

Denis Lage Ferreira da Silva

**Influência dos métodos de otimização e da
borda de absorção na inversão do campo de
onda acústico no domínio da frequência**

Niterói

2017

Denis Lage Ferreira da Silva

Influência dos métodos de otimização e da borda de absorção na inversão do campo de onda acústico no domínio da frequência

Tese apresentada ao Programa de Pós-graduação em Geologia e Geofísica Marinha da Universidade Federal Fluminense, como requisito parcial para a obtenção do título de Doutor em Geologia e Geofísica. Área de Concentração: Geofísica Aplicada

Universidade Federal Fluminense - UFF

Instituto de Geociências

Departamento de Geologia e Geofísica

Programa de Pós-graduação em Geologia e Geofísica Marinha

Orientador: Marco Antonio Cetale Santos

Niterói

2017

Denis Lage Ferreira da Silva

Influência dos métodos de otimização e da borda de absorção na inversão do campo de onda acústico no domínio da frequência/ Denis Lage Ferreira da Silva. – Niterói, 2017-

107 p. : il. (algumas color.) ; 30 cm.

Orientador: Marco Antonio Cetale Santos

Tese (Doutorado) – Universidade Federal Fluminense - UFF

Instituto de Geociências

Departamento de Geologia e Geofísica

Programa de Pós-graduação em Geologia e Geofísica Marinha, 2017.

1. FWI. 2. Modelagem acústica. 2. Otimização não linear. I. Marco Antonio Cetale Santos. II. UFF. III. LAGEMAR. IV. Influência dos métodos de otimização e da borda de absorção na inversão do campo de onda acústico no domínio da frequência

Denis Lage Ferreira da Silva

Influência dos métodos de otimização e da borda de absorção na inversão do campo de onda acústico no domínio da frequência

Tese apresentada ao Programa de Pós-graduação em Geologia e Geofísica Marinha da Universidade Federal Fluminense, como requisito parcial para a obtenção do título de Doutor em Geologia e Geofísica. Área de Concentração: Geofísica Aplicada

Trabalho aprovado. Niterói, 2017:

Marco Antonio Cetale Santos
Orientador

Prof. Djalma Manoel Soares Filho
Convidado 1

Prof. Jorge Leonardo Martins
Convidado 2

Prof. Rodrigo Bird Burgos
Convidado 3

Prof. Roger Matsumoto Moreira
Convidado 4

Niterói
2017

*Este trabalho é dedicado à minha mãe Rosane.
Você estará sempre comigo em meu coração.*

Agradecimentos

Gostaria de agradecer a todos que contribuíram para este trabalho.

A minha esposa, Marina, que esteve ao meu lado em todos os momentos, me apoiando e me incentivando. Obrigado, meu amor!

Ao meu pai Dario, meus irmãos Diogo e Daniel e minha vó Wanda.

A minha segunda família, Zé, Julieta, Fernanda, Rafael e Guga.

Ao meu orientador Marco Cetale, por todas as conversas e discussões que tornaram este trabalho possível.

Aos professores Roger, Jorge Leonardo, por toda a ajuda durante este trabalho.

Aos membros da banca, que com todas as sugestões e críticas, permitiram que este trabalho se tornasse ainda melhor.

A todos os amigos da UFF, em especial aos amigos do GISIS, Victor Martins, Felipe, Marcelle, Jaqueline e Victor Almeida

A todos os amigos da CGG.

A CAPES, pelo auxílio financeiro.

Resumo

A obtenção de uma boa estimativa das propriedades elásticas através do desenvolvimento de novas estratégias de inversão do campo total é de fundamental importância para a caracterização de reservatórios. A inversão do campo de onda completo (*Full Waveform Inversion*, FWI) é um procedimento complexo de ajuste de dados baseado na modelagem de onda completa visando extrair informações quantitativas de sismogramas. Isto é feito encontrando uma direção de busca no subespaço dos modelos e em seguida atualizando o modelo de forma a reduzir o resíduo dos sismogramas. Existem na literatura diversos modos de encontrar a direção de busca da FWI. Esta tese implementa a inversão do campo de onda acústica utilizando 3 técnicas, gradiente descendente, gradiente conjugado e L-BFGS. No processo de inversão, a modelagem sísmica é resolvida diversas vezes, tornando-a uma etapa fundamental para o desenvolvimento deste projeto, pois implica em custo computacional, como também na qualidade dos dados que serão usados durante a inversão. A modelagem sísmica foi implementada com a técnica de diferenças finitas de forma implícita e no domínio da frequência. Nesta modelagem novas técnicas de absorção em contornos artificiais foram desenvolvidas para reduzir a dispersão numérica e melhorar a absorção das frentes de onda. Esta tese apresenta a influência da borda de absorção desenvolvida no resultado da inversão. Além disso, uma comparação entre os diversos métodos de otimização é apresentada, analisando os efeitos de cada um deles no resultado da FWI.

Palavras-chave: FWI. Modelagem Acústica. Inversão Não-linear. Gradiente Descendente. Gradiente Conjugado. L-BFGS. Borda de absorção.

Abstract

A good estimation of the elastic properties through full waveform inversion (FWI) is essential for a proper reservoir characterization. FWI is a complex data adjustment flow based on the full wavefield modeling, with the objective of extracting quantitative information from seismograms. This is done by determining a search direction in the models subspace and after that updating the model so that the error between the seismograms is reduced. There are several ways of finding such direction. This thesis implements the acoustic FWI using 3 techniques, steepest descent, conjugate gradient and L-BFGS. During the inversion, the seismic modeling is solved multiple times, making it one of the key steps for the development of this thesis. The seismic modeling is solved using a frequency domain finite differences algorithm. For this modeling algorithm, we developed a new absorbing boundary technique that reduces the numerical dispersion and improves the wavefront absorption. This thesis present the influence of the new absorbing boundary in the result of the FWI. Furthermore, we present a comparison of the different optimization techniques in FWI, analysing its effect of the FWI result.

Keywords: FWI. Acoustic Modeling. Nonlinear Inversion. Steepest Descent. Conjugate Gradient. L-BFGS. Absorbing Boundary.

Lista de ilustrações

Figura 1 – Resultado para os métodos de otimização. Função Quadrática	31
Figura 2 – Resultado para os métodos de otimização. Função de Rosenbrock	33
Figura 3 – Esquema de inversão	36
Figura 4 – Modelo discretizado	37
Figura 5 – Modelo de duas camadas	39
Figura 6 – Modelo de Marmousi	40
Figura 7 – Matriz B	40
Figura 8 – Comparação bordas	42
Figura 9 – Comparação bordas	42
Figura 10 – Comparação bordas	43
Figura 11 – Comparação bordas	44
Figura 12 – Modelo Marmousi	46
Figura 13 – Fonte Ricker 20 Hz	46
Figura 14 – Exemplo de Sismograma.	47
Figura 15 – Esquema de inversão	49
Figura 16 – Modelo Anomalia	52
Figura 17 – Gradiente Anomalia	53
Figura 18 – Resultado da inversão para modelo de 4 camadas. Modelo inicial suavizado	64
Figura 19 – Resultado da inversão para modelo de 4 camadas. Modelo inicial homo- gêneo	65
Figura 20 – Evolução do modelo invertido para um fonte <i>Ricker</i> de frequência central de 20 Hz.	66
Figura 21 – Resíduo do modelo invertido para um fonte <i>Ricker</i> de frequência central de 20 Hz	67
Figura 22 – Comparação da FWI para diferentes frequências centrais da fonte sísmica	68
Figura 23 – Comparação da FWI para um número diferente de fontes sísmicas	69
Figura 24 – Comparação da FWI para diferentes modelos iniciais	70
Figura 25 – Resultado da inversão para as diferentes bordas de absorção.	75
Figura 26 – Espectro de frequência - Ruído Rosa	77
Figura 27 – Progressão de Erro - Exemplo A	78
Figura 28 – Total de Iterações - Exemplo A	78
Figura 29 – Total de Modelagens - Exemplo A	79
Figura 30 – Resultado da Inversão - Exemplo A	79
Figura 31 – Sismogramas - Exemplo A	80
Figura 32 – Progressão de Erro - Exemplo B	82
Figura 33 – Total de iterações - Exemplo B	82

Figura 34 – Total de modelagens - Exemplo B	83
Figura 35 – Resultado da Inversão - Exemplo B	83
Figura 36 – Sismogramas - Exemplo B	84
Figura 37 – Progressão de erro - Exemplo C	85
Figura 38 – Total de iterações - Exemplo C	86
Figura 39 – Total de modelagens - Exemplo C	86
Figura 40 – Resultado da Inversão - Exemplo C	87
Figura 41 – Sismogramas - Exemplo C	87
Figura 42 – Progressão de erro - Exemplo D	88
Figura 43 – Total de iterações - Exemplo D	89
Figura 44 – Total de modelagens - Exemplo D	89
Figura 45 – Resultado da Inversão - Exemplo D	90
Figura 46 – Sismogramas - Exemplo D	90
Figura 47 – Modelo discretizado	104

Lista de tabelas

Tabela 1 – Resultado função quadrática.	31
Tabela 2 – Resultado Rosenbrock	32
Tabela 3 – Resultado das técnicas de absorção	43
Tabela 4 – Parâmetros da Modelagem	45
Tabela 5 – Total de termos de frequência simulados	45
Tabela 6 – Parâmetros da Modelagem	64
Tabela 7 – Parâmetros de Simulação - Caso Exemplo 1	74
Tabela 8 – Parâmetros de Simulação - Caso exemplo 2	76
Tabela 9 – Resultados Inversão - Exemplo A	80
Tabela 10 – Resultados Inversão - Exemplo B	84
Tabela 11 – Resultados Inversão - Exemplo C	85
Tabela 12 – Resultados Inversão - Exemplo D	91

Lista de símbolos

χ	Função objetivo
\mathbf{m}	Modelo a ser otimizado
α	Passo de iteração
x, z	Coordenada horizontal, vertical
$\Delta x, \Delta z$	Amostragem horizontal, vertical
\mathbf{u}	Campo de pressão
$\nabla\chi$	Gradiente da função objetivo
\mathbf{H}	Matriz Hessiana
\mathbf{G}	Inversa da matriz Hessiana
\mathbf{s}	Fonte sísmica
\mathbf{c}	Velocidade de onda P
\mathbf{g}	Direção descendente

Sumário

1	INTRODUÇÃO	21
2	TÉCNICAS DE OTIMIZAÇÃO NÃO LINEAR	25
2.1	Otimização Não Linear	25
2.2	Métodos de cálculo da direção descendente	26
2.2.1	Gradiente Descendente	26
2.2.2	Newton	26
2.2.3	Quasinevton	27
2.2.4	Gradiente Conjugado	28
2.3	Busca do passo de iteração ótimo	29
2.4	Exemplo	29
2.5	Conclusão	34
3	MODELAGEM	35
3.1	Modelagem Acústica	35
3.2	Borda de absorção	38
3.3	Exemplo	44
3.4	Conclusão	45
4	INVERSÃO DO CAMPO DE ONDA ACÚSTICO	49
4.1	Gradiente da função objetivo	50
4.2	Método Adjunto	54
4.3	Regularização e pré-condicionamento	56
4.4	Métodos de otimização matemática em FWI	57
4.4.1	Gradiente Descendente	58
4.4.2	Gradiente Conjugado	58
4.4.3	Quasinevton	58
4.5	Critérios de Minimização	60
4.6	Métodos de linearização	60
4.7	Modelo Inicial	61
4.8	Estimativa da Fonte Sísmica	62
4.9	Exemplos	63
4.9.1	Exemplo 1 - Modelo camadas plano paralelas	63
4.9.2	Exemplo 2 - Modelo Marmousi	65
4.9.2.1	Frequência da fonte	68
4.9.2.2	Número de tiros	69

4.9.2.3	Modelo Inicial	70
4.10	Conclusão	71
5	RESULTADOS	73
5.1	Metodologia	73
5.2	Caso Exemplo 1 - Borda de absorção	73
5.2.1	Metodologia	73
5.2.2	Resultado	74
5.3	Caso exemplo 2 - Comparação dos métodos de otimização matemática	75
5.3.1	Metodologia	75
5.3.2	Resultados	76
6	DISCUSSÃO E CONCLUSÕES	93
6.1	Trabalhos futuros	93
	REFERÊNCIAS	95

APÊNDICES 101

APÊNDICE A – MODELAGEM ACÚSTICA NO DOMÍNIO DA FREQUÊNCIA 103

A.1	Interior do Modelo	103
A.2	Borda PML	104
A.3	Borda Superior	105
A.4	Borda Inferior	105
A.5	Borda Esquerda	105
A.6	Borda Direita	106
A.7	Canto Superior Esquerdo	106
A.8	Canto Superior Direito	106
A.9	Canto Inferior Esquerdo	106
A.10	Canto Inferior Direito	107

1 Introdução

Na produção de petróleo e gás, a simulação do fluxo de escoamento dos fluidos (óleo e água) nos reservatórios é de suma importância para estimar e otimizar a produção. Porém, para essas simulações, são necessárias propriedades petrofísicas e geométricas dos reservatórios. Esse conjunto de informações caracteriza o reservatório, e melhorar a estimativa dessas propriedades significa melhorar a produção de um campo petrolífero, agregando um alto valor econômico e justificando todos os investimentos nesta área.

Algumas propriedades importantes da petrofísica, tais como porosidade e saturação de fluidos estão relacionadas com as propriedades elásticas, que podem ser estimadas através da sísmica. Desta forma, obter uma boa estimativa das propriedades elásticas através do desenvolvimento de novas estratégias de inversão do campo de ondas total é de fundamental importância para a caracterização dos reservatórios. A inversão do campo de onda completo (*Full Waveform Inversion*, FWI) é um procedimento complexo de ajuste de dados baseado na modelagem de onda completa visando extrair informações quantitativas de sismogramas.

[Tarantola \(1984\)](#) inferiu modelos de subsuperfície de alta resolução utilizando técnicas de minimização do resíduo dos sismogramas observados e simulados. Esta técnica consiste em encontrar uma direção de busca no subespaço dos modelos a partir de um modelo inicial. Em seguida é realizada uma segunda busca na direção calculada por um modelo que reduza o resíduo dos sismogramas. Este procedimento é repetido até que haja convergência. Existem diversos modos de encontrar a direção de busca da FWI ([FLETCHER, 1980](#)), cada um com vantagens e desvantagens. Nos últimos anos, pesquisadores vem buscando formas de implementar a inversão do campo de onda completo de maneira robusta e eficiente computacionalmente.

Inicialmente, a FWI foi implementada realizando a busca através da direção contrária ao gradiente da função objetivo (método do gradiente descendente) ([TARANTOLA, 1984](#)). Esta técnica, embora de fácil implementação, necessita de um número elevado de iterações para atingir a convergência. [Hu et al. \(2011\)](#) utiliza o método do gradiente conjugado (MGC) apresentado por [Hestenes e Stiefel \(1952\)](#). A facilidade de implementação para problemas com diversos parâmetros e a convergência mais rápida que o gradiente descendente torna o MGC bastante popular como otimizador na FWI. Finalmente, uma aproximação de segunda ordem da função objetivo dá origem à otimização de Newton. Desta forma, a matriz das derivadas de segunda ordem da função objetivo (matriz Hessiana) é utilizada no cálculo da direção de busca da solução. No caso da FWI, onde o número de parâmetros é muito grande (>10000), esta matriz não pode ser calculada de forma

eficiente, inviabilizando o método. Contudo existem formas de encontrar a Hessiana de forma aproximada (DENNIS; MORÉE, 1977). Nocedal e Wright (2006) apresentam formas implícitas de calcular as atualizações do modelo sem a necessidade de armazenar a matriz Hessiana, viabilizando a aplicação na FWI. Alguns exemplos de aplicações de métodos baseados na técnica de Newton estão em Pratt, Shin e Hicks (1998), Brossier, Operto e Virieux (2009b), Brossier, Operto e Virieux (2009a), Métivier et al. (2013).

A FWI pode ser definido como um problema de minimização não linear mal-posto. Isto significa que existem diversas soluções para o mesmo problema. Além do método de otimização numérica utilizado, outros desafios devem ser atacados visando encontrar uma solução geologicamente plausível. Contribuem neste sentido a inversão sequencial de frequências específicas no domínio da frequência, ou de comprimentos de ondas específicos no domínio do tempo. Técnicas de regularização e pré-condicionamento dos dados de entrada também ajudam em tornar o problema mais robusto. Virieux e Operto (2009) apresenta um resumo completo de todas as etapas do processo de FWI.

No processo de inversão, a modelagem sísmica é resolvida diversas vezes, tornando-a uma etapa fundamental para o desenvolvimento deste projeto, pois implica em custo computacional, como também na qualidade dos dados que serão usados durante a inversão. A modelagem sísmica, uma das etapas principais dos métodos de inversão usados nesta tese, será implementada com a técnica de diferenças finitas de forma implícita e no domínio da frequência. !!!De modo a tratar o efeito da propagação de onda em um meio finito, precisamos implemenetar técnicas de absorção das frentes de ondas em contornos artificiais. Encontrar uma técnica de absorção que reduza significativamente o campo de onda refletido pelas bordas sem aumentar demasiadamente o custo computacional da modelagem é de suma importância para a FWI. Nesta tese foi desenvolvido uma nova técnica de absorção em contornos artificias capaz de melhoras a absorção das frentes de ondas com um relativamente baixo custo computacional. Esta técnica foi publicada pelos autores Roger M. Moreira, Marco Antonio Cetale Santos, Jorge Leonardo Martins, Denis Lage F da Silva, Raul B. V. Pessolani, Djalma Manuel Soares Filho e André Bucão, no periódico *Geophysics* (MOREIRA et al., 2014). !!!

Esta tese apresenta uma comparação entre métodos de otimização matemática utilizados na FWI. Para estas comparações, levou-se em consideração o número de iterações necessária para a convergência e o valor total da função objetivo encontrada para cada método. Os modelos de velocidades invertidos a partir de cada método serão comparados buscando encontrar o método de otimização mais robusto para FWI. Devido a sensibilidade da FWI, a qualidade do algoritmo da modelagem sísmica exerce grande influência no resultado final da inversão. A borda de absorção desenvolvida nesta tese contribui para um resultado melhor da inversão.

O capítulo 2 apresenta técnicas de otimização não linear para uma função objetivo

genérica. Estas técnicas são a base para a FWI. No capítulo 3, é apresentado o algoritmo de modelagem acústica utilizado nesta tese. A nova técnica de absorção utilizada pelo algoritmo de modelagem é apresentada neste capítulo. O capítulo 4 apresenta o algoritmo de FWI propriamente dito. A função objetivo é definida e o cálculo do gradiente da mesma é apresentado. Os métodos de otimização matemática apresentados no capítulo 2 são implementados no algoritmo de FWI. No capítulo 5 são apresentados alguns exemplos. Primeiro comparamos algumas bordas de absorção diferentes no algoritmo de FWI, demonstrando o benefício da borda de absorção desenvolvida durante esta tese. Em seguida comparamos os diversos métodos de otimização matemática. As características do problema, como a qualidade do modelo inicial e a qualidade do dado medido, exercem grande influência no resultado final de cada método. Finalmente, no capítulo 6, são apresentadas as principais conclusões obtidas durante esta tese e as sugestões de trabalhos futuros.

2 Técnicas de otimização não linear

Conforme foi dito na introdução, a inversão do campo de onda completo (FWI) pode ser definido como um processo de minimização não linear. Devido à complexidade dessas funções, em geral não podemos encontrar o mínimo da mesma de forma determinística. Para isto faz-se uma busca iterativa procurando, a cada passo, um conjunto de parâmetros que diminua o valor da função. Este capítulo descreve o processo matemático de otimização iterativa, apresentando algumas das formas de se buscar o caminho até o mínimo da função que se deseja otimizar, chamada de função objetivo.

2.1 Otimização Não Linear

Um minimizador iterativo consiste em buscar uma atualização para os parâmetros de uma função objetivo de modo a reduzir seu valor, tal qual descrito na Equação 2.1,

$$\chi(\mathbf{m}^{(k+1)}) < \chi(\mathbf{m}^{(k)}). \quad (2.1)$$

Para cada iteração k , encontra-se um novo conjunto de parâmetros $\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \boldsymbol{\delta}$ tal que a função objetivo $\chi(\mathbf{m}^{(k+1)})$ seja menor que na iteração anterior. $\boldsymbol{\delta}$ representa a variação nos parâmetros do modelo e pode ser definido como

$$\boldsymbol{\delta} = \alpha \mathbf{g}, \quad (2.2)$$

onde \mathbf{g} é o vetor que define a direção de atualização do modelo, ou direção de busca e α é o tamanho deste vetor, de forma a garantir que a condição definida pela Equação 2.1 seja verdadeira.

Naturalmente, para um determinado ponto da função objetivo, existem diversas direções de busca possíveis. Deseja-se encontrar um conjunto de direções capazes de atingir o mínimo da função o mais rápido possível. A velocidade do algoritmo pode ser definida em termos de número de iterações necessárias para atingir o mínimo, chamada de velocidade de convergência e custo computacional para o cálculo de cada iteração. Existem na literatura diversas formas de calcular esta direção de busca. Cada método possui características próprias envolvendo custo computacional e velocidade de convergência. Se, por exemplo, um determinado método for capaz de atingir o mínimo da função em um número muito pequeno de iterações, mas com um custo computacional necessário para calcular cada direção de busca muito elevado, o tempo total do algoritmo pode se tornar maior que um outro método que possua uma convergência mais lenta mas de menor custo computacional.

Em muitos casos os problemas estudados possuem uma complexidade tal que alguns métodos tornam-se computacionalmente inviáveis.

Todos os algoritmos aqui estudados utilizam o gradiente da função objetivo para calcular a direção de busca. O gradiente $\nabla\chi$ de uma função em um ponto indica a direção de maior decréscimo naquele ponto. Para funções conhecidas, o cálculo do gradiente pode ser trivial. Já no caso de funções mais complexas, como na FWI, este cálculo pode ser feito de forma aproximada e também pode representar um alto custo computacional. A seção seguinte apresenta alguns destes métodos de otimização.

2.2 Métodos de cálculo da direção descendente

Neste trabalho foram estudados os métodos do gradiente descendente, gradiente conjugado, Newton e Quasi-Newton BFGS.

2.2.1 Gradiente Descendente

No método do gradiente descendente, a direção de busca é o negativo do gradiente no ponto de estudo, conforme mostrado na Equação 2.3,

$$\mathbf{g}^{(k)} = -\nabla\chi^{(k)}, \quad (2.3a)$$

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha\mathbf{g}^{(k)}. \quad (2.3b)$$

A principal vantagem deste método é o baixo custo computacional necessário para calcular \mathbf{g} em cada iteração é baixo. Outra vantagem importante é que o negativo do gradiente sempre indica uma direção descendente de uma função, exceto nos pontos críticos ($\nabla\chi = \mathbf{0}$). A principal desvantagem do método gradiente descendente é o número de iterações necessárias para atingir a convergência. Em geral, para funções complexas, a convergência torna-se tão lenta que faz com o método seja muito pouco eficiente.

2.2.2 Newton

De forma a aumentar a convergência, pode-se utilizar uma aproximação de segunda ordem da série de Taylor da função objetivo, representada na Equação 2.4. Esta aproximação introduz o conceito da matriz Hessiana \mathbf{H} , definida como a segunda derivada da função objetivo, $H_{i,j} = \frac{\partial^2\chi}{\partial m_i \partial m_j}$,

$$\chi(\mathbf{m}^{(k)} + \boldsymbol{\delta}) \approx \chi^{(k)} + \nabla\chi^{(k)T} \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^T \mathbf{H}^{(k)} \boldsymbol{\delta}. \quad (2.4)$$

Deseja-se encontrar $\boldsymbol{\delta}$ de forma que $\chi(\mathbf{m}^{(k)} + \boldsymbol{\delta})$ seja mínimo. O desenvolvimento do minimizador de segunda ordem encontra-se nas Equações 2.5A,F,

$$\min_{\boldsymbol{\delta}} [\chi(\mathbf{m}^{(k)} + \boldsymbol{\delta})], \quad (2.5a)$$

$$\nabla_{\boldsymbol{\delta}} [\chi(\mathbf{m}^{(k)} + \boldsymbol{\delta})] = \nabla_{\boldsymbol{\delta}} \left[\chi(\mathbf{m}^{(k)}) + \nabla \chi^{(k)T} \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^T \mathbf{H}^{(k)} \boldsymbol{\delta} \right] = 0, \quad (2.5b)$$

$$\nabla \chi^{(k)} + \mathbf{H}^{(k)} \boldsymbol{\delta} = 0, \quad (2.5c)$$

$$\mathbf{H}^{(k)} \boldsymbol{\delta} = -\nabla \chi^{(k)}, \quad (2.5d)$$

$$\boldsymbol{\delta} = -\mathbf{H}^{(k)-1} \nabla \chi^{(k)} = -\mathbf{G}^{(k)} \nabla \chi^{(k)}, \quad (2.5e)$$

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \boldsymbol{\delta}. \quad (2.5f)$$

A direção descendente neste caso passa a ser expressa pela Equação 2.5e. $\mathbf{G} = \mathbf{H}^{-1}$ é definida como a inversa da matriz Hessiana. Quando \mathbf{H} é uma matriz positiva definida, a solução $\mathbf{m} + \boldsymbol{\delta}$ é o mínimo da função objetivo.

Este método é chamado de método de Newton e, embora de convergência muito rápida, possui um alto custo computacional, pois necessita calcular e inverter a matriz Hessiana \mathbf{H} para cada iteração. Em muitos casos, o cálculo da matriz Hessiana é inviável devido ao tamanho e a complexidade do problema. Isto é especialmente verdadeiro para os problemas de inversão do campo de onda completo estudados nesta tese. Com isto, torna-se necessário procurar uma aproximação para o minimizador de Newton que seja mais eficiente.

2.2.3 Quasinewton

Para diminuir o custo computacional do método de Newton, pode-se aproximar diretamente a matriz inversa da Hessiana (\mathbf{G}) iterativamente. Existem algumas formas possíveis de realizar esta estimativa, como por exemplo, o método apresentado por [Broyden \(1970\)](#), [Fletcher \(1970\)](#), [Goldfarb \(1970\)](#) e [Shanno \(1970\)](#). Este algoritmo é conhecido como BFGS devido às iniciais dos autores e está apresentado nas Equações 2.6A,E,

$$\mathbf{g}^{(k)} = -\mathbf{G}^{(k)} \nabla \chi^{(k)}, \quad (2.6a)$$

$$\mathbf{G}^{(k+1)} = \mathbf{G}^{(k)} + \left(1 + \frac{\boldsymbol{\gamma}^{(k)T} \mathbf{G}^{(k)} \boldsymbol{\gamma}^{(k)}}{\boldsymbol{\gamma}^{(k)T} \boldsymbol{\delta}^{(k)}} \right) \frac{\boldsymbol{\delta}^{(k)} \boldsymbol{\delta}^{(k)T}}{\boldsymbol{\delta}^{(k)T} \boldsymbol{\gamma}^{(k)}} - \left(\frac{\boldsymbol{\delta}^{(k)} \boldsymbol{\gamma}^{(k)T} \mathbf{G}^{(k)} + \mathbf{G}^{(k)} \boldsymbol{\gamma}^{(k)} \boldsymbol{\delta}^{(k)T}}{\boldsymbol{\delta}^{(k)T} \boldsymbol{\gamma}^{(k)}} \right), \quad (2.6b)$$

$$\boldsymbol{\delta}^{(k)} = \boldsymbol{\alpha}^{(k)} \mathbf{g}^{(k)}, \quad (2.6c)$$

$$\boldsymbol{\gamma}^{(k)} = \nabla \chi^{(k+1)} - \nabla \chi^{(k)}, \quad (2.6d)$$

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha^{(k)} \mathbf{g}^{(k)}. \quad (2.6e)$$

Este método, por ser uma aproximação de Newton, possui uma taxa de convergência menor. O custo computacional comparado ao Newton é significativamente menor, pois não necessita da inversão da matriz Hessiana. Porém, ainda existe a necessidade do armazenamento de \mathbf{G} . Para problemas muito grandes, isto exige uma capacidade de armazenamento muito elevada, tornando-o ineficiente. É possível implementar este algoritmo sem a necessidade de armazenar a matriz \mathbf{G} . Isto será visto no capítulo 4.

2.2.4 Gradiente Conjugado

O último método estudado foi o do gradiente conjugado. Este algoritmo consiste em encontrar o otimizador de uma função quadrática a partir de vetores conjugados, definidos a seguir.

Dado um conjunto de vetores não nulos $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(n)}$ e uma matriz positiva definida \mathbf{H} que atende a condição expressa em 2.7,

$$\mathbf{g}^{(i)T} \mathbf{H} \mathbf{g}^{(j)} = 0 \quad \forall i \neq j. \quad (2.7)$$

Diz-se que os vetores $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(k)}$ são vetores conjugados, ou H-ortogonais.

Os vetores conjugados podem ser encontrados da seguinte forma: faz-se $\mathbf{g}^{(0)} = -\nabla \chi^{(0)}$ e as direções seguintes $\mathbf{g}^{(k)}$ são calculadas iterativamente utilizando as Equações 2.8A,B,

$$\mathbf{g}^{(k)} = -\nabla \chi^{(k)} + \beta^{(k)} \mathbf{g}^{(k-1)}, \quad (2.8a)$$

$$\text{onde } \beta^{(k)} = \frac{\nabla \chi^{(k)T} \nabla \chi^{(k)}}{\nabla \chi^{(k-1)T} \nabla \chi^{(k-1)}}. \quad (2.8b)$$

O método do gradiente conjugado é capaz de encontrar o mínimo de uma função quadrática em \mathbb{R}^N em, no máximo, N iterações. Desta forma, a convergência do algoritmo

aumenta significativamente, sem aumentar demasiadamente o custo computacional para cada iteração, pois não é necessário calcular nenhuma matriz inversa e nem armazenar matrizes demasiadamente grandes.

2.3 Busca do passo de iteração ótimo

Conforme visto anteriormente, todos os métodos de otimização consistem em calcular um vetor direção de busca e em seguida atualizar os parâmetros da função objetivo de forma que a condição expressa em 2.1 seja atendida. Para isto multiplica-se o vetor de busca por um valor α , onde α é o passo de iteração. Este valor deve ser escolhido de forma a otimizar o passo de iteração, conforme descrito na Equação 2.9,

$$\chi(\mathbf{m}^{(k)} + \alpha_0 \mathbf{g}^{(k)}) < \chi(\mathbf{m}^{(k)} + \alpha \mathbf{g}^{(k)}) \text{ , para todo } \alpha \neq \alpha_0, \quad (2.9)$$

onde α_0 é o passo de iteração ótimo. O cálculo de α pode ser feito de diversas maneiras. Este trabalho busca um valor de α que atenda as condições de Wolfe (FLETCHER, 1980), expressas nas Equações 2.10A,B,

$$\chi(\mathbf{m}^{(k)} + \alpha^{(k)} \mathbf{g}^{(k)}) \leq \chi(\mathbf{m}^{(k)}) + \sigma_1 \alpha^{(k)} \nabla \chi^{(k)T} \mathbf{g}^{(k)}, \quad (2.10a)$$

$$|\nabla \chi^{(k+1)T} \mathbf{g}^{(k)}| \leq -\sigma_2 \nabla \chi^{(k)T} \mathbf{g}^{(k)}, \quad (2.10b)$$

sendo $0 < \sigma_1 \leq \sigma_2 < 1/2$. O algoritmo que realiza a busca na direção \mathbf{g}_k pelo α_k que atende estas condições é chamado de algoritmo de *line search*. Este algoritmo implica em recalcular a função objetivo algumas vezes na busca do passo ótimo, o que aumenta consideravelmente o custo computacional para cada passo de iteração, embora o objetivo de encontrar um α próximo do ótimo seja de aumentar a taxa de convergência, ou seja, reduzir o número total de iterações necessárias para atingir o mínimo da função objetivo. A seção seguinte apresenta um exemplo de otimização não linear utilizando os métodos aqui descritos.

