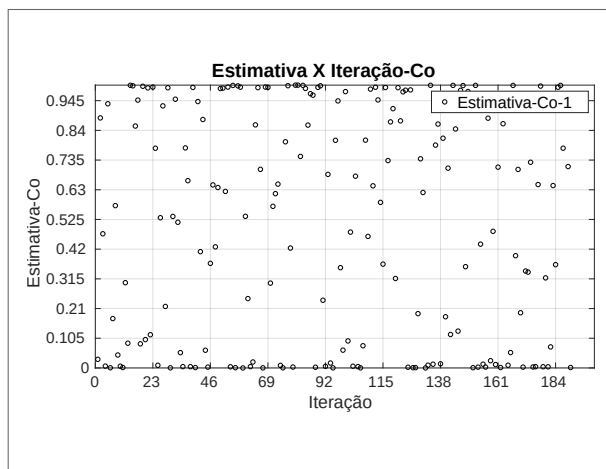


Fonte: Próprio Autor

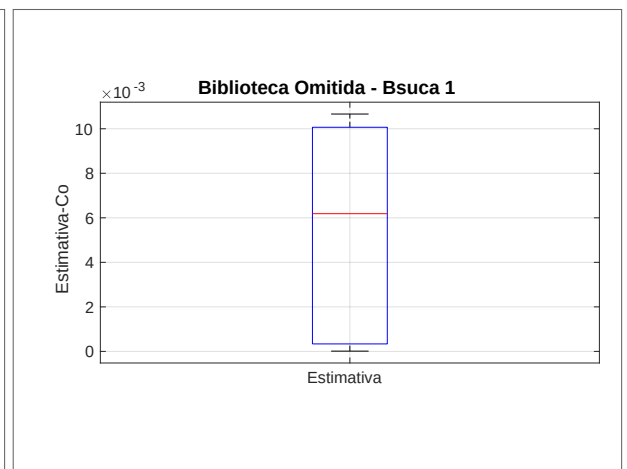
Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 85 – O gráfico 85a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 85b.

(a) Estimativas para o elemento omitido durante o passo de busca, instância dois, ausência do Co .



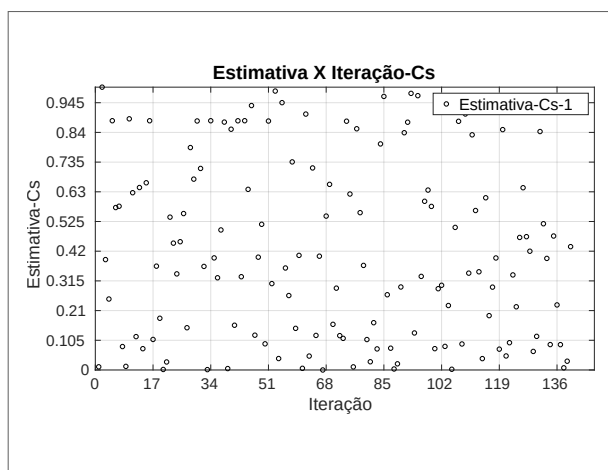
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de busca, instância dois.



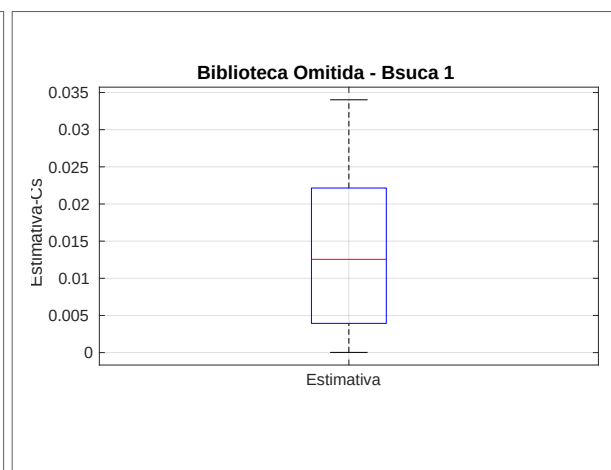
Fonte: Próprio Autor

Figura 86 – O gráfico 86a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 86b.

(a) Estimativas para o elemento omitido durante o passo de busca, instância três, ausência do C_s .



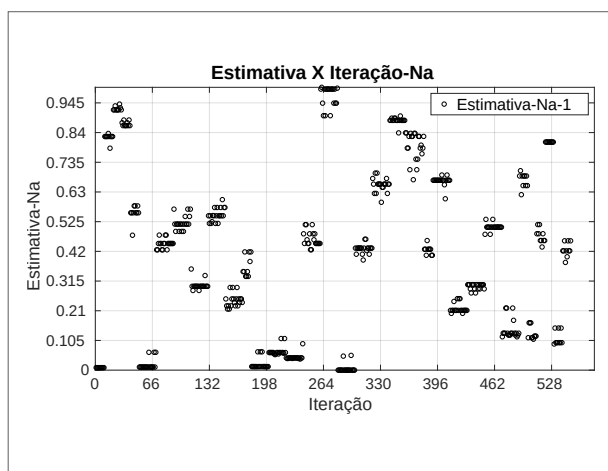
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de busca, instância três.



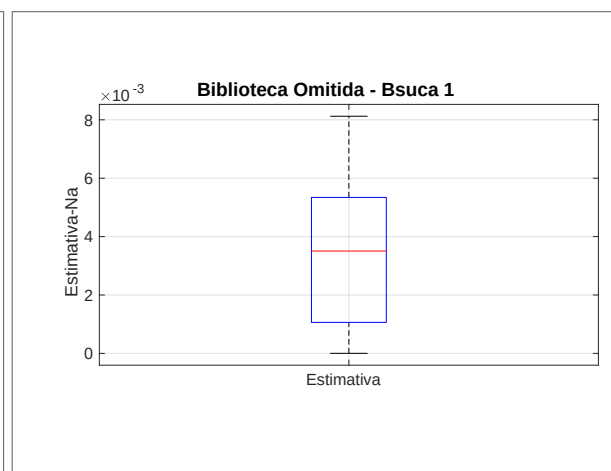
Fonte: Próprio Autor

Figura 87 – O gráfico 87a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 87b.

(a) Estimativas para o elemento omitido durante o passo de busca, instância quatro, ausência do Na .



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de busca, instância quatro.

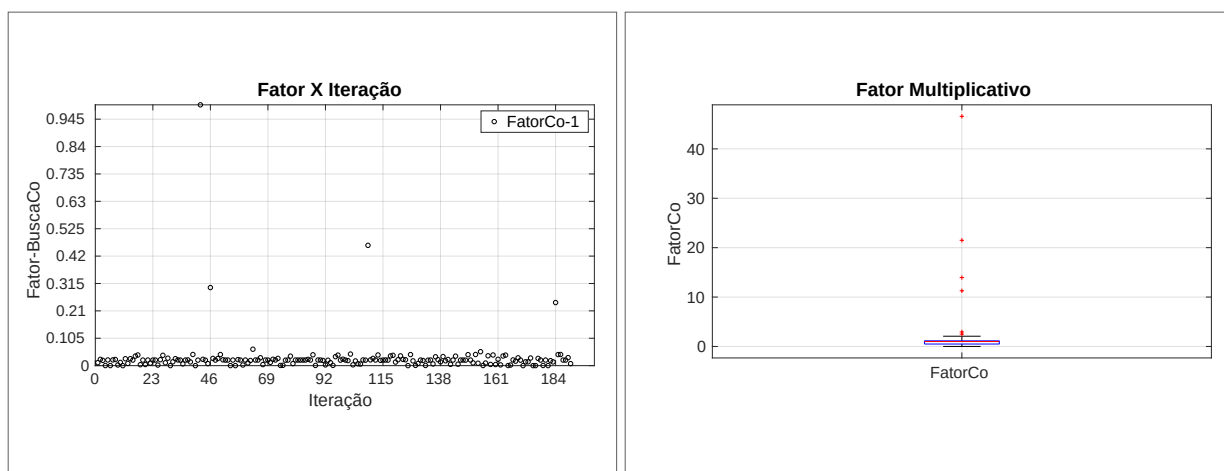


Fonte: Próprio Autor

Constante Modificadora da Biblioteca Omitida

Figura 88 – O gráfico 88a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 88b.

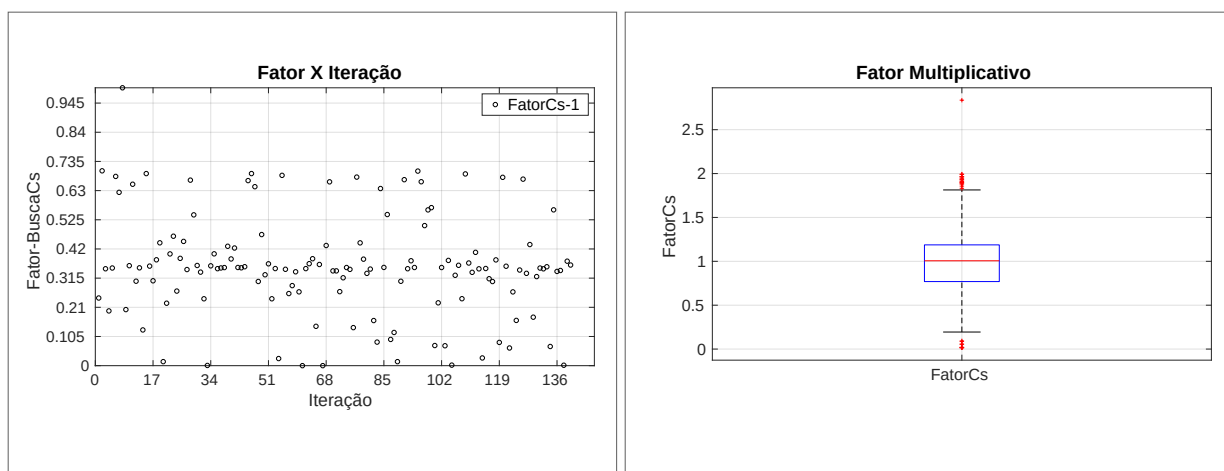
- (a) Valores da constante que modifica a estimativa da biblioteca omitida, instância dois, no passo da primeira busca.
- (b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância dois, passo de busca.



Fonte: Próprio Autor

Figura 89 – O gráfico 89a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 89b.

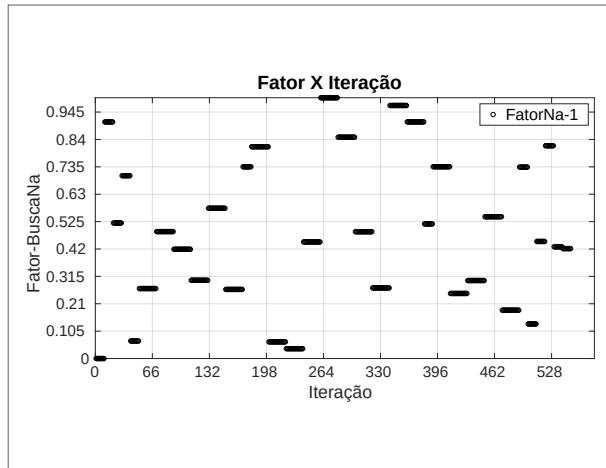
- (a) Valores da constante que modifica a estimativa da biblioteca omitida, instância três, no passo da primeira busca.
- (b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância três, passo de busca.



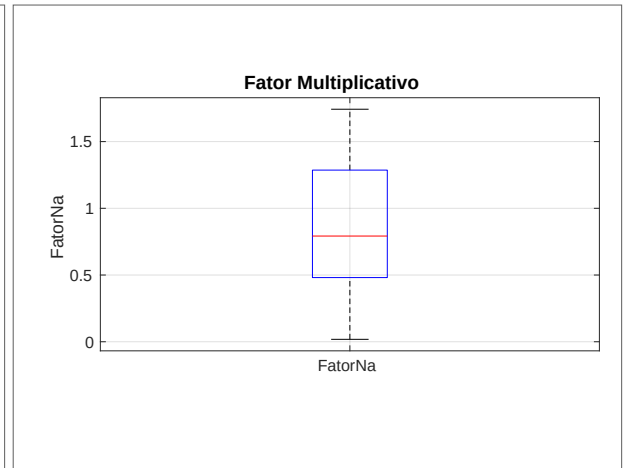
Fonte: Próprio Autor

Figura 90 – O gráfico 90a apresenta os valores de constantes que definem a biblioteca omitida, no passo de busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 90b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância quatro, no passo da primeira busca.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância quatro, passo de busca.



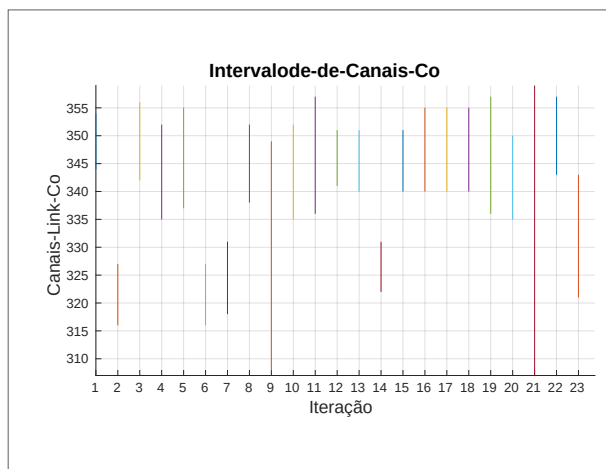
Fonte: Próprio Autor

I.1.3 Resultados do Passo de Linkagem

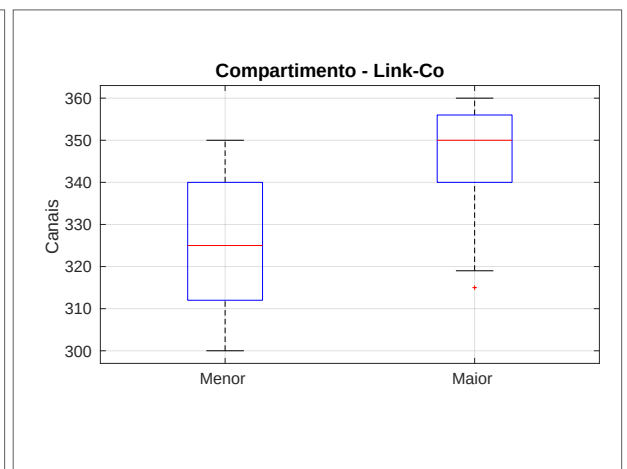
Escolha dos Canais nos Compartimentos

Figura 91 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do Co , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Co .

(a) Intervalo de canais como função da iteração, omitindo Co , no estágio de linkagem de uma solução.



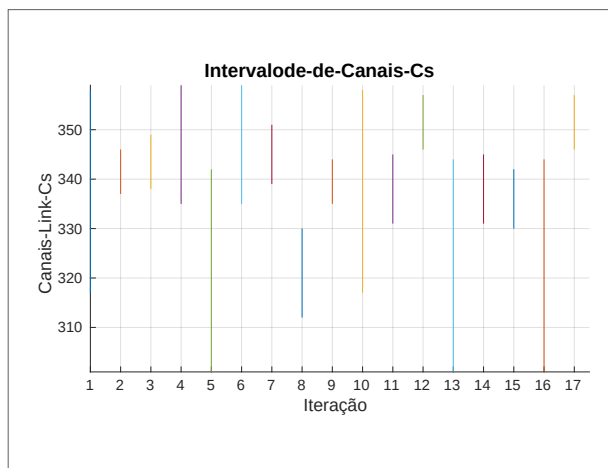
(b) Distribuição de canais que define o compartimento na partição de maior resíduo, estágio de linkagem, instância dois.



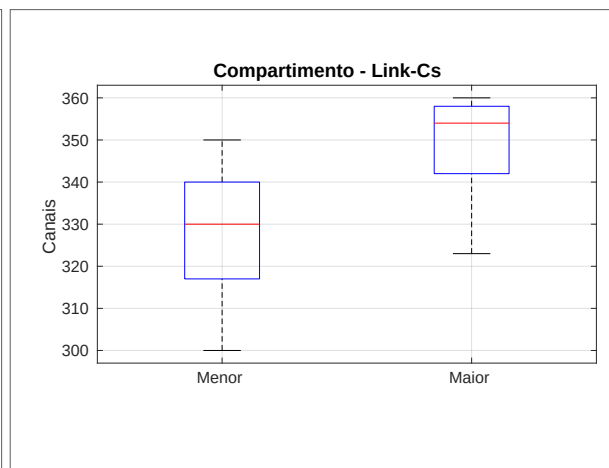
Fonte: Próprio Autor

Figura 92 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do C_s , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do C_s .

(a) Intervalo de canais como função da iteração, omitindo C_s , no estágio de linkagem de uma solução



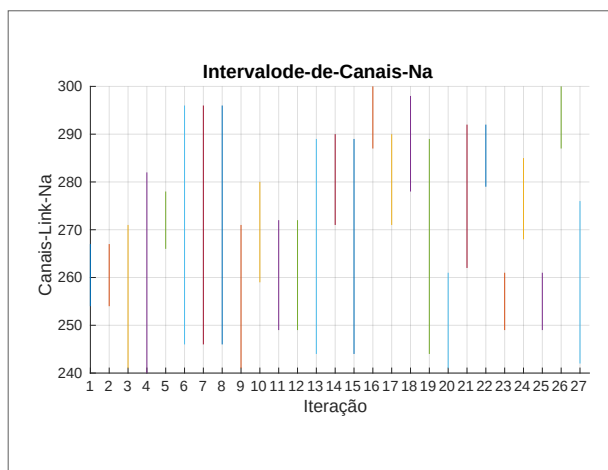
(b) Distribuição de canais que define o compartimento na partição de maior resíduo, estágio de linkagem, instância três.



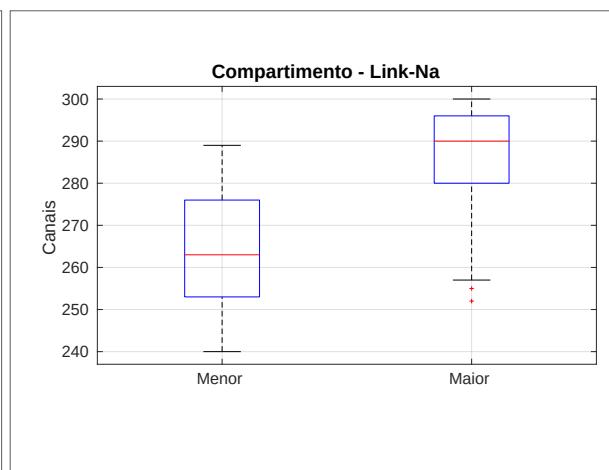
Fonte: Próprio Autor

Figura 93 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 32a, para omissão do Na , no estágio da linkagem. A figura 32b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Na .

(a) Intervalo de canais como função da iteração, omitindo Na , no estágio de linkagem de uma solução.



(b) Distribuição de canais que define o compartimento na partição de maior resíduo, estágio de linkagem, instância quatro.

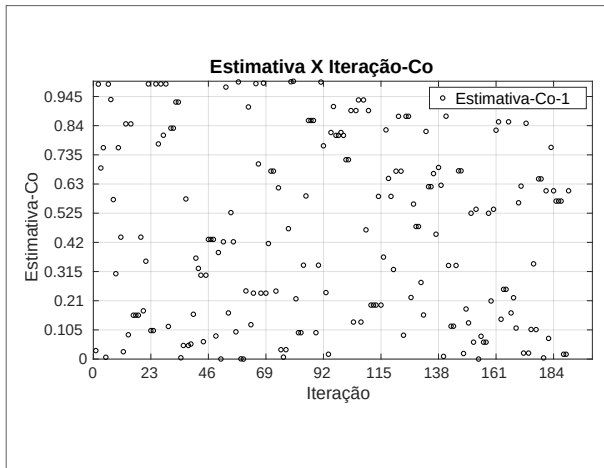


Fonte: Próprio Autor

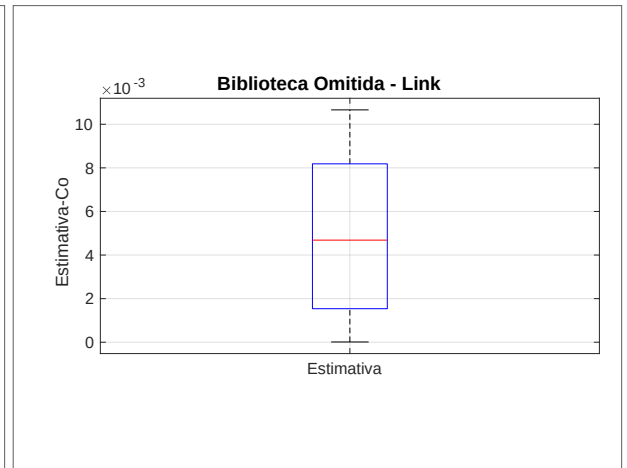
Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 94 – O gráfico 94a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 94b.

(a) Estimativas para o elemento omitido durante o passo de linkagem, instância dois, ausência do C_o .



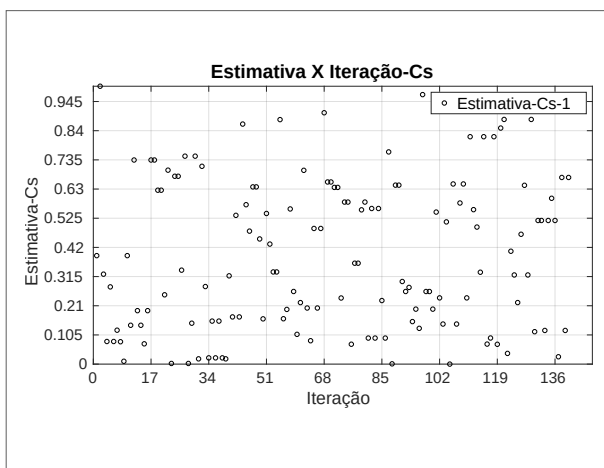
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de linkagem, instância dois.



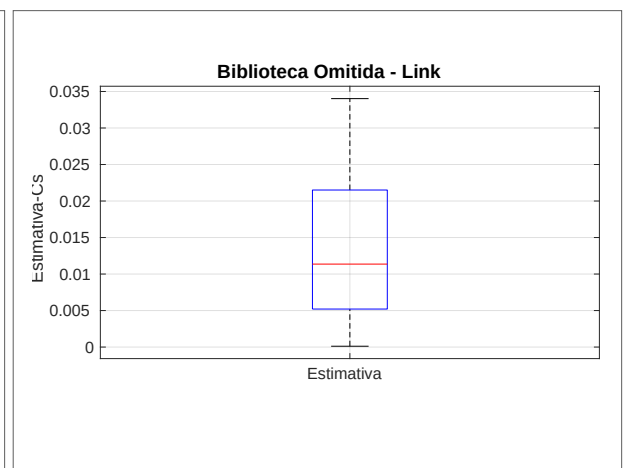
Fonte: Próprio Autor

Figura 95 – O gráfico 95a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 95b.

(a) Estimativas para o elemento omitido durante o passo de linkagem, instância três, ausência do C_s .



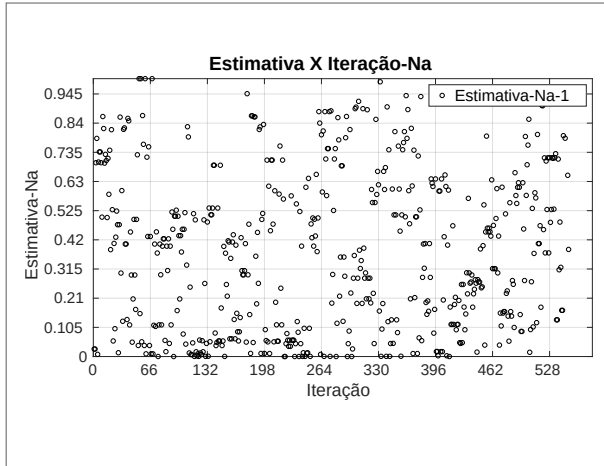
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de linkagem, instância três.



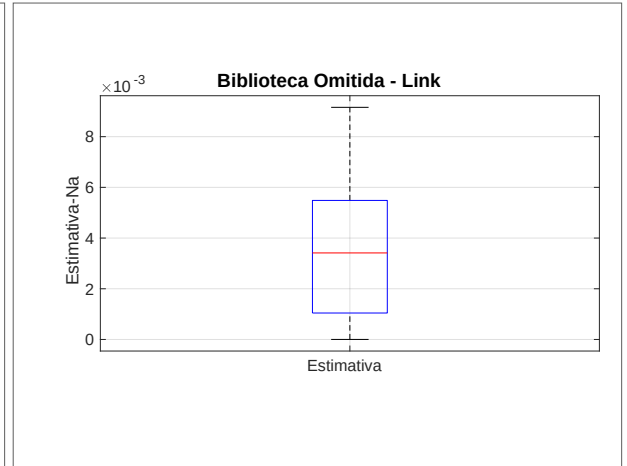
Fonte: Próprio Autor

Figura 96 – O gráfico 96a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 96b.

(a) Estimativas para o elemento omitido durante o passo de linkagem, instância quatro, ausência do N_a .



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de linkagem, instância quatro.

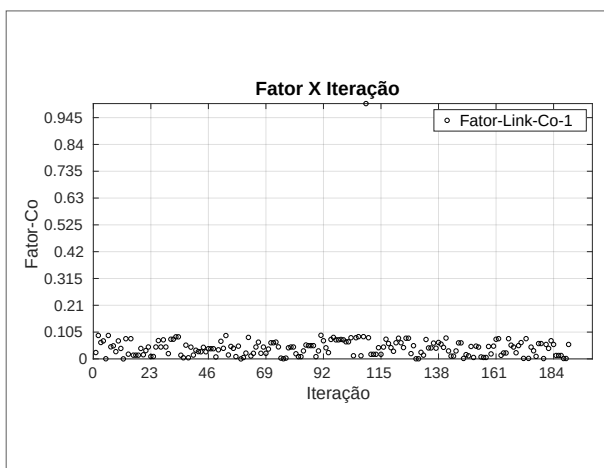


Fonte: Próprio Autor

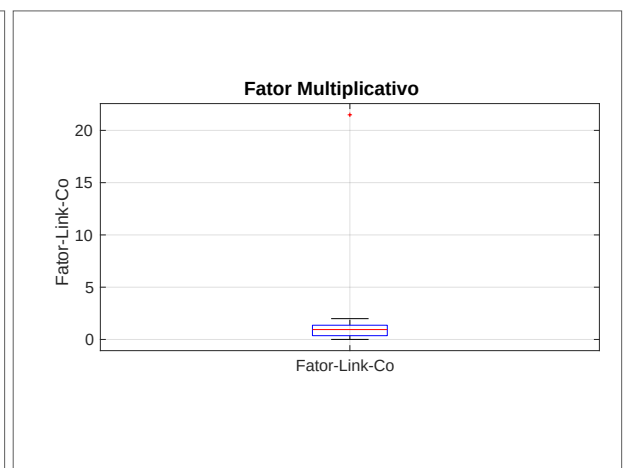
Constante Modificadora da Biblioteca Omitida

Figura 97 – O gráfico 97a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 97b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância dois, no passo da linkagem.



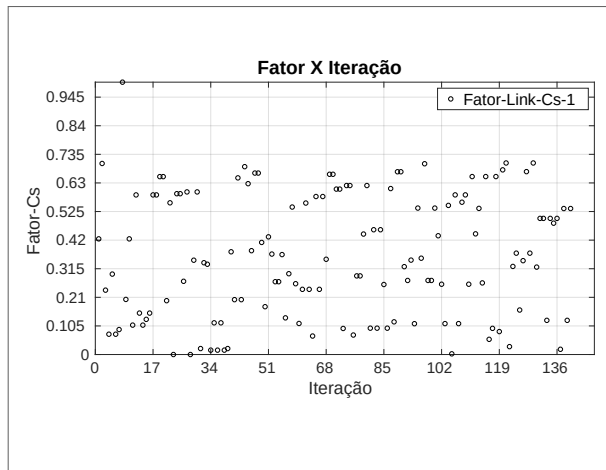
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância dois, passo de linkagem.



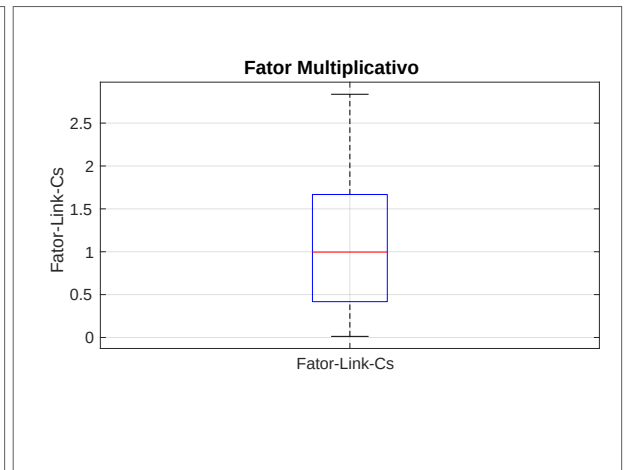
Fonte: Próprio Autor

Figura 98 – O gráfico 98a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 98b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância três, no passo da linkagem.



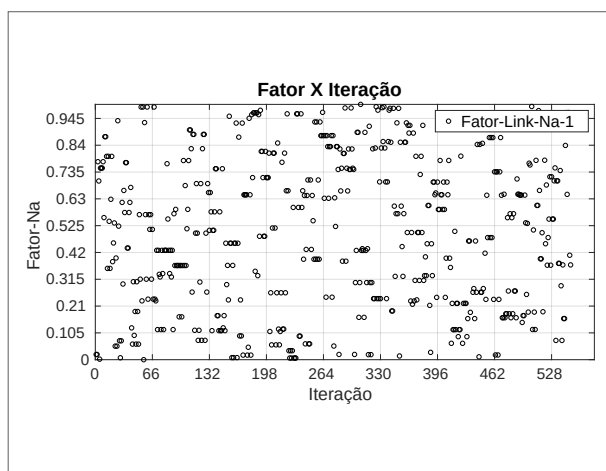
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância três, passo de linkagem.



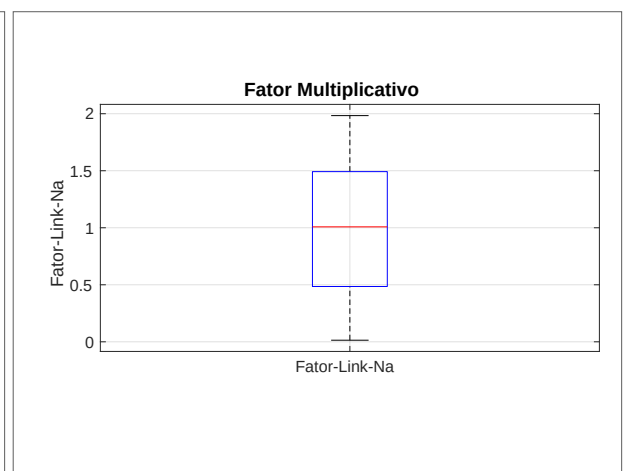
Fonte: Próprio Autor

Figura 99 – O gráfico 99a apresenta os valores de constantes que definem a biblioteca omitida, no passo de linkagem do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 99b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância quatro, no passo da linkagem.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância quatro, passo de linkagem.



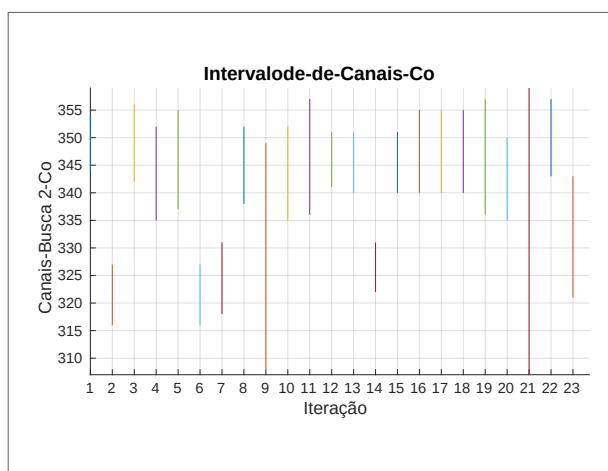
Fonte: Próprio Autor

I.1.4 Resultados do Passo da Segunda Busca

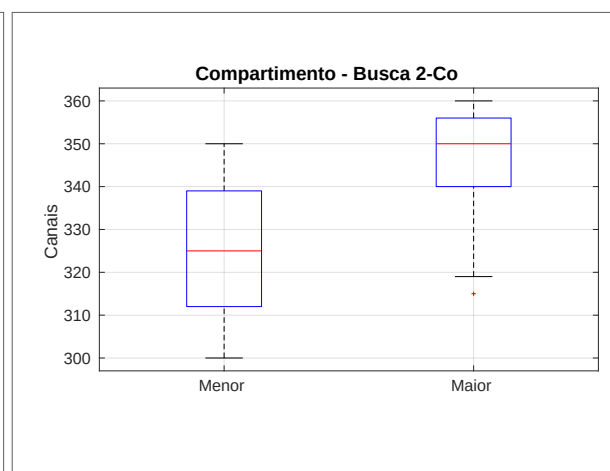
Escolha dos Canais nos Compartimentos

Figura 100 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 100a, para omissão do Co , no estágio da segunda busca. A figura 100b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Am .

(a) Intervalo de canais, como função da iteração, omitindo Co , no estágio de segunda busca do GRASP.



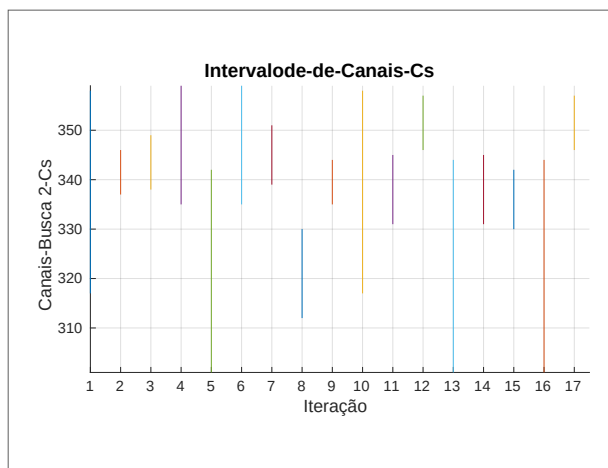
(b) Distribuição de canais que define o compartimento na partição de maior resíduo, segunda busca, instância dois.



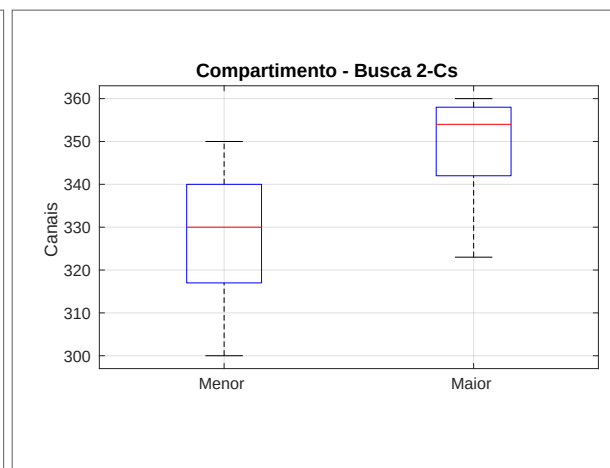
Fonte: Próprio Autor

Figura 101 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 101a, para omissão do Cs , no estágio da segunda busca. A figura 101b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Am .

(a) Intervalo de canais, como função da iteração, omitindo Cs , no estágio de segunda busca do GRASP.



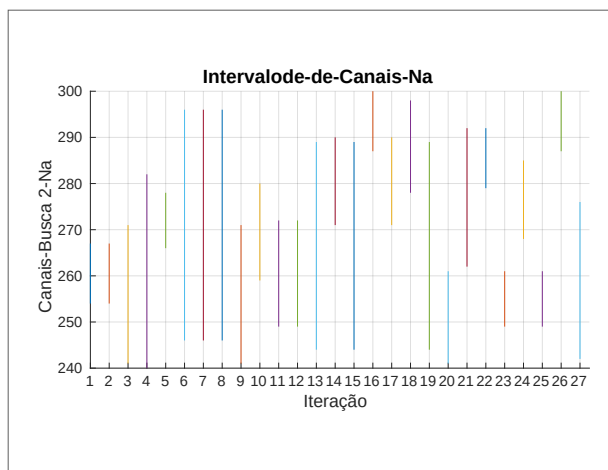
(b) Distribuição de canais que define o compartimento na partição de maior resíduo, segunda busca, instância três.



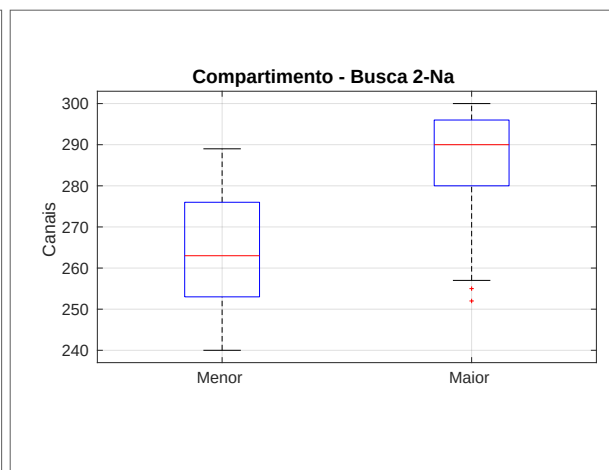
Fonte: Próprio Autor

Figura 102 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 102a, para omissão do Na , no estágio da segunda busca. A figura 102b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Na .

(a) Intervalo de canais, como função da iteração, omitindo Na , no estágio de segunda busca do GRASP.



(b) Distribuição de canais que define o compartimento na partição de maior resíduo, segunda busca, instância quatro.

