



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
DEPARTAMENTO DE FÍSICA TEÓRICA E EXPERIMENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

Tese de Doutorado

**Métodos otimizados para inversão completa das formas
de onda em 3D**

Edwin Humberto Fagua Duarte

Orientador: Prof. Dr. João Medeiros de Araújo

Natal, RN
2021



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
DEPARTAMENTO DE FÍSICA TEÓRICA E EXPERIMENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

Métodos otimizados para inversão completa das formas de onda em 3D

Edwin Humberto Fagua Duarte

Orientador: **Prof. Dr. João Medeiros de Araújo**

*Tese de Doutorado apresentada ao Programa de
Pós-Graduação em Física do Departamento de Fí-
sica Teórica e Experimental da Universidade Fede-
ral do Rio Grande do Norte como requisito parcial
para obtenção do título de **Doutor em Física**.*

Natal
Outubro, 2021

Resumo

A inversão de forma de onda completa (FWI) é uma das técnicas mais estudadas na atualidade para recuperar parâmetros do subsolo que afetam a propagação da onda. Uma das principais dificuldades da aplicação da FWI em levantamentos sísmicos 3D é o alto custo computacional desta. Este problema tem sido contornado com o avanço da tecnologia computacional mas ainda existem dificuldades para aplicar a FWI em dados reais 3D. Neste trabalho são estudadas estratégias para otimizar a FWI, as quais são baseadas nas propriedades físicas das ondas. Com este fim, este trabalho foi feito em duas partes. Na primeira parte foi usada a técnica de fontes simultâneas codificadas, na qual o cálculo do gradiente da FWI é feito propagando simultaneamente a informação de várias fontes agrupadas em um ou vários subgrupos aproveitando o princípio de superposição da equação da onda. Para reduzir o ruído de conversa cruzada ou “crosstalk” que aparece no cálculo do gradiente usando fontes simultâneas, foram propostos dois esquemas de codificação, por rotação de fase e por estática limitada a um período da frequência dominante. Na codificação por rotação de fase é aplicada uma rotação de fase em cada tiro onde o ângulo de rotação é escolhido aleatoriamente para cada tiro. No caso da codificação por estática, os traços de cada tiro são deslocados temporalmente e o valor deste deslocamento é escolhido aleatoriamente para cada tiro tendo um valor máximo de um período da frequência dominante do dado. Com ajuda de testes numéricos em dados sintéticos 2D foi demonstrado que os esquemas propostos apresentam uma convergência mais rápida que as codificações por polarização (multiplicar por +1 ou -1 aleatoriamente cada tiro) e por estática sem ser limitada. Na segunda parte foi desenvolvido um fluxograma para aplicar a FWI em um dado real adquirido na região do pré-sal brasileiro. Este dado foi obtido com aquisição tipo OBN 3D, onde os receptores são espalhados no assoalho oceânico. No fluxograma desenvolvido, são combinadas várias técnicas existentes para otimizar o uso de memória e de processamento, as quais aproveitam propriedades da propagação das ondas. Sobre o dado OBN foram feitas inversões usando estratégias de dizimação de fontes e receptores para reduzir o custo computacional. Também foram testados os esquemas de fontes codificadas estudadas na primeira parte. Esses testes foram feitos também usando várias funções objetivo com a intenção de melhorar a resolução das inversões. Foi verificado que as inversões recuperaram informação relacionada ao reservatório pré-sal e que tanto as estratégias de dizimação como de codificação de fontes reduziram o custo computacional da FWI sem comprometer a qualidade da inversão.

Palavras-chave: Inversão de forma de onda completa, Fontes simultâneas, estratégias de codificação, inversão em dado real 3D.

Abstract

Full waveform inversion (FWI) is one of the most studied techniques today to recover subsurface parameters that affect wave propagation. One of the main difficulties in applying FWI in 3D seismic surveys is its high computational cost. This problem has been solved with the advancement of computational technology, but there are still difficulties to apply a FWI to real 3D data. In this work are studied implemented to optimize a FWI, as which are based on the physical properties of waves. To this end, this work was done in two parts. In the first part, the encoded simultaneous sources technique was used, in which the FWI gradient is calculated by simultaneously propagating information from several sources grouped in one or several subgroups, taking advantage of the superposition principle of the wave equation. In order to reduce the crosstalk noise that appears in the gradient calculation using simultaneous sources, two encoding schemes have been proposed, by phase rotation and by static limited to a period of the dominant frequency. In phase rotation encoding a phase rotation is applied to each shot where the rotation angle is chosen randomly for each shot. In static coding the traces of each shot are temporally shifted and the value of this shift is chosen randomly for each shot having a maximum value of one period of the dominant frequency of the data. With the help of numerical tests on synthetic 2D data to demonstrate that the proposed schemes have a faster convergence than encodings by polarization (multiply by +1 or -1 randomly each shot) and by static without being limited. In the second part, a workflow was developed to apply a FWI to real data acquired in the Brazilian pre-salt region. This data was obtained with an Ocean Bottom Node (OBN) 3D acquisition, where the receivers are spread at the ocean bottom. In the developed workflow, several existing techniques are combined to optimize the use of memory and processing, which take advantage of wave propagation properties. On the OBN data for inversions made using sources and receivers decimation strategies to reduce computational cost. The encoded source schemes studied in the first part were also tested. These testicles also use various functions with the objective of an intention of melody and a resolution of inversions. It was verified that the inversions retrieved information related to the pre-salt reservoir and that both decimation and source coding strategies reduced the computational cost of FWI without compromising the quality of the inversion.

Keywords: Full Waveform Inversion, Simultaneous sources, encoding strategies, inversion using 3D field data.

Sumário

Sumário	i
Lista de Figuras	iii
1 Introdução	1
1.1 Método sísmico para exploração de óleo e gás	1
1.2 Inversão por forma de onda completa	2
1.3 Objetivos da tese	3
1.4 Estrutura da tese	5
2 Inversão por Forma de Onda Completa	6
2.1 Introdução	6
2.2 Amostragem de sinais digitais - Teorema de Shannon	6
2.3 Problema direto - Solução numérica da equação da onda	7
2.3.1 Método das diferenças finitas	8
2.3.2 Criterios de estabilidade	11
2.3.3 Simulação de um meio infinito	12
2.4 Formulação do problema inverso	14
2.4.1 Revisão de otimização não-linear	14
2.4.2 Métodos Newton e quase-Newton. Método L-BFGS	15
2.5 Função objetivo da FWI e seu gradiente	16
2.5.1 Função objetivo	16
2.5.2 Gradiente da função objetivo	17
3 Estratégias de Fontes Codificadas	21
3.1 Introdução	21
3.2 Função objetivo e gradiente usando fontes simultâneas	22
3.3 Esquemas de codificação de fontes	24
3.4 L-BFGS reinicializado	26
3.5 Experimentos numéricos	27
4 Construção de um código para FWI em 3D	33
4.1 Introdução	33
4.2 Física usada na modelagem sísmica	34
4.2.1 Uso de densidade variável - equação de Klein-Gordon	34
4.2.2 Condição de superfície livre	35
4.3 Posicionamento correto de fontes e receptores	36

4.4	Estimativa da assinatura da fonte	38
4.5	Cálculo de gradiente em 3D	40
4.5.1	Reamostragem temporal	40
4.5.2	Recálculo do campo direto	40
4.5.3	Filtragem do gradiente - Filtro de Laplace	41
4.6	Detalhes de implementação do código 3D	42
5	FWI aplicado em um dado do pré-sal brasileiro	45
5.1	Introdução	45
5.2	Descrição dos dados usados	46
5.3	Preparação de modelos iniciais	48
5.3.1	Modelo inicial de velocidade	48
5.3.2	Modelo de densidade	50
5.4	Fluxograma para aplicação do FWI	51
5.5	Dizimação de fontes e receptores no cálculo do gradiente	52
5.6	Funções objetivos testadas	53
5.7	Escolha de frequências de inversão	54
5.8	Resultados das inversões	56
5.8.1	Discussão	56
5.8.2	Custo computacional	61
5.8.3	Controle de qualidade das inversões	61
5.9	Resultados com fontes codificadas	63
5.9.1	Escolha de dizimação de fontes	64
5.9.2	Escolha de função objetivo	66
5.9.3	Escolha de quantidade de super-receptores	67
5.9.4	Escolha do esquema de codificação	68
6	Conclusões	73
Referências bibliográficas		75
A	Produção bibliográfica	81
A.1	Artigos completos publicados em periódicos	81
A.2	Artigos submetidos	81
A.3	Artigos publicados em anais de conferências	82

Lista de Figuras

1.1	Exemplo de levantamento sísmico marinho. Figura extraída da internet (www.pinterest.es/pin/234257618101516530/)	1
2.1	Representação do esquema de diferenças finitas para a equação da onda 2D usando um laplaciano de segunda ordem.	9
2.2	Extensão do modelo de velocidade para aplicar a borda absorvente CPML.	13
3.1	(a) Família de afastamento comum de 500m. Mesma família de afastamento comum com codificação (b) por polarização, (c) por rotação de fase aleatória e (d) por estática aleatória.	24
3.2	Esquema de L-BFGS reinicializado. (Figura extraida do artigo Rao & Wang 2017).	27
3.3	(a) Modelo SEG/EAGE overthrust. (b) Modelo inicial usado nos experimentos de inversão. (c) Resultado da inversão convencional.	28
3.4	FWI com fontes codificadas usando os diferentes esquemas de codificação. Resultado com codificação por (a) polarização aleatória (III) depois de 775 iterações, (b) rotação de fase aleatória (I) depois de 250 iterações, (c) estática aleatória limitada a um período (II) depois de 355 iterações e (d) estática aleatória limitada a 10 períodos (IV) depois de 690 iterações.	30
3.5	Convergência da função objetivo ao longo das iterações com os diferentes esquemas de codificação e a FWI convencional. Função objetivo como função (a) do número de iterações e (b) do número de propagações de onda.	31
3.6	Coeficiente de correlação localizada entre o modelo verdadeiro e o modelo invertido para (a) a FWI convencional e (b) a codificação de fontes por rotação de fase.	32
4.1	Esquema da implementação da superfície livre pelo método das imagens.	35
4.2	Efeito da interpolação de Hicks: a) Mapa com coordenadas reais. b) Traços calculados com interpolação de Hicks. c) Mapa com coordenadas aproximadas aos pontos da malha. d) Traços calculados sem interpolação.	37
4.3	Cálculo do gradiente usando a borda do domínio espacial para recalcular o campo direto. Esta figura foi feita baseada na Figura 1 de Yang et al. 2016.	42

5.1	(a) Planta do levantamento OBN usado nas inversões. As coordenadas usadas na figura não são geográficas, senão que são relativas ao modelo de velocidade. O nó realçado mostra a posição usada para modelar os sismogramas apresentados na seção de controle de qualidade. (b) Mapa 3D da aquisição.	47
5.2	Dado de campo de um receptor localizado no centro dos três círculos fontes. Cada retângulo corresponde a um círculo fonte com raios 6 km (esquerda), 8 km (meio), 10 km (direita). São destacados alguns sinais importantes como a onda direta (setas amarelas), as ondas refratadas (setas cianas) e as ondas espalhadas (setas azuis).	48
5.3	Modelo de velocidade de referência. A figura superior esquerda representa um corte em profundidade onde são desenhados as linhas fontes e um polígono indicando a área de receptores. A figura superior direita representa um corte crossline que vai de sul ao norte. A figura na parte inferior representa um corte inline onde as principais camadas geológicas são identificadas (Water: Camada d'água, PosS: Camada de sedimentos pós-sal, Salt: Corpo de sal, BaseS: Base do corpo do sal e Res: reservatório). Cada corte tem duas linhas roxas indicando o lugar por onde passam os outros dois cortes. No corte em profundidade estão as posições dos perfis das Figuras 5.13 e 5.20.	49
5.4	Modelos iniciais de (a) velocidade e (b) densidade.	49
5.5	Comparação de dado observado e filtrado com dados modelados usando o modelo de velocidade da Figura 5.4(a) e um modelo de densidade b) da Figura 5.4(b), c) constante e d) usando a equação de Gardner em todo o modelo.	50
5.6	(a) Mapa mostrando as fontes e receptores que contribuem ao cálculo de um gradiente. (b) Gradiente de um par fonte-receptor para um afastamento XY de 15.6 km. Pode ser observado que a zona de Fresnel possui uma largura de aproximadamente 2 km.	53
5.7	Diagrama de fase para a) 1 Hz b) 1.5 Hz c) 2 Hz d) 2.5 Hz.	56
5.8	Resultado da inversão usando a norma ℓ_2	57
5.9	Resultado da inversão usando a norma ℓ_1	58
5.10	Resultado da inversão usando NPR-FWI.	59
5.11	Comparação das diferentes funções objetivo. (a)-(c) Cortes em profundidade, inline e crossline da inversão com norma ℓ_2 . (d)-(f) Cortes em profundidade, inline e crossline da inversão com norma ℓ_1 . (g)-(i) Cortes em profundidade, inline e crossline da inversão com NPR-FWI. As posições dos cortes são as mesmas das Figuras 5.8 até 5.10	60
5.12	Comparação de (b) dados observados de campo com os dados modelados nos modelos (a) de referência, (c) inicial e como os resultados da inversão com (d) norma ℓ_2 , (e) norma ℓ_1 e (f) NPR-FWI.	62

5.13	Perfis verticais de velocidade comparando o modelo de referência com o modelo inicial e os modelos resultado da inversão com as diferentes funções objetivos. No mapa da direita são apresentadas as posições dos perfis.	62
5.14	Comparação de inversões de codificação por polarização usando 8 super-receptores variando a dizimação de fontes. Inversão (a)-(c) dizimando fontes 1 a cada 9, (d)-(f) dizimando fontes 1 a cada 5, (g)-(i) sem dizimar fontes. Os cortes em profundidade, inline e crossline são os mesmos da Figura 5.11	65
5.15	Comparação de inversões usando codificação por polarização usando 8 super-receptores usando as funções objetivo descritas na seção 5.6. (a)-(c) usando a norma ℓ_2 . (d)-(f) usando a norma ℓ_1 . (g)-(i) usando NPR-FWI. Os cortes em profundidade, inline e crossline são os mesmos das figuras anteriores.	66
5.16	Comparação de inversões usando codificação por polarização usando (a)-(c) 8 super-receptores, (d)-(f) 16 super-receptores e (g)-(i) 32 super-receptores. Os cortes em profundidade, inline e crossline são os mesmos das figuras anteriores.	67
5.17	Resultado da inversão usando norma ℓ_1 usando codificação por polarização.	69
5.18	Resultado da inversão usando norma ℓ_1 usando codificação por rotação de fase.	70
5.19	Comparação de inversões usando (a)-(c) codificação por polarização, (d)-(f) codificação por rotação de fase e (g)-(i) inversão convencional com a norma ℓ_1 da Figura 5.9. Os cortes em profundidade, inline e crossline são os mesmos das figuras anteriores.	71
5.20	Perfis verticais de velocidade comparando o modelo de referência com o modelo inicial e os modelos invertido com a norma ℓ_1 sem codificação (Figura 5.9), com codificação por polarização (Figura 5.17) e por rotação de fase (Figura 5.18). No mapa da direita são apresentadas as posições dos perfis.	71

Capítulo 1

Introdução

1.1 Método sísmico para exploração de óleo e gás

O método sísmico é um dos métodos geofísicos mais amplamente usados para a busca de óleo e gás. O método consiste na produção de ondas sísmicas através de fontes artificiais como explosivo ou caminhões vibradores no caso terrestre e canhões de ar no caso marinho. Estas ondas são refletidas ou refratadas pelas estruturas geológicas no subsolo. Os sinais refletidos ou refratados são registrados por um certo tempo em um conjunto de receptores localizados na superfície (Ver Sheriff 2002). Uma ilustração esquemática de um levantamento sísmico é mostrada na Figura 1.1. A informação registrada pelos receptores é gravada digitalmente em uma sequência de números denominada *traço sísmico*. O conjunto de traços que possuem a informação de uma fonte ou receptor comum se chama de *agrupamento¹* por tiro ou agrupamento de receptor, respectivamente.

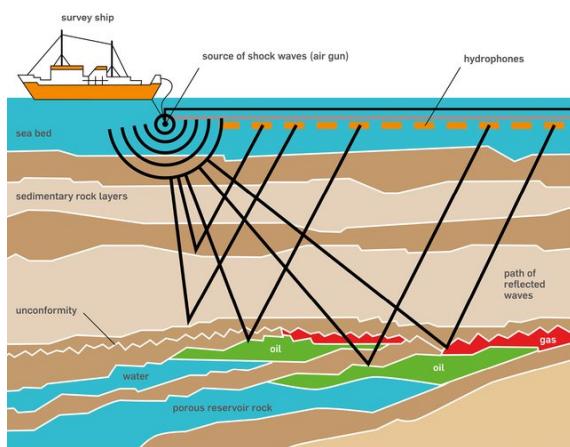


Figura 1.1: Exemplo de levantamento sísmico marinho. Figura extraída da internet (www.pinterest.es/pin/234257618101516530/).

Segundo a distribuição geométrica das fontes e receptores, um levantamento sísmico pode ser classificado como 2D se as fontes e receptores estão dispostos em

¹Tradução do termo em inglês *gather*, amplamente usado em sísmica.

uma linha reta e como 3D se estão dispostos ao longo de um área. Também existe uma classificação segundo o tipo de ondas a ser analisadas no processamento dos dados. Nessa classificação está a sísmica por refração, que é planejada para beneficiar as ondas refratadas e a sísmica de reflexão a qual usa as ondas refletidas nas interfaces entre as camadas e estruturas geológicas. Embora nos dois tipos de sísmica sejam gravadas todo tipo de ondas, refratadas ou refletidas, geralmente no processamento somente é usado um tipo de ondas.

Um estudo geológico usando o método sísmico é feito através de três etapas. Na primeira etapa denominada *aquisição* são gravados os dados das ondas sísmicas geradas na fonte. Na segunda etapa, denominada *processamento sísmico*, são analisados os dados registrados na etapa de aquisição para gerar uma *imagem sísmica*, ou seja, a distribuição e geometria das estruturas geológicas. Por último, na terceira etapa de *interpretação*, são analisadas as imagens obtidas na etapa de processamento, para identificar as possíveis regiões e estruturas de interesse.

A etapa de processamento é baseada em diferentes algoritmos computacionais os quais são baseados na solução de problemas inversos. A evolução desses algoritmos tem acompanhado o desenvolvimento da tecnologia computacional, permitindo que cada vez sejam melhoradas as suposições feitas sobre a física que descrevem os dados observados.

1.2 Inversão por forma de onda completa

Uma das técnicas de processamento sísmico que tem sido mais estudada para obter informação do subsolo a partir das ondas sísmicas é a denominada inversão por forma de onda completa ou FWI, pelas siglas em inglês de *Full Waveform Inversion*. O objetivo da FWI é obter um conjunto de parâmetros físicos ou modelo a partir da otimização de uma função objetivo, a qual compara os dados registrados em campo com dados simulados numericamente através da solução numérica da equação da onda. Os parâmetros que podem ser obtidos a partir da FWI são aqueles que afetam as amplitudes e a cinemática das ondas observadas, tais como a velocidade de onda ou a densidade entre outros. A função objetivo mais usada na literatura é a norma ℓ_2 da diferença entre os dados simulados e os dados observados.

Em geral, o estudo da FWI pode ser dividido em três grandes problemas:

1. **Função objetivo:** A função objetivo possui alguns problemas bem identificados como sensibilidade ao modelo inicial, ao ruído no dado observado e à falta de baixas frequências e afastamentos fonte-receptor longos os quais podem gerar um salto de ciclo ou *cycle skipping* (Virieux & Operto 2009). O salto de ciclo pode fazer com que a FWI convirja para valores incorretos de velocidade. Para resolver esses problemas têm sido propostas funções objetivo diferentes da norma ℓ_2 tais como a correlação cruzada (van Leeuwen & Mulder 2008), a AWI (Warner & Guasch 2016) ou transporte ótimo (Métivier et al. 2016).
2. **Física do problema:** O ajuste dos dados simulados com os observados depende também do modelo físico usado na simulação computacional. Uma simplifica-

ção ao problema é usar a equação da onda acústica com densidade constante a qual depende unicamente da velocidade da onda P. Essa simplificação permite obter o modelo de velocidades para uma ampla variedade de dados sísmicos, principalmente dados marinhos. Porém, em algumas situações, os efeitos de outros parâmetros elásticos como a densidade, velocidade de onda S, ou efeitos da atenuação da onda em um meio material, afetam significativamente os valores das amplitudes do sinal (Kamath et al. 2021). Esses problemas podem ser contornados a partir de equações de onda que incluem esses efeitos. Além disso, usar equações de onda mais completas permite obter mais parâmetros físicos do subsolo além da velocidade da onda P.

3. **Eficiência computacional:** No processo de inversão da FWI é necessário solucionar repetidas vezes a equação da onda, o que torna a FWI um problema caro computacionalmente. O custo computacional da FWI depende de fatores como extensão do modelo a ser invertido, a complexidade da equação da onda usada e o número de fontes sísmicas usadas na aquisição. Para resolver esses problemas existem estratégias computacionais, tal como a paralelização dos códigos, ou estratégias físicas como o uso das fontes simultâneas ou a dízimação do número de fontes e receptores, além de outras que também foram exploradas nessa tese.

1.3 Objetivos da tese

Como mencionado acima, um dos principais assuntos de pesquisa relacionados a FWI é a melhora da eficiência computacional da mesma, ou seja, reduzir o tempo de processamento e o consumo de memória. Além das melhorias computacionais tais como a paralelização do código e o uso de novas tecnologias de cômputo tais como a placa gráfica, também existem estratégias para reduzir o custo computacional que usam baseadas nas propriedades da propagação das ondas.

O principal objetivo desta tese é explorar o uso deste tipo de estratégias físicas para otimizar a FWI 3D aplicada em dados reais. Para isto, esta tese é dividida em duas partes. Na primeira parte é estudado o uso de fontes simultâneas (Krebs et al. 2009, Castellanos et al. 2014). Na FWI convencional o custo computacional da inversão é proporcional ao número de fontes sísmicas, devido a que as propagações de onda devem ser feitas independentemente para cada fonte. No caso das fontes simultâneas, a informação de todas as fontes é usada ao mesmo tempo na resolução da equação de onda, o qual pode ser feito por causa do princípio de superposição. Isto reduz consideravelmente o tempo da inversão uma vez que o número de propagações feitas na inversão já não é mais proporcional ao número de fontes.

Porém, quando o gradiente é calculado com fontes simultâneas, aparece um ruído devido à correlação cruzada de campos diretos e reversos gerados por fontes diferentes. Para atenuar esse ruído, chamado de conversa cruzada ou crosstalk, são usados os esquemas de codificação de fontes (Romero et al. 2000, Krebs et al. 2009, Castellanos et al. 2014), onde é aplicada a cada fonte um código, o qual transforma a amplitude e a fase do dado observado. A escolha do código é feita aleatoriamente

para cada fonte. Esses esquemas de codificação, junto com a escolha correta do algoritmo de otimização, permitem amortecer o crosstalk (Rao & Wang 2017, Rao et al. 2019).

Nesta tese foram propostos dois esquemas de codificação por rotação de fase e por estática temporal limitada a um período da frequência dominante. Os esquemas propostos apresentaram uma convergência mais rápida que os esquemas de codificação antes estudados na literatura como a codificação por polarização (Krebs et al. 2009) a codificação por estática temporal Schuster et al. 2011.

Na segunda parte foi desenvolvido um fluxograma para aplicar FWI em um dado real OBN² 3D adquirido no pré-sal brasileiro. Neste fluxograma são usadas também várias propriedades da propagação da onda para reduzir o custo computacional. Por exemplo, foi aplicado a reamostragem temporal para reduzir o custo de processamento, o qual é possível devido a que a FWI usa principalmente as frequências baixas dos dados observados (Virieux & Operto 2009).

Também foram usadas este tipo de estratégias para reduzir o uso de memória no cálculo do gradiente da FWI. Para calcular o gradiente é necessário manter o campo direto na memória enquanto é calculado o campo adjunto o reverso no tempo que para um levantamento OBN 3D pode ser da ordem de várias dezenas de gigabytes o qual supera a memória do computador. Para resolver isto foi usada a técnica de recalcular o campo direto simultaneamente ao campo reverso mantendo em memória somente o campo da borda computacional (Yang et al. 2014). Esta técnica é baseada no princípio de conservação da energia, uma vez que a equação de onda acústica não possui atenuação, o qual que permite que o eixo do tempo possa ser invertido na equação da onda.

Além das estratégias usadas no cálculo do gradiente, também foram aplicadas tanto estratégias de dizimação de fontes e receptores para aplicar a FWI convencional, como também foram aplicadas as estratégias de codificação desenvolvidas na primeira parte. Como a FWI usa frequências baixas, os comprimentos de onda que aparecem no gradiente são muito grandes o que implica que a zona de Fresnel é também muito grande e duas fontes ou dois receptores muito próximos vão contribuir ao gradiente praticamente com a mesma informação (Huang & Schuster 2014). Esta propriedade justifica a dizimação de fontes e receptores a qual reduziu o custo de processamento do gradiente sem perda considerável de informação. Com as inversões realizadas com as estratégias de dizimação e de codificação, foi demonstrado que a FWI no dado OBN usado consegue recuperar a informação relativa ao reservatório pré-sal. Além disso, foi demonstrado que o uso da codificação permite reduzir o tempo de inversão sem comprometer consideravelmente a qualidade da mesma.

²OBN é a sigla de *Ocean Bottom Node*, é um tipo de aquisição sísmica onde os receptores são posicionados fixos no assoalho oceânico.

1.4 Estrutura da tese

A estrutura da tese é a seguinte: no capítulo 2 é apresentada uma revisão da inversão de onda completa, mostrando os aspectos computacionais da solução dos problemas direto e inverso. No capítulo 3 são apresentados os primeiros resultados de melhora do desempenho da FWI através do uso de fontes simultâneas codificadas, onde são propostos dois novos esquemas de codificação de fontes. No capítulo 4 é descrito os métodos desenvolvidos para aplicar FWI em dados 3D. Estes métodos incluem estratégias para reduzir o consumo de memória e de processamento. Também é descrita a técnica usada na estimativa da assinatura da fonte e as modificações feitas na equação da onda para incluir o efeito da densidade e o efeito da superfície livre para a modelagem de múltiplas. No capítulo 5 são apresentados os resultados da FWI aplicada em um dado real tipo OBN 3D adquirido no pré-sal brasileiro, onde incluem-se os resultados com diferentes funções objetivo assim como também usando os esquemas de fontes simultâneas codificadas estudadas no capítulo 3. Por último, no capítulo 6 são apresentadas as conclusões da tese.

Capítulo 2

Inversão por Forma de Onda Completa

2.1 Introdução

Como mencionado no capítulo 1, a inversão por forma de onda completa ou FWI, visa obter propriedades do subsolo através da otimização de uma função objetivo, a qual compara os dados observados em campo com dados simulados computacionalmente. Devido à não linearidade e ao grande número de variáveis a serem obtidas na inversão, a FWI é proposta como um problema de otimização local, onde deve ser usado tanto o valor da função objetivo como o gradiente desta.

Nesse capítulo é feita uma revisão dos fundamentos da modelagem sísmica e da FWI. O capítulo está organizado da seguinte forma: Na seção 2.2 é descrito o teorema de Shannon o qual é muito usado no tratamento de sinais. A importância deste teorema é evidenciada na implementação do FWI 3D apresentada no capítulo 4. Na seção 2.3 é descrito como é resolvido o problema direto através da solução da onda. Depois, na seção 2.4 é feita uma revisão de otimização linear. Por último, na seção 2.5 é apresentado o conceito de função objetivo para a FWI assim como o cálculo do seu gradiente.

2.2 Amostragem de sinais digitais - Teorema de Shannon

O teorema de Shannon é um dos mais importantes na análise de sinais, uma vez que relaciona o conteúdo de frequência do sinal com a frequência da amostragem, além disso fornece uma fórmula para interpolar o sinal amostrado. O teorema de Shannon é formulado da seguinte maneira (Damelin & Miller 2009): Seja $g(t)$ um sinal de *banda limitada*, ou seja, seu espectro de frequência só possui valor diferente de zero para todo $f < f_{max}$, onde f_{max} é a frequência máxima do sinal. Se o sinal $g(t)$ é amostrado com uma frequência de amostragem $f_s = 1/\Delta t$, sendo Δt o intervalo de tempo entre amostras, então f_{max} deve satisfazer

$$f_{max} \leq f_{ny} = \frac{1}{2\Delta t} = \frac{f_s}{2}. \quad (2.1)$$

Onde a f_{ny} é conhecida como a frequência de Nyquist e fornece o limite de

frequência que pode ser medida dada certa taxa de amostragem. Além disso, se a condição da equação (2.1) é satisfeita, então o teorema de Shannon fornece uma fórmula para obter $g(t)$ em qualquer instante de tempo a partir das amostras temporais $g(k\Delta t)$, dada por:

$$g(t) = \sum_{k=-\infty}^{\infty} g(k\Delta t) \text{sinc}\left(\frac{t}{\Delta t} - k\right), \quad (2.2)$$

onde a função sinc é definida como:

$$\text{sinc}(t) = \begin{cases} \frac{\sin(\pi t)}{\pi t} & t \neq 0 \\ 1 & t = 0. \end{cases} \quad (2.3)$$

A equação (2.2) pode ser usada para reamostrar o sinal com um menor intervalo de amostragem Δt . Porém, aplicar a fórmula de interpolação é difícil na prática, devido ao custo computacional que é da ordem de $N_e * N_s$, sendo N_e o número de amostras do vetor de entrada e N_s o número de amostras do vetor de saída e $N_s \geq N_e$. Uma forma de usar a fórmula de interpolação do teorema de Shannon é aplicando um equivalente deste no domínio da frequência. Esta interpolação é descrita no algoritmo 1.

Algoritmo 1: Interpolação Sinc

Entrada: $x_e(t)$: vetor de entrada amostrado com Δt_e ; N_e : número amostras do vetor de entrada; Δt_s : Amostragem de saída.

- 1 Calcular número de amostras do vetor de saída $N_s = \Delta t_s / \Delta t_e$;
 - 2 Calcular a transformada de Fourier para vetores de entrada real (rfft em python) $x_e(f) = \mathcal{F}\{x_e(t)\}$. Este vetor possui $N_e/2 + 1$ amostras;
 - 3 Calcular a transformada $x_s(f)$ colocando zeros no final do vetor $x_e(f)$ até o vetor ter $N_s/2 + 1$ amostras;
 - 4 Calcular o vetor de saída $x_s(t) = \frac{N_s}{N_e} \mathcal{F}^{-1}\{x_s(f)\}$. \mathcal{F}^{-1} transformada inversa de Fourier ;
 - 5 **retorna** $x_s(t)$;
-

2.3 Problema direto - Solução numérica da equação da onda

Em qualquer problema de inversão, o problema direto é definido como a forma que são simulados os dados computacionalmente para serem comparados com os dados medidos a serem invertidos. O problema direto a ser resolvido depende, portanto, do tipo de dados que são medidos. Por exemplo, no caso da tomografia sísmica, onde o dado a ser invertido são os tempos de chegada das ondas sísmicas, o dado sintético é simulado através da solução numérica da equação eikonal (Nolet 2008).