2.4 Exemplo

De modo a demonstrar as diferentes taxas de convergências dos métodos estudados nas seções anteriores, serão apresentados a seguir dois exemplos de otimização em funções de duas variáveis. A partir de um ponto inicial, busca-se o mínimo da função utilizando os diferentes métodos de otimização. Cada método será avaliado em relação ao número de iterações necessárias para atingir o mínimo e o tempo total necessário, em segundos. É importante ressaltar que a medida de tempo neste caso simples não é precisa para

avaliar o tempo de convergência de um determinado método, porém estes valores ajudam a fornecer uma análise comparativa entre o custo computacional de cada iteração para cada um dos métodos estudados. Devido à simplicidade das funções objetivas estudadas, o cálculo do valor de α_0 é feito através de uma varredura de um número muito grande de α 's. Isto foi feito para todos os métodos, exceto o método de Newton, pois este não apresenta o cálculo de α explicitamente. Desta forma garante-se um valor ótimo de α e as convergências devem-se exclusivamente ao método de otimização. Para os exemplos aqui apresentados, as funções são conhecidas, assim como seu mínimo e o valor da função para o mínimo. O critério de parada escolhido foi o mínimo da função acrescido de uma pequena tolerância. Um segundo critério de parada foi utilizado limitando o número máximo de iterações. Para os exemplos a seguir, o número máximo de iterações permitidas foi de 500 iterações. Todos os algoritmos foram implementados utilizando o software MATLAB.

A primeira função utilizada é uma função quadrática. Neste caso, todos os métodos devem funcionar de maneira ótima. A função em duas variáveis $\mathbf{m} = (m_1 \ m_2)^T$ encontra-se representada na Equação 2.11,

$$f(\mathbf{m}) = 10m_1^2 + 28m_1m_2 + 20m_2^2. \quad (2.11)$$

O gradiente da função encontra-se na Equação 2.12 e a matriz Hessiana está na Equação 2.13,

$$\nabla \chi(\mathbf{m}) = \begin{pmatrix} 20m_1 + 28m_2 \\ 40m_2 + 28m_1 \end{pmatrix}, \quad (2.12)$$

$$\mathbf{H}(\mathbf{m}) = \begin{bmatrix} 20 & 28 \\ 28 & 40 \end{bmatrix}. \quad (2.13)$$

O mínimo da função é o ponto $(0; 0)$. O ponto inicial para a otimização é $(0, 5; -0.1)$. O resultado para cada método encontra-se na Tabela 1. A Figura 1 mostra as direções de busca do mínimo da função, fornecendo uma visão geométrica da convergência de cada método. Todos os métodos funcionaram muito bem. O método de Newton atingiu o mínimo em 1 iteração, como esperado. O método do gradiente conjugado foi capaz de atingir o mínimo em 2 iterações, conforme a propriedade dos vetores conjugados. Para este exemplo o método do gradiente descendente atingiu o mínimo em muitas iterações, e mesmo com um custo computacional baixo para cada iteração, tornou-se o método menos atrativo.

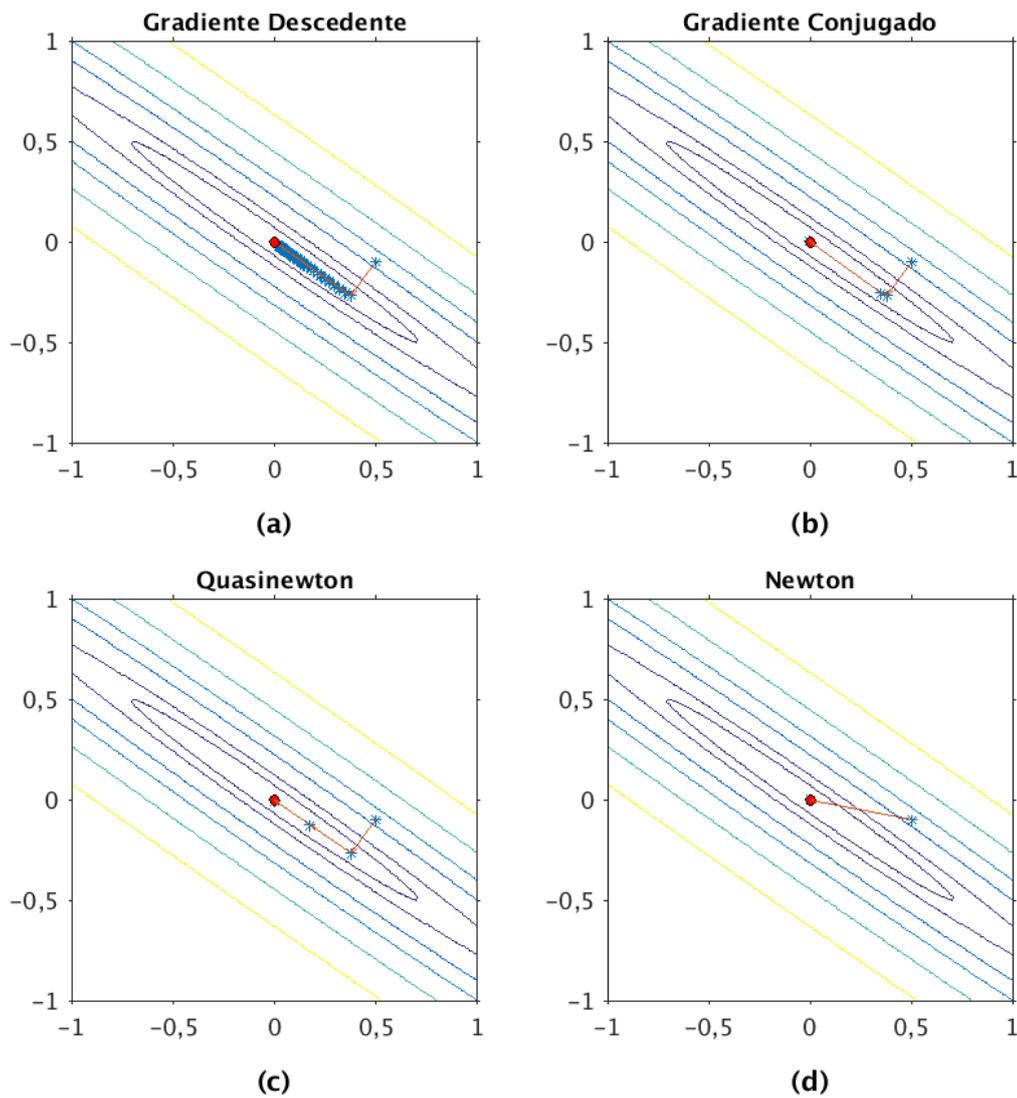
O segundo exemplo consiste na função de teste Rosenbrock, uma função muito utilizada na literatura para verificar a taxa de convergência de otimizadores. A função em

Tabela 1 – Resultado função quadrática.

Método	Iterações	Tempo (s)
Gradiente descendente	69	0,45
Gradiente conjugado	2	0,04
Newton	1	0,07
BFGS	3	0,12

Fonte: Produzida pelo autor.

Figura 1 – Resultado para os métodos de otimização. Função Quadrática



Fonte: Produzida pelo autor

duas variáveis $\mathbf{m} = (m_1 \ m_2)^T$ está representada na Equação 2.14,

$$f(\mathbf{m}) = 100(m_2 - m_1^2)^2 + (1 - m_1)^2. \quad (2.14)$$

O gradiente da função encontra-se na Equação 2.15 e a matriz Hessiana está na Equação 2.16,

$$\nabla\chi(\mathbf{m}) = \begin{pmatrix} -400m_1(m_2 - m_1^2) - 2(1 - m_1) \\ 200(m_2 - m_1^2) \end{pmatrix}, \quad (2.15)$$

$$\mathbf{H}(\mathbf{m}) = \begin{bmatrix} 1200m_1^2 - 400m_2 + 2 & -400m_1 \\ -400m_1 & 200 \end{bmatrix}. \quad (2.16)$$

O mínimo da função encontra-se no ponto (1; 1). De maneira a avaliar as taxas de convergência dos métodos apresentados neste capítulo, os otimizadores testados partiram do ponto (0; 0). O resultado encontra-se na Tabela 2. A Figura 2 mostra as direções de busca do mínimo da função, fornecendo uma visão geométrica da convergência de cada método.

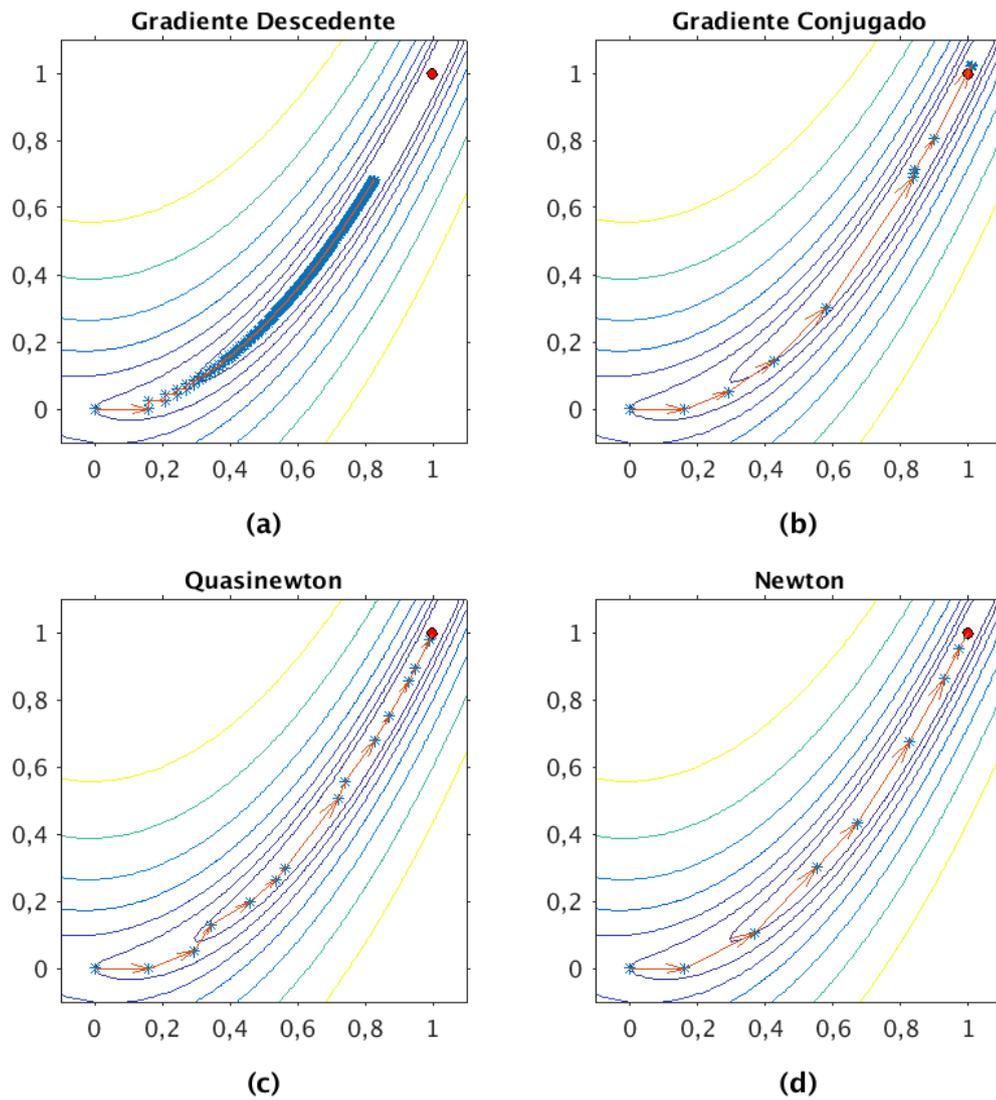
Tabela 2 – Resultado Rosenbrock

Método	Iterações	Tempo (s)
Gradiente descendente	500	0,29
Gradiente conjugado	14	0,05
Newton	8	0,65
BFGS	14	1,30

Fonte: Produzida pelo autor

Comparando os resultados da Tabela 2 e da Figura 2, é possível perceber algumas diferenças entre os métodos de otimização. O gradiente descendente é de fato um método bastante rápido de cálculo de otimização. O tempo necessário para calcular 500 passos é inferior ao tempo necessário para encontrar 8 passos do método de Newton e 14 passos do Quasinewton. Contudo, fica bem claro que a convergência do método não foi satisfatória, pois atingiu o máximo de iterações permitidas sem que o resultado chegasse perto do mínimo da função. Todos os outros métodos convergiram para o mínimo da função. O método de Newton obteve a convergência mais rápida de todos os métodos em termos de número de iterações, o que era esperado. Porém, conforme dito antes, o cálculo da Hessiana e a sua inversa não é sempre possível. Desta forma, por melhor que este método seja para funções mais simples, como é o caso deste exemplo, ele não pode ser utilizado

Figura 2 – Resultado para os métodos de otimização. Função de Rosenbrock



Fonte: Produzida pelo autor

para FWI. Ao analisar o resultado para os métodos de gradiente conjugado e quasinewton, nota-se que ambos possuem uma convergência satisfatória. Para cada iteração, o método do gradiente conjugado possui um custo computacional bem inferior ao método quasinewton. Sendo assim, para este exemplo, o método que obteve o resultado mais satisfatório levando em consideração taxa de convergência e custo computacional foi o método do gradiente conjugado.

2.5 Conclusão

Este capítulo apresentou o processo matemático da otimização não linear, e os diversos métodos de cálculo da direção de busca da função objetivo. O método mais simples estudado foi o gradiente descendente, embora a sua convergência não seja sempre satisfatória. Os métodos de quasinewton e gradiente conjugado são capazes de encontrar o mínimo da função em um número relativamente baixo de iterações, embora o método do gradiente conjugado tenha um custo computacional menor para cada iteração.

A seguir, é apresentado o problema da inversão do campo de onda completo aplicado à geofísica. O equacionamento do processo de propagação do campo de onda e como modelá-lo computacionalmente é apresentado. Em seguida define-se a função objetivo para a FWI. Finalmente, o cálculo do gradiente da função objetivo é desenvolvido, mostrando seu significado físico e como calculá-lo de forma eficiente. Os métodos matemáticos estudados neste capítulo são aplicados diretamente na FWI. O próximo capítulo trata da modelagem da propagação de ondas.

3 Modelagem

No capítulo anterior estudamos como encontrar o mínimo de uma função não linear genérica. A seguir vamos aplicar este minimizador à sísmica. Para isto, descrevemos o processo de inversão como na Figura 3. Por um lado temos o processo de aquisição sísmica. Uma fonte é detonada e a onda sísmica propagada é medida por sensores chamados geofones ou hidrofones, gerando os sismogramas medidos. Na aquisição *streamer* convencional, a fonte e os receptores encontram-se próximos à superfície da água. Já uma aquisição OBN (*Ocean Bottom Nodes*), a fonte é detonada próximo à superfície da água enquanto os receptores encontram-se no fundo do mar. Existem outras configurações de aquisições sísmicas, terrestres e marinhas. Nesta tese, trabalhamos com uma aquisição semelhante à uma aquisição marítima convencional, onde as fontes e os receptores encontram-se próximos à superfície da água. Este processo pode ser descrito matematicamente. Para isto é necessário definir um modelo das propriedades físicas no qual a onda simulada será propagada. Este processo é chamado de modelagem. Naturalmente, se o modelo sintético representa com exatidão a sub-superfície e se a modelagem é capaz de representar todo o processo de propagação da onda pelo meio, os sismogramas medido e calculado serão iguais. Desta forma, busca-se um modelo sintético que aproxime o sismograma simulado do sismograma medido, em outras palavras, que minimize as diferenças entre os sismogramas.

Antes de definir o problema inverso, é preciso estudar o problema direto. No caso da sísmica, o processo físico estudado é a equação da propagação de onda. Nesta tese optamos pela aproximação acústica da equação da onda. Na próxima seção, veremos a equação da onda acústica e como podemos modelar computacionalmente esta propagação.

3.1 Modelagem Acústica

A aproximação acústica 2D da equação da onda encontra-se na Equação 3.1,

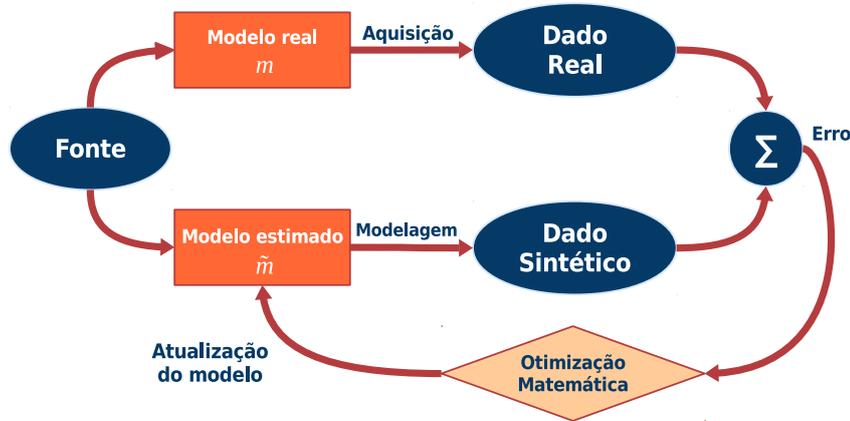
$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2}, \quad (3.1)$$

onde c é a velocidade de onda de pressão P e u é o campo de pressão. Esta tese utiliza o equivalente da Equação 3.1 no domínio da frequência. Para isto, tomamos a transformada de Fourier da Equação 3.1,

$$\frac{\omega^2}{c^2} u - \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2} = -s, \quad (3.2)$$

onde s é o termo da fonte. Em seguida, discretizamos as derivadas espaciais utilizando

Figura 3 – Esquema de inversão



Fonte: Produzida pelo autor

diferenças finitas de segunda ordem centrais. Todas as derivações utilizam uma grade retangular em duas dimensões. Referenciais locais para posições na grade são feitas utilizando os subscritos i e j . Substituindo os operadores de diferenças finitas na Equação 3.2 temos,

$$\frac{\omega^2}{c^2} u_{i,j} + \left[\frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\Delta x^2} \right] + \left[\frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{\Delta z^2} \right] = -s_{i,j}. \quad (3.3)$$

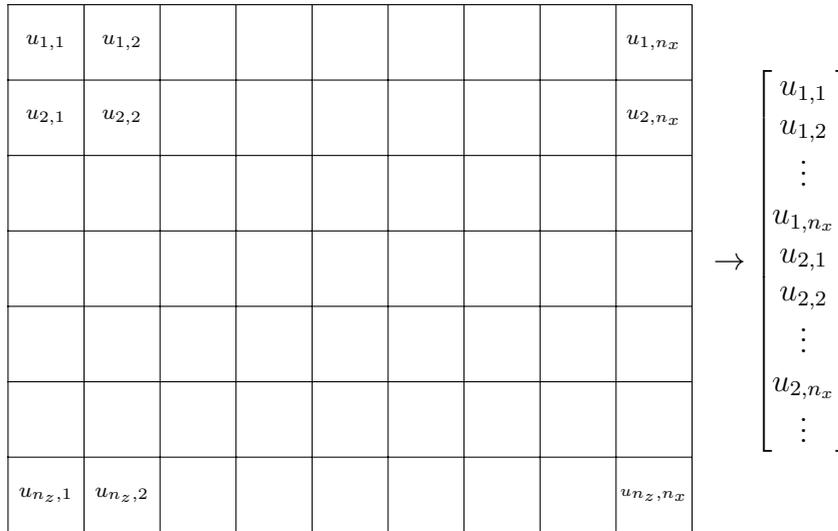
Este problema pode ser escrito em forma matricial se transformarmos a grade 2D em vetor conforme mostrado na Figura 4. A equação da onda no domínio da frequência em forma matricial encontra-se na Equação 3.4,

$$\mathbf{B}\mathbf{u} = -\mathbf{s}, \quad (3.4)$$

onde \mathbf{B} é o operador de diferenças finitas em forma matricial, que inclui todas as variações de velocidades, \mathbf{u} é o campo de pressão complexo e \mathbf{s} é o termo da fonte. A definição mais detalhada dos elementos da matriz \mathbf{B} encontra-se no Apêndice A. Desta forma, para obter o campo de pressão para uma determinada componente de frequência resolve-se a equação 3.4, conforme a Equação 3.5,

$$\mathbf{u} = -\mathbf{B}^{-1}\mathbf{s}. \quad (3.5)$$

Figura 4 – Modelo discretizado



Fonte: Produzida pelo Autor

Isto implica na solução de um sistema linear para cada componente de frequência. Existem algumas vantagens em se fazer a modelagem desta forma. A primeira delas é que as componentes de frequência são independentes. Podemos resolver a Equação 3.5 em qualquer ordem e o resultado final será o mesmo. Com isto podemos paralelizar o problema de forma simples. Dividimos o número total de componentes de frequências simuladas através de um *cluster* de processamento e com isto reduzimos o tempo total de processamento. Outra vantagem que permite diminuir o tempo total de modelagem vem da característica do sismograma no tempo. Por ser uma série real no tempo, temos a propriedade descrita na Equação 3.6 no domínio da frequência,

$$\mathbf{u}(\omega) = \mathbf{u}^*(-\omega). \quad (3.6)$$

Desta forma, sabendo metade dos termos na frequência, podemos calcular a outra metade a partir da equação 3.6. Isto reduz o custo computacional da modelagem pela metade. Além disso, existe uma propriedade da sísmica que nos permite reduzir ainda mais o custo computacional. Em geral, o espectro de frequência da fonte sísmica é limitado, o que significa que podemos aproximar o espectro do campo de ondas apenas pelas frequências presentes na fonte.

Sendo assim, para realizar a modelagem no domínio da frequência, primeiro avaliaremos a banda de frequência da fonte sísmica utilizada. Em seguida, modelamos as componentes de frequência entre 0 e f_{max} . Fazemos $\mathbf{u}(\omega) = 0$ para $2\pi f_{max}/f_s < \omega < \pi$, onde f_s é a frequência de amostragem da modelagem. Utilizamos a propriedade 3.6 para calcular $\mathbf{u}(\omega)$ para $-\pi < \omega < 0$. A transformada de Fourier inversa do campo de pressão resulta na série no domínio do tempo.

Como a transformada de Fourier de um sinal discreto é uma função contínua, precisamos discretizá-la no processo. Isto significa que podemos simular quantas componentes de frequência quisermos entre os valores de $0 < \omega < 2\pi f_{max}/f_s$. Se utilizarmos uma discretização muito pequena, poderemos ter um resultado de baixa qualidade. Mostraremos isto em mais detalhes no exemplo do final deste capítulo.

Na definição da modelagem, é preciso definir a equação na borda do modelo de velocidades. Como utilizamos um modelo finito, é necessário definir as equações de forma que a onda propagada não retorne para o interior do modelo ao encontrar a borda.

3.2 Borda de absorção

Nesta tese foram implementados alguns métodos de bordas de absorção. O primeiro método implementado utiliza uma condição simples de borda de absorção (*Absorbing Boundary Condition - ABC*) proposto por Clayton e Engquist (1977). A equação ABC encontra-se na Equação 3.7,

$$\frac{\partial \mathbf{u}}{\partial n} - ik\mathbf{u} = 0, \quad (3.7)$$

onde n é a direção normal à borda de absorção do modelo. Este método tem a desvantagem de uma absorção ruim para altos ângulos de incidência. Outro método implementado é a técnica de *Perfect Matching Layers* (PML) (BERENGER, 1994) que utiliza a equação de onda com o termo de pressão dividido por um fator α , e que oferece uma absorção quase completa da onda. Esta técnica envolve um termo de atenuação α incorporado na equação de onda 3.2. A nova equação encontra-se em 3.8,

$$-\frac{\omega^2}{c^2} \mathbf{u} = \frac{1}{\alpha} \left(\frac{\partial^2 \mathbf{u}}{\partial x^2} + \frac{\partial^2 \mathbf{u}}{\partial z^2} \right) + \mathbf{s}. \quad (3.8)$$

A equação 3.8 é válida para a região da borda PML ao redor do modelo. O termo α pode assumir diferentes formas. O termo proposto por Hustedt, Operto e Virieux (2004), apresentado na Equação 3.9 foi utilizado nesta tese,

$$\alpha = 1 + i \frac{C_{PML}}{\omega} \cos \left(\frac{\pi}{2} \frac{x_{PML}}{L} \right). \quad (3.9)$$

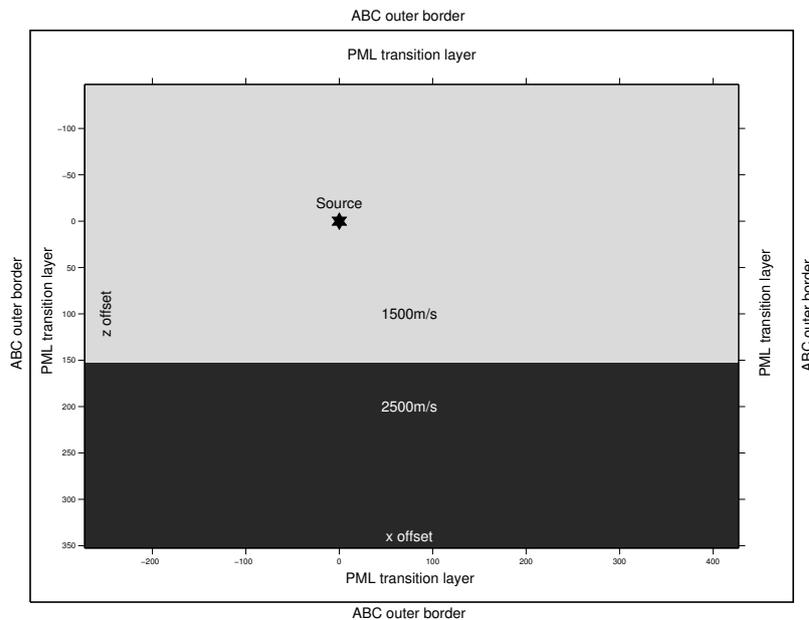
Na equação acima, L é o tamanho da borda PML e x_{PML} é a coordenada local, de modo que $\alpha = 1$ na parte interior da borda de absorção. O coeficiente escalar C_{PML} depende da largura da borda PML. O valor ótimo de C_{PML} é tal que minimize reflexões espúrias na borda.

O terceiro método desenvolvido nesta tese, definido em (MOREIRA et al., 2014) é um híbrido, que consiste em uma região de transição, implementada pela borda PML e,

ao final desta região, a borda ABC é implementada. O objetivo desta técnica é reduzir o tamanho da borda PML, minimizando os efeitos espúrios e o custo computacional simultaneamente.

Para testar este novo método, utilizamos três modelos de velocidade: Um modelo homogêneo, um de duas camadas e um janela do modelo Marmousi. Para os dois primeiros, as dimensões x e z são iguais a 500 m e o espaçamento da grade é de 5 m. O meio homogêneo possui velocidade $c = 2000$ m/s, enquanto que o modelo de duas camadas possui $c_{superior} = 1500$ m/s e $c_{inferior} = 2500$ m/s, conforme ilustrado na Figura 5. O modelo Marmousi encontra-se na Figura 6. As dimensões deste são 9204×3004 m com um espaçamento de grade de 4 m.

Figura 5 – Modelo de duas camadas

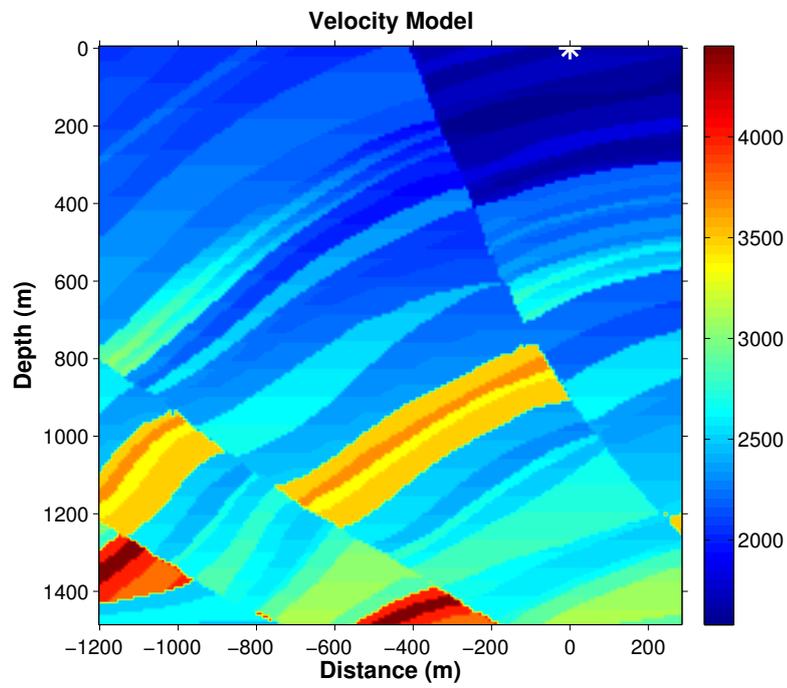


Fonte: Produzida pelo autor

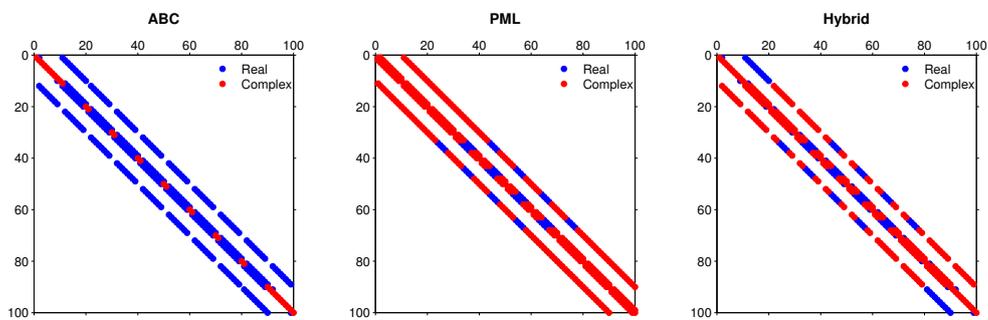
A matriz do operador de diferenças finitas \mathbf{B} apresentada na Equação 3.4 é uma matriz esparsa, não simétrica e complexa. A Figura 7 demonstra o padrão esparsa de \mathbf{B} , supondo uma grade 10×10 para as técnicas ABC, PML e híbrida, utilizando 3 termos da grade para a borda PML. Os círculos azul e vermelho correspondem a termos reais e complexos respectivamente. A técnica PML insere mais elementos complexos na matriz que as técnicas ABC e híbrida.