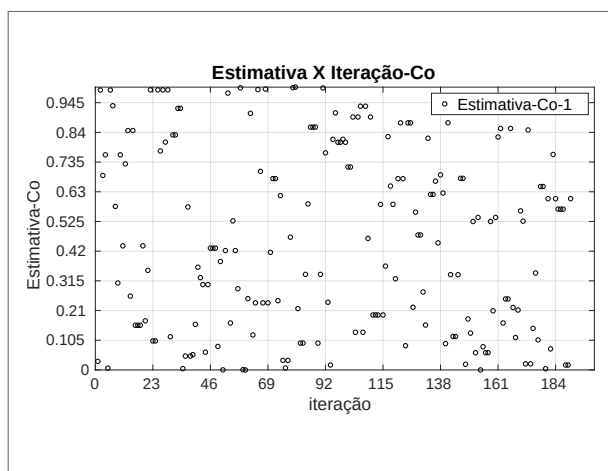


Fonte: Próprio Autor

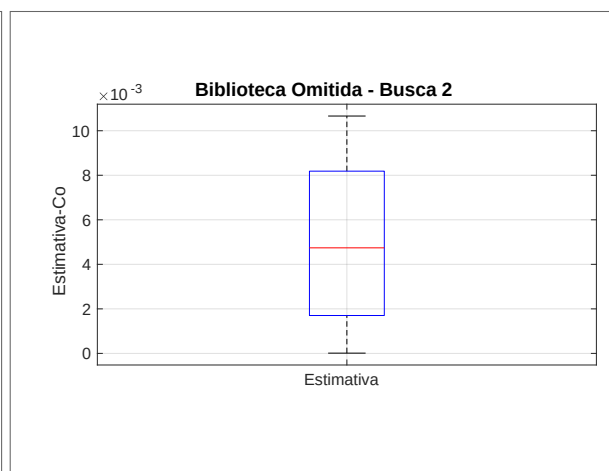
Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 103 – O gráfico 103a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 103b.

(a) Estimativas para o elemento omitido durante o da segunda busca, instância dois, ausência do Co .



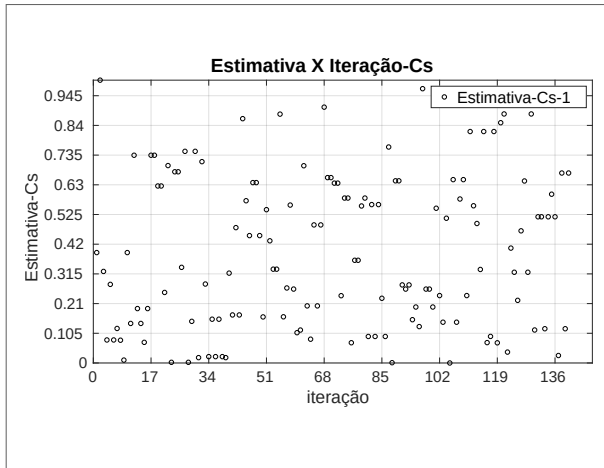
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de segunda busca, instância dois.



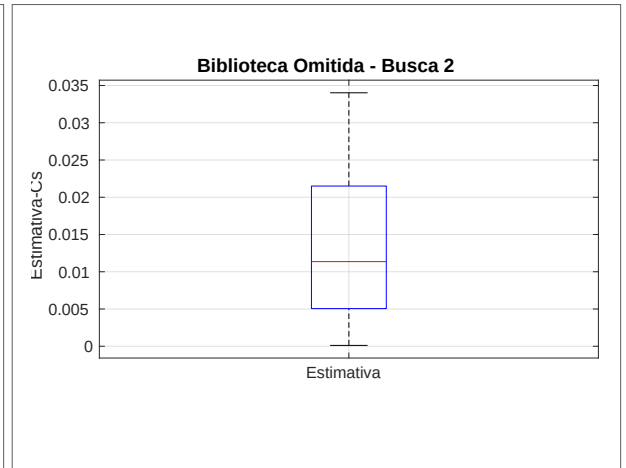
Fonte: Próprio Autor

Figura 104 – O gráfico 104a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 104b.

(a) Estimativas para o elemento omitido durante o da segunda busca, instância um, ausência do C_s .



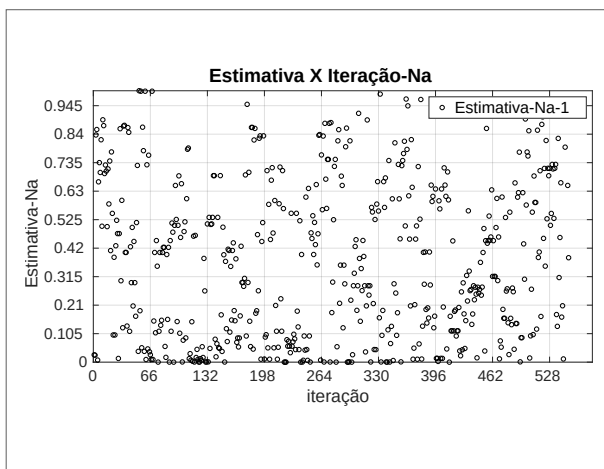
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de segunda busca, instância dois.



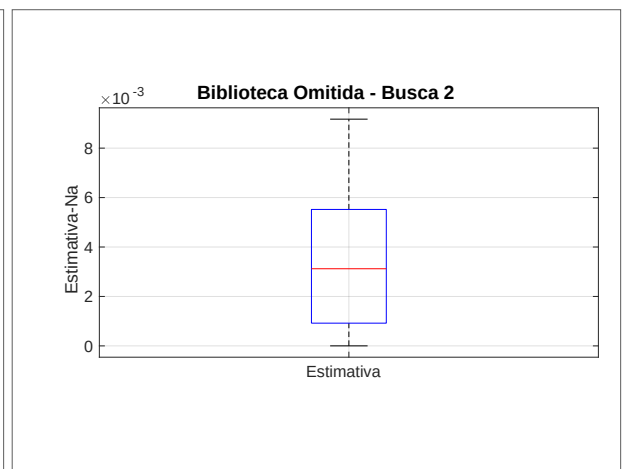
Fonte: Próprio Autor

Figura 105 – O gráfico 105a apresenta os valores de estimativas para a biblioteca omitida, no passo de segunda busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 105b.

(a) Estimativas para o elemento omitido durante o da segunda busca, instância um, ausência do N_a .



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de segunda busca, instância dois.

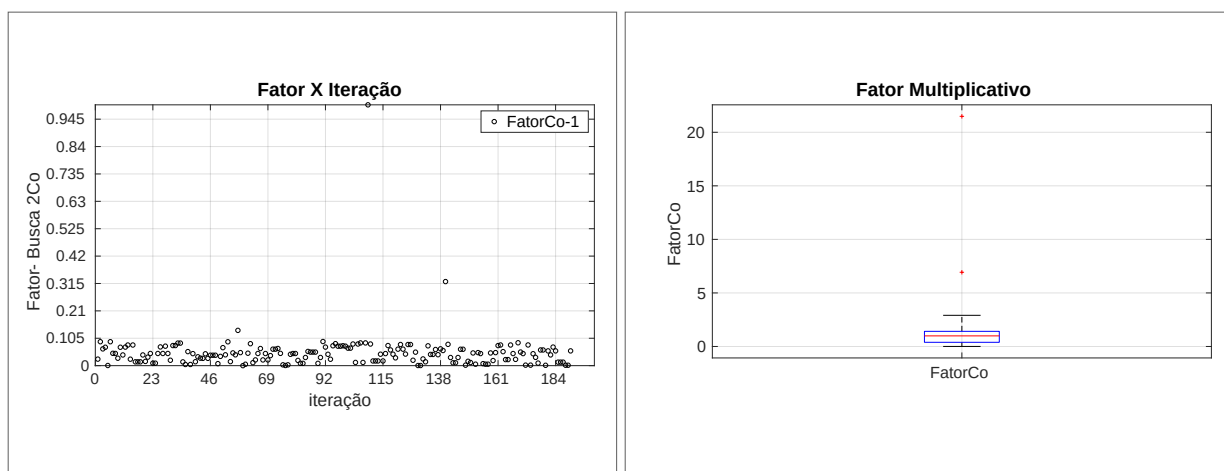


Fonte: Próprio Autor

Constante Modificadora da Biblioteca Omitida

Figura 106 – O gráfico 106a apresenta os valores de constantes que definem a biblioteca omitida, no passo de segunda busca do GRASP, instância dois. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 106b.

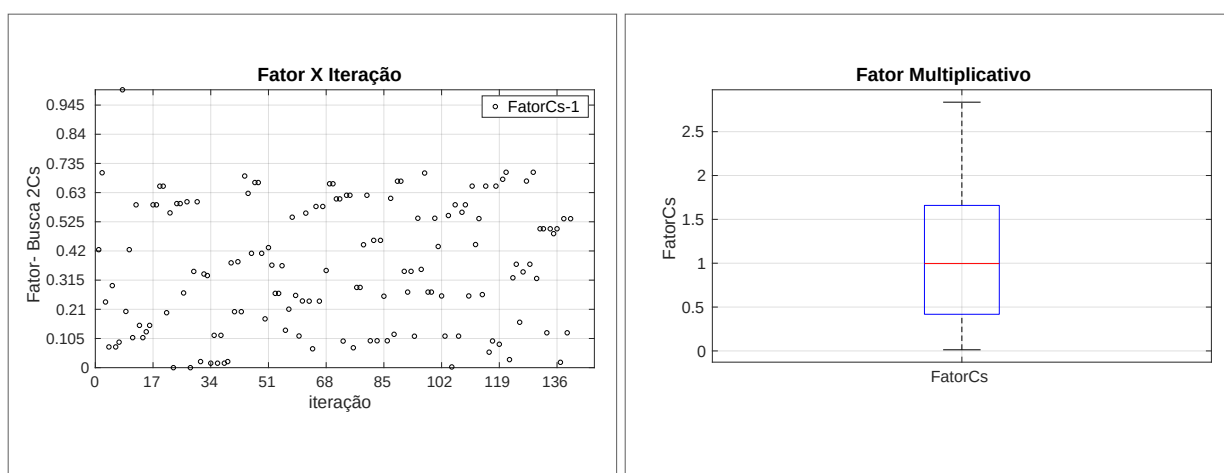
- (a) Valores da constante que modifica a estimativa da biblioteca omitida, instância dois, no passo da segunda busca.
- (b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância dois, passo da segunda busca.



Fonte: Próprio Autor

Figura 107 – O gráfico 107a apresenta os valores de constantes que definem a biblioteca omitida, no passo de segunda busca do GRASP, instância três. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 107b.

- (a) Valores da constante que modifica a estimativa da biblioteca omitida, instância três, no passo da segunda busca.
- (b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância três, passo da segunda busca.

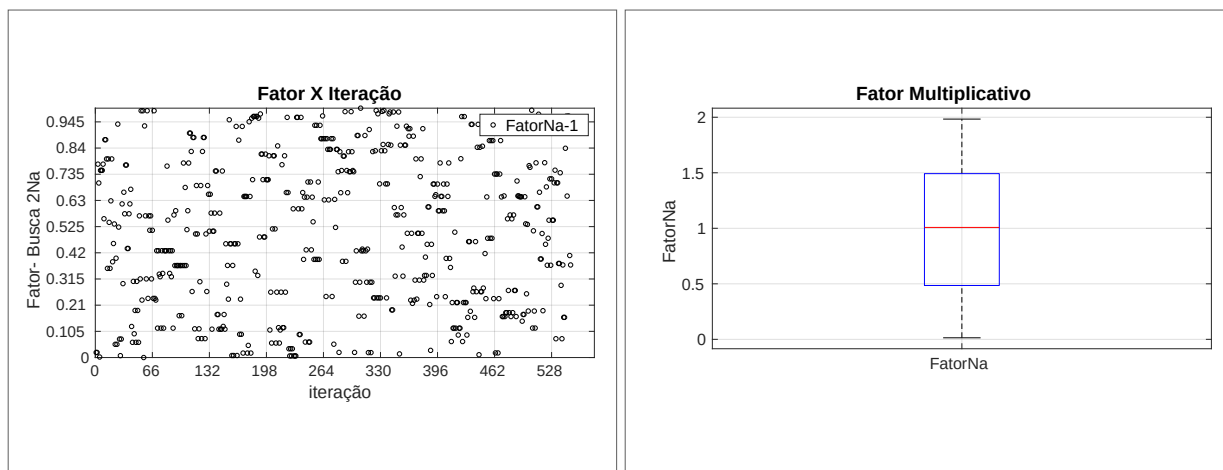


Fonte: Próprio Autor

Figura 108 – O gráfico 108a apresenta os valores de constantes que definem a biblioteca omitida, no passo de segunda busca do GRASP, instância quatro. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 108b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância quatro, no passo da segunda busca.

(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância quatro, passo da segunda busca.



Fonte: Próprio Autor

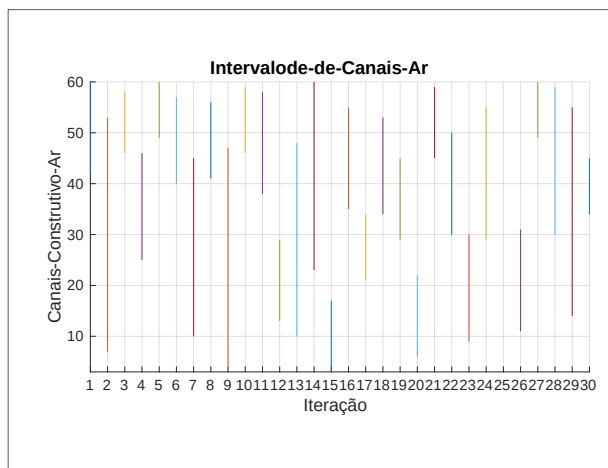
I.2 BASE DE DADOS 2

I.2.1 Resultados do Passo de Construção

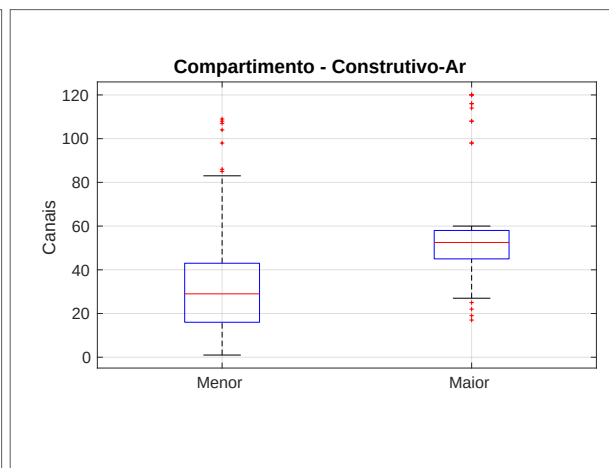
Escolha dos Canais nos Compartimentos

Figura 109 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 109a, para omissão do Ar, no estágio da construção. A figura 109b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.

(a) Intervalo de canais como função da iteração, omitindo Ar, no estágio de construção de uma solução.



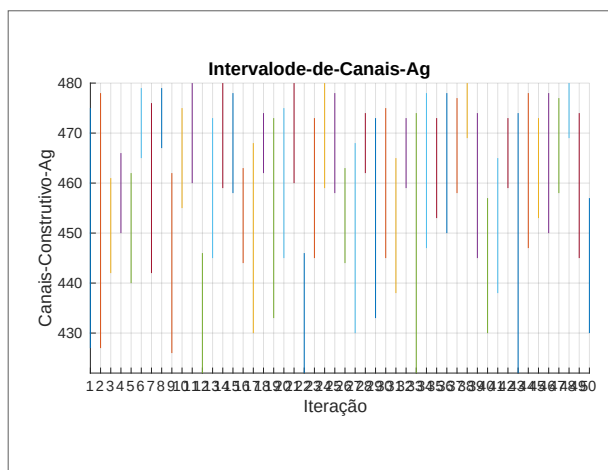
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de construção, instância cinco.



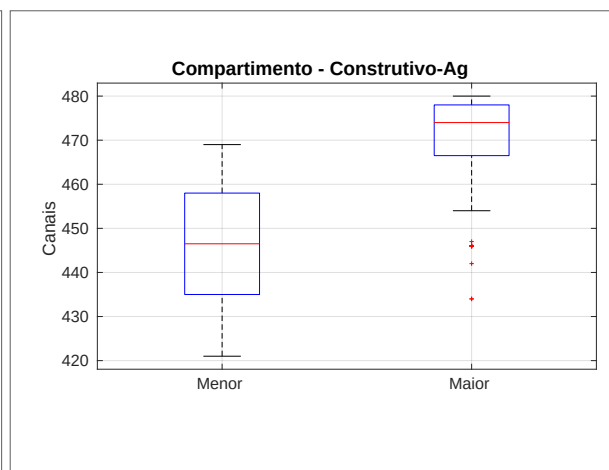
Fonte: Próprio Autor

Figura 110 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 110a, para omissão da Água, no estágio da construção. A figura 110b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.

(a) Intervalo de canais como função da iteração, omitindo Água, no estágio de construção de uma solução.



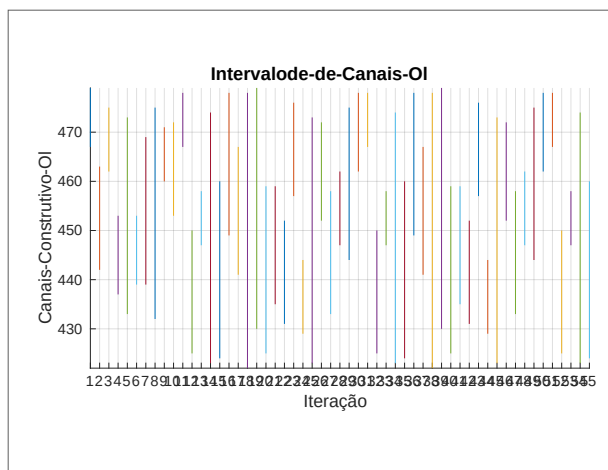
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de construção, instância seis.



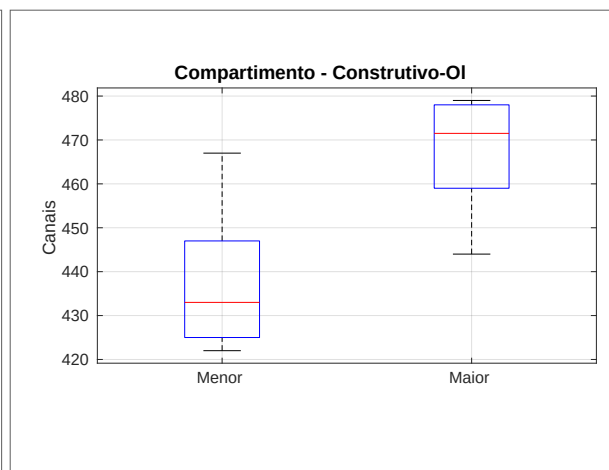
Fonte: Próprio Autor

Figura 111 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 111a, para omissão do Óleo, no estágio da construção. A figura 111b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.

(a) Intervalo de canais como função da iteração, omitindo Óleo, no estágio de construção de uma solução.



(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de construção, instância sete.

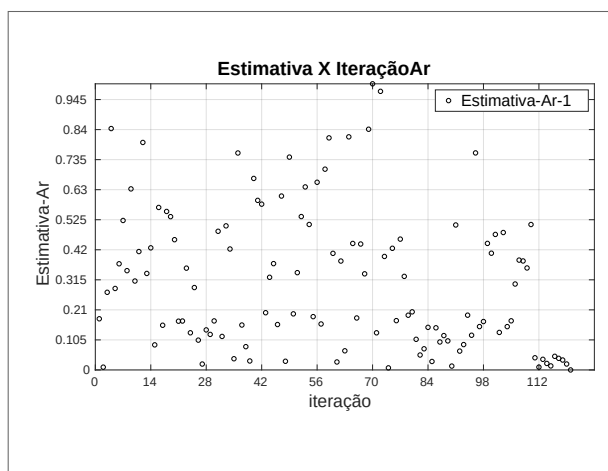


Fonte: Próprio Autor

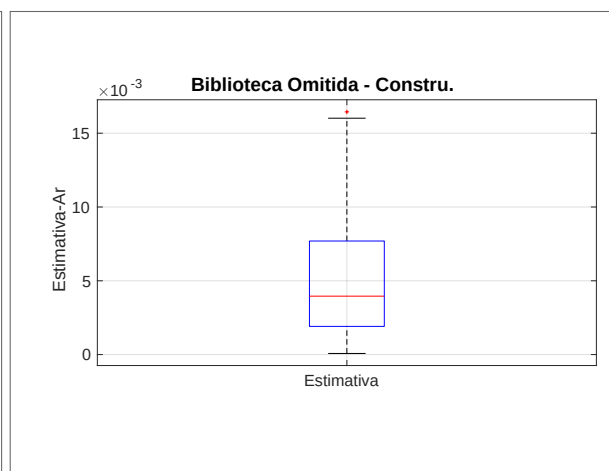
Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 112 – O gráfico 112a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 112b.

(a) Estimativas para o elemento omitido durante o passo construtivo, instância cinco, ausência do Ar.



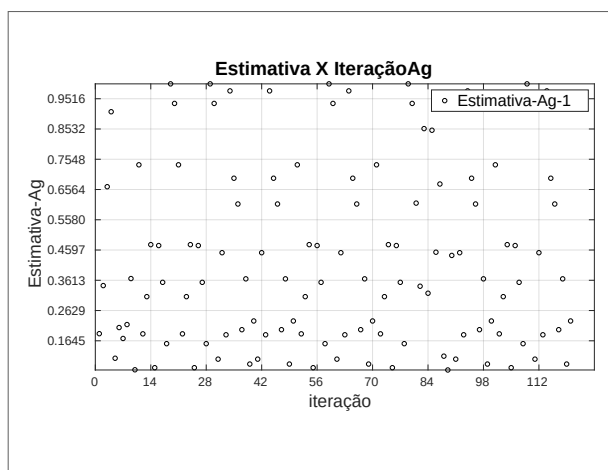
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo construtivo, instância cinco.



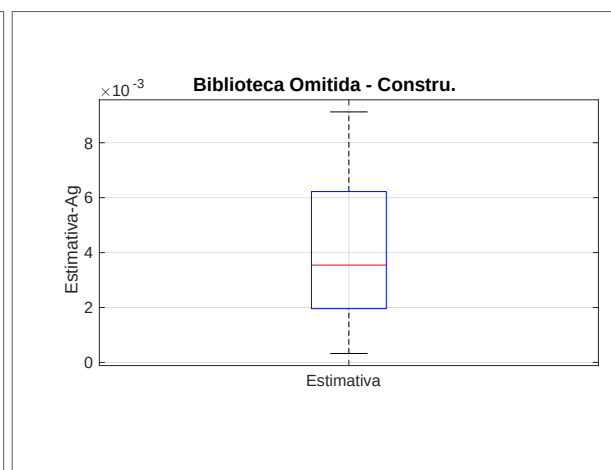
Fonte: Próprio Autor

Figura 113 – O gráfico 113a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 113b.

(a) Estimativas para o elemento omitido durante o passo construtivo, instância seis, ausência da Água.



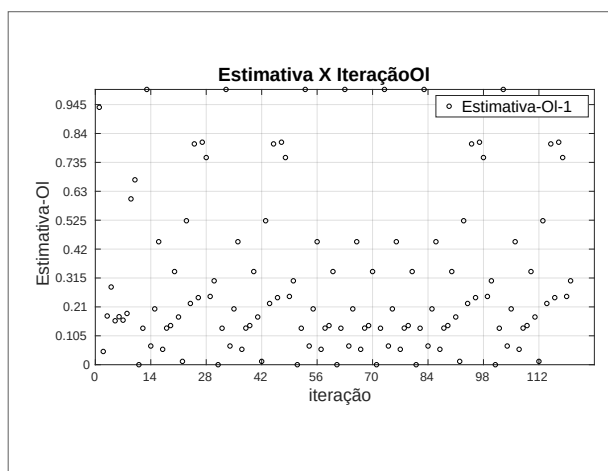
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo construtivo, instância seis.



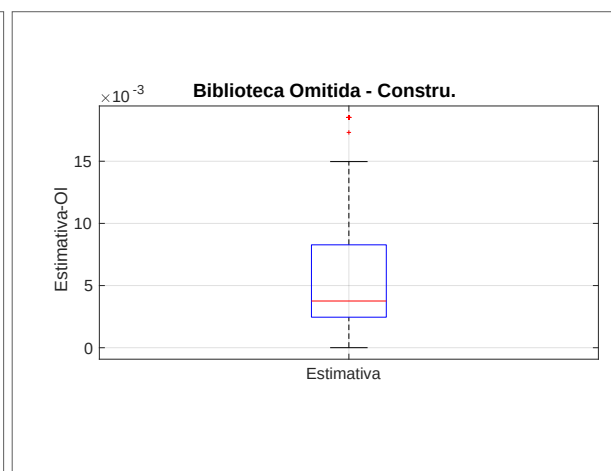
Fonte: Próprio Autor

Figura 114 – O gráfico 114a apresenta os valores de estimativas para a biblioteca omitida, no passo construtivo do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 114b.

(a) Estimativas para o elemento omitido durante o passo construtivo, instância sete, ausência do Óleo.



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo construtivo, instância sete.

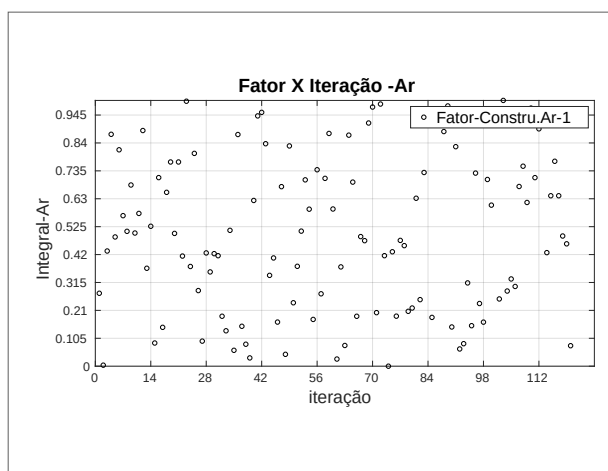


Fonte: Próprio Autor

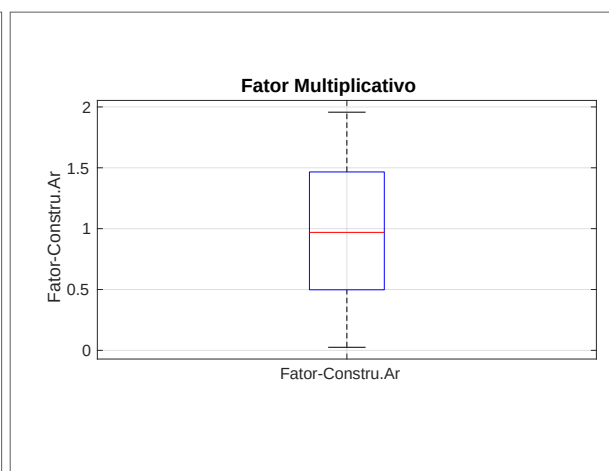
Constante Modificadora da Biblioteca Omitida

Figura 115 – O gráfico 115a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 115b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância cinco, no passo construtivo.



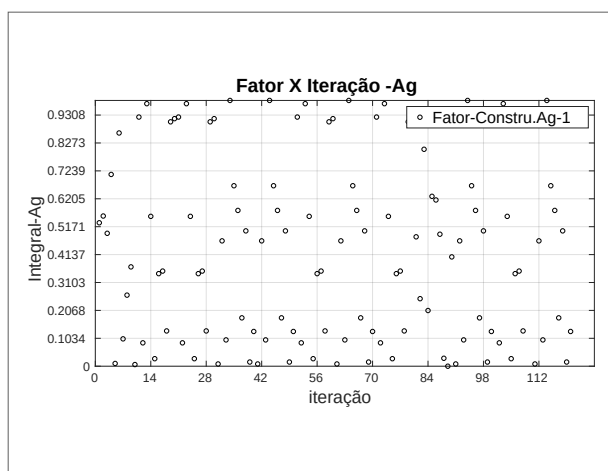
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância cinco, passo construtivo.



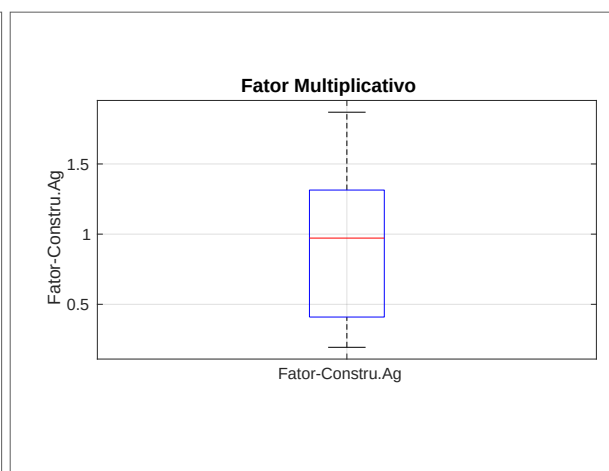
Fonte: Próprio Autor

Figura 116 – O gráfico 116a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 116b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância seis, no passo construtivo.



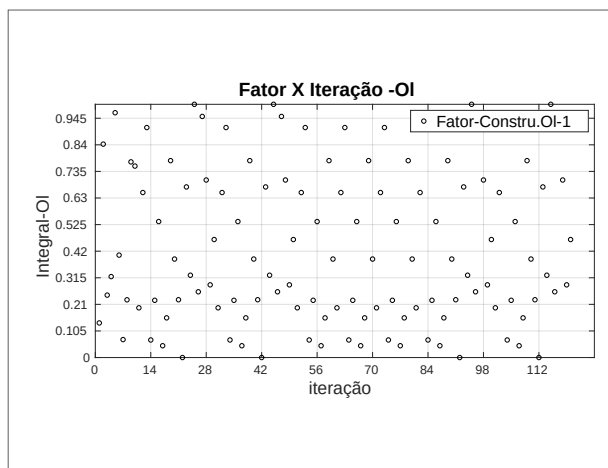
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância seis, passo construtivo.



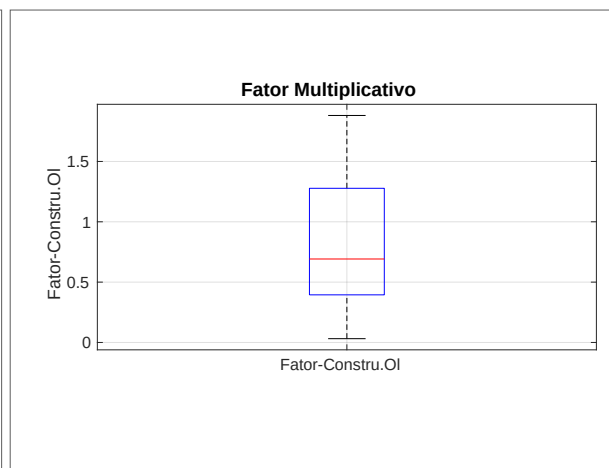
Fonte: Próprio Autor

Figura 117 – O gráfico 117a apresenta os valores de constantes que definem a biblioteca omitida, no passo construtivo do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 117b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância sete, no passo construtivo.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância sete, passo construtivo.



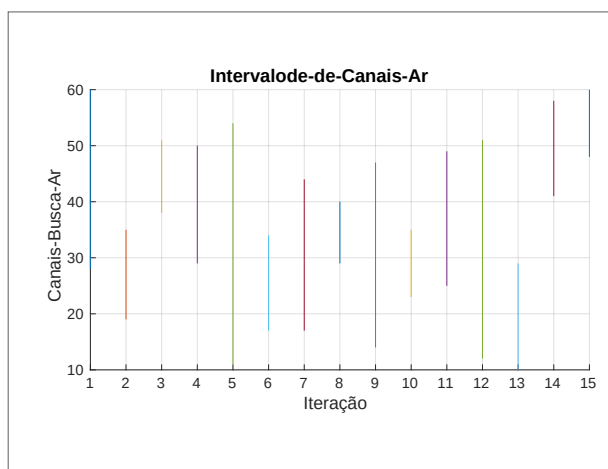
Fonte: Próprio Autor

I.2.2 Resultados do Passo de Busca

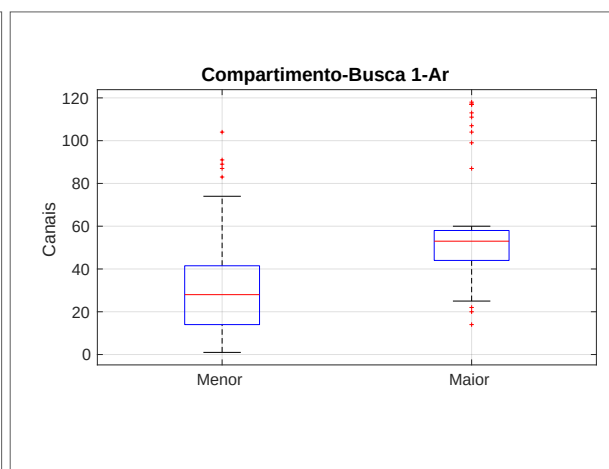
Escolha dos Canais nos Compartimentos

Figura 118 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 118a, para omissão do Ar, no estágio da busca. A figura 118b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.

(a) Intervalo de canais como função da iteração, omitindo Ar, no estágio de Busca de uma solução.



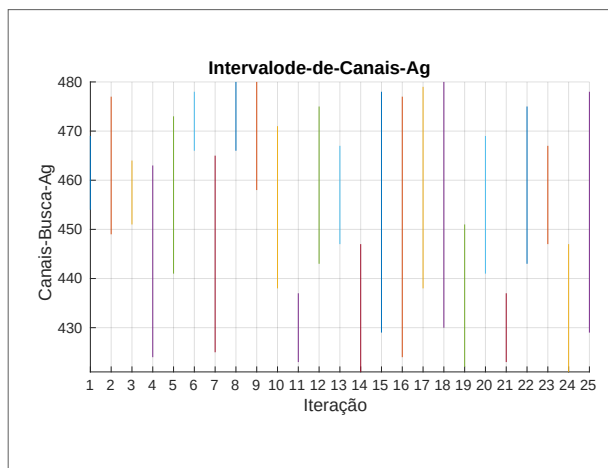
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de busca, instância cinco.



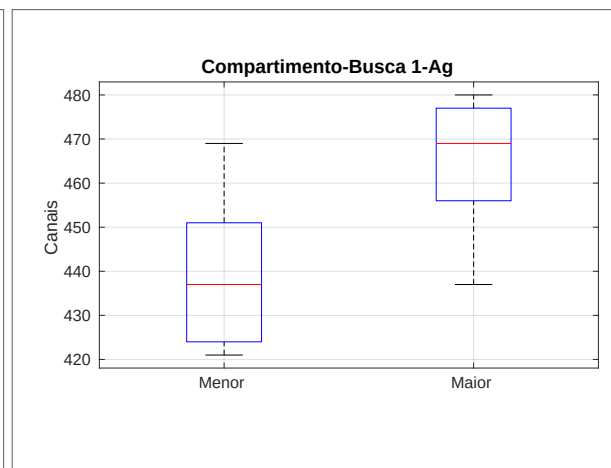
Fonte: Próprio Autor

Figura 119 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 119a, para omissão da Água, no estágio da busca. A figura 119b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.

(a) Intervalo de canais como função da iteração, omitindo Água, no estágio de Busca de uma solução.



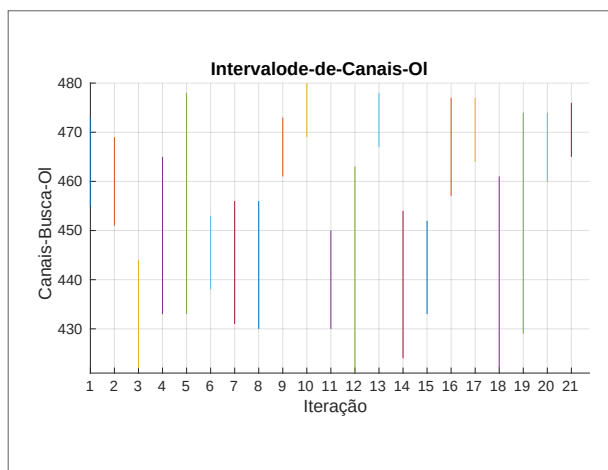
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de busca, instância seis.



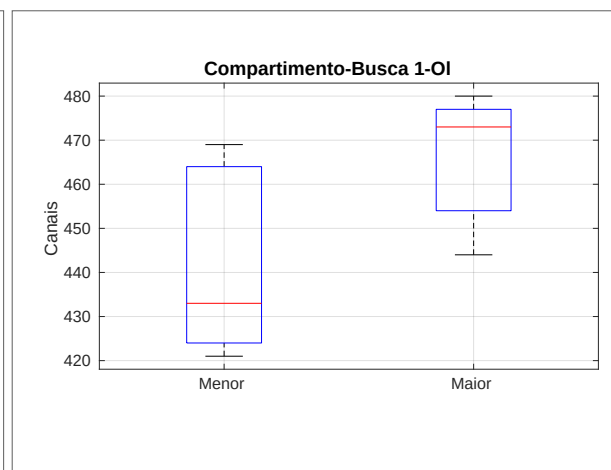
Fonte: Próprio Autor

Figura 120 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 120a, para omissão da Água, no estágio da busca. A figura 120b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.

(a) Intervalo de canais como função da iteração, omitindo Óleo, no estágio de Busca de uma solução.



(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de busca, instância sete.



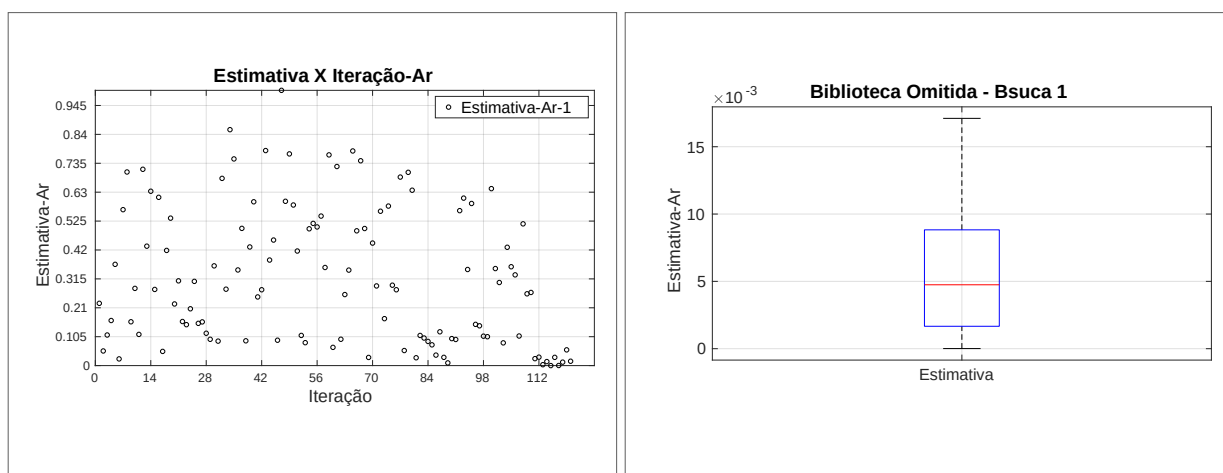
Fonte: Próprio Autor

Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 121 – O gráfico 121a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 121b.