No caso da FWI, os dados a serem invertidos consistem nas amplitudes da onda coletadas nos receptores espalhados ao longo da área a ser explorada. A suposição principal da FWI é que estas amplitudes podem ser descritas através da equação da onda. No caso dessa tese, é usada a equação de onda acústica, a qual, além de ser a mais fácil de implementar, é capaz de explicar uma ampla variedade de dados sísmicos com grande precisão.

A equação da onda acústica descreve a evolução temporal e espacial de uma onda que propaga-se através de um meio que possui densidade $\rho(\mathbf{x})$ e compressibilidade volumétrica $\kappa(\mathbf{x})$ espacialmente variáveis. A quantidade calculada pela equação da onda é o campo de pressão $p(\mathbf{x}, t)$ que é gerada por uma distribuição espacial e temporal de fontes $s(\mathbf{x}, t)$. Com isto a equação de onda pode ser formulada como

$$\nabla \cdot \left[\frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \right] - \kappa(\mathbf{x}) \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} = -s(\mathbf{x}, t). \quad (2.4)$$

Definindo a velocidade de onda acústica como $v^2(\mathbf{x}) = 1/(\kappa(\mathbf{x})\rho(\mathbf{x}))$, então a equação de onda é escrita como

$$\nabla \cdot \left[\frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \right] - \frac{1}{\rho(\mathbf{x})v(\mathbf{x})^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} = -s(\mathbf{x}, t). \quad (2.5)$$

No caso de um meio ter densidade constante ou aproximadamente constante a equação de onda acústica pode ser escrita como

$$\nabla^2 p(\mathbf{x}, t) - \frac{1}{v(\mathbf{x})^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} = -S(\mathbf{x}, t). \quad (2.6)$$

Onde $S(\mathbf{x}, t) = \rho s(\mathbf{x}, t)$ é o termo fonte corrigido pelo valor da densidade. A equação (2.5) é usada quando na inversão procura-se obter a densidade e a velocidade, enquanto que a equação (2.6) é usada no caso quando procura-se obter unicamente a velocidade de onda. Nos experimentos numéricos apresentados nessa tese o resultado da inversão são as velocidades de onda, e portanto, é apresentado na próxima seção a forma de resolver a equação (2.6) numericamente através do método das diferenças finitas. O material descrito a seguir é baseado na compilação de vários artigos, por exemplo Alford et al. 1974, Kelly et al. 1976 e Marfurt 1984, ou livros tais como Schuster 2017 e Ikelle & Amundsen 2005.

2.3.1 Método das diferenças finitas

Nessa seção é apresentado o método de diferenças finitas para resolver a equação da onda. Por simplicidade, e sem perda de generalidade, a metodologia é mostrada para o caso 2D, mas pode ser estendida diretamente para o caso 3D. Na metodologia de diferenças finitas o campo de onda é discretizado em uma malha computacional 2D como mostrado na Figura 2.1 onde cada quadrado representa uma matriz que contém um campo de onda em determinado instante de tempo. O tamanho da matriz de campo de onda deve ser igual ao tamanho da matriz de velocidades de onda.

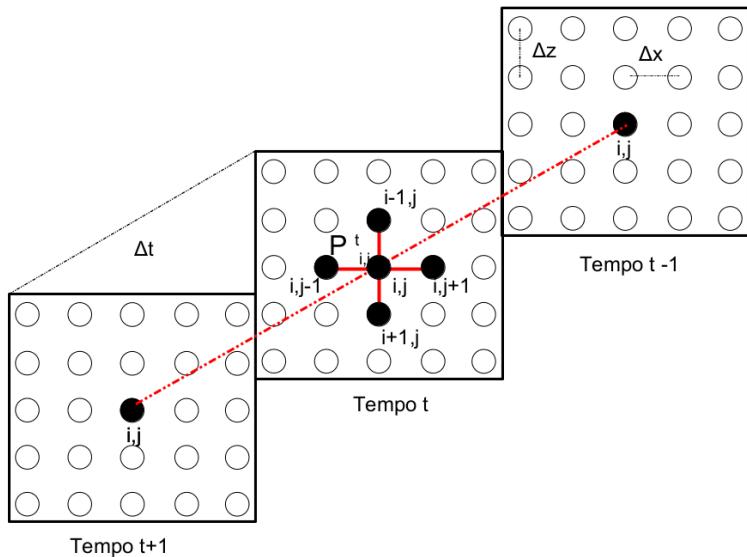


Figura 2.1: Representação do esquema de diferenças finitas para a equação da onda 2D usando um laplaciano de segunda ordem.

Cada círculo pequeno representa um elemento da matriz que vai conter um valor do campo de onda e esta identificado pelos índices (i, j) , com i relacionado à coordenada z e j relacionado à x . A distância entre células na direção horizontal é Δx e na vertical é Δz . O intervalo temporal entre duas matrizes representando dois campos de onda seguidos é Δt_p . No caso da equação da onda o método de diferenças finitas é dividido em três partes: o cálculo do laplaciano, o cálculo da derivada temporal e a definição da fonte.

Laplaciano

O cálculo do laplaciano é feito com as amplitudes de onda de um determinado instante de tempo. Para cada ponto é escolhida uma vizinhança de pontos cujo tamanho depende da ordem de aproximação da derivada. Por exemplo, na Figura 2.1, o laplaciano é calculado no ponto (i, j) usando os quatro vizinhos pretos, isto representando uma aproximação de segunda ordem. Para esta aproximação o laplaciano é dada pela seguinte equação

$$(\nabla^2 P)_{i,j}^t \approx \frac{P_{i,j+1}^t - 2P_{i,j}^t + P_{i,j-1}^t}{\Delta x^2} + \frac{P_{i+1,j}^t - 2P_{i,j}^t + P_{i-1,j}^t}{\Delta z^2}. \quad (2.7)$$

O índice t indica que as quantidades usadas correspondem ao campo no tempo t . A ordem usada para a aproximação de diferenças finitas depende do conteúdo de frequência no sinal e distância entre células. Entre maior for a frequência a ser propagada maior deve ser a ordem do laplaciano usado.

Derivada temporal

A derivada temporal é discretizada seguindo uma aproximação de segunda ordem. Isto implica que para calcular um campo de onda em um determinado instante é necessário ter disponíveis os campos de onda nos dois instantes anteriores como mostra na Figura 2.1. A discretização da derivada temporal é calculada da forma

$$\left(\frac{\partial^2 P}{\partial t^2} \right)_{i,j} = \frac{P_{i,j}^{t+1} - 2P_{i,j}^t + P_{i,j}^{t-1}}{\Delta t_p^2}. \quad (2.8)$$

Definição da assinatura da fonte

A assinatura da fonte é definida como a forma de onda que possui a fonte no instante que é gerada a onda sísmica. As fontes sísmicas são simuladas como pontuais, ou seja, são inseridas em um único ponto da matriz do campo de onda, o qual é equivalente à situação real que acontece em no levantamento sísmico. Uma das formas de onda mais usadas na pesquisa para simular a assinatura da fonte é a wavelet de Ricker (Wang 2015) a qual é definida pela equação

$$s(t) = (1 - 2\pi^2 f_p^2 (t - t_0)^2) \exp(-\pi^2 f_p^2 (t - t_0)^2), \quad (2.9)$$

onde $t_0 = \frac{6}{\pi f_p \sqrt{2}}$ é um tempo de deslocamento do tempo zero de tal forma que as amplitudes representativas da wavelet estejam em $t > 0$ e f_p é a frequência de pico, ou seja, a frequência com maior energia. A wavelet de Ricker é banda limitada e as frequências representativas estão entre 0 e $3f_p$. No caso da FWI aplicada em um dado real, a assinatura da fonte é extraída diretamente do próprio dado real, através do procedimento descrito na seção 4.4.

Equação de evolução temporal

Juntando a definição do laplaciano, da derivada temporal na equação (2.8) e da definição da assinatura da fonte, é obtida uma equação para calcular o campo futuro usando a informação dos campos presente e passado. Substituindo (2.8) em (2.6) é obtido que

$$\frac{1}{v_{i,j}^2} \frac{P_{i,j}^{t+1} - 2P_{i,j}^t + P_{i,j}^{t-1}}{\Delta t_p^2} = (\nabla^2 P)_{i,j}^t + s_{i_f,j_f}^t. \quad (2.10)$$

Isolando para $P_{i,j}^{t+1}$ é obtida equação básica de propagação.

$$P_{i,j}^{t+1} = 2P_{i,j}^t - P_{i,j}^{t-1} + v_{i,j}^2 \Delta t_p^2 \left((\nabla^2 P)_{i,j}^t + s_{i_f,j_f}^t \right), \quad (2.11)$$

onde o termo $(\nabla^2 P)_{i,j}^t$ pode ser qualquer aproximação do laplaciano. A equação (2.11) calcula explicitamente o campo pressão no instante t , usando a informação nos instantes t e $t - 1$ e tendo como parâmetros de propagação a matriz de velocidades $v_{i,j}$, o intervalo temporal Δt , e as distâncias entre células Δx e Δz como mostrado na Figura 2.1

2.3.2 Criterios de estabilidade

A equação (2.11) define como calcular um campo de onda a partir dos dois campos de onda anteriores. Como esta expressão calcula explicitamente os valores do campo sem necessidade de resolver um sistema de equações, então podem apresentar instabilidade numérica se os parâmetros de propagação não satisfazem dois critérios de estabilidade descritos a seguir.

Condição de estabilidade numérica

O critério de estabilidade numérica fornece uma condição que devem satisfazer os parâmetros de propagação para que a solução numérica possua convergência. A análise de Von Neumann é a técnica mais usada para analisar a estabilidade numérica de fórmulas explícitas como a equação (2.11) (de Moura & Kubrusly 2013). Para o caso 2D e fazendo uso da análise de estabilidade de Von Neumann é obtida a condição CFL para a propagação 2D ¹:

$$\Delta t_p \leq \sqrt{\frac{4}{\sum |C_k|}} \left(\frac{h_{efe}}{v_{max}} \right). \quad (2.12)$$

onde v_{max} é o valor máximo da matriz de velocidade, C_k representa os coeficientes de diferenças finitas da aproximação da segunda derivada espacial e h_{efe} é o tamanho de célula efetivo o qual é calculado em 2D como

$$\frac{1}{h_{efe}} = \sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta z^2}} \quad (2.13)$$

e em 3D como

$$\frac{1}{h_{efe}} = \sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}}. \quad (2.14)$$

A título de exemplo, usando uma aproximação de quarta ordem na derivada espacial e $\Delta x = \Delta z = h$ a condição CFL em 2D é

$$\Delta t_p < \sqrt{\frac{3}{4}} \left(\frac{h_{efe}}{v_{max}} \right). \quad (2.15)$$

Essa condição é usada para calcular o passo de tempo adequado Δt_p para a propagação, o qual geralmente é menor que o intervalo de tempo de amostragem do dado sísmico (Para uma discussão mais ampla deste fato, ver seção 4.5.1).

Condição de estabilidade espacial

Mesmo satisfazendo a condição CFL, a solução numérica da equação de onda pode gerar artefatos numéricos por causa de uma amostragem espacial pobre. Um

¹CFL refere-se às iniciais dos descobridores da condição, os matemáticos Courant, Friedrich e Lewy.

critério empírico usado na literatura é que o comprimento de onda mínimo gerado na propagação deve ser maior ou igual a certo número D de vezes o tamanho da célula (Ver Alford et al. 1974 ou Kelly et al. 1976). Isto implica que

$$\lambda_{min} = v_{min}/(f_{max}) \geq D \min(\Delta x, \Delta z). \quad (2.16)$$

Dado um determinado modelo de velocidade, esta condição permite calcular a máxima frequência f_{max} que pode ser propagada no modelo sem que haja dispersão numérica, como

$$f_{max} \leq v_{min}/(D \times \min(\Delta x, \Delta z)). \quad (2.17)$$

A equação (2.17) pode ser interpretada de duas formas: a primeira, é que dado um sinal com determinada frequência máxima, a equação (2.17) fornece uma fórmula para calcular o espaçamento entre células h que deve ter uma matriz de velocidade. A segunda forma é que dado um modelo de velocidade com certo $h = \max(\Delta x, \Delta z)$, a fórmula permite calcular a frequência máxima do sinal que pode ser propagado com o modelo. O valor de D depende da ordem aproximação de diferenças finitas do laplaciano. Para quarta ordem $D \approx 5$, para sexta ordem $D \approx 4$ e para oitava ordem $D \approx 3.5$.

2.3.3 Simulação de um meio infinito

Se somente fosse usada a equação (2.11) para resolver a equação da onda, seriam observadas ondas refletidas pelas bordas da matriz do modelo de velocidade, o qual não representa uma situação real onde a onda continua propagando-se. Para simular um meio infinito, ou seja, um meio onde não existe reflexão nas bordas, são usadas técnicas tais como condições de contorno absorventes (Clayton & Engquist 1977), onde é formulada uma condição de contorno especial para simular o meio infinito. Outra forma é usando a borda de esponja (Cerjan et al. 1985) onde o modelo de velocidade é estendido em todas as direções em um certo número de células. Na extensão do modelo as amplitudes das onda são multiplicadas por um função exponencial decrescente. Nesse trabalho é usado o método CPML,² descrito nos trabalhos Komatitsch & Martin 2007 e Pasalic & McGarry 2010, o qual é uma melhora do método PML original introduzido por Berenger 1994. O método CPML modifica a equação de onda dentro da borda absorvente, multiplicando as derivadas espaciais por um fator de estiramento s que depende da frequência angular ω definido como

$$s_i = 1 + \frac{d_i}{\alpha_i + i\omega}. \quad (2.18)$$

²Do inglês Convolutional Perfectly Marching Layer.

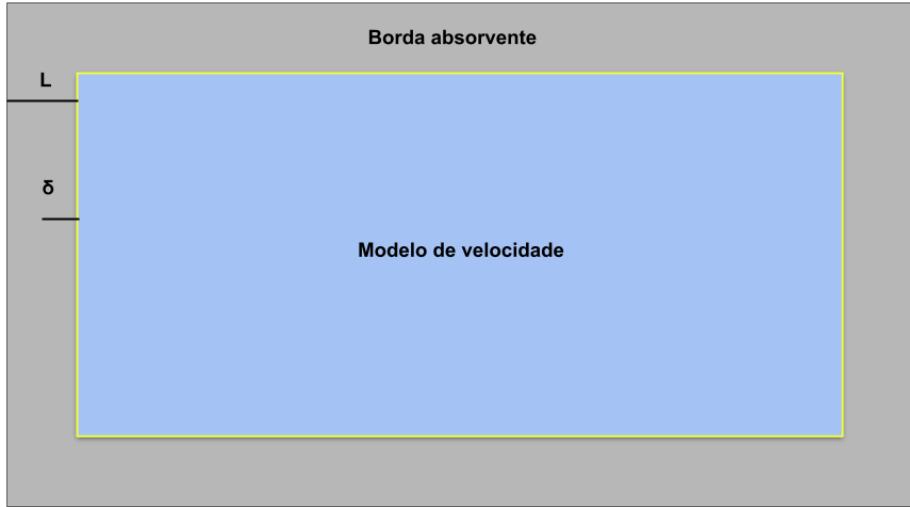


Figura 2.2: Extensão do modelo de velocidade para aplicar a borda absorvente CPML.

O índice i representa uma das coordenadas x ou z, d_i e α_i são definidos como

$$\begin{aligned} d_i &= d_0 \left(\frac{\delta}{L} \right)^2, \\ d_0 &= \frac{3}{2L} v_{max} \log(R), \\ \alpha_i &= \pi f_0 \frac{\delta}{L}, \end{aligned} \quad (2.19)$$

onde δ é a distância de um ponto na borda CPML até a borda da matriz de velocidade, L é a largura da borda PML (ver Figura 2.2). Além disso, v_{max} é a velocidade máxima do modelo de velocidades, R é um coeficiente de reflexão cujo valor é escolhido como $R = 10^{-5}$ e f_0 é a frequência dominante do sinal.

Como a derivada é multiplicada no domínio da frequência pelo fator s , isto no domínio do tempo é aplicado como uma convolução. Em princípio a convolução temporal deve ser feita com o campo de onda, mas como este campo é a variável a ser calculada, então o efeito da convolução é feito através das denominadas variáveis de memória. Nesse caso, no domínio do tempo a derivada é modificada da seguinte forma

$$\frac{\partial}{\partial i} = \frac{\partial}{\partial i} + \psi_i, \quad (2.20)$$

onde ψ_i é a variável de memória a qual é calculada de forma recursiva como

$$\psi_i^t = b_i \psi_i^{t-1} + a_i \left(\frac{\partial}{\partial i} \right)^t \quad (2.21)$$

onde o supraíndice t representa o valor no instante atual e os valores a_i e b_i é definido como (Ver Komatitsch & Martin 2007)

$$a_i = \frac{d_i}{d_i + \alpha_i} (b_i - 1); \quad b_i = e^{-(d_i + \alpha_i) \Delta t}, \quad (2.22)$$

onde os fatores d_i α_i são definidos na equação (2.19). Com essa modificação das derivadas, a equação de onda dentro da borda é calculada como

$$\frac{1}{v(\mathbf{x})^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} = \nabla^2 p(\mathbf{x}, t) + \frac{\partial \psi_x}{\partial x} + \frac{\partial \psi_z}{\partial z} + \zeta_x + \zeta_z \quad (2.23)$$

onde a variável ζ_i é também uma variável de memória para a segunda derivada, a qual também satisfaz uma relação recursiva

$$\zeta_i^t = b_i \zeta_i^{t-1} + a_i \left[\left(\frac{\partial^2 p}{\partial i^2} \right)^t + \left(\frac{\partial \psi_i}{\partial i} \right)^t \right] \quad (2.24)$$

2.4 Formulação do problema inverso

2.4.1 Revisão de otimização não-linear

Obter o mínimo de uma função de varias variáveis sem conhecer a forma analítica dela, ou seja, obter o mínimo numericamente, está relacionado com o ramo da matemática chamado *otimização não linear*. O método para otimizar uma função depende da quantidade de variáveis do problema. No caso da FWI, o número de variáveis é da ordem de $10^4 - 10^6$. Para este tipo de problemas os métodos mais usados são os de otimização local, os quais fornecem um mínimo global próximo ao valor inicial \mathbf{m}_0 , que não necessariamente é o mínimo global da função.

No método de otimização local é feita a suposição que o modelo corregido \mathbf{m} pode ser escrito como o modelo inicial mais uma correção $\Delta\mathbf{m}$. O objetivo dos algoritmos de otimização é calcular o valor da correção $\Delta\mathbf{m}$ a partir de valores no ponto \mathbf{m}_0 . Para isto é feita uma expansão em serie de Taylor da função $\Phi(\mathbf{m})$, ao redor do ponto \mathbf{m}_0 (Virieux & Operto 2009), obtendo que

$$\Phi(\mathbf{m}) = \Phi(\mathbf{m}_0) + \sum_{i=1}^M \frac{\partial \Phi(\mathbf{m}_0)}{\partial m_i} \Delta m_i + \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \frac{\partial^2 \Phi(\mathbf{m}_0)}{\partial m_i \partial m_j} \Delta m_i \Delta m_j + \mathcal{O}(\mathbf{m}^3). \quad (2.25)$$

Derivando esta equação com relação a uma componente genérica qualquer m_k e desprezando os termos de terceira ordem na derivada é obtido que

$$\frac{\partial \Phi(\mathbf{m}_0 + \Delta\mathbf{m})}{\partial m_k} = \frac{\partial \Phi(\mathbf{m}_0)}{\partial m_k} + \sum_{i=1}^M \frac{\partial^2 \Phi(\mathbf{m}_0)}{\partial m_i \partial m_k} \Delta m_i \quad (2.26)$$

Como a correção $\Delta\mathbf{m}$ deve fazer com que a função caia em um mínimo, então o

termo do lado esquerdo deve ser zero e, portanto

$$\frac{\partial \Phi(\mathbf{m}_0)}{\partial m_k} + \sum_{i=1}^M \frac{\partial^2 \Phi(\mathbf{m}_0)}{\partial m_i \partial m_k} \Delta m_i = 0 \quad (2.27)$$

Escrevendo em forma vetorial, a equação anterior fica como

$$\left(\frac{\partial \Phi(\mathbf{m}_0)}{\partial \mathbf{m}} \right)_{M \times 1} + \mathcal{H}_{M \times M} (\Delta \mathbf{m})_{M \times 1} = 0 \quad (2.28)$$

Onde $\mathcal{H}_{ik} = \frac{\partial^2 \Phi(\mathbf{m}_0)}{\partial m_i \partial m_k}$ é conhecida como a matriz Hessiana e o vetor $\left(\frac{\partial \Phi(\mathbf{m}_0)}{\partial \mathbf{m}} \right)$ é o gradiente da função no ponto inicial. Os subíndices representam as dimensões das matrizes. Da última equação pode-se calcular o termo de correção como

$$\Delta \mathbf{m} = -\mathcal{H}^{-1} \left(\frac{\partial \Phi(\mathbf{m}_0)}{\partial \mathbf{m}} \right), \quad (2.29)$$

a qual é a equação mais importante da otimização não linear. Em geral, a otimização é feita através de um processo iterativo onde uma nova interação é calculada a partir da última usando a seguinte equação

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha \mathcal{H}^{-1} \left(\frac{\partial \Phi(\mathbf{m}_k)}{\partial \mathbf{m}} \right), \quad (2.30)$$

onde α é conhecido como *tamanho do passo*. O tamanho do passo ótimo é calculado através dos denominados algoritmos de *busca em linha* (Nocedal & Wright 2006).

2.4.2 Métodos Newton e quase-Newton. Método L-BFGS

Quando a correção mostrada na equação 2.29 é calculada diretamente se denomina que o método de otimização é de tipo *Newton*. Como explicado em Pratt et al. 1998 a Hessiana pode ser decomposta em duas componentes linear e não linear. Quando somente a componente linear da Hessiana é calculada o método de otimização se denomina *Gauss-Newton*. Porém, como na FWI o número de variáveis obtida na inversão é da ordem de 10^4 a 10^6 , o cálculo da Hessiana e sua inversa resulta difícil na prática. Portanto, na FWI é comum usar os denominados métodos *quase-Newton*, os quais visam aproximar o produto da inversa da Hessiana pelo gradiente. Entre os métodos quase-Newton mais populares estão o de gradientes conjugados e o L-BFGS (Nocedal & Wright 2006). Nesse trabalho é explicado com mais detalhe o L-BFGS por ser o mais usado durante o desenvolvimento do mesmo.

O algoritmo L-BFGS, introduzido por Nocedal & Wright 2006, é a versão de baixa memória do algoritmo BFGS, nomeado pelos seus descobridores Broyden, Fletcher, Goldfarb e Shanno. Diferente do algoritmo BFGS, o L-BFGS não precisa estimar em algum momento o inverso da Hessiana. Ao invés disso calcula o produto inversa da Hessiana-gradiente de uma iteração determinada com os valores dos gradientes

e das soluções obtidas nas l iterações anteriores³. O cálculo do produto inversa da Hessiana-gradiente feito pelo L-BFGS é mostrado no algoritmo 2. Esse cálculo pode usar uma estimativa inicial da diagonal da inversa da Hessiana H_k^0 a qual é aplicada no passo 6. A estratégia completa de otimização usando o L-BFGS é apresentada no algoritmo 3. Os métodos de busca em linha mencionados no passo 5 do algoritmo 3 estão descritos em Nocedal & Wright 2006, capítulo 3.

Algoritmo 2: Algoritmo L-BFGS. Estimativa do produto Hessiana-gradiente.

Entrada: ∇f_k : gradiente na iteração k ; $s_i = x_{i+1} - x_i$: diferença duas soluções consecutivas ; $y_i = \nabla f_{i+1} - \nabla f_i$: diferença entre dois gradientes consecutivos; $\rho_i = (y_i^T s_i)$: escalar ; H_k^0 : Estimativa inicial da Hessiana; São usados os últimos l gradientes e soluções.

```

1  $q \leftarrow \nabla f_k;$ 
2 para  $i = k - l, k - l + 1, \dots, k - 1$  faca
3    $\alpha_i \leftarrow \rho_i s_i^T q;$ 
4    $q \leftarrow q - \alpha_i y_i;$ 
5 fim
6  $r \leftarrow H_k^0 q;$ 
7 para  $i = k - l, k - l + 1, \dots, k - 1$  faca
8    $\beta \leftarrow \rho_i y_i^T r;$ 
9    $r \leftarrow r - s_i(\alpha_i - \beta);$ 
10 fim
11 retorna  $H_k \nabla f_k = r;$ 
```

2.5 Função objetivo da FWI e seu gradiente

2.5.1 Função objetivo

A FWI visa obter um conjunto de parâmetros ou modelo da subsuperfície composto por quantidades que afetam as amplitudes e cinemática da onda. O modelo é representado por um vetor \mathbf{m} . O parâmetro mais usado é a velocidade da propagação da onda P, mas em geral, pode ser usado também a densidades, velocidades de ondas S e fatores de atenuação em cada ponto da subsuperfície.

A suposição principal é que o modelo correto é aquele que optimiza, ou seja minimiza ou maximiza, uma função $\Phi = \Phi(\mathbf{u}(\mathbf{m}), \mathbf{d})$ a qual compara os dados observados no campo \mathbf{d} com dados \mathbf{u} simulados computacionalmente, os quais dependem do modelo \mathbf{m} . A função Φ , chamada de *função objetivo*, é um quantidade escalar a qual é

³Na literatura esse parâmetro é chamado m mas por consistência com os próximos capítulos é chamado de l .

Algoritmo 3: Algoritmo L-BFGS. Procedimento completo.

Entrada: x_0 : Ponto inicial; l : número de gradientes e soluções a serem armazenadas na memória.

```

1  $k \leftarrow 0$ ;
2 repita
3   Escolher  $H_k^0$ ;
4   Calcular  $p_k \leftarrow -H_k \nabla f_k$  usando o algoritmo 2;
5   Calcular  $x_{k+1} \leftarrow x_k + \alpha_k p_k$ , onde  $\alpha_k$  satisfaz os critérios da busca em
      linha;
6   se  $k > m$  então
7     | Descartar os vetores  $s_{k-m}$  e  $y_{k-m}$  do armazenamento;
8   fim
9   Calcular e salvar  $s_k = x_{k+1} - x_k$  e  $y_i = \nabla f_{i+1} - \nabla f_i$ ;
10   $k \leftarrow k + 1$ ;
11 até Convergência;
```

definida matematicamente como:

$$\begin{aligned}\Phi: \mathbb{R}^M &\rightarrow \mathbb{R}^N \rightarrow \mathbb{R} \\ \mathbf{m} &\rightarrow \mathbf{u}(\mathbf{m}) \rightarrow \Phi(\mathbf{u}(\mathbf{m}), \mathbf{d}),\end{aligned}\tag{2.31}$$

onde M é o número de componentes do modelo \mathbf{m} e N é o número de componentes de \mathbf{u} . Os valores de M e N são na ordem de $M \approx 10^4$ e $N \approx 10^3$.

A função objetivo mais usada na FWI é o quadrado da norma ℓ_2 da diferença entre os dados simulados e os dados observados. Matematicamente, a função objetivo usando a norma ℓ_2 é definida como

$$\Phi_{\ell_2} = \frac{1}{2} \sum_{i=1}^{N_s} \|\mathbf{u}(\mathbf{m}) - \mathbf{d}\|_{\ell_2}^2,\tag{2.32}$$

onde a soma é feita ao longo dos N_s experimentos sísmicos ou fontes.

2.5.2 Gradiente da função objetivo

No domínio do tempo a função objetivo Φ para os traços relativos a uma fonte s , pode ser escrita em forma integral como

$$\Phi = \sum_{i=1}^{N_g} \int_0^T dt \phi(\mathbf{u}(\mathbf{m})_i, \mathbf{d}_i),\tag{2.33}$$

onde N_g é número de receptores que receberam as ondas geradas pela fonte s . Para calcular o gradiente da função custo, é calculada a derivada com relação a uma

componente genérica do modelo m_l obtendo-se pela regra da cadeia

$$\frac{\partial \Phi}{\partial m_l} = \sum_{i=1}^{N_g} \int_0^T dt \frac{\partial \mathbf{u}_i}{\partial m_l} \frac{\partial \phi}{\partial \mathbf{u}_i}. \quad (2.34)$$

O termo $\frac{\partial \phi}{\partial \mathbf{u}_i}$ pode ser calculado analiticamente da definição específica da função objetivo, enquanto o termo $\frac{\partial \mathbf{u}_i}{\partial m_l}$ é a denominada *derivada de Fréchet*. O gradiente poderia ser calculado diretamente a partir da equação (2.34), porém o custo computacional seria muito alto devido a que seria necessário realizar uma propagação de onda por cada componente da derivada de Fréchet. Para calcular o efeito da derivada de Fréchet pode ser usada a *aproximação de Born* (Virieux & Operto 2009 e Yang 2014). Partindo da equação da onda para a fonte s

$$m \frac{\partial^2 \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s, 0)}{\partial t^2} - \nabla^2 \mathbf{u} = \mathbf{s}(\mathbf{x}_s, 0), \quad (2.35)$$

onde $m = 1/v^2$ é o inverso do quadrado da velocidade. A suposição principal na aproximação de Born é que uma pequena perturbação Δm do modelo de velocidade m vai gerar uma perturbação nas amplitudes $\Delta \mathbf{u}$ do campo de onda. Nesse caso a equação da onda fica como

$$(m + \Delta m) \frac{\partial^2 \mathbf{u} + \Delta \mathbf{u}}{\partial t^2} - \nabla^2 (\mathbf{u} + \Delta \mathbf{u}) = \mathbf{s}(\mathbf{x}_s, 0). \quad (2.36)$$

Combinando as equações (2.35) e (2.36) e mantendo somente os termos de primeira ordem é obtida uma equação da onda para a perturbação $\Delta \mathbf{u}$

$$m \frac{\partial^2 \Delta \mathbf{u}}{\partial t^2} - \nabla^2 (\Delta \mathbf{u}) = -\Delta m \frac{\partial^2 \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s, 0)}{\partial t^2} \quad (2.37)$$

O termo ao lado direito é conhecido como fonte virtual. Agora, usando o fato que

$$\Delta \mathbf{u} \approx \frac{\partial \mathbf{u}}{\partial m} \Delta m. \quad (2.38)$$

então deve ser calculado $\Delta \mathbf{u}$ na posição do receptores em função de Δm e o termo que o multiplica este Δm será a derivada de Fréchet. Para calcular $\Delta \mathbf{u}$ é resolvida a equação (2.37) com ajuda do método da função de Green (Arfken & Weber 2005) a qual é definida como o campo de onda produzido por uma fonte pontual em uma posição arbitrária \mathbf{x}_i , ou seja que satisfaz a equação

$$m \frac{\partial^2 \mathbf{G}(\mathbf{x}, t; \mathbf{x}_i, 0)}{\partial t^2} - \nabla^2 \mathbf{G}(\mathbf{x}, t; \mathbf{x}_i, 0) = \delta(\mathbf{x} - \mathbf{x}_i) \delta(t). \quad (2.39)$$

Com a função de Green pode ser calculada a solução da equação (2.37) na posição

dos receptores como

$$\Delta \mathbf{u}(\mathbf{x}_i, t; \mathbf{x}_s, 0) = \left(- \int_{\Gamma} d\mathbf{x} \mathbf{G}(\mathbf{x}_i, t; \mathbf{x}, 0) * \frac{\partial^2 \mathbf{u}}{\partial t^2}(\mathbf{x}, t; \mathbf{x}_s, 0) \right) \Delta m. \quad (2.40)$$

onde o operador $*$ representa à convolução temporal. As posições de fonte e receptor na função de Green são trocadas por causa do princípio de reciprocidade. Na equação (2.40) o termo entre parêntese corresponde à derivada de Fréchet. Usando a equação (2.40) na equação (2.34) é obtido que

$$\frac{\partial \Phi}{\partial m_l} = - \sum_{i=1}^{N_g} \int_0^T dt \int_{\Gamma} d\mathbf{x} \mathbf{G}(\mathbf{x}_i, t; \mathbf{x}, 0) * \frac{\partial^2 \mathbf{u}}{\partial t^2} \left(\frac{\partial \phi}{\partial \mathbf{u}_i} \right). \quad (2.41)$$

Agora usando a seguinte propriedade da convolução

$$\int dt (f(t) * g(t)) h(t) = \int dt f(t) (g(-t) * h(t)), \quad (2.42)$$

a equação (2.41) pode ser expressada como

$$\frac{\partial \Phi(\mathbf{u}(\mathbf{m}))}{\partial m_l} = - \sum_{i=1}^{N_g} \int_0^T dt \frac{\partial^2 \mathbf{u}}{\partial t^2} \int_{\Gamma} d\mathbf{x} \left(\mathbf{G}(\mathbf{x}_i, -t; \mathbf{x}, 0) * \frac{\partial \phi}{\partial \mathbf{u}_i} \right) \quad (2.43)$$

Como o campo $\frac{\partial^2 \mathbf{u}}{\partial t^2}$ não depende da posição dos receptores, então pode ser definida a quantidade

$$\mathbf{R} = \sum_{i=1}^{N_g} \int_{\Gamma} d\mathbf{x} \mathbf{G}(\mathbf{x}_i, -t; \mathbf{x}, 0) * \frac{\partial \phi}{\partial \mathbf{u}_i}, \quad (2.44)$$

é obtida a forma mais conhecida do gradiente

$$\frac{\partial \Phi}{\partial m_l} = - \int_0^T \frac{\partial^2 \mathbf{u}}{\partial t^2} \mathbf{R} dt. \quad (2.45)$$

O campo \mathbf{R} representa a propagação reversa no tempo (sinal negativo no argumento t na função de Green) ou *retropropagação* da quantidade $\frac{\partial \phi}{\partial \mathbf{u}_i}$, a qual é conhecida como *fonte adjunta* a qual é interpretada como o objetivo entre os dados observados e os dados simulados derivado da função ϕ .