O tempo máximo de simulação dos sismogramas no tempo foi de 1,1 s, com as soluções calculadas para 62 componentes de frequência de 0 a 75 Hz. A fonte utilizada foi uma *Ricker* de frequência central de 30 Hz. A equação da fonte no domínio da frequência

Figura 6 – Modelo de Marmousi



Fonte: Produzida pelo autor

Figura 7 – Coeficientes Reais e Complexos da matriz **B**. a) ABC b) PML c) Híbrido

Fonte: Produzida pelo autor

encontra-se na Equação 3.10,

$$s = \frac{2f^2}{\pi^2 f_c^3} \exp\left(-\frac{f^2}{\pi f_c^2}\right). \quad (3.10)$$

Na equação acima, $f = \omega/(2\pi)$ é a frequência e $f_c = \bar{f}(3\sqrt{\pi})$ é a frequência central da fonte, e está relacionada a frequência de corte \bar{f} .

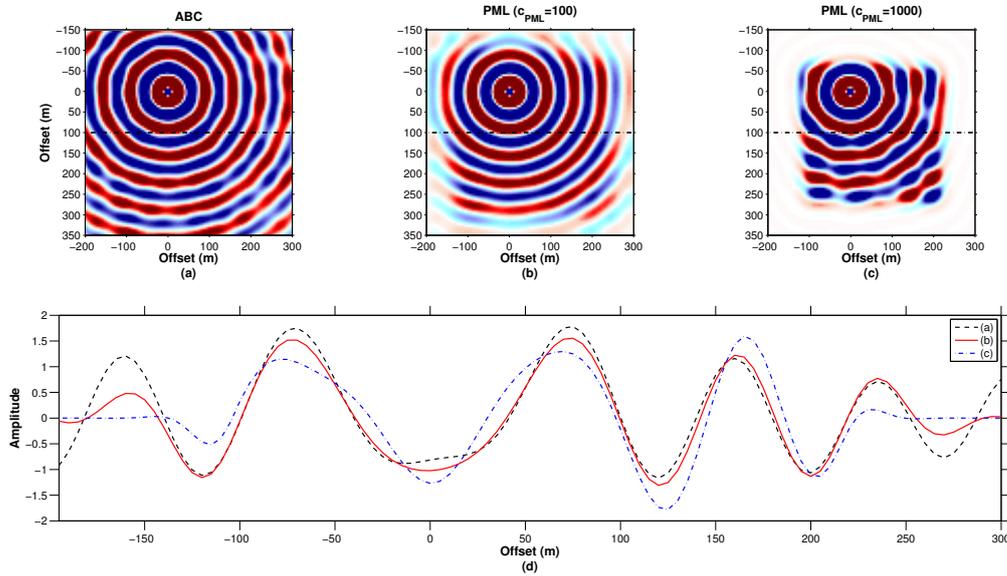
A Figura 8 mostra a parte real da componente de 30 Hz da solução do campo de pressão para o modelo homogêneo obtida utilizando as bordas tradicionais ABC e PML. A eficiência da borda PML é altamente dependente do coeficiente C_{PML} . Desta forma, foram testado diversos valores para C_{PML} para uma largura de borda de $L = 100$ m. O melhor resultado foi encontrado para $C_{PML} = 100$. Comparando este resultado com o método ABC, vemos que as reflexões de borda foram bastante minimizadas para $C_{PML} = 100$ (Figura 8b). A Figura 8c mostra o que acontece com a absorção ao escolher valores excessivamente altos para C_{PML} . Neste caso, cria-se uma "nova" borda e as reflexões causadas pela borda são ampliadas.

Conforme diminuimos a largura da borda PML, os efeitos colaterais tornam-se mais pronunciados. A Figura 9a apresenta a parte real da componente de 30 Hz da solução do campo de pressão para o modelo homogêneo obtido com uma borda PML mais curta ($L = 25$ m). Vemos claramente uma distorção da amplitude das ondas causada pela reflexão da borda menor. Comparando este resultado com a borda híbrida, percebemos que a última reduz consideravelmente as reflexões de borda para o mesmo comprimento da camada PML ($L = 25$ m). Este resultado pode ser observado na Figura 9b. Os resultados para a borda híbrida melhoram ainda mais se aumentarmos a largura da borda PML, como mostrado na Figura 9c para $L = 100$ m. Realizamos testes entre os três tipos de borda para frequências mais baixas e mais altas, sempre obtendo o melhor resultado para a borda híbrida.

A Figura 10 apresenta o sismograma no domínio do tempo para o modelo de duas camadas utilizando a borda ABC, PML otimizado (ou seja, com o melhor C_{PML}) e híbrido otimizado (ABC + PML otimizado). Podemos observar a onda direta e uma reflexão forte para os três casos. Uma pequena reflexão causada pela borda pode ser observada em todos os casos, embora a técnica híbrida apresente a menor reflexão dos três casos. Observando com maiores detalhes um traço sísmico (Figura 10d), um ruído residual de baixa frequência para a borda ABC é detectado. A eficiência da absorção de cada técnica pode ser estimada comparando a soma dos quadrados das amplitudes em tempo para um determinado traço sísmico. Esta métrica encontra-se na equação 3.11,

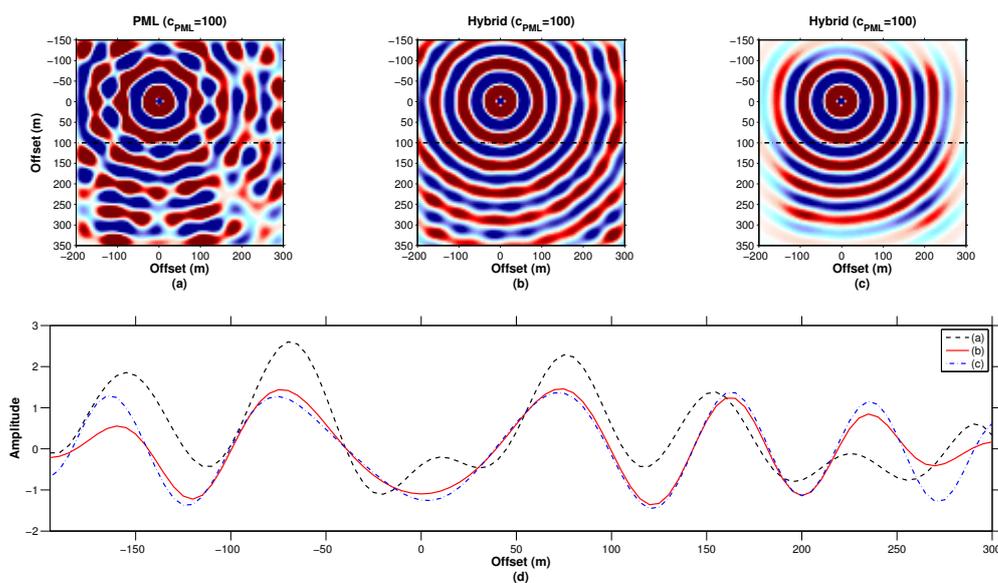
$$E = \sum_{t=T_0}^T \mathbf{u}^2. \quad (3.11)$$

Figura 8 – Comparação das bordas de absorção. a) ABC b) PML com $C_{PML} = 100$ c) PML com $C_{PML} = 1000$



Fonte: Produzida pelo autor

Figura 9 – Comparação das bordas de absorção. a) PML com $L = 25$ m b) Híbrido com $L = 25$ m c) Híbrido com $L = 100$ m



Fonte: Produzida pelo autor

Utilizamos a métrica acima para calcular o valor de E para cada técnica dentro do intervalo no tempo de $0,43 \text{ s} \leq t \leq 0,8 \text{ s}$ (área hachurada na Figura 10d). Os resultados encontram-se na Tabela 3. O modelo híbrido obteve o melhor resultado, com $0,60 \text{ Pa}^2$ seguido do ABC e PML com $1,64$ e $2,17 \text{ Pa}^2$ respectivamente.

Tabela 3 – Resultado das técnicas de absorção

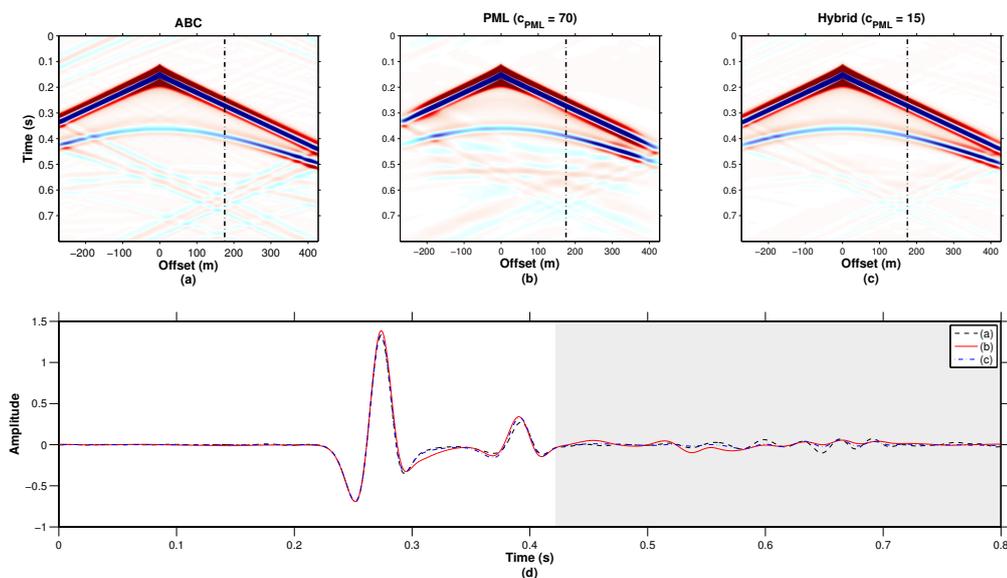
Borda	$E(\text{Pa}^2)$
ABC	1,64
PML otimizado	2,17
Híbrido otimizado	0,60

Fonte: Produzida pelo autor

Finalmente, a Figura 11 apresenta o resultado para o modelo Marmousi, utilizando as três técnicas aqui apresentadas. A borda híbrida apresenta as menores reflexões espúrias de todos os métodos.

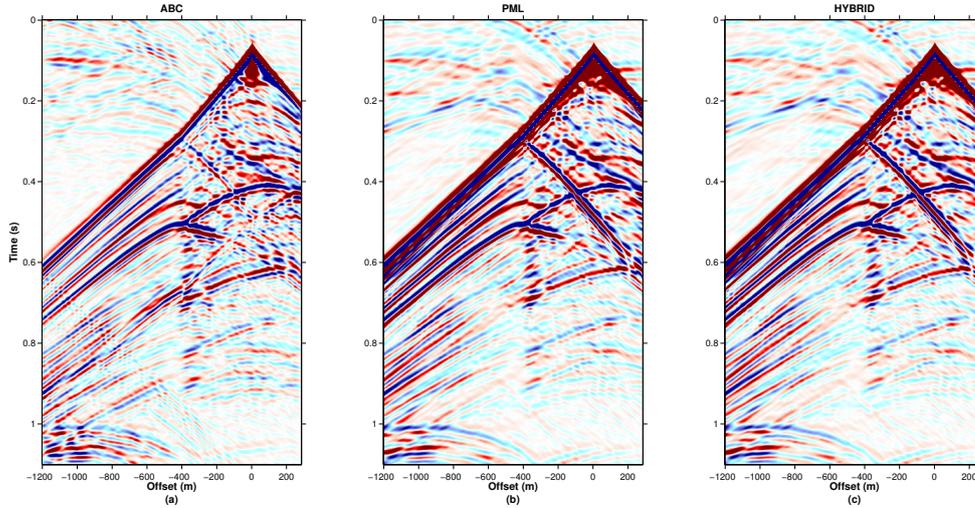
Desta forma, o algoritmo de modelagem apresentado neste capítulo utiliza a borda híbrida em sua implementação. Para uma descrição mais detalhada de como definir os termos da matriz \mathbf{B} utilizando a técnica de borda híbrida, veja o Apêndice A.

Figura 10 – Comparação das bordas de absorção. a) ABC b) PML otimizado c) Híbrido otimizado



Fonte: Produzida pelo autor

Figura 11 – Comparação das bordas de absorção. a) ABC b) PML c) Híbrido



Fonte: Produzida pelo autor

3.3 Exemplo

Para demonstrar a modelagem acústica implementada nesta tese, vamos simular a propagação de um fonte explosiva em um meio acústico. Vamos usar como exemplo o modelo de Marmousi de 1000×2004 m, apresentado na Figura 12. Para este exemplo, usamos um tamanho de grid quadrado de 4 m. Para a simulação, utilizamos um valor baixo de amostragem temporal, de 2×10^{-5} s para minimizar os efeitos de dispersão numérica. Simulamos 1,1 s de propagação. Os parâmetros da modelagem encontram-se resumidos na tabela 4. A fonte simulada foi do tipo Ricker (Equação 3.10). Para este exemplo, utilizamos uma fonte de frequência central $f_c = 20$ Hz. Simulamos uma fonte na superfície do modelo, logo abaixo da borda de absorção. O campo de pressão simulado foi medido também na superfície, na mesma profundidade de fonte. Os receptores foram espalhados por toda a superfície, exceto no interior das bordas de absorção. A Figura 13 apresenta a fonte utilizada. A Figura 13a apresenta a fonte no domínio do tempo, enquanto que a Figura 13b ilustra o espectro de frequência desta fonte. Podemos notar que o espectro de frequência da fonte é limitado em aproximadamente 60 Hz. Desta forma, somente necessitamos calcular as componentes de frequência entre 0 e 60 Hz. Para a discretização das componentes de frequência, utilizamos uma potência de 2 maior que o número de amostras no tempo que queremos simular. Neste caso, queremos simular 1,1 s com $\Delta t = 2 \cdot 10^{-5}$ s, ou um total de $n_t = 55000$ amostras no tempo. O algoritmo utilizado para realizar a transformada de Fourier da fonte e posteriormente do sismograma é o FFT. Este algoritmo é otimizado quando utilizamos um número de elementos na frequência em potências de 2. A potência de 2 mais próxima de n_t , maior que o mesmo é $2^{16} = 65536$. Chamaremos esta primeira potência maior de $NFFT = 0$. Simulando apenas as frequências entre 0 e 60 Hz, temos

um total de 7865 componentes de frequências simuladas, bem inferior aos 55000 instantes de tempo necessários para realizar a modelagem no domínio do tempo. Se diminuirmos a amostragem na frequência, ou seja, aumentarmos o número total de componentes de frequência entre 0 e f_{max} Hz, melhoramos a qualidade da modelagem. Podemos aumentar ainda mais a potência de 2 para determinar o novo número de componentes totais na frequência. Neste exemplo, teríamos $2^{17} = 131072$ termos. Chamaremos esta potência de $NFFT = 1$. Para demonstrar a influência de $NFFT$ e f_{max} , realizamos 4 modelagens utilizando valores de f_{max} de 60 e 100 Hz e valores de $NFFT$ de 2^{16} e 2^{17} . O número total de componentes de frequência simuladas para cada caso encontra-se na Tabela 5. Os sismogramas em tempo resultantes encontram-se na Figura 14. Podemos perceber que os sismogramas utilizando $f_{max} = 60$ Hz e $f_{max} = 100$ Hz são praticamente idênticos. Utilizando uma amostragem menor no domínio da frequência, podemos perceber um menor ruído de simulação acima da onda direta. Para os testes de FWI, realizaremos as modelagens utilizando $f_{max} = 3f_c$ e $NFFT = 1$.

Tabela 4 – Parâmetros da Modelagem

Parâmetro	Valor
Δx	4 m
Δy	4 m
Δt	$2 \cdot 10^{-5}$ s
Tempo total	1,1 s

Fonte: Produzida pelo autor

Tabela 5 – Total de termos de frequência simulados

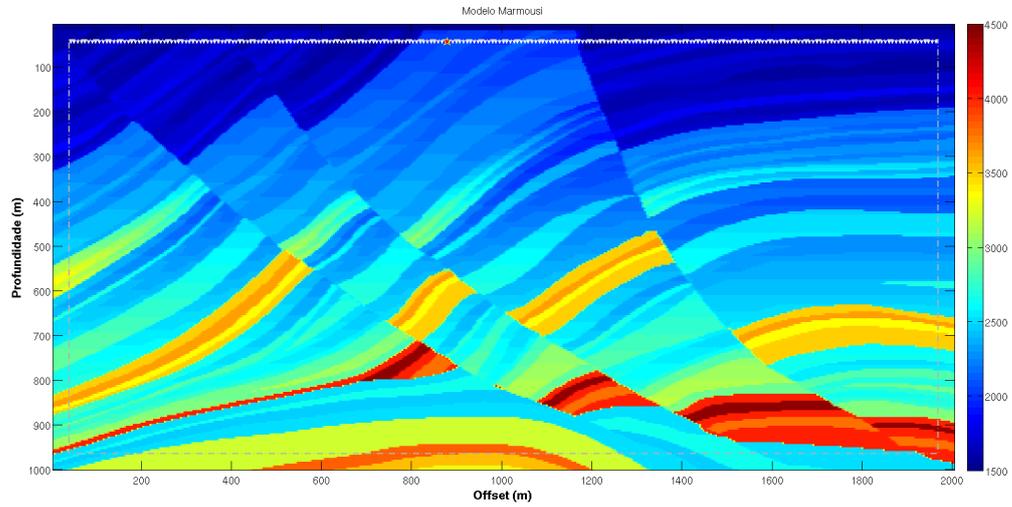
$f_{max}(\text{Hz})$	$NFFT = 0$	$NFFT = 1$
60	7865	15730
100	13108	26215

Fonte: Produzida pelo autor

3.4 Conclusão

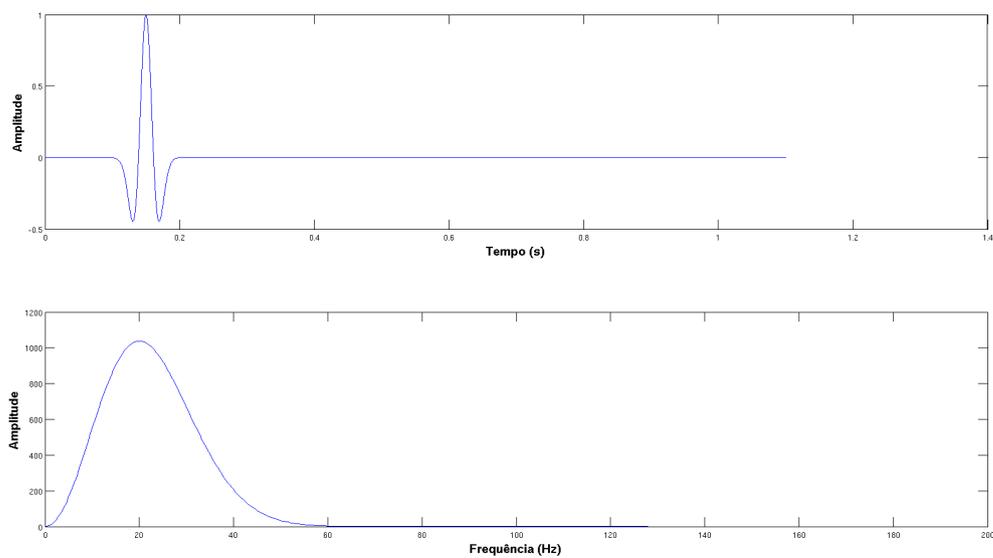
Este capítulo definiu o algoritmo de modelagem utilizado nesta tese. O mesmo utiliza a aproximação acústica da equação da onda no domínio da frequência. De maneira a minimizar o efeito resultante dos contornos artificiais do modelo, foi desenvolvida uma

Figura 12 – Modelo Marmousi



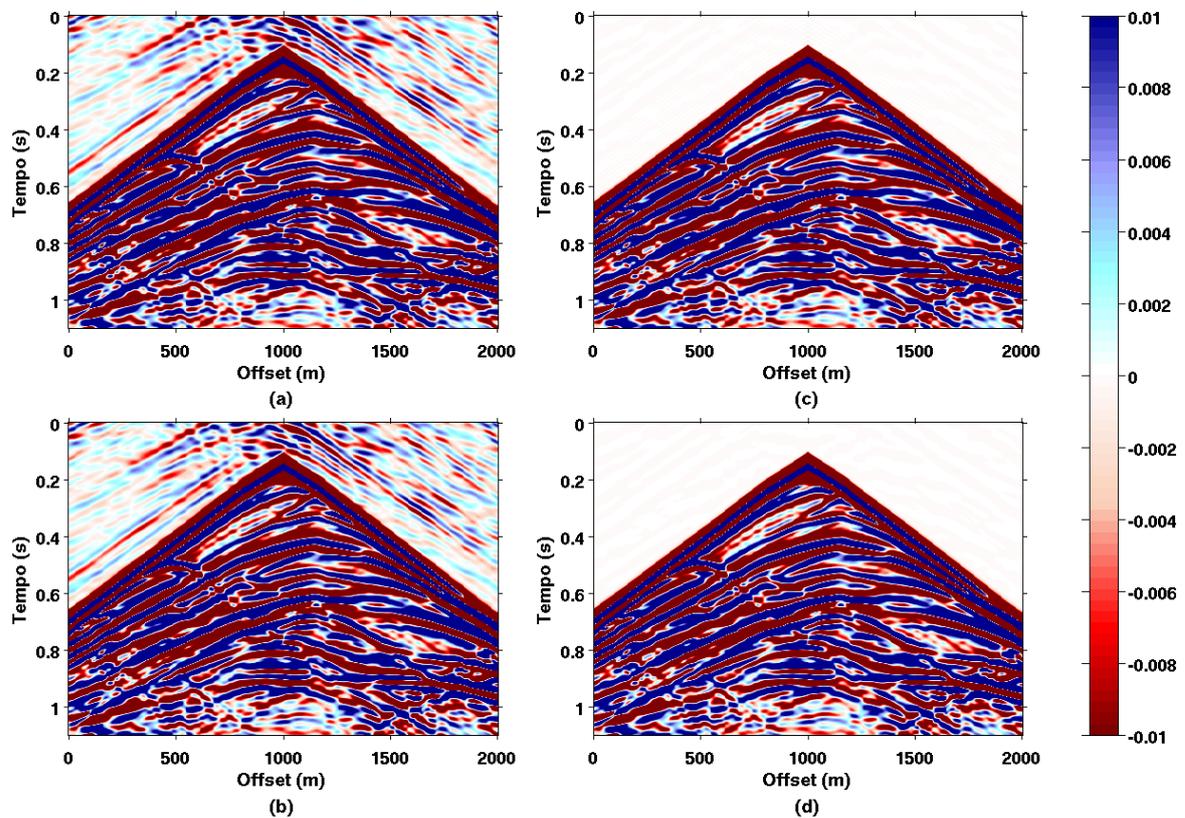
Fonte: Produzida pelo autor

Figura 13 – Fonte Ricker 20 Hz. O gráfico de cima mostra a forma da onda no tempo. O gráfico de baixo apresenta o espectro de amplitude da onda.



Fonte: Produzida pelo autor

Figura 14 – Exemplo de sismograma: a) Frequência máxima simulada de 60 Hz e NFFT = 0. b) Frequência máxima simulada de 60 Hz e NFFT = 1. c) Frequência máxima simulada de 100 Hz e NFFT = 0. d) Frequência máxima simulada de 100 Hz e NFFT = 1



Fonte: Produzida pelo autor

borda de absorção capaz de minimizar a energia refletida pelas mesmas. Em seguida apresentamos um exemplo de uma modelagem para o modelo de Marmousi.

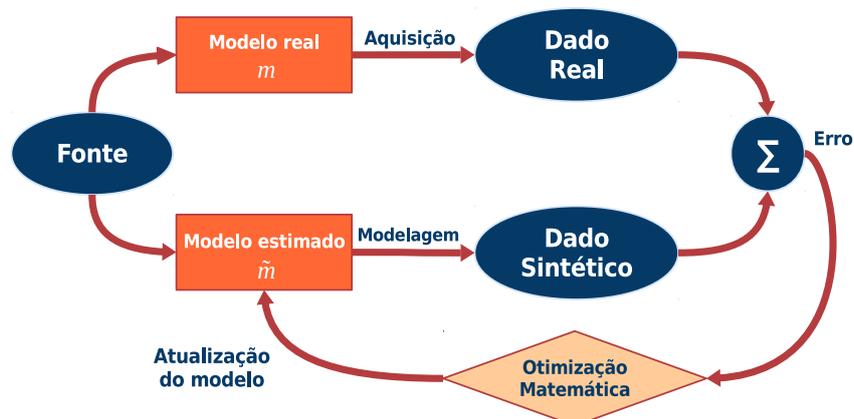
No próximo capítulo vamos estudar a inversão do campo de onda completo (FWI). Utilizando a implementação da modelagem introduzida neste capítulo vamos definir a função objetivo da inversão. Veremos como calcular o gradiente da função objetivo, seu significado físico e como calculá-lo de maneira mais eficiente. Em seguida veremos como aplicar os métodos de cálculo de direção de otimização estudados no Capítulo 2 para o caso da FWI.

4 Inversão do campo de onda acústico

Neste capítulo apresentamos o processo de inversão do campo de onda completo ou *Full Waveform Inversion* (FWI). A FWI é um processo de otimização matemática não linear que busca obter uma aproximação de propriedades físicas de subsuperfície através de medidas indiretas dessas propriedades. No capítulo 2, estudamos a otimização linear como um problema genérico. Vimos que, definida uma função objetivo, podemos encontrar seu ponto de mínimo local iterativamente. No caso da FWI acústica, estudada nesta tese, a propriedade que queremos obter é a velocidade das ondas P (\mathbf{v}_p) e a medida indireta que temos é o campo de pressão medido na superfície, chamado de sismograma. De maneira a generalizar o problema, vamos nos referir ao modelo de velocidades como \mathbf{m} , representando o modelo a ser invertido.

O objetivo da inversão é encontrar um modelo velocidades que melhor aproxime um sismograma sintético, gerado a partir do modelo de velocidades, de um sismograma medido em campo. Para gerar este sismograma sintético utilizamos o algoritmo definido no capítulo 3. Sendo assim podemos esquematizar o processo de inversão conforme a Figura 3, repetida na Figura 15 para facilitar a leitura. Em outras palavras, queremos encontrar um modelo de velocidade que minimize o erro entre os sismogramas medido e calculado.

Figura 15 – Esquema de inversão



Fonte: Produzida pelo autor

Definimos então o resíduo do sismograma, conforme a Equação 4.1

$$\Delta \mathbf{d}(\mathbf{m}) = \mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})_{calc}, \quad (4.1)$$

onde \mathbf{d}_{obs} é o sismograma medido e $\mathbf{d}(\mathbf{m})_{calc} = \mathcal{R}\mathbf{u}$ é o sismograma sintético modelado. \mathbf{u} é o campo de pressão modelado pela Equação 3.5 e \mathcal{R} é o operador de detecção, ou seja, \mathbf{d}_{calc} é o campo de pressão medido em alguns pontos predefinidos do modelo, geralmente próximos à superfície.

O resíduo definido em 4.1 é um vetor representado pelas diferenças, ponto a ponto, dos sismogramas. Para definir a função objetivo, precisamos de uma norma do resíduo. Nesta tese utilizamos a mesma definição de Pratt (1999), ou seja, a norma euclideana ou norma ℓ^2 . Desta forma podemos definir a nossa função objetivo χ , representada na Equação 4.2,

$$\chi(\mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^t \Delta \mathbf{d}^*, \quad (4.2)$$

onde o sobrescrito t indica transposição de um vetor ou matriz e o sobrescrito $*$ indica o complexo conjugado. Uma vez definida a função objetivo, vamos seguir o processo definido no capítulo 2. O próximo passo consiste em calcular o gradiente da função objetivo χ em relação ao modelo \mathbf{m} .

4.1 Gradiente da função objetivo

O gradiente de χ encontra-se definido na Equação 4.3,

$$\nabla_{\mathbf{m}} \chi(\mathbf{m}) = \left(\frac{\partial \mathbf{d}_{calc}}{\partial \mathbf{m}} \right)^t \Delta \mathbf{d}^*. \quad (4.3)$$

De modo a obter uma expressão para o primeiro termo do lado direito da Equação 4.3, primeiro lembramos a relação entre o sismograma medido e o campo de pressão simulado, expressa na Equação 4.4,

$$\mathbf{d}(\mathbf{m})_{calc} = \mathcal{R}\mathbf{u}. \quad (4.4)$$

Desta forma temos que $\partial \mathbf{d} / \partial m_i = \mathcal{R}(\partial \mathbf{u} / \partial m_i)$. Para encontrar $\partial \mathbf{u} / \partial m_i$ aplica-se a derivada parcial em relação a um elemento do modelo m_i nos dois lados da Equação 3.4, como descrito na Equação 4.5,

$$\frac{\partial \mathbf{B}\mathbf{u}}{\partial m_i} = 0, \quad (4.5a)$$

$$\frac{\partial \mathbf{B}}{\partial m_i} \mathbf{u} + \mathbf{B} \frac{\partial \mathbf{u}}{\partial m_i} = 0, \quad (4.5b)$$

$$\frac{\partial \mathbf{u}}{\partial m_i} = -\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m_i} \mathbf{u}. \quad (4.5c)$$

Comparando a Equação 4.5c com a Equação 3.5, podemos afirmar que a derivada do campo de pressão pode ser calculada como uma modelagem onde a fonte utilizada é uma fonte virtual definida na Equação 4.6,

$$\frac{\partial \mathbf{u}}{\partial m_i} = \mathbf{B}^{-1} \mathbf{f}^{(i)}, \quad (4.6a)$$

$$\text{onde, } \mathbf{f}^{(i)} = -\frac{\partial \mathbf{B}}{\partial m_i} \mathbf{u}. \quad (4.6b)$$

O gradiente da função objetivo fica então definido na Equação 4.7,

$$\nabla_{\mathbf{m}} \chi = \Re \{ \mathbf{J}_0 \Delta \mathbf{d}^* \}, \quad \nabla_{\mathbf{m}} \chi = \Re \{ \mathbf{F}^t [\mathbf{B}^{-1}]^t \Delta \mathbf{d}^* \}, \quad (4.7a)$$

onde \mathbf{F} é uma matriz cujas colunas são os termos de fontes virtuais para cada um dos N parâmetros, conforme apresentado na Equação 4.8. $\mathbf{J}_0 = \mathbf{F}^t [\mathbf{B}^{-1}]^t$ define a matriz de sensibilidade ou matriz derivada de Fréchet,

$$\mathbf{F} = \begin{bmatrix} \mathbf{f}^{(1)} & \mathbf{f}^{(2)} & \dots & \mathbf{f}^{(N)} \end{bmatrix}. \quad (4.8)$$

O termo $\partial \mathbf{B} / \partial m_i$ pode ser encontrado analisando a lei de formação da matriz \mathbf{B} apresentada no Apêndice A. Os únicos termos de \mathbf{B} que dependem do modelo encontram-se na diagonal principal de \mathbf{B} e possuem a expressão apresentada na Equação 4.9,

$$B_{i,i} = \left(\frac{\omega}{c_i} \right)^2 - 2 \left(\frac{1}{dx^2} + \frac{1}{dz^2} \right), \quad (4.9)$$

onde c_i é a velocidade de ondas P . Desta forma, basta aplicar a derivada para este termo. Podemos simplificar o cálculo desta derivada ao utilizar como parâmetro o quadrado da vagarosidade $m_i = 1/c_i^2$. Sendo assim, a derivada do termo 4.9 em relação ao quadrado da vagarosidade encontra-se na Equação 4.10,

$$\frac{\partial B_{i,i}}{\partial m_i} = \omega^2. \quad (4.10)$$

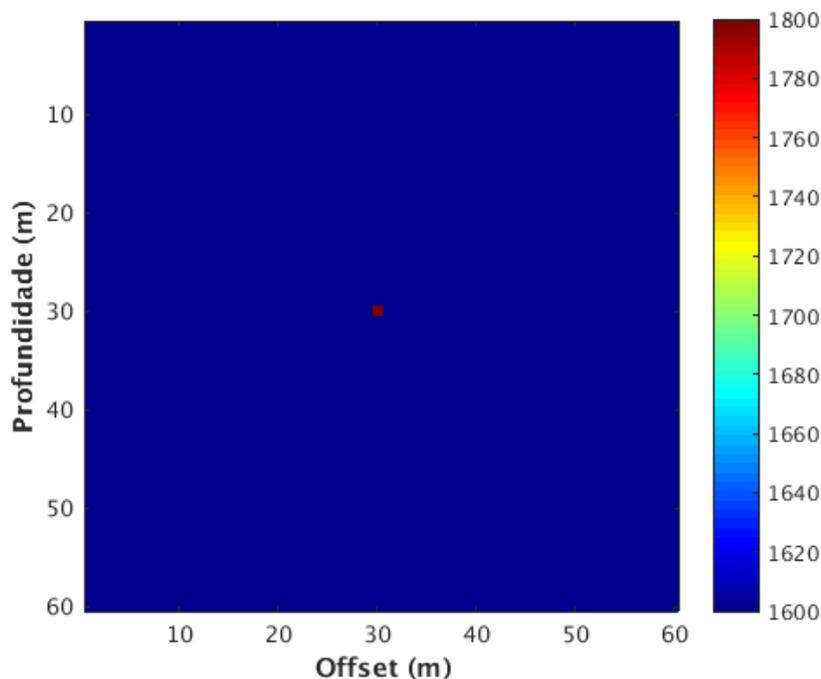
Fazendo $\partial \mathbf{B} / \partial m_i$, obtém-se uma matriz de zeros com apenas um termo não nulo. Desta forma temos a seguinte expressão para uma fonte virtual $\mathbf{f}^{(i)}$, expressa em 4.11,

$$\mathbf{f}^{(i)} = \frac{\partial \mathbf{B}}{\partial \mathbf{m}_i} \mathbf{u} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \omega^2 u_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (4.11)$$

Ou seja, a fonte virtual para um determinado ponto i do modelo está espacialmente centrada neste ponto e temporalmente centrada no tempo de chegada da onda incidente no ponto i .