(a) Estimativas para o elemento omitido durante o passo de busca, instância cinco, ausência do Ar.

(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de busca, instância cinco.

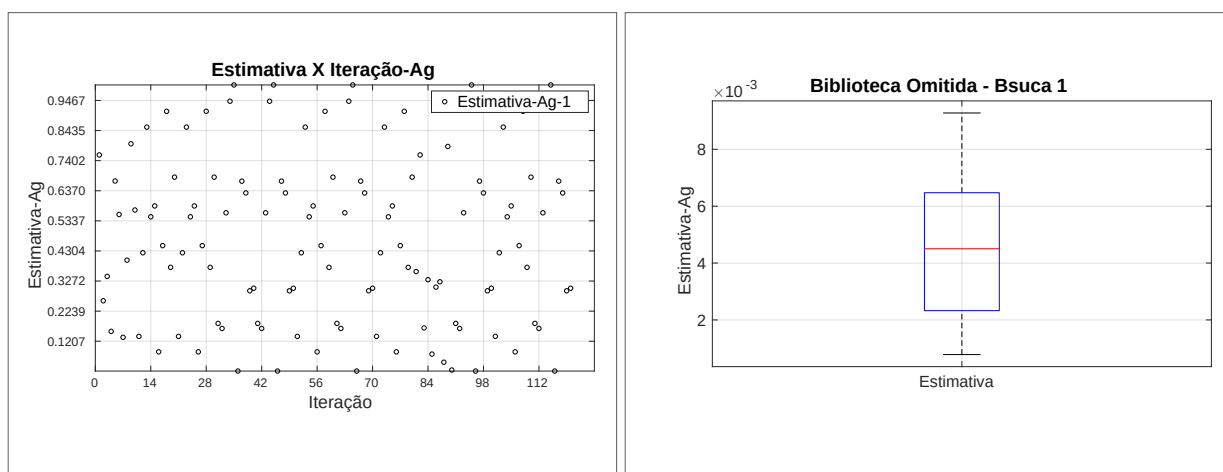


Fonte: Próprio Autor

Figura 122 – O gráfico 122a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 122b.

(a) Estimativas para o elemento omitido durante o passo de busca, instância seis, ausência da Água.

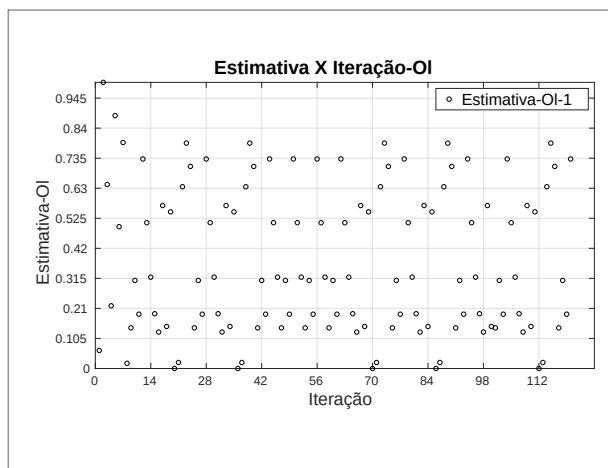
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de busca, instância seis.



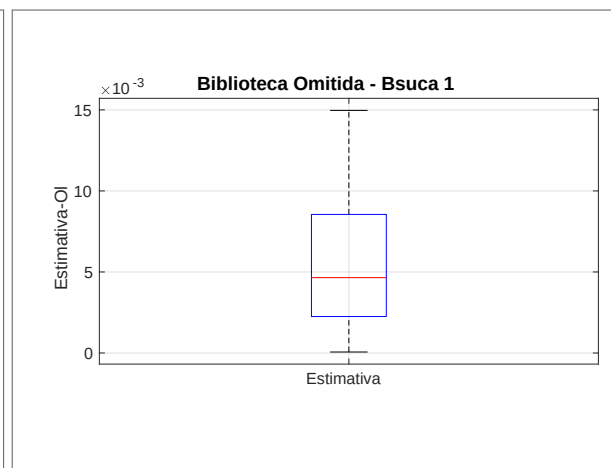
Fonte: Próprio Autor

Figura 123 – O gráfico 123a apresenta os valores de estimativas para a biblioteca omitida, no passo de busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 123b.

(a) Estimativas para o elemento omitido durante o passo de busca, instância sete, ausência do Óleo.



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de busca, instância sete.

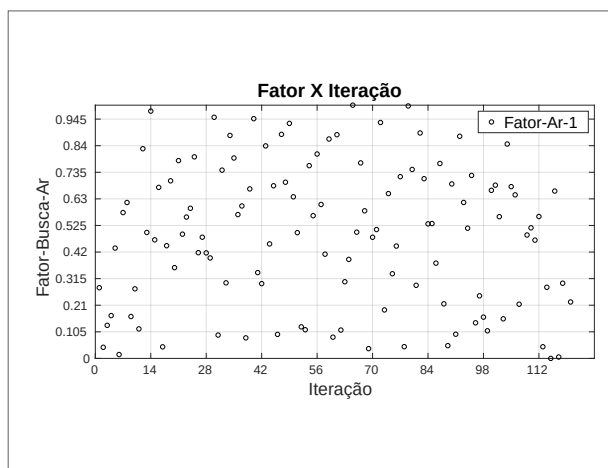


Fonte: Próprio Autor

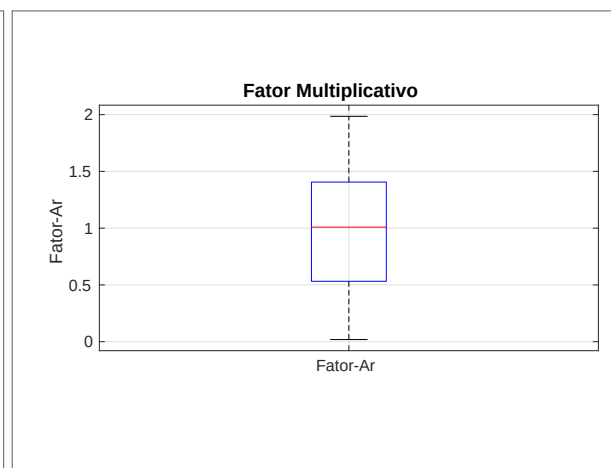
Constante Modificadora da Biblioteca Omitida

Figura 124 – O gráfico 124a apresenta os valores de constantes que definem a biblioteca omitida, no passo da primeira busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 124b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância cinco, no passo de busca.



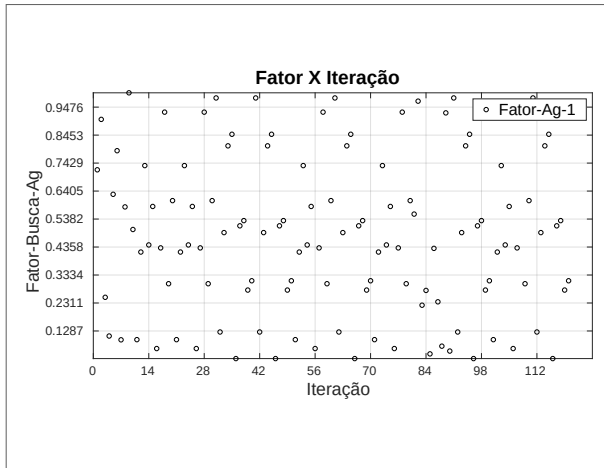
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância cinco, passo de busca.



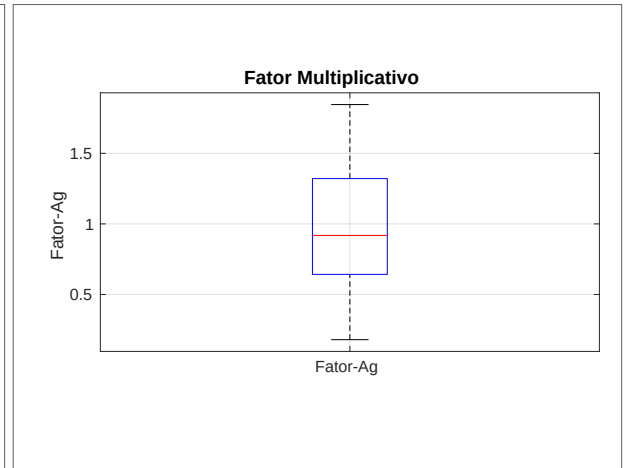
Fonte: Próprio Autor

Figura 125 – O gráfico 125a apresenta os valores de constantes que definem a biblioteca omitida, no passo da primeira busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 125b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância seis, no passo de busca.



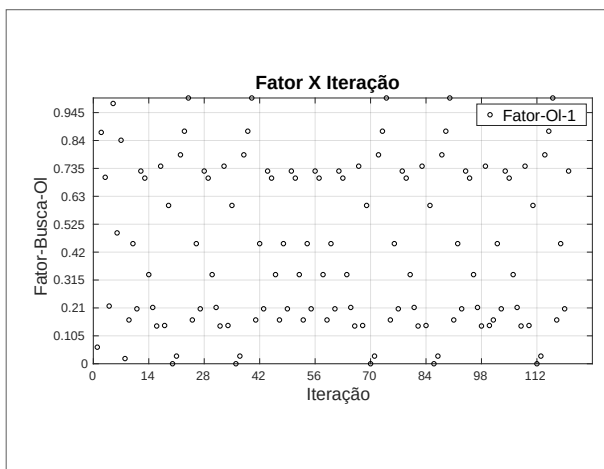
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância seis, passo de busca.



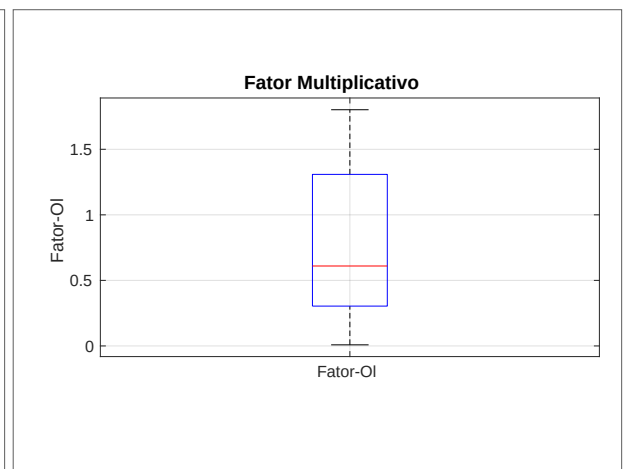
Fonte: Próprio Autor

Figura 126 – O gráfico 126a apresenta os valores de constantes que definem a biblioteca omitida, no passo da primeira busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 126b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância sete, no passo de busca.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância sete, passo de busca.



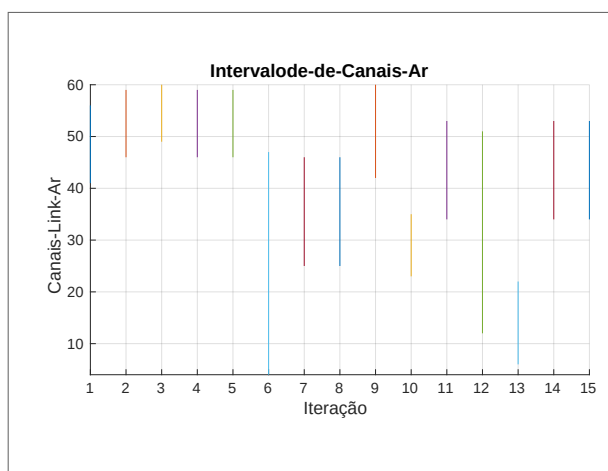
Fonte: Próprio Autor

I.2.3 Resultados do Passo de Linkagem

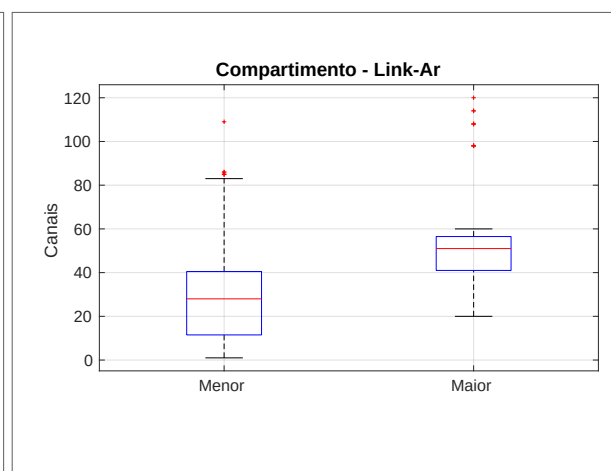
Escolha dos Canais nos Compartimentos

Figura 127 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 127a, para omissão do Ar, no estágio da linkagem. A figura 127b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.

(a) Intervalo de canais como função da iteração, omitindo Ar, no estágio de Linkagem de uma solução.



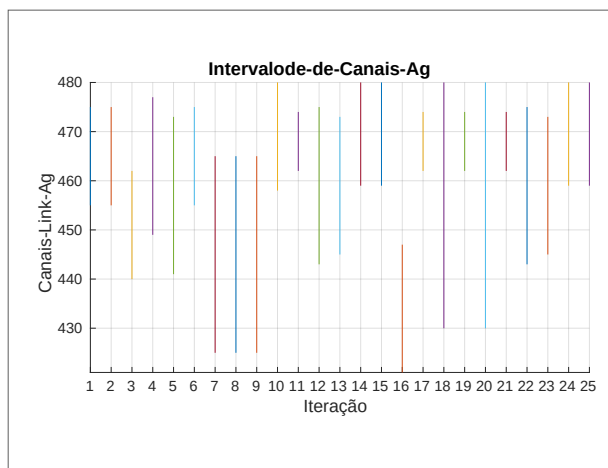
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de Linkagem, instância cinco.



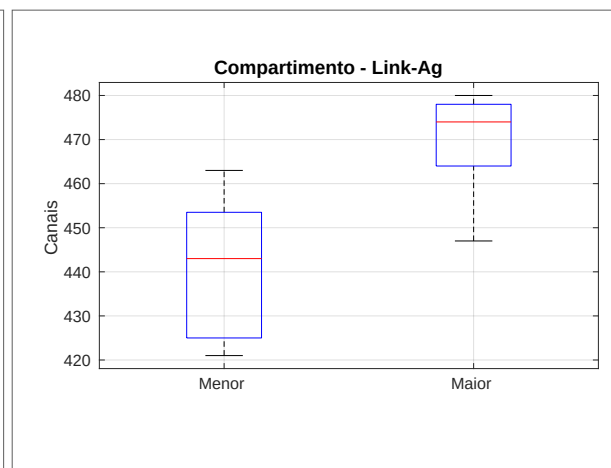
Fonte: Próprio Autor

Figura 128 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 128a, para omissão da Água, no estágio da linkagem. A figura 128b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.

(a) Intervalo de canais como função da iteração, omitindo Água, no estágio de Linkagem de uma solução.



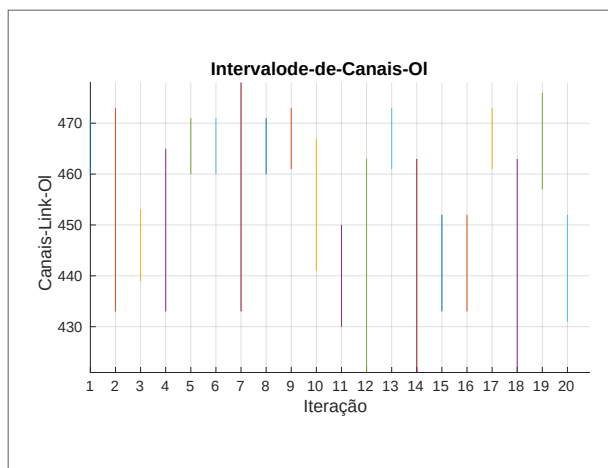
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de Linkagem, instância seis.



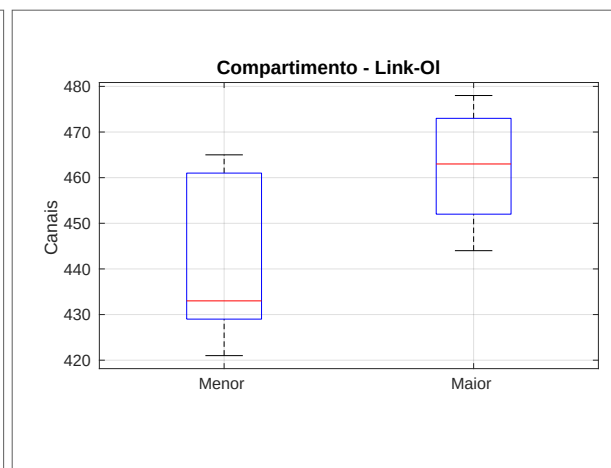
Fonte: Próprio Autor

Figura 129 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 129a, para omissão do Óleo, no estágio da linkagem. A figura 129b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.

(a) Intervalo de canais como função da iteração, omitindo Óleo, no estágio de linkagem de uma solução.



(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de Linkagem, instância sete.

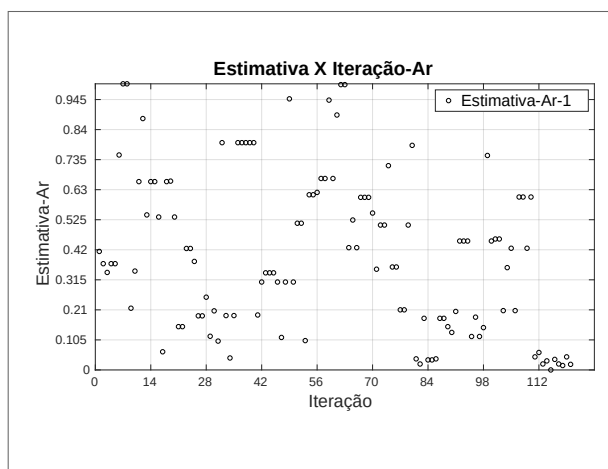


Fonte: Próprio Autor

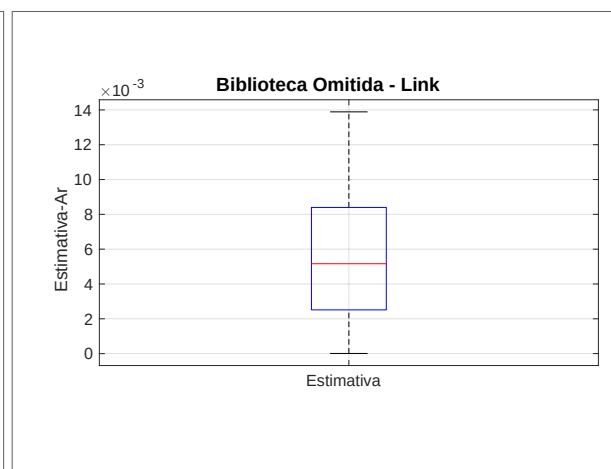
Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 130 – O gráfico 130a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 130b.

(a) Estimativas para o elemento omitido durante o passo de linkagem, instância cinco, ausência do Ar.



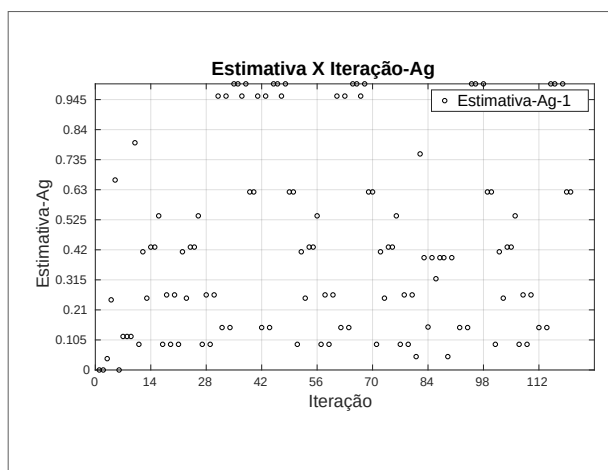
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de linkagem, instância cinco.



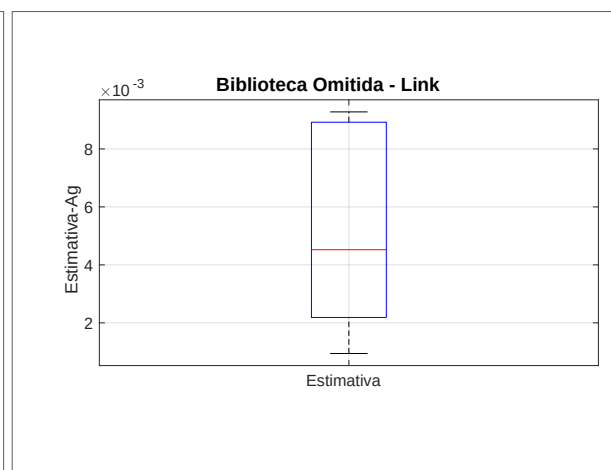
Fonte: Próprio Autor

Figura 131 – O gráfico 131a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 131b.

(a) Estimativas para o elemento omitido durante o passo de linkagem, instância seis, ausência da Água.



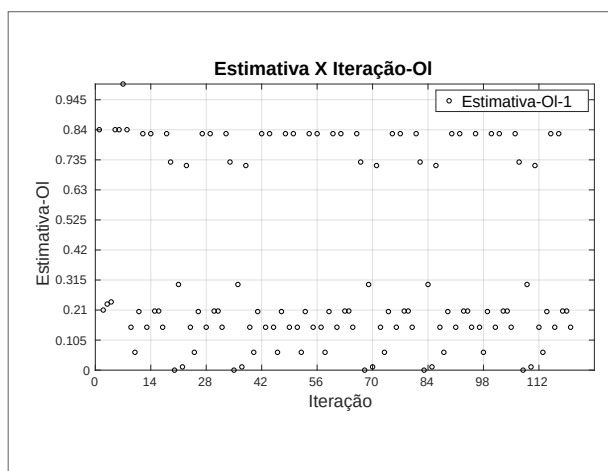
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de linkagem, instância seis.



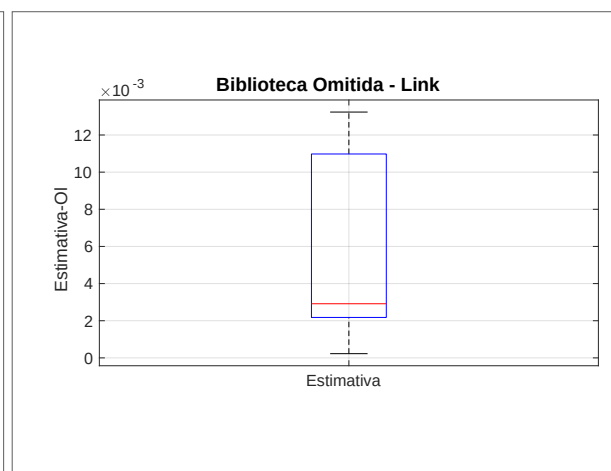
Fonte: Próprio Autor

Figura 132 – O gráfico 132a apresenta os valores de estimativas para a biblioteca omitida, no passo de linkagem do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 132b.

(a) Estimativas para o elemento omitido durante o passo de linkagem, instância sete, ausência do Óleo.



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de linkagem, instância sete.

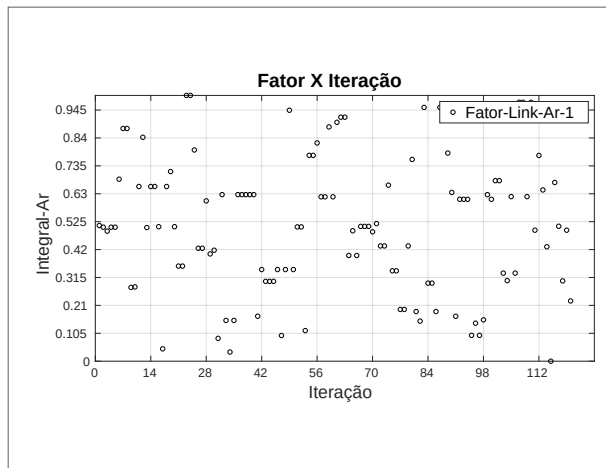


Fonte: Próprio Autor

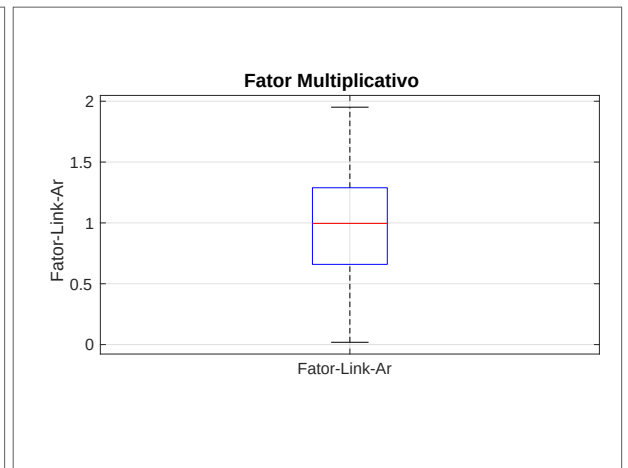
Constante Modificadora da Biblioteca Omitida

Figura 133 – O gráfico 133a apresenta os valores de constantes que definem a biblioteca omitida, no passo da linkagem do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 133b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância cinco, no passo de linkagem.



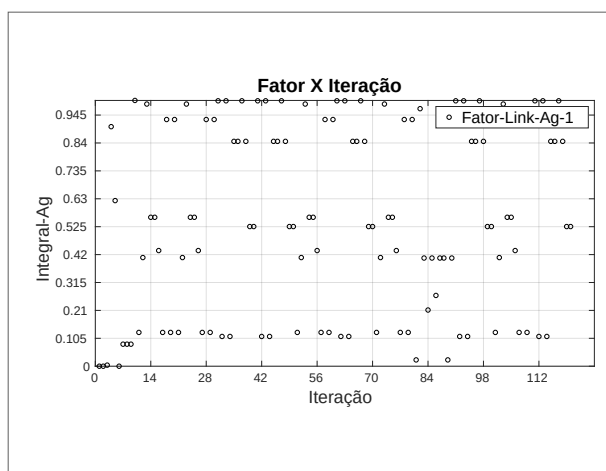
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância cinco, passo de linkagem.



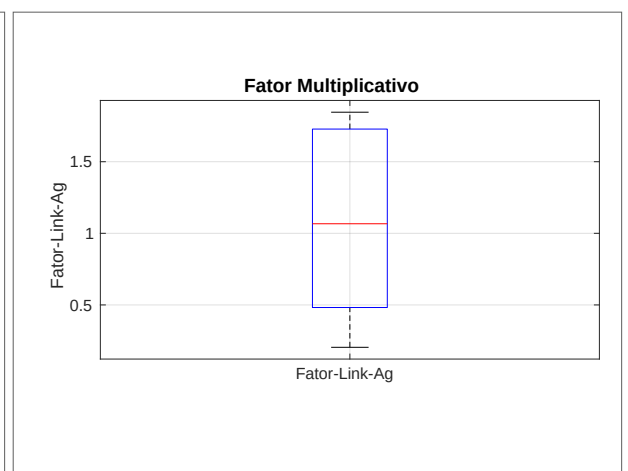
Fonte: Próprio Autor

Figura 134 – O gráfico 134a apresenta os valores de constantes que definem a biblioteca omitida, no passo da linkagem do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 134b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância seis, no passo de linkagem.



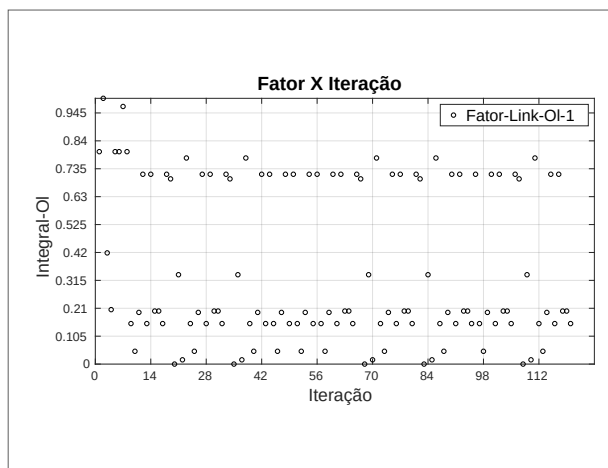
(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância seis, passo de linkagem.



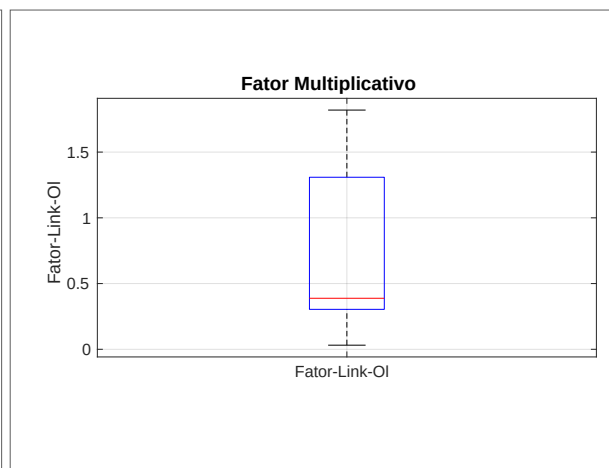
Fonte: Próprio Autor

Figura 135 – O gráfico 135a apresenta os valores de constantes que definem a biblioteca omitida, no passo da linkagem do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 135b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância sete, no passo de linkagem.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância sete, passo de linkagem.



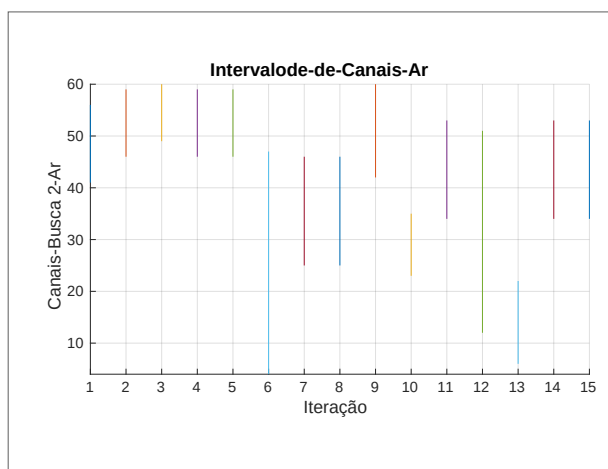
Fonte: Próprio Autor

1.2.4 Resultados do Passo da segunda Busca

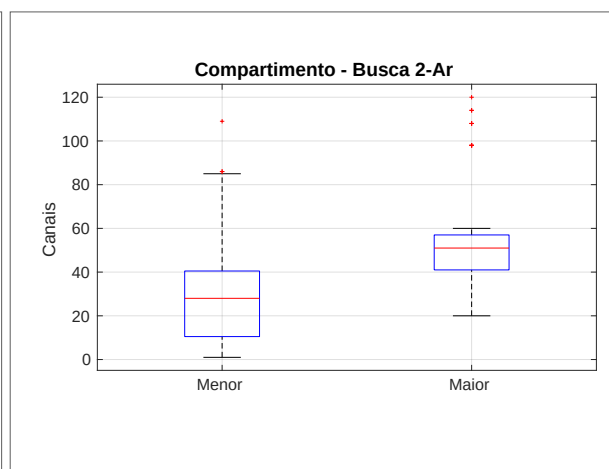
Escolha dos Canais nos Compartimentos

Figura 136 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 136a, para omissão do Ar, no estágio da Segunda Busca. A figura 136b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Ar.

(a) Intervalo de canais como função da iteração, omitindo Ar, no estágio da Segunda Busca de uma solução.



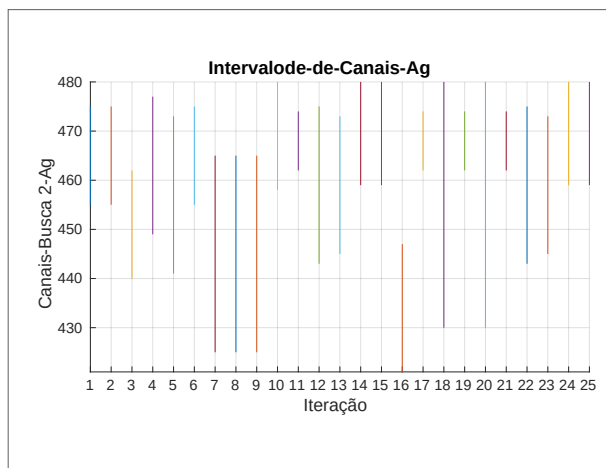
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de Segunda Busca, instância cinco.



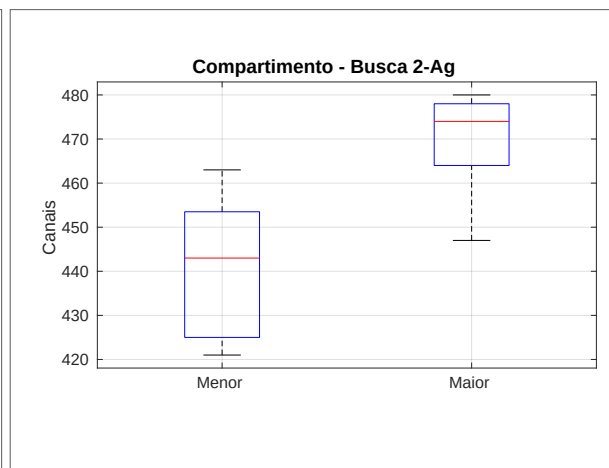
Fonte: Próprio Autor

Figura 137 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 137a, para omissão da Água, no estágio da Segunda Busca. A figura 137b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo da Água.

(a) Intervalo de canais como função da iteração, omitindo Água, no estágio de segunda Busca de uma solução.



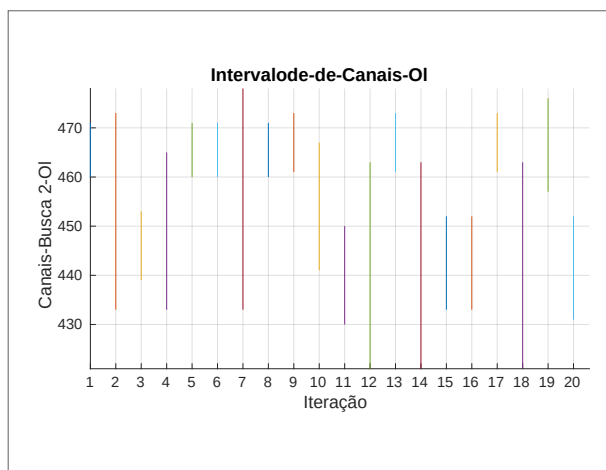
(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de Segunda Busca, instância seis.



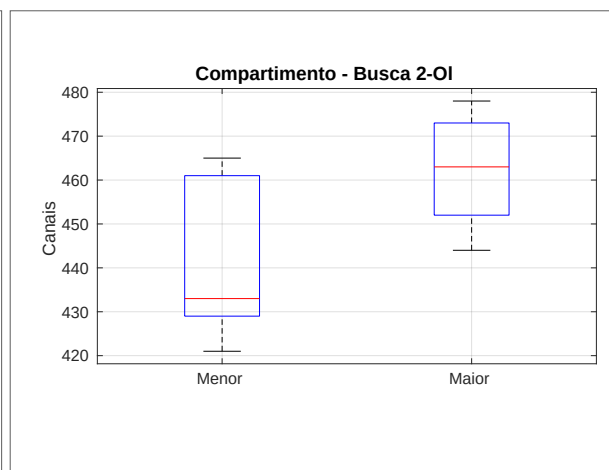
Fonte: Próprio Autor

Figura 138 – Linhas verticais indicam os canais usados na definição do compartimento na partição de maior resíduo, Figura 138a, para omissão do Óleo, no estágio da Segunda Busca. A figura 138b traz as distribuições dos canais que definem o compartimento da partição de maior resíduo do Óleo.

(a) Intervalo de canais como função da iteração, omitindo Óleo, no estágio de segunda Busca de uma solução.



(b) Distribuição de canais no compartimento na partição de maior resíduo, estágio de Segunda Busca, instância sete.

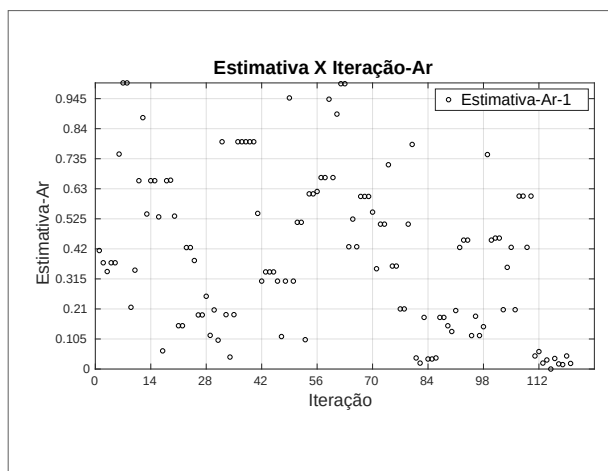


Fonte: Próprio Autor

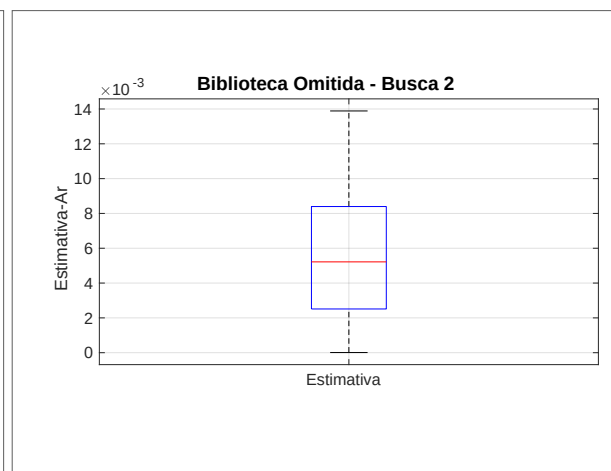
Estimativa da Biblioteca Omitida na Partição de Maior resíduo

Figura 139 – O gráfico 139a apresenta os valores de estimativas para a biblioteca omitida, no passo de Segunda Busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 139b.