Deve ser observado que a equação (2.45) representa uma correlação temporal entre um campo vindo do lado da fonte ($\frac{\partial^2 \mathbf{u}}{\partial t^2}$) e um campo vindo do lado do receptor (\mathbf{R}) o qual é equivalente matematicamente à migração reversa no tempo ou RTM⁴ (Pratt et al. 1998). Fazendo uma analogia com migração a qual posiciona o campo dos receptores na localização correta em profundidade, o gradiente posiciona em profundidade as diferenças entre os dados observados e simulados as quais estão em tempo.

⁴Siglas em inglês de Reverse Time Migration

Um exemplo do cálculo de fonte adjunta é o usado para a norma ℓ_2 . Usando a representação integral, a norma ℓ_2 é definida como

$$\Phi_{\ell_2} = \frac{1}{2} \int_0^T dt (\mathbf{u} - \mathbf{d})^2 \quad (2.46)$$

onde d é o dado observado. Para essa função objetivo a fonte adjunta é calculada como

$$\left(\frac{\partial \phi}{\partial \mathbf{u}} \right)_{\ell_2} = \frac{\partial}{\partial \mathbf{u}} \left(\frac{1}{2} (\mathbf{u} - \mathbf{d})^2 \right) = \mathbf{u} - \mathbf{d}. \quad (2.47)$$

Para a norma ℓ_2 a fonte adjunta é dada pela diferença entre dados simulados e observados.

Capítulo 3

Estratégias de Fontes Codificadas

Esse capítulo contém a publicação "Seismic shot-encoding schemes for waveform inversion" Duarte et al. 2020, que é requisito parcial para obter o título de doutor em física pela UFRN. As figuras usadas correspondem às usadas nessa publicação.

3.1 Introdução

Como mencionado no capítulo 1 e no capítulo 2, um dos procedimentos mais importantes para realizar a FWI é o cálculo do gradiente da função objetivo. O gradiente total é construído através da soma dos gradientes individuais para cada tiro, portanto, o custo computacional do cálculo do gradiente cresce linearmente com o número de tiros no levantamento sísmico.

O cálculo do gradiente de um tiro precisa de duas propagações de onda, como mencionado na seção 2.5.2: uma propagação direta gerada a partir da posição da fonte e uma propagação reversa no tempo onde a fonte adjunta é propagada com o eixo temporal invertido. Para um levantamento sísmico com N_S tiros é necessário, por tanto, realizar $2N_S$ propagações de onda para calcular o gradiente. Isto torna a FWI convencional um procedimento caro computacionalmente quando realizado sobre levantamento sísmico real com muitas fontes (Krebs et al. 2009, Wang & Rao 2009, Ben-Hadj-Ali et al. 2011, da Costa et al. 2018).

Uma estratégia para acelerar o cálculo do gradiente é através do uso dos denominados esquema de *fontes simultâneas*, onde os tiros no levantamento sísmicos são empilhados formando um super-tiro o qual é usado para realizar as propagações diretas e reversa (Krebs et al. 2009; Rao & Wang 2017). Usando fontes simultâneas, o número de propagações de onda é reduzido para somente dois por gradiente. A condição para usar a técnica de fontes simultâneas é que todas fontes compartilhem um mesmo conjunto de receptores como é o caso das aquisições tipo OBN. Porém, existem estratégias para usar fontes codificadas com tiros com diferentes conjuntos de receptores como as apresentadas em Choi & Alkhalifah 2012 e Son et al. 2014.

Apesar das vantagens computacionais, as estratégias de fontes codificadas possuem um dificuldade consistente na aparição de um ruído no gradiente chamado de

*crosstalk*¹ ou conversa cruzada em português, o qual aparece devido à correlação que existe entre os campos direto e reverso vindo de fontes diferentes, os quais são misturados na propagação simultânea. Para amenizar o ruído de crosstalk, Krebs et al. 2009, inspirado em uma técnica descrita por Romero et al. 2000 para migração, sugere usar a denominada *Codificação de fontes*, onde as amplitudes de cada tiro são modificadas de forma independente antes de ser somados para formar o super-tiro.

Os esquemas de codificação mais usados na FWI por fontes simultâneas são a codificação por polarização aleatória introduzida em Krebs et al. 2009 onde cada tiro é multiplicado aleatoriamente por +1 ou -1 e a codificação por estática aleatória onde a cada traço pertencente a um tiro é deslocado com um valor temporal fixo, escolhido aleatoriamente para cada tiro (Schuster et al. 2011). Além disso existem outras estratégias para amenizar o ruído de crosstalk como o uso de estratégias multiescala Wang 2016, agrupando os tiros em vários super-tiros ao invés de um único Castellanos et al. 2014 ou combinando os esquema de codificação com filtros sobre a imagem Schuster et al. 2010.

Nesse trabalho são propostas dois esquemas alternativos de codificação de fontes para a FWI no domínio do tempo. No primeiro esquema (I) é aplicada uma rotação de fase aos traços de cada tiro com ajuda da transformada Hilbert do sinal. O valor da fase é aleatório para cada tiro. O esquema por polarização aleatória (III) é um caso especial do esquema (I). No segundo esquema (II) é semelhante ao proposto em Schuster et al. 2011 (IV) mas o valor da estática é limitado até um período da frequência dominante dos dados observados. Nesse trabalho é demonstrado que esse novos esquemas melhoram a eficiência da FWI, acelerando a convergência da inversão. O esquema por polarização (III) é usado como referência para comparar o desempenho dos outros três esquemas (I, II e IV).

Devido a aleatoriedade inserida no processo de codificação, os algoritmos de otimização convencionais como o L-BFGS (descrito na seção 2.4.1) não podem ser usados como na FWI convencional. Portanto nesse trabalho foi usado o esquema do L-BFGS reinicializado proposto em Wang 2017. Esse capítulo esta organizado da seguinte maneira: Na seção 3.2 é descrita a função custo e o gradiente para fontes codificadas, identificando a origem do ruido de crosstalk. Na seção 3.3 são apresentadas em detalhe as novas estratégias de codificação. A seguir, na seção 3.4 é explicado o algoritmo do L-BFGS reinicializado, usado nesse trabalho. Por último na seção 3.5 são apresentados os experimentos numéricos mostrando a eficiência dos esquemas de codificação propostos.

3.2 Função objetivo e gradiente usando fontes simultâneas

Na FWI por fontes codificadas, a função objetivo é definida como

¹Por facilidade será usado o termo em inglês nesse trabalho.

$$\Phi = \frac{1}{2} \left\| P_{sim} - \sum_{i=1}^{N_s} e_i(t) \otimes p_{obs}(i) \right\|_{L2}^2, \quad (3.1)$$

onde $p_{obs}(i)$ é o dado observado para o tiro i e o termo $\sum_{i=1}^{N_s} e_i(t) \otimes p_{obs}(i)$ é o supertiro resultado do empilhamento dos tiros individuais codificadas usando do operador de codificação e_i , e P_{sim} é o supertiro sintético simulado através da solução da equação

$$\nabla^2 P_{sim} - \frac{1}{c^2(\mathbf{r})} \frac{\partial^2 P_{sim}}{\partial t^2} = \sum_{i=1}^{N_s} e_i(t) \otimes s(i, t). \quad (3.2)$$

O termo do lado direito corresponde ao empilhamento das assinaturas de cada fonte, colocadas nas suas respectivas posições e codificadas através do mesmo operador codificação e_i . O símbolo \otimes representa uma convolução e em geral o operador $e_i(t)$ depende do tempo.

Seguindo uma estratégia semelhante ao cálculo de gradiente apresentado na seção 2.5.2, o gradiente da função custo é calculado como

$$\nabla \Phi = - \int_0^T R_{sim}(\mathbf{r}, t) \frac{\partial^2 P_{sim}(\mathbf{r}, t)}{\partial t^2} dt, \quad (3.3)$$

onde R_{sim} é o campo de onda retropropagado o qual é calculado resolvendo o seguinte problema de propagação

$$\nabla^2 R_{sim} - \frac{1}{c^2(\mathbf{r})} \frac{\partial^2 R_{sim}}{\partial t^2} = P_{sim}(T-t) - \sum_{i=1}^{N_s} e_i(T-t) \otimes d(i, T-t). \quad (3.4)$$

O termo $T-t$ dentro do argumento significa que o termo fonte deve ser injetado em tempo reverso tal como descrito na seção 2.5.2. Devido à linearidade da equação da onda o gradiente na equação (3.3), pode ser decomposto da seguinte forma

$$\begin{aligned} \nabla \Phi = & - \sum_{i=1}^{N_s} \int_0^T e_i(-t) \otimes R_i(-t) \frac{\partial^2 e_i \otimes P_i(t)}{\partial t^2} dt \\ & - \sum_{i \neq j}^{N_s} \int_0^T e_i(-t) \otimes R_i(-t) \frac{\partial^2 e_j \otimes P_j(t)}{\partial t^2} dt \end{aligned} \quad (3.5)$$

O primeiro termo é a correlação dos campos direto e reverso relativos à mesma fonte i . Já o segundo termo mostra a correlação entre os campos direto e retropropagados vindo de fontes diferentes o qual corresponde ao ruído crosstalk.

Para que uma estratégia de codificação funcione ela deve satisfazer as duas propriedades a seguir:

- **Propriedade 1:** O primeiro termo na equação 3.5 deve reproduzir o gradiente da FWI convencional

$$\nabla \Phi_{conv} = - \sum_{i=1}^{N_s} \int_0^T R_i(\mathbf{r}, t) \frac{\partial^2 P_i(\mathbf{r}, t)}{\partial t^2} dt, \quad (3.6)$$

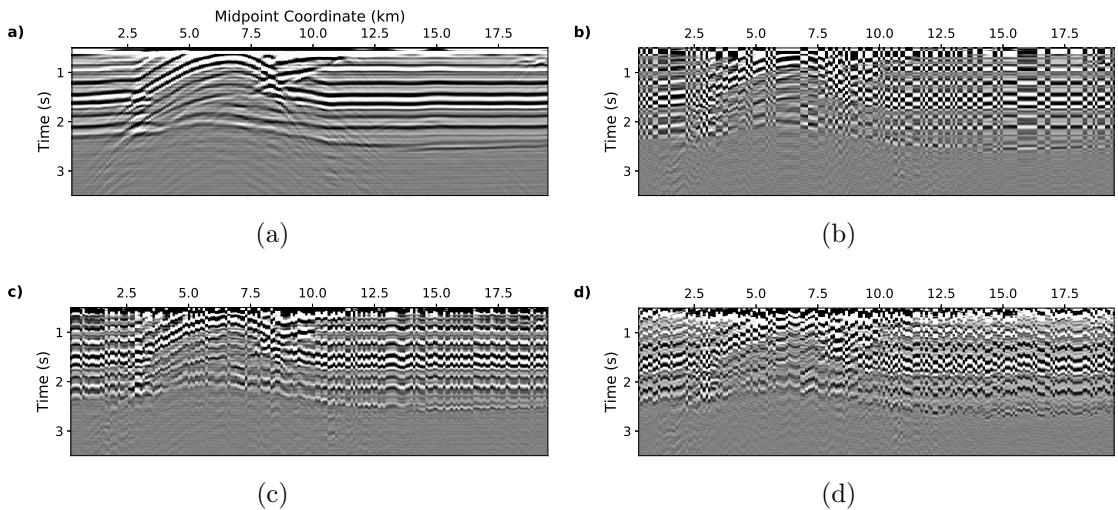


Figura 3.1: (a) Família de afastamento comum de 500m. Mesma família de afastamento comum com codificação (b) por polarização, (c) por rotação de fase aleatória e (d) por estática aleatória.

onde $P_i(\mathbf{r}, t)$ e $R_i(\mathbf{r}, t)$ são os campos direto e retropropagado para o tiro i . Na seguinte seção é demonstrado que as estratégias de codificação propostas satisfazem essa propriedade.

- **Propriedade 2:** O segundo termo na equação (3.5) deve cancelar-se de alguma forma. Em geral, como explicado em Krebs et al. 2009, Schuster et al. 2011 e Castellanos et al. 2014, o termo de crosstalk não é cancelado pela codificação mas é amortecido pelo esquema de otimização como é mostrado nos experimentos numéricos.

3.3 Esquemas de codificação de fontes

Como mencionado na introdução, nesse trabalho foram propostos dois esquemas de codificação. O primeiro é o esquema por rotação de fase aleatória (I) e o outro é a codificação por estática limitada a um período da frequência dominante (II). Esses dois esquemas são comparados com o esquema por polarização aleatória (III) e a codificação por estática sem ser limitada (IV).

Para ilustrar o efeito do esquemas de codificação de tiro é usado um agrupamento de traços de afastamento comum² sendo este igual a 500m. Cada traço do agrupamento representa um tiro. A Figura 3.1(a) é a família sem codificar onde pode ser observada a coerência dos refletores. Quando são aplicados os esquemas de codificação os refletores perdem continuidade, tal como observado nas Figuras 3.1(b) - 3.1(d). Essa perda de continuidade mostra a falta de coerência que possuem os campos de onda quando é feita as correlações cruzadas que fazem aparecer o ruído de crosstalk.

²Em inglês *Common Offset Gather*.

Codificação por rotação de fase (I)

No esquema de codificação por rotação de fase (I) todos os traços pertencentes a um tiro são rotacionados por um valor de fase aleatória φ variando no intervalo $0 \leq \varphi < \pi$. A rotação de fase de uma série temporal $x(t)$ é feita usando uma rotação no domínio do tempo usando a equação

$$x_{rot}(t; \varphi) = x(t) \cos \varphi + \mathcal{H}\{x(t)\} \sin \varphi, \quad (3.7)$$

onde \mathcal{H} é o operador da transformada de Hilbert (Barnes 2007). A codificação por polarização (III) é um caso especial da codificação por rotação de fase para os valores extremos $\cos \varphi = \pm 1$, ou seja quando os ângulos de fase são $\varphi = 0$ e $\varphi = \pi$. O valor de fase é constante para todos os traços dentro de um mesmo tiro mas é escolhido aleatoriamente para cada tiro. Depois de aplicar a rotação de fase para cada tiro, estes são somados seguindo a equação 3.1. Da mesma forma as assinaturas das fontes para cada tiro são rotacionadas com a mesma fase aplicada no dado observado correspondente.

Codificação por estática limitada (II)

No esquema de codificação por estática aleatória, os traços pertencentes a um tiro são deslocados temporalmente por um valor Δs . O valor Δs varia entre os tiros do levantamento sísmico. O nome estática provém do processamento de dados sísmicos terrestres (Yilmaz 2001), onde um valor de estática é aplicado nos traços para corrigir o efeito da topografia. A codificação por estática aleatória sem limitar o valor de estática foi proposta em Schuster et al. 2011 e nesse trabalho é dominada como a estratégia (IV). A codificação proposta aqui consiste em limitar o valor de estática aplicada de tal forma que o valor de estática satisfaça a condição

$$0 \leq \Delta s \leq \frac{1}{f_0}, \quad (3.8)$$

onde f_0 é a frequência dominante do dado sísmico observado.

Equivalência com codificação no domínio da frequência

Ambos os esquemas de codificação por rotação de fase e por estática aleatória são versões no domínio do tempo da codificação por fase utilizada na FWI no domínio da frequência, introduzido em Ben-Hadj-Ali et al. 2011, Castellanos et al. 2014, onde as amplitudes da transformada de Fourier dos traços sísmicos são multiplicadas por o termo $e^{i\theta}$ com θ escolhido aleatoriamente. Para demonstrar isto, é usado o fato que uma função com deslocamento temporal satisfaz a seguinte propriedade no domínio da frequência (ver Bracewell 1999)

$$\mathcal{F}\{x(t - \Delta s)\}(\omega) = e^{i\omega\Delta s} \mathcal{F}\{x(t)\}(\omega). \quad (3.9)$$

onde \mathcal{F} é o operador de transformada de Fourier. Da mesma maneira, a rotação de fase na equação (3.7) também possui uma versão no domínio na frequência (ver Purves 2014)

$$\mathcal{F}\{x_{rot}(t; \varphi)\}(\omega) = \mathcal{F}\{x(t)\}(\omega) \cdot \begin{cases} e^{i\varphi} & \omega > 0 \\ 0 & \omega = 0 \\ e^{-i\varphi} & \omega < 0. \end{cases} \quad (3.10)$$

Ou seja, ambas as estratégias de codificação, como mostram as equações (3.10) e (3.9), são uma multiplicação por um termo de fase no domínio da frequência e, portanto, são versões no domínio do tempo da codificação no domínio da frequência. Comparado com o esquema de codificação por polarização (III), os esquemas por rotação de fase e por estática aleatória permitem mais valores de códigos e, portanto, aumentam a aleatoriedade da codificação. Esse efeito também pode ser observado na Figura 3.1.

Cumprimento da propriedade 1

Para demonstrar que a propriedade 1 descrita na seção 3.2 é aplicada a relação de Parseval (Arfken & Weber 2005) no primeiro termo da equação (3.5) obtendo a seguinte relação

$$(\nabla\Phi)_1 = - \sum_{i=1}^{N_s} \int \omega^2 \hat{e}_i^*(\omega) \hat{e}_i(\omega) \hat{R}_i^*(\mathbf{r}, \omega) \hat{P}_i(\mathbf{r}, \omega) d\omega, \quad (3.11)$$

onde $\hat{e}_i(\omega)$ é o operador de codificação no domínio da frequência e $\hat{e}_i^*(\omega)$ corresponde com a transformada de Fourier do operador invertido no tempo. No caso da codificação por polarização é evidente que $\hat{e}_i^*(\omega) \hat{e}_i(\omega) = 1$. No caso das codificações por rotação de fase e estática aleatória, pode-se observar que também satisfazem que $\hat{e}_i^*(\omega) \hat{e}_i(\omega) = 1$, isto usando as equações (3.9) e (3.10).

3.4 L-BFGS reinicializado

Os esquemas de codificação discutidos na seção anterior satisfazem a propriedade 1 mas não satisfazem diretamente a propriedade 2 que faz que o segundo termo da equação (3.5), ou seja o crosstalk, seja zero. Isto implicaria satisfazer uma propriedade de ortogonalidade dos campos de onda de fontes diferentes. A forma de amortecer o ruído de crosstalk é através da mudança da codificação ao longo das iterações. Por exemplo, Krebs et al. 2009 propuseram mudar a codificação a cada iteração e calcular a direção de busca através do método de gradientes conjugados.

Esta estratégia de é problemática para algoritmos como L-BFGS descrito na seção 2.4.1, nos quais a estimativa do produto inversa da Hessiana-gradiente é feita através de subtrações das soluções e dos gradientes de um conjunto de iterações consecutivas. Para resolver isto, Castellanos et al. 2014 propuseram manter a mesma

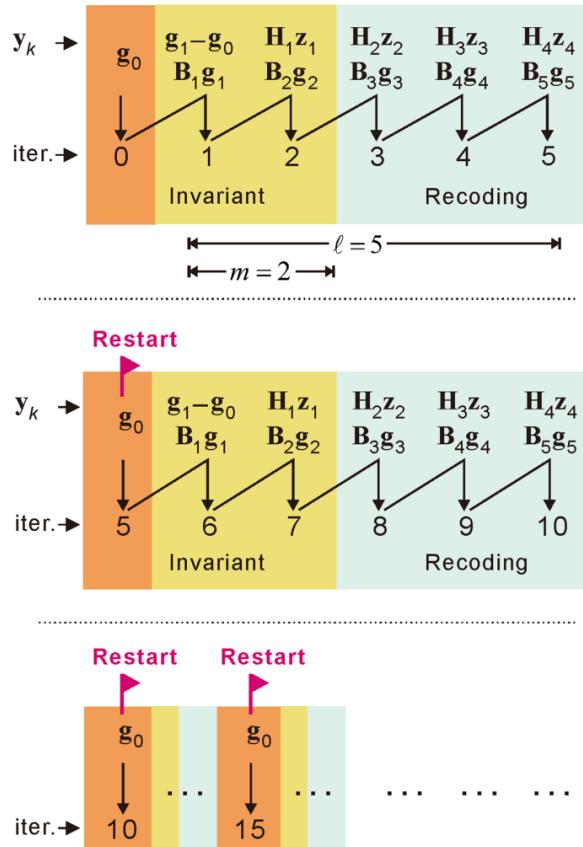


Figura 3.2: Esquema de L-BFGS reinicializado. (Figura extraída do artigo Rao & Wang 2017).

codificação por um número ℓ de iterações de depois reinicializar o L-BFGS mudando a codificação.

Nos trabalhos de Rao & Wang 2017 e Rao et al. 2019 foi introduzido o algoritmo de L-BFGS reinicializado o qual é uma variação do algoritmo em Castellanos et al. 2014. No L-BFGS reinicializado a codificação das fontes é modificada ao longo das iterações seguindo a estratégia apresentada na Figura 3.2. A codificação inicial (Faixa laranja na Figura 3.2) é mantida durante m iterações (Representada pela faixa verde na Figura 3.2). Depois a codificação ao longo das iterações (Faixa celeste na Figura 3.2) é modificada até atingir ℓ iterações e no final desse ciclo de iterações o L-BFGS é reinicializado. Em Rao et al. 2019 é demonstrado que esta estratégia, obtida de forma empírica, possui uma convergência mais rápida que o algoritmo de Castellanos et al. 2014.

3.5 Experimentos numéricos

Para demonstrar a eficiência das estratégias de codificação propostas, foram feitos alguns testes numéricos sobre uma fatia 2D do modelo SEG/EAGE overthrust

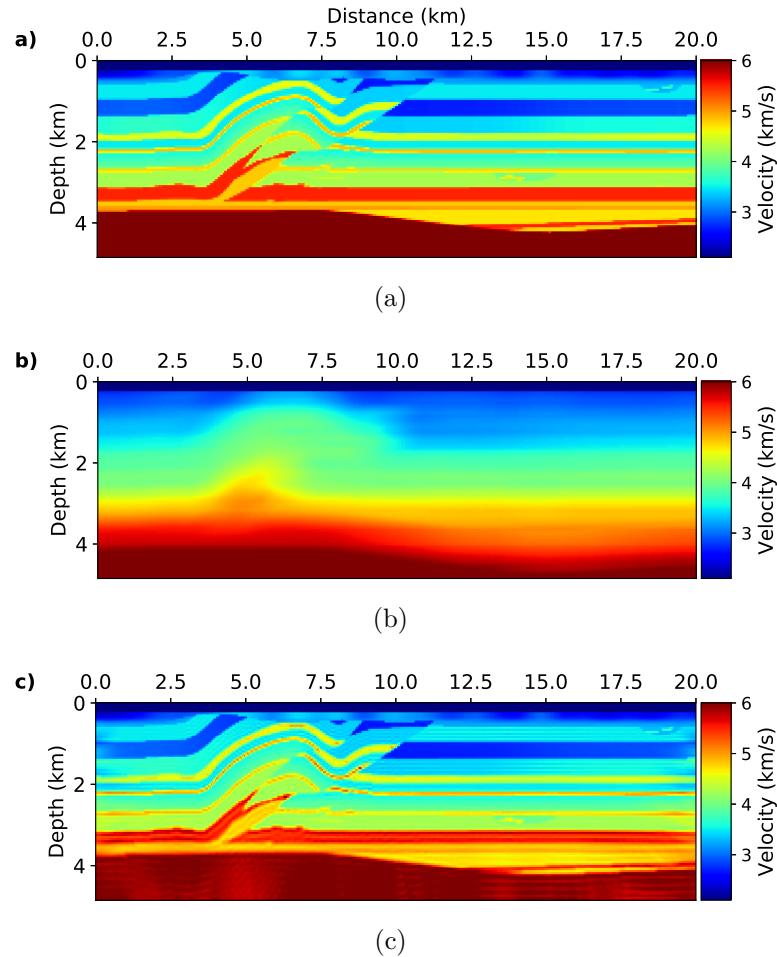


Figura 3.3: (a) Modelo SEG/EAGE overthrust. (b) Modelo inicial usado nos experimentos de inversão. (c) Resultado da inversão convencional.

(Figura 3.3(a)). O modelo tem dimensão de 401×98 celas com tamanho de cela igual a 50m o que equivale a uma dimensão de $20\text{km} \times 4.85\text{km}$. As primeiras 5 celas possuem uma velocidade constante de 2.1km/s. Essa primeira camada é fixada também no modelo inicial na Figura 3.3(b) e é suposto como conhecida ao longo da inversão. Isso é feito para evitar as altas amplitudes que aparecem no gradiente próximo às posições de receptores e fontes.

Para simular os dados observados foi usado um propagador de onda por diferenças finitas como descrito na seção 2.3.1. Para simular um meio infinito é usada uma implementação da borda absorvente CPML baseada nos trabalhos de Pasalic & McGarry 2010 e Komatitsch & Martin 2007 e descrita na seção 2.3.3. O número de celas da borda é de 30 a cada lado das bordas do modelo. Foi simulada uma aquisição de 201 tiros na superfície com 401 receptores por tiro. A assinatura da fonte foi uma wavelet de Ricker (Wang 2015) com frequência pico de 6Hz e cada traço possui 1500 amostras temporais com intervalo de amostragem de 4ms.

O modelo inicial para a inversão, apresentado na Figura 3.3(b), é uma suavização

do modelo real. Para ter uma referência da qualidade da inversão dos esquemas de codificação propostos, foi feita uma inversão convencional, onde o gradiente é calculado tiro por tiro. Esta inversão é apresentada na Figura 3.3(c).

Para resolver o problema de inversão foi usado o algoritmo L-BFGS reinicializado descrito na seção 3.4, com parâmetros $\ell = 5$ e $m = 2$ os quais foram usados também no artigo de Rao & Wang 2017. As Figuras 3.4(a)-3.4(d) apresentam os resultados com os esquemas de polarização (III), rotação de fase (I), estática limitada a um período (II), estática aleatória limitada a 10 períodos (III). A estática limitada a 10 períodos foi feita para simular a condição de não limitar o valor da estática como proposto no trabalho de Schuster et al. 2011. Para que os resultados das inversões sejam comparáveis, foi usado um critério de parada discutido em Castellanos et al. 2014 o qual impõe que a inversão pare na iteração (k) cujo valor de função custo $\Phi^{(k)}$ satisfaça que

$$\frac{\Phi^{(k)}}{\Phi^{(0)}} < \varepsilon^2 \quad (3.12)$$

onde $\Phi^{(0)}$ é o valor da função custo para o modelo inicial e ε é o valor limiar para parar a inversão. Para os experimentos apresentados o valor desse limiar é $\varepsilon = 0.03$.

Como observado na Figura 3.4, todas as estratégias possuem uma qualidade semelhante entre si e ao resultado da FWI convencional na Figura 3.3(c). Isto implica que todas as estratégias de codificação mitigam o ruído de crosstalk. A inversão com a polarização aleatória leva 775 iterações para alcançar o critério de parada. Os esquemas de codificação propostos por rotação de fase e por estática limitada a um período convergem mais rápido. Vale a pena notar que o esquema (IV) o qual não limita o valor da estática leva duas vezes mais iterações que quando o valor da estática é limitada, razão que justifica o fato de limitar o valor da estática. A Figura 3.5(a) mostra o valor da função objetivo em função do número de iterações para os diferentes esquemas de codificação. Todos os esquemas apresentam uma convergência mais lenta quando comparada com a FWI convencional. Isto é devido ao fato que além de diminuir a função custo, o processo de inversão também mitiga o ruído de crosstalk. Porém quando avaliada a função objetivo em função do número de propagações de onda que devem ser feitas para cada estratégia, pode-se observar que o número de propagações é aproximadamente duas ordens de grandeza menor que a FWI convencional o que mostra que as estratégias de codificação reduzem consideravelmente o tempo de execução da FWI.

A Figura 3.5 mostra que os esquemas (III) e (IV) possuem taxas de convergência semelhantes. Porém, as inversões propostas neste trabalho, (I) e (II) possuem taxas de convergência mais rápidas os esquemas mencionados anteriormente. Também pode ser observado na Figura 3.5 que o valor da função objetivo não é sempre decrescente, apresentando alguns picos. Isto é devido ao fato que o processo de busca linear para calcular o comprimento de passo ótimo é afetado pelo ruido de crosstalk. Porém, como explicado em Nocedal & Wright 2006, a equação secante usada no método L-BFGS escala o gradiente corretamente de tal forma que o tamanho de passo pode ser assumido igual a 1 na maioria dos casos.