Para entender melhor a definição do gradiente da função objetivo, podemos utilizar o seguinte exemplo. Considera-se um modelo de um campo de velocidades constante de 1600 m/s com apenas um ponto difrator no centro, com uma variação de velocidade de $12,5\%$. Este modelo encontra-se na Figura 16. A fonte está representada por * enquanto que os receptores são representados por \blacktriangledown . Partindo de um modelo inicial constante de 1600 m/s , busca-se encontrar o gradiente da função objetivo definida pela Equação 4.2.

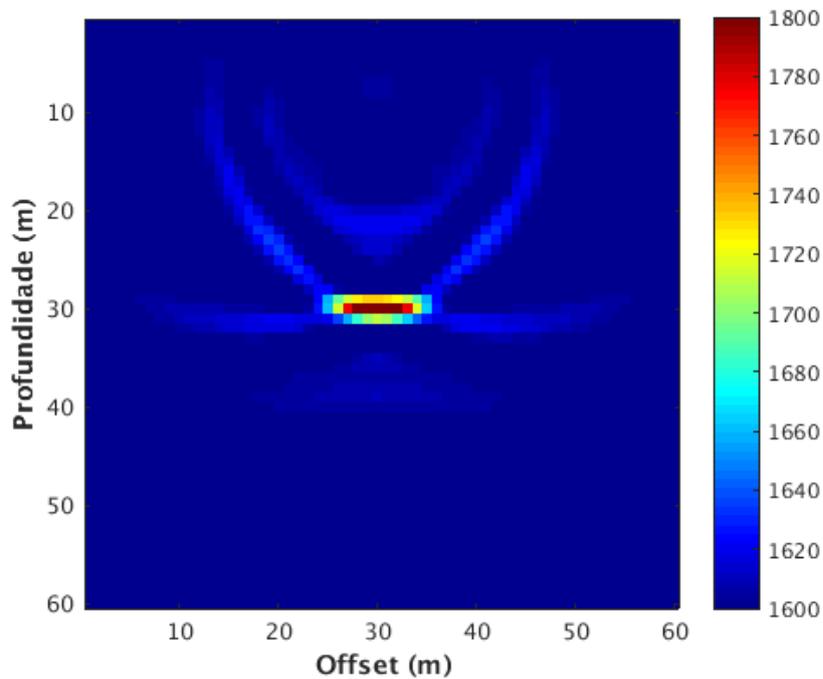
Figura 16 – Modelo de velocidade



Fonte: Produzida pelo autor

A correlação entre o sismograma gerado pela fonte virtual em pontos onde o modelo está correto é muito próxima de zero, enquanto a correlação no ponto da anomalia é próximo de um. Se repetirmos o procedimento acima para todos os pontos do modelo, temos então o nosso gradiente da função objetivo, apresentado na Figura 17.

Figura 17 – Gradiente Anomalia



Fonte: Produzida pelo autor

Resumindo, o cálculo do gradiente a partir das fontes virtuais pode ser definido da seguinte forma:

1. Calcula o sismograma sintético utilizando um modelo inicial
2. Obtém o resíduo entre os sismogramas sintético e observado
3. Para um determinado ponto do modelo, encontramos a derivada de \mathbf{B} e multiplicamos pelo campo simulado. Esta será a fonte virtual para este determinado ponto
4. Propaga a fonte virtual a partir deste ponto e registra nos receptores.
5. A correlação entre o sismograma gerado pela fonte virtual e o resíduo será o valor do gradiente para este ponto.
6. Repete o cálculo acima para cada ponto do modelo.

O procedimento acima, embora de fácil entendimento, apresenta um problema prático. A determinação do gradiente exigiria uma modelagem para cada ponto do modelo,

tornado o processo muito caro computacionalmente. Para modelos muito grandes, esta abordagem é inviável. Podemos reduzir o número total de modelagens através do método adjunto do cálculo do gradiente, apresentado a seguir.

4.2 Método Adjunto

Nesta seção vamos demonstrar como calcular o gradiente da função objetivo a partir do método de estado adjunto e multiplicadores de Lagrange. Nocedal e Wright (2006) apresenta uma revisão de problemas de otimização com restrição e Plessix (2006) apresenta o método adjunto com aplicações em FWI.

Considera-se a função lagrangiana \mathcal{L} correspondente a função objetivo aumentada da restrição de igualdade definida na Equação 4.12,

$$\mathcal{L}(\bar{\mathbf{u}}, \mathbf{m}, \lambda) = \frac{1}{2} \langle \mathcal{P}(d_{obs} - \mathcal{R}\bar{\mathbf{u}}) | (\mathcal{P}(d_{obs} - \mathcal{R}\bar{\mathbf{u}})) \rangle - \langle \lambda | \mathbf{B}(\mathbf{m})\bar{\mathbf{u}} - \mathbf{s} \rangle, \quad (4.12)$$

onde o produto escalar $\langle \cdot | \cdot \rangle$ representa o operador da norma ℓ^2 . \mathcal{P} representa o operador que expande o vetor dos dados medidos com zeros para o tamanho do modelo, de forma que a dimensão do vetor aumentado seja compatível com a dimensão de \mathbf{B} . A restrição de igualdade corresponde a equação do problema direto, ou equação de estado, que deve ser satisfeita pelo campo de onda. A realização \mathbf{u} da equação de estado é chamada de variável de estado. Introduzimos a variável $\bar{\mathbf{u}}$ na Equação 4.12 para distinguir um elemento qualquer do espaço de variáveis de estado de uma realização da equação de estado (PLESSIX, 2006).

O vetor λ , cuja dimensão é a mesma do campo de onda u , é o multiplicador de Lagrange, e corresponde às variáveis de estado adjunto. No âmbito da teoria de otimização com restrições, as condições ótimas de primeira ordem, conhecidas como condições de Karush-Kuhn-Tucker (KKT) dizem que a solução do problema de otimização é obtida em pontos estacionários de \mathcal{L} .

A primeira condição $[\partial\mathcal{L}/\partial\lambda]_{\bar{\mathbf{u}}=cste, \mathbf{m}=cste} = \mathbf{0}$ leva à equação do problema direto $\mathbf{B}(\mathbf{m})\mathbf{u} = \mathbf{s}$. Resolvendo a equação de estado para $\mathbf{m} = \mathbf{m}_0$ temos o campo de onda incidente para FWI.

A segunda condição $[\partial\mathcal{L}/\partial\bar{\mathbf{u}}]_{\lambda=cste, \mathbf{m}=cste} = 0$ leva à chamada equação de estado adjunto, expressa na Equação 4.13,

$$\frac{\partial\mathcal{L}(\bar{\mathbf{u}}, \mathbf{m}, \lambda)}{\partial\bar{\mathbf{u}}} = \mathcal{P}(d_{obs} - \mathcal{R}\bar{\mathbf{u}}) - \mathbf{B}^\dagger(\mathbf{m})\lambda = 0. \quad (4.13)$$

Para derivar a Equação 4.13, utilizamos o fato que $\langle \lambda, \mathbf{B}(\mathbf{m})\bar{\mathbf{u}} \rangle = \langle \mathbf{B}^\dagger(\mathbf{m})\lambda, \bar{\mathbf{u}} \rangle$ e que a fonte não depende de $\bar{\mathbf{u}}$. Escolhendo $\mathbf{m} = \mathbf{m}_0$ e $\bar{\mathbf{u}} = \mathbf{u}(\mathbf{m}_0)$ na Equação 4.13 leva à

Equação 4.14,

$$\lambda^* = \mathbf{B}^{-1}(\mathbf{m}_0)\mathcal{P}(\Delta\mathbf{d}_0^*), \quad (4.14)$$

onde exploramos o fato que \mathbf{B}^{-1} é simétrica em virtude da reciprocidade espacial da função de Green. As variáveis de estado adjunto são calculadas resolvendo um problema direto para uma composição de fontes formada pelo conjugado dos resíduos, o que é equivalente a retro-propagação dos resíduos pelo modelo.

A terceira condição $[\partial\mathcal{L}/\partial\mathbf{m}]_{\bar{\mathbf{u}}=cste,\lambda=cste} = 0$ define o mínimo de \mathcal{L} de maneira comparável à minimização sem restrições da função objetivo χ . Temos assim a Equação 4.15,

$$\frac{\partial\mathcal{L}(\bar{\mathbf{u}}, \mathbf{m}, \lambda)}{\partial\mathbf{m}} = \left\langle \lambda \left| \frac{\partial\mathbf{B}(\mathbf{m})}{\partial\mathbf{m}} \bar{\mathbf{u}} \right. \right\rangle. \quad (4.15)$$

Para qualquer realização do problema direto \mathbf{u} , $\mathcal{L}(\mathbf{u}, \mathbf{m}, \lambda) = \chi(\mathbf{m})$. Logo, a Equação 4.15 nos dá a expressão desejada do gradiente da função objetivo como uma função das variáveis de estado e variáveis de estado adjunto quando $\bar{\mathbf{u}} = \mathbf{u}$, conforme apresentado na Equação 4.16,

$$\frac{\partial\chi}{\partial\mathbf{m}} = \left\langle \lambda \left| \frac{\partial\mathbf{B}(\mathbf{m})}{\partial\mathbf{m}} \mathbf{u} \right. \right\rangle. \quad (4.16)$$

Inserindo λ expresso na Equação 4.14 em 4.16 e escolhendo $\mathbf{m} = \mathbf{m}_0$ nos dá a expressão do gradiente de χ em \mathbf{m}_0 , como explicitado na Equação 4.17,

$$\frac{\partial\chi(\mathbf{m}_0)}{\partial\mathbf{m}} = \mathbf{u}^t(\mathbf{m}_0) \frac{\partial\mathbf{B}^t(\mathbf{m}_0)}{\partial\mathbf{m}} \mathbf{B}^{-1}(\mathbf{m}_0)\mathcal{P}(\Delta\mathbf{d}_0^*), \quad (4.17)$$

onde \mathcal{P} é o operador que aumenta o vetor de resíduos para o mesmo tamanho do modelo. Da mesma forma que fizemos anteriormente, podemos agora resumir o processo da geração do gradiente da função objetivo da seguinte forma:

1. Calcula o sismograma sintético utilizando um modelo inicial
2. Obtém o resíduo entre os sismogramas sintético e observado
3. Propaga o conjugado do resíduo através do modelo inicial, o que equivale a retro-propagar o resíduo.
4. O produto escalar entre o campo retro-propagado e $(\partial\mathbf{B}/\partial\mathbf{m})\mathbf{u}$ será o valor do gradiente da função objetivo.

Podemos perceber que o processo de obtenção do gradiente envolve apenas duas modelagens, independente do tamanho do modelo. O processo descrito acima apresenta o cálculo do gradiente para uma componente de frequência e para uma posição da fonte sísmica. A equação completa do gradiente para todas as posições de fonte e para todas as frequências encontra-se na Equação 4.18,

$$\nabla_m \chi = \sum_{\mathbf{s}} \sum_{\omega} \Re \left\{ \mathbf{u}^t(\mathbf{m}) \frac{\partial \mathbf{B}^t}{\partial \mathbf{m}} \mathbf{B}^{-1} \mathcal{P}(\Delta \mathbf{d}^*) \right\}. \quad (4.18)$$

A Equação 4.18 sugere que o gradiente seja calculado para todo o espectro de frequências do dado. Contudo, a aproximação utilizada nesta tese é baseada na aproximação de Born, que considera que o modelo atualizado \mathbf{m} pode ser descrito como o soma entre o modelo inicial \mathbf{m}_0 e uma perturbação $\Delta \mathbf{m}$. Devido a não linearidade do problema de FWI, o modelo inicial deve ser escolhido próximo o suficiente do modelo real. Caso os erros nos tempos de trânsito observados sejam maiores que metade de um período, ocorrerá o chamado *cycle skipping* e o modelo irá convergir para um mínimo local. Para mitigar este problema, podemos pensar no nível de detalhes que temos no nosso problema para cada componente de frequência. Se pensarmos em um modelo de baixas frequências, o nosso sismograma só será capaz de representar variações lentas de velocidade e com isso, o nosso modelo será mais suave. Naturalmente, o número de mínimos locais diminui neste caso, auxiliando na convergência. Esta estratégia é uma prática bastante comum nos algoritmos de inversão. Sendo assim, ao calcularmos o gradiente, inserimos as componentes de frequência aos poucos, da mais baixa para a mais alta.

4.3 Regularização e pré-condicionamento

Por ser um problema mal-posto, existem infinitos modelos que aproximam os sismogramas. Para tornar o problema mais bem-posto, introduzem-se informações adicionais, na forma de regularização (MENKE, 2012; TARANTOLA, 1987). A função objetivo pode ser aumentada como na Equação 4.19,

$$\chi(\mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^\dagger \mathbf{W}_d \Delta \mathbf{d} + \frac{1}{2} \epsilon (\mathbf{m} - \mathbf{m}_{\text{priori}})^\dagger \mathbf{W}_m (\mathbf{m} - \mathbf{m}_{\text{priori}}), \quad (4.19)$$

onde $\mathbf{W}_d = \mathbf{S}_d^t \mathbf{S}_d$ e $\mathbf{W}_m = \mathbf{S}_m^t \mathbf{S}_m$ são os operadores de pesos. O operador \mathbf{S}_d pode ser implementado como um operador diagonal que controla o peso de cada elemento do vetor de resíduos. Em aplicações em geofísica, é comum utilizar um operador de regularização \mathbf{S}_m de tal forma que penalize grandes variações no modelo de velocidades, geralmente implementado como uma matriz de diferenças de primeira ou segunda ordem.

Este regularizador é chamado de suavidade ou Tikhonov de ordem 1. A perturbação do modelo regularizada encontra-se na Equação 4.20,

$$\Delta \mathbf{m} = -\alpha \mathbf{W}^{-1} \Re \left[\mathbf{J}_0^\dagger \mathbf{W}_d \Delta \mathbf{d}_0 \right]. \quad (4.20)$$

Nesta tese, aplicamos um filtro passa-baixa gaussiano no gradiente da função objetivo para aproximar o regularizador suavidade descrito na Equação 4.20. Conforme a frequência da inversão aumenta, o tamanho do filtro diminui de forma a manter as informações de alta frequência invertidas. Nesta tese não utilizamos pré-condicionamento do dado.

Com o gradiente calculado, agora podemos atualizar o modelo utilizando um dos métodos apresentados no capítulo 2.

4.4 Métodos de otimização matemática em FWI

Conforme vimos anteriormente, vamos utilizar as equações no domínio da frequência para calcular as atualizações do modelo. Utilizando a técnica descrita acima, o modelo inicial é utilizado para a inversão da componente de frequência mais baixa do dado. A saída desta etapa é utilizada como entrada para a frequência subsequente e assim sucessivamente. A frequência inicial e a final dependem da fonte utilizada. Chama-se a isso um rodada completa do algoritmo da FWI. Este processo pode ser repetido diversas vezes até atingir a convergência. Sendo assim, os métodos aplicados nesta tese seguem todos a mesma sequência, definida a seguir:

```

for  $I_{rodada} = 1 : N_{rodada}$ 
  for  $\omega = \omega_i : \omega_f$ 
    Para um dado  $\mathbf{m}_0$ 
    Calcula-se  $\chi_0$  e  $\nabla_m \chi_0$ 
    While  $(\chi_{k-1} - \chi_k) \geq \epsilon$ 
      Calcula a direção de busca  $\Delta \mathbf{m}_k$ 
      Encontra  $\alpha_k$  satisfatório
       $\mathbf{m}_{k+1} \leftarrow \mathbf{m}_k + \alpha_k \Delta \mathbf{m}_k$ 
       $k = k + 1$ 
    end
     $\mathbf{m}_0 \leftarrow \mathbf{m}_k$ 
  end end

```

O cálculo da direção de busca $\Delta \mathbf{m}_k$ segue a mesma forma dos métodos apresentados no capítulo 2. Para a FWI, implementamos três métodos, descritos a seguir.

4.4.1 Gradiente Descendente

No método do gradiente descendente, a direção descendente é o negativo do gradiente no ponto de estudo,

$$\Delta \mathbf{m}_k = -\nabla_m \chi_k. \quad (4.21)$$

4.4.2 Gradiente Conjugado

O método do gradiente conjugado utiliza a mesma definição do capítulo 2, sendo apresentado abaixo.

if $k = 0$

$$\Delta \mathbf{m}_k = -\nabla \chi_k$$

else

$$\beta_{k+1} = \frac{\nabla \chi_{k+1}^T \nabla \chi_{k+1}}{\nabla \chi_k^T \nabla \chi_k}$$

$$\Delta \mathbf{m}_{k+1} = -\nabla \chi_{k+1} + \beta_{k+1} \Delta \mathbf{m}_k$$

end

Para o método gradiente conjugado, caso o passo de iteração α_k não satisfaça certas condições, as direções de busca podem não ser descendentes. Desta forma, torna-se necessário que o cálculo de α_k atenda as condições de Wolfe (2.10). Isto torna o método do gradiente conjugado menos estável, pois diferenças no cálculo de α_k vão resultar em diferenças no cálculo de Δm_{k+1} e estas diferenças serão propagadas ao longo do algoritmo. A utilização de regularizadores minimiza esta instabilidade do método.

4.4.3 Quasinewton

Conforme vimos no capítulo 2, o método de Newton é inviável para problemas muito grandes, devido ao alto custo computacional necessário para calcular e inverter a matriz Hessiana. O método BFGS resolve um destes problemas, estimando diretamente a inversa da Hessiana, porém a necessidade de armazenar a matriz completa continua tornando este método inviável para um problema grande como a FWI. A solução encontrada foi implementar o método quasinewton L-BFGS (Limited-memory BFGS), que é uma aproximação do algoritmo Broyden-Fletcher-Goldfarb-Shanno utilizando uma quantidade limitada de memória computacional. Este método encontra-se descrito em Nocedal e Wright (2006).

Neste método, utiliza-se uma aproximação da direção de busca $\Delta \mathbf{m}_k = \mathbf{H}_k^{-1} \nabla \chi_k$ utilizando L informações passadas. Resultados satisfatórios são obtidos com valores baixos de L entre 3 e 20. O algoritmo é apresentado abaixo.

```

q =  $\nabla \chi_k$ 
for  $i = k - 1, k - 2, \dots, k - l$ 
     $\alpha_i = \rho_i \mathbf{z}_i^T \mathbf{q}$ 
     $\mathbf{q} = \mathbf{q} - \alpha_i \mathbf{y}_i$ 
end(for)
 $\Delta \mathbf{m}_k = \mathbf{G}_k^0 \mathbf{q}$ 
for  $i = k - l, k - l + 1, \dots, k - 1$ 
     $\beta = \rho_i \mathbf{y}_i^T \Delta \mathbf{m}_k$ 
     $\Delta \mathbf{m}_k = \Delta \mathbf{m}_k + \mathbf{z}_i (\alpha_i - \beta)$ 
end(for)

```

Os termos do algoritmo descrito acima encontram-se na Equação 4.22,

$$\rho_k = \frac{1}{\mathbf{y}_k^T \mathbf{z}_k}, \quad (4.22a)$$

$$\mathbf{z}_k = \mathbf{m}_{k+1} - \mathbf{m}_k, \quad (4.22b)$$

$$\mathbf{y}_k = \nabla \chi_{k+1} - \nabla \chi_k, \quad (4.22c)$$

$$\mathbf{G}_k^0 = \frac{\mathbf{z}_{k-1}^T \mathbf{y}_{k-1}}{\mathbf{y}_{k-1}^T \mathbf{y}_{k-1}} \mathbf{I}, \quad (4.22d)$$

$$\mathbf{G} = \mathbf{H}^{-1}. \quad (4.22e)$$

Cada um dos métodos descritos nesta seção possui suas particularidades em termos de custo computacional por iteração, velocidade de convergência e resultado final. Entender as vantagens e limitações de cada um dos métodos é essencial para um entendimento mais profundo da FWI. Devido ao alto custo computacional relacionado a problemas de inversão, é necessário que saibamos determinar a melhor estratégia a ser seguida, incluindo o número máximo de iterações utilizadas para cada componente de frequência e o limite de frequência superior a ser invertido. Nesta tese foi feita uma comparação entre os diferentes métodos de otimização implementados buscando obter informações que nos ajudem a escolher o melhor método para a FWI. Esta comparação encontra-se detalhada no capítulo 5.

Neste momento temos o processo de FWI completamente definido. Uma vez calculada a perturbação do modelo, atualiza-se o mesmo e repete-se o processo até a convergência. Existem na literatura, diversas variações do processo desenvolvido neste capítulo (VIRIEUX; OPERTO, 2009). A título de informação, vamos apresentar a seguir algumas destas variações.

4.5 Critérios de Minimização

A norma dos mínimos quadráticos assume uma distribuição normal da função objetivo (TARANTOLA, 1987). Resultados indesejáveis podem surgir quando, por exemplo, existem ruídos de grande amplitude no dado. Logo, um controle de qualidade rigoroso deve ser feito antes de uma inversão de mínimos quadráticos. Crase et al. (1990) investigaram o uso de diversas normas, entre elas a norma de mínimos quadráticos ℓ^2 , a norma de mínimos valores absolutos ℓ^1 , o critério de Cauchy e o critério da secante hiperbólica (*sech*) em FWI. A norma ℓ^1 especificamente ignora a amplitude dos resíduos ao realizar a retro-propagação dos resíduos ao calcular o gradiente, tornando este critério menos sensível a grandes ruídos no dado. Os critérios de Cauchy e *sech* podem ser considerados como uma combinação das normas ℓ^1 e ℓ^2 , com diferentes transições entre elas. A conclusão obtida por Crase et al. (1990) diz que a inversão mais confiável resulta da utilização do critérios de Cauchy ou de *sech*.

4.6 Métodos de linearização

A matriz de sensibilidade é geralmente calculada utilizando a aproximação de Born, que assume um relação linear tangente entre o modelo e a perturbação do campo de onda (WOODWARD, 1992). Essa relação linear entre as perturbações pode ser inferida assumindo que o campo de onda calculado a partir do modelo atualizado é o campo de onda calculado a partir do modelo inicial somado à perturbação do campo de onda.

A aproximação de Rytov considera a fase generalizada como campo de onda (WOODWARD, 1992). A aproximação de Rytov gera uma relação linear entre a perturbação de fase complexa e a perturbação do modelo assumindo que o campo de onda calculado a partir do modelo atualizado está relacionado ao campo de onda calculado no modelo inicial através de $\mathbf{u}(\mathbf{x}, \omega) = \mathbf{u}_0 \exp(\Delta\phi(\mathbf{x}, \omega))$, onde $\Delta\phi(\mathbf{x}, \omega)$ é a perturbação de fase complexa. A sensibilidade do núcleo de Rytov é zero no caminho de raio de Fermat pois o tempo de trânsito é estacionário ao longo do caminho. Uma relação linear entre a perturbação do modelo e o logaritmo da razão de amplitude $\ln[A(\omega)/A_0(\omega)]$ também é atingida pela aproximação de Rytov ao utilizar a parte real do núcleo da sensibilidade da integral de Rytov ao invés da parte imaginária, que gera a perturbação de fase.

A aproximação do Born é válida no caso de perturbações pequenas e fracas, mas a aproximação de Rytov supostamente é válida para grandes ângulos de abertura e pequenas quantidades de espalhamento por comprimento de onda, isto é, pequenas perturbações ou variações suaves no gradiente de perturbação de fase (BEYDOUN; TARANTOLA, 1988). Embora diversas análises da aproximação de Rytov tenham sido feitas, continua incerto até que ponto seu domínio de validade difere significativamente do domínio de validade da aproximação de Born. Uma comparação entre a aproximação de Born e de

Rytov em FWI elástico no domínio da frequência foi feito por [Gélis, Virieux e Grandjean \(2007\)](#). A principal vantagem da aproximação de Rytov pode ser que esta providencia uma separação natural entre fase e amplitude ([WOODWARD, 1992](#)). Esta separação permite a implementação de inversões de amplitude e de fase ([PYUN; SHIN; BEDNAR, 2007](#); [BEDNAR; SHIN; PYUN, 2007](#)) a partir do código de FWI no domínio da frequência utilizando uma norma logarítmica ([SHIN; MIN, 2006](#); [SHIN; PYUN; BEDNAR, 2007](#)) cujo uso gera uma aproximação de Rytov no contexto de problemas de otimização locais.

4.7 Modelo Inicial

O objetivo final do imageamento sísmico é poder aplicar a FWI com confiança a partir do zero, isto é, sem a necessidade de informações a priori sofisticadas. Infelizmente, como a FWI multidimensional no momento só pode ser abordada como uma minimização local, a construção de um modelo inicial confiável para FWI permanece um dos assuntos mais abordados na literatura.

O modelo inicial para FWI pode ser construído através de tomografia de reflexão e análise de velocidade a partir de migrações como as que são feitas na indústria de óleo e gás. Uma revisão deste processo é dado em ([WOODWARD et al., 2008](#)). Outras abordagens possíveis para a criação de modelos iniciais, que tendem para um procedimento mais automatizado são a tomografia de tempo de trânsito de primeira chegada (*First-Arrival Traveltime Tomography* ou FATT), tomografia estéreo e inversões no domínio Laplace.

A tomografia de tempo de trânsito de primeira chegada (FATT) realiza uma inversão não linear do tempo de trânsito de primeira chegada para produzir modelos suavizados de subsuperfície ([NOLET, 1987](#); [HOLE, 1992](#); [ZELT; BARTON, 1998](#)). Resíduos de tempo de trânsito são retro-propagados ao longo dos caminhos de raios para calcular a matriz de sensibilidade. O sistema tomográfico, aumentado da regularização suavidade são geralmente resolvidos utilizando um algoritmo do tipo gradiente conjugado como o LSQR ([PAIGE; SAUNDERS, 1982](#)). Alternativamente, o método adjunto pode ser aplicado a FATT, evitando construir explicitamente a matriz de sensibilidade, assim como na FWI ([TAILLANDIER et al., 2009](#)). A resolução espacial da FATT é estimada como sendo a largura da primeira zona de Fresnel ([WILLIAMSON, 1991](#)).

Exemplos de aplicação de FWI com dados reais utilizando um modelo inicial construído a partir de FATT são mostrados, por exemplo, em [Ravaut et al. \(2004\)](#), [Operto et al. \(2006\)](#), [Jaiswal et al. \(2008\)](#), [Jaiswal et al. \(2009\)](#) e [Malinowski e Operto \(2008\)](#) para aquisições de superfície.

Diversos testes cegos que correspondem a aquisições de superfície vem sendo atacados por uma inversão conjunta utilizando FATT e FWI. Resultados em escala de exploração de óleo e gás e em escala litosférica são apresentados em [Brenders e Pratt](#)

(2007a), [Brenders e Pratt \(2007b\)](#), [Brenders e Pratt \(2007c\)](#) e sugerem que frequências muito baixas e afastamentos muito grandes são requisitos para obter um resultado confiável de FWI quando o modelo inicial é construído a partir de FATT. Por exemplo, apenas a parte superior do modelo BP benchmark foi imageado com sucesso por [Brenders e Pratt \(2007c\)](#) usando uma frequência inicial tão pequena quanto $0,5\text{Hz}$ e um afastamento máximo de 16Km .

Métodos de tomografia de tempo de trânsito que levam em conta refrações e reflexões costumam produzir modelos iniciais mais consistentes para FWI. Entre estes métodos, a estéreo tomografia é provavelmente uma dos mais promissores pois utiliza informações de inclinação de eventos localmente coerentes junto com um procedimento semiautomático de determinação desses eventos ([LAMBARÉ, 2008](#)). Aplicações de estéreo tomografia em dados sintéticos e reais encontram-se em [Billette e Lambaré \(1998\)](#), [Alerini et al. \(2002\)](#), [Billette et al. \(2003\)](#), [Lambaré e Alerini \(2005\)](#) e ([DüMMONG et al., 2008](#)).

Outro método utilizado para construir um modelo inicial para FWI é a inversão no domínio de Laplace e no domínio de Laplace-Fourier. ([SHIN; CHA, 2008](#); [SHIN; CHA, 2009](#); [SHIN; HA, 2008](#)). A inversão no domínio de Laplace pode ser vista como um inversão do campo de onda no domínio da frequência utilizando frequências complexas cuja parte real é zero e a parte imaginária controla a atenuação do campo de onda no tempo. A informação contida no dado é semelhante a amplitude do campo de onda ([SHIN; CHA, 2008](#)). A inversão no domínio Laplace gera uma reconstrução suavizada do modelo de velocidades, que pode ser o modelo inicial para inversões Laplace-Fourier ou FWI no domínio da frequência.