(a) Estimativas para o elemento omitido durante o passo de Segunda Busca, instância cinco, ausência do Ar.



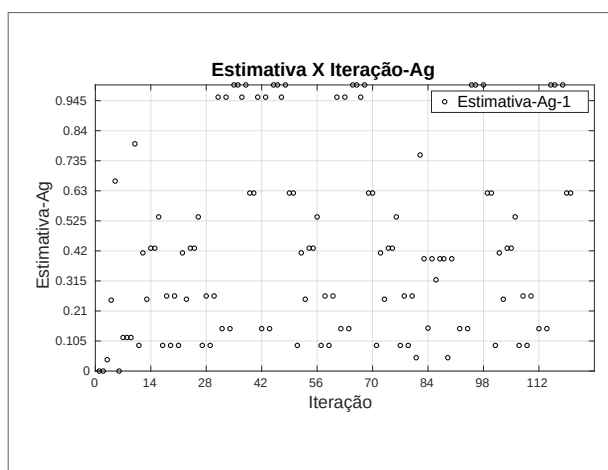
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de Segunda Busca, instância cinco.



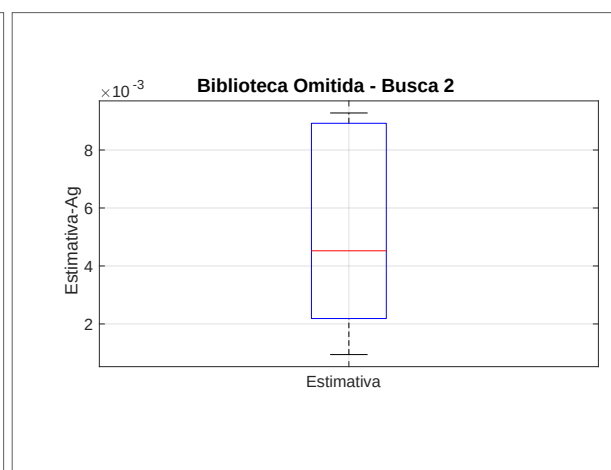
Fonte: Próprio Autor

Figura 140 – O gráfico 140a apresenta os valores de estimativas para a biblioteca omitida, no passo de Segunda Busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 140b.

(a) Estimativas para o elemento omitido durante o passo de Segunda Busca, instância seis, ausência da Água.



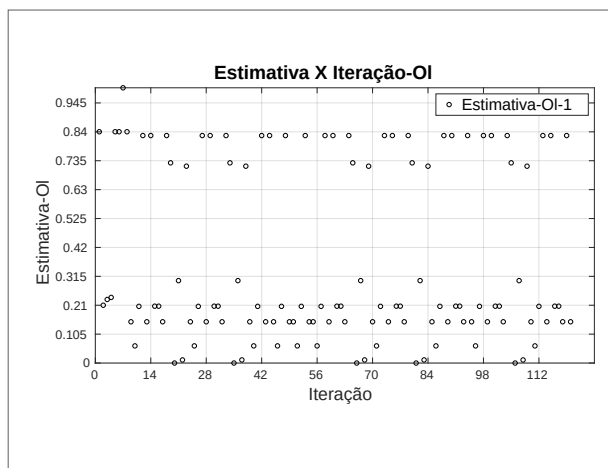
(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de Segunda Busca, instância seis.



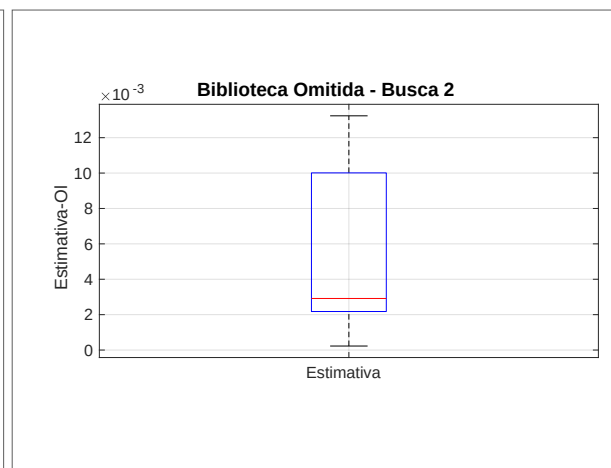
Fonte: Próprio Autor

Figura 141 – O gráfico 141a apresenta os valores de estimativas para a biblioteca omitida, no passo de Segunda Busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 141b.

(a) Estimativas para o elemento omitido durante o passo de Segunda Busca, instância sete, ausência do Óleo.



(b) Distribuição para os valores de estimativa da biblioteca omitida durante o passo de Segunda Busca, instância sete.

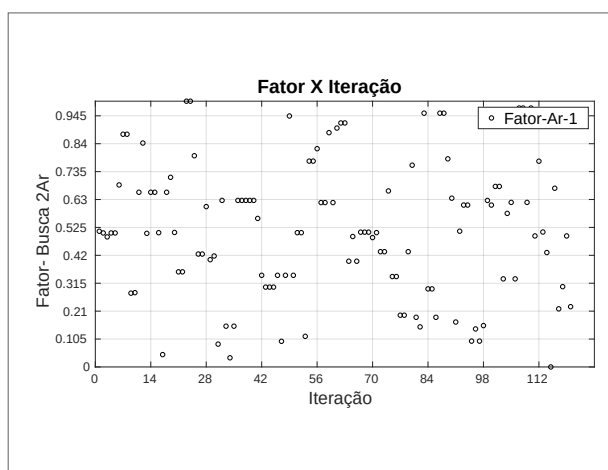


Fonte: Próprio Autor

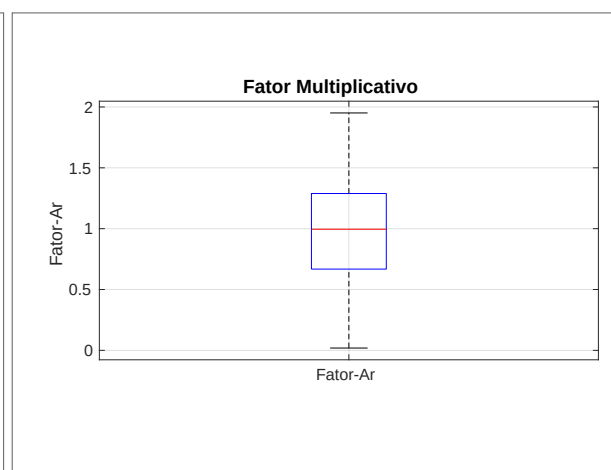
Constante Modificadora da Biblioteca Omitida

Figura 142 – O gráfico 142a apresenta os valores de constantes que definem a biblioteca omitida, no passo da Segunda Busca do GRASP, instância cinco. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 142b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância cinco, no passo da Segunda Busca.



(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância cinco, passo da Segunda Busca.

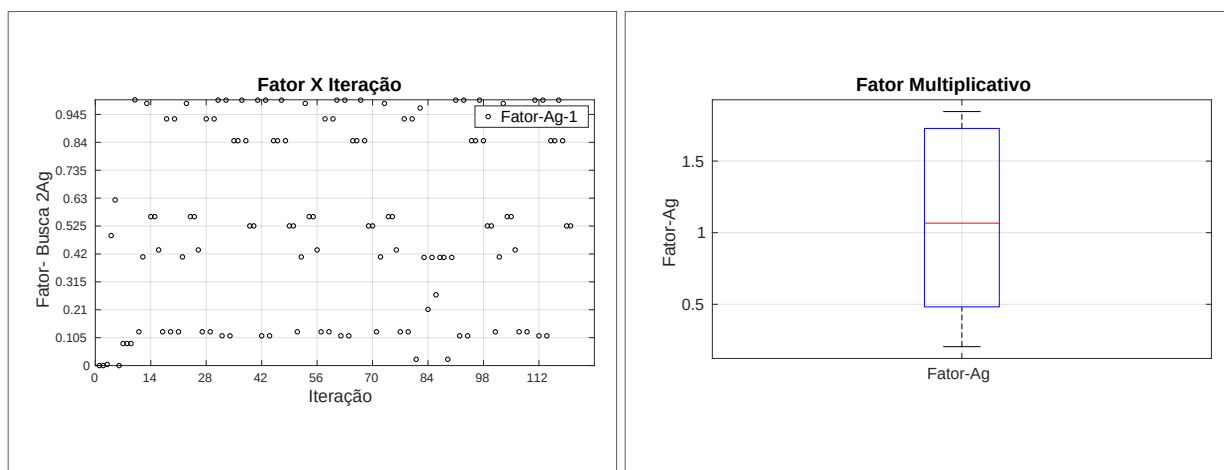


Fonte: Próprio Autor

Figura 143 – O gráfico 143a apresenta os valores de constantes que definem a biblioteca omitida, no passo da Segunda Busca do GRASP, instância seis. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 143b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância seis, no passo da Segunda Busca.

(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância seis, passo da Segunda Busca.

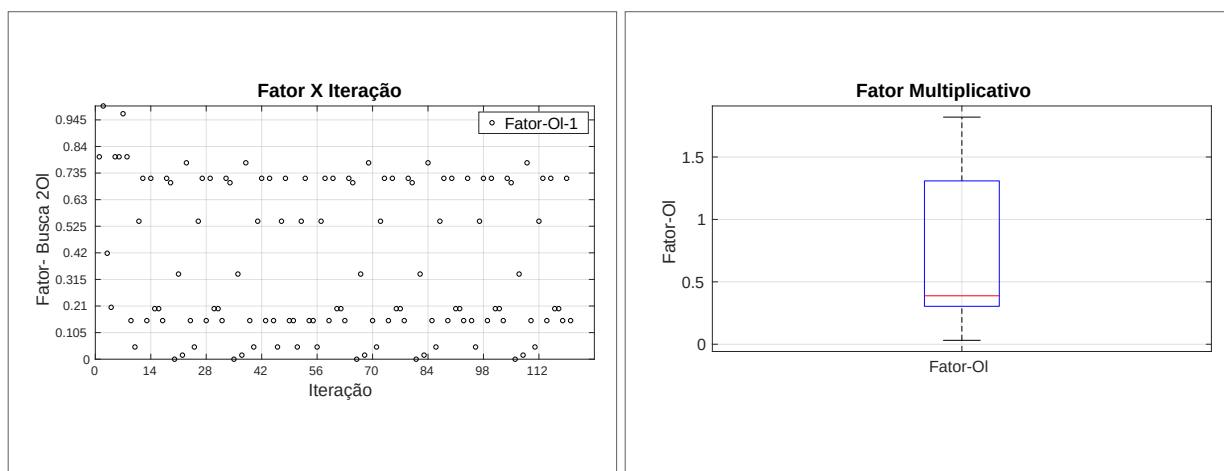


Fonte: Próprio Autor

Figura 144 – O gráfico 144a apresenta os valores de constantes que definem a biblioteca omitida, no passo da Segunda Busca do GRASP, instância sete. As distribuições para esses valores, durante 30 execuções do nosso algoritmo, está demonstrado na Figura 144b.

(a) Valores da constante que modifica a estimativa da biblioteca omitida, instância sete, no passo da Segunda Busca.

(b) Distribuição das constantes que modificam as estimativas da biblioteca omitida, instância sete, passo da Segunda Busca.



Fonte: Próprio Autor

APÊNDICE J – RESULTADOS CGRASP

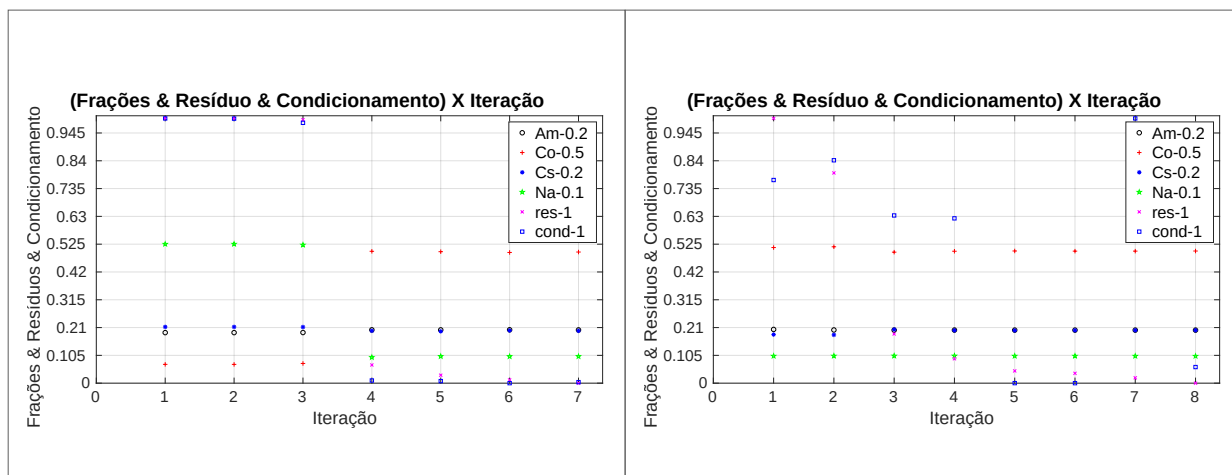
J.1 RESULTADOS CGRASP - BASE DE DADOS 1

J.1.1 Resultados CGRASP - Execução única

Figura 145 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 145a e 145b os valores de frações de contagem percentual, resíduo e condicionamento, como função da iteração, para omissão do Co . Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. O ponto $\circ$ indica as aproximações para o Am ; $+$ as aproximações para Co ; $*$ as aproximações para o Cs ; $\star$ as aproximações para o Na ; $\times$ os valores para o resíduo e $\square$ os valores para número de condicionamento.

(a) Percentuais dos elementos e respectivas métricas. Aproximações Co

(b) Percentuais dos elementos e respectivas métricas. Aproximações Co

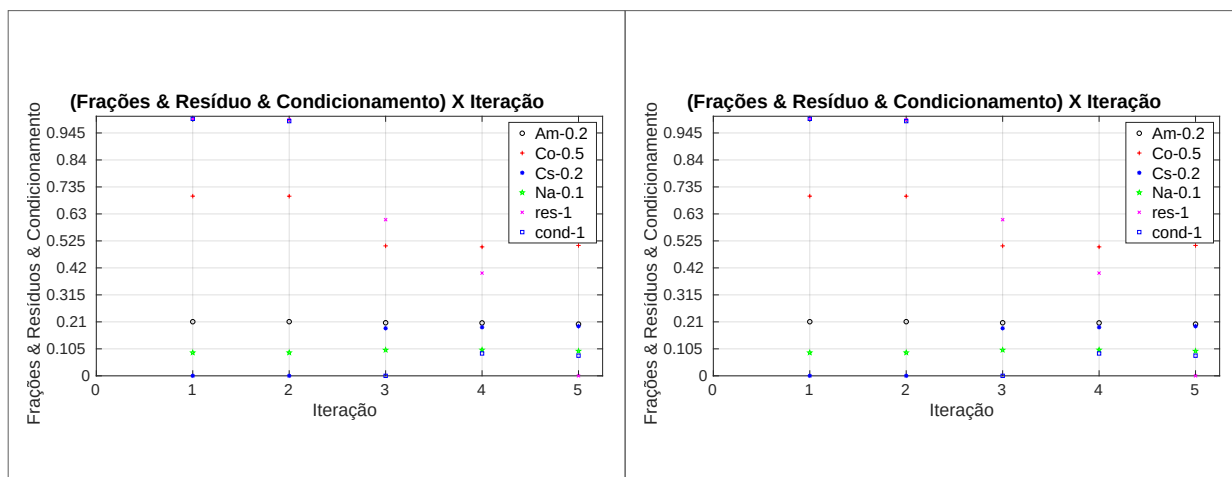


Fonte: Próprio Autor

Figura 146 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 146a e 146b os valores de frações de contagem percentual, resíduo e condicionamento, como função da iteração, para omissão do Cs . Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. O ponto $\circ$ indica as aproximações para o Am ; $+$ as aproximações para Co ; $*$ as aproximações para o Cs ; $\star$ as aproximações para o Na ; $\times$ os valores para o resíduo e $\square$ os valores para número de condicionamento.

(a) Percentuais dos elementos e respectivas métricas. Aproximações Cs

(b) Percentuais dos elementos e respectivas métricas. Aproximações Cs .

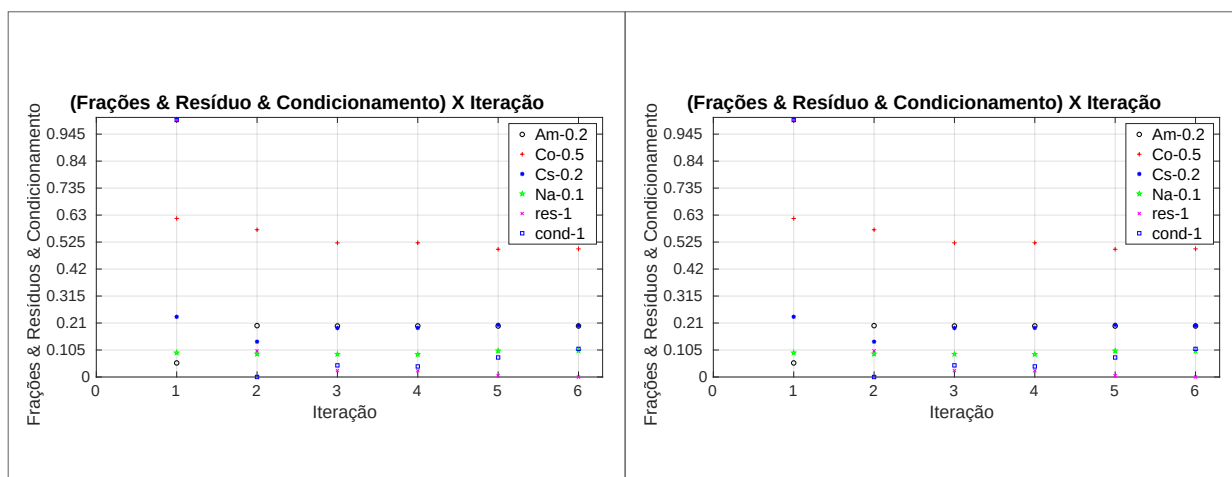


Fonte: Próprio Autor

Figura 147 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas Figuras 147a e 147b os valores de frações de contagem percentual, resíduo e condicionamento, como função da iteração, para omissão do Na . Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. O ponto $\circ$ indica as aproximações para o Am ; $+$ as aproximações para Co ; $*$ as aproximações para o Cs ; $\star$ as aproximações para o Na ; $\times$ os valores para o resíduo e $\square$ os valores para número de condicionamento.

(a) Percentuais dos elementos e respectivas métricas. Aproximações Na

(b) Percentuais dos elementos e respectivas métricas. Aproximações Na

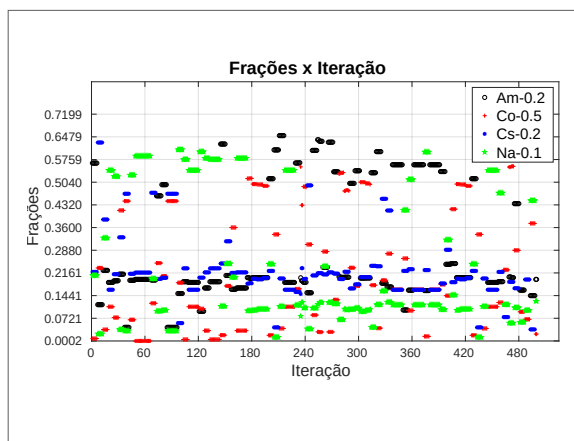


Fonte: Próprio Autor

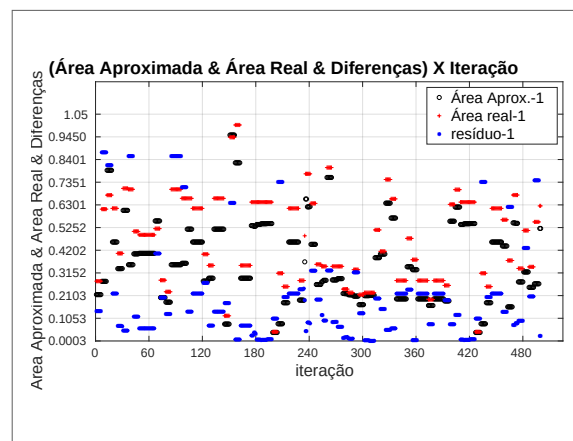
J.1.2 Resultados CGRASP - Etapa Construtiva

Figura 148 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 148a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (Am ○; Co +; Cs *); na Figura 148b(○); área real +; resíduo *) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 148c tem-se os valores de estimativa do composto omitido (Co) e a aproximação dos demais compostos.

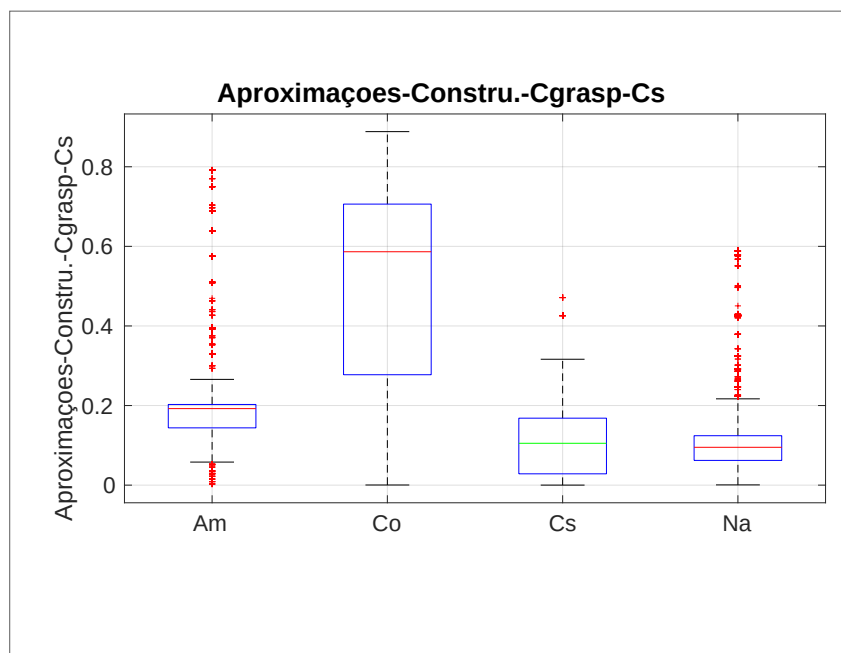
(a) Frações de contagem percentual dos isótopos inseridos no espectro total na omissão do Co . Pontos: Am ○; Co +; Cs *; Na ☆.



(b) Medidas da área abaixo da curva do espectro total. Área calculada ○; área real +; resíduo *.



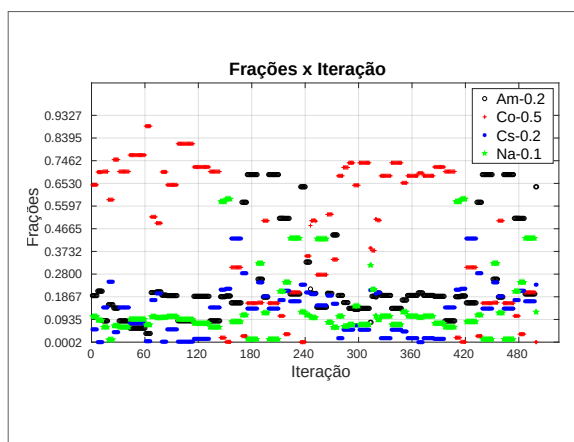
(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP.



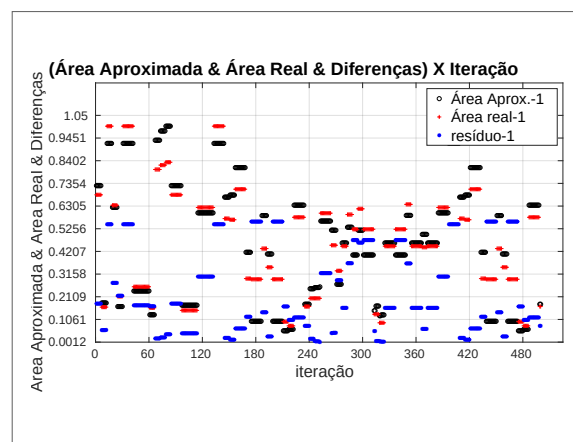
Fonte: Próprio Autor

Figura 149 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 149a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (Am ○; Co +; Cs *); na Figura 149b(○); área real +; resíduo *) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 149c tem-se os valores de estimativa do composto omitido (Cs) e a aproximação dos demais compostos.

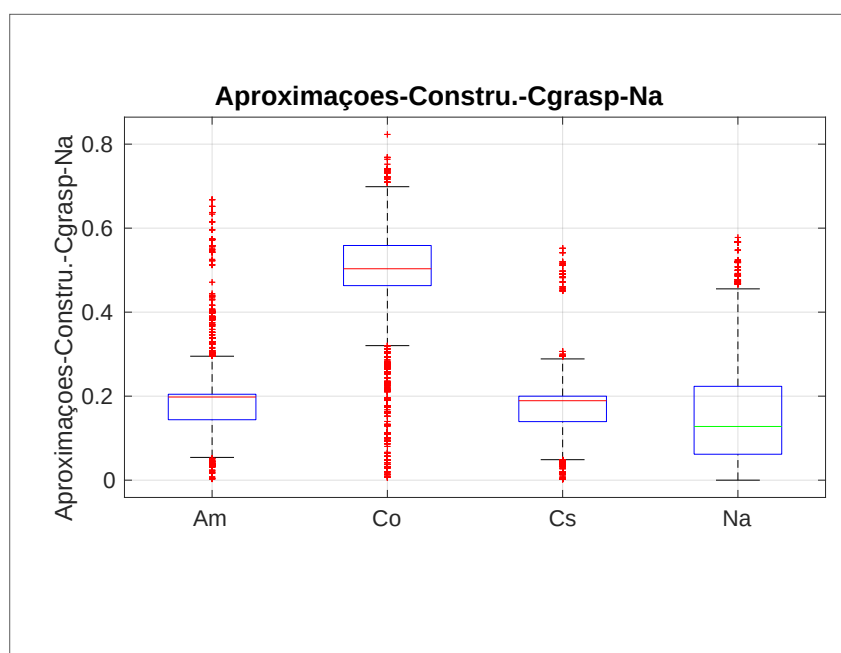
(a) Frações de contagem percentual dos isótopos inseridos no espectro total na omissão do Cs . Pontos Am ○; Co +; Cs *; Na ☆.



(b) Medidas da área abaixo da curva do espectro total. Área calculada ○; área real +; resíduo *.



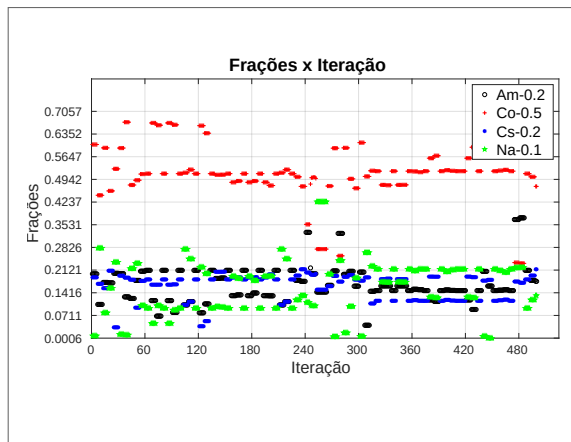
(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP.



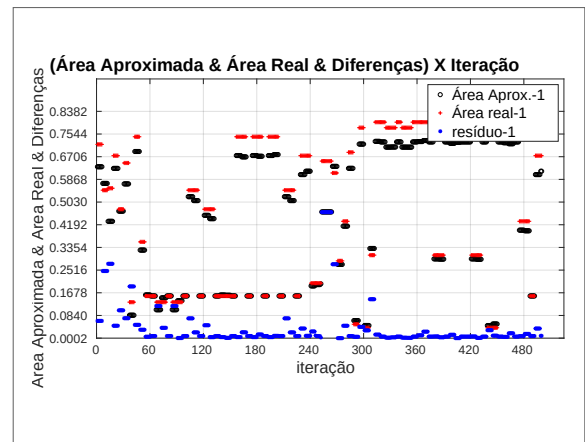
Fonte: Próprio Autor

Figura 150 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 150a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (Am ○; Co +; Cs *); na Figura 150b(○); área real +; resíduo *) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 150c tem-se os valores de estimativa do composto omitido (Na) e a aproximação dos demais compostos.

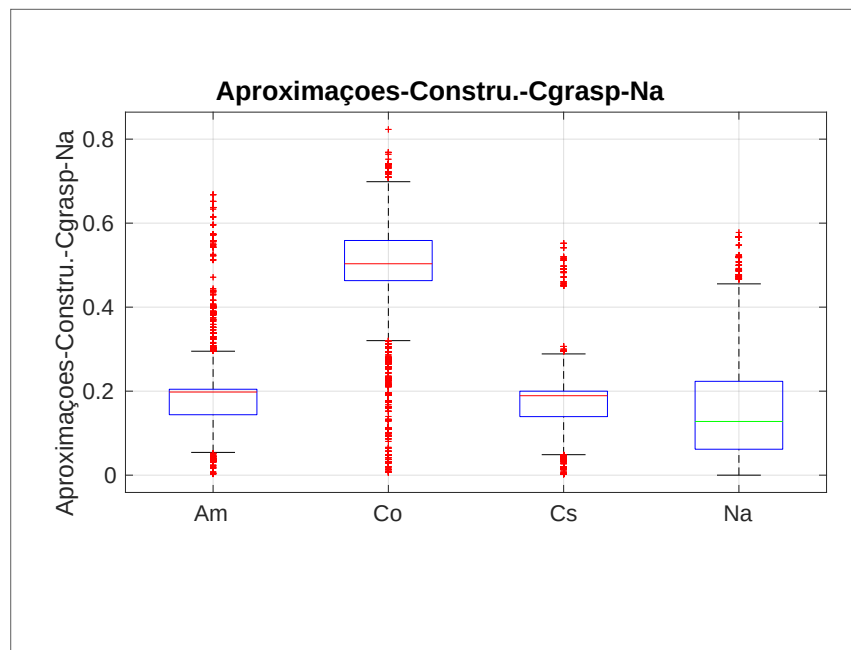
(a) Frações de contagem percentual dos isótopos inseridos no espectro total na omissão do Na . Pontos Am ○; Co +; Cs *; Na ☆.



(b) Medidas da área abaixo da curva do espectro total. Área calculada ○; área real +; resíduo *.



(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP.

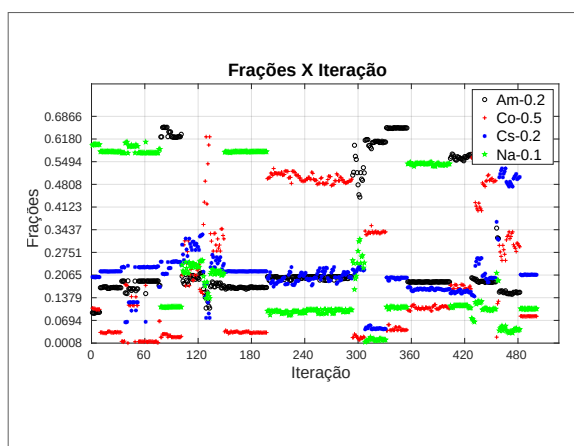


Fonte: Próprio Autor

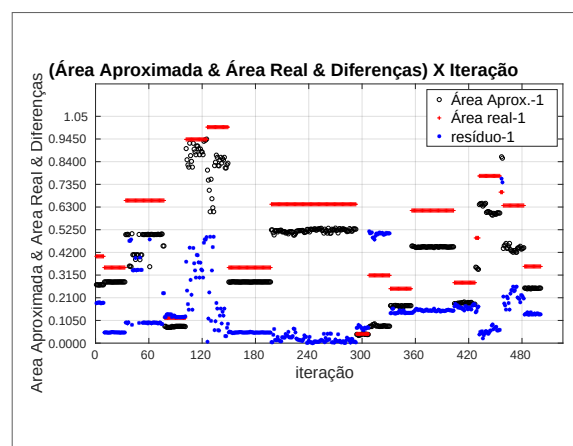
J.1.3 Resultados CGRASP - Etapa de Busca

Figura 151 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 151a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (Am ○); Co +; Cs *); na Figura 151b são demonstrados: área calculada(○); área real +; resíduo *), demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 151c tem-se a distribuição das frações de contagem percentual, omitido Co (mediana em verde), e a aproximação dos demais compostos.

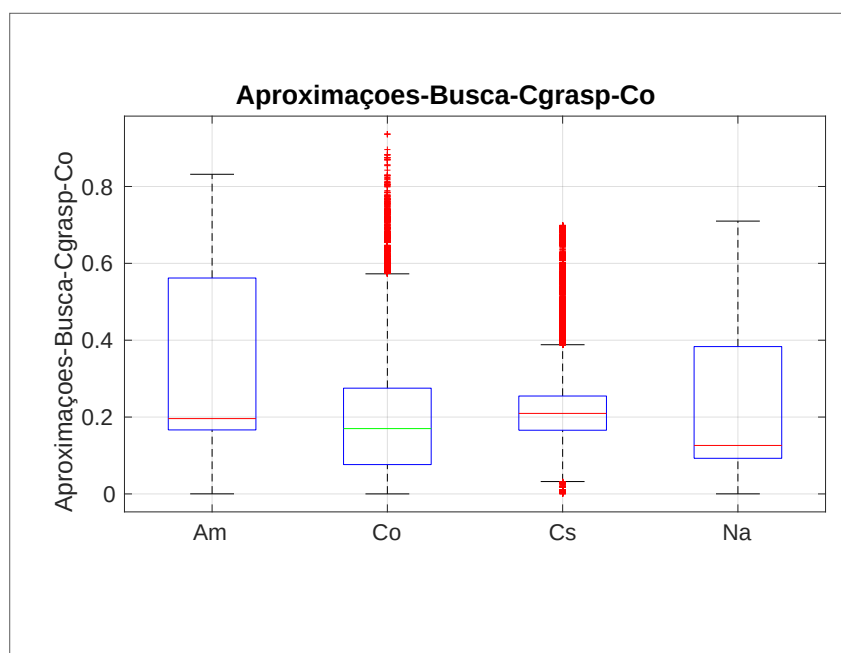
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão do Co (Am ○); Co +; Cs *; Na ☆). Passo de busca local.



(b) Medidas da área abaixo da curva do espectro total. Área calculada (○); área real +; resíduo *). Passo de busca local.



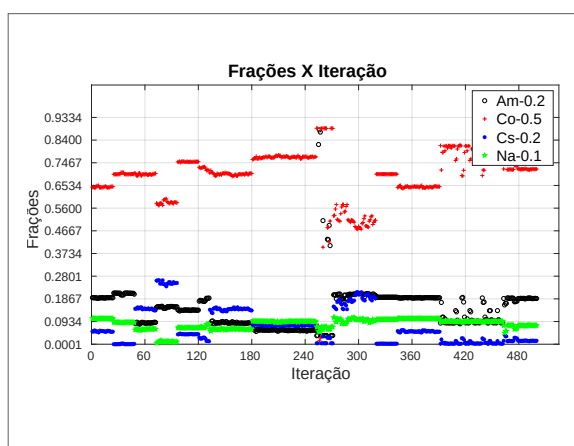
(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP.



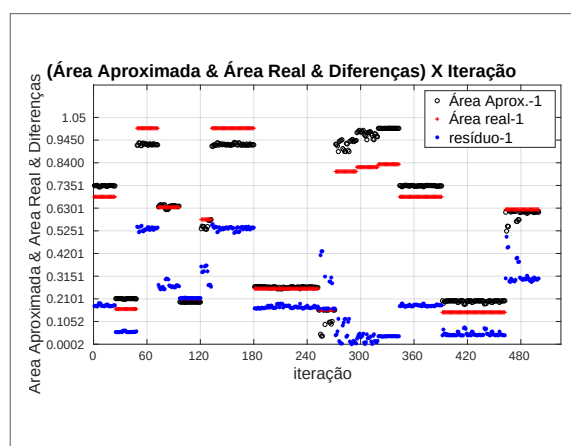
Fonte: Próprio Autor

Figura 152 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 152a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (Am ○); Co +; Cs *); na Figura 152b são demonstrados: área calculada(○); área real +; resíduo *), demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 152c tem-se a distribuição das frações de contagem percentual, omitido Cs (mediana em verde), e a aproximação dos demais compostos.

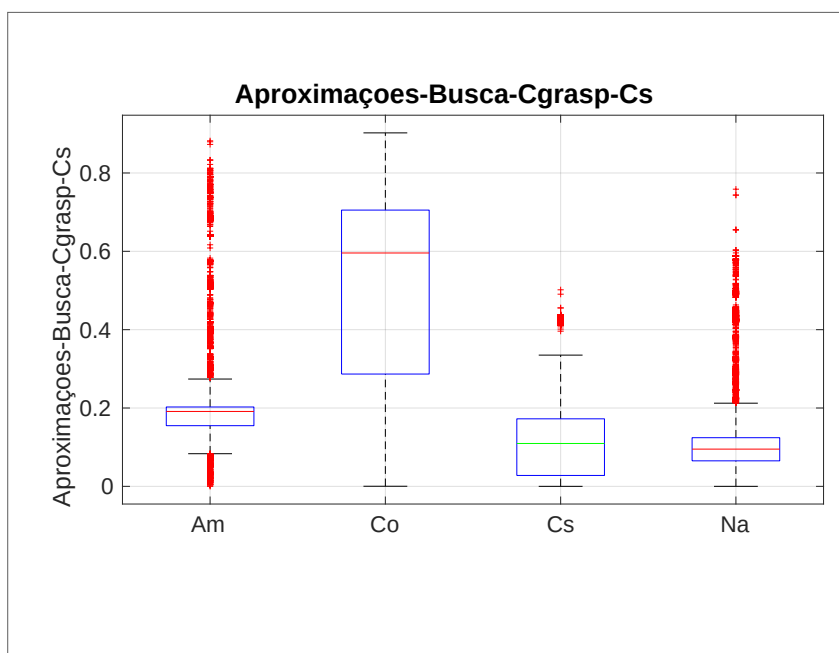
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão do Cs (Am ○); Co +; Cs *); Na ☆). Passo de busca local.



(b) Medidas da área abaixo da curva do espectro total. Área calculada (○); área real +; resíduo *). Passo de busca local.