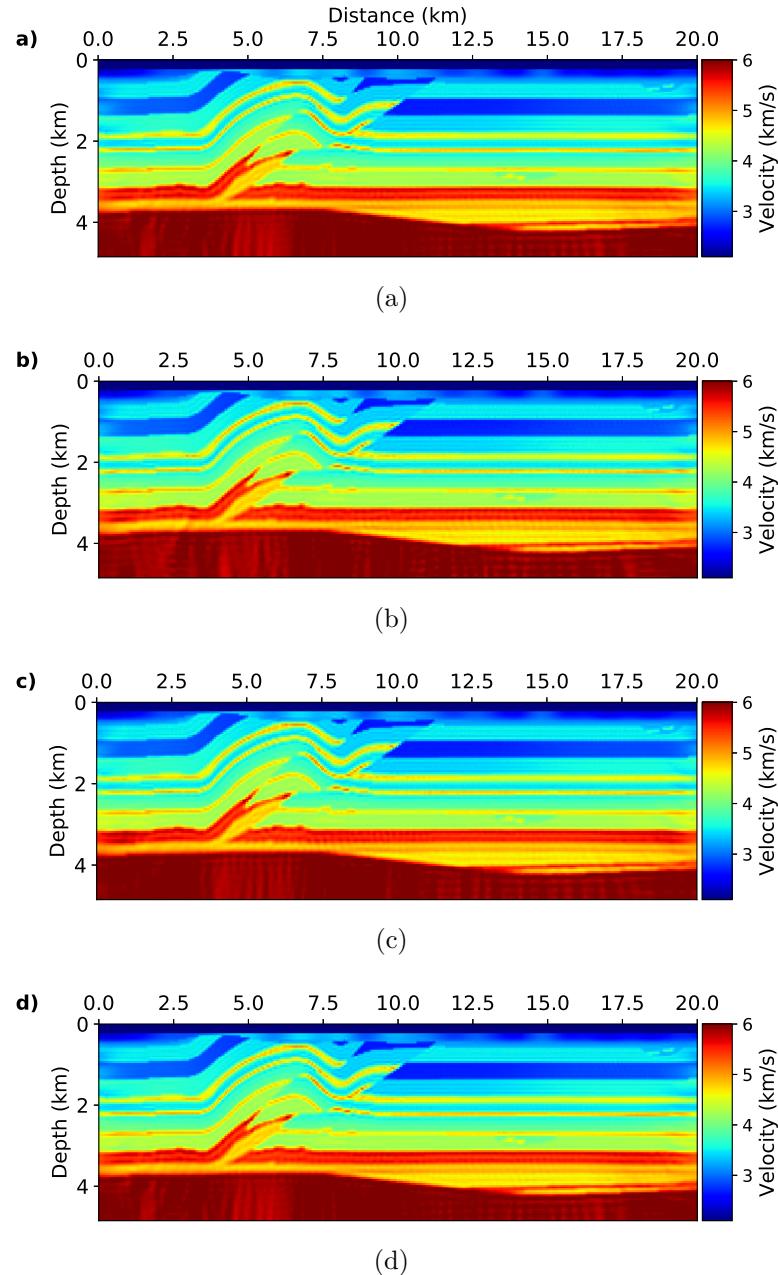


Figura 3.4: FWI com fontes codificadas usando os diferentes esquemas de codificação. Resultado com codificação por (a) polarização aleatória (III) depois de 775 iterações, (b) rotação de fase aleatória (I) depois de 250 iterações, (c) estática aleatória limitada a um período (II) depois de 355 iterações e (d) estática aleatória limitada a 10 períodos (IV) depois de 690 iterações.

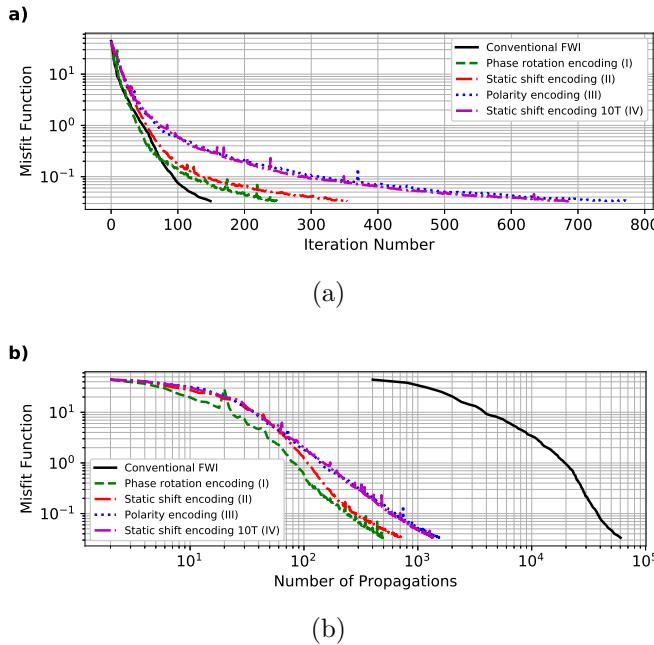


Figura 3.5: Convergência da função objetivo ao longo das iterações com os diferentes esquemas de codificação e a FWI convencional. Função objetivo como função (a) do número de iterações e (b) do número de propagações de onda.

Para avaliar a qualidade da inversão é calculado um coeficiente de correlação calculado a cada ponto do modelo de velocidade entre o resultado da inversão e o modelo verdadeiro como feito em Gao & Wang 2019. O coeficiente de correlação local para o ponto (i, j) é definido como

$$C_{i,j} = \frac{\sum \mathcal{A} \cdot \mathcal{B}}{\max(\sum \mathcal{A} \cdot \mathcal{A}, \sum \mathcal{B} \cdot \mathcal{B})}, \quad (3.13)$$

onde \mathcal{A} e \mathcal{B} são matrizes 5×5 centradas na posição (i, j) e extraídas do modelo real e do resultado da inversão, respectivamente. Esse coeficiente é mostrado na Figura 3.6(a) para a FWI convencional e na Figura 3.6(b) para a inversão usando a codificação por polarização. O coeficiente de correlação localizada indica que (i) não há uma diferença marcante entre a inversão obtida usando a FWI convencional ou usando um esquema de codificação e (ii) não existe uma diferença marcante entre os resultados dos diferentes esquemas de codificação e, por isto, somente é apresentada a correlação localizada para o esquema de codificação por rotação de fase.

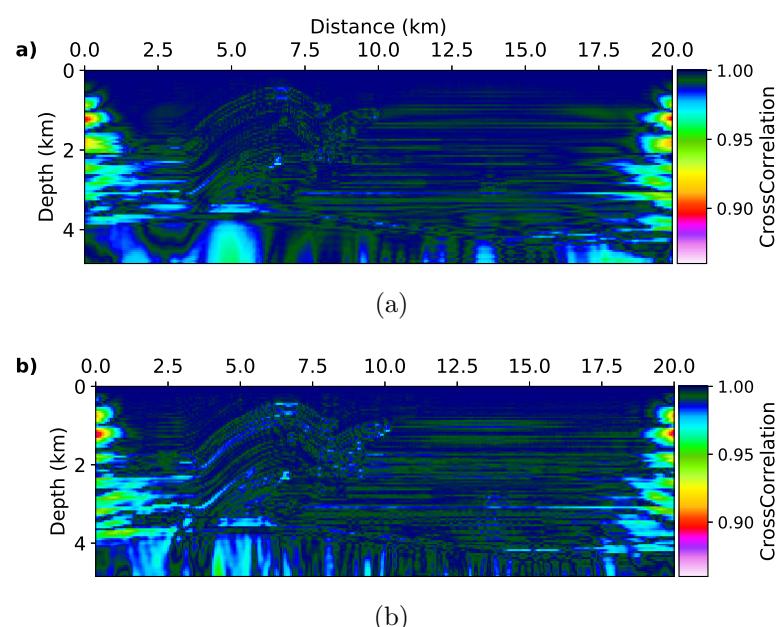


Figura 3.6: Coeficiente de correlação localizada entre o modelo verdadeiro e o modelo invertido para (a) a FWI convencional e (b) a codificação de fontes por rotação de fase.

Capítulo 4

Construção de um código para FWI em 3D

4.1 Introdução

Neste capítulo é descrita a implementação de um código para modelagem e inversão para levantamentos 3D. Este código foi implementado com o intuito de lidar com dados tanto sintéticos quanto reais. No caso dos dados reais relacionados a dados marinhos, é necessário implementar uma equação da onda que lide tanto com efeitos da velocidade e da densidade. Além disso, nos dados marinhos reais pode-se observar também um marcado efeito das múltiplas geradas pela superfície do mar. Para lidar com esses dois fenômenos, na seção 4.2 é apresentada a formulação da equação de Klein-Gordon, a qual é um equivalente da equação acústica com densidade variável. Também é apresentada a condição de superfície livre na modelagem, a qual permite modelar o efeito das múltiplas na superfície do mar.

Como em um caso geral, as posições de fontes e receptores não seguem um padrão regular, ao modelar ou inverter estas posições não podem ser encaixadas em uma malha computacional. Para posicionar corretamente as fontes e receptores dentro da malha computacional, foi implementada a interpolação sinc com uma janela de Kaiser a qual é conhecida como interpolação de Hicks (Hicks 2002). Esta implementação é apresentada na seção 4.3.

Devido ao tamanho dos modelos de velocidade 3D, o cálculo do gradiente possui um grande custo computacional tanto em processador quanto de memória. Na seção 4.5 são apresentadas várias estratégias para reduzir o custo computacional no cálculo do gradiente entre as quais estão a amostragem temporal e o recálculo do campo direto usando a informação da borda computacional do campo de onda. Além disso, devido a que os dados reais possuem ruído, em muitos casos é necessário filtrar o gradiente para facilitar a convergência do FWI e para que os modelos invertidos sejam menos ruidosos. Portanto, também na seção 4.5, é descrito o filtro de Laplace o qual permite obter um gradiente mais suavizado. Por último na seção 4.6 se apresentam algumas características computacionais do código 3D.

4.2 Física usada na modelagem sísmica

4.2.1 Uso de densidade variável - equação de Klein-Gordon

A aproximação acústica é usada com maior frequência para inverter dados marinhos, uma vez que esta aproximação permite explicar os eventos observados nos dados de campo. A equação que permite calcular o campo de pressão na aproximação acústica é

$$\frac{1}{v^2(\mathbf{x})} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} - \rho(\mathbf{x}) \nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \right) = s(\mathbf{x}_s, t). \quad (4.1)$$

Para resolver esta equação numericamente, na literatura é usada a método de malha intercalada descrito em Virieux 1986. Neste trabalho, foi usada uma alternativa diferente a qual consiste em usar a equação de *Klein-Gordon* (KG). A formulação do uso da equação KG é apresentada nos trabalhos de Anno et al. 1992 e Mojica & Maciel 2020. Definindo um novo campo $q(\mathbf{x}, t)$ como

$$q(\mathbf{x}, t) = \frac{p(\mathbf{x}, t)}{\sqrt{\rho(\mathbf{x})}}, \quad (4.2)$$

na equação (4.1), é obtida uma nova da equação de segunda ordem para o campo q

$$\left[\frac{1}{v^2(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla^2 + M(\mathbf{x}) \right] q(\mathbf{x}, t) = \sqrt{\rho(\mathbf{x})} s(\mathbf{x}_s, t) \quad (4.3)$$

com $M(\mathbf{x}) = \sqrt{\rho(\mathbf{x})} \nabla^2 (1/\sqrt{\rho(\mathbf{x})})$. A equação (4.3) possui a mesma forma matemática da equação de Klein-Gordon a qual é usada para aplicar a relatividade restrita na mecânica quântica.

Existem duas vantagens ao usar a equação KG: a primeira é que por ser de segunda ordem, pode-se calcular o campo de pressão diretamente sem necessidade de calcular o campo de velocidades de partícula como acontece na formulação de primeira ordem. A segunda vantagem é que a equação KG não mistura as derivadas espaciais do campo de pressão e da densidade como acontece na equação (4.1), o qual facilita a implementação numérica. A equação (4.3) é fácil de implementar adicionando o termo M ao termo (2.11). Fazendo isto, a equação de evolução temporal para a equação KG pode ser escrita como

$$Q_{i,j,k}^{t+1} = 2Q_{i,j,k}^t - Q_{i,j,k}^{t-1} + v_{i,j}^2 \Delta t_p^2 \left((\nabla^2 Q_{i,j,k})^t + M_{i,j,k} Q_{i,j,k} + s_{i_f,j_f,k_f}^t \right), \quad (4.4)$$

Para checar a implementação da equação KG, foi feita uma comparação com o código FDELMODC, descrito no trabalho de Thorbecke & Draganov 2011, o qual esta disponível na internet. Este código é baseado na formulação de malha intercalada. Comparando traço a traço foi encontrado que os resultados entre as duas formulações são equivalentes.

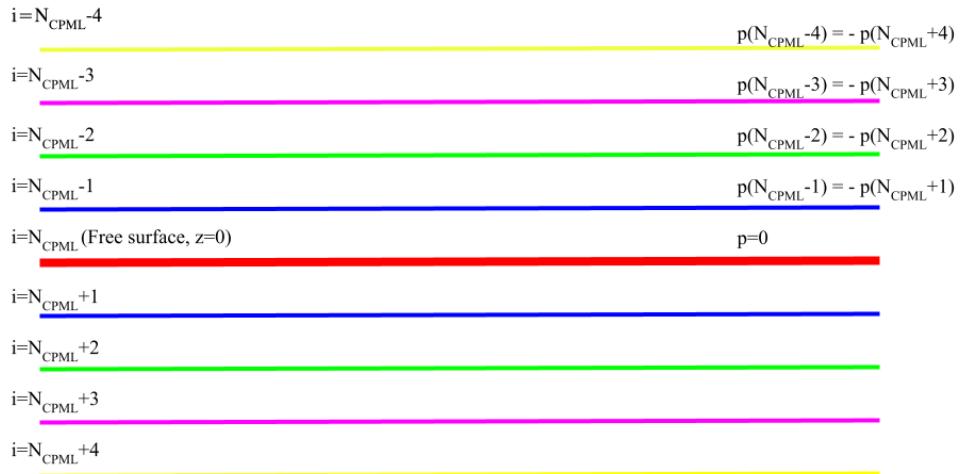


Figura 4.1: Esquema da implementação da superfície livre pelo método das imagens.

4.2.2 Condição de superfície livre

Os dados marinhos são fortemente afetados pela presença de múltiplas causadas pela reflexão na superfície do mar. Para modelar o efeito das múltiplas é usada a denominada condição de *superfície livre*. Existem vários métodos para impôr a condição de superfície livre tais como o método do vácuo e o método das imagens. Um resumo das diferentes técnicas para implementar a superfície livre é encontrado em Fichtner 2011.

Neste trabalho foi implementado o método das imagens (Levander 1988) e é representado na Figura 4.1. Neste método deve-se garantir que tanto a pressão como a sua derivada sejam zero na superfície. Para fazer isto no método de diferenças finitas, cada vez que é feito um passo de tempo, a componente de pressão na linha da matriz correspondente é igualada a zero. Como na implementação da diferenças é usado um modelo estendido para aplicar a CPML (Ver seção 2.3.3 e Figura 2.2), então a superfície livre vai corresponder com a linha $i = N_{cpml}$ que é o número de elementos da borda.

Para garantir que a derivada seja zero na superfície, os valores das pressões nas linhas da matriz acima de $i = N_{cpml}$ são calculados da seguinte forma

$$p(i = N_{cpml} - l, j, k) = -p(i = N_{cpml} + l, j, k) \quad (4.5)$$

onde j, k são os índices correspondentes às coordenadas x e y e l indica a posição da linha a ser calculada. O número de linhas onde deve ser feito esse procedimento é igual à metade da ordem da aproximação de diferenças finitas σ , ou seja, $l = 1, \dots, \sigma/2$. A condição da derivada ser zero na superfície é análoga fisicamente a ter um coeficiente de reflexão igual a -1 que é o caso da interface água-ar.

4.3 Posicionamento correto de fontes e receptores

No método de diferenças finitas, o espaço é discretizado de tal forma que o cálculo do campo de onda é feito para certas posições que são múltiplos da distância entre celas da malha. Se a posição da fonte ou do receptor não coincide com um elemento da malha, uma opção seria aproximar a posição real ao elemento mais próximo dentro da malha. Porém, se a malha for muito grosseira, esta aproximação leva a erros tanto na cinemática quanto na amplitude das ondas modeladas. Para resolver esse problema, no código é usada a metodologia de interpolação sinc ponderada por uma janela de Kaiser, a qual é apresentada em Hicks 2002.

Em uma malha de diferenças finitas, um ponto fonte ou receptor pode-se expressar como

$$s_i = A_s \delta(i + \alpha) \quad (4.6)$$

onde S é a amplitude da fonte, i numera um elemento do grid, $\delta(i + \alpha)$ é uma função delta de Dirac espacial e $-0.5 < \alpha \leq 0.5$ e a distância do ponto fonte ao ponto da malha mais próximo. Para que uma fonte seja corretamente amostrada pela malha de diferenças finitas, deve-se substituir a função delta por uma versão banda limitada desta. Para este fim, é usado o teorema de Shannon descrito na seção 2.2, e portanto, a função fonte fica expressada como

$$s_i = A_s \text{sinc}(i + \alpha). \quad (4.7)$$

Se $\alpha = 0$, ou seja, o ponto fonte coincide com um ponto da malha, a amplitude da fonte vai ser zero em todos os pontos da malha menos no ponto fonte uma vez que $\text{sinc}(0) = 1$ e $\text{sinc}(i) = 1$ quando $i \neq 0$. Caso contrário, para injetar uma fonte numa posição fora da malha seria necessário injetar em todos os pontos da malha com um peso dado pela equação 4.7, o qual aumentaria consideravelmente o custo computacional de injetar a fonte. Para contornar isto, em Hicks 2002, os coeficientes s_i são truncados por uma janela tipo Kaiser de tal forma que a fonte seja injetada somente em alguns pontos vizinhos cuja distância seja menor a um raio r . Com esta modificação a fonte agora é injetada como

$$s_i = A_s W(i + \alpha) \text{sinc}(i + \alpha). \quad (4.8)$$

onde W está definida como

$$W(x) = \begin{cases} \frac{I_0(b\sqrt{1-(x/r)^2})}{I_0(b)} & |x| \leq r \\ 0 & |x| > r \end{cases} \quad (4.9)$$

onde I_0 é a função de Bessel modificada de ordem zero e b é um parâmetro de controle da largura da janela. No código implementado neste trabalho foram usados os valores sugeridos em Hicks 2002 que são $b = 4.14$ e $r = 4$, o que implica que se a fonte não coincide com a malha computacional, então são usados oito pontos para injetar a fonte (quatro para cada lado da posição da fonte) seguindo a equação 4.8. Para coletar a informação do receptor que não esteja posicionado em algum ponto da

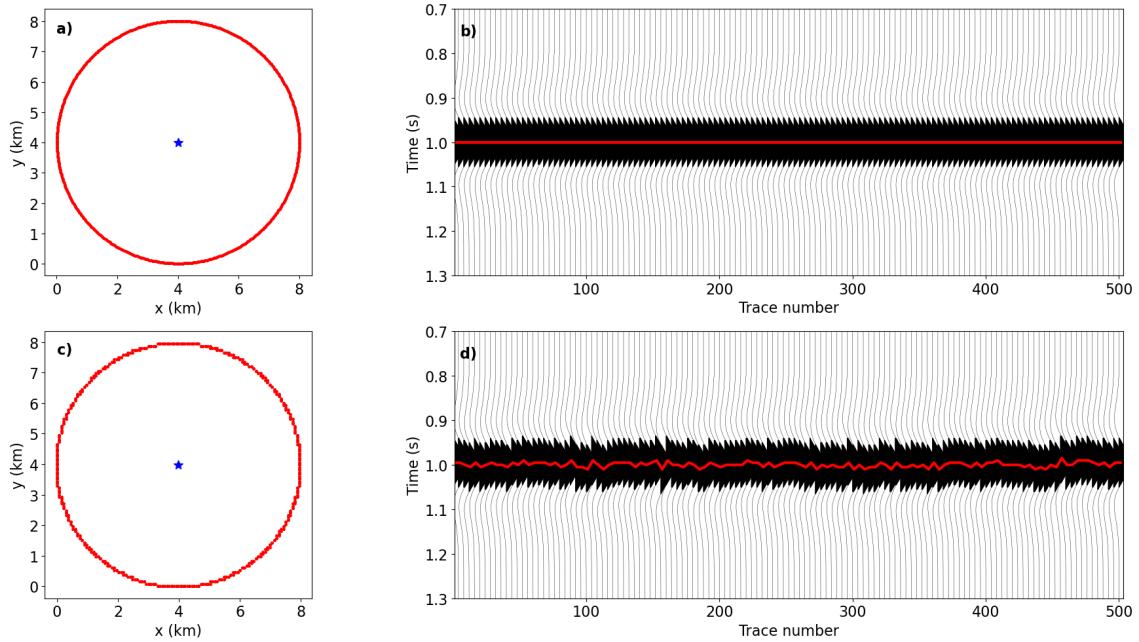


Figura 4.2: Efeito da interpolação de Hicks: a) Mapa com coordenadas reais. b) Traços calculados com interpolação de Hicks. c) Mapa com coordenadas aproximadas aos pontos da malha. d) Traços calculados sem interpolação.

malha computacional, deve-se seguir o mesmo princípio da fonte, mas os termos da equação 4.8 agora devem ser somados, ou seja a amplitude A_r do receptor, coletada a partir do campo de onda calculado $u(n)$, é calculada como:

$$A_r = \sum_i u(i) W(i + \alpha) \text{sinc}(i + \alpha) \quad (4.10)$$

No caso de duas e três dimensões, a construção do delta de Dirac deve ser feita para cada dimensão. Uma vez que o delta de Dirac em três dimensões é o produto do delta em cada dimensão, então a fonte em 3D deve ser injetada de tal que

$$s_{ijk} = A_s W(i + \alpha_i) \text{sinc}(i + \alpha_i) W(j + \alpha_j) \text{sinc}(j + \alpha_j) W(k + \alpha_k) \text{sinc}(k + \alpha_k). \quad (4.11)$$

e o receptor é gravado como

$$A_r = \sum_{ijk} u(i, j, k) W(i + \alpha_i) \text{sinc}(i + \alpha_i) W(j + \alpha_j) \text{sinc}(j + \alpha_j) W(k + \alpha_k) \text{sinc}(k + \alpha_k). \quad (4.12)$$

onde $\alpha_i, \alpha_j, \alpha_k$ são a separação ao elemento da malha mais próximo nas dimensões i, j e k respectivamente. Isto implica que no caso 3D, a fonte deve ser injetada em $8^3 = 512$ pontos, enquanto que para coletar a informação de um único receptor devem ser usados 512 pontos. Este fato aumenta o custo computacional da propagação.

Para demonstrar a importância de aplicar a interpolação de Hicks, foi feito um teste onde foram colocados receptores ao longo de um círculo com raio de 4 km e uma fonte no meio do círculo como mostrado na Figura 4.2(a). Foi usado também um modelo de velocidade constante de 4 km/s. O resultado da modelagem usando a interpolação de Hicks é apresentado na Figura 4.2(b). Pode-se observar que como esperado os tempos de viagem coincidem com o valor teórico ($t = 4 \text{ km}/4 \text{ km/s} = 1 \text{ s}$) e também observa-se que todos os traços são iguais o que também é esperado uma vez que o afastamento fonte-receptor é o mesmo para todos os traços. Isto não acontece quando as coordenadas são aproximadas aos pontos da malha (Figura 4.2(c)). No sismograma modelado com estas coordenadas (Figura 4.2(d)) os tempos de viagem já não coincidem com o valor teórico, e os traços seguem um padrão “serrilhado”, o que significa que aproximar as coordenadas aos pontos da malha pode alterar a cinemática observada e isto pode ser um problema na hora de ajustar os dados observados com os dados reais.

4.4 Estimativa da assinatura da fonte

Estimar a assinatura da fonte, ou *wavelet*, é um dos passos mais importantes para a aplicação, uma vez que esta é a que determina a amplitude e a fase do sismograma modelado. Uma wavelet bem estimada faz com que os sismogramas modelados ajustem melhor com os sismogramas de campo e, portanto, faz com que desajuste seja melhor.

Para estimar a wavelet, neste trabalho é usada a metodologia apresentada em Pratt 1999, a qual pode ser implementada tanto no domínio do tempo quanto no domínio da frequência, que a apresentada a seguir. Para obter uma componente de frequência da wavelet $s(\mathbf{x}_s, f)$ de uma fonte localizada em \mathbf{x}_s , é resolvido através da minimização do erro quadrático entre o dado real e a o dado calculado através da função de Green. A formulação no domínio da frequência deste erro esta dada por

$$\phi(s(\mathbf{x}_s, f)) = \frac{1}{2} \sum_i |d(\mathbf{x}_i, f) - s(\mathbf{x}_s, f)G(\mathbf{x}_i, \mathbf{x}_s, f)|^2, \quad (4.13)$$

onde o índice i indica uma soma sobre receptores com posições \mathbf{x}_i e $d(\mathbf{x}_i, f)$ é a componente da frequência f no receptor i . $G(\mathbf{x}_i, \mathbf{x}_s, f)$ é a função de Green entre a fonte e o receptor para a frequência f . Por ser ter uma dependência linear do argumento a ser minimizado, o valor $s(\mathbf{x}_s, f)$ pode ser calculado diretamente, derivando a equação (4.13) com relação a s , igualando a zero e deixando em evidência o termo s , obtendo que

$$s(\mathbf{x}_s, f) = \frac{\sum_i G^*(\mathbf{x}_i, \mathbf{x}_s, f)d(\mathbf{x}_i, f)}{\sum_i G^*(\mathbf{x}_i, \mathbf{x}_s, f)G(\mathbf{x}_i, \mathbf{x}_s, f)}. \quad (4.14)$$

Para aplicar a equação (4.14) em dado real é recomendável usar os traços mais próximos á fonte para evitar que possíveis efeitos de atenuação sejam dominantes. Para isto devem ser usados os primeiros N traços ordenados por afastamento fonte-

receptor. O parâmetro N deve ser escolhido pelo usuário. Para obter a $d(\mathbf{x}_i, f)$ é aplicada a transformada de Fourier sobre os dados escolhidos.

O cálculo da função de Green é feito no domínio do tempo propagando uma delta de Dirac com um propagador de onda, o qual deve incluir a física que mais se ajuste com o dado observado. Por exemplo, se o efeito da densidade e das múltiplas de superfície livre é importante nos dados observados, estes devem ser incluídos no propagador. Uma vez que a função de Green é calculada no domínio do tempo, é aplicada a transformada de Fourier para obter $G(\mathbf{x}_i, \mathbf{x}_s, f)$. Depois é aplicada a equação (4.14) frequência por frequência, e por último é aplicada a transformada de Fourier inversa para obter a wavelet no domínio do tempo. O procedimento para a estimativa da wavelet está resumido no algoritmo 4.

Algoritmo 4: Estimativa da assinatura da fonte $s(\mathbf{x}_s, t)$

Entrada: $d(\mathbf{x}_i, t)$: dado observado no domínio do tempo (com filtro passa-banda aplicado); N : número de traços usados no cálculo; f_{max} : frequência máxima do dado (determinada pelo filtro passa-banda).

- 1 Ordenar os traços $d(\mathbf{x}_i, t)$ por afastamento fonte-receptor;
 - 2 Escolher os primeiros N traços;
 - 3 Aplicar transformada de Fourier para obter $d(\mathbf{x}_i, f)$
 - 4 Construir delta de Dirac: Vetor de zeros com o número de amostras temporais colocando 1 na primeira entrada;
 - 5 Cálculo da função de Green $G(\mathbf{x}_i, \mathbf{x}_s, t)$: Calcular um sismograma usando como fonte a delta de Dirac;
 - 6 Aplicar transformada de Fourier para obter $G(\mathbf{x}_i, \mathbf{x}_s, f)$;
 - 7 **para** $f = 0, \dots, f_{max}$ **faz**
 - 8 | Calcular $s(\mathbf{x}_s, f)$ com a equação (4.14) usando os N traços escolhidos.
 - 9 **fim**
 - 10 Aplicar transformada de Fourier inversa sobre $s(\mathbf{x}_s, f)$;
 - 11 **retorna** $s(\mathbf{x}_s, t)$;
-

A delta de Dirac construída seguindo o algoritmo 4 possui energia para todas as frequências inferiores à frequência de Nyquist, e portanto, a função de Green calculada também possui o mesmo conteúdo de frequência. Porém, pela condição de estabilidade espacial descrito na seção 2.3.2, as frequências que não satisfazem o critério da equação (2.17) possuem dispersão numérica e, portanto, a função de Green para estas frequências não é confiável. Este problema é contornado limitando o conteúdo de frequência do dado observado até uma frequência que satisfaça a equação (2.17). Isto é feito aplicando um filtro passa-banda. Portanto, a equação (4.14) não seria calculada até a frequência de Nyquist, senão até a frequência máxima f_{max} (ver o algoritmo 4) do dado, a qual é selecionada pelo usuário através do filtro passa-banda.

4.5 Cálculo de gradiente em 3D

O cálculo do gradiente da função objetivo é uma das etapas do FWI 3D mais caras computacionalmente. A razão disto pode ser identificada na equação (2.45). O cálculo do gradiente requer guardar na memória do computador o campo de onda direto em espaço e tempo enquanto o campo reverso no tempo é calculado. Para modelos de velocidade em 2D o campo direto manter o campo de onda na memória é viável. Porém, no caso 3D o campo de onda completo pode ocupar mais algumas dezenas ou centenas de gigabytes o qual é maior que a memoria disponível de qualquer computador. Portanto, na literatura foram desenvolvidas algumas estratégias para contornar este problema as quais foram combinadas neste trabalho para fazer o cálculo do gradiente e as quais são descritas a seguir.

4.5.1 Reamostragem temporal

Para aplicar a FWI é necessário favorecer o conteúdo de baixas frequências nos dados observados, como por exemplo para aplicar a estratégia de inversão multi-escala (Bunks et al. 1995). Isto é feito aplicando um filtro passa-banda ao dado observado para limitar o conteúdo de frequência. Após a filtragem, o dado possui uma frequência máxima muito inferior à frequência de Nyquist associada à amostragem. Por exemplo, se os dados observados estão amostrados com $\Delta t = 0.002\text{ s}$ (taxa de amostragem usada comumente na indústria), portanto de acordo com a equação 2.1, $f_{ny} = 250\text{ Hz}$. Se o dado for filtrado até 10 Hz, então poderia-se se usar uma amostragem cuja f_{ny} fosse apenas um pouco maior que 10 Hz. Neste caso, se o dado for reamostrado com $\Delta t = 0.02\text{ s}$, então $f_{ny} = 25\text{ Hz}$ a qual seria ainda maior que frequência máxima e, portanto, continuaria respeitando o teorema de Shannon (Ver seção 2.2). No cálculo do gradiente nessa nova reamostragem, o campo de onda ocupa 10 vezes menos memória, e o cálculo da equação (2.45) é mais rápido, uma vez que envolve menos instantes de tempo sem perda de acurácia.

Como a FWI envolve propagações de onda pelo método de diferenças finitas, o intervalo de temporal de propagação Δt_p deve respeitar o critério CFL descrito na seção (2.3.2). Como dito em Yang et al. 2016, o Δt_p geralmente é muito menor que a amostragem Δt necessária para o cálculo do gradiente. Para conciliar isto, o código foi planejado para receber como parâmetro o Δt de amostragem e automaticamente calcula um Δt_p de propagação de tal forma que $\Delta t_p = \Delta t/n$ sendo n um inteiro e que satisfaça a condição CFL (Equação 2.12). O dado observado $d(\mathbf{x}_i, t)$ e a wavelet $s(\mathbf{x}_s, t)$ devem estar amostrados com a amostragem Δt_p . Para injetar a wavelet $s(\mathbf{x}_s, t)$, esta deve ser reamostrada para ter um intervalo de amostragem Δt_p . Portanto, a wavelet que é fornecida pelo usuário é interpolada dentro do código usando a interpolação sinc descrita no algoritmo 1.