A inversão Laplace-Fourier é equivalente a inversão de sismogramas atenuados em tempo. O resultado mostrado em [Shin e Cha \(2009\)](#) sugere que este método pode ser aplicado para frequências menores que a menor frequência do espectro da fonte. Uma aplicação em dados reais do golfo do México é apresentada em [Shin e Cha \(2009\)](#). Para aplicações em dados reais, frequências entre 0 e 2 Hz combinadas com nove constantes de atenuação de Laplace foram usadas para inversão. O modelo final foi utilizado como modelo inicial de uma FWI no domínio da frequência padrão.

4.8 Estimativa da Fonte Sísmica

A forma de onda da fonte sísmica geralmente não é conhecida e precisa ser considerada como uma variável do problema ([PRATT, 1999](#)). A *wavelet* da fonte pode ser estimada resolvendo um problema linear pois a relação entre o campo de onda sísmico e a

fonte é linear (Equação 3.4). A solução para a fonte é dada pela Equação 4.23,

$$\mathbf{s} = \frac{\langle \mathbf{g}_{cal}(\mathbf{m}_0) | \mathbf{d}_{obs} \rangle}{\langle \mathbf{g}_{cal}(\mathbf{m}_0) | \mathbf{g}_{cal}(\mathbf{m}_0) \rangle}, \quad (4.23)$$

onde $\mathbf{g}_{cal}(\mathbf{m}_0)$ é a função de Green calculada a partir do modelo inicial \mathbf{m}_0 . A função da fonte pode ser estimada diretamente no algoritmo da FWI uma vez que o campo de onda incidente já foi modelado. A fonte e o meio são atualizados em turnos através das iterações da FWI. É possível aproveitar a estimativa da fonte para projetar uma função objetivo alternativa baseado em otimização diferencial por semelhança (PRATT; SYMES, 2002) ou para definir um critério de parada mais heurístico para inversão (JAISWAL et al., 2009).

Alternativamente, novas funções objetivo foram projetadas de forma que a inversão torne-se independente da função da fonte (LEE; KIM, 2003; ZHOU; GREENHALGH, 2003). A ideia fundamental deste método é normalizar cada sismograma de tiro comum pela soma de todos os sismogramas. Isto remove a dependência do dado normalizado com respeito a fonte e modifica a função objetivo. A desvantagem dessa abordagem é que este requer uma estimativa explícita da matriz de sensibilidade; Os resíduos normalizados não podem ser retro-propagados pois eles não satisfazem a equação da onda.

4.9 Exemplos

Nesta seção vamos apresentar alguns exemplos de FWI, com o objetivo de ilustrar alguns dos conceitos apresentados neste capítulo.

4.9.1 Exemplo 1 - Modelo camadas plano paralelas

O primeiro exemplo consiste de um modelo de 4 camadas plano paralelas. O modelo tem dimensões $800m \times 800m$ com amostragem espacial $\Delta x = 4m$, $\Delta z = 4m$. Todas as camadas possuem $200m$ de largura e as velocidades são, de cima para baixo, $1500m/s$, $3500m/s$, $4500m/s$ e $3000m/s$. A fonte utilizada na modelagem caso foi uma *Ricker* com frequência central de $8Hz$. Esta fonte tem espectro de frequência limitada em aproximadamente $25Hz$. Para a inversão vamos inverter as frequências de $1Hz$ até $25Hz$ sucessivamente, com incrementos de $1Hz$. Foram feitas 10 rodadas completas de inversão. Os parâmetros da modelagem e inversão para este exemplo encontram-se na tabela 6. Vamos estudar aqui a influência do modelo inicial na FWI. Serão utilizados dois modelos iniciais. O primeiro é um modelo próximo do modelo real, obtido através da filtragem do modelo real por um filtro média móvel. O segundo modelo inicial é um modelo homogêneo de velocidade $1500m/s$. Os resultados encontram-se nas Figuras 18 e 19.

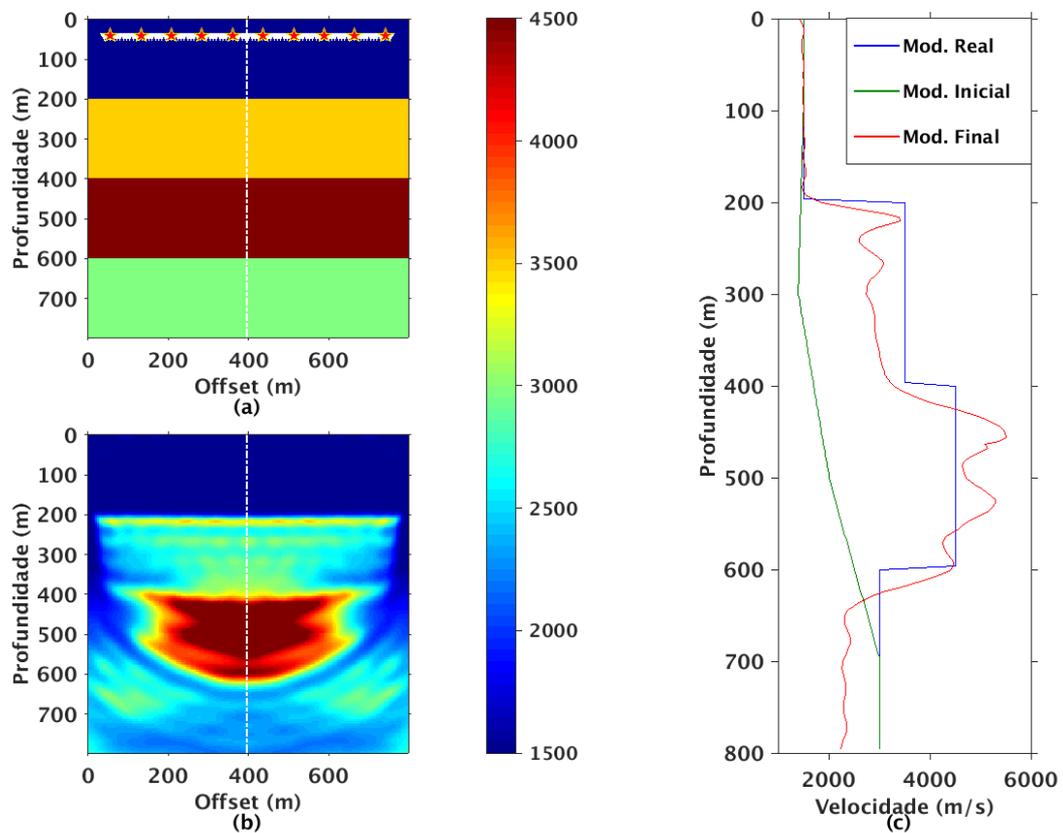
Observando o resultado, podemos ver claramente a influência do modelo inicial no resultado da inversão. Por ser um exemplo com variações bruscas entre as camadas, é

Tabela 6 – Parâmetros da Modelagem

Parâmetro	Valor
Δx	4m
Δy	4m
Δt	$2 \times 10^{-5} s$
Tempo total	1, 1s
Nx	800m
Nz	800m
Fonte sísmica	Ricker 8Hz
Nº rodadas FWI	10

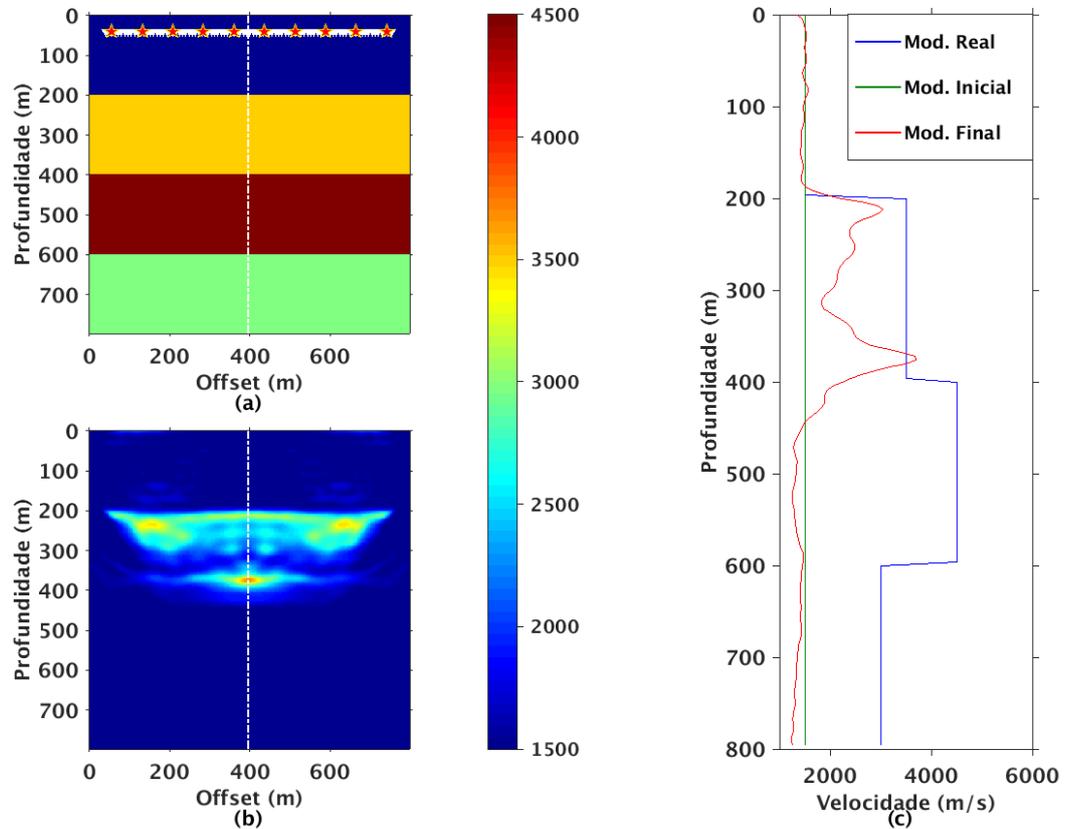
Fonte: Produzida pelo autor

Figura 18 – Resultado da inversão para modelo de 4 camadas. Modelo inicial suavizado. Os traços pontilhados em (a) e (b) indicam a localização do traço em (c). a) Modelo Real. b) Modelo invertido. c) Comparação para um traço do modelo.



Fonte: Produzida pelo autor

Figura 19 – Resultado da inversão para modelo de 4 camadas. Modelo inicial homogêneo. Os traços pontilhados em (a) e (b) indicam a localização do traço em (c). a) Modelo Real. b) Modelo invertido. c) Comparação para um traço do modelo.



Fonte: Produzida pelo autor

natural que o resultado não fique perfeito para nenhum dos casos, pois a resolução espacial da FWI depende da resolução do sismograma. Mesmo assim vemos que para o modelo suave, a FWI foi capaz de melhorar o modelo de velocidades, identificando melhor as camadas e suas interfaces. Já para o modelo homogêneo, vemos que a FWI não foi capaz de convergir para um modelo próximo do modelo real. Percebe-se que a inversão identificou apenas as interfaces entre as camadas, mantendo um artefato de baixa velocidade entre as interfaces.

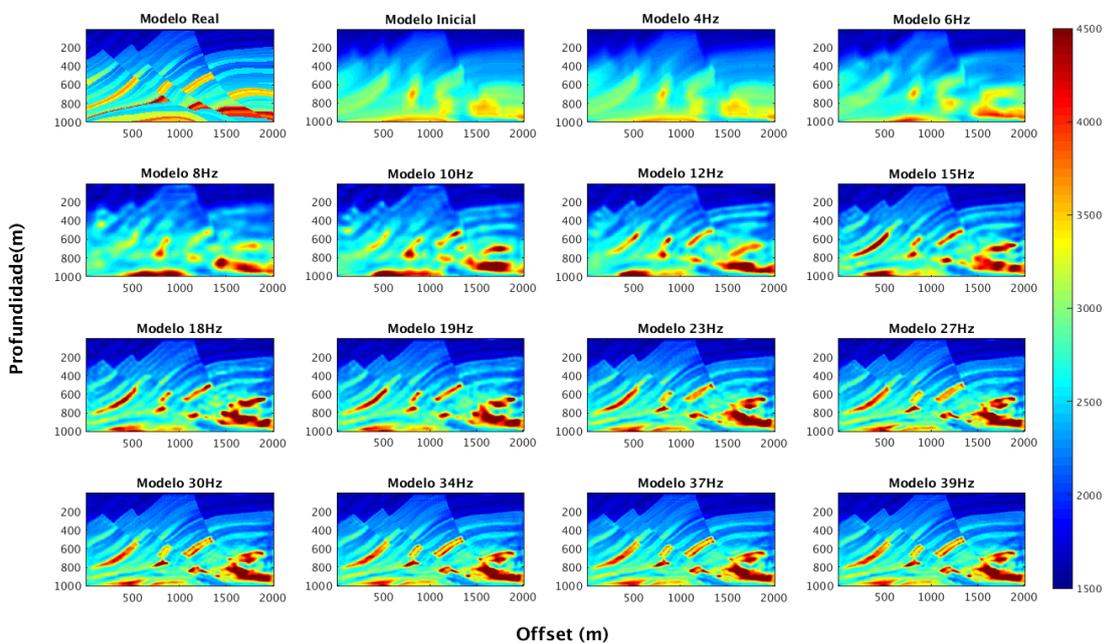
4.9.2 Exemplo 2 - Modelo Marmousi

O segundo exemplo utiliza o modelo de Marmousi como modelo de velocidades real. Com este modelo, podemos estudar a influência de alguns parâmetros da inversão, como por exemplo o número de tiros, o espectro de frequência da fonte sísmica, e o modelo inicial. O modelo de Marmousi utilizado possui um tamanho de $1000 \times 2004m$, discretizado por um grid quadrado de $4m \times 4m$. Para a simulação, utilizamos um valor baixo de amostragem

temporal, de $2 \times 10^{-5} s$ para minimizar os efeitos de dispersão numérica. Simulamos 1, 1s de propagação.

Primeiramente vamos ilustrar a evolução do modelo invertido ao longo do processo de FWI. O resultado da inversão para algumas frequências invertidas encontra-se na Figura 20 para um fonte do tipo *Ricker* de frequência central $20Hz$. A primeira figura (canto superior esquerdo) mostra o modelo real. A figura seguinte (a direita) mostra o modelo inicial. Podemos ver a evolução do modelo invertido (da esquerda para a direita e de cima para baixo). Podemos perceber que a medida que a frequência aumenta, a definição do modelo invertido aumenta. O modelo final é visivelmente muito mais próximo do modelo real que o modelo inicial. Podemos ver isso melhor comparando as diferenças entre os modelos iniciais e finais contra o modelo real, na Figura 21 para a *Ricker* de $20Hz$. Podemos perceber pelas diferenças que a o modelo final representa muito bem o modelo real, especialmente na parte mais rasa do modelo. Podemos ver também que nas interfaces das camadas ainda existe um resíduo. Isto se deve pela limitação da frequência da fonte. A inversão só é capaz de representar aquilo que pode ser medido no sismograma.

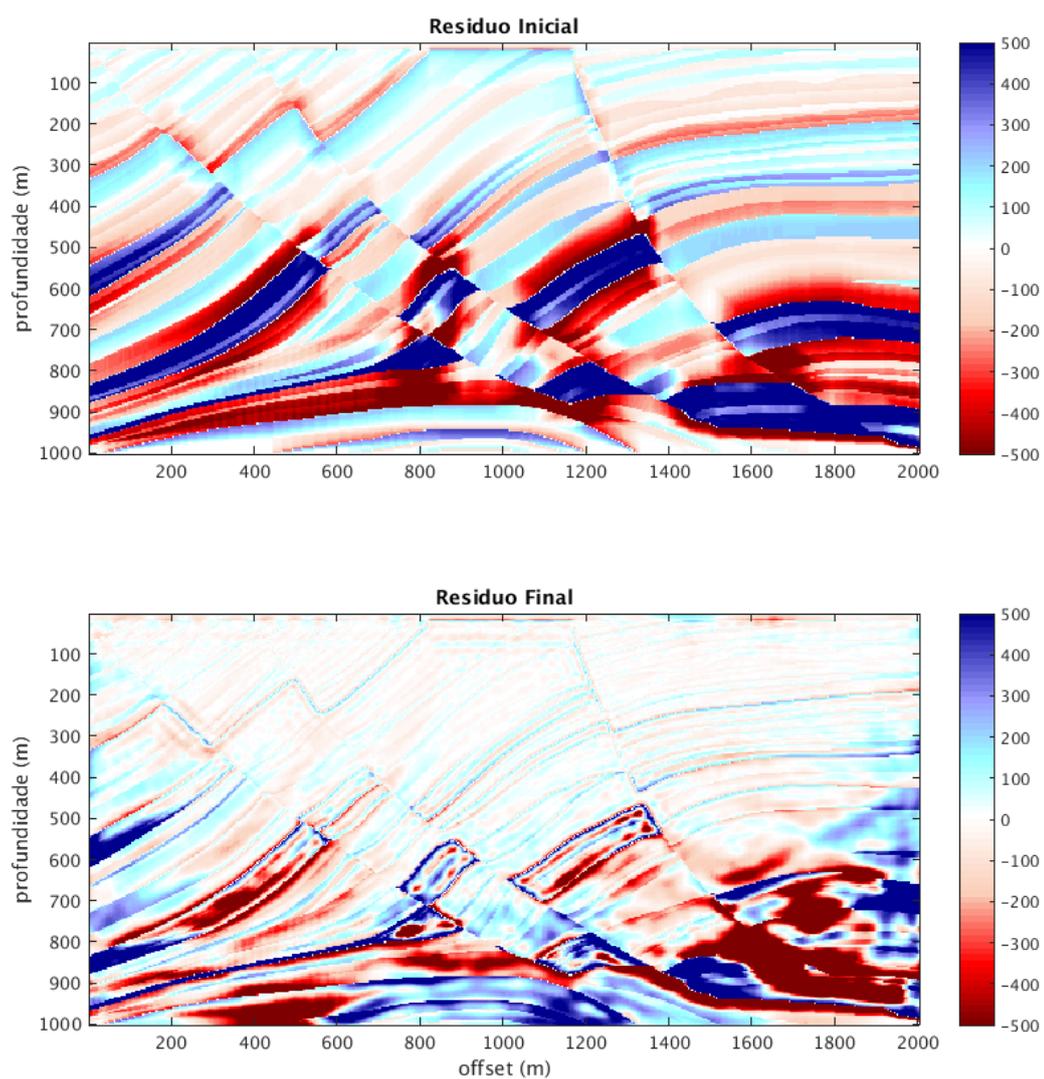
Figura 20 – Evolução do modelo invertido para um fonte *Ricker* de frequência central de 20 Hz.



Fonte: Produzida pelo autor

A seguir vamos analisar a influência de alguns parâmetros no resultado da inversão.

Figura 21 – Resíduo do modelo invertido para um fonte *Ricker* de frequência central de 20 Hz.

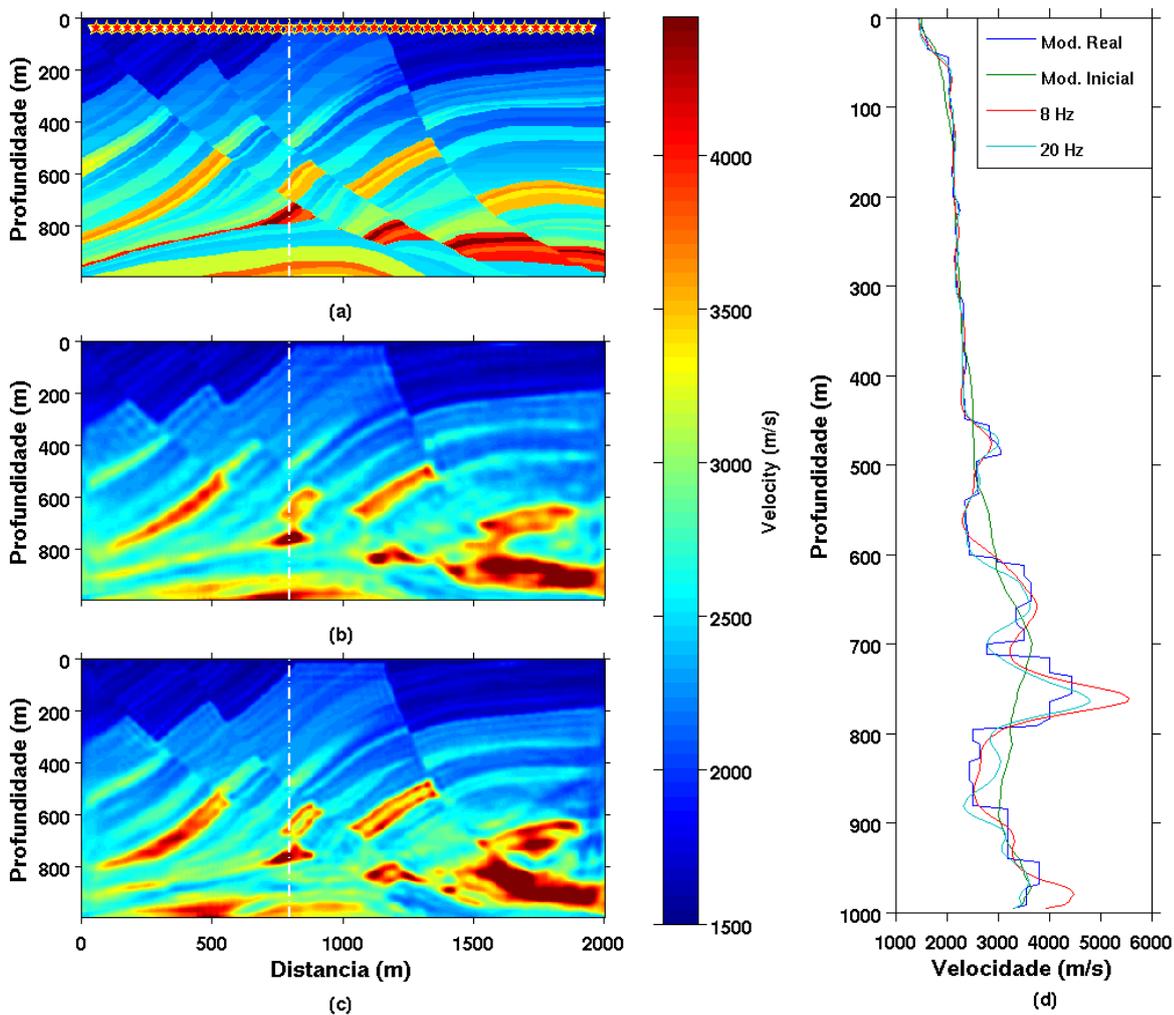


Fonte: Produzida pelo autor

4.9.2.1 Frequência da fonte

Para demonstrar a influência do espectro de frequência da fonte no resultado da inversão, vamos comparar a inversão para duas fontes com frequências centrais diferentes. No primeiro caso vamos usar uma fonte *Ricker* com frequência central de 8Hz . O segundo exemplo será o de uma fonte *Ricker* com frequência central de 20Hz . Naturalmente a fonte de 20Hz possui um espectro de frequências mais amplo que a de 8Hz . Isto permite que a inversão possa prosseguir até uma frequência mais alta. O resultado final encontra-se na Figura 22. Podemos perceber que as interfaces estão mais bem definidas para a *Ricker* de 20Hz . Para ambos os testes, foram utilizados 47 tiros distribuídos ao longo da superfície do modelo (1 tiro a cada 40m).

Figura 22 – Comparação da FWI para diferentes frequências centrais da fonte sísmica. a) Modelo Real. b) Modelo invertido utilizando uma *Ricker* de 8Hz . c) Modelo invertido utilizando uma *Ricker* de 20Hz . d) Comparação para um traço do modelo.

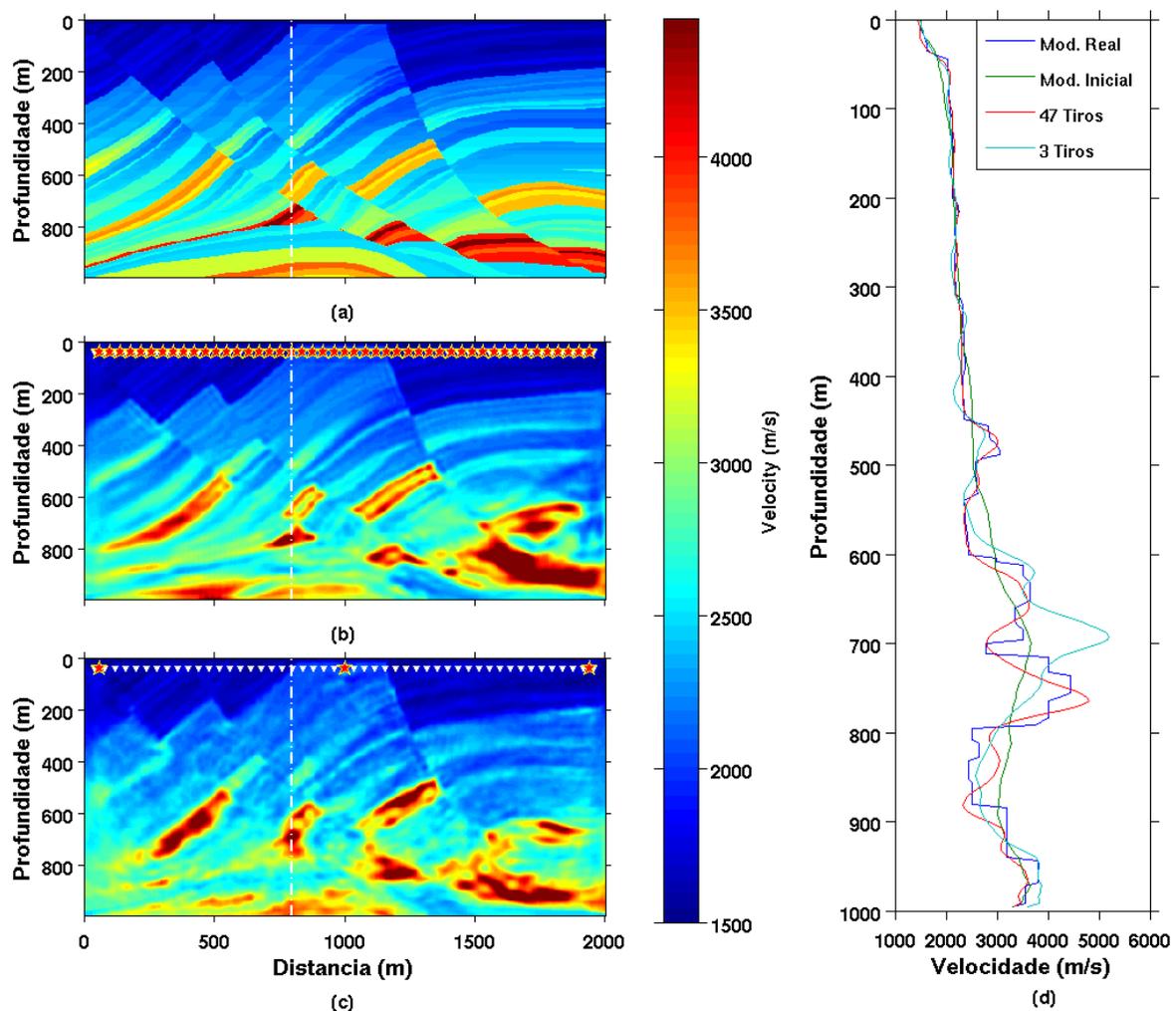


Fonte: Produzida pelo autor

4.9.2.2 Número de tiros

Em seguida vamos analisar a influência do número de tiros na inversão. Vamos utilizar como exemplo uma inversão utilizando 47 tiros (1 tiro a cada 40m) comparando com uma inversão utilizando apenas 3 tiros. Ambos testes foram modelados utilizando como fonte uma *Ricker* de 20Hz. O número de tiros corresponde a redundância de informação para cada ponto de sub superfície. Quanto maior o número de tiros, maior a informação para um determinado ponto, pois este será imageado através de diferentes afastamentos. O resultado encontra-se na Figura 23. Naturalmente, podemos perceber que o resultado da inversão melhora consideravelmente quando aumentamos o número de tiros. Este exemplo demonstra a importância de uma geometria de aquisição apropriada para a FWI.

Figura 23 – Comparação da FWI para um número diferente de fontes sísmicas. a) Modelo Real. b) Modelo invertido utilizando 47 fontes. c) Modelo invertido utilizando 3 fontes. d) Comparação para um traço do modelo.

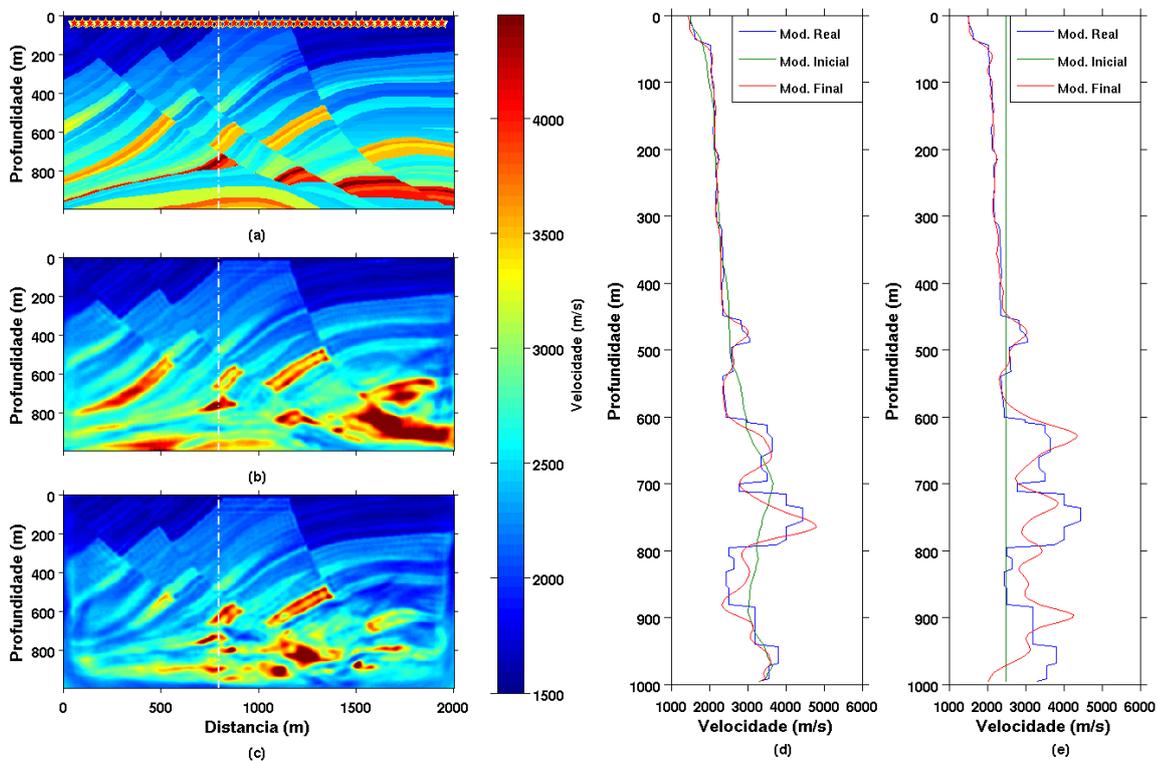


Fonte: Produzida pelo autor

4.9.2.3 Modelo Inicial

No exemplo anterior vimos a influência do modelo inicial na FWI. Vamos analisar novamente a influência do modelo inicial no resultado da inversão para o modelo de Marmousi. Vamos utilizar como modelos iniciais um modelo real filtrado por um filtro média móvel comparado a um modelo inicial homogêneo. Os testes foram realizados utilizando uma geometria de aquisição com 47 tiros e uma fonte sísmica do tipo *Ricker* de $20Hz$. O resultado encontra-se na Figura 24. Podemos perceber novamente que o modelo suave obteve um resultado melhor que o modelo homogêneo, embora neste caso o modelo homogêneo também tenha sido capaz de atingir um resultado satisfatório. Isto acontece pois o modelo Marmousi é um modelo que apresenta pequenas variações de velocidade entre as camadas. Neste caso o modelo homogêneo encontra-se próximo o suficiente do modelo real permitindo a convergência.