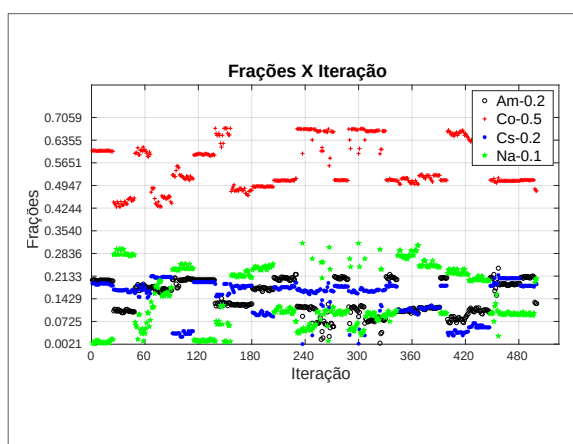
(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP.



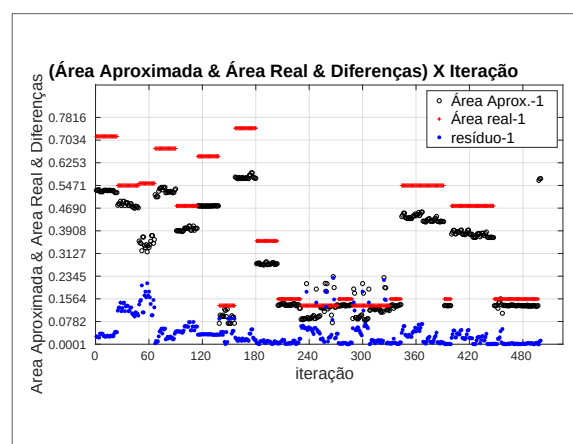
Fonte: Próprio Autor

Figura 153 – Resultados acumulados para a etapa da busca local do CGRASP. Na Figura 153a, tem-se os valores quantitativos dos isótopos que constituem o espectro total (Am ○); Co +; Cs *); na Figura 153b são demonstrados: área calculada(○); área real +; resíduo *), demonstrando a evolução da métrica que determina a qualidade da solução encontrada pelo CGRASP. Por fim, em 153c tem-se a distribuição das frações de contagem percentual, omitido Na (mediana em verde), e a aproximação dos demais compostos.

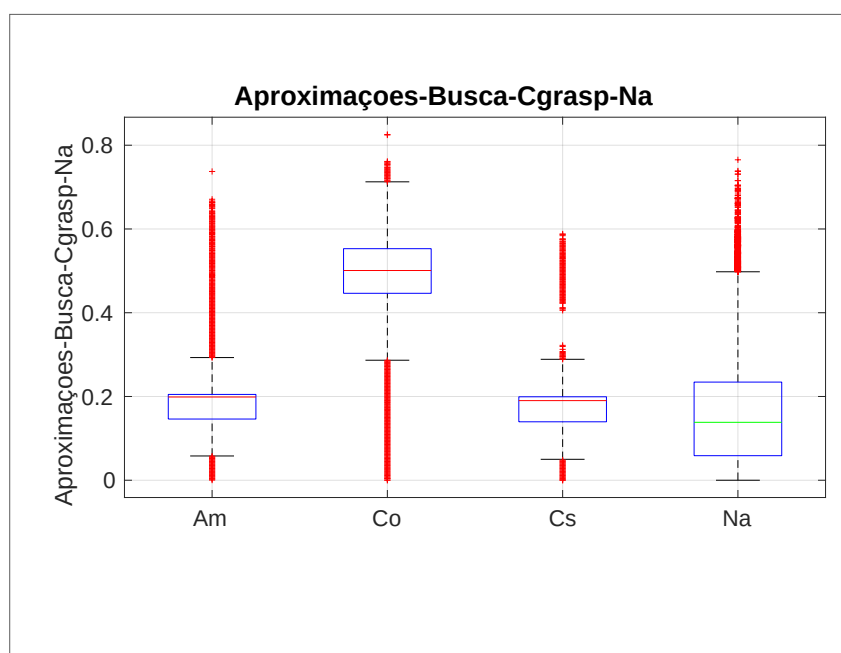
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão do Na (Am ○); Co +; Cs *); Na ☆). Passo de busca local.



(b) Medidas da área abaixo da curva do espectro total. Área calculada ○; área real +; resíduo *). Passo de busca local.



(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP.



Fonte: Próprio Autor

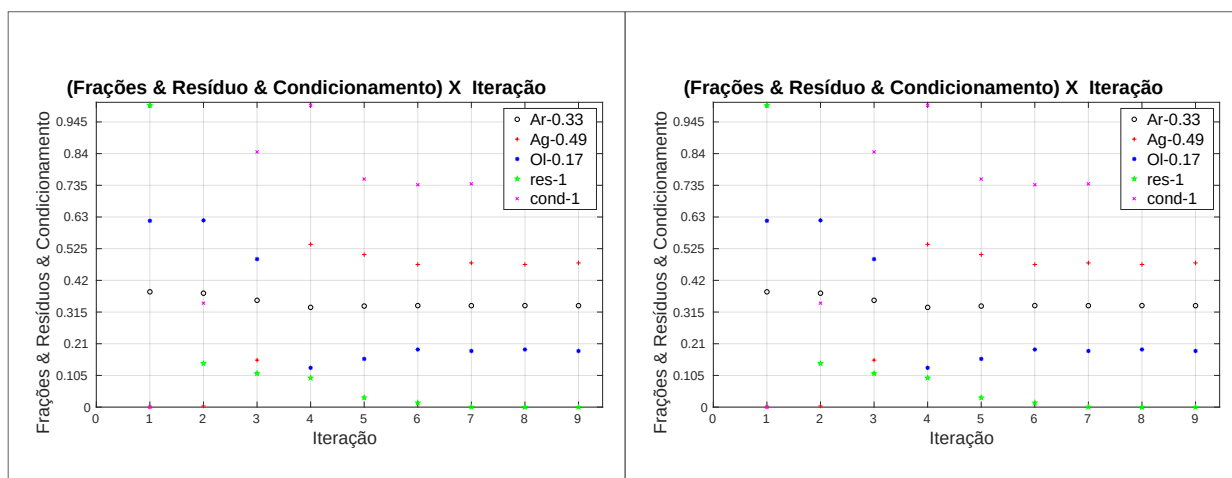
J.2 RESULTADOS CGRASP - BASE DE DADOS 2

J.2.1 Resultados CGRASP - Execução única

Figura 154 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas figuras 154a e 154b valores referentes à omissão da Água. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. ○ indicam as aproximações para o Ar; + as aproximações para a Água; * as aproximações para o Óleo; ☆ os valores para o resíduo ✕ os valores para número de condicionamento.

(a) Percentuais dos elementos e respectivas métricas. Aproximações Água.

(b) Percentuais dos elementos e respectivas métricas. Aproximações Água.

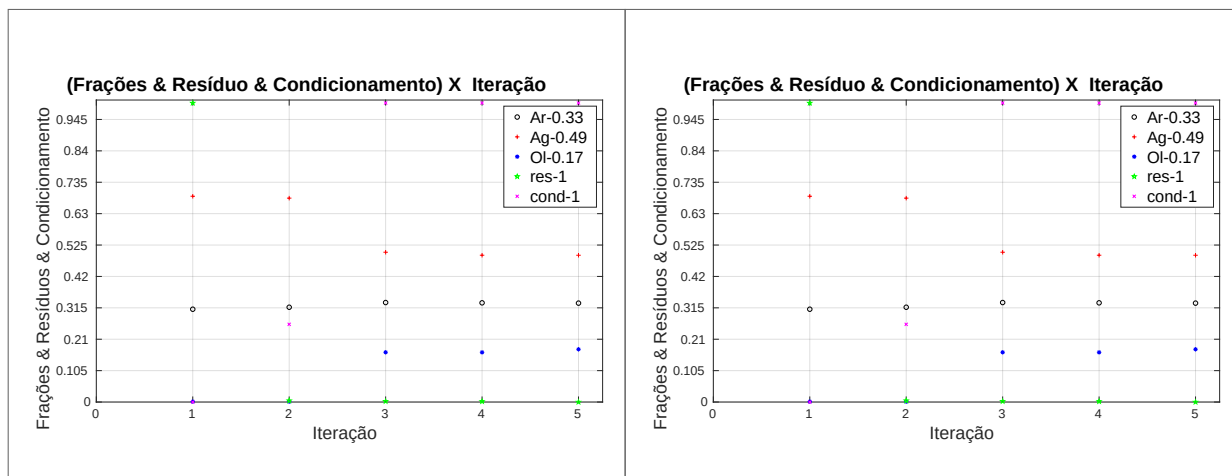


Fonte: Próprio Autor

Figura 155 – Execução do CGRASP para quantificação dos elementos inseridos no espectro total. Nas figuras 155a e 155b valores referentes à omissão do Óleo. Os percentuais aproximam-se dos valores nominais à medida que os resíduos reduzem. ○ indicam as aproximações para o Ar; + as aproximações para a Água; * as aproximações para o Óleo; ☆ os valores para o resíduo ✕ os valores para número de condicionamento.

(a) Percentuais dos elementos e respectivas métricas. Aproximações Óleo.

(b) Percentuais dos elementos e respectivas métricas. Aproximações Óleo.

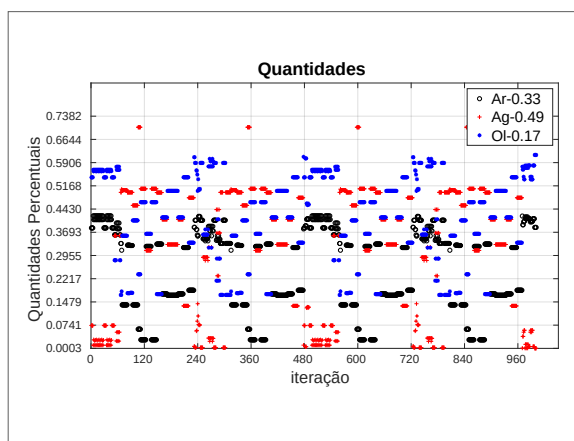


Fonte: Próprio Autor

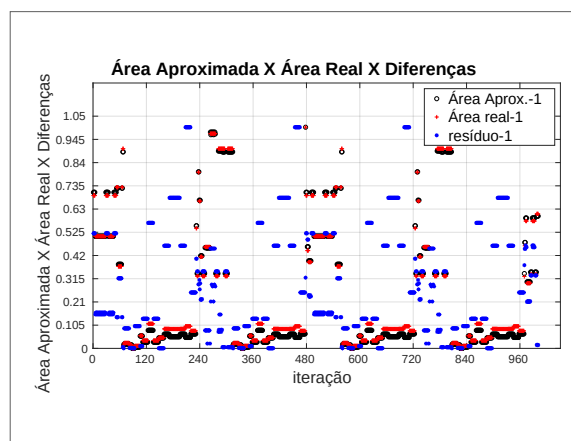
J.2.2 Resultados CGRASP - Etapa Construtiva

Figura 156 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○); (Água +); (Óleo *); na Figura 157b tem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Água) e a aproximação dos demais compostos.

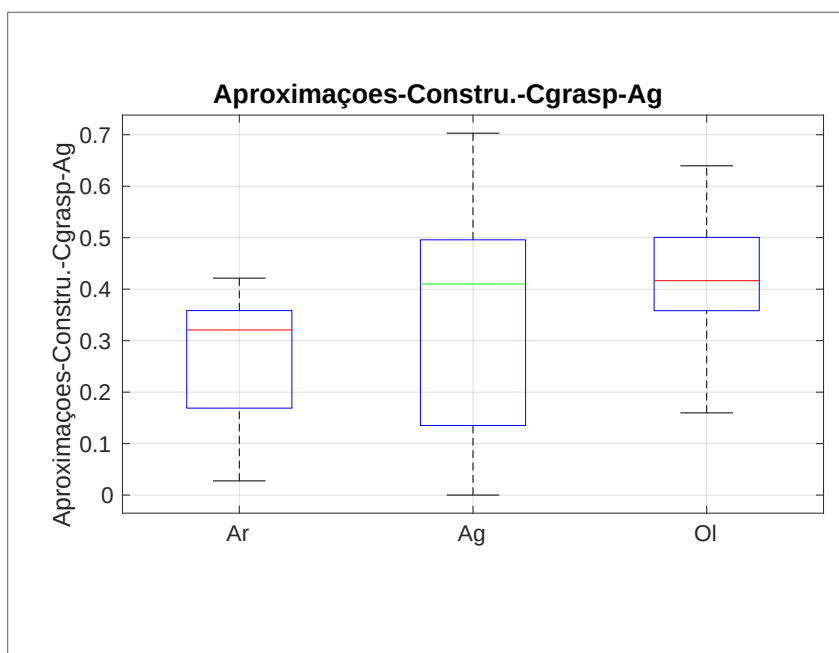
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão da Água (Ar ○); Água +; Óleo *.



(b) Medidas da área abaixo da curva do espectro total. Área calculada (○); área real +; resíduo *.



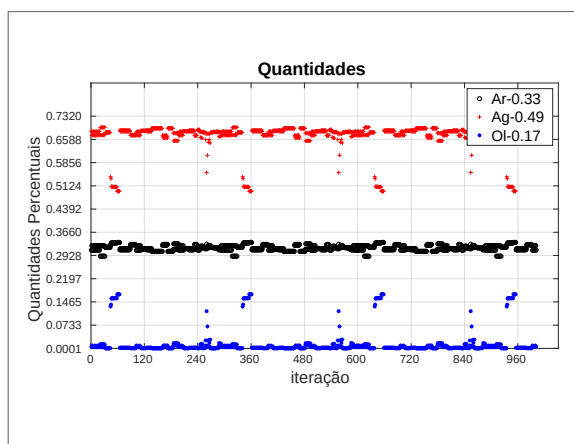
(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP. Omissão Água.



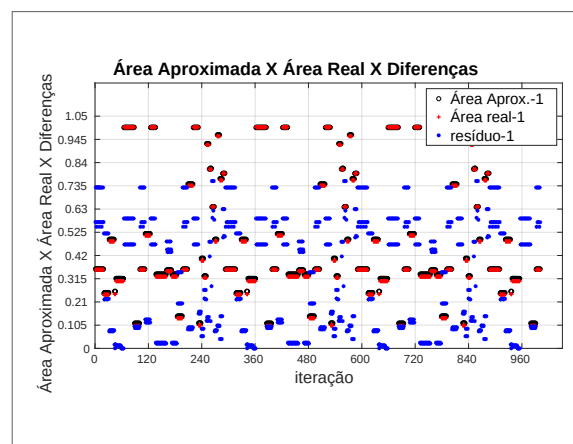
Fonte: Próprio Autor

Figura 157 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○);(Água +);(Óleo *); na Figura 157b tem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) é demonstrada a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos.

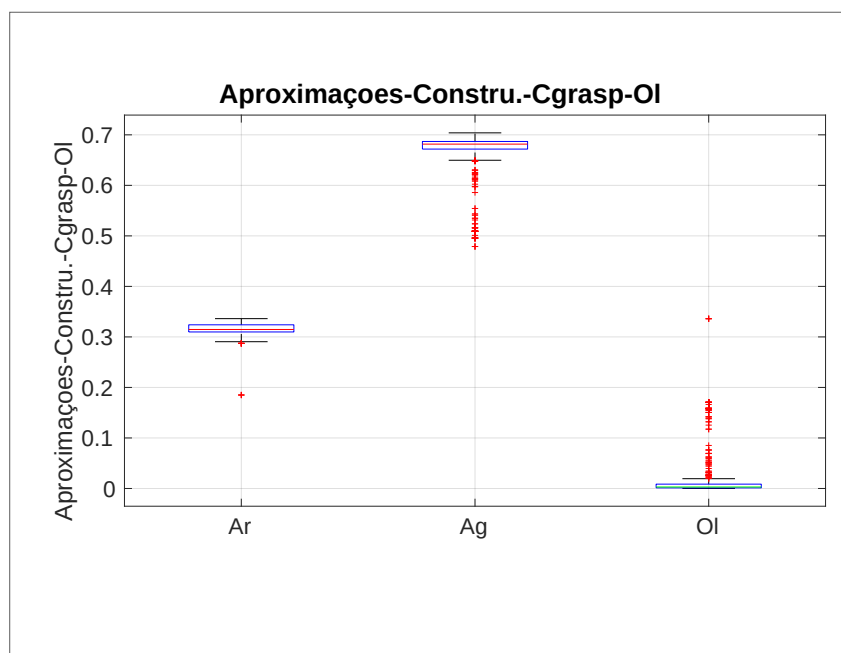
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão do Óleo (Ar ○); Água +; Óleo *.



(b) Medidas da área abaixo da curva do espectro total. Área calculada (○); área real +; resíduo *.



(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP. Omissão Óleo.

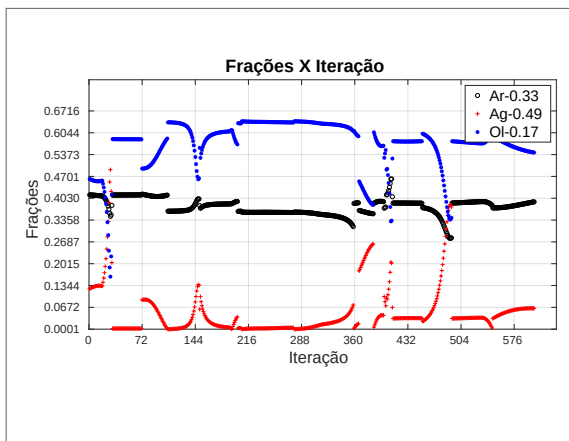


Fonte: Próprio Autor

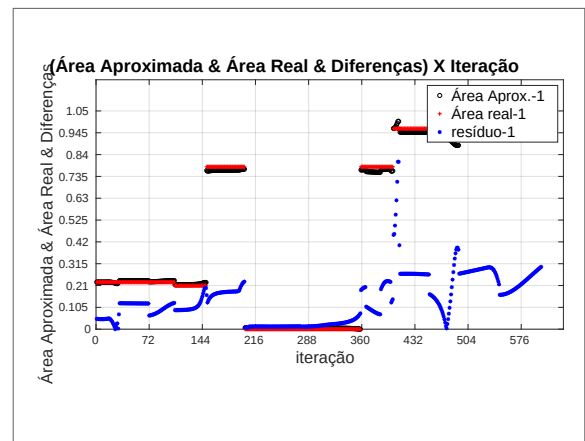
J.2.3 Resultados CGRASP - Etapa de Busca

Figura 158 – Resultados acumulados para a etapa de busca do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○);(Água +);(Óleo *); na Figura 157b tem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) demonstram a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos.

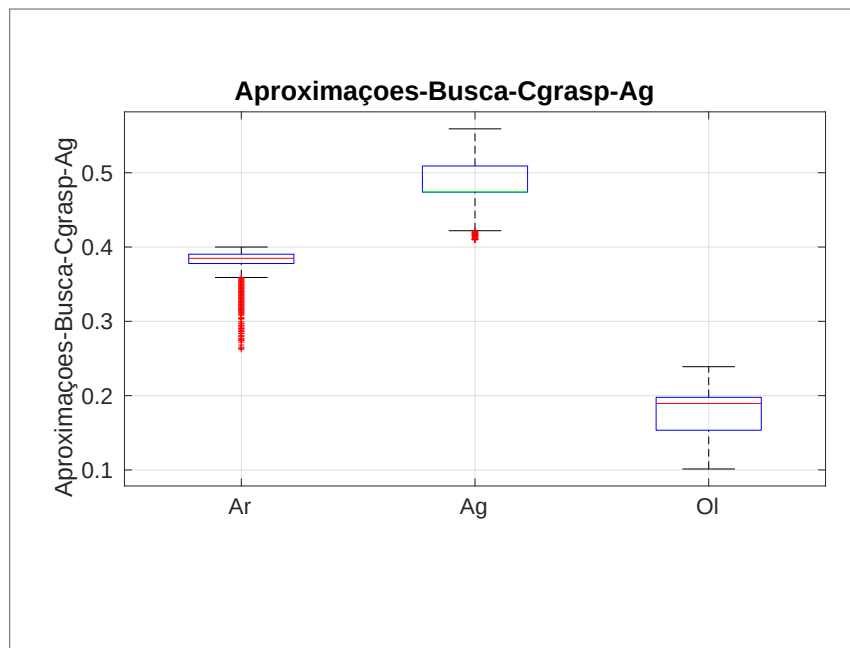
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão da Água (Ar ○); Água +; Óleo *.



(b) Medidas da área abaixo da curva do espectro total. Área calculada (○); área real +; resíduo *.



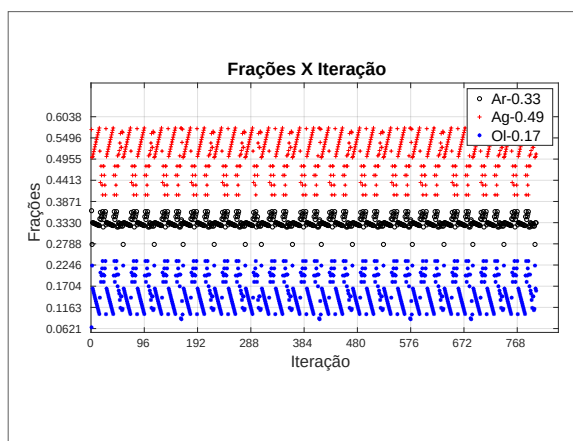
(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP. Omissão Água.



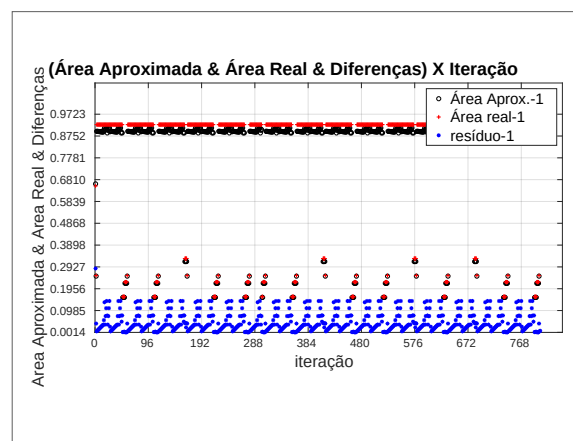
Fonte: Próprio Autor

Figura 159 – Resultados acumulados para a etapa construtiva do CGRASP. Na Figura 61a, tem-se os valores quantitativos dos isótopos que constituem o espectro total ((Ar ○);(Água +);(Óleo *); na Figura 157b tem-se a seguinte descrição: Área estimada (○); área real (+); resíduo (*) demonstram a evolução da métrica que avaliará a qualidade da solução encontrada pelo CGRASP. Por fim, em 157c tem-se os valores de estimativa do composto omitido (Óleo) e a aproximação dos demais compostos.

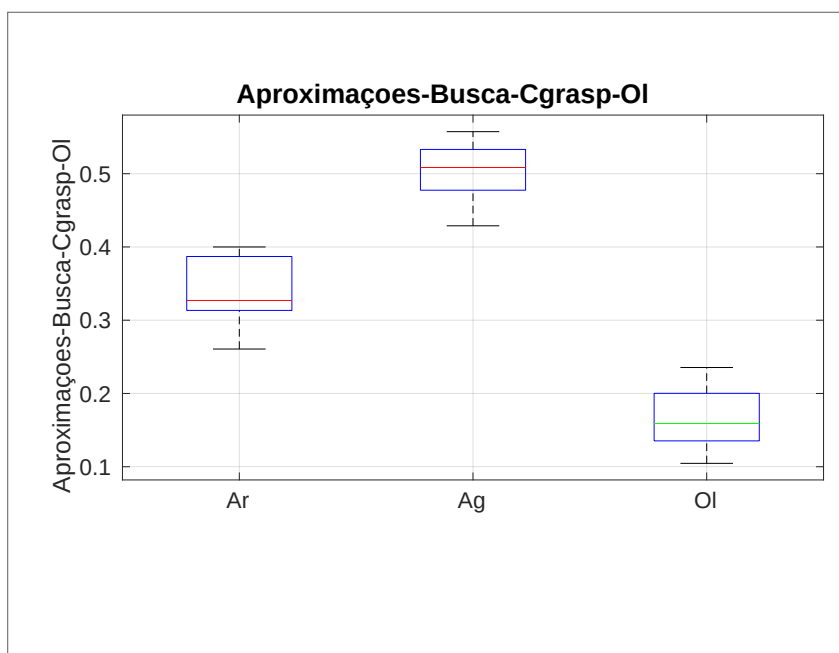
(a) Valores das quantidades dos isótopos inseridos no espectro total na omissão do Óleo (Ar ○); Água +; Óleo *.



(b) Medidas da área abaixo da curva do espectro total. Área calculada (○); área real +; resíduo *.



(c) Box contendo as flutuações nas determinações percentuais dos isótopos contidos no espectro. O box com mediana verde indica o isótopo estimado pelos ajustes do CGRASP. Omissão Óleo.



Fonte: Próprio Autor

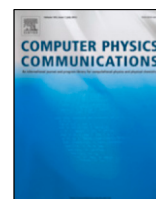
APÊNDICE K – PUBLICAÇÕES

K.1 ARTIGOS ACEITOS E PUBLICADOS



Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpcSpectral analysis and optimization of the condition number problem[☆]

Halisson Alberdan Cavalcanti Cardoso^{a,*}, Silvio de Barros Melo^a,
Ricardo Martins de Abreu Silva^a, Sidartha Azevedo Lobo de Carvalho^{a,b},
Silas Garrido Teixeira de Carvalho Santos^a, Carlos Costa Dantas^c

^a Centro de Informática Campus Recife - Universidade Federal de Pernambuco, Av. Jornalista Aníbel Fernandes - s/n - Cidade Universitária - PE, ZIP Code: 50.740 - 560, Brazil

^b Campus Quixadá - Universidade Federal do Ceará - Av. José de Freitas Queiroz - 5003 - Cedro - Quixadá - Ceará, Zip Code: 63.902-580, Brazil

^c Departamento de Energia Nuclear Campus Recife - Universidade Federal de Pernambuco - Av. Prof. Luiz Freire - 1000 - Cidade Universitária - PE, ZIP Code: 50.740-545, Brazil



ARTICLE INFO

Article history:

Received 26 December 2019

Received in revised form 26 July 2020

Accepted 28 August 2020

Available online 6 September 2020

Keywords:

Prompt Gamma Neutron Analysis

Activation (PGNAA)

Condition number

Monte Carlo Library Least Square (MCLLS)

Greedy Randomized Adaptive Search

Procedure (GRASP)

Least square

Singular Value Decomposition (SVD)

ABSTRACT

Spectral analysis systems using gamma radiation to determine the percentage fraction of several compounds in a sample need several stages. The conditioning number of covariance matrices associated with the least-squares system needs to be very close to one so that the determined values are as close as possible to the actual values. Scientific evidence gives us reason to believe that these matrices make it possible to obtain fractions of compounds in the sample close to the true values. This work focuses on the use of the metaheuristic Greedy Randomized Adaptation Search Procedure (GRASP) to estimate percent counts of constituents of a compound represented by prompt gamma ray spectra, based on the minimization of the condition number of the covariance matrix derived from the underlying linear system. For this purpose, steps of GRASP were modified: Algorithms are used for building a solution and for searching in the vicinity of that built solution. Following the stages of building a solution and searching around the built solution, carried out by GRASP, there are attempts to achieve a better quality solution than determined by local search by executing the linking algorithm between two solutions. The results presented in our work improve the condition numbers of the covariance matrices found in articles that are currently published, observing the characteristics of the application data under study. The proposed algorithm obtained the spectral count matrices as input, obtaining average improvements of 31.30% in the condition number of the covariance matrices in the execution of GRASP in the partitioned data. The objective was reached in 7 of the 9 instances. The improvement of the values in condition number was on average of 84.62% in cases of orthogonalized covariance matrices. The average error of the spectral count percentages for each element was $7.2 \times 10^{-3}\%$. Results lead us to the conclusion that it is possible to obtain covariance matrices that minimize numerical instability in the least squares solution.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Prompt Gamma-ray Activation Analysis (PGNAA) is an efficient non-destructive technique widely used for elemental analysis of samples, such as: (1) elemental analysis of coal and minerals [1]; (2) oil well exploration [2,3]; (3) natural measurement of radioactivity [4–6] and (4) measurements in the marine environment

for radiologic purposes [7]. PGNAA is considered a suitable measurement technique for determining and monitoring impurity concentration in crude oil samples. Gamma rays, when interacting with neutrons, are used to identify and quantify chemical elements. Specific chemical properties of oxygen, carbon, and nitrogen are used in the EURITRACK system [8,9] for the detection and identification of explosives and illicit drugs in the context of the PGNAA study.

Quantitative Analyses in PGNAA have been addressed by such methods as Monte Carlo Library Least Squares (MCLLS) [10] and the Single Peak Analyses [11]. Additionally, linear modeling of this intrinsically non-linear problem has been proposed, such as the LLS (Library Least Squares). In [12], the authors use Monte Carlo simulations to overcome this approximation difficulty, producing the MCLLS iterative technique. The MCLLS approach has an

[☆] The review of this paper was arranged by Prof. Hazel Andrew.

* Corresponding author.

E-mail addresses: hacc@cin.ufpe.br (H.A.C. Cardoso), sbm@cin.ufpe.br (S. de Barros Melo), rmas@cin.ufpe.br (R.M. de Abreu Silva), sidartha@ufc.br (S.A.L. de Carvalho), sgtcs@cin.ufpe.br (S.G.T. de Carvalho Santos), ccd@ufpe.br (C.C. Dantas).

advantage over the single peak analysis technique for utilizing the full information contained in the spectrum (Whole Spectrum Analysis (WSA)) [12]. MCLS, assuming that the constituents define the total spectrum by a linear combination, can be applied to the quantitative study of mass samples, including multiphase hydrocarbon flow [2,13]. Spectroscopy analysis is classified as an overdetermined problem [14], since the number of energy frequencies, or channels, in the MCLS approximation is higher than the number of members that make up the total spectrum.

Classic multiple linear regression (MLR) methods are used in [15] to solve the question of overdetermination from the WSA approach. Despite obtaining appropriate results, improper results are reported, such as the determination of negative values for percentage fractions of unknown spectrum constituents. The difficulty in correctly estimating percentage fractions occurs because the inverse problem presents a numerically ill condition [16] or the correct libraries are not provided [17].

Several methods are proposed to overcome these difficulties: (1) minimizing the condition number of the covariance matrix [13,16]; (2) application of the nonnegative least squares method to gamma ray spectra [5,18] and (3) application of Full Spectral Analysis (FSA) to multiple energy windows [19], controlling the adaptation process by a comparison module [20].

The previously described methods can also be highlighted: (1) fireworks algorithm [21]; (2) Particle Swarm Optimization (PSO) [17]; (3) genetic intelligence techniques based on genetic algorithms [22]; and (4) neural networks, when applied to gamma ray spectrum analysis [23]. In [20], a fuzzy system was successfully applied to identify the pertinent libraries using one-spectrum photopeaks for further genetic algorithm adjustment using the identified isotopes. The application of fuzzy logic may not yield the appropriate results when the number of libraries considered increases considerably [24].

The linear combination of detector counts that contribute to the library spectra (R) forms the basic assumption of the construction of an total spectrum (B) [17], described in Eq. (1).

$$B = R\alpha + \epsilon \quad (1)$$

In Eq. (1), $B = (B_1, B_2, \dots, B_n)^T$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$ and R stands for the $n \times m$ response matrix. Since the number of channels is usually much larger than the number of libraries ($n \gg m$), the system becomes overdetermined. ϵ is the statistical uncertainty vector related to the counts recorded. α is the vector that contains the percentage counts of the elements present in the sample [25,26] and [27].

Inverting matrix R gives the appropriate coefficient values. In many cases, when the matrix is invertible, inversion may insert numerical instabilities that imply solutions sensitive to small perturbations in the system of linear equations. Instabilities may be due to an ill-conditioned matrix, whose inversion involves algebraic operations that are prone to introduce non-negligible uncertainties in the numerical results [13,28]. MCLS process instabilities will occur mainly when [1]:

1. Two or more libraries are linearly correlated, implying the linear dependence of the matrix columns.
2. A library spectrum has an irrelevant contribution to the total gamma-ray spectrum. The counting rates of one library are very small orders of magnitude smaller than the relative values of another library.

In this paper, an approach based on spectral partitioning and optimization in the calculation of covariance matrices is presented. This approach takes advantage of the fact that the quantity of compounds is considerably smaller than the available energy channels, allowing a selection of appropriate sets of these energy channels aiming at the linear independence of the input

data. The focus was on the use of the LLS technique for PGNAA analyzers. The major motivation for this work was a publication by [29] that demonstrated that the covariance matrix can be singular or ill-conditioned when inverted in the least squares. In many cases, the ill-conditioned of the covariance matrix will pose an important challenge in MCLS implementation. So, a new method can be used to overcome cases of ill-conditioning in the MCLS approach applied to the multiphase flow problem [29]. The algorithm proposed in this work aims to use GRASP to obtain, from a subset of the input data, covariance matrix that have the least possible condition number. Linear independence between the covariance matrix columns will be tested after the partitioning of the input data to reduce the condition number in the gamma-ray analysis problem. The results of the current approach are also compared with those obtained by [29]. In addition, there is evidence that our approach can suppress ill-conditioned, leading to improved sensitivities.

The present work is a complement of the work presented in [30]. As proposed improvements, we highlight the implementation of Path-Relinking (Algorithm 4) as a new stage in the execution of GRASP (Algorithm 1). For this, it takes into account the space of solutions comprised between two base solutions: the one with the lowest condition number between the two is named target solution, while the other is named start solution. The results showed potential for the application of Path-Relinking in the domain of the problem studied in reducing the condition number of matrix covariance.

It is essential to highlight that one of the main contributions of the present work is to use the GRASP optimization algorithm allied to input data division based on linear independence criterion. This proposal will contribute to the construction of two-dimensional covariance matrices with a lower condition number than those presented in [29].

The results for the other GRASP steps are also described in Section 5, demonstrating an efficient overall calculation to minimizing the condition number of covariance matrices.

The definitions presented in Section 2 provide the basis for understanding the problem, such as: Section 2.1 introduces definitions on condition number and states an example that demonstrates numerical instability associated with ill-conditioning problems and Section 2.2 will bring essential elements for definition of the heuristic algorithm used in this work. In Section 3, there is the description of the proposed algorithm to analyze the input data and overcome the ill-conditioning problems. Section 4 provides details on the hardware configuration used for the computational experiments carried out in this work. Section 5 describes the results achieved by applying the proposal and comparing it with the reference work, finishing up with the conclusions and future work in Section 6.

2. Background

This section will cover the related works and the necessary concepts that will give support to this work, including: the concepts of condition number, GRASP algorithm and some definitions for PGNAA and MCLS.

2.1. Condition number

The condition number of a nonsingular matrix A is a quantitative indicator of the sensitivity to perturbation of a linear system involving A [31]. It is possible to interpret the condition number of a matrix A as its proximity to a singularity.

Consider matrix A to be non-singular such that the inverse matrix A^{-1} exists and is unique. Consider the linear system:

$$Ax = b \quad (2)$$

Definition 1. The Condition Number for linear systems can be defined by:

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}^{-1}\| \|\mathbf{A}\| \quad (3)$$

The identity matrix is a rank 1 norm, so $\mathbf{I} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ and $\|\mathbf{A}^{-1}\mathbf{A}\| = 1 \leq \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$. One concludes that $\text{cond}(\mathbf{A}) \geq 1$. A well-conditioned matrix has a condition number close to the unity.

The matrix's SVD (Singular Value Decomposition) makes possible to precisely compute the condition number. The condition number corresponds to the ratio of the largest to the smallest Singular Values.

Thus, a matrix ill-conditioning may occur when its lowest non-zero singular value is too close to zero [32].

2.2. GRASP

The problem presented in the current work is computationally intractable because of its size. Suppose n is the number of distinct values a particular channel may assume. Usually, a number of channels m much smaller than 2048 is enough to completely represent all information present in a given spectrum. Therefore, the size of the solutions space is $C_m^{2048} \times n^m$. Heuristic methods are used to find good solutions, but do not necessarily find the optimal solution. The effectiveness of these methods depends on their ability to avoid entrapment in local minima and to explore the basic structure of the problem. These notions provide tools to develop heuristic search techniques that have proven to improve the ability to get reasonable good solutions to severe problems. One of the most promising methods includes GRASP [29]. GRASP is a metaheuristic to find approximate solutions to combinatorial optimization problems formulated as follows [33]:

$$\begin{cases} \text{minimize } f(x) \\ \text{subject to } x \in \chi \end{cases} \quad (4)$$

where $f(\cdot)$ is an objective function to be minimized, and χ is a discrete set of feasible solutions.

In [34], there is a description of the GRASP algorithm as an iterative methodology, that provides a solution to the problem under study through iterative processes.

GRASP has experienced continuous development [35,36] and has been applied in a broad set of areas [37,38].

In this paper, we present a GRASP with path-relinking heuristics for the problem of determination of covariance matrices with the lowest condition number. The discrete set of feasible solutions is composed of radiation countings in function of energy levels (channels) supplied as input to the problem. From the selection of these channels, the objective function becomes the condition number of the covariance matrix produced from these channels least squares (LS) system.

2.3. PGNA

Extracting information from gamma-ray spectra is a very frequently found aspect in the research and development activities with a focus on inverse radiation analyzers that have found applications within a variety of fields such as medicine [39–41], industry [41,42], as well as in homeland security [43].

PGNA analysis consists of determining the chemical and physical properties of a sample based on prompt-gamma spectra recorded appropriately by a spectrometer.

The most parsimonious approach to this problem is to assume linearity between the total spectrum and the constituent spectra. Noise may be present in the input spectra, which suggests the use of a Least Squares Method. However, its results usually are

seriously affected by highly nonlinear data, a problem that can be mitigated by the iterative use of Monte Carlo (MC) [10].

MCLS has a significant advantage over other techniques as it makes use of the full gamma-ray spectrum, providing more accurate analysis [12]. The data used in this work is composed of five libraries of spectral counts (water, salt, oil, gas, and background), organized according to 2048 energy channels [29].