4.5.2 Recálculo do campo direto

Existem duas técnicas principais para evitar manter o campo direto na memória enquanto é calculado o campo reverso. A primeira técnica é conhecida como *check-*

pointing ótimo (Symes 2007) onde somente são mantidos na memória os campos correspondentes a certas amostras de tempo e quando é feito o cálculo reverso no tempo, as amostras que não foram guardadas na memória são recalculadas. Na segunda técnica, somente são mantidos em memória a informação correspondente com a fronteira do domínio físico do campo (Clapp 2009, Yang et al. 2014) e a partir desta informação o campo direto é recalculado ao mesmo tempo que o campo reverso no tempo.

A técnica de check-pointing tem como vantagem que pode ser usada para equações de onda com efeito de atenuação, uma vez que o recálculo é feito com o tempo aumentando, o qual não acontece com a segunda técnica uma vez que o recálculo do campo de onda é feito em tempo reverso e por conservação da energia a amplitude não pode ser recuperada. Porém, se a equação de onda não inclui o efeito de atenuação, a segunda técnica consegue recuperar fielmente o campo direto e, além disso, tem mais vantagens que o check-pointing, devido a que é mais fácil de implementar e também porque o recálculo dos campos de onda possui um custo computacional menor.

Neste trabalho foi usada a técnica do recálculo do campo direto baseado nos trabalhos de Yang et al. 2014 e Yang et al. 2016. Para explicar a técnica é usada a Figura 4.3. Primeiro é calculado o campo direto e é guardada a informação da fronteira, a qual é representada pela área amarela em cada um dos instantes de tempo da Figura 4.3. A fronteira deve ter uma largura igual à metade da ordem de diferenças finitas usada para aproximar o laplaciano e dever ser a fronteira do domínio físico e não do modelo estendido pela borda CPML (área cinza). É importante lembrar que é necessário guardar a fronteira a cada instante da propagação Δt_p (instantes azuis na Figura 4.3) e não somente nos instantes da amostragem Δt (instantes verdes).

Uma vez terminada a propagação direta, o campo direto é reconstruído a partir da fronteira injetando a informação desta como se fosse um termo fonte. Isto é feito simultaneamente ao cálculo convencional do campo reverso. A reinjeção da informação da fronteira é feita em todos os instantes da propagação Δt_p , mas o cálculo do gradiente somente é feito nos instantes de tempo da amostragem Δt (instantes verdes). O procedimento do cálculo do gradiente usando esta técnica é resumido no algoritmo 5.

4.5.3 Filtragem do gradiente - Filtro de Laplace

Aplicar filtros sobre o gradiente tem dois grandes objetivos. O primeiro objetivo é remover ou amortecer as amplitudes anômalas que aparecem no cálculo do gradiente ao redor das posições de fontes e receptores. O segundo objetivo é regularizar o gradiente, isto é, remover ruído do gradiente causado pela presença de ruído no dado real que faria com a solução fosse conduzida para uma solução ruidosa. Embora existem vários filtros na literatura, como o filtro gaussiano usado que é usado em Kamath et al. 2021, neste trabalho foi implementado o filtro de Laplace seguindo o artigo de Trinh et al. 2017.

O filtro de Laplace é calculado aplicando duas vezes o filtro de Bessel (Trinh

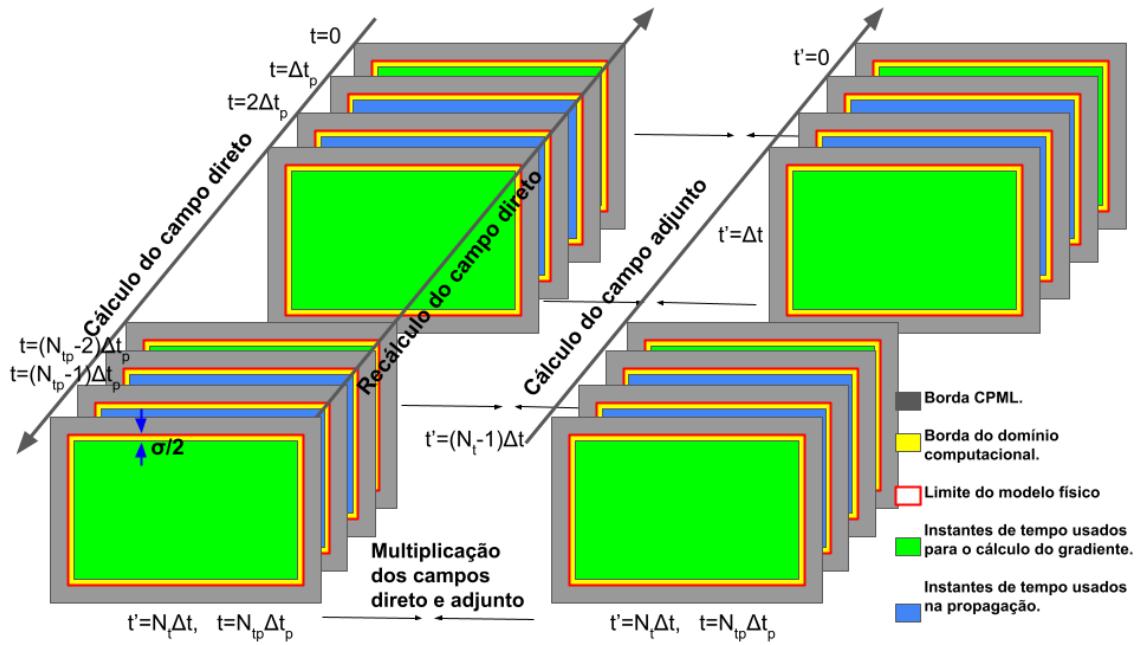


Figura 4.3: Cálculo do gradiente usando a borda do domínio espacial para recalcular o campo direto. Esta figura foi feita baseada na Figura 1 de Yang et al. 2016.

et al. 2017). O filtro de Bessel por sua vez é aplicado resolvendo a seguinte equação diferencial:

$$\sigma_{out}(x, y, z) - \left(L_x^2 \frac{\partial^2}{\partial x^2} + L_y^2 \frac{\partial^2}{\partial y^2} + L_z^2 \frac{\partial^2}{\partial z^2} \right) \sigma_{out}(x, y, z) = \sigma_{in}(x, y, z), \quad (4.15)$$

onde $\sigma_{in}(x, y, z)$ é o gradiente de entrada que deseja-se filtrar e $\sigma_{in}(x, y, z)$ é o gradiente de saída já filtrado. Os valores L_x , L_y e L_z são os operadores de coerência para cada dimensão e estes podem ser constantes ou variáveis espacialmente e estes valores controlam a intensidade do filtro. Entre maior seja o operador de coerência mais suavizada vai ser a saída. Como equação 4.15 envolve um operador linear, então esta é resolvida através do método de gradientes conjugados lineares onde o chute inicial é o próprio gradiente de entrada $\sigma_{in}(x, y, z)$.

No algoritmo 5 está o procedimento do cálculo do gradiente onde são misturados o reamostragem temporal, o recálculo do campo direto e a filtragem do gradiente.

4.6 Detalhes de implementação do código 3D

A implementação dos códigos para inversão de dados 3D descritos neste capítulo, foi feita na linguagem Julia desenvolvida no MIT (Massachusetts Institute of Technology) (Bezanson et al. 2017), a qual segundo os desenvolvedores consegue fornecer velocidades de processamento semelhantes com a linguagem C, mas com a facilidade de programação semelhante ao MATLAB. Devido ao tamanho dos modelos 3D e à quantidade de dados que são usados para inverter, no código 3D desenvolvido neste

Algoritmo 5: Cálculo do gradiente

Entrada: $d(\mathbf{x}_i, t)$: dado observado no domínio do tempo; $s(\mathbf{x}_s, t)$: Assinatura da fonte; A amostragem Δt de $d(\mathbf{x}_i, t)$ e $s(\mathbf{x}_s, t)$ deve ser a mesma e não necessariamente satisfazer a equação (2.12); Coordenadas de fonte e receptor.

- 1 Achar um fator $n \in \mathbb{Z}$ tal que $\Delta t_p = \Delta t/n$ e Δt_p satisfaça a equação (2.12) ;
- 2 Interpolar $s(\mathbf{x}_s, t)$ para ter a amostragem Δt_p usando a interpolação sinc (Algoritmo 1);
- 3 **para** $t = 0, \dots, N_{t_p} \Delta t_p$ **faca**
 - 4 Avançar um passo de tempo na propagação;
 - 5 Salvar $\mathbf{u}(\mathbf{x}, t; \mathbf{x}_s, 0)$ nas bordas do modelo físico (seção 4.5.2);
 - 6 **se** $t = 0, \Delta t, \dots, N_t \Delta t$ **então**
 - 7 Salvar dado simulado na posição dos receptores.
 - 8 **fim**
- 9 **fim**
- 10 Calcular a fonte adjunta e a função desajuste;
- 11 Interpolar a fonte adjunta usando a interpolação sinc (Algoritmo 1);
- 12 Gradiente $\frac{\partial \Phi}{\partial m_l} \leftarrow 0$;
- 13 **para** $t = N_{t_p} \Delta t_p, \dots, 0$ **faca**
 - 14 Avançar um passo de tempo do campo reverso no tempo $\mathbf{R}(\mathbf{x}, t)$;
 - 15 Recalcular o campo direto $\mathbf{u}(\mathbf{x}, t; \mathbf{x}_s, 0)$ através da borda computacional usando o procedimento 4.5.2);
 - 16 **se** $t = 0, \Delta t, \dots, N_t \Delta t$ **então**
 - 17 Calcular $\frac{\partial^2 \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s, 0)}{\partial t^2}$;
 - 18 $\frac{\partial \Phi}{\partial m_l} \leftarrow \frac{\partial \Phi}{\partial m_l} - \frac{\partial^2 \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s, 0)}{\partial t^2} \times \mathbf{R}(\mathbf{x}, t)$
 - 19 **fim**
- 20 **fim**
- 21 Aplicar filtro de Laplace no gradiente;
- 22 **retorna** Função desajuste, gradiente;

trabalho foram implementadas estratégias de paralelização para que o código possa distribuir o trabalho em um cluster com vários computadores.

Para este fim, foram combinadas diferentes funções de paralelização próprias da linguagem Julia. Para calcular o laplaciano foi usado um paralelismo estilo OPENMP o qual distribui os cálculos entre vários threads de um mesmo computador. Também foi usado um paralelismo estilo MPI usando a biblioteca Distributed incluída no Julia. Com este tipo de paralelismo, conseguiu-se calcular simultaneamente o gradiente de várias fontes usando vários nós de um cluster.

Capítulo 5

FWI aplicado em um dado do pré-sal brasileiro

5.1 Introdução

Nos últimos anos, o uso da FWI em dados reais tem aumentado consideravelmente, isto por causa de dois motivos principais. O primeiro, é o aumento do poder computacional e o segundo é o planejamento e aquisição de levantamentos sísmicos mais adequados para a FWI como é o caso das aquisições tipo *Ocean Bottom Node* ou OBN, onde os sensores são colocados no fundo marinho. Como a FWI é um método baseado na equação da onda, a FWI esta sendo usada principalmente em áreas de geologia complexa, onde outros métodos de análise de velocidade de migração, por exemplo os métodos baseados em raios, podem falhar.

Na literatura podem achar-se exemplos de aplicação da FWI em zonas de geologia complexa como por exemplo o mar do Norte (Sirgue et al. 2010, Warner et al. 2013, Operto et al. 2015) e o golfo do México (Shen et al. 2018). Nestas regiões o uso de aquisições de OBN é fundamental, uma vez que estas fornecem tanto uma boa contribuição de afastamentos fonte-receptor longos, como também frequências mais baixas que as aquisições de cabo puxado ou *streamer*, o qual é muito importante para o sucesso da aplicação da FWI (Virieux & Operto 2009).

O pré-sal brasileiro tem se tornado uma das áreas de maior interesse de exploração devido a que possui uma das maiores reservas de petróleo do mundo. O pré-sal brasileiro oferece dificuldades, tanto na exploração quanto na exploração de recursos, por causa da espessura da camada d'água, que na média é de aproximadamente 2 km, como também da localização do reservatório, o qual encontra-se abaixo de uma camada de sal que de espessura variável.

Por causa disto, no pré-sal brasileiro tem aumentado o uso de aquisições OBN em comparação ao uso de aquisições streamer. O OBN tem sido usado tanto para imageamento (Jouno et al. 2019, Fruehn et al. 2019) como para monitoramento de reservatórios (Cypriano et al. 2019). Em Martinez et al. 2020 é proposta uma metodologia para melhorar a aquisição e o processamento de dados OBN na área do pré-sal, através do aumento da distância de receptores do fundo marinho. Esta estratégia facilita a execução do levantamento sem comprometer a qualidade do processamento final. Em Lopez et al. 2020 é apresentada uma modificação experimental

desta estratégia onde, além de usar uma distribuição esparsa de receptores, também é usada uma distribuição esparsa de fontes, o qual é diferente dos OBN convencionais onde o distribuição de fontes é altamente densa.

Neste capítulo é apresentado um fluxograma para aplicar FWI na aquisição OBN apresentada em Lopez et al. 2020. Na seção 5.2 é descrita detalhes da aquisição dos dados usados. Na seção 5.3 é mostrado como foi feita a construção dos modelos iniciais de velocidade e densidade. Na seção 5.4, é resumido o fluxograma da FWI aplicada nos dados. Na seção 5.6 são descritas as funções objetivo testadas para inverter o dado. Na seção 5.7, é apresentada a estratégia para escolher as frequências máxima e mínima para filtrar os dados observados. Na seção 5.8 são apresentados os resultados da inversão usando as diferentes funções objetivo. Por último, na seção 5.9 são apresentados os testes feitos para obter inversões com fontes simultâneas codificadas. As estratégias de codificação apresentaram um resultado semelhante ao resultado sem codificação.

5.2 Descrição dos dados usados

O levantamento usado nesta tese é um OBN cuja aquisição em 2019 no campo gato do mato localizado na parte central da bacia de Santos, a qual é uma das áreas com maior produção do pré-sal. A distribuição de fontes e receptores é apresentada no mapa da Figura 5.1(a). Nesta aquisição foram usados 1029 nós receptores de quatro componentes espalhados no fundo do mar, o qual possui uma profundidade media de 2 km como pode ser observado no mapa 3D na figura 5.1(b). Os receptores foram distribuídos em linhas receptoras as quais seguem a padrão intercalado. A distância entre linhas e nós é de 400 m. Também foram atiradas 3006 fontes com uma profundidade de 8 m, seguindo três círculos de raios 6 km, 9 km e 10 km. A distância entre duas fontes consecutivas foi de 50 m.

No planejamento original todas as fontes deveriam ter sido escutadas por todos os receptores, mas devido a um problema logístico, as fontes correspondentes ao círculo de 10 km não foram por 85 receptores pertencentes às 4 linhas receptoras no extremo sul. Portanto, para mais facilidade e para fazer os testes com fontes codificadas, foram usados os 944 receptores que tem a informação completa dos três círculos fontes. Os dados foram recebidos em formato SEG-Y agrupados por receptor. Foi recebida somente a componente de pressão sem nenhum filtro aplicado a não ser os próprios do equipamento de aquisição (filtros anti-alias). O comprimento de traço é de 10 s com uma taxa de amostragem de 2 ms.

A Figura 5.2 mostra os traços correspondentes ao receptor localizado no centro dos três círculos. Esta figura é apresentada sem nenhum filtro. Os traços dentro deste receptor seguem a ordem cronológica de aquisição e é possível identificar os traços correspondentes a cada círculo fonte. Foi observado que no receptor existe um ruído de baixa frequência o qual é associado às oscilações da coluna d'água.

Como mencionado em Lopez et al. 2020, este levantamento foi desenhado para ser um levantamento por refração, no qual os afastamentos longos são beneficiados. Os afastamentos longos possuem informação dominante das ondas refratadas (*diving*

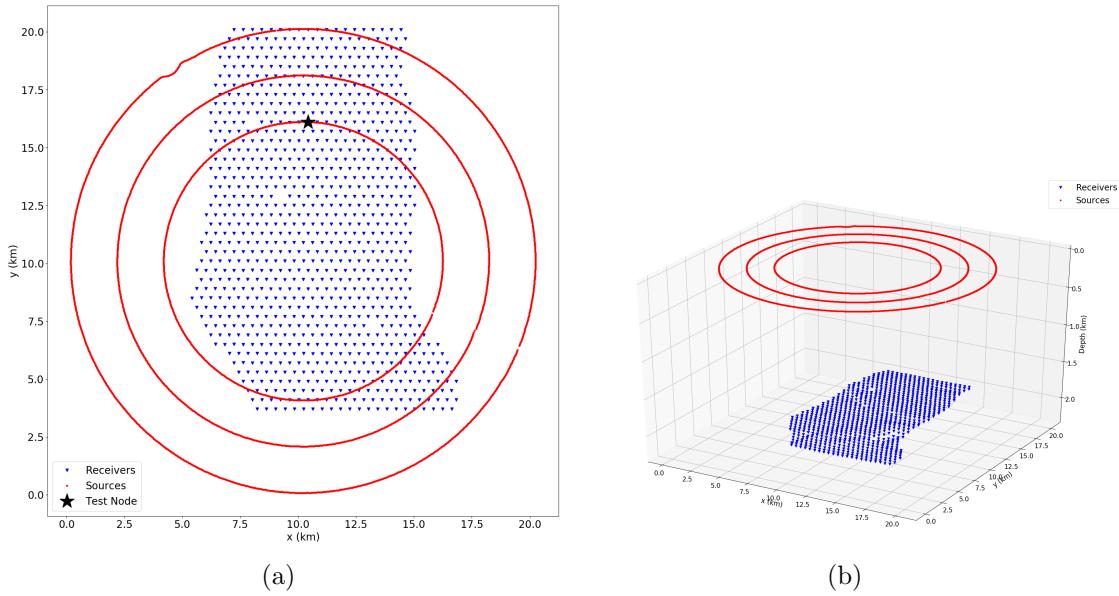


Figura 5.1: (a) Planta do levantamento OBN usado nas inversões. As coordenadas usadas na figura não são geográficas, senão que são relativas ao modelo de velocidade. O nó realçado mostra a posição usada para modelar os sismogramas apresentados na seção de controle de qualidade. (b) Mapa 3D da aquisição.

waves) tal como observado na Figura 5.2, as quais conseguem iluminar as regiões no reservatório do pré-sal. Mais detalhes sobre o levantamento e a iluminação desta aquisição podem ser encontrado em Lopez et al. 2020. Além das onda refratadas, o receptor da Figura 5.2 apresenta informação dominante das múltiplas causada pela superfície do mar. O efeito das múltiplas é levado em consideração nas inversões feitas nas seguintes seções.

Vale a pena mencionar que, diferentemente de outros levantamentos OBN onde a distribuição de fontes é altamente densa, neste OBN a distribuição de fontes é esparsa, o qual dificulta a aplicação de outros métodos de análise de velocidade baseados em tempos de viagem, tais como a tomografia. Isto porque a tomografia atualiza o modelo somente pelos pontos por onde passa o raio principal entre fonte e receptor (Williamson 1991), enquanto que o FWI atualiza em todos os pontos que pertencem à primeira zona de Fresnel a qual é bastante ampla, uma vez que são usadas frequências muito baixas (Huang & Schuster 2014). A sobreposição das amplas zonas de Fresnel compensam a falta de cobertura da aquisição. Esta ideia é mais discutida na seção 5.5.

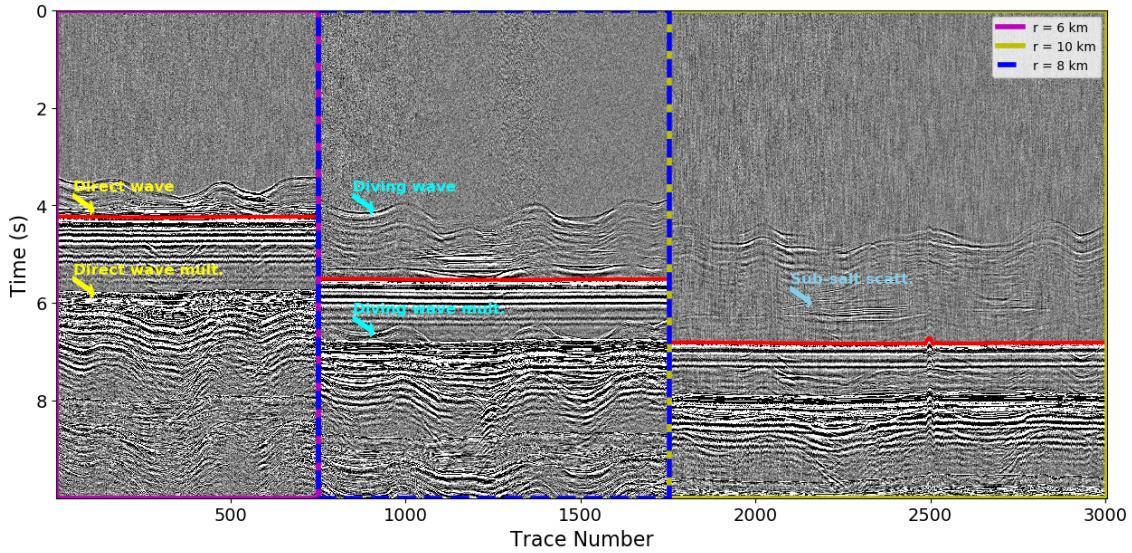


Figura 5.2: Dado de campo de um receptor localizado no centro dos três círculos fontes. Cada retângulo corresponde a um círculo fonte com raios 6 km (esquerda), 8 km (meio), 10 km (direita). São destacados alguns sinais importantes como a onda direta (setas amarelas), as ondas refratadas (setas cianas) e as ondas espalhadas (setas azuis).

5.3 Preparação de modelos iniciais

5.3.1 Modelo inicial de velocidade

Foi usado como modelo de referência para comparar com os modelos obtidos com a FWI, um modelo de velocidade obtido na mesma área usando um outro OBN o qual era mais denso em fontes e receptores que o OBN usando nesta tese. Na Figura 5.3 é apresentado o modelo de referência. Este modelo é um cubo com distância entre celas de 75 m e cujas números de celas em cada dimensão são $n_x = 274$, $n_y = 270$ e $n_z = 120$, cobrindo uma área de $20.55 \text{ km} \times 20.25 \text{ km}$ com 9 km de profundidade.

A Figura 5.3 apresenta três cortes do modelo, um corte inline o qual vai de leste a oeste, um corte crossline que vai de sul ao norte e um corte em profundidade. No corte inline são identificadas cinco camadas principais: a camada d'água, a qual possui uma largura aproximada de 2 km, uma camada de sedimentos pos-sal, a camada de sal, a camada da base do sal e por último a camada do reservatório.

Uma vez que a aquisição OBN foi desenhada para iluminar o reservatório na área do pré-sal, neste trabalho foi testado se o fluxograma proposto consegue recuperar informação relacionada ao reservatório. Portanto, foi colocado um valor de velocidade constante nas camadas da base do pré-sal e do reservatório. Além disso, o modelo foi suavizado para evitar componentes de alto número de onda no modelo. A Figura 5.4(a) mostra o modelo final resultante desta operação.

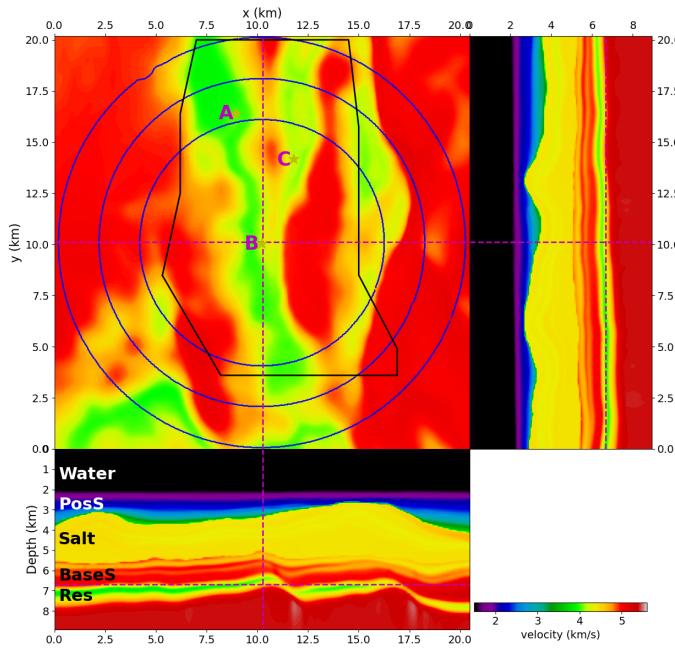


Figura 5.3: Modelo de velocidade de referência. A figura superior esquerda representa um corte em profundidade onde são desenhados as linhas fontes e um polígono indicando a área de receptores. A figura superior direita representa um corte crossline que vai de sul ao norte. A figura na parte inferior representa um corte inline onde as principais camadas geológicas são identificadas (Water: Camada d'água, PosS: Camada de sedimentos pós-sal, Salt: Corpo de sal, BaseS: Base do corpo do sal e Res: reservatório). Cada corte tem duas linhas roxas indicando o lugar por onde passam os outros dois cortes. No corte em profundidade estão as posições dos perfis das Figuras 5.13 e 5.20.

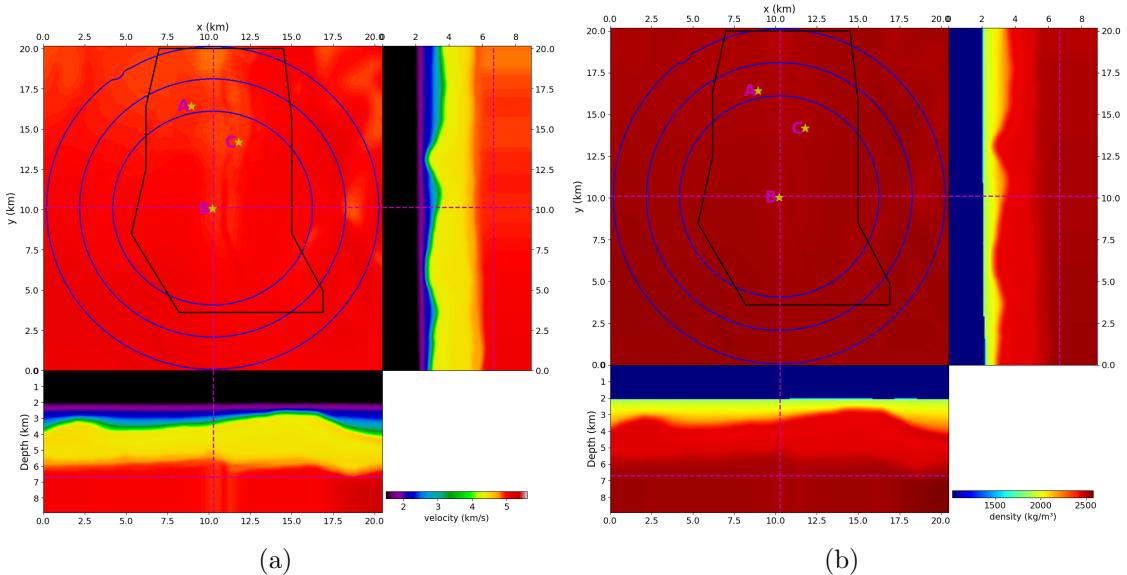


Figura 5.4: Modelos iniciais de (a) velocidade e (b) densidade.

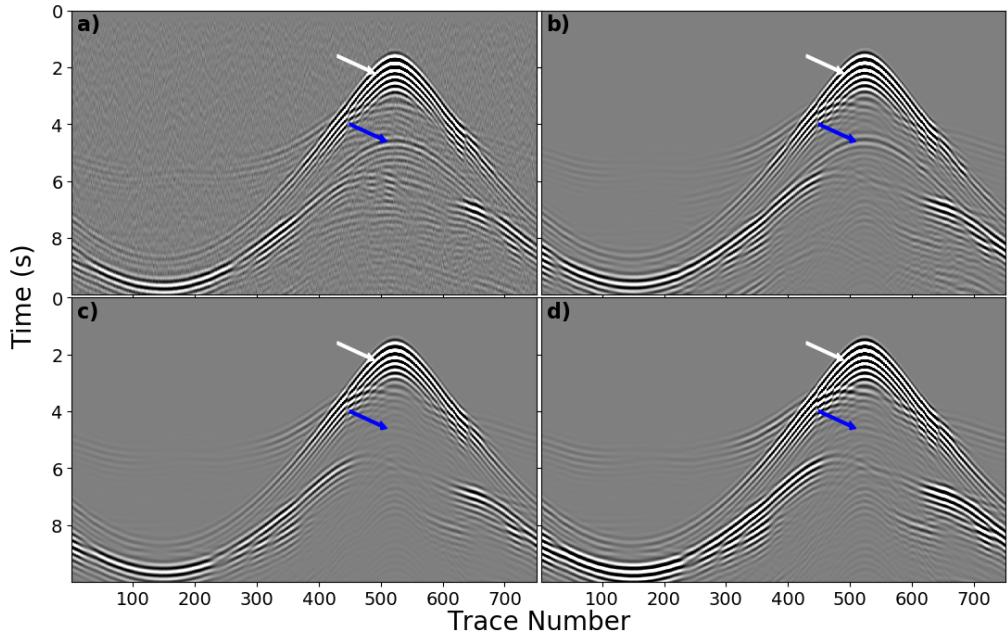


Figura 5.5: Comparação de dado observado e filtrado com dados modelados usando o modelo de velocidade da Figura 5.4(a) e um modelo de densidade b) da Figura 5.4(b), c) constante e d) usando a equação de Gardner em todo o modelo.

5.3.2 Modelo de densidade

Seguindo a metodologia apresentada em Operto et al. 2015, a densidade é calculada a partir do modelo inicial usando uma modificação da equação de Gardner, a qual é proposta em Castagna et al. 1993. Em esta equação de Gardner modificada, a densidade ρ é calculada a partir da velocidade de onda P v_p como

$$\rho = a v_p^2 + b v_p + c \quad (5.1)$$

onde os coeficientes a, b e c dependem do tipo de rocha. Na equação 5.1 é medida em km/s. No caso das inversões feitas nesta tese, foram escolhidos os valores $a = -11.5$, $b = 261.0$ e $c = 1515$ que segundo Castagna et al. 1993 corresponde com uma rocha tipo arenito. Como a densidade d'água não satisfaz a lei de Gardner, então foi colocado uma densidade constante de 1027 kg/m^3 em toda a camada d'água. A densidade resultante é apresentada na Figura 5.4(b).

Para avaliar o efeito da densidade foram feitas três modelagens, usando o modelo de velocidade da Figura 5.4(a), mas usando a densidade da Figura 5.4(b) (Figura 5.5(b)), densidade constante (Figura 5.5(c)) e densidade calculada com a equação (5.1) sem a modificação na camada d'água. Estas modelagens são comparadas com o dado observado e filtrado o qual em apresentado na Figura 5.5(a).

Nos sismogramas da Figura 5.5 tem duas setas para identificar dois eventos de referência. A seta branca aponta para a onda direta primária e a seta azul aponta para a primeira有多重的“múltipla” da onda direta. Pode-se observar que embora todos os

modelos de densidade explicam bem a onda direta, a múltipla somente aparece quando é criado o contraste de densidade água-rocha. Ajustar a onda direita e sua múltipla é importante para o FWI uma vez que estes possuem a maior energia.