Figura 24 – Comparação da FWI para diferentes modelos iniciais. a) Modelo Real. b) Modelo invertido utilizando um modelo inicial suavizado. c) Modelo invertido utilizando um modelo inicial homogêneo. d) Comparação para um traço do modelo, para o modelo inicial suavizado. e) Comparação para um traço do modelo, para o modelo inicial homogêneo.



Fonte: Produzida pelo autor

4.10 Conclusão

Este capítulo definiu o algoritmo de FWI desenvolvido nesta tese. Mostramos como definir a função objetivo para o problema da sísmica e como calcular o gradiente da mesma. Em seguida definimos os métodos de cálculo de direção de busca utilizados na otimização não linear. Finalmente apresentamos a regularização utilizada de forma a estabilizar o processo de inversão.

No próximo capítulo, mostraremos os resultados obtidos com o algoritmo de inversão aqui apresentado. Apresentaremos diversos testes e analisaremos seus resultados, proporcionando assim um maior entendimento das características da inversão do campo de onda completo.

5 Resultados

No capítulo anterior, implementamos o algoritmo de inversão do campo de onda completo. Neste capítulo, apresentaremos os resultados obtidos utilizando o algoritmo de FWI acústico no domínio da frequência. Será apresentado um estudo da influência da borda de absorção utilizada no algoritmo de modelagem no resultado final da inversão. Apresentaremos também uma comparação entre os diversos métodos de otimização matemática, estudando a sensibilidade de cada método a ruído e a diferentes modelos iniciais.

5.1 Metodologia

Conforme foi visto no Capítulo 4, para garantir uma boa convergência da FWI, é necessário tomar alguns cuidados no que diz respeito a representação do dado medido. Uma boa estimativa da forma de onda da fonte sísmica e um algoritmo de modelagem capaz de representar a propagação de onda no meio real são fundamentais para que o resultado da inversão seja satisfatório. Neste trabalho estamos interessados em estudar a influência de alguns parâmetros isolados da FWI no resultado final. Desta forma, vamos simular dados reais modelando a propagação de uma fonte sísmica conhecida através de um modelo de velocidades conhecido utilizando o algoritmo de modelagem acústica no domínio da frequência apresentado no Capítulo 3. A mesma fonte e o mesmo algoritmo de modelagem será utilizado na FWI. A borda de absorção utilizada na criação do dado "real" foi a borda híbrida desenvolvida neste trabalho.

Serão apresentados a seguir dois estudos de caso. O primeiro compara diferentes bordas de absorção dentro do algoritmo de FWI. O segundo compara as diferentes técnicas de otimização matemática apresentadas no Capítulo 4.

5.2 Caso Exemplo 1 - Borda de absorção

O primeiro exemplo estuda o efeito da borda de absorção utilizada no algoritmo de modelagem no resultado da inversão.

5.2.1 Metodologia

Os parâmetros da simulação encontram-se na Tabela 7. O modelo real de velocidade utilizado foi o Marmousi, com 4 m de amostragem espacial. A taxa de amostragem utilizada foi 2×10^{-5} s. A fonte utilizada foi uma *Ricker* de 20 Hz de frequência central (3.10).

Foram modelados 47 tiros ao longo da superfície, representando um tiro a cada 40 m aproximadamente. A borda de absorção utilizada na modelagem foi tipo híbrida de $N_B = 10$ e $C_{PML} = 100$. As frequências modeladas foram de 0 até 60 Hz. A inversão foi realizada invertendo frequências de 4 a 50 Hz sucessivamente, com incremento $1Hz$. Foi feita uma rodada completa de inversão.

Tabela 7 – Parâmetros de Simulação - Caso Exemplo 1

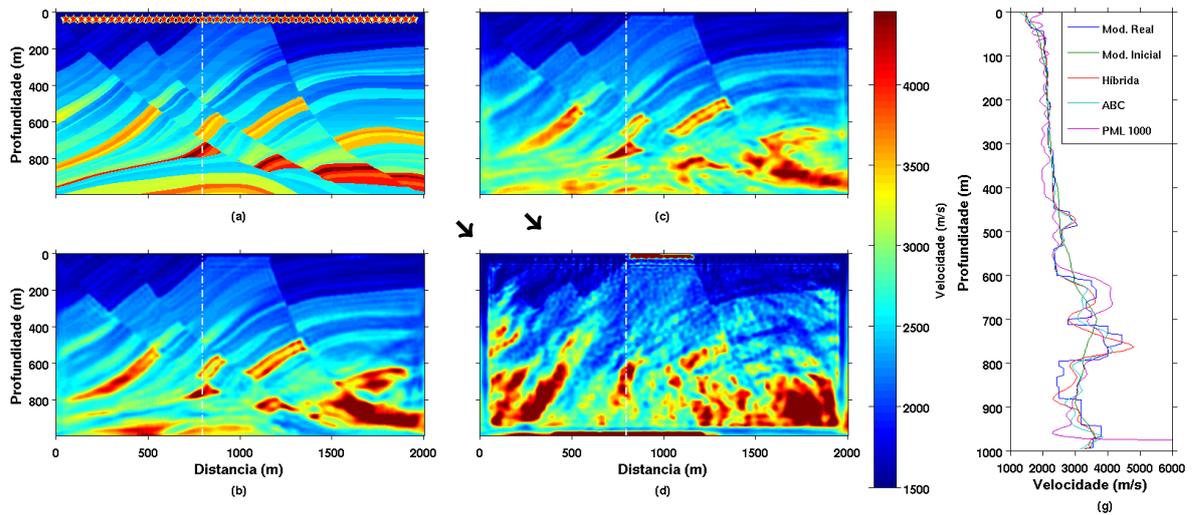
Parâmetros Modelagem	
Parâmetro	Valor
Δx	4 m
Δy	4 m
Δt	$2 \cdot 10^{-5}$ s
Tempo total	1,1 s
N_x	2004 m
N_z	1000 m
Fonte sísmica	Ricker 20 Hz
Parâmetros Inversão	
Nº rodadas FWI	1
Frequências Invertidas	4 a 50 Hz

Fonte: Produzida pelo autor

5.2.2 Resultado

Para o primeiro exemplo, vamos comparar o efeito de diferentes bordas de absorção utilizadas no algoritmo de modelagem na FWI. Vamos comparar uma borda de absorção ABC, com uma borda PML mal dimensionada, ou seja, com um valor excessivamente alto de C_{PML} , e com a borda híbrida desenvolvida neste projeto. A Figura 25 mostra o resultado do modelo invertido para cada um dos casos. A Figura 25a mostra o modelo real. As Figuras 25b-d mostram o resultado da inversão para cada uma das bordas utilizadas na FWI. A Figura 25e mostra a comparação de todas as bordas para um perfil de velocidades correspondente à linha tracejada nas imagens. Podemos perceber claramente que a borda PML mal dimensionada gerou artefatos no modelo de velocidade invertido muito fortes, devido as reflexões espúrias causadas pela borda mal dimensionada. A borda ABC, embora tenha obtido um resultado melhor do que a borda PML, também é possível perceber alguns artefatos no modelo causado pelas reflexões da borda, identificados pelas setas em 25c. Para a borda híbrida, não percebemos nenhum artefato. Sabemos que neste caso, esta

Figura 25 – Resultado da inversão para diferentes bordas de absorção. Os traços pontilhados em branco em (a)-(d) indicam a posição do traço apresentado em (e). As setas pretas em (c) indicam pontos onde o modelo de velocidades está pior que em (b). a) Modelo real. \star indica a posição das fontes simuladas. ∇ indica a posição dos receptores. b) Resultado utilizando a borda híbrida. c) Resultado utilizando a borda ABC. d) Resultado utilizando a borda PML. e) Comparação de todas as bordas para um traço.



Fonte: Produzida pelo autor

nova borda reduz consideravelmente reflexões espúrias das bordas, tornando-a a melhor borda para ser utilizada na inversão no domínio da frequência.

5.3 Caso exemplo 2 - Comparação dos métodos de otimização matemática

O segundo exemplo compara os diferentes tipos de otimização matemática. Cada exemplo foi invertido utilizando o método do gradiente descendente, o método do gradiente conjugado e o método L-BFGS. Estudamos como os métodos se comportam com a presença ou não de ruído no dado modelado e com diferentes modelos iniciais.

5.3.1 Metodologia

Os parâmetros da simulação encontram-se na Tabela 8. O modelo de velocidade utilizado foi o Marmousi, com 4 m de amostragem espacial. A taxa de amostragem utilizada foi $2 \cdot 10^{-5}$ s. A fonte utilizada foi uma *Ricker* de 8 Hz de frequência central. Foram modelados 47 tiros ao longo da superfície, representando um tiro a cada 40 m aproximadamente. A borda de absorção utilizada na modelagem foi a tipo híbrida de

$N_B = 10$ e $C_{PML} = 100$. As frequências modeladas foram de 0 até 30 Hz. A inversão foi realizada invertendo frequências de 1 a 25 Hz sucessivamente, com incrementos de 1 Hz. Foi feita uma rodada completa de inversão. Para alguns exemplos, inserimos um ruído rosa nos dados medidos. O Ruído Rosa ou Ruído de $1/f$ é um sinal ou um processo onde o espectro de frequências como a densidade espectral de potência é inversamente proporcional à frequência do sinal. O espectro de frequências de um ruído rosa encontra-se na Figura 26.

Tabela 8 – Parâmetros de Simulação - Caso exemplo 2

Parâmetros de Modelagem	
Parâmetro	Valor
Δx	4 m
Δy	4 m
Δt	$2 \cdot 10^{-5}$ s
Tempo total	1,1 s
N_x	2004 m
N_z	1000 m
Fonte sísmica	Ricker 8 Hz
Parâmetros de Inversão	
Nº rodadas FWI	1
Frequências invertidas	1 a 25 Hz

Fonte: Produzida pelo autor

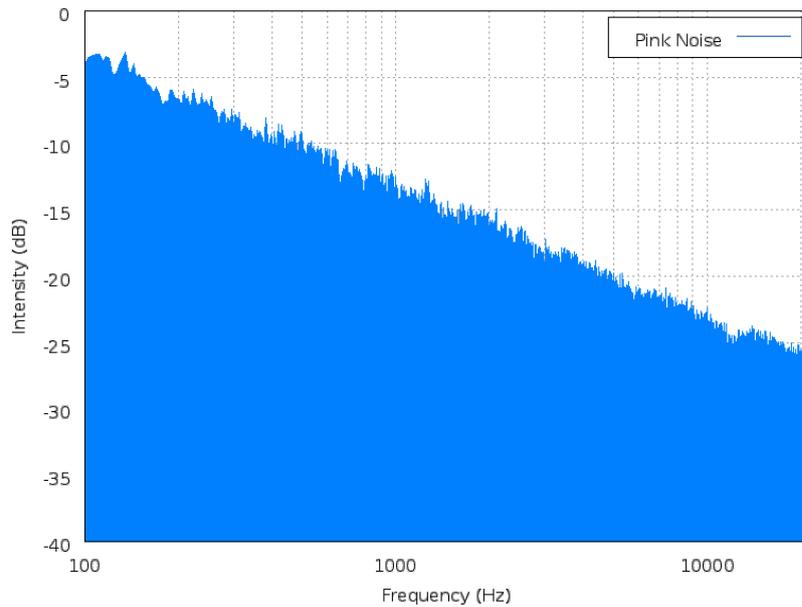
5.3.2 Resultados

A comparação entre os métodos será feita a partir de 4 exemplos, descritos a seguir:

- A - Modelo inicial suavizado sem ruído.
- B - Modelo inicial homogêneo sem ruído.
- C - Modelo inicial suavizado com ruído.
- D - Modelo inicial homogêneo com ruído.

Para cada um dos exemplos vamos demonstrar o resultado a partir de três figuras. A primeira mostra a evolução do erro para cada frequência invertida. Mostramos isso

Figura 26 – Espectro de frequência - Ruído Rosa

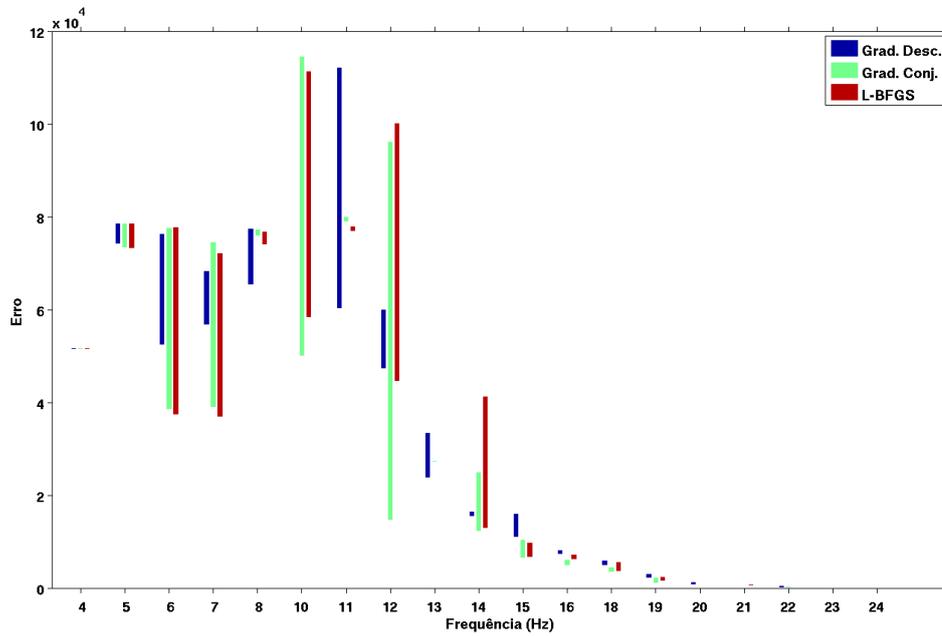


Fonte: <https://pt.wikipedia.org/wiki/Ru%C3%ADdo_rosa>

através de um gráfico de barras. Cada barra indica a variação do erro para cada frequência. O erro é definido conforme a Equação 4.2, ou seja, a soma das normas l_2 dos resíduos para todas as fontes. A extremidade superior da barra indica o erro inicial para uma determinada frequência, enquanto que a extremidade inferior indica o erro final. Este gráfico nos permite avaliar a evolução de cada método. Por exemplo, caso algum método não tenha sido capaz de realizar uma inversão muito eficiente para uma determinada frequência, a barra será curta, indicando que o método não foi capaz de atualizar o modelo consideravelmente. Outro ponto importante de notar neste gráfico é a relação dos erros na função objetivo com a amplitude do espectro de frequência da fonte. A segunda figura mostra o número total de modelagens realizado por cada método para cada frequência. Esta métrica é uma boa maneira de medir a velocidade de cada método. Durante a inversão, o maior custo computacional é o cálculo da modelagem direta. Como vimos no capítulo 4, para cada cálculo do gradiente, necessitamos de duas modelagens. Para o cálculo do passo de iteração ótimo, realizamos mais algumas modelagens até encontrarmos o passo que atenda as condições de Wolfe. Desta forma, o número total de modelagens nos dá uma métrica mais real do que somente o número total de iterações para cada frequência. A terceira figura mostra o resultado do modelo de velocidade invertido. Nela podemos ver o resultado final da inversão.

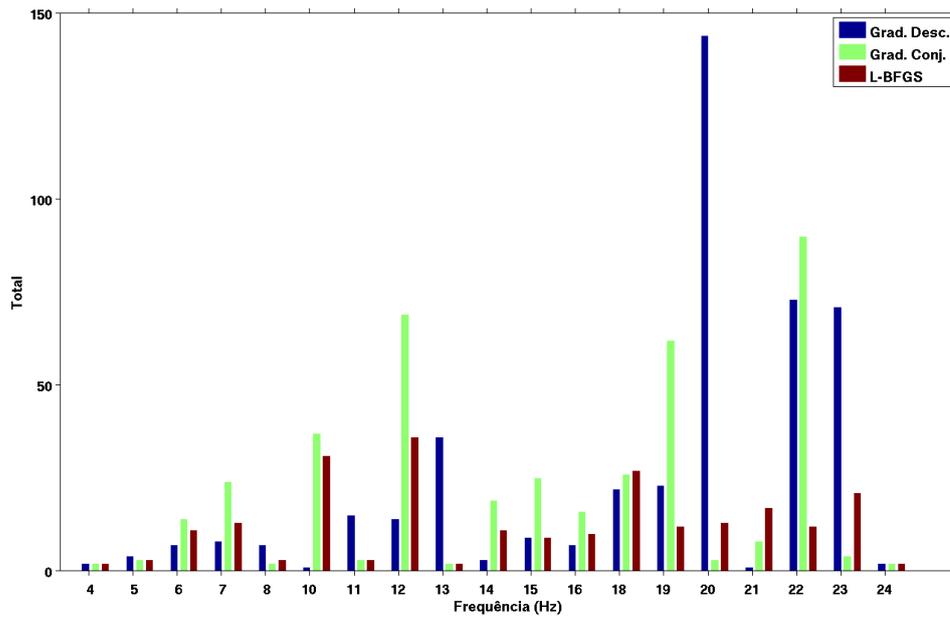
Para o primeiro exemplo, utilizamos como modelo inicial o modelo Marmousi suavizado e os sismogramas medidos sem a presença de ruído. A evolução do erro encontra-se na Figura 27, enquanto que o total de iterações encontra-se na Figura 28. O total

Figura 27 – Progressão de Erro - Exemplo A



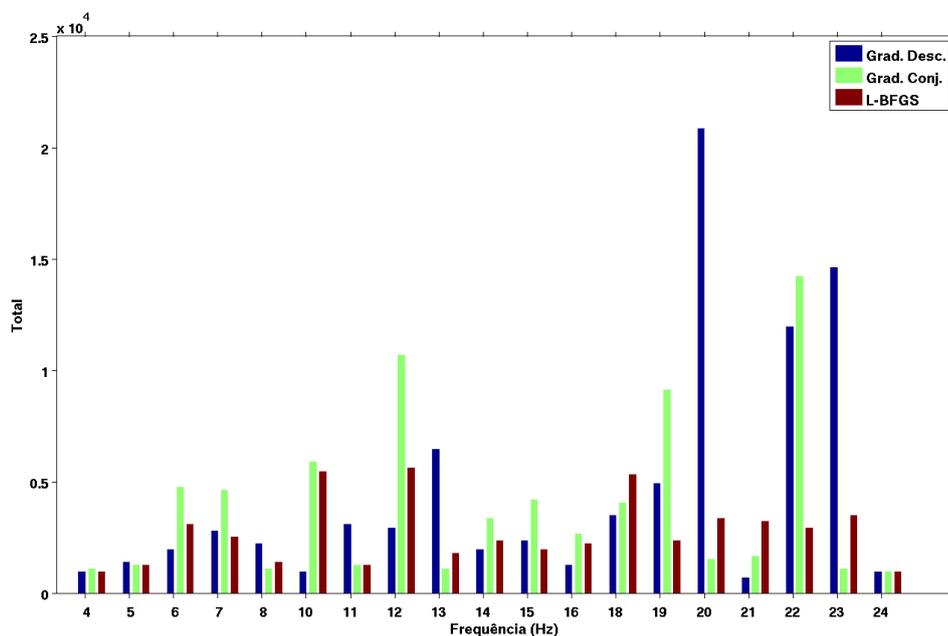
Fonte: Produzida pelo autor

Figura 28 – Total de Iterações - Exemplo A



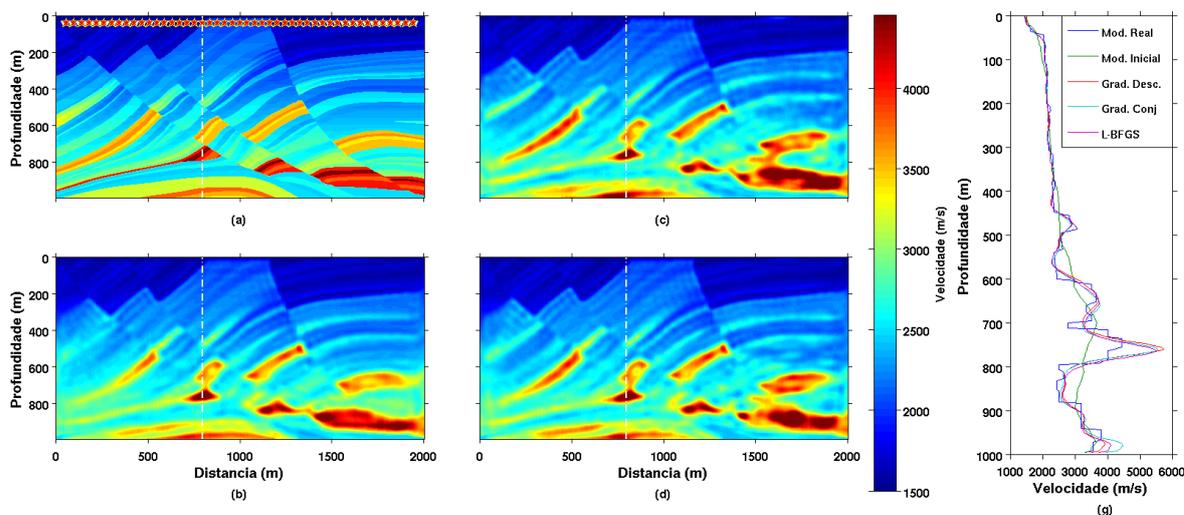
Fonte: Produzida pelo autor

Figura 29 – Total de Modelagens - Exemplo A



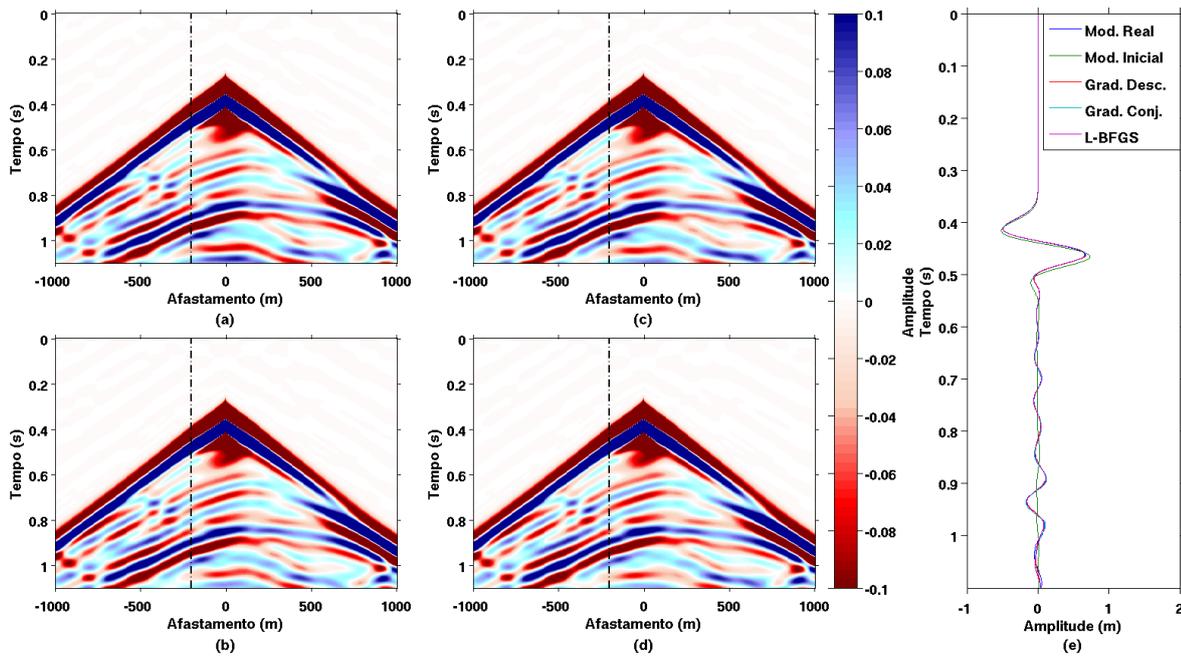
Fonte: Produzida pelo autor

Figura 30 – Resultado da inversão - Exemplo A. Os tracos pontilhados em branco em (a)-(d) indicam a posição do traço apresentado em (e). a) Modelo real. \star indica a posição das fontes simuladas. ∇ indica a posição dos receptores. b) Resultado Gradiente Descendente. c) Resultado Gradiente Conjugado. d) Resultado L-BFGS. e) Comparação para um traço do modelo.



Fonte: Produzida pelo autor

Figura 31 – Sismogramas - Exemplo A. Os tracos pontilhados em (a)-(d) indicam a posição do traço apresentado em (e). a) Sismograma medido. b) Sismograma - Gradiente Descendente. c) Sismograma - Gradiente Conjugado. d) Sismograma - L-BFGS. e) Comparação para um traço do sismograma.



Fonte: Produzida pelo autor

Tabela 9 – Resultados Inversão - Exemplo A

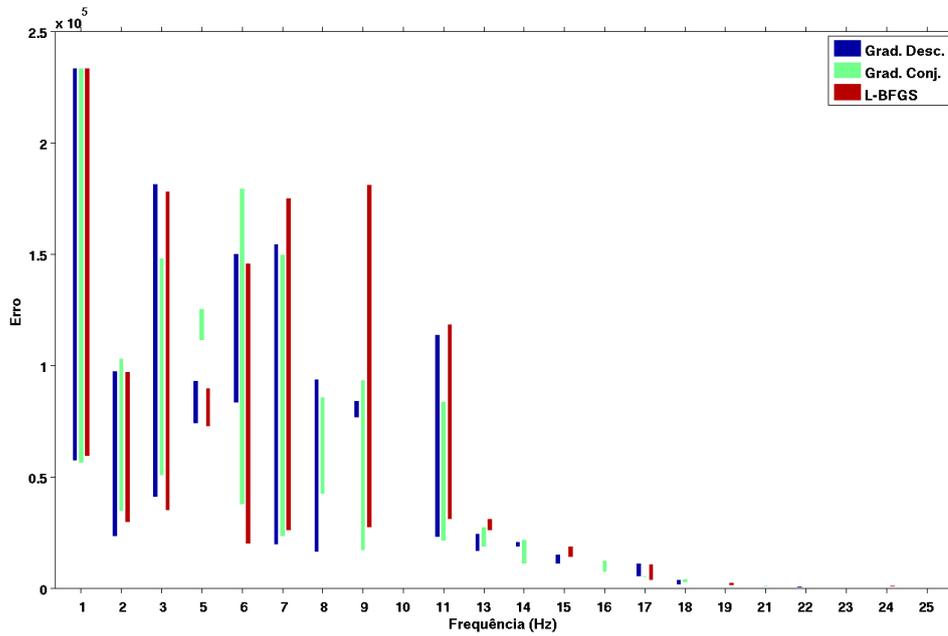
Método	Total de iterações	Total de modelagens	Erro Final
Gradiente Descendente	449	86292	15.96
Gradiente Conjugado	411	75153	26.61
L-BFGS	238	52029	25.60

Fonte: Produzida pelo autor

de modelagens encontra-se na Figura 29. O resultado final encontra-se na Figura 30. Os sismogramas encontram-se na Figura 31. A Figura 31a apresenta o sismograma real. As Figuras 31b-d apresentam os sismogramas sintéticos obtidos a partir dos modelos invertidos. A Figura 31e apresenta a comparação para um traço dos sismogramas. A Tabela 9 apresenta o total de iterações e modelagens para cada método, além da norma final do resíduo entre os sismogramas real e sintético. Podemos perceber pelo gráfico do erro que em geral o gradiente conjugado e o L-BFGS tiveram uma convergência melhor que o gradiente descendente. Na maior parte das frequências invertidas, a barra dos outros dois métodos é maior que a barra do gradiente descendente. Outro ponto importante no gráfico de evolução do erro é a diferença dos erros finais e iniciais para componentes de frequências adjacentes. Por exemplo, para o método gradiente descendente, não conseguimos enxergar a barra para a frequência de $9Hz$, ao ponto que as barras para os outros dois métodos são grandes para essa frequência. Quando olhamos a frequência seguinte ($10Hz$), vemos o comportamento oposto, a barra é grande para o método gradiente descendente e pequena para os outros dois métodos. Isto indica que o incremento de $1Hz$ nas frequências invertidas pode não ser ótimo. Já no número total de modelagens, vemos que o gradiente conjugado teve mais modelagens que os outros métodos na maioria das frequências invertidas, porém o método do gradiente descendente teve um número maior de iterações e modelagens que os outros métodos. O método do gradiente conjugado teve um número total de iterações um pouco menor que o gradiente descendente. O método L-BFGS foi o método que precisou do menor número de iterações de todos os testados. Já comparando o gradiente descendente com o L-BFGS vemos que, em grande parte das frequências, o gradiente descendente obteve mais modelagens que o L-BFGS. Isto indica que o gradiente descendente foi mais lento para obter uma convergência pior. Os sismogramas sintéticos são semelhantes aos sismogramas reais. O erro total dos sismogramas é menor para o método do gradiente descendente. Observando o resultado final, percebemos que todos os métodos tiveram um resultado bastante semelhante. Isto é esperado pois este é o melhor caso possível, pois o dado de entrada está livre de ruídos e o modelo inicial é bem próximo do modelo real. Neste exemplo, considerando o resultado final junto com a convergência e o número total de iterações, escolheríamos o método L-BFGS, mesmo o erro menor sendo do gradiente descendente.