2.4. MCLS

The most effective approach in the quantitative analysis of PGNA uses the MCLS method proposed by [10]. The amount of each pertinent library is then calculated in a linear least-squares formulation whereby a set of linear equations needs to be solved simultaneously. More detailed description of the MCLS approach is given in [13]. The application of MCLS for quantitative analyses of gamma-ray spectra has shown higher sensitivity than conventional single peak analyses. In the study of single peaks, the qualitative and quantitative analyses of the spectrum make use only of photopeaks [12]. In addition, to being more sensitive and robust than single peak analyses, the MCLS approach has the following advantages:

- Automatic handling is provided to unresolved peaks;
- All the information contained in the spectrum is used, both the photopeaks and its *Continua Compton*.
- Residuals can reveal if libraries are missing in the analysis.

The main challenge in implementing MCLS is that the coefficient matrix of the equation system is ill-conditioned [3], exhibiting a high degree of sensitivity to small perturbations in the measurement data. Ill-conditioning implies numerically unstable solutions in the least-squares approximation process. The ill-conditioning in the MCLS approach may come about as a result of the following:

1. A library spectrum is similar in shape to any of the other libraries or to linear combinations of these libraries, i.e., a strong degree of linear correlation between any of two or more libraries.
2. Library spectrum is presenting a near-negligible contribution to the prompt gamma-ray spectrum.

The point 1 will be related to the low energy resolution of scintillation detectors, frequently applied in field applications [13]. Point 2 is related to the fact that it finds elements that have low participation in the sample of interest or find a compound with low neutron capture, such as oxygen in an aqueous sample [1].

Two approaches were proposed for the case of ill-conditioning covariance matrices in the MCLS approach with emphasis on the analysis of multiphase oil, gas, salt, water and samples. The first approach is iterative, considering: (1) separate an arbitrary library; (2) the rest of the libraries are added together, forming a single library; (3) The covariance matrix is constructed from the separate spectral library and the matrix composed of the sum of the other spectral counts, for later calculation of the conditioning number; (4) repeat (1) to (3) for each library; (5) solve the system with the matrix that presents the lowest condition number and record the percentage of counts from each element; (6) remove this separate compound multiplied by its fraction from the sample; (7) return to step (1) with the remaining compounds (if there are still more than one compound) [13].

The second approach is based on the introduction of a densitometer configuration using a Cs-137 radioisotope, in addition to the neutron source [44]. The results showed that both methods successfully suppressed the ill-conditioned of covariance matrix in the MCLS approach. On the other hand, the same sensitivity analyses have also revealed that there is room for improvement

with respect to further reduction in the ill-conditioning of the MCLS approach, ensuring improved sensitivity of the method for inverse radiation analysis.

3. Methodology and proposed approach

The determination of the percentage of counts from each element of a sample subjected to gamma radiation is not an exact process. The solution is approximate and depends on the condition number of the covariance matrix of the system. Lower condition numbers impart more dependable approximations.

The mathematical constraints of the present problem can be summarized by:

$$\sum x_{ij} = 1, \forall i \in N, \forall j \in K \quad (5)$$

$$B_i = \sum_{j=1}^I \alpha_j R_{ij} + \epsilon_i \quad (6)$$

$$0 \leq \alpha_j \leq 1 \quad (7)$$

where the constraint (5) determines the insertion of the i th channel in the j th partition, for all i in the input data set N and for all j th in the partition set K . The variable $x_{ij} = 1$ if i th channel is inserted in j th partition. Restriction (6) states that the unknown spectral count on the i th channel is a linear combination of the spectral counts of the j th compound on the i th channel and its respective j th percentage abundance α_j . Constraint in Eq. (7) determines that the percentage fraction of the j th compound is between 0 and 1 subject to the following objective function:

$$\text{minimize } \frac{\sigma_1(\mathbf{R})}{\sigma_n(\mathbf{R})} \quad (8)$$

where $\sigma_1(\mathbf{R})$ in Eq. (8) is the largest Singular Value of $\mathbf{R}$, and $\sigma_n(\mathbf{R})$ is the smallest Singular Value of $\mathbf{R}$.

The procedures performed in the present work to partition data propose to overcome the currently main challenges in the PGNA study: (1) overdetermination of the linear equation systems that are present in the work of [13]; (2) strong degree of linear correlation between one or more libraries; (3) insignificant contribution of some library spectra to the full gamma-ray spectrum.

In the next sections the GRASP algorithm and a method of improving matrix computations are presented. Some actions are performed before the most essential steps of Algorithm 1: the input data are separated into partitions based on null spectral counts of the compounds (line 6 of Algorithm 1), calculating the proportion of channels inserted in each of the partitions (line 7). This procedure is carried out by seeking the linear independence of the data. Finally, these proportions are used to determine the number of channels in each of the partitions of a new solution with an arbitrary number of channels (line 8).

In the routine of constructing a solution, an arbitrary number of channels is chosen to build the covariance matrix (Section 3.1.1), aiming at a decrease in the dimensionality of the input data. The investigation for better quality solutions than that brought by the construction algorithm will be carried out by the search algorithm. The neighborhoods of the solution determined by the construction algorithm will be the target of the search algorithm. A new solution is defined as neighbor of the current solution that is different from the current by a single channel (Section 3.1.2). The result of the search will be linked to a solution chosen arbitrarily from the GRASP list (Section 3.1.3). Each of the outputs brought by the Algorithms 2, 3, and 4 contains the covariance matrices and the orthogonalized covariance matrix (Section 3.2).

3.1. The GRASP algorithm

The algorithm used to achieve the purposes of the present work, briefly described in Section 2.2, is basically composed of: (1) step of building a solution (Algorithm 2); (2) searches in regions surrounding the built solution (Algorithm 3) and (3) linking between the newly built solution and an arbitrary solution that is present in the GRASP list (Algorithm 4).

Algorithm 1 GRASP

```

1: input { $M, T, Max, Num, Obj$ };
2: output { $resulGrasp$ };
3: initialize the algorithm list as empty;
4:  $minCond \leftarrow \text{inf}$ ;
5:  $maxCond \leftarrow 0$ ;
6:  $estInt \leftarrow \text{intervals}(M)$ ;
7:  $estPro \leftarrow \text{proportion}(M)$ ;
8:  $numCha \leftarrow \text{numChalnt}(M)$ ;
9: while  $Obj < minCond$  or current iteration  $< Max$  do
10:   while the solution is not distinct from the others do
11:      $d \leftarrow \text{constr.}(M, estInt, Num, numCha, estPro)$ ;
12:   end while
13:   if current iteration  $< T$  then
14:     insert  $d$  in the list;
15:      $x \leftarrow \text{condition number of } d$ ;
16:     if  $x < minCond$  then
17:        $resulGrasp \leftarrow d$ ;
18:        $minCond \leftarrow x$ ;
19:     end if
20:     if  $x > maxCond$  then
21:        $maxCond \leftarrow x$ ;
22:     end if
23:   else
24:      $k \leftarrow \text{Loc.Sear}(estConst, T, iter, estInt, M, numCha)$ ;
25:      $u \leftarrow \text{randomly select an element of the list}$ ;
26:      $e \leftarrow \text{linking}(s_1, s_2, numChaint, M, total)$ ;
27:      $l \leftarrow \text{Loc.Sear}(e, T, iter, estInt, M, numCha)$ ;
28:      $g \leftarrow \text{condition number of } l$ ;
29:     if  $g < maxCond$  then
30:       replacing the list element most similar;
31:       if  $g < minCond$  then
32:          $resulGrasp \leftarrow l$ ;
33:       end if
34:     end if
35:   end if
36: end while

```

In Algorithm 1, the variables T , Max , Num and Obj indicated in line 1 respectively represent the size of the solution list; the maximum number of iterations of Algorithm 1; the number of channels to be inserted in a new solution and the target value for the condition number of the covariance matrix. The result is a data structure containing the selected channels and the covariance matrix associated with the lowest condition number obtained in the execution of Algorithm 1.

The variables that monitor the highest and lowest number of condition number values are initialized in Algorithm 1 lines 4 and 5, respectively. The functions *intervals*, *proportion* and *numChalnt* are called in lines 6, 7 and 8, respectively, returning the variables $estInt$, $estPro$ and $numCha$. These lines make use of the problem structure in order to find data that are linearly independent. The quest for linear independence in the data is a modification of the basic steps of GRASP [33–35,45] in order to achieve the desired results. The function *intervals* (line 6) partitions the input data into intervals. A partition starts at the channel where the spectral count is zero for a compound, ending at the channel position

where a null spectral count is found. The function *proportion* (line 7) determines the fraction of channels inserted in each of the ranges. The execution of the function *numChalnt* (line 8) calculates how many channels are used in each of the intervals for a new solution with an arbitrary number of *num* channels, based on the values generated in the call of the *estPro*.

A solution is created from the input data in Algorithm 1 from lines 9 to 12. The construction of a solution in the context of the problem is made by a matrix that indexes the channels to be inserted in each one of the partitions. In this matrix, the rows are the channels to be added, while the columns represent the partitions. Each randomly constructed solution is assigned to a list (line 14). The pointers that check the output of the algorithm and the lowest and highest values of condition numbers obtained that far are updated in lines 17, 18 and 21. Whether this solution bears the highest or the lowest condition number during the current run is checked there.

The neighborhood search of the built solution is performed in line Algorithm 1, line 24, following a linkage process (explained later) between the search result and some randomly chosen solution from the GRASP list (line 26), finishing a new search based on the result of linking (line 27). The solution defined by the second local search replaces the solution on the GRASP list, which has a greater condition number and which is more similar to the result of the second search. At lines 29 to 36, the variable *g* will monitor the output of the algorithm when the production of the second fetch has a lower condition number up to the current iteration.

The search for better solutions, performed in Algorithm 1 lines 24 and 27, will occur from the definition of neighbors of the current solution. In the present work, neighbors are calculated from the substitution of a channel in the matrices of channels used to construct the covariance matrices.

The linkage process will verify the differences between the two given solutions: the target solution (better condition number) and the initial solution (worst condition number). The verification of the differences will occur one by one, choosing a new initial solution whenever the modified initial solution presents better condition number.

3.1.1. The algorithm constructive

In order to satisfy the constructive step of the Algorithm 1, line 11, the solutions are built by the Algorithm 2, starting the GRASP iterations. The build algorithm will be based on a new solution with fewer channels than the one present in the GRASP Algorithm global input data. This new solution will have the same number of partitions as the original data, and the channels to be inserted into the partition will be chosen randomly.

The input parameters are: a table containing the spectral counts of the compounds as function of the channels (*M*); a data structure containing partition ranges in the global data (*estInt*); the number *Num* of channels inserted into the construction of an arbitrary solution; number *numChalnt* of channels in each partition of a new solution containing a random amount of channels; the proportion *estPro* of channels in the original data partitions.

The data structure of the channels starts in Algorithm 2, line 3. The elements of the *numChalnt* structure are iterated beginning in line 4. The indices that determine the start and end of each partition with an arbitrary number *Num* of channels are retrieved by using the *numChalnt* variable in line 5. From line 7 to line 14, channels are chosen from the original data partitions to fill the partitions in a new arbitrary solution. From line 16 to line 19 the following variables are defined: covariance matrix *N* of the new solution constructed with the number of channels *Num* (line 16); the orthogonalized covariance matrix computed from the *N* matrix (line 17); determination of the condition number of the matrices *N* (line 18) and *O* (line 19). Then, variables *N*, *O*, *C* and *P* are assigned to the data structure *estConst* (line 20).

Channel matrices will be defined to index the construction of covariance matrices. These channel arrays will have dimension $m \times n$, where *m* is a number bigger than the number of constituents, and *n* is the number of components in the full spectrum. The construction of a new channel matrix will be compared to the channel matrices associated with the solutions introduced in the GRASP Algorithm list. Diverging channels from two compared solutions should not represent less than 90% of all corresponding channels.

Algorithm 2 constructive

```

1: input {M, estInt, Num, numChalnt, estPro};
2: output estConst;
3: channels  $\leftarrow$  size matrix M;
4: for each i in numChalnt do
5:   [start, stop]  $\leftarrow$  numChalnt{i};
6:   k  $\leftarrow$  start;
7:   while k < stop do
8:     j  $\leftarrow$  estInt{i};
9:     l  $\leftarrow$  estInt {i+1};
10:    value  $\leftarrow$  random value between j and l;
11:    if value  $\notin$  channels then
12:      insert value in channels;
13:    end if
14:  end while
15: end for
16: N  $\leftarrow$  matrix of covariance calculated from M indexed by channels;
17: O  $\leftarrow$  orthogonalization of N;
18: C  $\leftarrow$  condition number of N;
19: P  $\leftarrow$  condition number of O;
20: estConst  $\leftarrow$  [N, O, C, P];

```

An arbitrary solution is inserted into the list after the execution of the constructive step. This insertion is immediate if the list is not filled and the solution created is sufficiently different from the solutions present in the list. A local search in the vicinity of the built solution is initiated if the condition of full completion of the list is verified. The locating algorithm is described in Algorithm 3.

3.1.2. The algorithm local search

After the construction phase, the search phase begins around the constructed solution. The algorithm associated with this phase is described in Algorithm 3. The search algorithm will define solutions that are adjacent to the solution determined in the construction stage. The following steps define a neighboring solution to the current solution: (1) A channel that is not present in the current solution is chosen from the universe of input channels of the algorithm; (2) The channel inserted in step (1) is added in the appropriate data partition; (3) The newly inserted channel will replace a randomly chosen channel that belongs to the same partition as that new channel. Indexes are used to point out elements in the global data. The channel matrix used to construct the covariance matrix contains indexes that point to the global input data.

The local search algorithm will modify the elements inserted in the channel matrix. A channel present in global data and absent in channel matrix is randomly chosen, and that channel is inserted into the appropriate partition.

Algorithm 3 Local Search

```

1: input {estConst, T, Iter, estInt, M, numChalnt };
2: output{resultSearch };
3: estConst  $\leftarrow$  current solution ;

```

```

4: cond ← condition number of current solution ;
5: list ← initialize the list with size T ;
6: posLi ← 1 ;
7: while list is not empty do
8:   count ← 0 ;
9:   k ← 1 ;
10:  while k ≤ T and count < Iter do
11:    cop ← copies of the current solution channels ;
12:    while c ∈ channels do
13:      c ← Randomly choose a channel of M not present in
estConst ;
14:    end while
15:    tamEstInt ← size of estInt ;
16:    q ← 1 ;
17:    while q < tamEstInt do
18:      if c > estInt[q] and c < estInt[q+1] then
19:        column ← q ;
20:        h ← insert in an arbitrary line of cop ;
21:        cop{h,column} ← c ;
22:      end if
23:    end while
24:    meEs ← submatrix M index by cop ;
25:    meCo ← covariance matrix of meEs ;
26:    conMe ← condition number of meCov ;
27:    if conMe < cond then
28:      list{posLi} ← [meEs, meCo, conMe] ;
29:      posLi ← posLi + 1 ;
30:    end if
31:  end while
32:  if the list is not empty then
33:    current solution ← select an element in the list ;
34:    restart the list ;
35:    resultSearch ← current solution ;
36:  end if
37: end while

```

The following are given as input parameters for Algorithm 3: *estConst*, solution determined in the construction step; *T*: size of the list used in the local search; *Iter*: maximum number of iterations to be completed; *estInt*: data structure that defines the channels inserted in the partitions of the original data; *M*: the matrix containing the spectral counts in all channels; *numChalnt*: a data structure containing the channels in each partition, given a subset with a number of channels *Num*.

The solution calculated in the constructive step (Algorithm 2) becomes the current initial solution of Algorithm 3. The local search routine will contain a list with all solutions that are better than current solution. A local minimum will be determined when this list is empty after a local search iteration. The channels that define the current solution are copied in Algorithm 3, line 11. An arbitrary channel *c* is chosen from global data if that channel is not inserted into the current solution (line 13). This channel will be inserted into the appropriate partition, thus defining a neighbor for the current solution (lines 19 to 21).

In Algorithm 3, line 20, a random value is determined for the position where the channel will be inserted. In line 24, the covariance matrix is constructed from the values provided by the copy of the current modified solution (*cop*). The covariance matrix and its condition number are calculated in line 25. The new solution obtained by modifying the current solution is added to the local search algorithm list if the new calculated condition number is lower than the current solution's condition number. A new value is chosen randomly from the local search list (line 32) to be the new current solution in the next iteration, if the list is not empty. If the list is empty, the current solution is a local minimum, and the local search will end.

The set of channels selected, their respective covariance matrix and the associated condition number are used to define the *resultSearch* data structure provided in the output of the Algorithm 3 (line 35).

The solution found by the local search will be subjected to a linking with a solution in the GRASP list.

3.1.3. The algorithm path-relinking

The Path-Relinking algorithm will search for a better quality solution in the region between two different solutions. The best quality solution is called *target* solution, while the second solution is called *start* solution. The differences between the two solutions are used to produce a search in the solution space between them. For each of the differences, a list is populated. Subsequently, a solution is randomly chosen from the list of Path-Relinking. This chosen solution will be the start solution in the next iteration, if the condition number associated with it is less than the condition number of the target solution. However, if the condition number of the solution removed from the list has a lower value than that of the target solution, the solutions invert their roles (see Fig. 1).

In Algorithm 4, line 5, the differences between the solutions are calculated. In lines 6 and 7 the highest and the lowest condition numbers are initialized. Line 8 will be performed while there are differences between the two solutions. Line 9 initializes the list of the Algorithm, whereas line 12 scans the channels present in the solution *s*₂ (called *start* solution) and absent in *s*₁ (called *target* solution). These differences are described by a matrix *di1* of the same size of the matrices that define the channels inserted in the solutions that have used *target* and *start*. The matrix *di1* will describe the partitions in the columns and the channels of each partition in the rows. A position will be filled with 1 in case the channel is common to both solutions. Otherwise, it will be filled with zero to indicate channel present in *start* that is absent in *target*. The channels present in one of the solutions (*target*) and absent in the other solution (*start*) are determined in line 12. A copy of the channels that define the solution *start* is made in line 15. Line 16 checks the *n*th position of the target, verifying whether there is a missing channel in the start solution or not.

Algorithm 4, line 17 inserts a channel present in the *i*, *j*-th target position and that is absent in start. The following are calculated: in line 18, matrix *mat* of spectral counts index from *copC1*; in line 19, matrix *tot* of spectral counts total index for *copC1*; covariance matrix *matCov* from *mat* in line 20; orthogonal matrix *matOr* (line 21) from *matCov*; the condition number of the covariance matrix (line 22) and orthogonal matrix (line 23). In the line 24, a structure *struct* containing the values *matCov*, *matOrt*, *c*₁, *c*₂ is defined.

If the condition number obtained by modifying the solution named start is lower than the condition number found so far (line 25), it is inserted in the list (line 26). The values of the variables are updated (lines 27 to 33). If the list is filled, the element that is most similar to the current element and has the highest condition number is removed (lines 35 to 38).

After the differences between the two solutions are verified one by one, one element is chosen from the list to be the new solution *start* (line 43 to line 45). From that moment, the Algorithm performs another iteration after choosing a new solution *start* (line 44), decrementing the variable *quaDif*.

Algorithm 4 Path-relinking

```

1: input{s1, s2, numChalnt, M, total} ;
2: output{resulPath} ;
3: c1 ← channels of s1 ;
4: c2 ← channels of s2 ;
5: quaDif ← number of differences between s1 and s2 ;

```

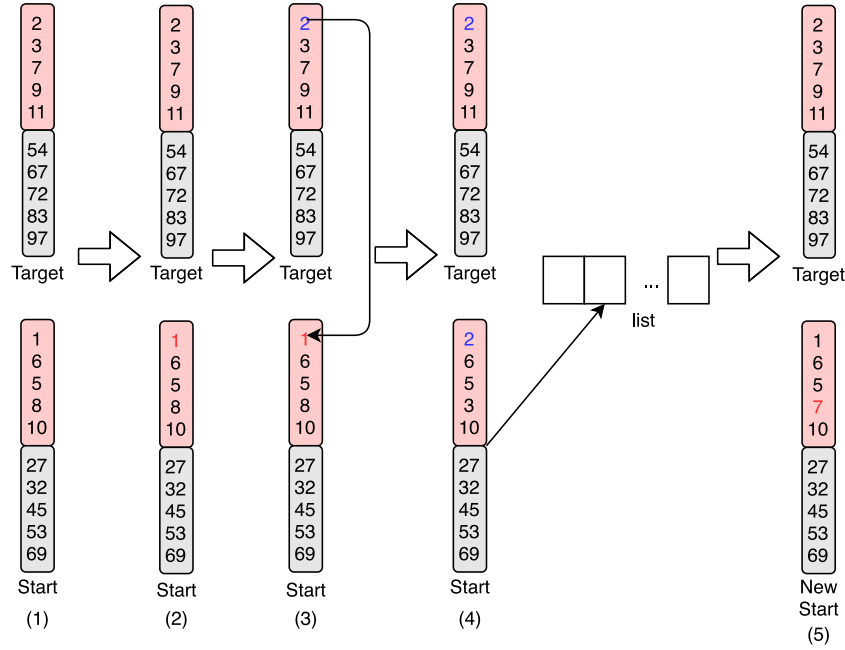


Fig. 1. (1) Two solutions provided as parameters to the Path-Relinking (PR) function with partitions in different colors; (2) In the start solution, a channel is chosen that is not present in the target solution (channel represented in red); (3) choose a channel index in the target solution (channel represented in blue) that is not present in the start solution and replace the channel chosen in step (2); (4) When replacing the channel, the covariance matrix is constructed, the condition number of this modified solution is verified. If this condition number is less than the condition number of the current start solution, the new solution is inserted in the list; (5) At random, a solution is chosen from the list to be the new current start solution and the whole process is restarted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

```

6:  $menCond \leftarrow \inf$ ;
7:  $maxCond \leftarrow 0$ ;
8: while  $quaDif > 0$  do
9:   start the list;
10:   $n \leftarrow 1$ ;
11:   $iter \leftarrow 1$ ;
12:   $di_1 \leftarrow$  channels presents in  $s_2$  and absents in  $s_1$ ;
13:  for all rows  $i$  in  $di_1$  do
14:    for all columns  $j$  in  $di_1$  do
15:       $copC1 \leftarrow c_1$ ;
16:      if  $di_1(i, j) = 0$  then
17:         $copC1 \leftarrow c_2(i, j)$ ;
18:         $mat \leftarrow$  matrix  $M$  index for  $copC1$ ;
19:         $tot \leftarrow$  matrix total index for  $copC1$ ;
20:         $matCov \leftarrow$  covariance matrix from  $mat$ ;
21:         $matOrt \leftarrow$  orthogonalization of  $matCov$ ;
22:         $c_1 \leftarrow$  condition number of  $matCov$ ;
23:         $c_2 \leftarrow$  condition number of  $matOrt$ ;
24:         $struct \leftarrow [matCov, matOrt, c_1, c_2]$ ;
25:        if  $c_1 < menCond$  then
26:           $lista\{iter\} \leftarrow struct$ ;
27:          if  $c_1 < menCond$  then
28:             $menCond \leftarrow c_1$ ;
29:             $resulpath \leftarrow struct$ ;
30:          end if
31:          if  $c_1 > maxCond$  then
32:             $maxCond \leftarrow c_1$ ;
33:          end if
34:        else
35:          if  $c_1 < maxCond$  then
36:             $point \leftarrow$  list element most similar to cur-
37:            rent;
38:             $list\{point\} \leftarrow struct$ ;
39:          end if

```

```

40:   end if
41:   end for
42:   end for
43:   if list is not empty then
44:      $c_1 \leftarrow$  randomly select an element from the list;
45:   end if
46:    $quaDif \leftarrow quaDif - 1$ ;
47: end while

```

3.2. QR decomposition with Gram–Schmidt

The covariance matrices calculated in the Algorithm 1 in lines 11, 24, 26 and 27 is based on the linear independence and partitioning of the data are subjected to orthogonalization of its columns by the QR decomposition method.

A rectangular matrix $A \in \mathbb{R}^{m \times n}$ can be factors in a product of an orthogonal matrix $Q \in \mathbb{R}^{m \times n}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ so that $A = QR$ [46].

The QR-factorization splits the problem $Ax = b$, with $A = QR$, into two subproblems:

$$\begin{aligned} Qy &= b \\ Rx &= y \end{aligned} \quad (9)$$

where $y = Q'b$ and R is upper triangular. Since Q is orthogonal, $\|Q\| = 1$ and $\|Q^{-1}\| = \|Q'\| = 1$, where:

$$cond(Q) = 1 \longrightarrow cond(A) = cond(QR) = cond(R) \quad (10)$$

QR-factorization decouples the original problem $Ax = b$ into two problems. One of them ($Qy = b$) has no error amplification, while the other one ($Rx = y$) has the minimal possible error amplification allowed by the inherent error amplification of the original problem. Orthogonal transformations are stable.

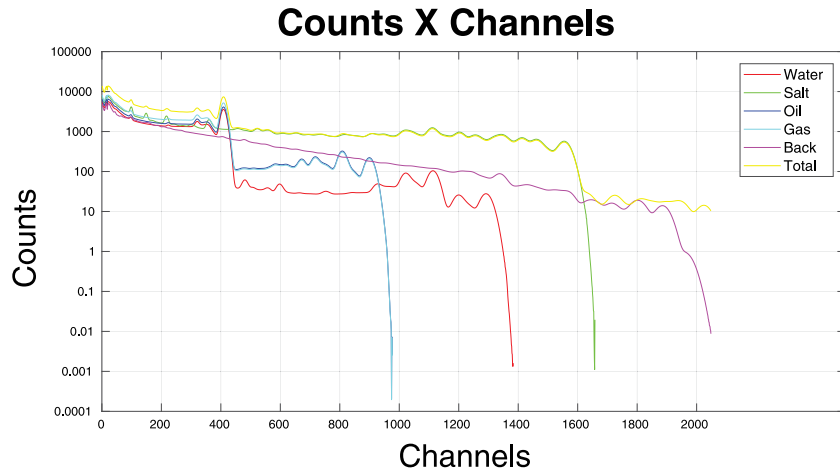


Fig. 2. The prompt gamma-ray library spectra of the oil, salt, water and gas phases. The total background library is also shown.

Table 1
Experimental Configuration.

Components	Description
Processor	Intel® Core™ i7-4770K CPU
Clock processor	3.50 GHz
Core	8
Memory	16GB in dual channel
Software	MATLABR2012a
Operational System	Debian GNU/Linux 8 (jessie) 64-bit

Table 2
Instances.

Instance	Compound	Sum
1	Salt	Oil + Gas + Water + Background
2	Oil	Salt + Gas + Water + Background
3	Water	Salt + Oil + Gas + Background
4	Gas	Salt + Oil + Water + Background
5	Water	Oil + Gas + Background
6	Gas	Water + Oil + Background
7	Oil	Gas + Water + Background
8	Water	Gas + Background
9	Gas	Water + Background

4. Experimental setup

The experimental setup used five spectral library counts (water, salt, oil, gas and background) in order to obtain covariance matrices with less condition number. Details on the characteristics of the data and the conditions under which it was generated can be found at [29]. The characterization of produced water samples is a crucial part of the subsea process and production optimizations and monitoring (see Fig. 2).

The number of channels chosen to compose a new solution was 100, out of a total of 2048, since this is approximately the size of the support of a photopeak of any of the constituents used.

All experiments were performed by using the MATLABR2012a [47] tool, which includes much of the functionality needed to work problems with this specification. The hardware used in the experiments is described in Table 1:

The technique used to construct the covariance matrix A is present in [3]. This technique works by isolating the most significant library from the others, which are combined to work as just one compound. The composed library is produced simply by adding their spectra counts, following the calculation of the covariance matrix:

$$A_{i,j} = \sum_{k=1}^N \frac{1}{\sigma_i^2} f_k(x_i) f_k(x_j) \quad (11)$$

The combinations of libraries are described in Table 2.

5. Results

The results presented in this section consider the procedures described in Section 3, and also presented in [29]. These procedures can be summarized as the sum of all the libraries against one of them separately, so the condition number can be calculated. The condition number of values of the covariance matrices of this problem that guide this work can be found in [29].

Table 3 shows the results and their comparisons with some current approaches [29]. The first and second columns of Table 3, respectively, indicate the order number of the instance and the respective condition number values that are to be minimized. The values reached by the algorithm are described in the third column in Table 3, and the percentage improvement values for the condition number of covariance matrices obtained by partitioning the input data in relation to [3] are described in the fourth column. In the fifth column, the values of the condition number for orthogonalized covariance matrices by the QR decomposition method. The percentage of improvement in values in the condition number obtained in the sixth column compared to those in the second column are described in the sixth column. In the last column, the results for the runtime of the algorithm. The parameters presented in Table 3 were obtained by defining the following parameters: (1) The size of the GRASP list set to 10; (2) 14 iterations of the GRASP algorithm; (3) A solution built with 100 channels chosen at random.

Instances 4, 6 and 9 exceeded the target values (second column in Table 3) in [29] after the constructive step of Algorithm 1. Instances 5, 7 and 8 reach the goal in the first GRASP iteration. For instances 2 and 3 in Table 2, the objective of minimizing the conditioning of covariance matrices in data partitioning was not achieved. The covariance matrices orthogonalized by the QR method of instances 2 and 3 in Table 2 achieved the goals described in the work of [29]. The lines, or columns, of covariance matrices that have almost linear dependence, may be responsible for ill-conditioning. The difference in magnitude between the eigenvalues of the covariance matrix can also be enumerated as responsible for configurations presenting unwanted conditioning number values. Conditioning number is also related to orthogonality. Matrices whose rows, or columns, show a tendency to orthogonality are more likely to demonstrate a low condition number.

Table 4 shows the average results for thirty GRASP executions. For each of the experiments, different seeds are used for the

Table 3
Results of the experiments.

Instance	Goal [29]	Achieved	Improvement	Orthogonalized	Improvement	Time execution (s)
1	3.09	2.64	14.56%	2.52	18.44%	190.52
2	21.82	29.09	−33.31%	1.12	94.86%	272.83
3	50.62	65.35	−29.09%	1.15	97.72%	91.2
4	34.58	7.76	77.55%	1.06	96.93%	388.44
5	47	13.95	70.31%	1.38	97.06%	85.52
6	35.1	13.23	62.30%	1.19	96.60%	300.00
7	21.6	10.23	52.63%	1.0	95.37%	194.32
8	18	12.93	28.16%	1.49	91.72%	69.08
9	12.5	7.67	38.64%	2.14	72.88%	174.17

Table 4

Result of experiments for thirty executions with different seeds and with 95% confidence intervals. It is important to note that the different iterations are associated with the instances as follows: Background | Water + Salt + Oil + Gas (iteration 1); Salt | Water + Oil + Gas (iteration 2); Water | Oil + Gas (iteration 3); Oil | Gas and Gas | Oil (iteration 4).

Iteration	References [29]	Condition number	Coefficients	Error(%)
1	N/A	$1.01 \times 10^0 \pm 3.47 \times 10^{-4}$	$1.01 \times 10^{-2} \pm 4.81 \times 10^{-6}$	N/A
2	0.01	$1.62 \times 10^2 \pm 3.12 \times 10^{-2}$	$1.00 \times 10^{-2} \pm 1.32 \times 10^{-18}$	$2.49 \times 10^{-6} \pm 1.66 \times 10^{-7}$
3	0.34	$1.02 \times 10^0 \pm 1.62 \times 10^{-3}$	$3.21 \times 10^{-1} \pm 2.16 \times 10^{-3}$	$2.86 \times 10^{-2} \pm 6.55 \times 10^{-3}$
4	0.45	$4.12 \times 10^1 \pm 1.83 \times 10^{-2}$	$4.50 \times 10^{-1} \pm 7.37 \times 10^{-17}$	$1.02 \times 10^{-14} \pm 3.25 \times 10^{-15}$
4	0.20	$4.12 \times 10^1 \pm 7.43 \times 10^{-3}$	$2.00 \times 10^{-1} \pm 2.84 \times 10^{-17}$	$2.18 \times 10^{-14} \pm 5.64 \times 10^{-15}$

GRASP randomization stage, such as the choice of channels for the matrix that will index the covariance matrix. The lines in Table 4 define: (1) the iteration of the GRASP algorithm (first column); (2) Percentage reference counts for each element found in [29] (second column); (3) Average values obtained for the condition number after thirty executions of GRASP (third column); (4) Spectral count percentages of each element, and respective confidence interval (fourth column); (5) Errors associated with the values in column 4, and respective confidence intervals (fifth column).

Line three of Table 4 reports that, in the third iteration, where all the possibilities of isolating a library of spectral counts in the first column were analyzed, against the sum of spectral counts of all other libraries in the second column, the thirty executions found the least condition number for instance composed of the separation of the water against the sum of all other compounds. For this scenario, the average conditioning number value was 1.02 for the thirty runs; the average percentage coefficient value was 3.21×10^{-1} and the error was $2.86 \times 10^{-2}\%$.

Line four of Table 4 determines that the least conditioning number was found in two different instances, for example, thirteen and fourteen. For each of the cases of iteration four, approximately the same average condition number values were reported. The warning to be made is that in line 1 of Table 4 the abbreviation Not Applicable (N/A) was used because there is no such instance in work used as a Ref. [29].

Details of built solutions in instances 2 and 3 in Table 3 are shown in Figs. 3–8. The orthogonalization algorithm calculates the conditioning number of two matrices: (1) determines the condition number of the orthogonalized covariance matrix and (2) the condition number of the matrix that transforms the canonical covariance matrix to the covariance matrix on the orthonormal basis.

The high variability in the number of covariance matrices justifies the use of heuristic algorithms to obtain the covariance matrices that have the lowest condition number (Figs. 3–8). The orthogonalization process of covariance matrices drastically reduces the associated condition number. In Figs. 3 and 6 the condition number of the covariance matrices calculated in the constructive step, comparatively larger than the condition number values for the orthogonalized covariance matrices obtained in 4 and 7 is possible to be observed.

Figs. 9–14 demonstrate the results of the first local search for solutions in the vicinity of the solution constructed in the

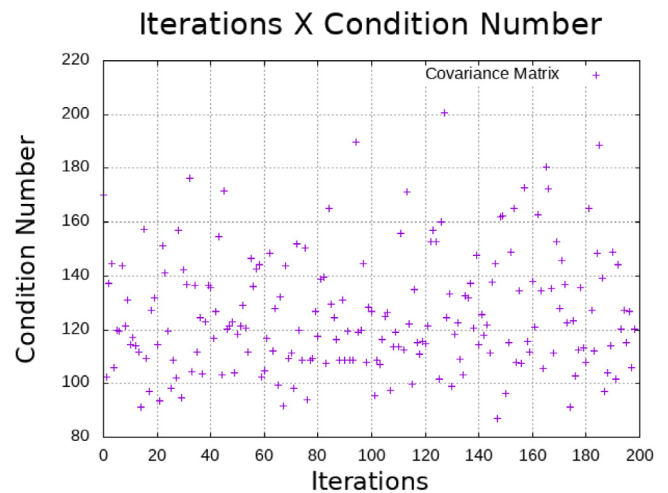


Fig. 3. Condition number of covariance matrices calculated during the construction phase of a solution.

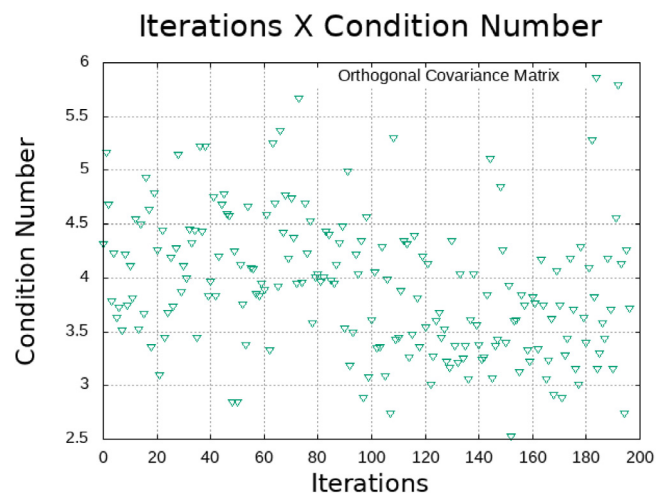


Fig. 4. Condition number of the covariance matrices presented in Fig. 3 orthogonalized by Gram–Schmidt process.

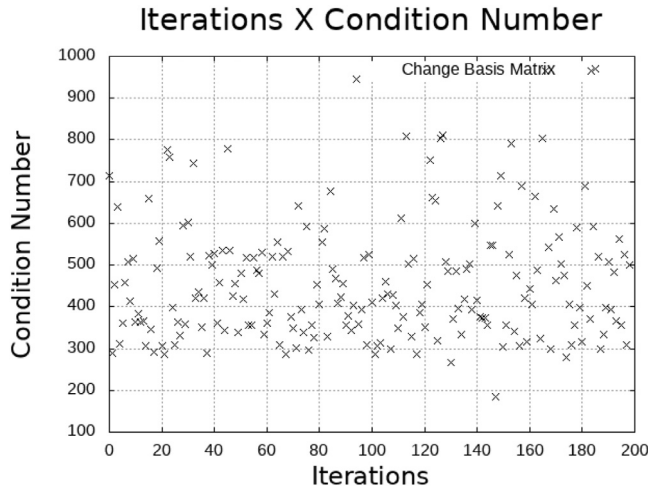


Fig. 5. Condition number of the matrices that make the change to the Gram-Schmidt base of the matrices presented in Fig. 3.

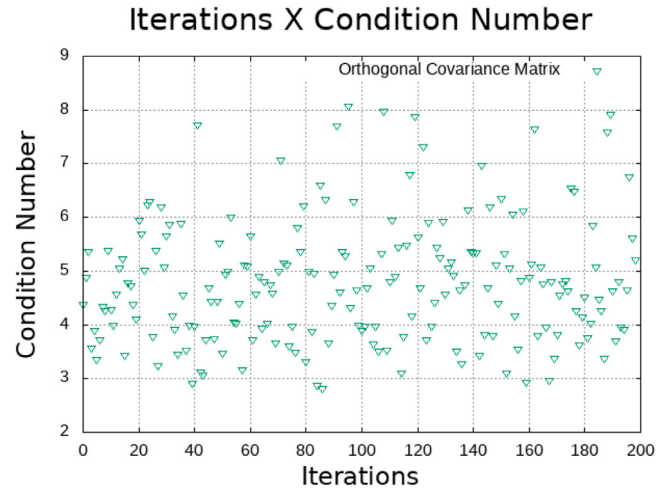


Fig. 7. Condition number of the covariance matrices presented in Fig. 6 orthogonalized by Gram-Schmidt process.