5.4 Fluxograma para aplicação do FWI

O fluxograma aplicado no dado OBN consiste dos seguintes passos:

1. Filtragem: Os dados observados foram filtrado usando um filtro passa-banda tipo Orsmy fase-mínima. O filtro deve ser máxima para assegurar a causalidade da wavelet estimada. O filtro Orsmy usado tem frequências de 1-2-4-5.5 Hz, as quais correspondem com as esquinas do trapézio usado para filtrar o sinal. A seleção de estas frequências é explicada na seção 5.7.
2. Reamostragem temporal: Como explicado na seção 4.5.1, a reamostragem temporal, ou seja a diminuição do número de amostras temporais através do aumento do intervalo de amostragem, acelera o cálculo do gradiente e também reduz o consumo de memória e de disco duro, sem perder informação do dado e sem comprometer a qualidade da FWI. Neste caso, o dado originalmente estava amostrado a 2ms e após a filtragem foi reamostrado para 10 ms.
3. Reciprocidade: Devido a que o número de fontes é muito maior, é comum aplicar o princípio de reciprocidade na FWI de dados OBN, tal como é feito nos trabalhos de Sirgue et al. 2010, Warner et al. 2013, Operto et al. 2015 e Kamath et al. 2021. Na reciprocidade, as posições de fontes e receptores são trocadas de tal forma que assume-se que as fontes estão na posição nos receptores e vice-versa. Isto é uma vantagem para os dados usados neste trabalho uma vez que eles estão agrupados por receptores e, portanto, cada receptor pode ser processado como se fosse uma fonte.
4. Estimativa da wavelet fonte: A estimativa da wavelet fonte foi feita seguindo o algoritmo 4 descrito na seção 4.4. Foram testadas duas opções: a primeira, estimar uma wavelet efetiva para toda a aquisição e a segundo foi estimar uma wavelet para cada receptor (lembrando que o princípio de reciprocidade é usado e cada receptor atua como uma fonte). Foi encontrado que estimar uma wavelet para cada fonte fornece um melhor resultado na inversão. O dado observado para estimar a wavelet foi o dado filtrado com o filtro descrito no item 1. Cabe ressaltar que nenhum outro filtro foi aplicado como o filtro de bolha ou o deghosting uma vez que na estimativa da wavelet todos esse efeitos estão incluídos.
5. Cálculo do gradiente e filtragem: O cálculo de gradiente foi feito usando a estratégia de dizimação de fontes usada em Warner et al. 2013 e Kamath et al. 2021 e seguindo o fluxograma apresentado no algoritmo 5. Além da dizimação de fontes foi também aplicada uma dizimação por receptor. Estas dizimações são explicadas em detalhe na seção 5.5. Após o cálculo do gradiente, este foi filtrado com um filtro de Laplace o qual é descrito na seção 4.5.3. Após vários testes os operadores de coerência usados foram $Lx = Ly = Lz = 100\text{ m}$.

6. Otimização: O método de otimização usado foi o algoritmo l-BFGS o qual é descrito no algoritmos 2 e 3. Foram feitas 35 iterações para cada função objetivo a ser testada, as quais são descritas na seção 5.6.

5.5 Dizimação de fontes e receptores no cálculo do gradiente

O custo computacional de gradiente dependem diretamente do número de fontes, ou no caso deste trabalho, do número de receptores, que por causa do princípio de reciprocidade são usados como fontes. Na literatura existem estratégias para dizimar fontes e receptores, as quais visam diminuir o custo computacional do cálculo do gradiente.

A estratégia de dizimação usada nas inversões deste trabalho foi a mesma usada nos trabalhos de Warner et al. 2013 e Kamath et al. 2021 a qual esta baseada no trabalho de van Leeuwen & Herrmann 2012, na qual o cálculo de gradiente é feito com um subgrupo de receptores (que atuam como fontes pela reciprocidade) escolhido aleatoriamente, tal como é mostrado na Figura 5.6(a). Para as inversões feitas neste trabalho foram usados subgrupos de 80 receptores. Este número foi escolhido porque foi usado no trabalho de Warner et al. 2013 e deu um bom resultado com um tempo de 1 hora por iteração em um cluster com 10 nós, calculando simultaneamente o gradiente de quatro receptores por cada nó.

Em van Leeuwen & Herrmann 2012 e em Warner et al. 2013 o subgrupo de receptores é trocado a cada iteração. No caso deste trabalho, como é usado o l-BFGS como minimizador, trocar o subgrupo de receptores a cada iteração não conveniente, uma vez que o efeito da hessiana sobre o gradiente defende do histórico dos gradientes de algumas iterações anteriores. Portanto, o subgrupo de fontes no cálculo do gradiente é mantido por um número de iterações N_{grad} e depois o l-BFGS é reinicializado e o subgrupo de receptores é trocado. Foram testados os valores $N_{grad} = 5$ e $N_{grad} = 10$, obtendo que a primeira opção foi a melhor.

Além da dizimação dos receptores, foi aplicada uma dizimação das fontes. Neste caso, foi testada uma dizimação regular, pegando uma de cada nove fontes sobre os círculos. O padrão de fontes resultantes é apresentado na Figura 5.6(a). Embora a dizimação foi escolhida empiricamente, esta pode ser justificada com a ajuda de um gradiente de um par fonte-receptor o qual é apresentado na Figura 5.6(b).

Como pode ser observado nesta figura, a primeira zona de Fresnel possui uma largura de 2 km. A zona de Fresnel é a região ao redor do raio principal (representado pela linha tracejada na Figura 5.6(b)) cujos pontos contribuem com interferência construtiva para uma onda gravada no receptor. Portanto, usar fontes com distância muito menor à primeira zona de Fresnel não vai contribuir com muita mais informação o que implica que aumentar o intervalo entre fontes não representa um problema.

Esta largura da zona de Fresnel depende da frequência dominante do dado. No caso da FWI como são usadas frequências baixas, a largura da zona de Fresnel é

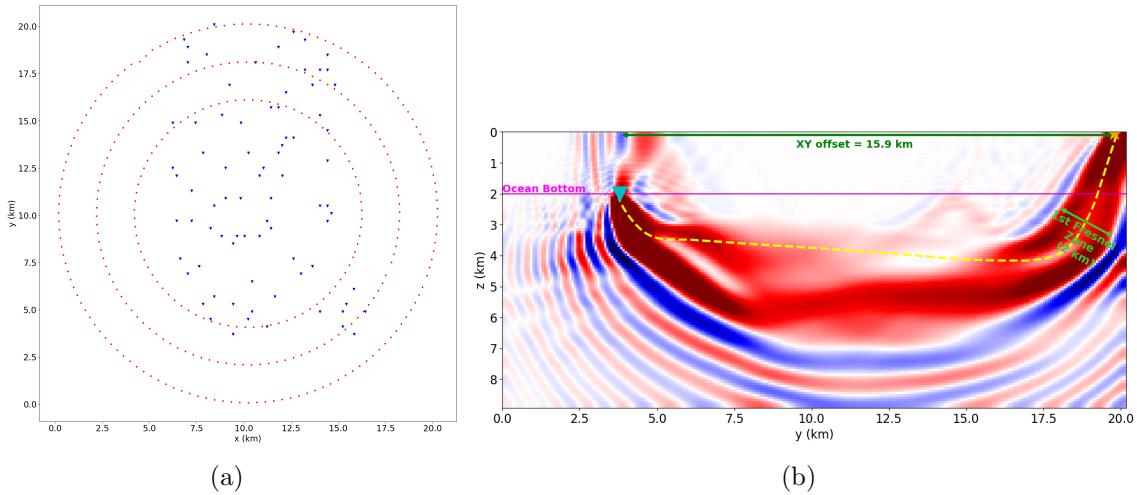


Figura 5.6: (a) Mapa mostrando as fontes e receptores que contribuem ao cálculo de um gradiente. (b) Gradiente de um par fonte-receptor para um afastamento XY de 15.6 km. Pode ser observado que a zona de Fresnel possui uma largura de aproximadamente 2 km.

grande. No caso da migração como as frequências usadas são mais altas, a zona de Fresnel seria mais fina e esta dizimação ão poderia ser feita.

Para checar se esta dizimação de fontes não afetava a qualidade da inversão, foram feitas duas inversões com e sem dizimação obtendo praticamente o mesmo resultado. Porém, com esta a dizimação de fontes, o tempo de cálculo do gradiente foi reduzido à metade. Como esta estratégia funcionou, não foram testadas outras estratégias de dizimação, por exemplo, dizimar aleatoriamente como feito nos receptores.

5.6 Funções objetivos testadas

No contexto da FWI existem vários tipos de funções objetivo para comparar os dados observados \mathbf{d} com os dados simulados \mathbf{u} usando um modelo teste. Neste trabalho foram testadas três funções objetivo descritas a seguir:

1. Norma ℓ_2 : A norma ℓ_2 é calculada como

$$\Phi_{\ell_2} = \frac{1}{2} \sum_s^{N_s} \sum_r^{N_r} \sum_{i=1}^{N_t} (u(r, s, i) - d(r, s, i))^2 \quad (5.2)$$

onde s , r e t são índices para somar sobre as fontes, receptores e amostras temporais respetivamente. N_s , N_r e N_t correspondem com os números de fontes, receptores e amostras temporais respetivamente. A norma ℓ_2 é a mais usada na literatura nos exemplos de inversão de dados reais.

2. Norma ℓ_1 : A norma ℓ_1 é calculada como:

$$\Phi_{\ell_1} = \sum_s^{N_s} \sum_r^{N_r} \sum_{i=1}^{N_t} |u(r, s, i) - d(r, s, i)|. \quad (5.3)$$

Como descrito em Brossier et al. 2010, a norma ℓ_1 é mais tolerante à presença de ruído quando comparada com a norma ℓ_2 uma vez que o ajuste entre os dados observado e simulados é feito através da mediana e não da média como é o caso da norma ℓ_2 .

3. Função objetivo de resíduos não paramétricos (Daqui em diante mencionada como NPR-FWI): Nas funções objetivo convencionais usadas na FWI, (por exemplo as normas ℓ_2 e ℓ_1), impõem uma distribuição de probabilidade específica sobre os resíduos, entendidos como as diferenças entre dados observados e simulados. Por exemplo a norma ℓ_2 assume que os resíduos seguem uma distribuição de probabilidade gaussiana. (Tarantola 2005). Em contrapartida, a NPR-FWI, proposta em Carvalho et al. 2021, não busca-se impôr uma distribuição de probabilidade específica sobre os resíduos. Para este fim, a NPR-FWI usa a técnica de estimativa de densidade por kernel (KDE pela sigla em inglês) (Silverman 1986), a qual permite estimar a distribuição de probabilidade exclusivamente dos próprios resíduos impondo a menor quantidade de suposições. A fórmula da NPR-FWI é

$$\Phi_{NPR} = N_s N_r N_t \log(N_t \sqrt{2\pi}) - \sum_s^{N_s} \sum_r^{N_r} \sum_{i=1}^{N_t} \log \left[\sum_{j=1}^{N_t} \frac{1}{h(r, s, j)} \exp\left(-\frac{1}{2} v^2(i, j)\right) \right] \quad (5.4)$$

onde

$$v(i, j) = \frac{(u(r, s, i) - d(r, s, i)) - (u(r, s, j) - d(r, s, j))}{h(r, s, j)} \quad (5.5)$$

e $h(s, r, j)$ é conhecido como bandwidth e é um parâmetro relacionado com a técnica KDE. Do mesmo jeito que a norma ℓ_1 , a NPR-FWI é menos sensível ao ruído uma vez que ajusta automaticamente a densidade de probabilidade que seguem os resíduos. Além disso, a NPR-FWI é menos sensível ao salto de ciclo que as normas ℓ_2 e ℓ_1 , uma vez que compara os resíduos entre diferentes amostras dos dados observados e calculados e não somente entre as mesmas amostras. Mais detalhes sobre a definição e cálculo do gradiente da NPR-FWI podem ser encontrado em Carvalho et al. 2021.

5.7 Escolha de frequências de inversão

Um dos passos mais importantes para a aplicação da FWI consiste em estimar o intervalo de frequência a ser invertido. Idealmente, para a FWI convergir a um modelo confiável, os dados reais devem ser ricos em frequências baixas (Virieux &

Operto 2009). Porém, nos dados reais estas frequências baixas possuem baixa razão sinal-ruído devido a que a fonte não consegue gerar estas frequências baixas e também pela presença de ruídos fortes nestas frequências baixas como ruído de mares. Portanto, devem ser filtradas as frequências muito baixas onde o ruído é mais dominante. Para escolher a frequência de filtragem são usados os diagramas de fase como feito em Warner et al. 2013.

Obter o diagrama de fase de um OBN convencional é muito simples uma vez que a densidade de fontes é muito grande e portanto, pode ser usada a informação de um único receptor para construir o diagrama (Warner et al. 2013, Jouno et al. 2019). Porém, no caso desta aquisição, isto não é possível, porque o diagrama de fase não seria facilmente interpretável uma vez que os traços dentro do diagrama ficariam muito esparsos.

Portanto, foi desenhada uma metodologia para gerar o diagrama de fase combinando a informação de vários receptores da seguinte forma: É escolhido um grupo de receptores que seja vizinho entre si. No caso deste trabalho foram escolhidos os receptores pertencentes às seis linhas receptoras no centro do levantamento. Foram calculadas coordenadas xy dos pontos médios entre tiro-receptor para todos os traços. Foi formada uma malha computacional com células de tamanho $100m \times 100m$ a qual cobre toda a área de pontos médios. Os traços cujos pontos médios pertencem a determinada célula são empilhados verticalmente sem aplicar nenhuma correção cinemática como normal move-out. Para cada traço empilhado é aplicada a transformada de Fourier, obtendo o valor de fase para a frequência desejada. Esta componente de fase é posicionada dentro da malha de pontos médios, formando o diagrama de fase final.

A Figura 5.7 mostra o diagrama de fase para quatro frequências: 1 Hz, 1.5 Hz, 2 Hz e 2.5 Hz. O diagrama de fase mostra que não existe nenhuma energia coerente para 1 Hz (Ver Figura 5.7(a)). Para 1.5 Hz começa aparecer alguma energia coerente formando um padrão de anéis (Ver Figura 5.7(b)). Este padrão de coerência é ainda mais forte para 2 Hz (Figure 5.7(c)) e 2.5 Hz (Figure 5.7d). A coerência fraca 1.5Hz é devida ao ruido presente de baixa frequência presente no dado real, o qual é mais intenso no círculo mais externo (Ver painel direito na Figura 5.2). Com a informação fornecida pelos diagramas de fase foi decidido filtrar o dado aplicando um corte de frequência em 1 Hz e subindo gradualmente até o filtro ter amplitude máxima em 2 Hz. Isto foi feito para aproveitar a informação em 1.5 Hz.

Para escolher o corte superior de frequência, foi usado a condição de estabilidade espacial descrita na seção 2.3.2. Nos testes realizado, foi usado um laplaciano de oitava ordem o qual permite propagar comprimentos de onda maiores que 3.5 vezes distância entre celas do modelo de velocidade. Com este critério, foi encontrado que a frequência máxima que pode ser propagada sem dispersão com a discretização do modelo usada (Ver seção 5.3) é 5.5 Hz. Para evitar a aparição do fenômeno de Gibbs, foi escolhido uma frequência máxima de 4 Hz onde o filtro possui amplitude máxima e partir deste o valor da amplitude vai diminuindo até zerar em 5.5 Hz. Portanto, o filtro aplicado no dado tem as seguintes frequências: 1-2-4-5.5 Hz.

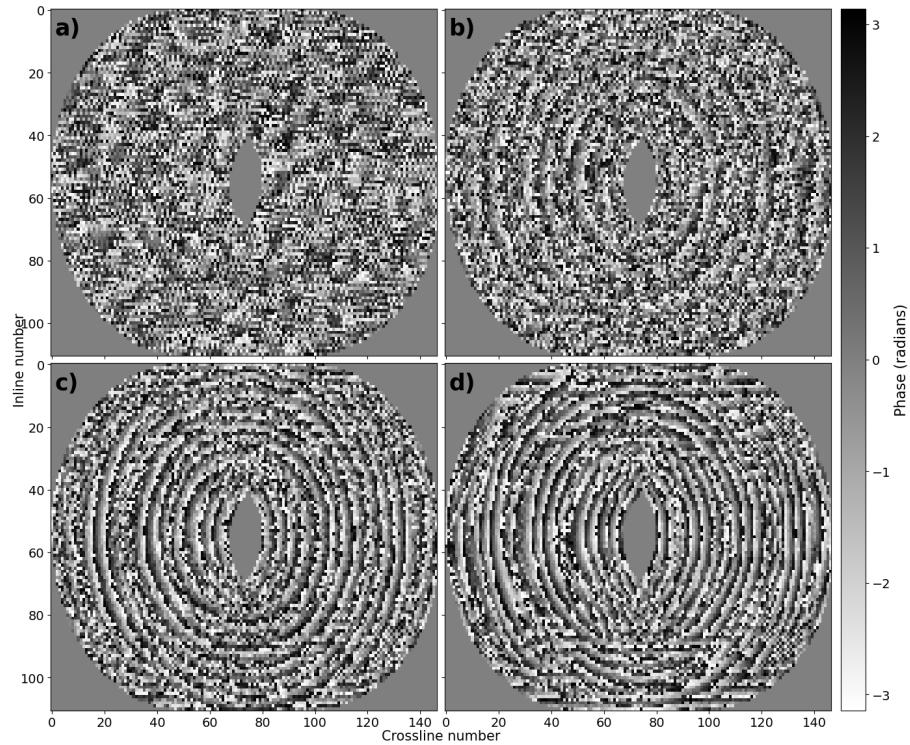


Figura 5.7: Diagrama de fase para a) 1 Hz b) 1.5 Hz c) 2 Hz d) 2.5 Hz.

5.8 Resultados das inversões

5.8.1 Discussão

Forma feitas inversões para cada uma das três funções objetivo descritas na seção 5.6. Todas as inversões foram feitas seguindo o fluxograma descrito na seção 5.4. Na Figura 5.8 é mostrado o resultado da inversão para a norma ℓ_2 . Para efeito de comparação são usados os cortes inline, crossline e em profundidade das Figuras 5.3 e 5.4(a). A norma ℓ_2 não atualiza consideravelmente o modelo de velocidade na camada d'água e nos sedimentos pós-sal o qual indica que as velocidades nestas camadas já estavam bem definidas no modelo inicial. Na camada de sal, a velocidade foi ligeiramente reduzida em comparação ao modelo inicial. A maior atualização no modelo de velocidade foi na área de pré-sal onde as camada de base do sal e principalmente o reservatório foram recuperados. Isto é observado nas seções inline e crossline.

Pode-se observar que a recuperação da camada do reservatório no resultado na norma ℓ_2 não é tão extensa quanto a observada no modelo de referência (Figure 5.3). Isto é por causa da limitação da iluminação do levantamento OBN usado. Como observado na Figura 5.8 a recuperação do reservatório acontece somente dentro e nas regiões próximas ao polígono que delimita os receptores. A geometria da camada de reservatório obtida na inversão é semelhante à observada no modelo de referência, mas o valor de velocidade é mais alto.

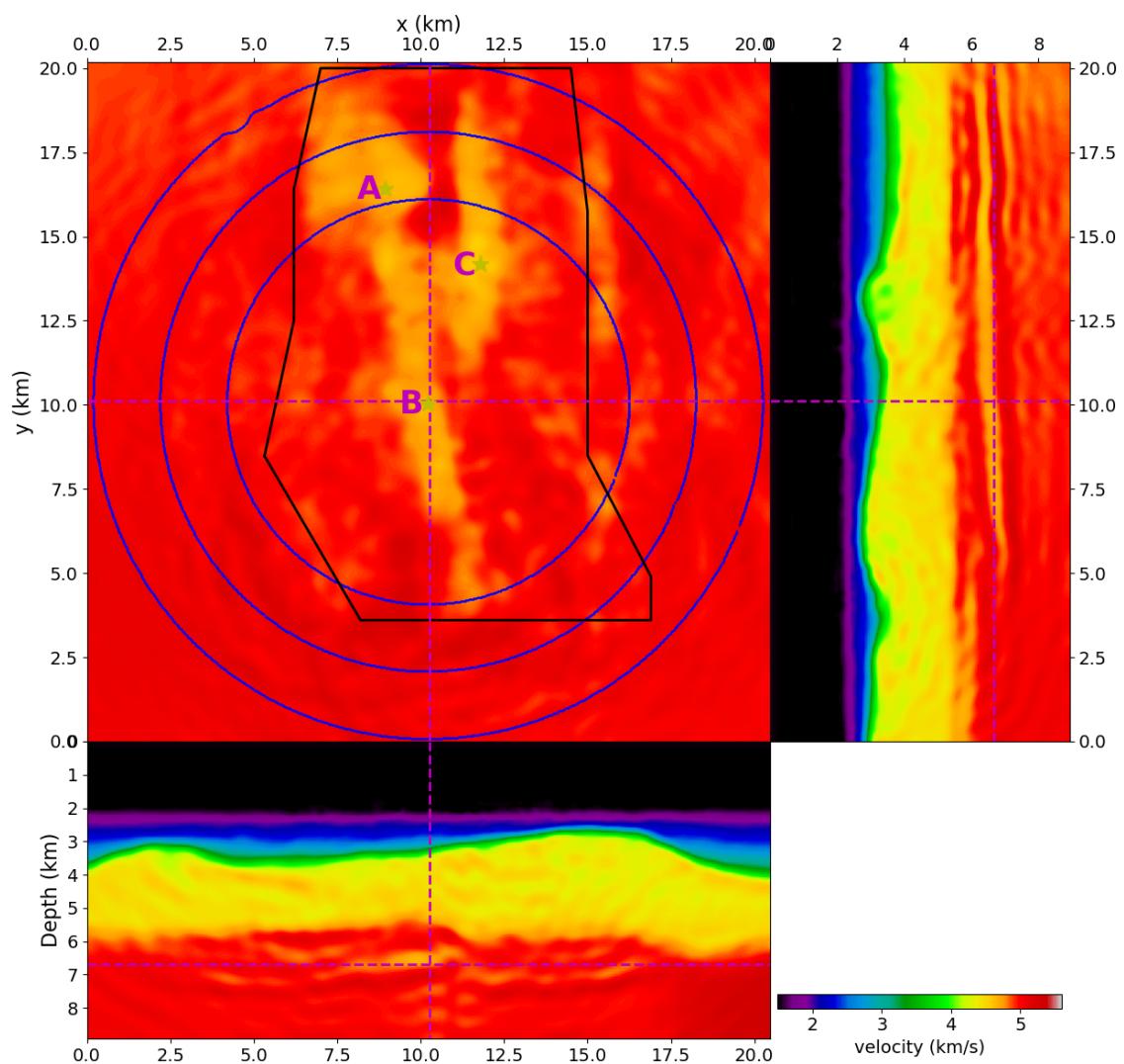


Figura 5.8: Resultado da inversão usando a norma ℓ_2

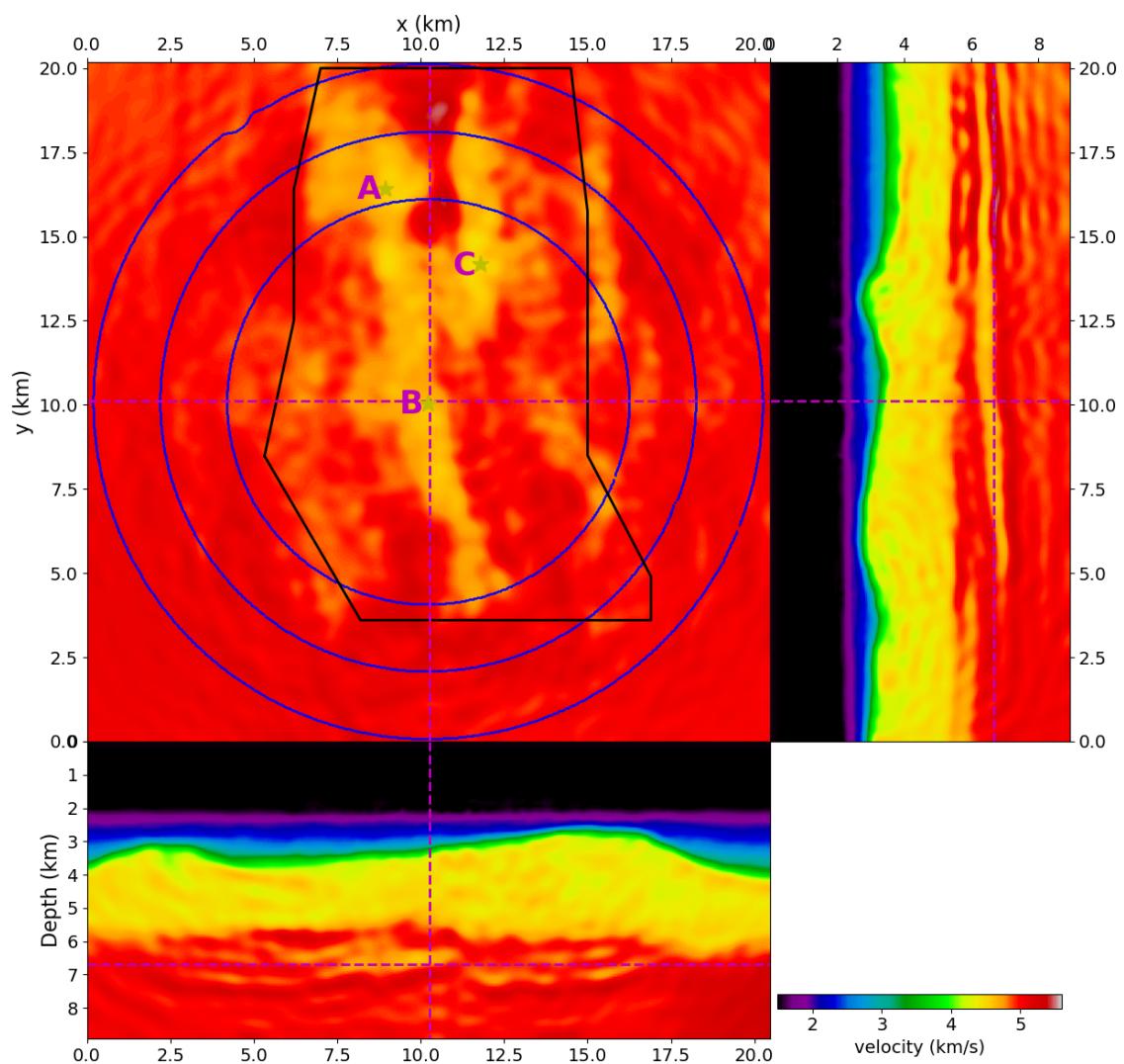


Figura 5.9: Resultado da inversão usando a norma ℓ_1

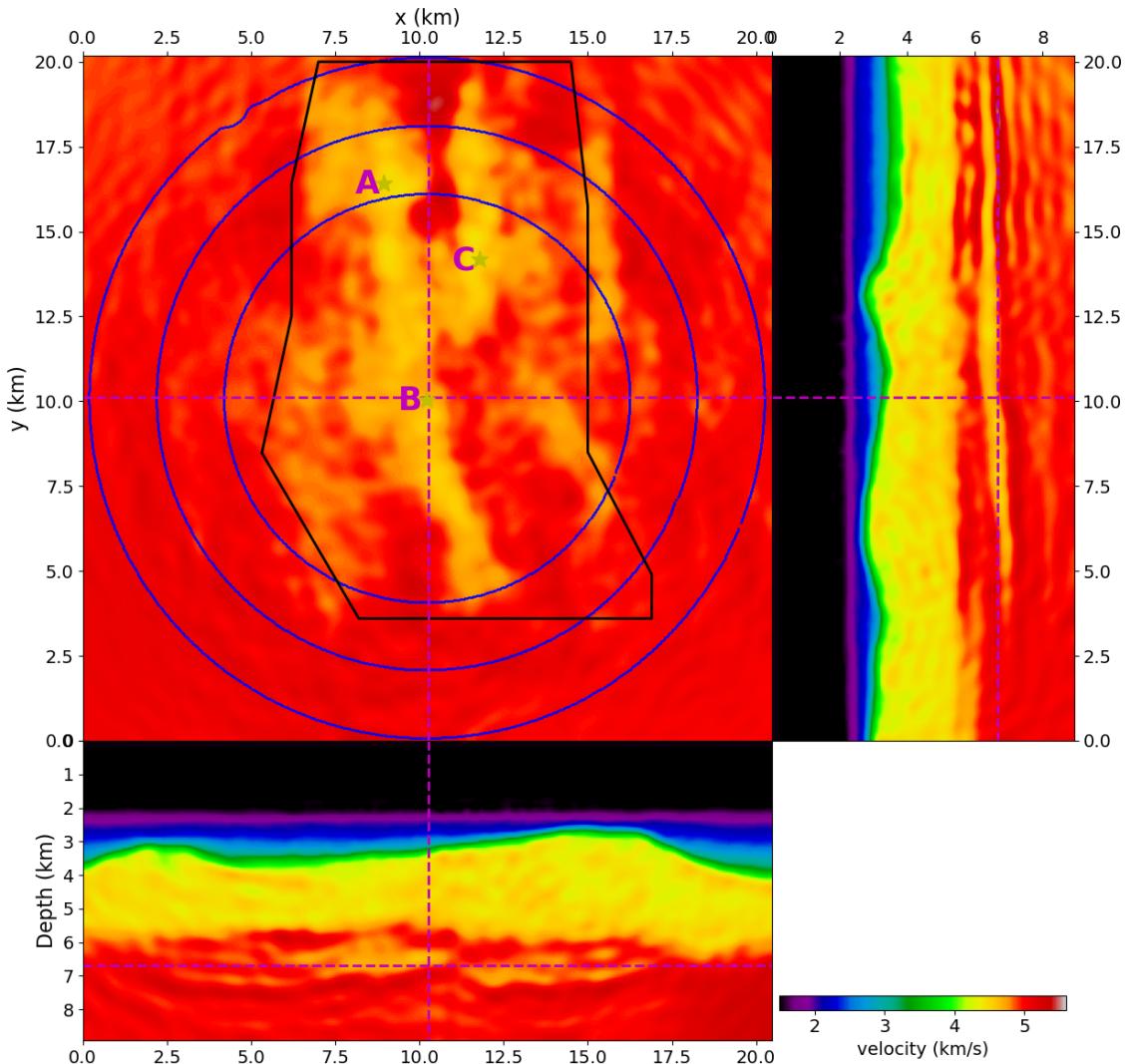


Figura 5.10: Resultado da inversão usando NPR-FWI.

A Figura 5.9 mostra o resultado da FWI usando a norma ℓ_1 . As atualizações do modelo de velocidade nas camadas d'água, de sedimentos pós-sal e o sal, possuem um comportamento semelhante com o observado para a norma ℓ_2 . Em contraste, a atualização na camada do reservatório possui uma maior extensão que a observada no resultado com a norma ℓ_2 . Isto pode ser observado nos cortes inline e crossline onde a camada do reservatório é melhor definida na norma ℓ_1 do que na norma ℓ_2 . No corte em profundidade observamos que no resultado da norma ℓ_1 aparecem novos detalhes no reservatório, principalmente na parte inferior, tanto na esquerda quanto na direita do polígono de receptores.

A Figura 5.10 mostra o resultado para a função objetivo NPR-FWI, onde é observado que a área recuperada do reservatório é ligeiramente maior que a obtida na norma ℓ_1 , mas em geral pode-se considerar que os resultados da NPR-FWI e a norma ℓ_1 são semelhantes entre si e melhores que os obtidos com a norma ℓ_2 .