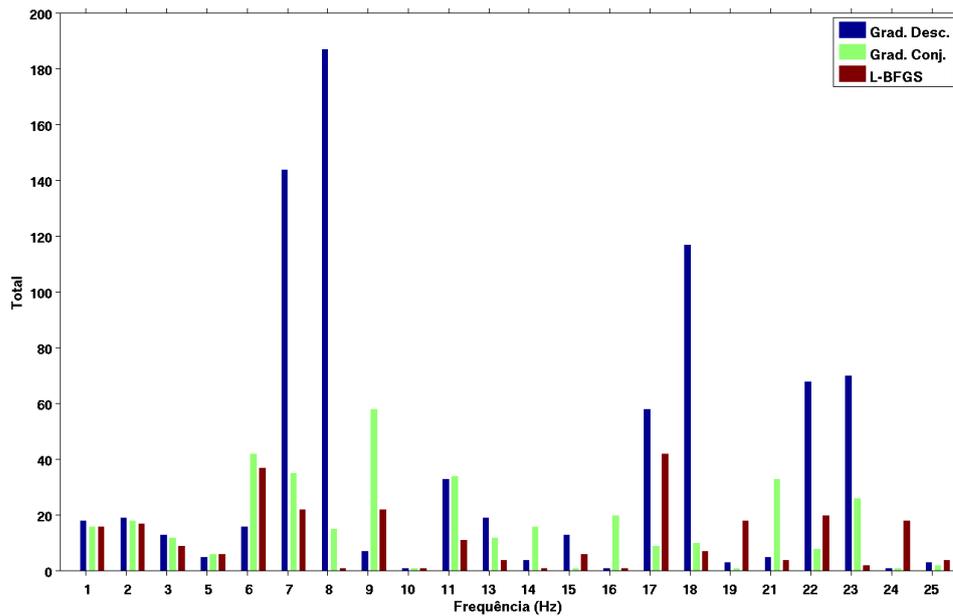
Para o segundo exemplo, utilizamos como modelo inicial um modelo homogêneo. Novamente não adicionamos ruído ao dado sísmico real. Pelo gráfico do erro, apresentado na Figura 32, vemos que os métodos tiveram uma convergência semelhante. O gradiente descendente obteve uma convergência mais comparável aos outros métodos neste exemplo. Olhando o gráfico do total de iterações da Figura 33 e do total de modelagens da Figura 34, vemos claramente que o gradiente descendente utilizou um número maior de modelagens durante a inversão. Isto torna-o o mais lento para esse exemplo. Comparando o resultado dos três métodos, que encontra-se na Figura 35, vemos que o L-BFGS não obteve um resultado

Figura 32 – Progressão de Erro - Exemplo B



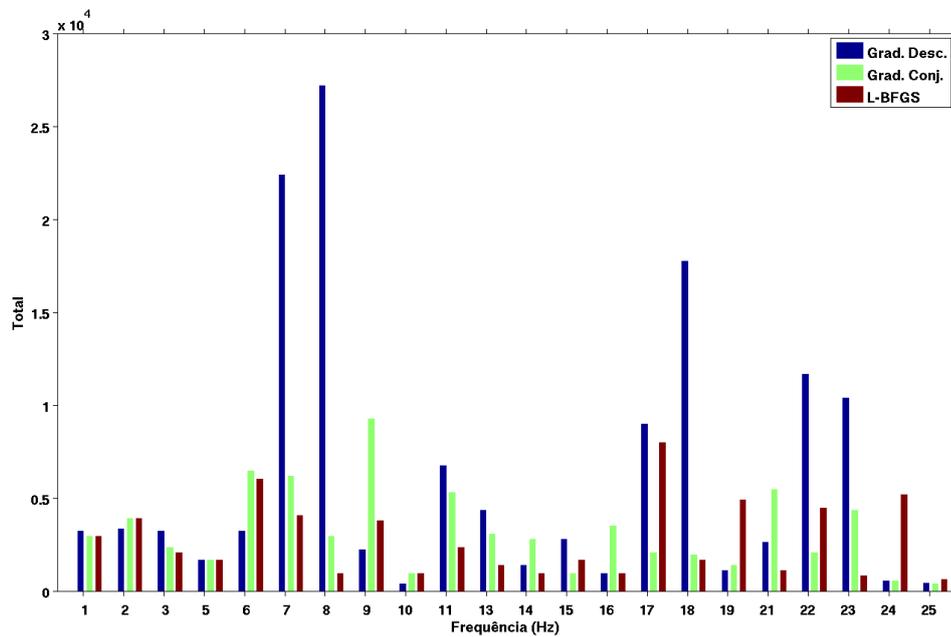
Fonte: Produzida pelo autor

Figura 33 – Total de iterações - Exemplo B



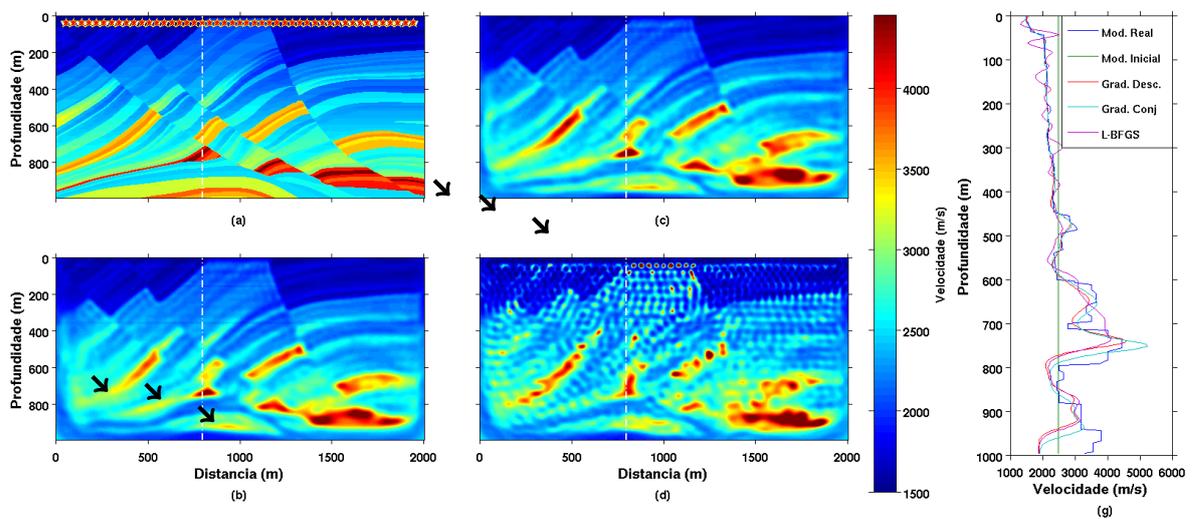
Fonte: Produzida pelo autor

Figura 34 – Total de modelagens - Exemplo B



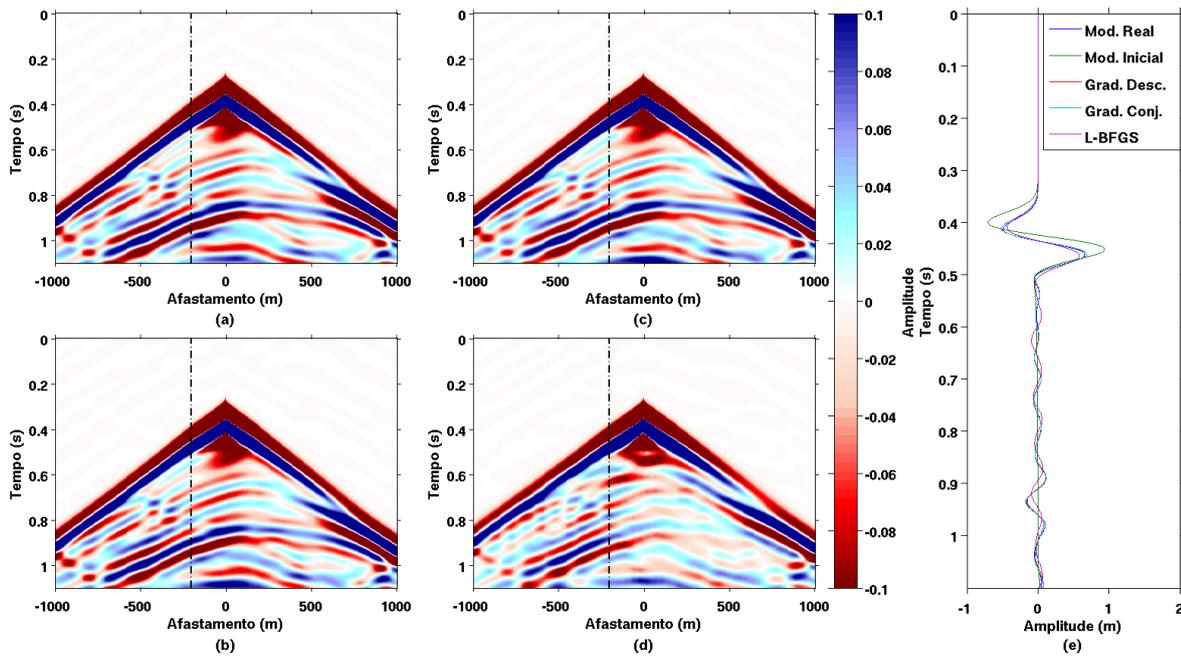
Fonte: Produzida pelo autor

Figura 35 – Resultado da inversão - Exemplo B. Os traços pontilhados em branco em (a)-(d) indicam a posição do traço apresentado em (e). As setas pretas indicam pontos de interesse. a) Modelo real. \star indica a posição das fontes simuladas. ∇ indica a posição dos receptores. b) Resultado Gradiente Descendente. c) Resultado Gradiente Conjugado. d) Resultado L-BFGS. e) Comparação para um traço do modelo.



Fonte: Produzida pelo autor

Figura 36 – Sismogramas - Exemplo B. Os traços pontilhados em (a)-(d) indicam a posição do traço apresentado em (e). a) Sismograma medido. b) Sismograma - Gradiente Descendente. c) Sismograma - Gradiente Conjugado. d) Sismograma - L-BFGS. e) Comparação para um traço do sismograma.



Fonte: Produzida pelo autor

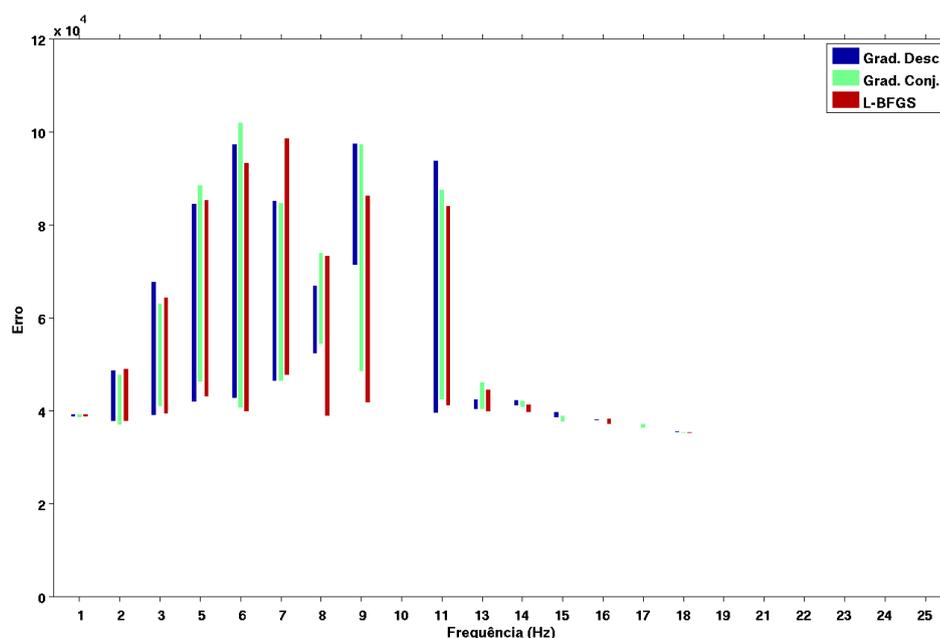
Tabela 10 – Resultados Inversão - Exemplo B

Método	Total de iterações	Total de modelagens	Erro Final
Gradiente Descendente	805	137240	27.31
Gradiente Conjugado	376	71205	21.23
L-BFGS	269	61147	118.15

Fonte: Produzida pelo autor

satisfatório na inversão. Comparando o gradiente descendente e o gradiente conjugado, vemos que o segundo obteve um resultado ligeiramente melhor, com as camadas mais profundas melhor representadas que o gradiente descendente. Os sismogramas apresentados na Figura 36 e os resultados da Tabela 10 confirmam o resultado do modelo de velocidades. Neste caso, a escolha óbvia foi o método do gradiente conjugado para a inversão.

Figura 37 – Progressão de erro - Exemplo C



Fonte: Produzida pelo autor

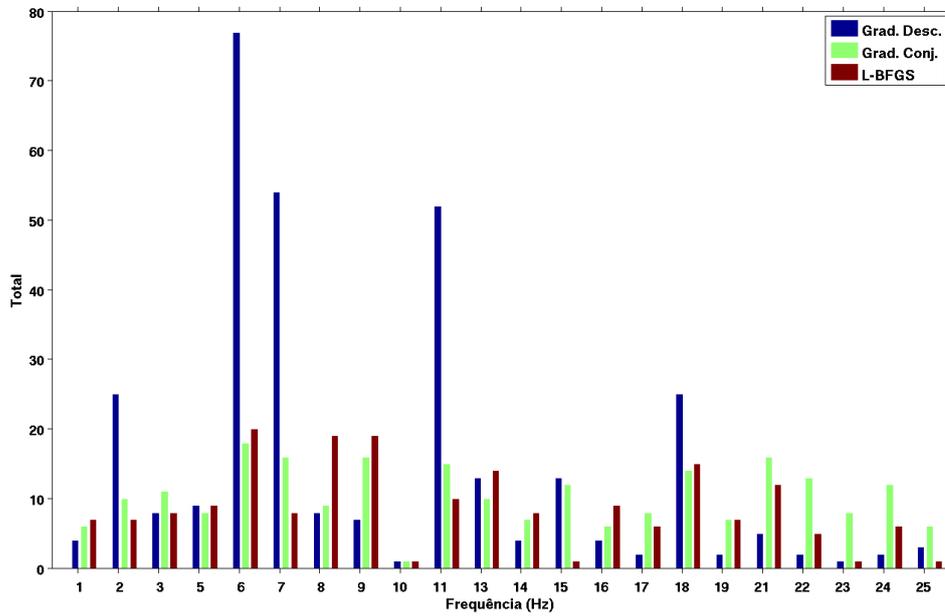
Tabela 11 – Resultados Inversão - Exemplo C

Método	Total de iterações	Total de modelagens	Erro Final
Gradiente Descendente	321	61241	62.47
Gradiente Conjugado	229	43052	93.04
L-BFGS	193	43005	126.75

Fonte: Produzida pelo autor

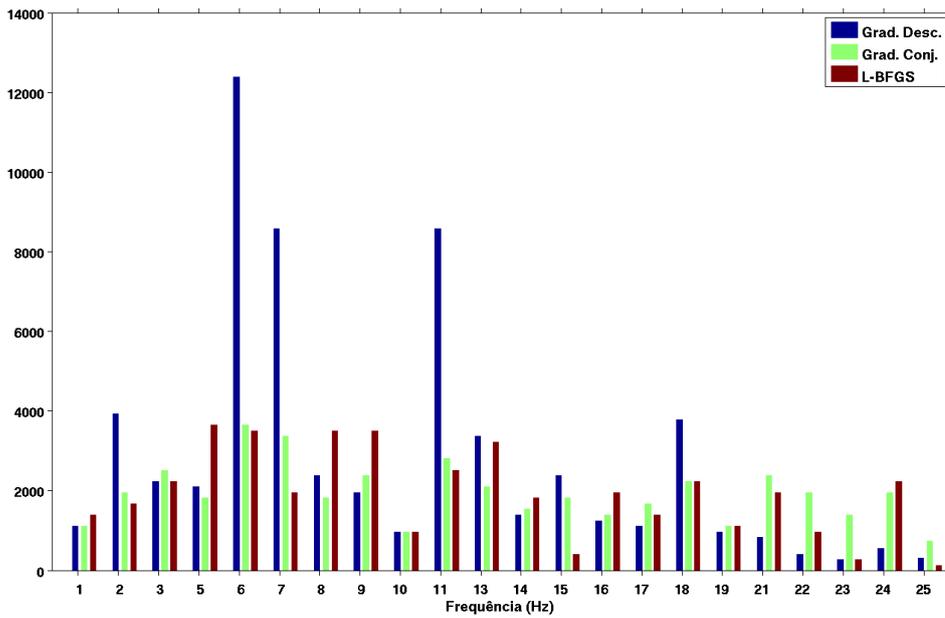
Vamos agora analisar a presença de ruído na inversão. Primeiro vamos avaliar a influência do ruído para um modelo inicial bem próximo do modelo real. Neste caso vamos utilizar como modelo inicial o modelo Marmousi suavizado igual ao primeiro exemplo. O gráfico da evolução do erro para cada método encontra-se na Figura 37. Podemos ver que em geral o método do gradiente descendente obteve uma convergência um pouco pior que os outros dois métodos. O número total de iterações encontra-se na Figura 38 e o total de

Figura 38 – Total de iterações - Exemplo C



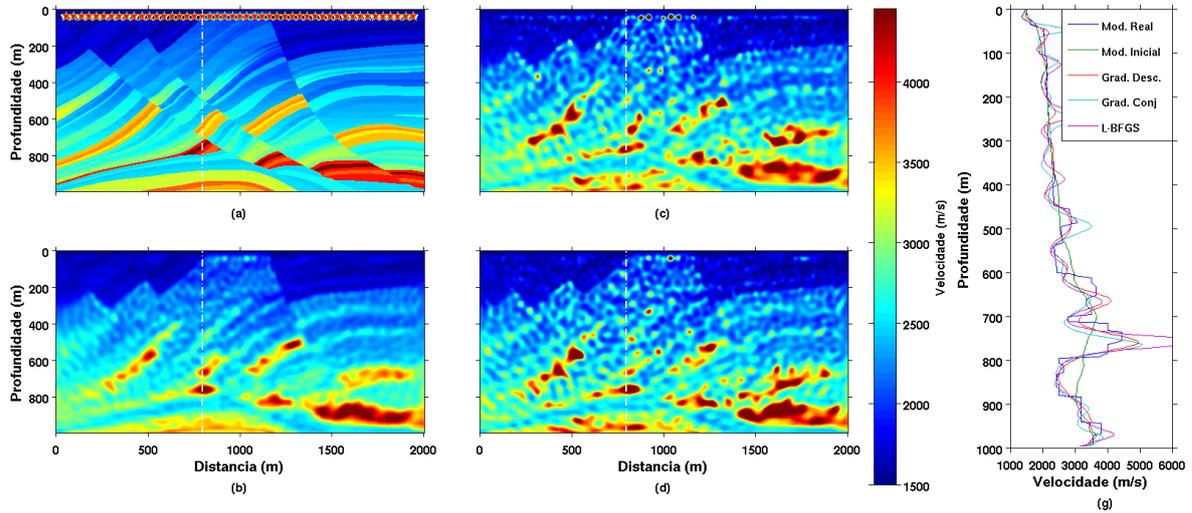
Fonte: Produzida pelo autor

Figura 39 – Total de modelagens - Exemplo C



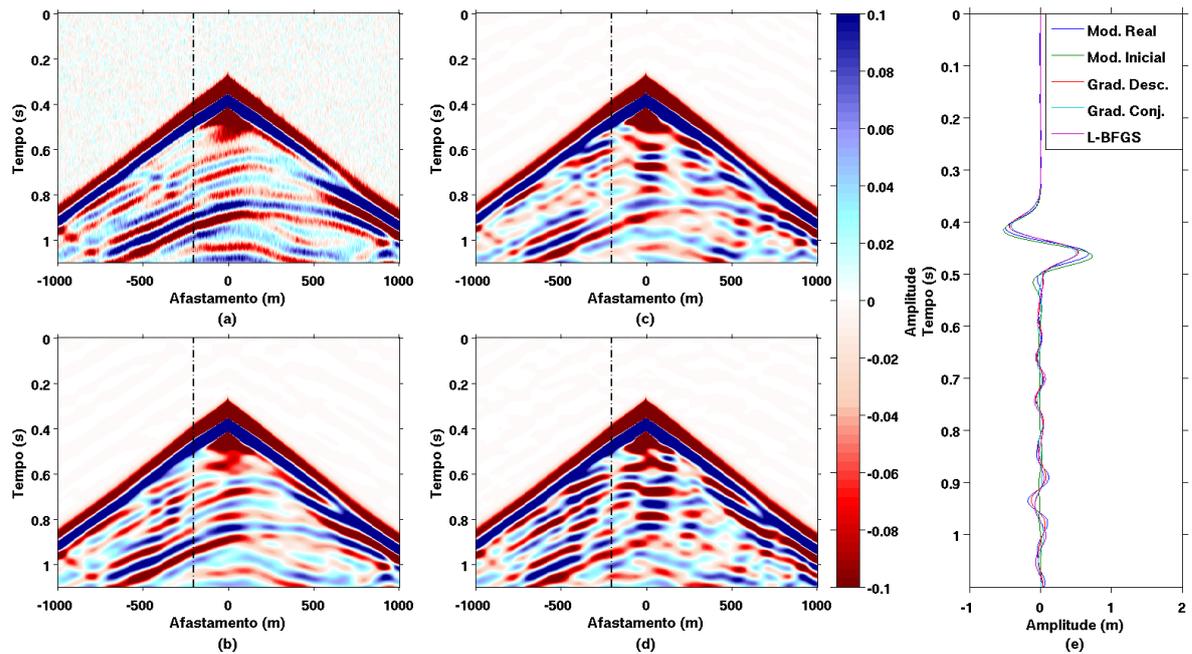
Fonte: Produzida pelo autor

Figura 40 – Resultado da inversão - Exemplo C. Os traços pontilhados em branco em (a)-(d) indicam a posição do traço apresentado em (e). a) Modelo real. \star indica a posição das fontes simuladas. ∇ indica a posição dos receptores. b) Resultado Gradiente Descendente. c) Resultado Gradiente Conjugado. c) Resultado L-BFGS. d) Comparação para um traço do modelo.



Fonte: Produzida pelo autor

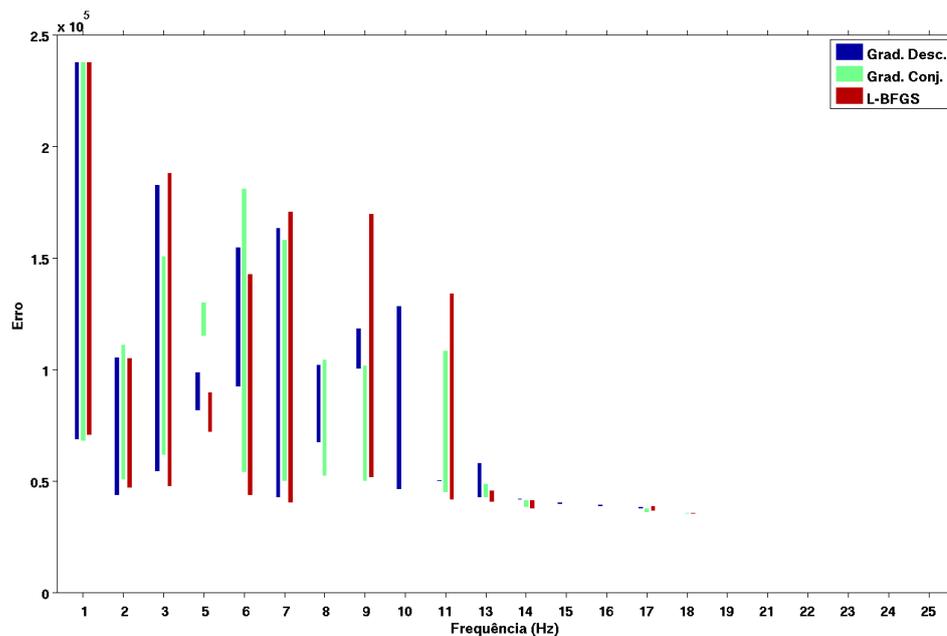
Figura 41 – Sismogramas - Exemplo C. Os traços pontilhados em (a)-(d) indicam a posição do traço apresentado em (e). a) Sismograma medido. b) Sismograma - Gradiente Descendente. c) Sismograma - Gradiente Conjugado. c) Sismograma - L-BFGS. d) Comparação para um traço do sismograma.



Fonte: Produzida pelo autor

modelagens encontra-se na Figura 39. Podemos perceber claramente que o número total de modelagens do gradiente descendente foi muito maior que os outros métodos, tornando-o o mais lento com a pior convergência. Porém, ao analisarmos o resultado final da Figura 40 percebemos que o melhor resultado foi o do método do gradiente descendente. Uma possível causa para este resultado pior é que os outros métodos possuem uma maior capacidade de incorporar artefatos presentes no dado de entrada no modelo invertido. Neste caso, os métodos L-BFGS e gradiente conjugado tiveram mais sucesso em encontrar um modelo de velocidades que representasse o ruído presente nos sismogramas reais. Observando os sismogramas na Figura 41 podemos confirmar isto. O sismograma obtido a partir do gradiente conjugado e do L-BFGS são mais próximos do sismograma real com ruído, ou seja, o ruído presente no dado foi incorporado no modelo de velocidades. Observando os resultados da Tabela 11, vemos que o erro para o gradiente descendente foi menor que os outros métodos. Neste caso o gradiente descendente seria a escolha natural.

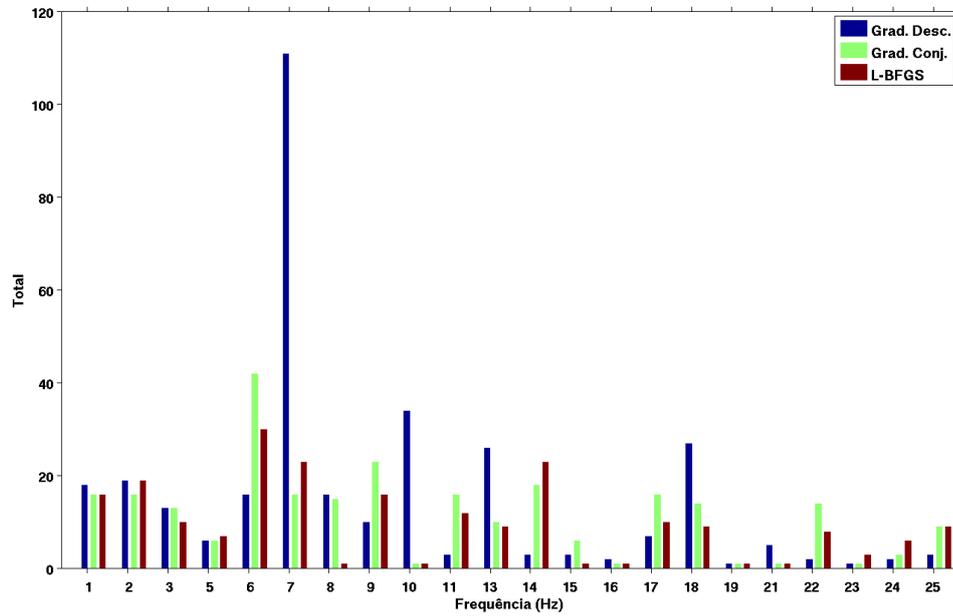
Figura 42 – Progressão de erro - Exemplo D



Fonte: Produzida pelo autor

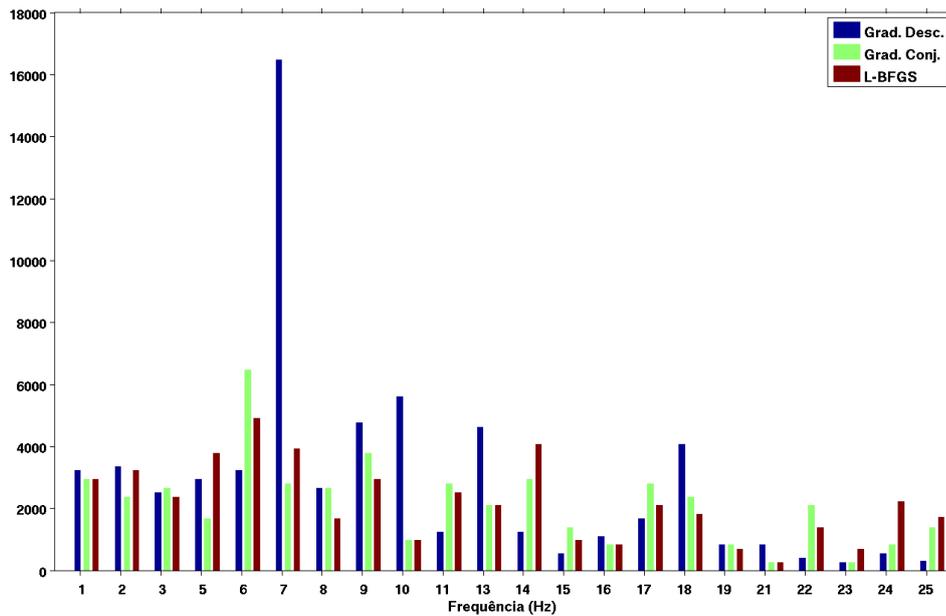
Para o último exemplo, vamos utilizar como modelo inicial um modelo homogêneo e o dado de entrada com ruído. A evolução do erro encontra-se na Figura 42. Novamente, o L-BFGS obteve a maior convergência. O número de iterações na Figura 43 e o número de modelagens encontra-se na Figura 44. O método do gradiente descendente novamente teve o maior número de modelagens para grande parte das frequências. Já para o resultado final, mostrado na Figura 45, vemos que todos os métodos não tiveram um bom resultado e assim como no exemplo anterior, os métodos gradiente conjugado e L-BFGS tiveram o pior resultado. Os sismogramas mostrados na Figura 46 confirmam o resultado semelhante

Figura 43 – Total de iterações - Exemplo D



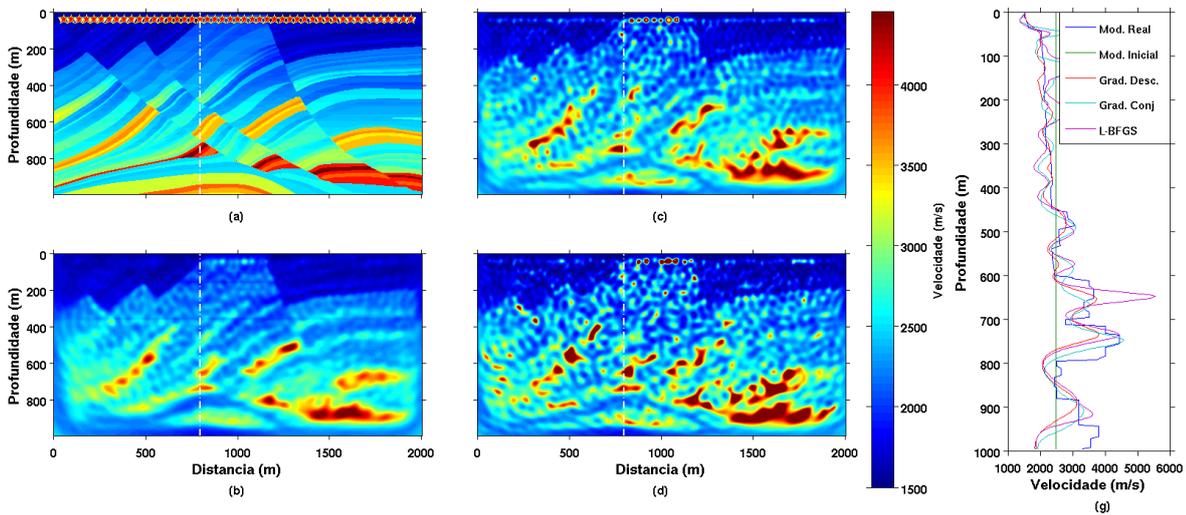
Fonte: Produzida pelo autor

Figura 44 – Total de modelagens - Exemplo D



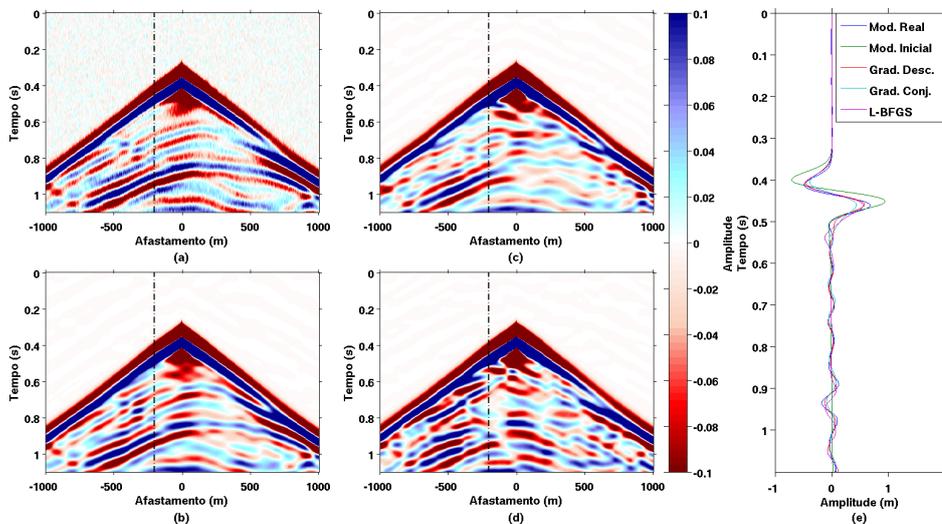
Fonte: Produzida pelo autor

Figura 45 – Resultado da inversão - Exemplo D. Os tracos pontilhados em branco em (a)-(d) indicam a posição do traço apresentado em (e). a) Modelo real. \star indica a posição das fontes simuladas. ∇ indica a posição dos receptores. b) Resultado Gradiente Descendente. c) Resultado Gradiente Conjugado. c) Resultado L-BFGS. d) Comparação para um traço do modelo.