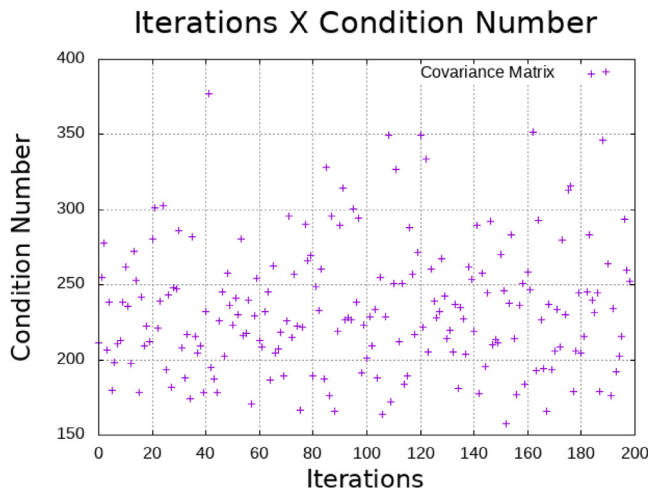


Fig. 6. Condition number of covariance matrices calculated during the construction phase of a solution.

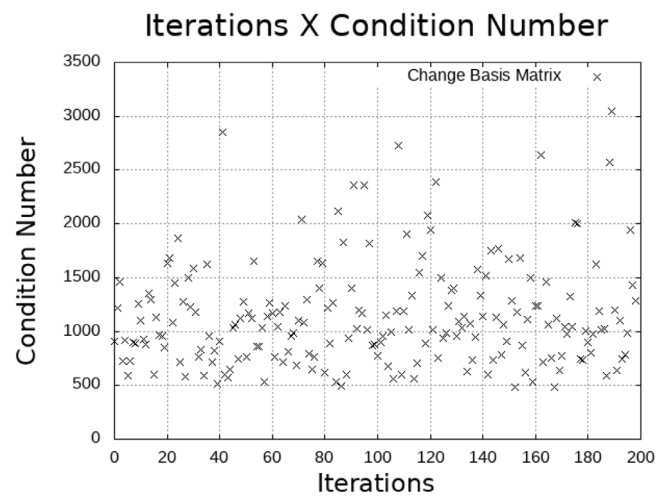


Fig. 8. Condition number of the matrices that make the change to the Gram-Schmidt base of the matrices presented in Fig. 6.

constructive step (Figs. 3–8). The results presented are from three local search calls during the execution of Algorithm 1. The condition numbers presented in Figs. 10 and 13 for the orthogonalized covariance matrix are lower than the condition numbers presented in Figs. 9 and 12. The local search is efficient in reducing the condition number of the covariance matrix constructed by Algorithm 2.

The results of three different calls to Algorithm 4 are shown in Figs. 15–20. Similarly as occurred in the execution of other sub-routines of GRASP, covariance matrices that have suffered QR factorization (Figs. 16 and 19) show lower condition numbers when compared to the condition number of other (Figs. 15 and 18).

After running the Algorithm 4, a second local search is performed around the result that has been returned (Figs. 21–26). This second search follows the same criteria established in the first local search. Although, there is no clear trend of convergence as shown in Figs. 9–14, the bounded range for the values of the condition number tends to be more restricted as the algorithm increases the number of iterations. Once again, the condition numbers of values of the covariance matrices that are subjected to orthogonalization for QR factorization (Figs. 22 and 25) have a

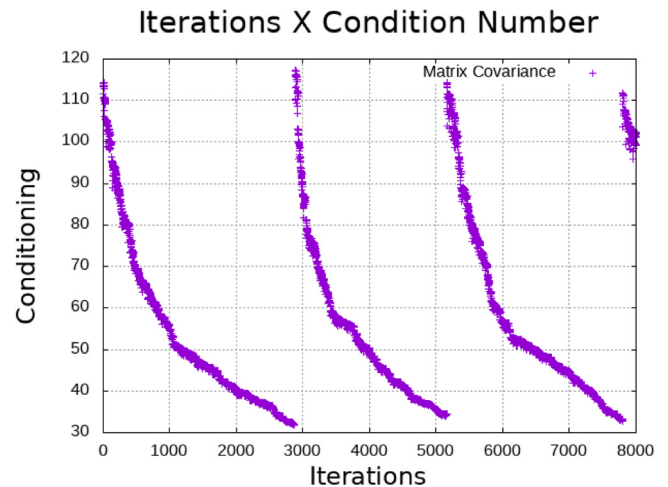


Fig. 9. Condition numbers of covariance matrices calculated during the first local search phase in the vicinity of the solution constructed in the constructive step.

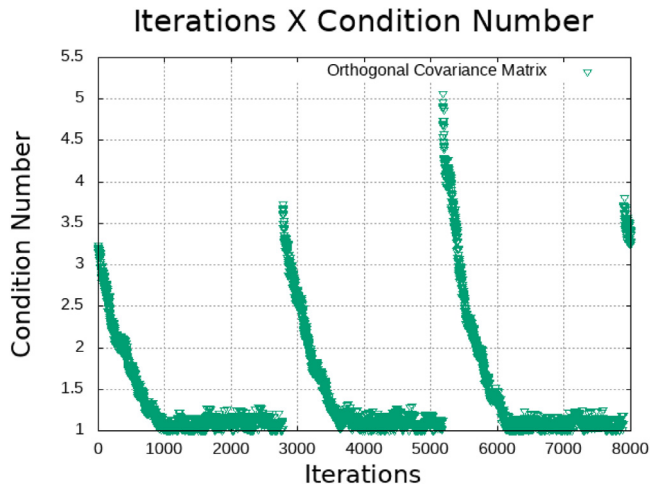


Fig. 10. Condition number of the covariance matrices presented in Fig. 9 orthogonalized by Gram–Schmidt process.

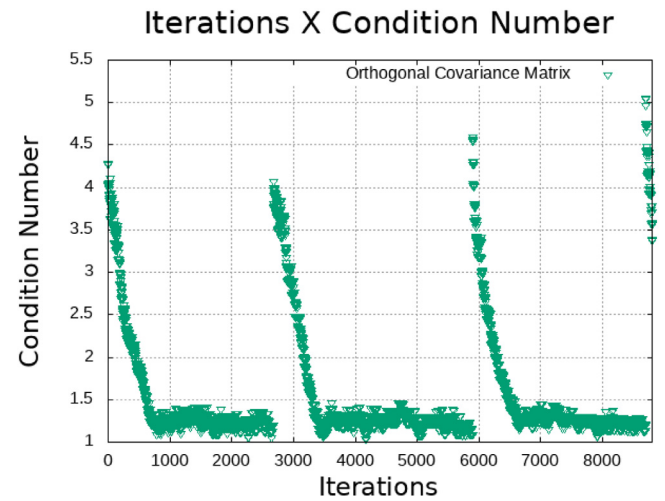


Fig. 13. Condition number of the covariance matrices presented in Fig. 12 orthogonalized by Gram–Schmidt process.

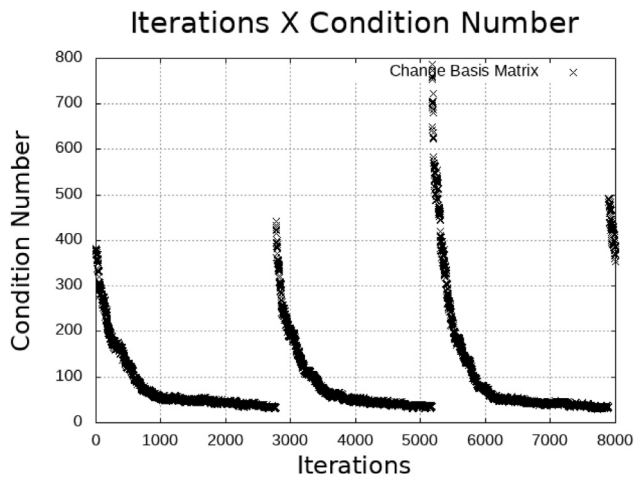


Fig. 11. Condition number of the matrices that make the change to the Gram–Schmidt base of the matrices presented in Fig. 9.

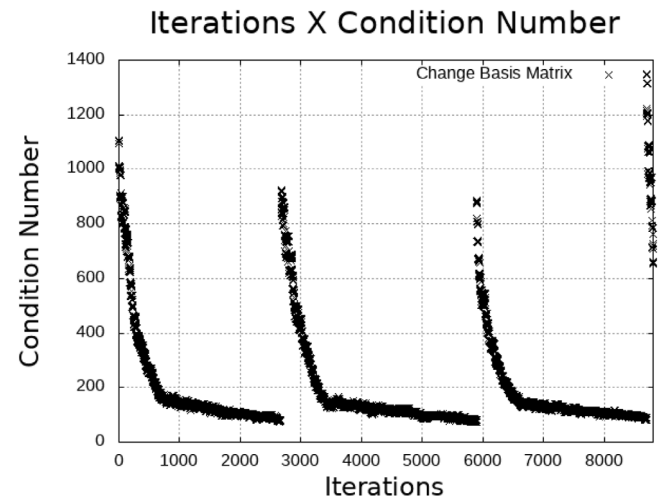


Fig. 14. Condition number of the matrices that make the change to the Gram–Schmidt base of the matrices presented in Fig. 12.

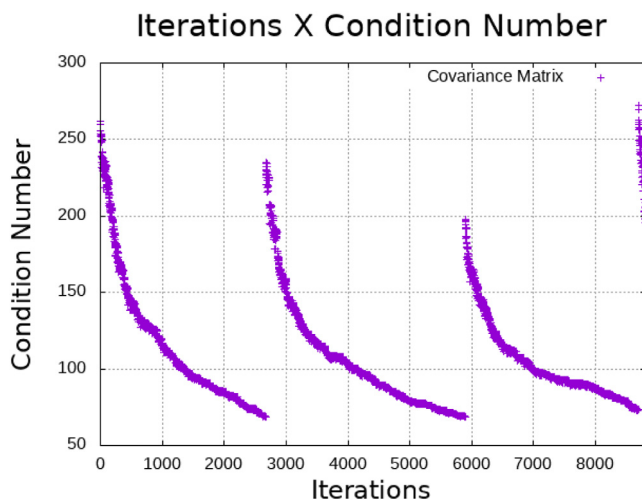


Fig. 12. Condition numbers of covariance matrices calculated during the first local search phase in the vicinity of the solution constructed in the constructive step.

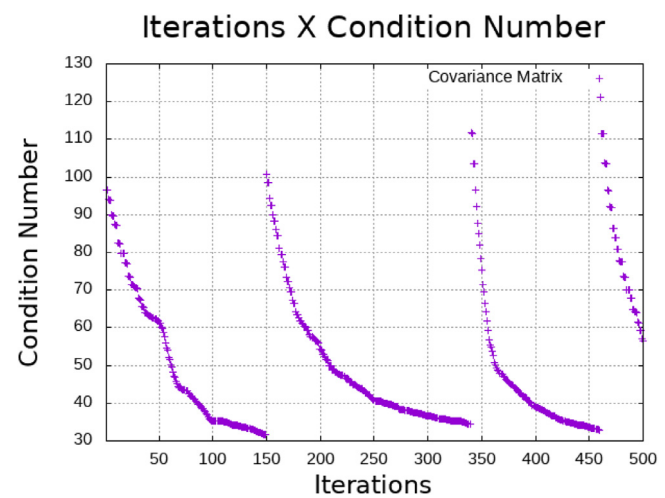


Fig. 15. Condition number presented by the covariance matrices during path-relinking process.

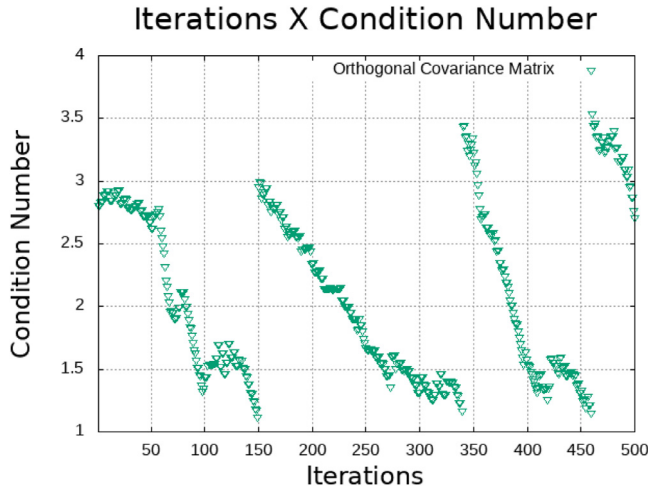


Fig. 16. Condition number of covariance matrices in Fig. 15 orthogonalized by Gram–Schmidt process.

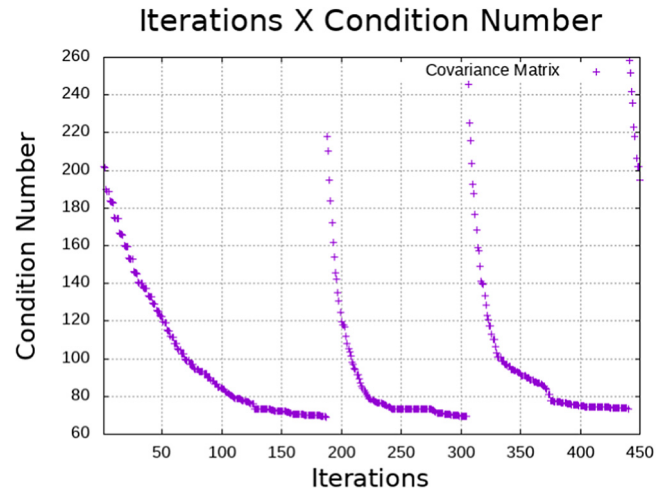


Fig. 18. Condition number presented by covariance matrices during the path-relinking process.

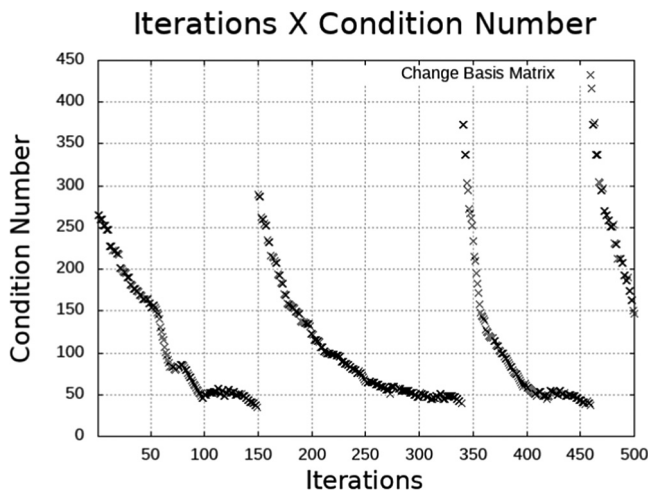


Fig. 17. Condition number of the matrix that transforms the covariance matrix of the canonical basis for the Gram–Schmidt basis during the Path-Relinking process of the matrices presented in Fig. 15.

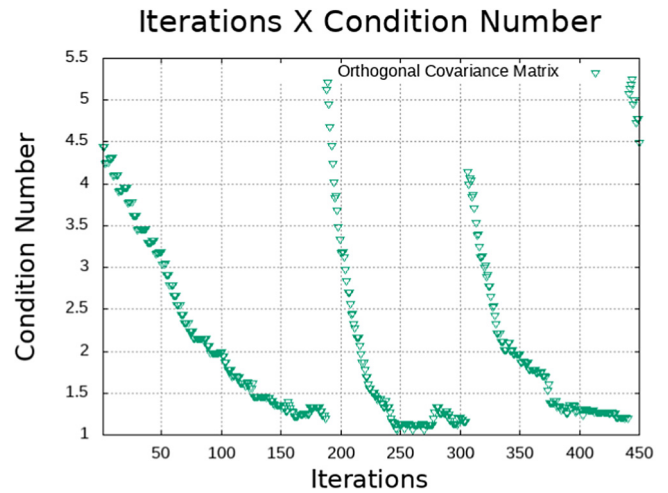


Fig. 19. Condition number of covariance matrices in Fig. 18 orthogonalized by Gram–Schmidt process.

lower condition number than other matrices (Figs. 21, 23, 24, and 26).

In Figs. 21–26, the outputs of the iterations of the GRASP algorithm for instance three are presented. The outputs of instance two will be omitted due to the similarity of the results. The outputs to GRASP subroutines (Algorithm 3, GRASP List and Algorithm 4) are shown in Figs. 27–32 for the covariance matrix only. In Fig. 30 are also demonstrated the results (Second local search) for the canonical base covariance matrices (Fig. 30), orthogonalized covariance matrices (Fig. 31) and transformation matrices from the canonical base to the orthonormal base (Fig. 32). In order to avoid redundancy, it was decided to present the results in this way: the results shown in each step can be generalized to the other matrices presented, and the outputs for a single subroutine can be generalized to the other steps of the algorithm.

Due to the efficiency of the Local Search and Path-Relinking algorithms, it is possible to observe how the condition numbers tend to converge with an average of 100 iterations. It is also perceived as a threshold from which the condition number will not decrease, regardless of the number of iterations to be completed.

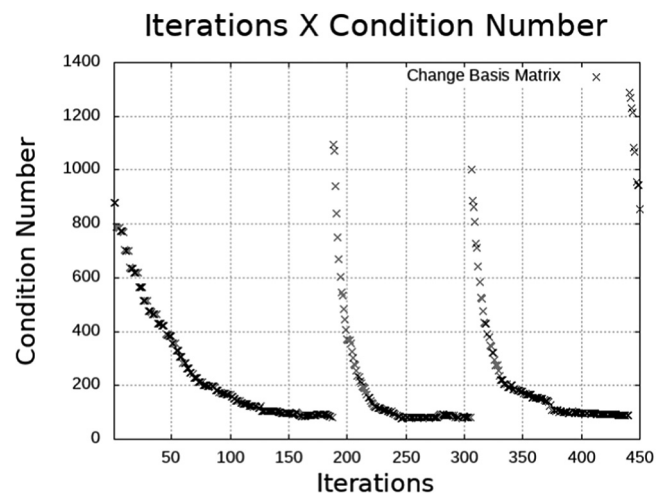


Fig. 20. Condition number of the matrix that transforms the covariance matrix of the canonical basis for the Gram–Schmidt basis during the path-relinking process of the matrices presented in Fig. 18.

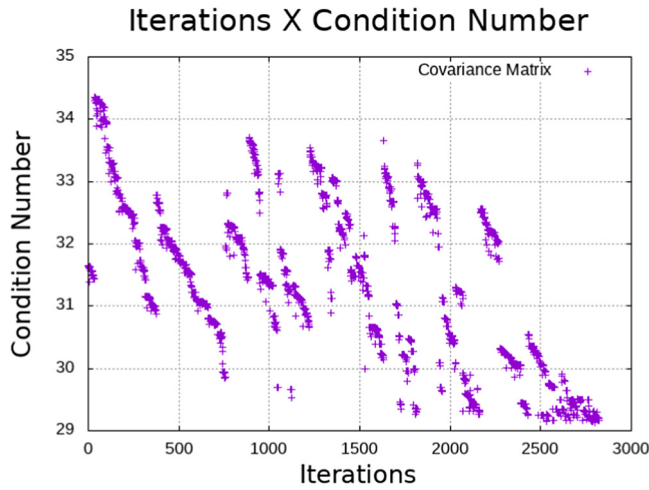


Fig. 21. Condition number presented by the covariance matrices during the second search.

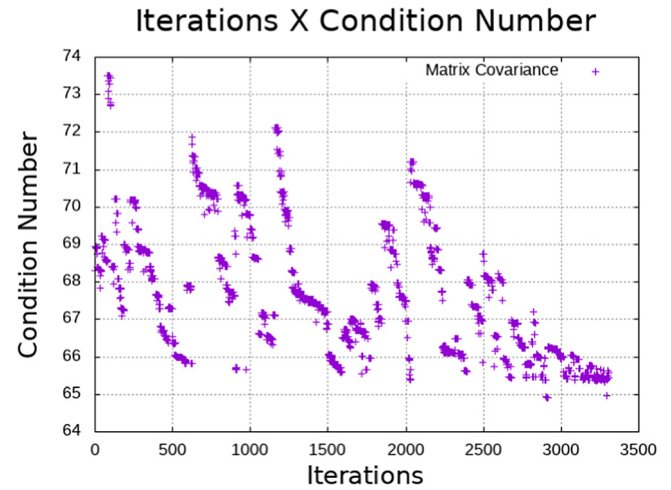


Fig. 24. Condition number presented by the covariance matrices during the second search.

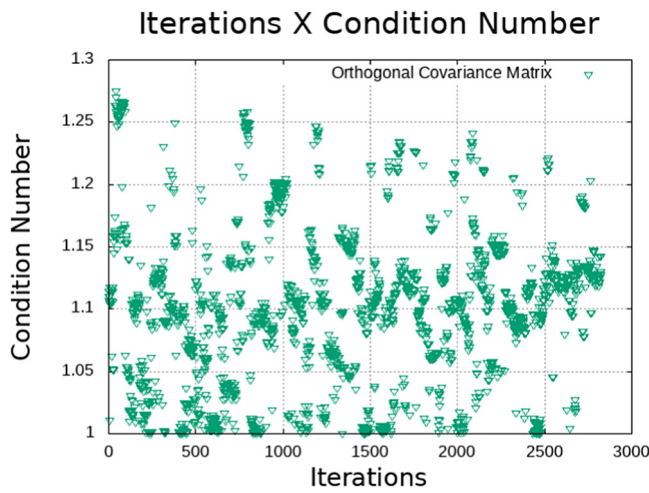


Fig. 22. Condition number of covariance matrices in Fig. 21 orthogonalized by Gram–Schmidt process.

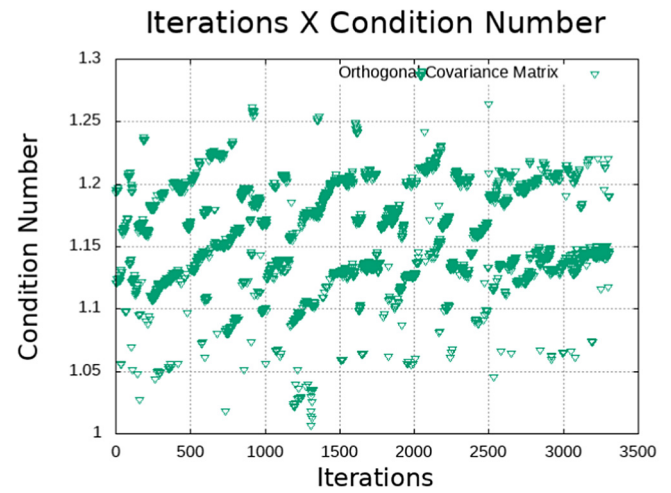


Fig. 25. Condition number of covariance matrices in Fig. 24 orthogonalized by Gram–Schmidt process.

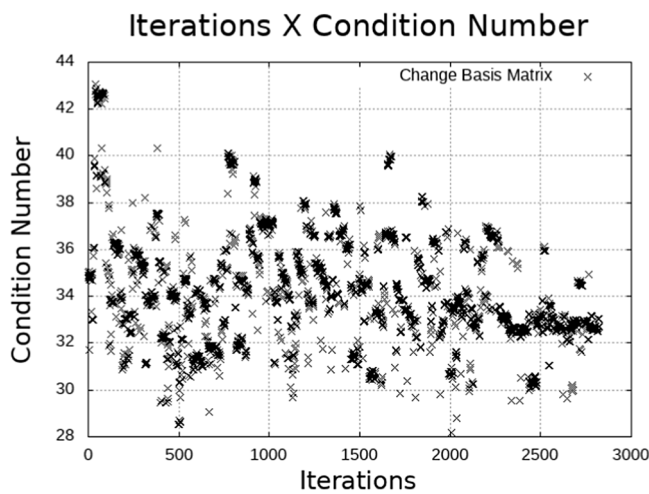


Fig. 23. Condition number of the matrix that transforms the covariance matrix of the canonical basis for Gram–Schmidt basis during the second search of the matrices presented in Fig. 21.

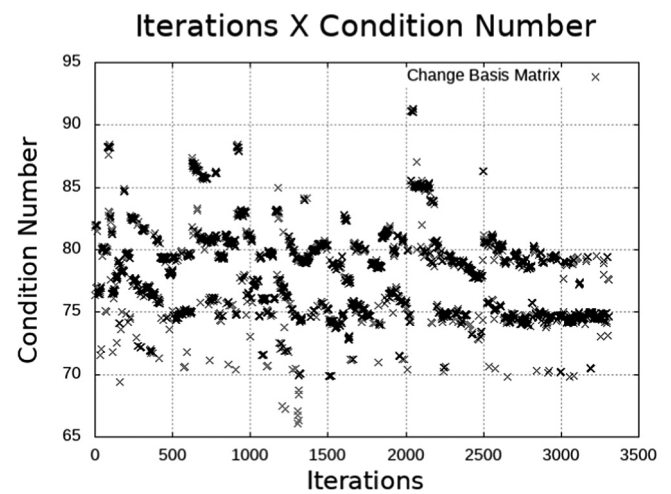


Fig. 26. Condition number of the matrix that transforms the covariance matrix of the canonical basis for the Gram–Schmidt basis during the second search of the matrices presented in Fig. 24.

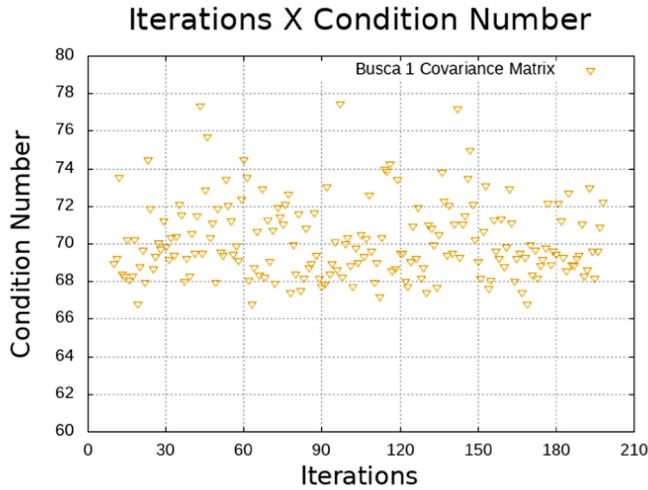


Fig. 27. Condition Number for the covariance matrices on first search of GRASP.

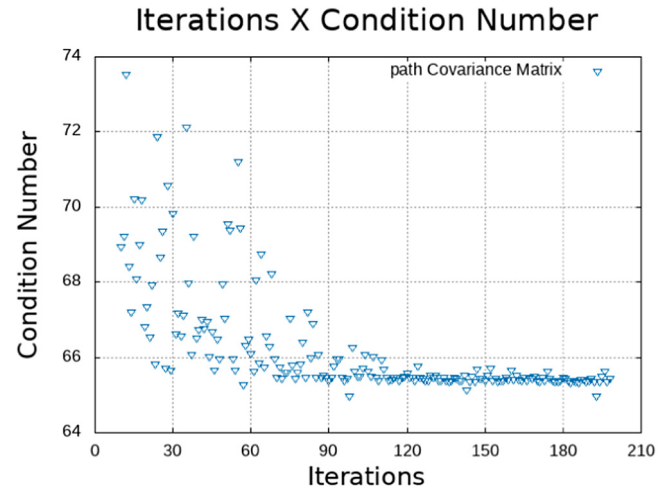


Fig. 29. Condition number values of covariance matrices calculated by Path-Relinking.

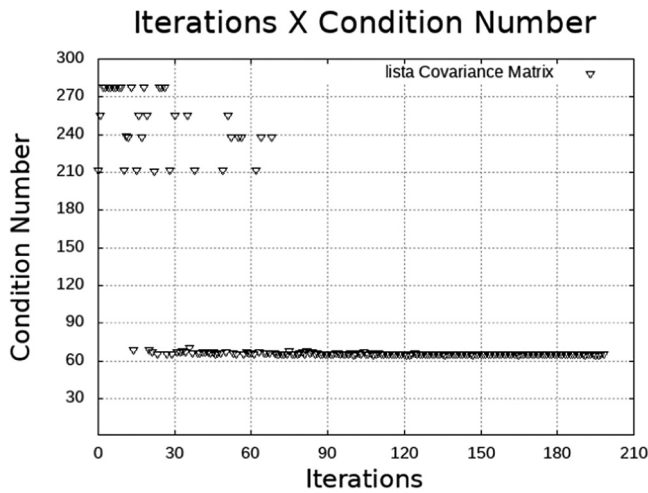


Fig. 28. Condition number values of the covariance matrices chosen from the list of GRASP chosen to be one of the solutions for Path-Relinking.

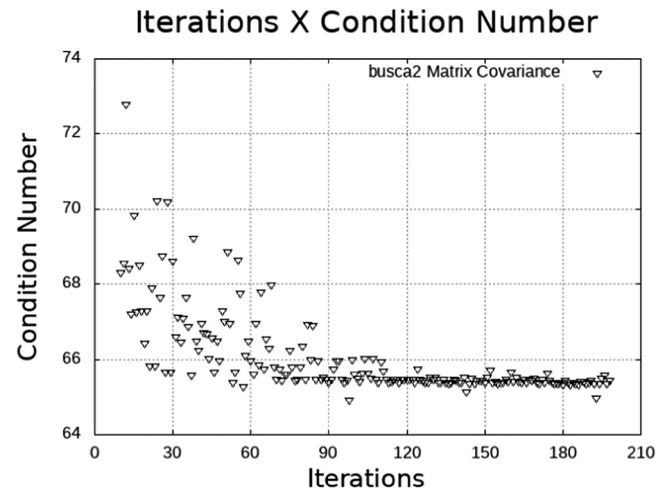


Fig. 30. Condition number values of covariance matrices calculated by second Local Search.

6. Conclusions

PGNAA is a powerful and sensitive quantitative analysis tool when used in conjunction with MCLS methods applied to estimation of quantities of different phases in multiphase flow. Unfortunately, the accuracy of MCLS depends on: (1) the existence of two or more correlated libraries; (2) the sample containing traces of elements with relatively low thermal neutron capture cross sections. In these cases, the condition number of the covariance matrix will be too high, implying significant uncertainties in the estimated parameters. The current work is based on iterations in which the success criteria are the array conditioning number. The results indicate that the approach may be adequate to treat cases of ill-conditioning. This fact can be confirmed by estimation of the proposed method compared to the results of [29].

The GRASP algorithm consists of several steps that aim to reach the objective function, that is, to minimize the condition number of the covariance matrices. In these steps, one might find: (1) a construction of a candidate solution; (2) searches around that candidate solution; (3) linkage between the search result and a solution constructed by the GRASP algorithm.

The search for covariance matrices that present a lower condition number than those shown by [29] was successful in seven

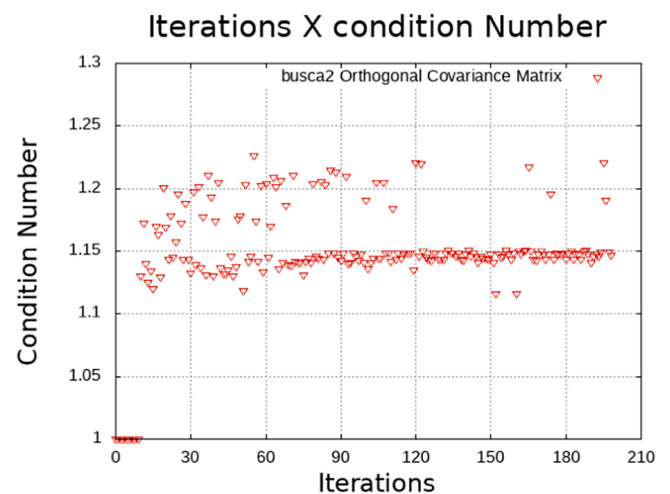


Fig. 31. Condition number values of orthogonal covariance matrices calculated by second Local Search.

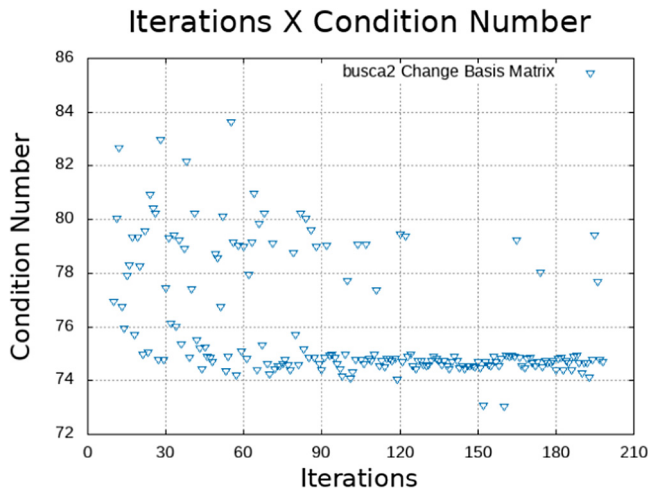


Fig. 32. Condition number values in the matrices that change a canonical basis for orthogonal basis, calculated by second Local Search.

of the nine instances presented in Table 3, partitioning the input data by following criteria of linear independence and algorithm concept heuristics (GRASP). The two cases that were not solved through the partitioning of the data had their values exceeded by orthogonalization QR of the covariance matrices, immediately after being partitioned and submitted to GRASP.

The randomness in the condition number of the covariance matrices is possible to be noticed in Figs. 3–8. This feature confirms the inherent random behavior of the GRASP algorithm. In contrast, search (Figs. 9–14 and Figs. 21–26) and linkage algorithms (Figs. 15–20) have a strong tendency to decrease the condition number of covariance matrices from a solution in the construction step, proving the GRASP algorithm's "greedy" characteristic. Emphasizing that each plot in Figs. 9 to 14, 15 to 20 and 21 to 26 is there for a single execution of the algorithm search and linking.

The outputs of the search and linkage algorithms during the execution of GRASP algorithm can be checked in Figs. 27–32. In these graphs, it is possible to observe that the condition number value is inserted in a more extensive range at the beginning of GRASP algorithm. This amplitude decreases as the output of the algorithm occurs for each of the GRASP subfunctions.

The execution of local search is useful in finding solutions with less condition number. The decrease in magnitude of the condition number at the end of this stage may be as high as 10. This decrease can be seen in Figs. 9 and 12 for the case of matrices in the canonical basis, as for matrices in the base Gram–Schmidt, Figs. 10 and 13. The matrices that perform the transformation to the Gram–Schmidt base show a more significant decrease in the condition number, in order of 10^2 (Figs. 11 and 14). The same conclusions for the first local search can be extended to the results brought by path relinking (Figs. 15–20).

In future works, might consider adjusting some parameters used in the present work:

1. Parameter adjustments used in the current work, such as iterations of the algorithm and list size used in GRASP execution. Research into these issues can be done by applying artificial intelligence techniques.
2. Quantitative analysis of the matrices built in the present work. A gap in the current work that may be the subject of future investigations lies in this question of the effects of minimizing the conditioning number on the PGNA quantative analysis.

3. Building a solution with a new amount of channels implies new neighborhood definitions in the execution of the local search algorithm.
4. From the description found in the paper [29], it is possible to suggest new library combinations.

The method proposed in the present work shows potential in determining the lowest number of conditioning; in determining the percentage counts of the elements and in using fewer channels of the input data, differently from the methodology used in the [29] method.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Special thanks to Mr. Severino Cardoso dos Santos (In memoriam) and Governo do Estado de Pernambuco. Silas Santos is supported by a postdoctorate grant number 88887.374884/2019-00 from CAPES.

References

- [1] H.P. Chang, I. Meric, D. Sudac, K. Nađ, J. Obhodaš, G. Hou, Y. Zhang, R.P. Gardner, *Radiat. Phys. Chem.* 155 (2019) 197–201.
- [2] J. Wang, F. Li, R.P. Gardner, *Meas. Sci. Technol.* 19 (9) (2008) 094005.
- [3] I. Meric, *Nuclear Methods for Subsea Fluid Characterization* (Ph.D. thesis), University of Bergen, 2012.
- [4] A. Caciolli, M. Baldoncin, G. Bezzon, C. Broggini, G. Buso, I. Callegari, T. Colonna, G. Fiorentini, E. Guastaldi, F. Mantovani, et al., *Sci. Total Environ.* 414 (2012) 639–645.
- [5] M. Baldoncin, M. Albéri, C. Bottardi, E. Chiarelli, K.G.C. Raptis, V. Strati, F. Mantovani, *J. Environ. Radioact.* 192 (2018) 105–116.
- [6] Z. Morsy, M.A. El-Wahab, N. El-Faramawy, *Ann. Nucl. Energy* 44 (2012) 8–11.
- [7] G. Eleftheriou, C. Tsabaris, E. Androulakis, D. Patiris, M. Kokkoris, C. Kalfas, R. Vlastou, *Appl. Radiat. Isot.* 82 (2013) 268–278.
- [8] W. El Kanawati, C. Carasco, B. Perot, A. Mariani, A.-C. Raoux, V. Valkovic, D. Sudac, J. Obhodaš, M. Baricevic, *IEEE Trans. Nucl. Sci.* 57 (5) (2010) 2879–2885.
- [9] W. El Kanawati, B. Perot, C. Carasco, C. Eleon, V. Valkovic, D. Sudac, J. Obhodaš, *Nucl. Instrum. Methods Phys. Res. A* 654 (1) (2011) 621–629.
- [10] R.P. Gardner, L. Xu, *Radiat. Phys. Chem.* 78 (10) (2009) 843–851.
- [11] R. Gardner, A. Sood, Y. Wang, L. Liu, P. Guo, R. Gehrke, *Appl. Radiat. Isot.* 48 (10) (1997) 1331–1335, [http://dx.doi.org/10.1016/S0969-8043\(97\)00127-9](http://dx.doi.org/10.1016/S0969-8043(97)00127-9), URL <http://www.sciencedirect.com/science/article/pii/S0969804397001279>.
- [12] R.P. Gardner, A. Sood, Y.Y. Wang, L. Liu, P. Guo, R.J. Gehrke, *Appl. Radiat. Isot.* 48 (10) (1997) 1331–1335.
- [13] I. Meric, G.A. Johansen, M.B. Holstad, J. Mattingly, R.P. Gardner, *Meas. Sci. Technol.* 23 (5) (2012) 055603.
- [14] A. Cichocki, S.-i. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley & Sons, 2002.
- [15] L. Salmon, *Nucl. Instrum. Methods* 14 (1961) 193–199.
- [16] I. Meric, G. Johansen, M. Holstad, J. Wang, R.P. Gardner, *Meas. Sci. Technol.* 22 (12) (2011) 125701.
- [17] H. Shahabinejad, N. Vosoughi, *Nucl. Instrum. Methods Phys. Res. A* 911 (2018) 123–130.
- [18] M. Baldoncin, *Applicazione del metodo non negative least square alla full spectrum analysis nel processo di calibrazione di uno spettrometro di raggi gamma portatile* (Graduate thesis), University of Ferrara, 2010.
- [19] E. Androulakis, M. Kokkoris, C. Tsabaris, G. Eleftheriou, D. Patiris, F. Pappa, R. Vlastou, *Appl. Radiat. Isot.* 114 (2016) 76–86.
- [20] M. Alamaniotis, T. Jevremovic, *IEEE Trans. Nucl. Sci.* 62 (3) (2015) 1262–1277.
- [21] M. Alamaniotis, L.H. Tsoukalas, *Critical Developments and Applications of Swarm Intelligence*, IGI Global, 2018, pp. 155–181.
- [22] M. Alamaniotis, J. Mattingly, L.H. Tsoukalas, *IEEE Trans. Nucl. Sci.* 60 (3) (2013) 2222–2231.
- [23] E. Yoshida, K. Shizuma, S. Endo, T. Oka, *Nucl. Instrum. Methods Phys. Res. A* 484 (1–3) (2002) 557–563.
- [24] H. Shahabinejad, N. Vosoughi, *Ann. Nucl. Energy* 132 (2019) 369–380.