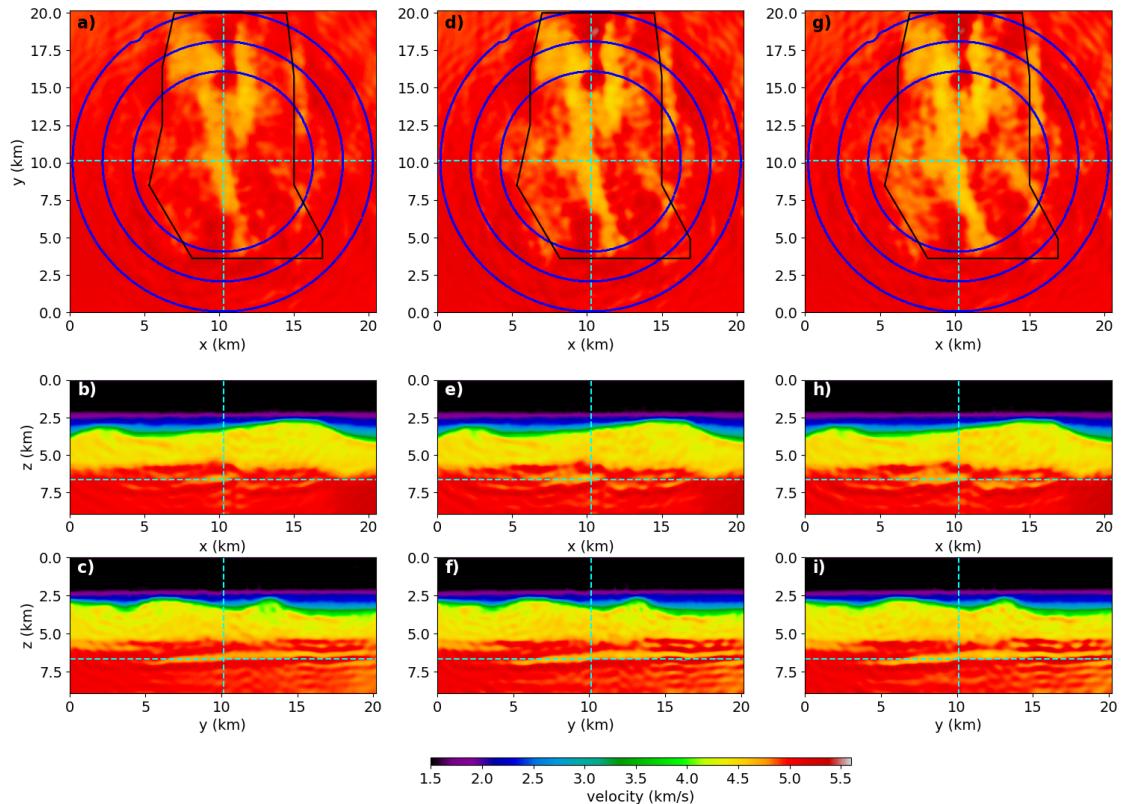


Figura 5.11: Comparaçao das diferentes funções objetivo. (a)-(c) Cortes em profundidade, inline e crossline da inversão com norma ℓ_2 . (d)-(f) Cortes em profundidade, inline e crossline da inversão com norma ℓ_1 . (g)-(i) Cortes em profundidade, inline e crossline da inversão com NPR-FWI. As posições dos cortes são as mesmas das Figuras 5.8 até 5.10

Nenhuma das três funções objetivo testadas conseguiu recuperar informação nas regiões mais afastadas do polígono de receptor, o qual implica que a geometria do levantamento limita a área do reservatório a ser recuperada. Isto é muito importante na hora do planejamento deste tipo de aquisição que seja focada para uso da FWI.

A Figura 5.11 mostra uma comparação direta entre os resultados das Figuras 5.8 até 5.10. Esta figura confirma a vantagem que possui a norma ℓ_1 e a NPR-FWI sobre a norma ℓ_2 na recuperação da extensão do reservatório.

5.8.2 Custo computacional

Para cada uma das inversões realizadas, foram usados quatro nós de um cluster de computadores. Cada nó possui 125 Gb de memória e um processador de 2.7 GHz com 80 Threads, dos quais 40 são reais e os outros 40 são virtuais. Para evitar a sobre carga do processador ou hyper-threading, foram usados somente a metade dos threads, ou seja 40 threads. Estes threads foram distribuídos de tal forma que cada nó faz cálculos para quatro receptores simultaneamente usando dez threads para calcular o gradiente de cada receptor. Isso significa que foram realizados cálculos para 16 receptores simultaneamente.

Para o tamanho e discretização do modelo usado, o cálculo de um gradiente sem dizimar as fontes levou aproximadamente 185 minutos tal como mostrado na Tabela 5.1. Já com a dizimação de fontes o tempo foi reduzido para 85 minutos por iteração. O cálculo do gradiente usando a estratégia do recálculo do campo de onda direto descrito na seção 4.5.2, usa aproximadamente 9.5 Gb de memória RAM por receptor. Guardando todos os campos de onda na propagação direta o gradiente de um receptor usaria aproximadamente 34 Gb o que representa uma economia de memória de 72%.

5.8.3 Controle de qualidade das inversões

Para checar se os resultados das inversões fornecem um modelo de velocidade que seja confiável ou que esteja melhorando o modelo inicial, foram calculados sismogramas sintéticos usando o modelo de referência da Figura 5.3, o modelo inicial da Figura 5.4(a) e os modelos invertidos das Figuras 5.8-5.10. Estos sismogramas foram comparados com o sismograma filtrado de campo. A posição do sismograma esta mostrada na Figura 5.1(a). A wavelet usada para calcular os sismogramas sintéticos é calculada a partir do dado real usando a metodologia do algoritmo 4.

A Figura 5.12 mostra a comparação do dado real filtrado (Figura 5.12(b)) com os sismogramas calculados nos diferentes modelos. Em geral, os sismogramas modelados nos modelos invertidos (Figuras 5.12(d)-(f)) são similares entre si e não é possível decidir qual inversão é melhor usando estes sismogramas. Porém, esses sismogramas apresentam mais semelhança com o dado observado de campo (Figura 5.12(b)) que os sismogramas modelados nos modelos de referência (Figura 5.12(a)) e inicial (Figura 5.12(c)). As principais diferenças são apontadas com as setas coloridas.

A seta branca indica a reflexão primária no topo do corpo do sal. Pode-se

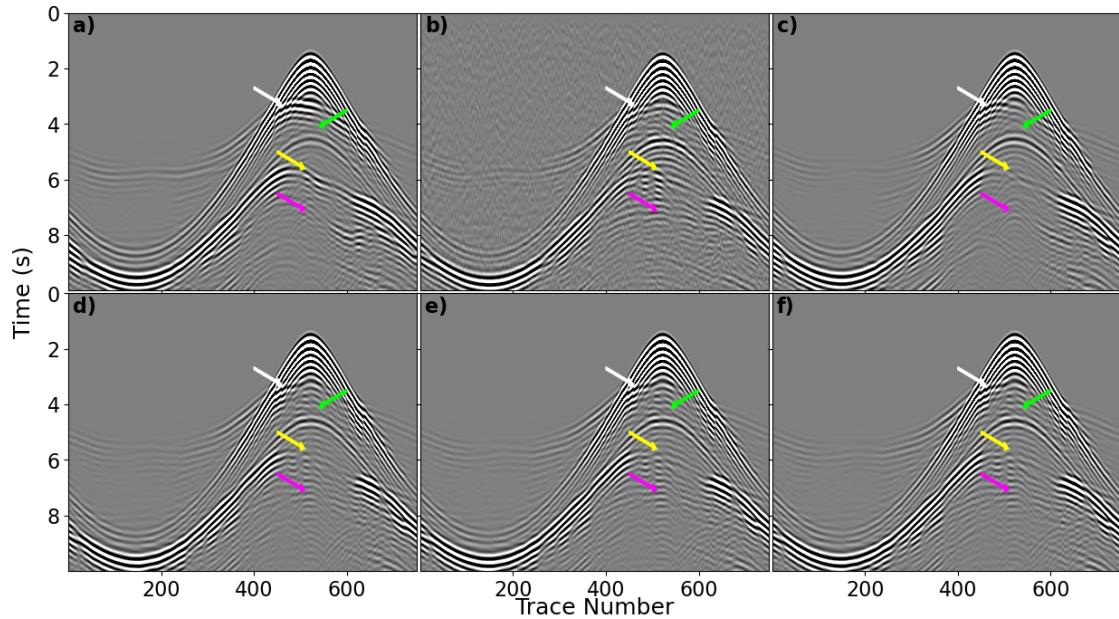


Figura 5.12: Comparação de (b) dados observados de campo com os dados modelados nos modelos (a) de referência, (c) inicial e como os resultados da inversão com (d) norma ℓ_2 , (e) norma ℓ_1 e (f) NPR-FWI.

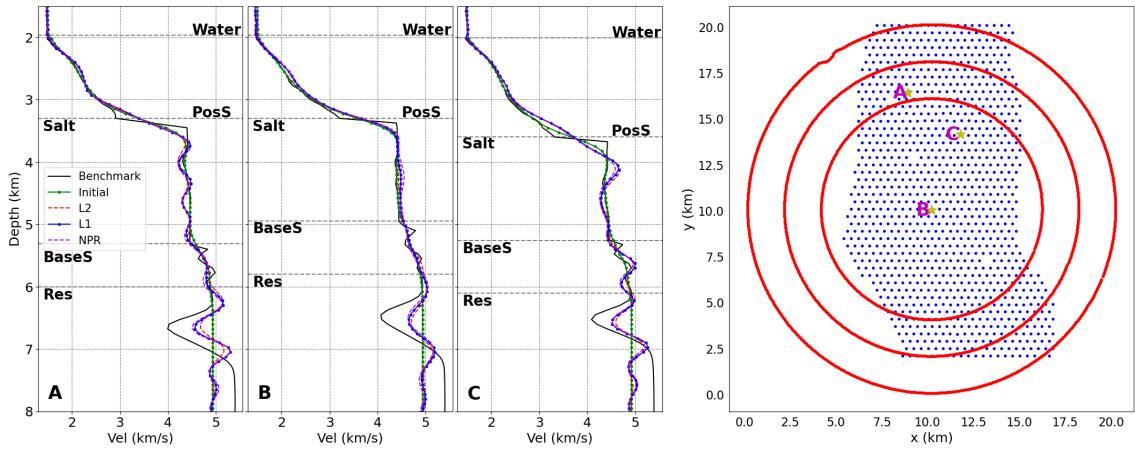


Figura 5.13: Perfis verticais de velocidade comparando o modelo de referência com o modelo inicial e os modelos resultado da inversão com as diferentes funções objetivos. No mapa da direita são apresentadas as posições dos perfis.

observar que existe uma diferença nos tempos de viagem entre as modelagens feitas no modelos de referência e inicial quando comparadas com o dado observado. Esta diferença aparece corregida nos sismogramas modelados nos modelos invertidos. Esta diferença de cinemática pode ser explicada pela redução na velocidade no topo do sal observada na seção crossline em $y = 13\text{ km}$.

A seta verde indica a reflexão primária causada pelas camadas do pré-sal. Como no anterior caso, os sismogramas calculados nos modelos invertidos mostram mais similaridade com os dados observados. As setas amarela e roxa mostram as primeiras múltiplas das reflexões do topo do sal e das camadas pré-sal. Estes eventos também apresentam uma correção na cinemática da onda, mostrando maior semelhança com o dado observado. Vale a pena notar que, embora os sismogramas calculados nos modelos invertidos aproximam-se mais ao dado observado, ainda existem diferenças de amplitude, as quais podem indicar que a física usada na inversão é insuficiente para explicar o dado observado.

Também comparamos alguns perfis de velocidade vertical na Figura 5.13. Nos pontos escolhidos, os perfis invertidos mostram que as três funções objetivo fornecem resultados semelhantes. Conforme apontado na discussão, as funções objetivo não atualizam a velocidade na camada de água e sedimentos pós-sal, e atualizam levemente a camada de sal. A principal atualização está na camada de reservatório, onde as inversões obtêm uma velocidade de reservatório maior do que o modelo de referência. Observa-se também que o contraste de alta velocidade entre as camadas pós-sal e sal não pode ser recuperado pelas inversões devido à falta de altas frequências nos dados usados durante a inversão.

Pode-se observar no perfil vertical no ponto **C** da Figura 5.13, que as inversões recuperam alguns eventos na base da camada de sal que também estão presentes no modelo de referência. Esta informação não é obtida nos pontos **A** e **B** porque o comprimento de onda dos refletores é menor que o comprimento de onda mínimo fornecido pelos dados na banda de frequência usada na inversão.

5.9 Resultados com fontes codificadas

A dizimação dos receptores e das fontes seguindo a estratégia da seção 5.5, pode levar a não aproveitar a informação de todos os receptores simultaneamente. Como os receptores usados nas inversões são iluminados pelas mesmas fontes, então foram feitos testes com as estratégias de fontes simultâneas codificadas estudadas no capítulo 3.

Na literatura existem alguns exemplos de trabalhos usando fontes codificadas em dado real como o trabalho de Schiemenz & Igel 2013. Devido a que o princípio de reciprocidade é aplicado nas inversões, daqui em diante é usado o termo *super-receptor* para indicar o sismograma gerado pelo empilhamento de vários receptores os quais são propagados simultaneamente no cálculo do gradiente. O termo super-receptor é análogo ao termo super-tiro usado quando não é aplicada a reciprocidade na inversão.

Uma das observações feitas em Schiemenz & Igel 2013, que também foi comprovada neste trabalho, é que em dado real não é conveniente combinar todos os receptores em um único super-receptor como foi feito nos testes de dados sintéticos do capítulo 3, mas devem ser usados vários super-receptores. Para isto os receptores da aquisição (por estar usando a reciprocidade) são divididos em vários grupos e os dados observados para cada grupo são empilhados para formar o super-receptor.

O fluxograma de inversão para os testes com fontes codificadas é o mesmo usado para a inversão convencional com exceção da dizimação dos receptores. O modelo inicial também é o mesmo dos testes sem codificação. Do mesmo jeito que foi feito na inversão convencional, a wavelet fonte foi estimada independentemente para cada receptor. A codificação foi aplicada seguindo o procedimento descrito na seção 3.3. Para obter um fluxograma de inversão ótimo foram testados empiricamente quatro parâmetros: A dizimação de fontes, a função objetivo a ser usada, o número de super-receptores usado para calcular o gradiente e o esquema de codificação. Estes parâmetros influem na qualidade das inversões e no custo computacional das mesmas. Em todos os testes com fontes codificadas apresentados nesta seção foram feitas 50 iterações de FWI.

5.9.1 Escolha de dizimação de fontes

Como visto na seção 5.5, o efeito da dizimação ajuda a reduzir o tempo de cálculo do gradiente sem comprometer a qualidade da inversão. Porém, no caso das fontes simultâneas existe o ruído de crosstalk o qual é amortecido pela combinação da codificação e pela estratégia de inversão, que no caso deste trabalho é o l-BFGS reinicializado descrito na seção 3.4.

O nível do ruído crosstalk depende da quantidade de traços do sismograma do super-tiro (ou super-receptor) mas a eficiência da codificação para eliminar o crosstalk também aumenta com o número de traços. Portanto, foi necessário fazer testes sobre o parâmetro de dizimação de fontes diz para ver o efeito desta sobre a qualidade da inversão. O parâmetro diz indica a cada quantas fontes é feita para dizimação. Por exemplo, no caso da inversão convencional $diz = 9$ uma vez que foi usado 1 a cada 9 fontes. Nesta seção foram feitas três inversões com $diz = 9$ como foi feito na FWI convencional, $diz = 5$ e $diz = 1$, ou seja, sem dizimar. Os três testes foram feitos usando 8 super-receptores para calcular o gradiente e a norma ℓ_2 como função objetivo e com a codificação por polarização (multiplicar cada receptor aleatoriamente por 1 ou -1).

Na Figura 5.14 é apresentada a comparação das três dizimações. As inversões apresentam bastante ruido de crosstalk o qual pode ser identificado como as “manchas” que aparecem na definição do reservatório tal como visto nos cortes em profundidade. Porém, na avaliação visual das inversões é observado que a inversão com $diz = 9$ possui mais ruído de crosstalk que as outras duas dizimações e que as inversões com $diz = 5$ e $diz = 1$ tem praticamente o mesmo resultado. Como mostrado na Tabela 5.1, o cálculo do gradiente com $diz = 5$ demora um pouco menos da metade que usando $diz = 1$. Portanto, como as duas dizimações tem um resultado

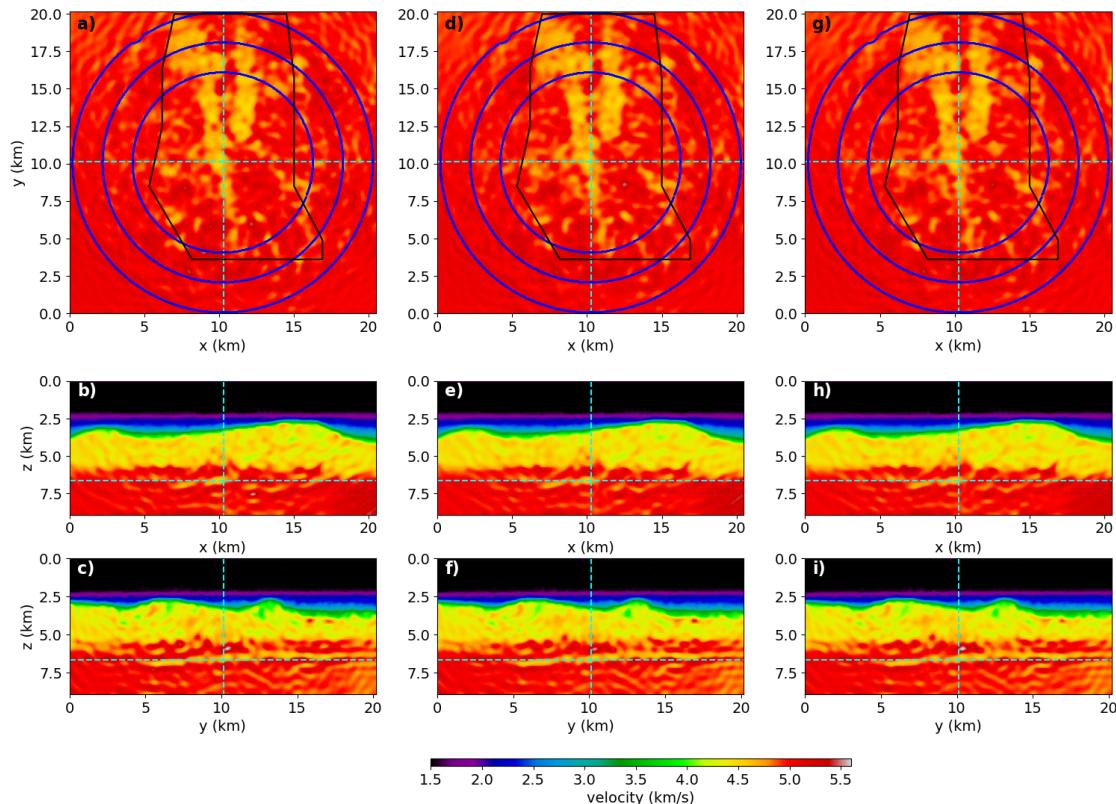


Figura 5.14: Comparaçāo de inversões de codificação por polarizaçāo usando 8 super-receptores variando a dizimação de fontes. Inversão (a)-(c) dizimando fontes 1 a cada 9, (d)-(f) dizimando fontes 1 a cada 5, (g)-(i) sem dizimar fontes. Os cortes em profundidade, inline e crossline são os mesmos da Figura 5.11

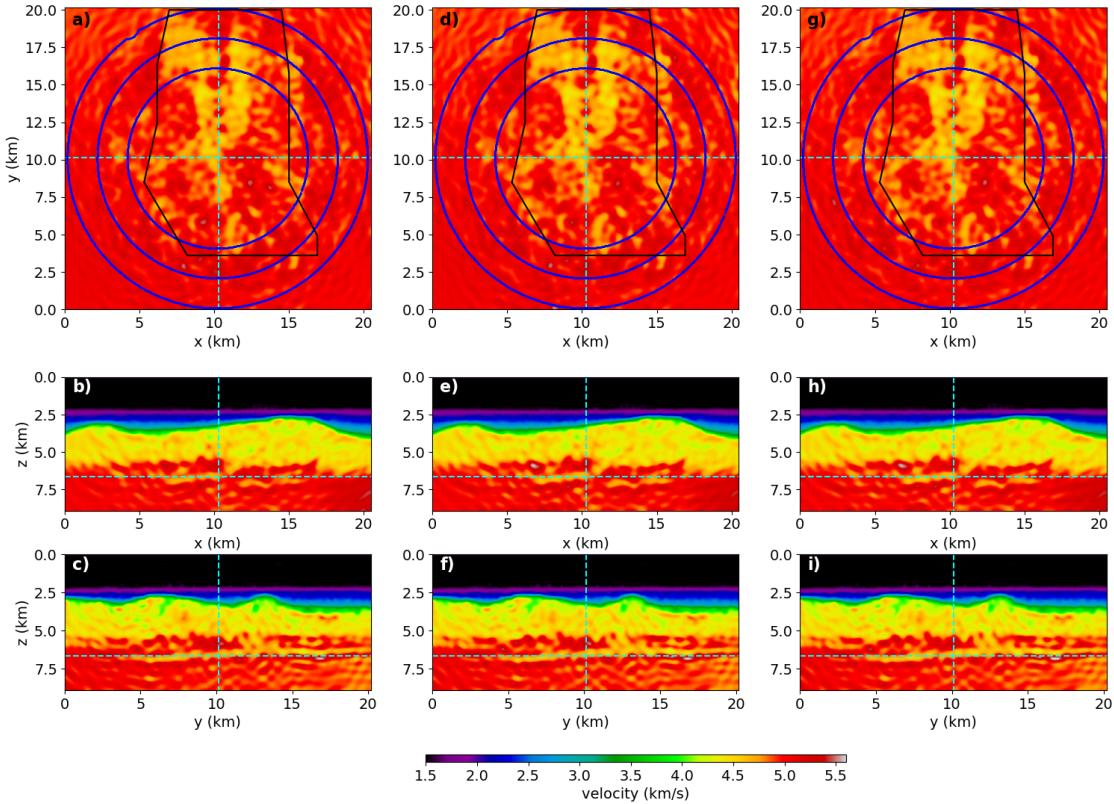


Figura 5.15: Comparação de inversões usando codificação por polarização usando 8 super-receptores usando as funções objetivo descritas na seção 5.6. (a)-(c) usando a norma ℓ_2 . (d)-(f) usando a norma ℓ_1 . (g)-(i) usando NPR-FWI. Os cortes em profundidade, inline e crossline são os mesmos das figuras anteriores.

parecido, foi usado $diz = 5$ para todas as inversões daqui em diante.

5.9.2 Escolha de função objetivo

O ruído de crosstalk corresponde com o segundo termo da equação 3.5, ou seja

$$\nabla \Phi_{xtalk} = - \sum_{i \neq j}^{N_s} \int_0^T e_i(-t) \otimes R_i(-t) \frac{\partial^2 e_j \otimes P_j(t)}{\partial t^2} dt. \quad (5.6)$$

Desta equação pode-se inferir que o ruído de crosstalk é inherent ao cálculo do gradiente sem importar a função objetivo usada, mas a forma e quantidade do ruído crosstalk vai depender da fonte adjunta, que por sua vez depende da função objetivo (ver seção 2.5.2). Pode-se intuir que uma função objetivo cuja fonte adjunta seja menos sensível a amplitude, pode gerar menos ruído de crosstalk. Portanto, além da função objetivo norma ℓ_2 , foram testadas a norma ℓ_1 e NPR-FWI que são menos sensíveis às variações de amplitude tal como explicado na seção 5.6.

Na Figura 5.15 são apresentados os resultados das três funções objetivo. Nos testes foram usados 8 super-receptores para calcular o gradiente, uma dizimação de

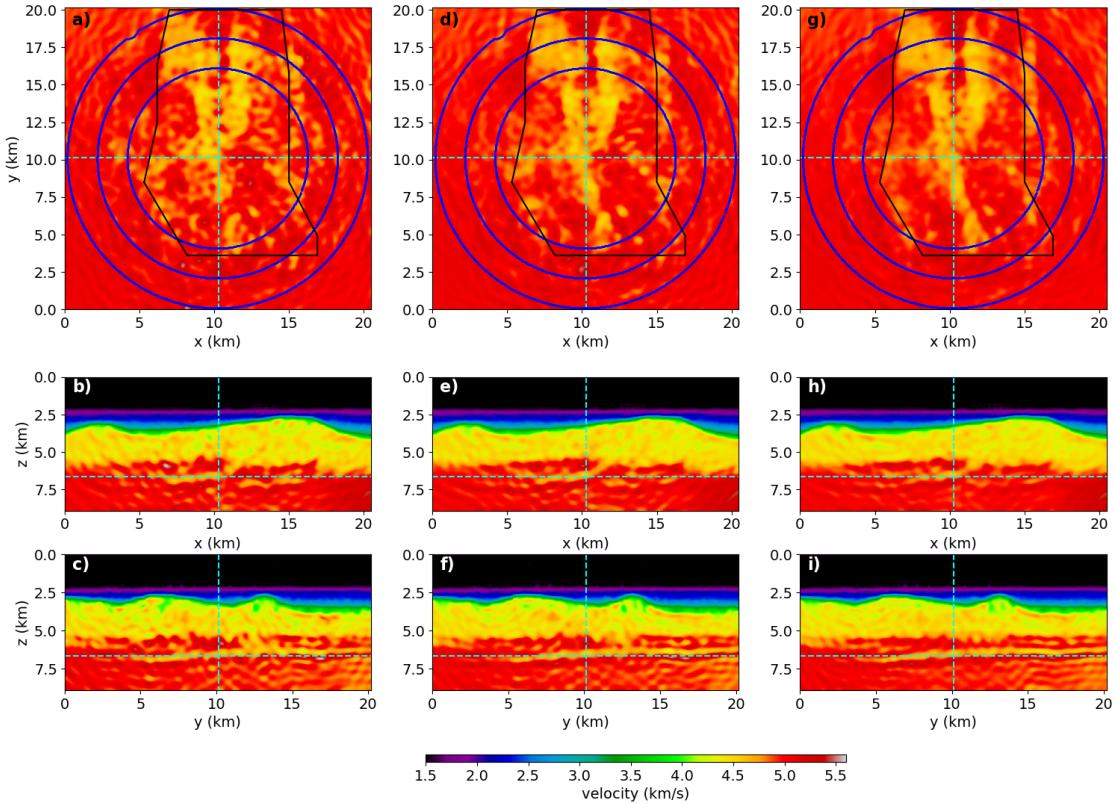


Figura 5.16: Comparação de inversões usando codificação por polarização usando (a)-(c) 8 super-receptores, (d)-(f) 16 super-receptores e (g)-(i) 32 super-receptores. Os cortes em profundidade, inline e crossline são os mesmos das figuras anteriores.

fontes $diz = 5$ e foi usada a codificação por polarização. Embora os três resultados possuem bastante ruído de crosstalk, foi observado que a norma ℓ_1 possui mais definição da área no reservatório, o qual pode apreciar-se nos cortes em profundidade. Portanto, foi usada a norma ℓ_1 para os testes com diferentes quantidades de super-receptores.

5.9.3 Escolha de quantidade de super-receptores

A quantidade de super-receptores ou super-fontes ajuda ao amortecimento do ruído de crosstalk, porém diminui a eficiência computacional do método de fontes codificadas. Portanto, deve ser testado o número de super-receptores usados de tal forma que o ruído de crosstalk seja amortecido sem afetar tanto o custo computacional.

Na Figura 5.16 são mostrados os resultados para 8, 16 e 32 super-receptores. As inversões foram feitas usando a dizimação de fontes com $diz = 5$, foi usada a norma ℓ_1 com função objetivo e a codificação por polarização. Pode-se observar que entre mais super-receptores foram usados mais amortecido é o ruido de crosstalk fazendo com que a recuperação da informação do reservatório fosse melhor quando

foram usados 32 super-receptores. Também houve uma melhora na continuidade das camadas recuperadas na área do pré-sal.

Os testes de número de super foram feitos aumentando o recurso computacional de tal forma que o número de super-receptores por nó de computo fosse a mesma. Isto quer dizer, que como o número de super-receptores foi duas e quatro vezes mais que os testes anteriores, então o recurso computacional foi aumentado na mesma proporção tal como mostrado na Tabela 5.1. Em particular a inversão com 32 super-receptores usa o mesmo recurso computacional que a inversão convencional, ou seja, sem codificação, fazendo o cálculo do gradiente em um tempo quase três vezes menor.

5.9.4 Escolha do esquema de codificação

Por último, foram testados os esquemas de codificação por polarização e por rotação de fase a qual é descrita na seção 3.3. A codificação por estática também foi testada mas não foi obtido um bom resultado. Os testes foram feitos usando uma dizimação com $diz = 5$, usando norma ℓ_1 , e usando 32 super-receptores para calcular o gradiente. Os resultados destas inversões são apresentadas nas Figuras 5.17 e 5.18.

A Figura 5.19 apresenta a comparação entre a codificação por polarização, por rotação de fase e junto a estes resultados foi colocado o resultado da norma ℓ_1 sem codificação apresentado na Figura 5.9. Os cortes em profundidade mostram que o resultado da codificação por rotação de fase apresenta menos de ruído crosstalk que a codificação por polarização. Isto pode ser observado no corte de profundidade onde o reservatório recuperado é mais parecido com o obtido na inversão sem codificação. Também pode-se observar na seção inline no cruzamento das linhas verdes na Figura 5.19(b), (e), (h), que o reservatório recuperado na codificação por polarização ainda é afetado pelo ruído crosstalk o que não acontece na codificação por rotação de fase.

O custo de calcular o gradiente na codificação por polarização e por rotação de fase foi praticamente o mesmo. Após os diferentes testes para obter os melhores parâmetros da codificação, o custo do cálculo do gradiente foi reduzido aproximadamente 2.57 vezes quando comparado com a inversão sem codificação obtendo uma qualidade semelhante, como pode-se observar na Tabela 5.1.

Na Figura 5.20 é apresentada uma comparação de perfis velocidade semelhante à Figura 5.13, mas comparando os modelos na Figura 5.19 com o modelo inicial e o modelo de referência. A recuperação na área do reservatório é semelhante nas inversões sem e com codificação. O ruído de crosstalk afeta principalmente o resultado dentro da camada de sal (ver perfil C) e na base do sal.