Fonte: Produzida pelo autor

Figura 46 – Sismogramas - Exemplo D. Os tracos pontilhados em (a)-(d) indicam a posição do traço apresentado em (e). a) Sismograma medido. b) Sismograma - Gradiente Descendente. c) Sismograma - Gradiente Conjugado. c) Sismograma - L-BFGS. d) Comparação para um traço do sismograma.



Fonte: Produzida pelo autor

Tabela 12 – Resultados Inversão - Exemplo D

Método	Total de iterações	Total de modelagens	Erro Final
Gradiente Descendente	328	62933	52.09
Gradiente Conjugado	258	47658	114.31
L-BFGS	216	48551	146.09

Fonte: Produzida pelo autor

ao exemplo anterior. Os resultados da Tabela 12 apontam novamente que o método do gradiente descendente obteve o melhor resultado, embora tenha tido a convergência mais lenta. Para este exemplo, a escolha natural seria o método do gradiente descendente.

6 Discussão e conclusões

Primeiro, vimos que o algoritmo de modelagem influencia bastante no resultado final da inversão. Durante este trabalho, desenvolvemos um novo tipo de borda de absorção capaz de reduzir as reflexões das bordas. Vimos no exemplo prático que uma borda mal dimensionada pode prejudicar bastante o resultado final da inversão. E mesmo a presença de reflexões de borda durante o algoritmo de inversão que não encontram-se no dado medido já são suficientes para prejudicar o resultado final.

O segundo exemplo compara os métodos de otimização matemática implementados neste trabalho. Vimos que, em geral, o método do gradiente descendente é o que possui a pior convergência com o maior número de modelagens. Porém, em casos que o dado medido esteja muito contaminado com ruído, o gradiente descendente pode se tornar o melhor método, pois este se torna mais robusto, não incorporando os artefatos presentes no dado no modelo de velocidades, ao contrário dos outros métodos. Para um dado livre de ruídos, o gradiente conjugado e o L-BFGS obtiveram um resultado parecido, sendo que o gradiente conjugado apresentou-se mais estável nos dois exemplos apresentados. Uma análise da evolução dos erros por componente de frequência invertida indica que a escolha de incremento de $1Hz$ não é ideal. Em alguns casos, vemos que determinadas componentes de frequência introduziam pouca informação no modelo invertido. Uma análise mais profunda da evolução do erro pode levar à escolha ótima do incremento de frequência na inversão, aumentando a eficiência computacional do processo. Os resultados presentes neste trabalho demonstram as diferentes características dos diversos métodos de otimização presentes na literatura. Estas diferenças sugerem a possibilidade da utilização de algoritmos híbridos de inversão que sejam capazes de analisar os dados de entrada e selecionar qual dos métodos disponíveis melhor se aplica em cada momento, juntando assim, robustez e eficiência.

6.1 Trabalhos futuros

Para trabalhos futuros, podemos enumerar alguns possíveis pontos de estudo.

Primeiro, este trabalho não utiliza nenhuma forma de pré-condicionamento de dados na inversão. Existem na literatura algumas formas diferentes de realizar isto, como descrito no capítulo 4. A regularização implementada neste trabalho é feita através da filtragem do gradiente. A implementação do pré-condicionamento, juntamente com outros métodos de regularização podem melhorar o resultado da inversão.

Outro ponto importante de estudo, ainda no caso acústico, é estudar os aspectos

práticos da inversão em casos mais realísticos. Utilizar diferentes algoritmos de modelagem para o dado medido e para a inversão, implementar o algoritmo de inversão da forma da onda sísmica e utilizar modelos iniciais obtidos a partir do próprio dado (através de tomografia) são alguns dos possíveis estudos que devem ser feitos para validar o algoritmo de inversão do campo de onda completo.

Finalmente, podemos estender as técnicas aqui apresentadas para incluir outros fenômenos presentes na propagação de ondas em subsuperfície. A inclusão dos parâmetros elásticos e de anisotropia sísmica nos algoritmos de modelagem e inversão são um desafio enfrentado por pesquisadores hoje e indicam o caminho natural de evolução dos métodos desenvolvidos neste trabalho.

Referências

- ALERINI, M.; BÉGAT, S. L.; LAMBARÉ, G.; BAINA, R. 2D pp- and ps-stereotomography for a multicomponent dataset. *SEG Technical Program Expanded Abstracts 2002*, Society of Exploration Geophysicists, Jan 2002. Disponível em: <http://dx.doi.org/10.1190/1.1817391>.
- BEDNAR, J. B.; SHIN, C.; PYUN, S. Comparison of waveform inversion, part 2: phase approach. *Geophysical Prospecting*, Wiley-Blackwell, v. 55, n. 4, p. 465–475, Jul 2007. ISSN 1365-2478. Disponível em: <http://dx.doi.org/10.1111/j.1365-2478.2007.00618.x>.
- BERENGER, J.-P. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, Elsevier BV, v. 114, n. 2, p. 185–200, Oct 1994. ISSN 0021-9991. Disponível em: <http://dx.doi.org/10.1006/jcph.1994.1159>.
- BEYDOUN, W. B.; TARANTOLA, A. First born and rytov approximations: Modeling and inversion conditions in a canonical example. *The Journal of the Acoustical Society of America*, Acoustical Society of America (ASA), v. 83, n. 3, p. 1045–1055, Mar 1988. ISSN 0001-4966. Disponível em: <http://dx.doi.org/10.1121/1.396537>.
- BILLETTE, F.; BÉGAT, S. L.; PODVIN, P.; LAMBARÉ, G. Practical aspects and applications of 2D stereotomography. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 68, n. 3, p. 1008–1021, May 2003. ISSN 1942-2156. Disponível em: <http://dx.doi.org/10.1190/1.1581072>.
- BILLETTE, F.; LAMBARÉ, G. Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophysical Journal International*, Oxford University Press (OUP), v. 135, n. 2, p. 671–690, Nov 1998. ISSN 1365-246X. Disponível em: <http://dx.doi.org/10.1046/j.1365-246X.1998.00632.x>.
- BRENDERS, A. J.; PRATT, R. G. Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data. *Geophysical Journal International*, Oxford University Press (OUP), v. 168, n. 1, p. 152–170, Jan 2007. ISSN 1365-246X. Disponível em: <http://dx.doi.org/10.1111/j.1365-246X.2006.03096.x>.
- BRENDERS, A. J.; PRATT, R. G. Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model. *Geophysical Journal International*, Oxford University Press (OUP), v. 168, n. 1, p. 133–151, Jan 2007. ISSN 1365-246X. Disponível em: <http://dx.doi.org/10.1111/j.1365-246X.2006.03156.x>.
- BRENDERS, A. J.; PRATT, R. G. Waveform tomography of marine seismic data: What can limited offset offer? *SEG Technical Program Expanded Abstracts 2007*, Society of Exploration Geophysicists, Jan 2007. Disponível em: <http://dx.doi.org/10.1190/1.2793099>.
- BROSSIER, R.; OPERTO, S.; VIRIEUX, J. Robust elastic frequency-domain full-waveform inversion using the l1 norm. *Geophysical Research Letters*, v. 36, n. 20, p. n/a–n/a, 2009. ISSN 1944-8007. Disponível em: <http://dx.doi.org/10.1029/2009GL039458>.

BROSSIER, R.; OPERTO, S.; VIRIEUX, J. Seismic imaging of complex onshore structures by 2d elastic frequency-domain full-waveform inversion. *Geophysics*, v. 74, n. 6, p. WCC105–WCC118, 2009. Disponível em: <<http://library.seg.org/doi/abs/10.1190/1.3215771>>.

BROYDEN, C. G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, v. 6, n. 1, p. 76–90, 1970. Disponível em: <<http://imamat.oxfordjournals.org/content/6/1/76.abstract>>.

CLAYTON, R.; ENGQUIST, B. Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the Seismological Society of America*, Bulletin of the Seismological Society of America, v. 67, n. 6, p. 1529–1540, 1977. ISSN 0037-1106. Disponível em: <<http://bssa.geoscienceworld.org/content/67/6/1529>>.

CRASE, E.; PICA, A.; NOBLE, M.; MCDONALD, J.; TARANTOLA, A. Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, v. 55, n. 5, p. 527–538, 1990. Disponível em: <<http://library.seg.org/doi/abs/10.1190/1.1442864>>.

DENNIS; MORÉE, J. J. Quasi-newton methods, motivation and theory. *SIAM Review*, Society for Industrial and Applied Mathematics, v. 19, n. 1, p. 46–89, 1977. Disponível em: <<http://dx.doi.org/10.2307/2029325>>.

DÜMMONG, S.; MEIER, K.; GAJEWSKI, D.; HÜBSCHER, C. Comparison of prestack stereotomography and nip wave tomography for velocity model building: Instances from the messinian evaporites. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 73, n. 5, p. VE291–VE302, Sep 2008. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.2950306>>.

FLETCHER, R. A new approach to variable metric algorithms. *The Computer Journal*, Oxford University Press (OUP), v. 13, n. 3, p. 317–322, Mar 1970. ISSN 1460-2067. Disponível em: <<http://dx.doi.org/10.1093/comjnl/13.3.317>>.

FLETCHER, R. *Practical Methods of Optimization. Volume 1: Unconstrained Optimization*. 1st. ed. John Wiley and Sons Ltd, 1980. Hardcover. ISBN 0471277118. Disponível em: <<http://www.worldcat.org/isbn/0471277118>>.

GÉLIS, C.; VIRIEUX, J.; GRANDJEAN, G. Two-dimensional elastic full waveform inversion using born and rytov formulations in the frequency domain. *Geophysical Journal International*, Oxford University Press (OUP), v. 168, n. 2, p. 605–633, Feb 2007. ISSN 1365-246X. Disponível em: <<http://dx.doi.org/10.1111/j.1365-246X.2006.03135.x>>.

GOLDFARB, D. A family of variable-metric methods derived by variational means. *Mathematics of Computation*, American Mathematical Society (AMS), v. 24, n. 109, p. 23–23, Jan 1970. ISSN 0025-5718. Disponível em: <<http://dx.doi.org/10.1090/S0025-5718-1970-0258249-6>>.

HESTENES, M. R.; STIEFEL, E. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, v. 49, n. 6, p. 409+, dez. 1952. ISSN 0091-0635. Disponível em: <<http://dx.doi.org/10.6028/jres.049.044>>.

HOLE, J. A. Nonlinear high-resolution three-dimensional seismic travel time tomography. *Journal of Geophysical Research*, Wiley-Blackwell, v. 97, n. B5, p. 6553, 1992. ISSN 0148-0227. Disponível em: <<http://dx.doi.org/10.1029/92JB00235>>.

- HU, W.; ABUBAKAR, A.; HABASHY, T. M.; LIU, J. Preconditioned non-linear conjugate gradient method for frequency domain full-waveform seismic inversion. *Geophysical Prospecting*, Blackwell Publishing Ltd, v. 59, n. 3, p. 477–491, 2011. ISSN 1365-2478. Disponível em: <<http://dx.doi.org/10.1111/j.1365-2478.2010.00938.x>>.
- HUSTEDT, B.; OPERTO, S.; VIRIEUX, J. Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling. *Geophysical Journal International*, Oxford University Press (OUP), v. 157, n. 3, p. 1269–1296, Jun 2004. ISSN 1365-246X. Disponível em: <<http://dx.doi.org/10.1111/j.1365-246X.2004.02289.x>>.
- JAISWAL, P.; ZELT, C. A.; BALLY, A. W.; DASGUPTA, R. 2-D travelttime and waveform inversion for improved seismic imaging: Naga thrust and fold belt, India. *Geophysical Journal International*, Oxford University Press (OUP), v. 173, n. 2, p. 642–658, May 2008. ISSN 1365-246X. Disponível em: <<http://dx.doi.org/10.1111/j.1365-246X.2007.03691.x>>.
- JAISWAL, P.; ZELT, C. A.; DASGUPTA, R.; NATH, K. K. Seismic imaging of the naga thrust using multiscale waveform inversion. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 74, n. 6, p. WCC129–WCC140, Nov 2009. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.3158602>>.
- LAMBARÉ, G. Stereotomography. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 73, n. 5, p. VE25–VE34, Sep 2008. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.2952039>>.
- LAMBARÉ, G.; ALERINI, M. Semi automatic pp-ps stereotomography: application to the synthetic valhall dataset. *SEG Technical Program Expanded Abstracts 2005*, Society of Exploration Geophysicists, Jan 2005. Disponível em: <<http://dx.doi.org/10.1190/1.2148316>>.
- LEE, K. H.; KIM, H. J. Source-independent full-waveform inversion of seismic data. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 68, n. 6, p. 2010–2015, Nov 2003. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.1635054>>.
- MALINOWSKI, M.; OPERTO, S. Quantitative imaging of the permo-mesozoic complex and its basement by frequency domain waveform tomography of wide-aperture seismic data from the polish basin. *Geophysical Prospecting*, Wiley-Blackwell, v. 56, n. 6, p. 805–825, Nov 2008. ISSN 1365-2478. Disponível em: <<http://dx.doi.org/10.1111/j.1365-2478.2007.00680.x>>.
- MENKE, W. *Geophysical Data Analysis: Discrete Inverse Theory: MATLAB Edition*. Elsevier Science, 2012. (International Geophysics). ISBN 9780123977847. Disponível em: <https://books.google.com.br/books?id=sUDAqFV_oRAC>.
- MÉTIVIER, L.; BROSSIER, R.; VIRIEUX, J.; OPERTO, S. Full waveform inversion and the truncated newton method. *SIAM Journal on Scientific Computing*, v. 35, n. 2, p. B401–B437, 2013. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/120877854>>.
- MOREIRA, R. M.; SANTOS, M. A. C.; MARTINS, J. L.; SILVA, D. L. F.; PESSOLANI, R. B. V.; FILHO, D. M. S.; BULCÃO, A. Frequency-domain acoustic-wave modeling with hybrid absorbing boundary conditions. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 79, n. 5, p. A39–A44, Sep 2014. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/geo2014-0085.1>>.

NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. 2nd. ed. Springer New York, 2006. (Springer Series in Operations Research and Financial Engineering). Disponível em: <http://dx.doi.org/10.1007/978-0-387-40065-5>.

NOLET, G. *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*. Springer Netherlands, 1987. (Modern Approaches in Geophysics). ISBN 9789027725219. Disponível em: <https://books.google.com.br/books?id=vM23Ux3TsnwC>.

OPERTO, S.; VIRIEUX, J.; DESSA, J.-X.; PASCAL, G. Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography: Application to the eastern nankai trough. *Journal of Geophysical Research*, Wiley-Blackwell, v. 111, n. B9, 2006. ISSN 0148-0227. Disponível em: <http://dx.doi.org/10.1029/2005JB003835>.

PAIGE, C. C.; SAUNDERS, M. A. Algorithm 583: LSQR: Sparse linear equations and least squares problems. *ACM Transactions on Mathematical Software*, Association for Computing Machinery (ACM), v. 8, n. 2, p. 195–209, Jun 1982. ISSN 0098-3500. Disponível em: <http://dx.doi.org/10.1145/355993.356000>.

PLESSIX, R.-E. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, Oxford University Press (OUP), v. 167, n. 2, p. 495–503, Nov 2006. ISSN 1365-246X. Disponível em: <http://dx.doi.org/10.1111/j.1365-246X.2006.02978.x>.

PRATT, G.; SHIN, C.; HICKS. Gauss–newton and full newton methods in frequency–space seismic waveform inversion. *Geophysical Journal International*, Blackwell Science Ltd, v. 133, n. 2, p. 341–362, maio 1998. Disponível em: <http://dx.doi.org/10.1046/j.1365-246x.1998.00498.x>.

PRATT, R. G. Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics*, SEG, v. 64, n. 3, p. 888–901, 1999. Disponível em: <http://dx.doi.org/10.1190/1.1444597>.

PRATT, R. G.; SYMES, W. Semblance and differential semblance optimisation for waveform tomography: A frequency domain implementation. In: *Sub-basalt imaging conference, Expanded Abstracts*. [S.l.: s.n.], 2002. p. 183–184.

PYUN, S.; SHIN, C.; BEDNAR, J. B. Comparison of waveform inversion, part 3: amplitude approach. *Geophysical Prospecting*, Wiley-Blackwell, v. 55, n. 4, p. 477–485, Jul 2007. ISSN 1365-2478. Disponível em: <http://dx.doi.org/10.1111/j.1365-2478.2007.00619.x>.

RAVAUT, C.; OPERTO, S.; IMPROTA, L.; VIRIEUX, J.; HERRERO, A.; DELL’AVERSANA, P. Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt. *Geophysical Journal International*, Oxford University Press (OUP), v. 159, n. 3, p. 1032–1056, Dec 2004. ISSN 1365-246X. Disponível em: <http://dx.doi.org/10.1111/j.1365-246X.2004.02442.x>.

SHANNO, D. F. Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation*, American Mathematical Society (AMS), v. 24, n. 111, p. 647–647, Sep 1970. ISSN 0025-5718. Disponível em: <http://dx.doi.org/10.1090/S0025-5718-1970-0274029-X>.

SHIN, C.; CHA, Y. H. Waveform inversion in the Laplace domain. *Geophysical Journal International*, Oxford University Press (OUP), v. 173, n. 3, p. 922–931, Jun 2008. ISSN 1365-246X. Disponível em: <<http://dx.doi.org/10.1111/j.1365-246X.2008.03768.x>>.

SHIN, C.; CHA, Y. H. Waveform inversion in the Laplace-Fourier domain. *Geophysical Journal International*, Oxford University Press (OUP), v. 177, n. 3, p. 1067–1079, Jun 2009. ISSN 1365-246X. Disponível em: <<http://dx.doi.org/10.1111/j.1365-246X.2009.04102.x>>.

SHIN, C.; HA, W. A comparison between the behavior of objective functions for waveform inversion in the frequency and Laplace domains. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 73, n. 5, p. VE119–VE133, Sep 2008. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.2953978>>.

SHIN, C.; MIN, D.-J. Waveform inversion using a logarithmic wavefield. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 71, n. 3, p. R31–R42, May 2006. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.2194523>>.

SHIN, C.; PYUN, S.; BEDNAR, J. B. Comparison of waveform inversion, part 1: conventional wavefield vs logarithmic wavefield. *Geophysical Prospecting*, Wiley-Blackwell, v. 55, n. 4, p. 449–464, Jul 2007. ISSN 1365-2478. Disponível em: <<http://dx.doi.org/10.1111/j.1365-2478.2007.00617.x>>.

TAILLANDIER, C.; NOBLE, M.; CHAURIS, H.; CALANDRA, H. First-arrival traveltime tomography based on the adjoint-state method. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 74, n. 6, p. WCB1–WCB10, Nov 2009. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.3250266>>.

TARANTOLA, A. Linearized inversion of seismic reflection data. *Geophysical Prospecting*, Blackwell Publishing Ltd, v. 32, n. 6, p. 998–1015, 1984. ISSN 1365-2478. Disponível em: <<http://dx.doi.org/10.1111/j.1365-2478.1984.tb00751.x>>.

TARANTOLA, A. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier, 1987. ISBN 9780444427656. Disponível em: <<https://books.google.com.br/books?id=AdUQAQAIAAJ>>.

VIRIEUX, J.; OPERTO, S. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, Society of Exploration Geophysicists, v. 74, n. 6, p. WCC1–WCC26, nov. 2009. Disponível em: <<http://dx.doi.org/10.1190/1.3238367>>.

WILLIAMSON, P. R. A guide to the limits of resolution imposed by scattering in ray tomography. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 56, n. 2, p. 202–207, Feb 1991. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.1443032>>.

WOODWARD, M. J. Wave-equation tomography. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 57, n. 1, p. 15–26, Jan 1992. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.1443179>>.

WOODWARD, M. J.; NICHOLS, D.; ZDRAVEVA, O.; WHITFIELD, P.; JOHNS, T. A decade of tomography. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 73, n. 5, p. VE5–VE11, Sep 2008. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.2969907>>.

ZELT, C. A.; BARTON, P. J. Three-dimensional seismic refraction tomography: A comparison of two methods applied to data from the faeroe basin. *Journal of Geophysical Research: Solid Earth*, Wiley-Blackwell, v. 103, n. B4, p. 7187–7210, Apr 1998. ISSN 0148-0227. Disponível em: <<http://dx.doi.org/10.1029/97JB03536>>.

ZHOU, B.; GREENHALGH, S. A. Crosshole seismic inversion with normalized full-waveform amplitude data. *GEOPHYSICS*, Society of Exploration Geophysicists, v. 68, n. 4, p. 1320–1330, Jul 2003. ISSN 1942-2156. Disponível em: <<http://dx.doi.org/10.1190/1.1598125>>.

Apêndices

APÊNDICE A – Modelagem Acústica no Domínio da Frequência

Criação da matriz de modelagem no domínio da frequência.

Conforme vimos no capítulo 3, a equação da modelagem no domínio da frequência pode ser escrita de forma matricial

$$\mathbf{B}\mathbf{u} = \mathbf{s} \quad (\text{A.1})$$

A matriz \mathbf{B} deve incluir o tratamento da borda de absorção da modelagem. Conforme vimos no capítulo 3, a borda utilizada neste trabalho foi uma borda híbrida que inclui tanto a borda PML como a borda ABC. Os elementos da matriz podem ser separados então em três partes:

1. Interior do modelo
2. Borda PML
3. Borda ABC

A discretização por diferenças finitas de segunda ordem utilizada neste trabalho utiliza no máximo cinco elementos do campo de pressão e do modelo de velocidades para o cálculo de cada elemento do campo de pressão, conforme descrito na equação abaixo.

$$\frac{\omega^2}{\mathbf{c}^2}u_{i,j} + \left[\frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\Delta x^2} \right] + \left[\frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{\Delta z^2} \right] = -s_{i,j} \quad (\text{A.2})$$

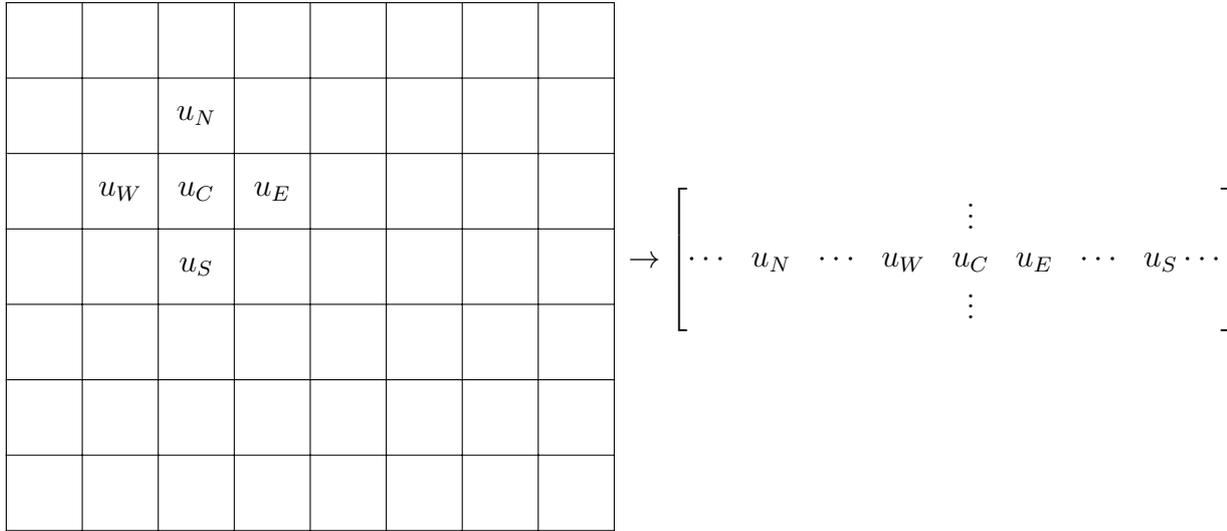
Podemos definir estes elementos como central, norte, sul, leste e oeste. Na criação da matriz B , esta relação deve ser vetorizada, conforme mostrado na figura 47.

Desta forma, os índices da matriz \mathbf{B} podem ser descritos conforme as equações abaixo.

A.1 Interior do Modelo

O interior do modelo corresponde aos valores de $N_b + 1 \geq x \leq N_x - N_b - 1$ e $N_b + 1 \geq z \leq N_z - N_b - 1$

Figura 47 – Modelo discretizado



Fonte: Produzida pelo Autor

$$B_C = B_{i,j} = \left(\frac{\omega}{c}\right)^2 - 2 \left(\frac{1}{dx^2} + \frac{1}{dz^2} \right) \quad (\text{A.3})$$

$$B_N = B_{i,j-Nx} = \frac{1}{dz^2} \quad (\text{A.4})$$

$$B_S = B_{i,j+Nx} = \frac{1}{dz^2} \quad (\text{A.5})$$

$$B_W = B_{i,j-1} = \frac{1}{dx^2} \quad (\text{A.6})$$

$$B_E = B_{i,j+1} = \frac{1}{dx^2} \quad (\text{A.7})$$

A.2 Borda PML

A parte do modelo que corresponde a borda PML encontra-se nos índices do modelo $2 \geq x \leq N_b$ e $N_b + 1 \geq z \leq N_z - N_b - 1$ para a borda esquerda, $2 \geq z \leq N_b$ e $N_b + 1 \geq x \leq N_z - N_b - 1$ para a borda superior, $N_b \geq x \leq N_x - 1$ e $N_b + 1 \geq z \leq N_z - N_b - 1$ para a borda direita, $N_b \geq z \leq N_z$ e $N_b + 1 \geq x \leq N_x - N_b - 1$ para a borda inferior, $2 \geq x \leq N_b$ e $2 \geq z \leq N_b$ para a quina superior esquerda, $2 \geq x \leq N_b$ e $N_z - N_b \geq z \leq N_z - 1$ para a quina inferior esquerda, $N_x - N_b \geq x \leq N_x - 1$ e $2 \geq z \leq N_b$ para a quina superior direita e $N_x - N_b \geq x \leq N_x - 1$ e $N_z - N_b \geq z \leq N_z - 1$ para a quina inferior esquerda,

$$B_C = B_{i,j} = \left(\frac{\omega}{c}\right)^2 - \frac{2}{\alpha} \left(\frac{1}{dx^2} + \frac{1}{dz^2}\right) \quad (\text{A.8})$$

$$B_N = B_{i,j-N_x} = \frac{1}{\alpha dz^2} \quad (\text{A.9})$$

$$B_S = B_{i,j+N_x} = \frac{1}{\alpha dz^2} \quad (\text{A.10})$$

$$B_W = B_{i,j-1} = \frac{1}{\alpha dx^2} \quad (\text{A.11})$$

$$B_E = B_{i,j+1} = \frac{1}{\alpha dx^2} \quad (\text{A.12})$$

$$\alpha = 1 + i \frac{C_{PML}}{\omega} \cos\left(\frac{\pi L}{2 N_b}\right) \quad (\text{A.13})$$

A.3 Borda Superior

A borda superior corresponde aos índices $2 \geq x \leq N_x - 1$ e $z = 1$

$$B_C = B_{i,j} = -\frac{1}{dz} - i \frac{\omega}{c} \quad (\text{A.14})$$

$$B_S = B_{i,j+N_x} = \frac{1}{dz} \quad (\text{A.15})$$

A.4 Borda Inferior

A borda inferior corresponde aos índices $2 \geq x \leq N_x - 1$ e $z = N_z$

$$B_C = B_{i,j} = -\frac{1}{dz} - i \frac{\omega}{c} \quad (\text{A.16})$$

$$B_N = B_{i,j-N_x} = \frac{1}{dz} \quad (\text{A.17})$$

A.5 Borda Esquerda

A borda esquerda corresponde aos índices $2 \geq z \leq N_z - 1$ e $x = 1$

$$B_C = B_{i,j} = -\frac{1}{dx} - i \frac{\omega}{c} \quad (\text{A.18})$$

$$B_E = B_{i,j+1} = \frac{1}{dx} \quad (\text{A.19})$$

A.6 Borda Direita

A borda direita corresponde aos índices $2 \geq z \leq N_z - 1$ e $x = N_x$

$$B_C = B_{i,j} = -\frac{1}{dx} - i\frac{\omega}{c} \quad (\text{A.20})$$

$$B_W = B_{i,j-1} = \frac{1}{dx} \quad (\text{A.21})$$

A.7 Canto Superior Esquerdo

A quina superior esquerda corresponde aos índices $x = 1$ e $z = 1$

$$B_C = B_{i,j} = -\frac{1}{dx} - \frac{1}{dz} - 2i\frac{\omega}{c} \quad (\text{A.22})$$

$$B_E = B_{i,j+1} = \frac{1}{dx} \quad (\text{A.23})$$

$$B_S = B_{i,j+N_x} = \frac{1}{dz} \quad (\text{A.24})$$

A.8 Canto Superior Direito

A quina superior direita corresponde aos índices $x = N_x$ e $z = 1$

$$B_C = B_{i,j} = -\frac{1}{dx} - \frac{1}{dz} - 2i\frac{\omega}{c} \quad (\text{A.25})$$

$$B_W = B_{i,j-1} = \frac{1}{dx} \quad (\text{A.26})$$

$$B_S = B_{i,j+N_x} = \frac{1}{dz} \quad (\text{A.27})$$

A.9 Canto Inferior Esquerdo

A quina inferior esquerda corresponde aos índices $x = 1$ e $z = N_z$

$$B_C = B_{i,j} = -\frac{1}{dx} - \frac{1}{dz} - 2i\frac{\omega}{c} \quad (\text{A.28})$$

$$B_E = B_{i,j+1} = \frac{1}{dx} \quad (\text{A.29})$$

$$B_N = B_{i,j-N_x} = \frac{1}{dz} \quad (\text{A.30})$$

A.10 Canto Inferior Direito

A quina inferior direita corresponde aos índices $x = N_x$ e $z = N_z$

$$B_C = B_{i,j} = -\frac{1}{dx} - \frac{1}{dz} - 2i\frac{\omega}{c} \quad (\text{A.31})$$

$$B_W = B_{i,j-1} = \frac{1}{dx} \quad (\text{A.32})$$

$$B_N = B_{i,j-N_x} = \frac{1}{dz} \quad (\text{A.33})$$