- [25] J. Galford, J. Truax, A. Hrametz, C. Haramboure, et al., SPWLA 50th Annual Logging Symposium, Society of Petrophysicists and Well-Log Analysts, 2009.
- [26] R. Radtke, M. Lorente, B. Adolph, M. Berheide, S. Fricke, J. Grau, S. Herron, J. Horkowitz, B. Jorion, D. Madio, et al., SPWLA 53rd Annual Logging Symposium, Society of Petrophysicists and Well-Log Analysts, 2012.
- [27] R. Pemper, A. Pereira, G. Hou, D. Dolliver, J. Tudge, J. Kharrazi, H. Chok, G. Schmid, N. Mekic, T. Blankinship, et al., SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers, 2018.
- [28] R.C. Aster, B. Borchers, C.H. Thurber, *Parameter Estimation and Inverse Problems*, Elsevier, 2018.
- [29] I. Meric, G.A. Johansen, M.B. Holstad, J. Mattingly, R.P. Gardner, *Meas. Sci. Technol.* 23 (5) (2012) 055603, <http://dx.doi.org/10.1088/0957-0233/23/5/055603>.
- [30] H. Cardoso, S. de Barros Melo, R.M. de Abreu Silva, 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC, IEEE, 2018, pp. 4411–4416.
- [31] P.E. Gill, W. Murray, M.H. Wright, *Numerical Linear Algebra and Optimization*, in: *Numerical Linear Algebra and Optimization*, no. 1, Addison-Wesley Pub. Co., 1991, URL <http://books.google.com.br/books?id=ng3vAAAAAAAJ>.
- [32] D.S. Watkins, *Fundamentals of Matrix Computations*, Vol. 64, John Wiley & Sons, 2004.
- [33] G.R. Mateus, M.G. Resende, R.M. Silva, *J. Heuristics* 17 (5) (2011) 527–565, <http://dx.doi.org/10.1007/s10732-010-9144-0>.
- [34] T.A. Feo, M.G.C. Resende, *J. Global Optim.* 6 (2) (1995) 109–133, <http://dx.doi.org/10.1007/BF01096763>.
- [35] T.A. Feo, M.G.C. Resende, *Oper. Res. Lett.* 8 (2) (1989) 67–71, [http://dx.doi.org/10.1016/0167-6377\(89\)90002-3](http://dx.doi.org/10.1016/0167-6377(89)90002-3), URL <http://www.sciencedirect.com/science/article/pii/0167637789900023>.
- [36] M.G.C. Resende, C.C. Ribeiro, *Handbook of Metaheuristics*, Springer, 2010, pp. 283–319.
- [37] P. Festa, M.G.C. Resende, *Essays and Surveys in Metaheuristics*, Springer, 2002, pp. 325–367.
- [38] P. Festa, M.G.C. Resende, *Int. Trans. Oper. Res.* 16 (1) (2009) 1–24.
- [39] C. Oliveira, J. Salgado, F.G. Carvalho, *Encyclopedia of Analytical Chemistry*, American Cancer Society, 2006, <http://dx.doi.org/10.1002/9780470027318.a6212>, arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470027318.a6212>, URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470027318.a6212>.
- [40] R.G. Zamenhof, O.L. Deutsch, B.W. Murray, *Med. Phys.* 6 (3) (1979) 179–192, <http://dx.doi.org/10.1118/1.594631>, arXiv:<https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1118/1.594631>, URL <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.594631>.
- [41] M. Reginatto, *Radiat. Meas.* 45 (10) (2010) 1323–1329.
- [42] J. Mattingly, D.J. Mitchell, *IEEE Trans. Nucl. Sci.* 57 (6) (2010) 3734–3743.
- [43] H.-J. Im, K. Song, *Appl. Spectrosc. Rev.* 44 (4) (2009) 317–334.
- [44] I. Meric, G.A. Johansen, J. Mattingly, R.P. Gardner, *Radiat. Phys. Chem.* 95 (2014) 401–404.
- [45] N. Paciornikand, J.A. Ferland, R. Cl  roux, *Int. Trans. Oper. Res.* 10 (1) (2003) 77–87, <http://dx.doi.org/10.1111/1475-3995.00394>, arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-3995.00394>, URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-3995.00394>.
- [46] G.H. Golub, C.F.V. Loan, *Matrix Computations*, in: *Matrix Computations*, Johns Hopkins University Press, 2012, URL <https://books.google.com.br/books?id=5U-l8U3P-VUC>.
- [47] G.O. Toolbox, Matlab r2012a, 2012, Archived MathWorks Documentation.

Optimizing the Condition Number in Spectral Analysis

Halisson Cardoso

Centro de Informática

Universidade Federal de Pernambuco

Governo de Pernambuco

Recife, Pernambuco. Email: hacc@cin.ufpe.br

Silvio de Barros Melo

Centro de Informática

Universidade Federal de Pernambuco

Email: sbm@cin.ufpe.br

Ricardo Martins

de Abreu Silva

Centro de Informática

Universidade Federal de Pernambuco

Email: rmas@cin.ufpe.br

Abstract—It is well known that covariance matrices employed to estimate a solution to an overdetermined linear system, such as the ones based on the least squares, are usually poorly conditioned. This is particularly prevalent in the case of Prompt Gamma-ray Neutron Activation Analysis, where the data present additional challenges, such as noise introduced by the acquisition process, excessive magnitude differences among constituents, their non-linearity relation, external radiation interferences, and the necessity to provide an initial solution. The most effective framework proposed to overcome some of these difficulties is the Monte Carlo Library Least Squares, which produces iteratively and alternately the spectra and their most probable set of coefficients assumed to be the constituents weight fractions with respect to an input sample. This work focuses on the use of the metaheuristic GRASP to minimize the condition number of the covariance matrix used in the MCLLS. This modified GRASP was able to mostly improve the results from the most recent work that addressed this issue.

I. INTRODUCTION

The use of PGNA [7], [9], [4] [15] [13] and [11] for compositional analysis of bulk samples represents an inverse problem in the sense that the chemical and physical properties of the sample of interest would have to be determined based on the gamma-ray spectrum recorded by an appropriate spectrometer whereas there is a need for robust spectrum unfolding methods for the quantitative analysis. In this case, the inverse PGNA analysis is carried out by utilizing the Monte Carlo Library Least-Squares (MCLLS) [11] approach which consists of assuming a composition for the unknown object being investigated and generating library spectra using accurate forward Monte Carlo (MC) simulations. The application of the MCLLS approach for quantitative analysis shown to be more sensitive than the more conventional single peak analysis [5]. The MCLLS approach has the following advantages:

- Unresolved peaks are treated automatically.
- The spectral information contained in the entire spectrum is utilized, i.e., both the photopeaks and their Compton continua.
- Residuals may reveal whether or not there are missing libraries in the analysis.

Proper MCLLS implementation is a critical challenge in that the method represents an inverse problem in which a linear system to be solved presents poor conditioning, showing

sensitivity to small disturbances in the measurement data. Numerically unstable solutions to the least-squares problem may arise as an implication of the fact described above. The ill-conditioning in the MCLLS approach may come about as a result of the following: (1) Substantial degree of linear correlation between any two or more libraries; (2) Set of libraries may present values with too wide magnitude range (caused by background radiation [6]). Recently, two different approaches were proposed for the treatment of such ill-conditioned cases in the MCLLS approach with emphasis on analysis of multiphase oil/gas saltwater samples. One of these is based on an iterative method where library combinations that yield the lowest condition number of the covariance matrix at each iteration are accepted as a solution until all libraries are stripped away and their respective amounts determined [11]. The second method is based on introducing a densitometer set-up using a Cs-137 radioisotope with appropriate intensity in addition to the neutron source. The Cs-137 densitometer is used to determine mixture densities and through that, the amounts of oil and gas in a multiphase mixture [14]. Results have shown that both methods successfully suppressed the ill-conditioning in the MCLLS approach, leading to improved sensitivities. In this paper, an approach based on optimizing the partitioning of the library spectra is presented. This approach takes advantage of the fact that the dimension of the libraries vector space is much lower than the size of their spectra, allowing a selection of the appropriate sets of channels while keeping the linear independence of the working compounds. In the remainder of the paper, the method is described in details and applied to the problem of multiphase flow measurements by PGNA where the required libraries are generated through Monte Carlo (MC) simulations. The results indicate that this new approach can overcome the ill-conditioned of the inverse problem, leading to better results. The present proposal can be seen as a way to automate and improve the results from the least squares in the iterative process present in the MCLLS. The results calculated from the least squares for the percent fractions of the compounds, present in the iterative process of the MCLLS [6], can be improved from the proposal of the current work.

II. METHODOLOGY

A. power libraries

The determination of the weight fraction for a sample of compounds constituents subjected to gamma radiation is not an exact solution. The solution is approximate, having better quality if the number of covariance matrix conditioning is as small as possible.

The treatment to be given in this work for the library spectrum is presented in [10]. The steps used can be explained as follows, taking into account that the background radiation ([1]): (1) New libraries are formed by the individual junction of the spectral count of a compound and the sum of all other compound libraries. The library joining procedure is done for all possible combinations of compounds, except for the background library as a first compound; (2) Among these combinations mentioned in the previous step, it is checked which compound presents a covariance matrix with the smallest condition number. For this compound, it is calculated its fraction in the total spectrum; (3) The library of the compound described in the above step, after being multiplied by its fraction, gets removed from the full spectrum. The process then returns to the first step with the remaining library spectra.

The challenge set to this work is to select a subset of peaks through some heuristics to produce the best-conditioned covariance matrix possible.

B. GRASP

The heuristic used is the Adaptive Greedy Randomized Search Procedures (GRASP-[2],[3],[8]) to proceed with the peaks/frequencies selection to achieve restrictions mentioned above.

Algorithm 1 GRASP

```

1: input { $M, T, Max, Num, Obj$ }
2: output { $resulGrasp$ }
3: initialize the Algorithm list as empty
4:  $minCond \leftarrow \inf$ 
5:  $maxCond \leftarrow 0$ 
6:  $estInt \leftarrow intervals(M)$ 
7:  $estPro \leftarrow proportion(M)$ 
8:  $numCan \leftarrow numCanInt(M)$ 
9: while  $Obj < minCond$  or current iteration  $< Max$  do
10:   while the solution is not distinct for the others do
11:      $d \leftarrow$  new built solution
12:   end while
13:   if current iteration  $< T$  then
14:     insert  $d$  in the list
15:      $x \leftarrow$  condition number of  $d$ 
16:     if  $x < minCond$  then
17:        $resulGrasp \leftarrow d$ 
18:        $minCond \leftarrow x$ 
19:     end if
20:     if  $x > maxCond$  then
21:        $maxCond \leftarrow x$ 
22:     end if
23:   else
```

```

24:      $k \leftarrow$  search in the vicinity of  $d$ 
25:      $g \leftarrow$  condition number of  $k$ 
26:     if  $k < maxCond$  then
27:       replacing the list element most similar to  $k$  with
       higher condition number
28:     if  $k < minCond$  then
29:        $resulGrasp \leftarrow k$ 
30:     end if
31:   end if
32: end if
33: end while
```

In Algorithm 1, parameters M , T , Max , Num , Obj are, respectively, the matrix consisting of compounds as a function of the channels, the solutions list size, the maximum number of iterations, the number of channels to compose the new solution and the desired value for the condition number of the working matrix. The result is a data structure containing the selected channels, the covariance matrix constructed for the smallest condition number obtained during the execution.

On Lines 4 and 5, the algorithm initializes the variables which monitoring of fitness values obtained throughout the execution, and more specifically the lowest and highest fitness values. On lines 6, 7 and 8, respectively, the functions *intervals* (Algorithm 2), *proportion* (Algorithm 3), *numCanInt* (Algorithm 4) are called, returning the structures *estInt*, *estPro* and *numCan*. These outputs are used in the subsequent steps of GRASP. The routine *estInt* partitions the input data in intervals. A partition begins where the spectral count is zero for a compound and ends where the spectral count is zero for another compound. The routine *estPro* determines which fraction of the number of channels each interval is composed of. The routine *numCan* computes how many channels are used in each of the intervals determined by *estInt*, assuming *estPro* and the random number *num*, given as parameters in the call of GRASP.

On Line 9 up to line 12, a solution is built from the global solution space. From Line 13 to Line 23, this solution is assigned to a list, set initially to empty. It is checked whether this solution bears the highest or the lowest condition number during the current run.

On Line 24, the solution built on Line 11 is subject to a search in its vicinity. If the result of the local search is better than some of these solutions in the list, this result gets inserted, causing the removal of the solution that has greater fitness value and is most similar to the search result. On Lines 28-30, the algorithm's expected result is assigned to the variable *resulGrasp*.

The result of the algorithm (*resulGrasp*) is a data structure containing various pieces of information, (covariance matrix needed and a vector containing the selected channels to form the covariance matrix). The sub-routines *intervals*, *proportion*, *local search*, *path-relinking* which make up the Algorithm 1 explained later.

C. The sub-routine intervals

The intervals algorithm partitions the input data contained in the input data matrix.

Partitioning takes into account the channel in which a spectral count of a compound is zero, up to the previous channel where there is zero count for some other compound. Hence, then, a particular part is defined for the data.

The output of this algorithm will be a column matrix containing the structure of intervals in each of the positions. This data structure will be called *EstInt*.

D. The sub-routine proportion

The algorithm described below will serve to indicate the proportion of channels in each of the ranges. Remembering that the intervals were defined in the previous algorithm.

This algorithm will take into account the number of channels in each interval and the number of total channels. The proportion of channels in each of the ranges is determined by dividing the number of channels in the range by the number of total channels.

The output of this algorithm will be a column array containing the number ratio values of channels in each of the ranges. This array will be referred to as *estPro*.

E. The sub-routine numCanInt

The algorithm will determine the number of channels of each partition, given a subset of the total set of channels. The number of channels in this subset is passed as a parameter in the call of *GRASP*. This number of channels in each of the partitions will obey the same proportionality of the partitions of the original data.

The input data consists of the proportionality of the partitions of the original data and the number of channels of the subset.

The output will consist of a data structure containing the channel range of each of the partitions called *numCanInt*.

F. The sub-routine constructive

The Algorithm 2 construct the solutions from which *GRASP* begins. The input data are: a table containing the counts as a function of channels (*M*); a data structure containing the intervals of global data (*estInt*); the number of channels with respect to which the solutions are built (*Num*); the number of channels of each interval in a solution with *num* channels (*numCanInt*); the proportion of channels in each interval (*estPro*).

On Line 3 the data structure *channels* are initialized. On Line 4, a loop runs on the elements in *numCanInt*. On Line 5, the start and stop of each interval are recovered out of *numCanInt*, indexed by *i*. From Line 7 to Line 14 channel values from the respective partitions of the original data are drawn, to fill the intervals of the new solution. From Line 15 to Line 18 the following are built respectively: The Covariance matrix of the new solution built with *num* channels (line 15, *N*); the orthogonalization covariance matrix (line 16, *O*); Condition number of the covariance matrices *N* and *O*. Then

the assignment of *N*, *O*, *C*, *P* to a data structure *estConst* occurs on Line 19.

The covariance matrices which pass through the construction process need to have a difference in threshold relative to all other covariance matrices constructed by *GRASP*. Here, the criterion used to construct a solution sufficiently distinct from the solutions inserted in the *GRASP* list is that it has at least 90% of its channels different from the channels of the solutions inserted in the list, comparing these solutions two by two.

Algorithm 2 constructive

```

1: input {M, estInt, Num, numCanInt, estPro}
2: output estConst
3: channels  $\leftarrow$  size matrix M;
4: for each i in numCanInt do
5:   [start, stop]  $\leftarrow$  numCanInt{i}
6:   k  $\leftarrow$  start
7:   while k < stop do
8:     j  $\leftarrow$  estInt{i}
9:     l  $\leftarrow$  estInt{i+1}
10:    value  $\leftarrow$  random value between j and l
11:    if value  $\notin$  channels then
12:      insert value in channels
13:    end if
14:  end while
15: N  $\leftarrow$  matrix from M indexed by channels.
16: O  $\leftarrow$  orthogonalization of N
17: C  $\leftarrow$  condition number of N
18: P  $\leftarrow$  condition number of O
19: estConst  $\leftarrow$  [N, O, C, P]
20: end for
```

Once the solution is built, it is inserted in the list of candidate solutions. This insertion is automatic if the list is not filled. If the list is populated, a new stage of the algorithm is started. The local search algorithm is described in Algorithm 3.

G. The sub-routine Local Search

Algorithm 3 Local Search

```

1: input {estConst, T, Iter, estInt, M, numCanInt}
2: output {resulSearch}
3: estConst  $\leftarrow$  current solution
4: cond  $\leftarrow$  condition number of current solution
5: list  $\leftarrow$  instantiates the list with size T
6: posLi  $\leftarrow$  1
7: while list is not empty do
8:   count  $\leftarrow$  0
9:   k  $\leftarrow$  1
10:  while k  $\leq$  T and count < Iter do
11:    cop  $\leftarrow$  copies of the current solution channels
12:    while c  $\in$  channels do
13:      c  $\leftarrow$  Randomly choose a channel of M
14:    end while
15:    tamEstInt  $\leftarrow$  size of estInt
16:    q  $\leftarrow$  1
```

```

17:   while  $q < tamEstInt$  do
18:     if  $c > estInt\{q\}$  and  $c < estInt\{q+1\}$  then
19:       column  $\leftarrow q$ 
20:        $h \leftarrow$  random value between  $estInt\{q\}$  and
        $estInt\{q+1\}$ 
21:        $cop\{h, column\} \leftarrow c$ 
22:     end if
23:   end while
24:    $meEs \leftarrow$  submatrix  $M$  index by  $cop$ 
25:    $meCo \leftarrow$  covariance matrix of  $meEs$ 
26:    $conMe \leftarrow$  condition number of  $meCo$ 
27:   if  $conMe < cond$  then
28:     list{ $posLi$ }  $\leftarrow [meEs, meCo, conMe]$ 
29:      $posLi = posLi + 1$ 
30:   end if
31: end while
32: if the list is not empty then
33:   current solution  $\leftarrow$  select an element in the list
34:   restart the list
35:    $resulSearch \leftarrow$  current solution
36: end if
37: end while

```

The input data in Algorithm 3 are: $estConst$, built in the current interaction GRASP; T : size of list to be used in local search; $Iter$: maximum number of local search iterations to be applied; $estInt$: data structure that determines which channels are inserted in each of the intervals of the 2048 original channels; M : matrix containing the spectral counts for all compounds in 2048 channels; $numCanInt$: data structure containing channels in each of the intervals, given a subset of channels Num called in GRASP.

The initial value of the current solution of Algorithm 3 is the solution built by Algorithm 2. One of the requirements to end of the local search is that the list containing better solutions than the current solution is empty, indicating a local minimum. On line 13 a channel c is chosen from the global data that is not present in the current solution. This selected channel needs to be inserted into its appropriate partition in the new solution (on lines 19 to 21).

A random value is requested on Line 20 to determine what position the channel will be inserted. On Line 24 the desired matrix is constructed from the data provided in cop . On Line 25 the covariance matrix and its condition number are computed. If the condition number is less than $cond$, the constructed solution ($Mespec, numb$) is added to the list. On Line 32, if the list is not empty, a new value to the current solution is randomly chosen in the next iteration of the algorithm. Otherwise, the current solution is a local minimum, and the search is terminated. The neighborhood is defined by the difference of a channel relative to the current solution.

Once the system builds the matrix containing the channels that the solution must have, on Line 24 a new matrix $meEs$ (indexed by cop) is obtained from M . On Line 25, the covariance matrix of $meEs$ is calculated, followed by the determination of its condition number (Line 26). One data structure containing

the values of Lines 24, 25 and 26 is inserted into a position in the list (Line 27-30). From Line 32 to 36, it is checked whether the list is filled with some value after the search; it is randomly selected an element in the list to be the current solution (Line 33); and the list is reset (Line 34).

The output of the algorithm is a data structure $resulSearch$ that contains the matrix in which some channels are selected, the respective covariance matrix and condition number.

III. EXPERIMENTAL SETUP

In the configuration of the experiments, the spectral counting of five libraries (water, oil, salt, gas, background) was used [12].

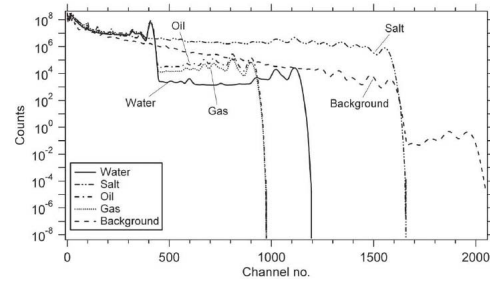


Fig. 1: The prompt gamma-ray library spectra of the oil, salt, water and gas phases.

The number of channels to be used as the sliding interval size was chosen to be 100, out of a total of 2048 units, since this is about the size of the support of a photopeak of any of the utilized constituents.

The technique used to construct the covariance matrix A is present in [10].

$$A_{i,j} = \sum_{i=1}^N \frac{1}{\sigma_i^2} f_k(x_i) f_l(x_i) \quad (1)$$

The combinations of libraries are described below:

instance	compound	sum
1	salt	oil + gas + water + background
2	oil	sal + gas + water + background
3	water	salt + oil + gas + background
4	gas	salt + oil + water + background
5	water	oil + gas + background
6	gas	water + oil + background
7	oil	gas+water+background
8	water	gas + background
9	gas	water + background

A. Results

Table 2 shows the results and their comparisons with [1].

Instance	goal	achieved	orthogonalized
1	3.09	3.08	2.91
2	21.82	29.11	1.12
3	50.62	64.91	1.14
4	34.58	30.57	3.79
5	47	14.07	1.35
6	35.1	11.99	1.16
7	21.6	10.90	1.0
8	18	13.55	1.42
9	12.5	8.30	3.17

Some instances reached the goal by just building a random solution of 100 channels (1, 4, 6 and 9). Other instances, in the first iteration of GRASP, can already get a good solution (5, 7 and 8). Instances 2 and 3 reached all possible iterations allowed for GRASP, but failed to achieve the goal despite the normalized covariance matrix corresponding value presenting a number of very low fitness. The orthogonalized covariance matrix can be used to solve the system of equations since the matrix used for their manufacture is used to return the coefficient of the orthogonalized base for QR factorization to the canonical base of the compounds. Instances 2, due to the failure to achieve the goal, will be studied in more detail. The drilldown can be extended to instance 3.

A detailing of built solutions (instances two and three) is shown in Figures 2a, 2b and 2c. Each built solution, calculate the conditioning of the covariance matrix, of the orthogonalized covariance matrix for QR factorization and the matrix to convert the covariance matrix in the canonical basis for orthogonalized covariance matrix for QR factorization. From the Figures 2a it is possible to notice the randomness in the construction of the solutions. It is important to mention that the covariance matrix condition number (Figures 2a) has its values abruptly reduced to be orthogonalized by QR factorization process (figure 2b)

The results of the search for a better response from the randomly built solution are described in Figure 3a for the first search described in GRASP. Graphs of local search are described for three different calls. Just as happened in the construction process of the solutions, the condition numbers of orthogonal covariance matrices for QR factorization (Figure 3b) and showed the best values over the condition number of the covariance matrices (Figure 3a)

The outputs of each of the iterations of GRASP, after the execution of all of its subroutines in each iteration, can be checked in figure 4, for instance two. The fitness values tend to converge. This convergence usually occurs well before it is reached the limit of 100 iterations. There is a certain threshold beyond which the condition number is not lowered no matter how many further iterations are done.

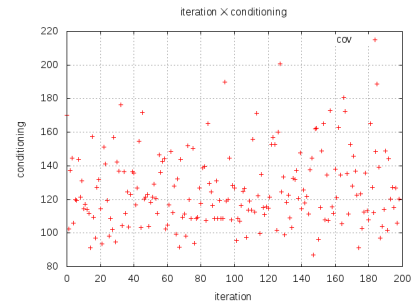
IV. CONCLUSIONS

In this study, covariance matrices with lower number of conditioning were presented than those in [11].

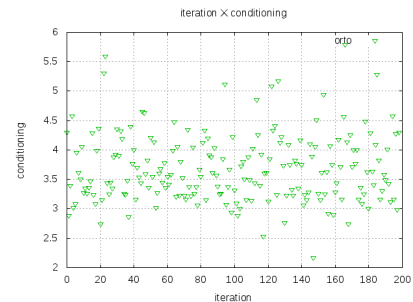
GRASP consists of several other subroutines, and it is possible to cite the *construct* and *local search* routines.

The results presented were successful in seven of the nine instances presented. In instances two and three, where it was not possible to reach the reference values, QR factorization methods demonstrated that it is possible to reach the desired number of conditioning values.

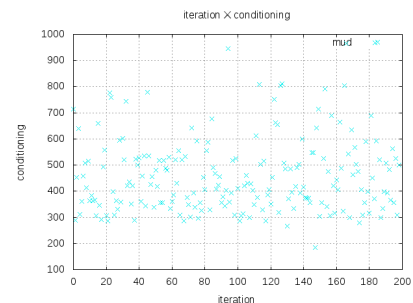
It is possible to note the strength of heuristics presented in an attempt to reduce the number of conditioning of the covariance matrix. Instances that have not had the goals achieved during the construction phase of GRASP managed to do that in the first or second iteration of the algorithm.



(a) Conditioning numbers presented by the covariance matrices during the construction process. Instance 2



(b) Conditioning of covariance matrices in (a), the Gram-Schmidt basis during the construction process. Instance 2

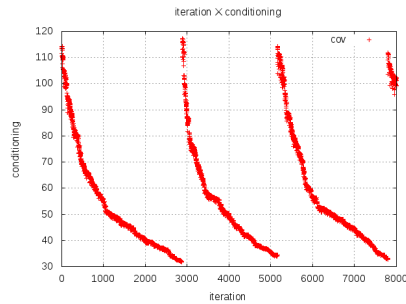


(c) Conditioning of the matrix that transforms the covariance matrix of the canonical basis for the Gram-Schmidt basis during the construction process. Instance 2

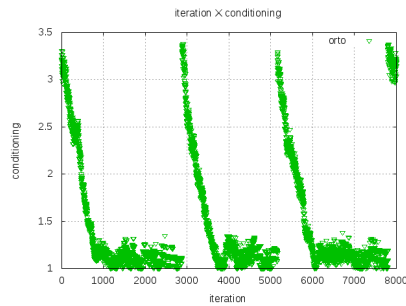
Fig. 2

It was revealed during the process of construction of the solutions of the instances 2 and 3: they are totally random, not having thus bias for building them. Remembering that in building the solutions channels of all intervals were used in the same proportion as the corresponding subpartition's size concerning the original partitioning. This initiative avoids arguments that the data were skewed in an attempt to achieve the results described in the paper.

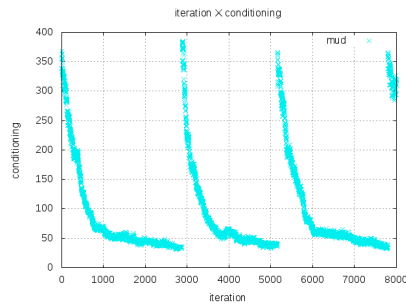
The local search process has proven to be one of the most effective in improving the solutions from the construction process. The decreasing conditioning order of magnitude of this algorithm step became the order of 10^2 .



(a) Conditioning numbers presented by the covariance matrices during the Local Search process. Instance 2



(b) Conditioning of covariance matrices in (a), the Gram-Schmidt basis during the Local Search process. Instance 2



(c) Conditioning of the matrix that transforms the covariance matrix of the canonical basis for the Gram-Schmidt basis during the Local Search process. Instance 2

Fig. 3

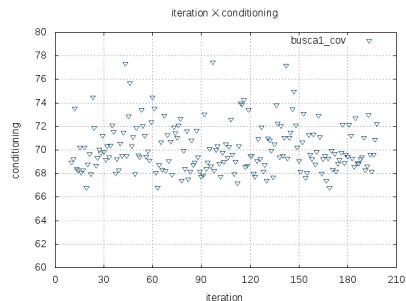


Fig. 4: Outputs for the first search of GRASP, instance 2

ACKNOWLEDGMENT

Special thanks: Government of The State of Pernambuco; Ilker Meric, teacher in Western Norway University, who graciously ceded research support data, and Severino Cardoso dos Santos (In memoriam).

REFERENCES

- [1] "On the ill-conditioning of the multiphase flow measurement by prompt gamma-ray neutron activation analysis," *Radiation Physics and Chemistry*, 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.radphyschem.2012.12.047>
- [2] T. A. Feo and M. G. Resende, "A probabilistic heuristic for a computationally difficult set covering problem," *Operations Research Letters*, vol. 8, no. 2, pp. 67–71, 1989. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/016737789900023>
- [3] T. Feo and M. Resende, "Greedy randomized adaptive search procedures," *Journal of Global Optimization*, vol. 6, no. 2, pp. 109–133, 1995. [Online]. Available: <http://dx.doi.org/10.1007/BF01096763>
- [4] R. Gardner, X. Han, P. Guo *et al.*, "On using monte carlo generated libraries for applying the library least-squares analysis approach to the c/o tool," in *SPWLA 47th Annual Logging Symposium*. Society of Petrophysicists and Well-Log Analysts, 2006.
- [5] R. P. Gardner, A. Sood, Y. Wang, L. Liu, P. Guo, and R. Gehrke, "Single peak versus library least-squares analysis methods for the pgnaa analysis of vitrified waste," *Applied radiation and isotopes*, vol. 48, no. 10, pp. 1331–1335, 1997.
- [6] R. P. Gardner and L. Xu, "Status of the monte carlo library least-squares (mclls) approach for non-linear radiation analyzer problems," *Radiation Physics and Chemistry*, vol. 78, no. 10, pp. 843–851, 2009.
- [7] H.-J. Im and K. Song, "Applications of prompt gamma ray neutron activation analysis: detection of illicit materials," *Applied Spectroscopy Reviews*, vol. 44, no. 4, pp. 317–334, 2009.
- [8] G. R. Mateus, M. G. Resende, and R. M. Silva, "Grasp with path-relinking for the generalized quadratic assignment problem," *Journal of Heuristics*, vol. 17, no. 5, pp. 527–565, Oct. 2011. [Online]. Available: <http://dx.doi.org/10.1007/s10732-010-9144-0>
- [9] J. Mattingly and D. J. Mitchell, "A framework for the solution of inverse radiation transport problems," *Nuclear Science, IEEE Transactions on*, vol. 57, no. 6, pp. 3734–3743, 2010.
- [10] I. Meric, "Nuclear methods for subsea fluid characterization," Ph.D. dissertation, 2012.
- [11] I. Meric, G. A. Johansen, M. B. Holstad, J. Mattingly, and R. P. Gardner, "On the treatment of ill-conditioned cases in the monte carlo library least-squares approach for inverse radiation analyzers," *Measurement Science and Technology*, vol. 23, no. 5, p. 055603, 2012.
- [12] —, "On the treatment of ill-conditioned cases in the monte carlo library least-squares approach for inverse radiation analyzers," *Measurement Science and Technology*, vol. 23, no. 5, p. 055603, 2012. [Online]. Available: <http://stacks.iop.org/0957-0233/23/i=5/a=055603>
- [13] I. Meric, G. A. Johansen, M. B. Holstad, J. Wang, and R. P. Gardner, "Produced water characterization by prompt gamma-ray neutron activation analysis," *Measurement Science and Technology*, vol. 22, no. 12, p. 125701, 2011.
- [14] I. Meric, G. A. Johansen, J. Mattingly, and R. Gardner, "On the ill-conditioning of the multiphase flow measurement by prompt gamma-ray neutron activation analysis," *Radiation Physics and Chemistry*, vol. 95, pp. 401–404, 2014.
- [15] M. Reginatto, "Overview of spectral unfolding techniques and uncertainty estimation," *Radiation Measurements*, vol. 45, no. 10, pp. 1323–1329, 2010.

K.2 ARTIGOS EM PRODUÇÃO

Determination and Prediction of Data Partitioning Compounds in the Gamma Ray Study.

Halisson Alberdan Cavalcanti Cardoso^{a,*}, Sidartha Azevedo Lobo de Carvalho^{a,b}, Silvio de Barros Melo^a, Ricardo Martins de Abreu Silva^a, Silas Garrido Teixeira de Carvalho Santos^a and Carlos Costa Dantas^c

^aCentro de Informática Campus Recife - Universidade Federal de Pernambuco, Av. Jornalista Aníbel Fernandes - s/n - Cidade Universitária - PE - Brazil - ZIP Code: 50.740 - 560.

^bCampus Quixadá - Universidade Federal do Ceará - Av. José de Freitas Queiroz - 5003 - Cedro - Quixadá - Ceará - Brazil. Zip Code: 63.902-580.

^cDepartamento de Energia Nuclear Campus Recife - Universidade Federal de Pernambuco - Av. Prof. Luiz Freire - 1000 - Cidade Universitária - PE - Brazil: ZIP Code: 50.740-545.

ARTICLE INFO

Keywords:

Prompt Gamma Neutron Analysis Activation (PGNAA), Condition Number, Monte Carlo Library Least Square (MCLS), Greedy Randomized Adaptive Search Procedure (GRASP), Least Square, Singular Value Decomposition (SVD)

ABSTRACT

Context: Prompt Gamma-Ray Neutron Activation Analysis is widely used for the elementary analysis of small samples, providing a fast and non-destructive medium of the material under study. **GAP:** The linear combination of the spectra of the constituents of this sample defines the spectrum of an arbitrary sample. The application of several techniques to determine the spectra inserted in an example are reported, such as the use of genetic algorithms or the use of optimization by Particle Swarm Optimization (PSO). Many of these techniques fall into the problem of least-squares system over determination. **Purpose:** The present work describes the least-squares approximation in PGNAA through data partitioning. The data will be partitioned based on the peaks of the libraries that make up the total spectrum of the unknown sample. It will also be taken into account the modeling of an unknown compound that has participated in the full spectrum of the unknown sample through constant functions in the different partitions. **Methodology:** The present proposal consists of: (1) partitioning the data by locating the peaks of the libraries that define the full spectrum and which will assist in determining the location of the largest residue and the reference partitions containing the peaks of known compounds; (2) perform the basic steps of GRASP (solution building, searching in the vicinity of a current solution and linking two solutions) in the quantitative study of data partitioning and approximation of a suppressed compound of the full spectrum. **Results:** The results suggest a close to 100% determination in the quantitative determination of compounds by partitioning the data, and an excellent approximation (put a numerical value here) in the prediction of the whole spectrum suppressed compound. **Conclusions:** It is therefore suggested that data partitioning be an alternative to the methods hitherto presented that use dimension two covariance matrices, in addition to the determination of compounds that are present in the full spectrum and are not inserted therein.

1. Introduction

Identification and application of isotopes by analyzing gamma ray spectra recorded in a Multi-Channel Analyzer (MCA) has numerous applications, including: (1) measurements in marine environments for measuring salt content in crude oil [4] [43] [10]; (2) measurement of natural radioactivity [6] [7] [19]; (3) elemental analysis using Prompt Gamma Neutron Analysis (PGNAA) [30]; (4) industry [38]; (5) therapy [25]; (6) security [42]; (7) radiography [26] and (8) biophysics [27].

Gamma radiation analysis is a non-destructive technique widely used in the industry. In solving inverse gamma radiation problems, the Library Least Squares [39] method is often used to provide rapid solutions.

The photopeaks inserts in the gamma-ray spectra are superimposed on the background radiation, Compton Continua, and other parts related to different physical processes. The

most intense photopeaks, after subtracting the background radiation counts, determine the type of element inserted into the full spectrum in the gamma-ray analysis.

A major challenge in performing peak analysis lies in the energy restricted windows that have a limited amount of information used. It is also possible to enumerate the difficulties in determining radioisotopes that have low counting rates, as well as situations where peak overlap occurs due to the low resolution of detectors.

Alternatively, to peak analysis, quantitative analysis in PGNAA by the Monte Carlo Library Least Squares (MCLS) [17] approximation appears. Quantitative PGNAA analysis by MCLS has an advantage over conventional peak analysis in that it makes use of the full gamma-ray spectrum. Significant improvements were obtained in the implementation of the Whole Spectrum Analysis (WSA) or Full Spectrum Analysis (FSA) methods [39] [16] [22].

The central assumption of MCLS is that the total gamma-ray spectrum is defined as a linear combination of the constituent contributions individually. The MCLS method lies in finding the best fit for the unknown gamma-ray spectrum

*Corresponding author

✉ halissonalberdan@gmail.com (H.A.C. Cardoso);

silobocarvalho@gmail.com (S.A.L.d. Carvalho); sbm@cin.ufpe.br (S.d.B.

Melo); rmas@cin.ufpe.br (R.M.d.A. Silva); silas@silasgarrido.com

(S.G.T.d.C. Santos); ccd@ufpe.br (C.C. Dantas)

ORCID(s): 0000-0002-1166-481X (H.A.C. Cardoso)