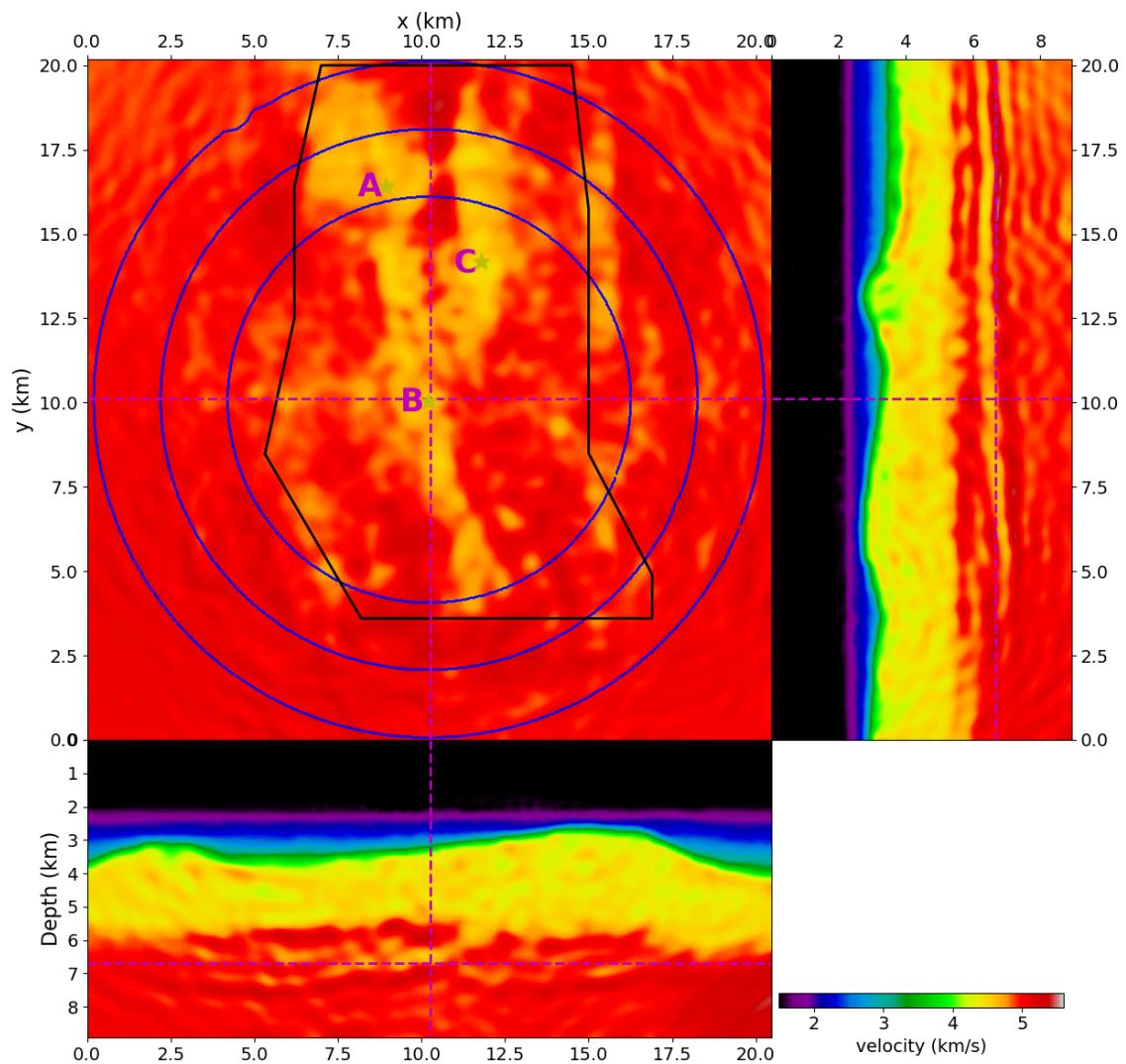


Figura 5.17: Resultado da inversão usando norma ℓ_1 usando codificação por polarização.

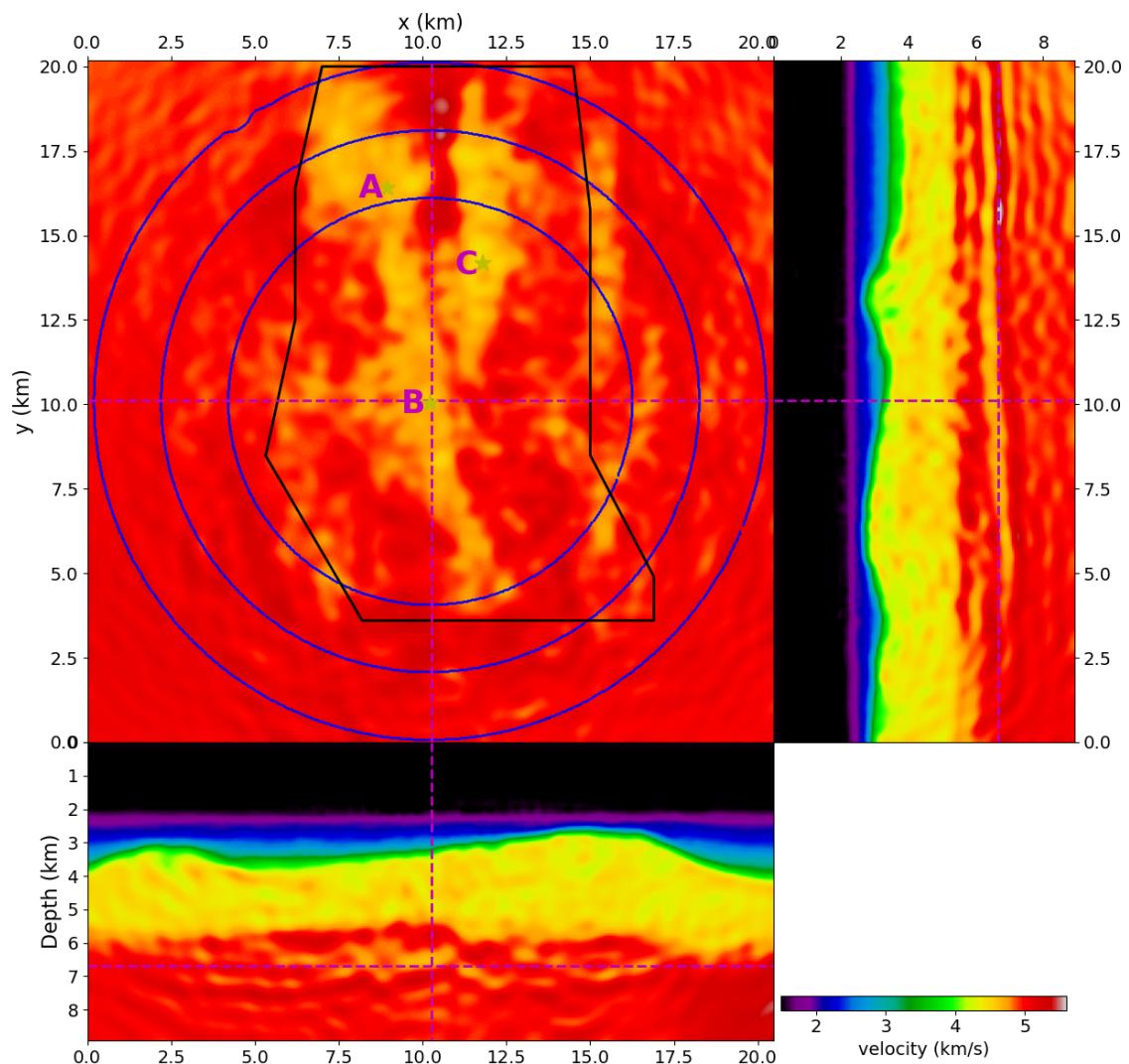


Figura 5.18: Resultado da inversão usando norma ℓ_1 usando codificação por rotação de fase.

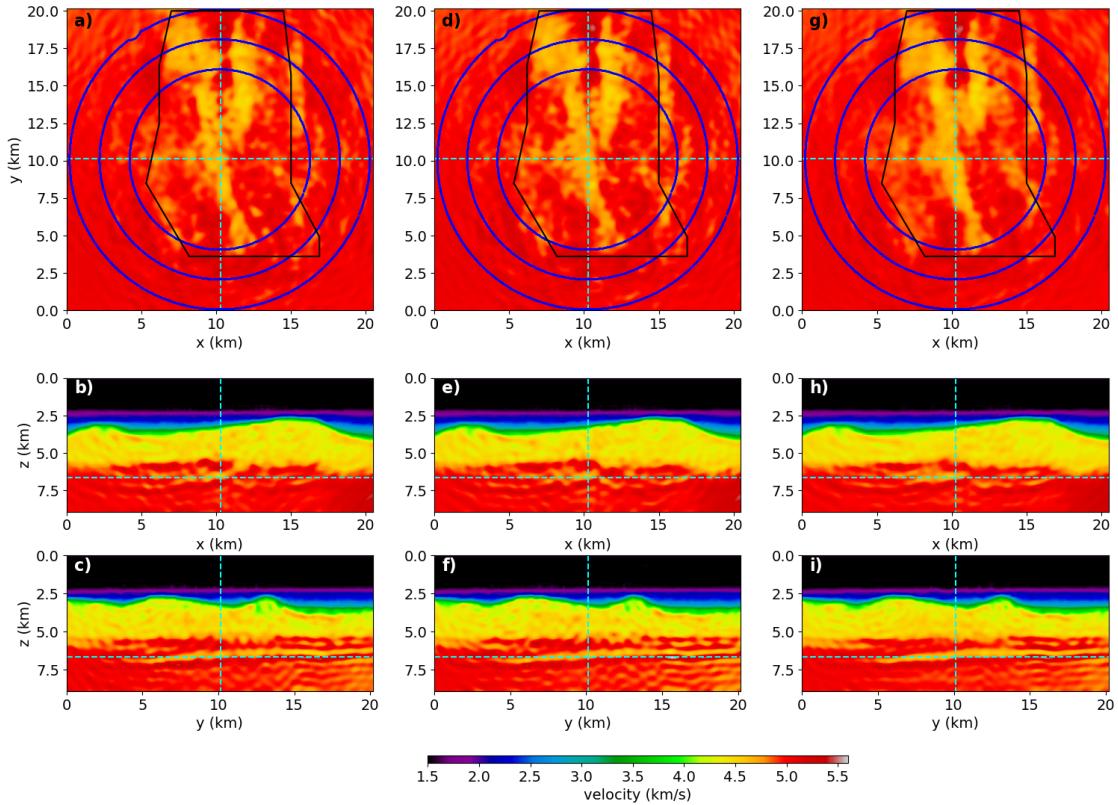


Figura 5.19: Comparaçāo de inversões usando (a)-(c) codificação por polarizaçāo, (d)-(f) codificação por rotaçāo de fase e (g)-(i) inversão convencional com a norma ℓ_1 da Figura 5.9. Os cortes em profundidade, inline e crossline são os mesmos das figuras anteriores.

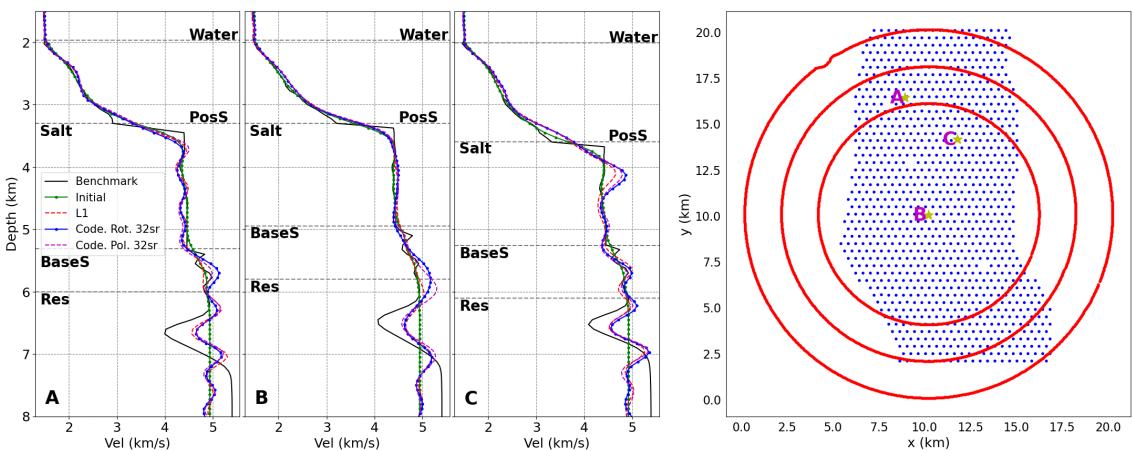


Figura 5.20: Perfis verticais de velocidade comparando o modelo de referência com o modelo inicial e os modelos invertido com a norma ℓ_1 sem codificação (Figura 5.9), com codificação por polarizaçāo (Figura 5.17) e por rotaçāo de fase (Figura 5.18). No mapa da direita são apresentadas as posições dos perfis.

Tipo de inversão	Número de nós	Super-receptores simultâneos	Tempo (min)
Conv. 80 receptores por gradiente, diz=1 (seção 5.8)	4	16	185
Conv. 80 receptores por gradiente, diz=9 (seção 5.8)	4	16	85
Codificado 8 super-receptores, diz=1	1	4	70
Codificado 8 super-receptores, diz=9	1	4	28
Codificado 8 super-receptores, diz=5	1	4	31
Codificado 16 super-receptores, diz=5	2	8	34.75
Codificado 32 super-receptores, diz=5	4	16	33

Tabela 5.1: Comparaçāo de tempo de cálculo de um gradiente para os diferentes tipos de inversões feitas neste trabalho. É especificado também o número de nós usados e número de receptores ou super-receptores que são calculados simultaneamente. O parāmetro diz indica o nível de dizimação das fontes. O número de threads por receptor (ou super-receptor) é de 10 e cada nó de computo faz cálculos para 4 receptores (ou super-receptores) simultaneamente.

Capítulo 6

Conclusões

Nesta tese foram estudados diferentes estratégias que visaram reduzir o custo computacional da FWI baseadas em propriedades físicas da propagação de onda. Isto foi feito em duas partes. Na primeira parte, foram propostos dois esquemas de codificação de fontes para realizar a FWI por fontes simultâneas. Ditos esquemas são por rotação de fase aleatória e por estática aleatória limitada a um período da frequência dominante. Na codificação por rotação de fase é aplicada uma rotação de fase em cada tiro onde o ângulo de rotação é escolhido aleatoriamente para cada tiro. No caso da codificação por estática, os traços de cada tiro são deslocados temporalmente e o valor deste deslocamento é escolhido aleatoriamente para cada tiro tendo um valor máximo de um período da frequência dominante do dado. Os testes foram feitos com dados sintéticos 2D calculados no modelo overthrust. Foi demonstrado que todos os esquemas de codificação, junto com o algoritmo de L-BFGS reinicializado, podem mitigar o ruído de crosstalk que aparece no cálculo do gradiente por fontes simultâneas obtendo resultados semelhantes aos obtidos usando a FWI convencional. Porém, os esquemas de codificação propostos apresentam uma convergência mais rápida em comparação aos esquemas já existentes por polarização aleatória (multiplicar por -1 e 1 aleatoriamente cada tiro) e por estática aleatória sem ter valor limitada.

Na segunda parte da tese foi desenvolvido um fluxograma para aplicar a FWI em um conjunto de dados de um levantamento 3D OBN adquirido na área do pré-sal brasileiro. Este levantamento foi projetado para beneficiar os longos traços de deslocamento onde a informação de ondas de mergulho é dominante. O fluxograma desenvolvido usa estratégias baseadas em propriedades da equação da onda, as quais visam reduzir o tempo computacional e uso de memória. Entre as estratégias usadas estão a reamostragem temporal e a dizimação de fontes e receptores, a quais reduzem o tempo computacional e o recálculo do campo direto através do uso dos valores do campo na fronteira espacial a qual reduz o uso de memória.

Com o fluxograma desenvolvido, foram testadas três funções objetivo, a saber, norma ℓ_2 , norma ℓ_1 e NPR-FWI. Embora as inversões com as três funções objetivo recuperam informação do reservatório, a norma ℓ_1 e NPR-FWI recuperam um área mais extensa do reservatório que a norma ℓ_2 . Com todas as funções objetivo a melhor recuperação ocorre na área dos receptores, o qual demonstra que a resultado da inversão é fortemente influenciada pela geometria de aquisição. A geometria do

reservatório recuperada pela inversão é semelhante à observada em um modelo de referência obtido anteriormente com outro levantamento OBN 3D mais denso que o usado neste trabalho. O valor da velocidade no reservatório obtido nas inversões é superior ao valor no modelo de referência.

Nos controles de qualidade feitos sobre as inversões, foi observado que os sismogramas modelados nos modelos invertidos estão mais próximos em fase e amplitude aos dados de campo, do que os sismogramas modelados nos modelos inicial e de referência, demonstrando que o fluxograma de FWI 3D usado melhora o modelo de velocidade.

Com o fluxograma desenvolvido também foram feitos diferentes testes para aplicar os esquemas de fontes codificadas estudadas na primeira parte da tese. Os testes visaram obter os parâmetros ótimos de dizimação de fontes, função objetivo, número de super-receptores usado no cálculo do gradiente e de esquema de codificação. Os melhores parâmetros foram obtidos por avaliação visual das diferentes inversões. Embora o melhor resultado obtido com as fontes codificadas ainda apresenta ruído de crosstalk, este resultado possui uma qualidade semelhante à inversão sem codificação mas com uma redução considerável do custo computacional.

Mais estudos devem ser feitos sobre a aplicação da FWI 3D neste OBN 3D para melhorar a qualidade da inversão obtido. Por exemplo, é necessário estudar o efeito da atenuação sísmica na modelagem assim como também o efeito de atualizar o modelo de densidade durante a inversão, isto com o intuito de explicar algumas diferenças observadas entre os dados modelados e de campo que não foi possível modelar com a física usada neste trabalho. Uma das conclusões mais interessantes dos experimentos realizados neste trabalho é que a geometria do levantamento usado permite recuperar informações da área do pré-sal. Esta geometria tem como vantagem ser de baixo custo, o que implica que este tipo de levantamento podem ser usado para monitoramento contínuo na área do pré-sal.

Referências Bibliográficas

- Alford, R. M., K. R. Kelly & D. M. Boore 1974, ‘ACCURACY OF FINITE-DIFFERENCE MODELING OF THE ACOUSTIC WAVE EQUATION’, *GEOPHYSICS* **39**(6), 834–842.
- Anno, Phil, Norman Bleistein & Jack K. Cohen 1992, Waves and rays in acoustic variable density media, *em* ‘SEG Technical Program Expanded Abstracts 1992’, Society of Exploration Geophysicists.
- Arfken, George & Hans Weber 2005, *Mathematical methods for physicists*, Academic press.
- Barnes, Arthur E. 2007, ‘A tutorial on complex seismic trace analysis’, *Geophysics* **72**(6), W33–W43.
- Ben-Hadj-Ali, Hafedh, Stéphane Operto & Jean Virieux 2011, ‘An efficient frequency-domain full waveform inversion method using simultaneous encoded sources’, *Geophysics* **76**(4), R109–R124.
- Berenger, Jean-Pierre 1994, ‘A perfectly matched layer for the absorption of electromagnetic waves’, *Journal of Computational Physics* **114**(2), 185–200.
- Bezanson, Jeff, Alan Edelman, Stefan Karpinski & Viral B Shah 2017, ‘Julia: A fresh approach to numerical computing’, *SIAM review* **59**(1), 65–98.
URL: <https://doi.org/10.1137/141000671>
- Bracewell, Ronald N. 1999, *The Fourier Transform and Its Applications*, third^a edição, McGraw-Hill.
- Brossier, Romain, Stéphane Operto & Jean Virieux 2010, ‘Which data residual norm for robust elastic frequency-domain full waveform inversion?’, *GEOPHYSICS* **75**(3), R37–R46.
- Bunks, Carey, Fatimetou M. Saleck, S. Zaleski & G. Chavent 1995, ‘Multiscale seismic waveform inversion’, *GEOPHYSICS* **60**(5), 1457–1473.
- Carvalho, P.T., S.L.E.F. Silva, E.F. Duarte, R. Brossier, G. Corso & J.M. de Araujo 2021, ‘Full waveform inversion based on the non-parametric estimate of the probability distribution of the residuals’, *Geophysical Journal International* p. On revision.

- Castagna, John P., Michael L. Batzle, Kenneth M. Tubman, James E. Gaiser, Michael D. Burnett, M. L. Batzle & T. K. Kan 1993, 2. petrophysics, em ‘Offset-Dependent Reflectivity—Theory and Practice of AVO Analysis’, Society of Exploration Geophysicists, pp. 113–172.
- Castellanos, Clara, Ludovic Métivier, Stéphane Operto, Romain Brossier & Jean Virieux 2014, ‘Fast full waveform inversion with source encoding and second-order optimization methods’, *Geophysical Journal International* **200**(2), 720–744.
- Cerjan, Charles, Dan Kosloff, Ronnie Kosloff & Moshe Reshef 1985, ‘A nonreflecting boundary condition for discrete acoustic and elastic wave equations’, *GEOPHYSICS* **50**(4), 705–708.
- Choi, Yunseok & Tariq Alkhalifah 2012, ‘Application of multi-source waveform inversion to marine streamer data using the global correlation norm’, *Geophysical Prospecting* **60**(4), 748–758.
- Clapp, Robert G. 2009, Reverse time migration with random boundaries, em ‘SEG Technical Program Expanded Abstracts 2009’, Society of Exploration Geophysicists.
- Clayton, Robert & Björn Engquist 1977, ‘Absorbing boundary conditions for acoustic and elastic wave equations’, *Bulletin of the Seismological Society of America* **67**(6), 1529–1540.
- Cypriano, Luis, Ziqin Yu, Denis Ferreira, Benjamin Huard, Roberto Pereira, Florian Jouno, Adel Khalil, Eduardo Naomitsu Agapito Urasaki, Nathalia Martinho Souto Muniz da Cruz, Ailin Yin, Dan Clarke & Carlos Cardeal Jesus 2019, Obn for pre-salt imaging and reservoir monitoring - potential and road ahead, em ‘16th International Congress of the Brazilian Geophysical Society & EXPOGEF, Rio de Janeiro, Brazil, 19–22 August 2019’, Brazilian Geophysical Society.
- da Costa, Carlos A. N., Jessé C. Costa, Walter E. Medeiros, D. J. Verschuur & Alok K. Soni 2018, ‘Target-level waveform inversion: a prospective application of the convolution-type representation for the acoustic wavefield’, *Geophysical Prospecting* **67**(1), 69–84.
- Damelin, Steven B. & Willard Jr Miller 2009, *The Mathematics of Signal Processing*, Cambridge University Press.
- de Moura, Carlos A. & Carlos S. Kubrusly, eds. 2013, *The Courant–Friedrichs–Lewy (CFL) Condition*, Birkhäuser Boston.
- Duarte, Edwin Fagua, Carlos A N da Costa, João M de Araújo, Yanghua Wang & Ying Rao 2020, ‘Seismic shot-encoding schemes for waveform inversion’, *Journal of Geophysics and Engineering* **17**(5), 906–913.

- Fichtner, Andreas 2011, *Full Seismic Waveform Modelling and Inversion*, Springer Berlin Heidelberg.
- Fruehn, Juergen, Stuart Greenwood, Ross O'Driscoll, Ian Jones & John Brittan 2019, A strategy for regional-scale FWI in the salt provinces offshore brazil, *em 'SEG Technical Program Expanded Abstracts 2019'*, Society of Exploration Geophysicists.
- Gao, Fengxia & Yanghua Wang 2019, 'Radiation pattern analyses for seismic multi-parameter inversion of HTI anisotropic media', *Journal of Geophysics and Engineering*.
- Hicks, Graham J. 2002, 'Arbitrary source and receiver positioning in finite-difference schemes using kaiser windowed sinc functions', *GEOPHYSICS* **67**(1), 156–165.
- Huang, Yunsong & Gerard T. Schuster 2014, 'Resolution limits for wave equation imaging', *Journal of Applied Geophysics* **107**, 137–148.
- Ikelle, Luc T. & Lasse Amundsen 2005, *Introduction to Petroleum Seismology*, Society of Exploration Geophysicists.
- Jouno, Florian, Adriano Martinez, Denis Ferreira, Daniela Donno & Adel Khalil 2019, Illuminating santos basin's pre-salt with OBN data: Potential and challenges of FWI, *em 'SEG Technical Program Expanded Abstracts 2019'*, Society of Exploration Geophysicists.
- Kamath, Nishant, Romain Brossier, Ludovic Métivier, Arnaud Pladys & Pengliang Yang 2021, 'Multiparameter full-waveform inversion of 3d ocean-bottom cable data from the valhall field', *GEOPHYSICS* **86**(1), B15–B35.
- Kelly, K. R., R. W. Ward, Sven Treitel & R. M. Alford 1976, 'SYNTHETIC SEISMOGRAMS: A FINITE -DIFFERENCE APPROACH', *GEOPHYSICS* **41**(1), 2–27.
- Komatitsch, Dimitri & Roland Martin 2007, 'An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation', *Geophysics* **72**(5), SM155–SM167.
- Krebs, Jerome R., John E. Anderson, David Hinkley, Ramesh Neelamani, Sunwoong Lee, Anatoly Baumstein & Martin-Daniel Lacasse 2009, 'Fast full-wavefield seismic inversion using encoded sources', *Geophysics* **74**(6), WCC177–WCC188.
- Levander, Alan R. 1988, 'Fourth-order finite-difference p-SV seismograms', *GEOPHYSICS* **53**(11), 1425–1436.
- Lopez, Jorge, Francisco Neto, Marlon Cabrera, Sarah Cooke, Samantha Grandi & Deane Roehl 2020, Refraction seismic for pre-salt reservoir characterization and monitoring, *em 'SEG Technical Program Expanded Abstracts 2020'*, Society of Exploration Geophysicists.

- Marfurt, Kurt J. 1984, ‘Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations’, *GEOPHYSICS* **49**(5), 533–549.
- Martinez, Adriano, Adel Khalil, Werter Oliveira, Eduardo Naomitsu & Ana Moltenero 2020, The potential of sparse nodes for exploration in santos basin, *em ‘SEG Technical Program Expanded Abstracts 2020’*, Society of Exploration Geophysicists.
- Métivier, L., R. Brossier, Q. Mérigot, E. Oudet & J. Virieux 2016, ‘Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion’, *Geophysical Journal International* **205**(1), 345–377.
- Mojica, Oscar F. & Jonathas S. Maciel 2020, Seismic modeling from scratch using devito: A demonstration with a typical brazilian pre-salt model, *em ‘SEG Technical Program Expanded Abstracts 2020’*, Society of Exploration Geophysicists.
- Nocedal, Jorge & Stephen J. Wright 2006, *Numerical Optimization*, 2^a edição, Springer, New York.
- Nolet, Guust 2008, *A Breviary of Seismic Tomography*, Cambridge University Press.
- Operto, S., A. Miniussi, R. Brossier, L. Combe, L. Métivier, V. Monteiller, A. Riboldetti & J. Virieux 2015, ‘Efficient 3-d frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to valhall in the visco-acoustic vertical transverse isotropic approximation’, *Geophysical Journal International* **202**(2), 1362–1391.
- Pasalic, Damir & Ray McGarry 2010, Convolutional perfectly matched layer for isotropic and anisotropic acoustic wave equations, *em ‘SEG Technical Program Expanded Abstracts’*, Society of Exploration Geophysicists.
- Pratt, Gerhard, Changsoo Shin & Hicks 1998, ‘Gauss-newton and full newton methods in frequency-space seismic waveform inversion’, *Geophysical Journal International* **133**(2), 341–362.
- Pratt, R. Gerhard 1999, ‘Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model’, *GEOPHYSICS* **64**(3), 888–901.
- Purves, Steve 2014, ‘Phase and the hilbert transform’, *The Leading Edge* **33**(10), 1164–1166.
- Rao, Y., Y. Wang & D. Han 2019, ‘Seismic waveform tomography with simplified restarting scheme’, *IEEE Geoscience and Remote Sensing Letters* **16**(1), 135–139.
- Rao, Ying & Yanghua Wang 2017, ‘Seismic waveform tomography with shot-encoding using a restarted l-BFGS algorithm’, *Scientific Reports* **7**(8494).

- Romero, L. A., D. C. Ghiglia, C. C. Ober & S. A. Morton 2000, ‘Phase encoding of shot records in prestack migration’, *Geophysics* **65**(2), 426–436.
- Schiemenz, Alan & Heiner Igel 2013, ‘Accelerated 3-d full-waveform inversion using simultaneously encoded sources in the time domain: application to valhall ocean-bottom cable data’, *Geophysical Journal International* **195**(3), 1970–1988.
- Schuster, G. T., G. Zhan, W. Dai & C. Boonyasiriwat 2010, Acoustic multi-source waveform inversion with deblurring, *em* ‘72nd EAGE Conference’.
- Schuster, G. T., X. Wang, Y. Huang, W. Dai & C. Boonyasiriwat 2011, ‘Theory of multisource crosstalk reduction by phase-encoded statics’, *Geophysical Journal International* **184**(3), 1289–1303.
- Schuster, Gerard T. 2017, *Seismic Inversion*, Society of Exploration Geophysicists.
- Shen, Xukai, Imtiaz Ahmed, Andrew Brenders, Joe Dellinger, John Etgen & Scott Michell 2018, ‘Full-waveform inversion: The next leap forward in subsalt imaging’, *The Leading Edge* **37**(1), 67b1–67b6.
- Sheriff, Robert E. 2002, *Encyclopedic Dictionary of Applied Geophysics*, Society of Exploration Geophysicists.
- Silverman, B. W. 1986, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.
- Sirgue, L., O.I. Barkved, J. Dellinger, J. Etgen, U. Albertin & J.H. Kommedal 2010, ‘Thematic set: Full waveform inversion: the next leap forward in imaging at valhall’, *First Break* **28**(4).
- Son, Woohyun, Sukjoon Pyun, Changsoo Shin & Han-Joon Kim 2014, ‘An algorithm adapting encoded simultaneous-source full-waveform inversion to marine-streamer acquisition data’, *Geophysics* **79**(5), R183–R193.
- Symes, William W. 2007, ‘Reverse time migration with optimal checkpointing’, *GEOPHYSICS* **72**(5), SM213–SM221.
- Tarantola, Albert 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics.
- Thorbecke, Jan W. & Deyan Draganov 2011, ‘Finite-difference modeling experiments for seismic interferometry’, *GEOPHYSICS* **76**(6), H1–H18.
- Trinh, P.T., R. Brossier, L. Métivier, J. Virieux & P. Wellington 2017, ‘Bessel smoothing filter for spectral-element mesh’, *Geophysical Journal International* **209**(3), 1489–1512.

- van Leeuwen, T. & W.A. Mulder 2008, ‘Velocity analysis based on data correlation’, *Geophysical Prospecting* **56**(6), 791–803.
- van Leeuwen, Tristan & Felix J. Herrmann 2012, ‘Fast waveform inversion without source-encoding’, *Geophysical Prospecting* **61**, 10–19.
- Virieux, Jean 1986, ‘P-sv wave propagation in heterogeneous media: Velocity-stress finite-difference method’, *GEOPHYSICS* **51**(4), 889–901.
- Virieux, Jean & Stéphane Operto 2009, ‘An overview of full-waveform inversion in exploration geophysics’, *Geophysics* **74**(6), WCC1–WCC26.
- Wang, Yanghua 2015, ‘Frequencies of the Ricker wavelet’, *Geophysics* **81**(2), A31–A37.
- Wang, Yanghua 2017, *Seismic Inversion, Theory and Applications*, Wiley Blackwell, Oxford, UK.
- Wang, Yanghua, ed. 2016, *Seismic Inversion*, John Wiley & Sons, Ltd.
- Wang, Yanghua & Ying Rao 2009, ‘Reflection seismic waveform tomography’, *Journal of Geophysical Research* **114**.
- Warner, Michael, Andrew Ratcliffe, Tenice Nangoo, Joanna Morgan, Adrian Umpleby, Nikhil Shah, Vetle Vinje, Ivan Štekl, Lluís Guasch, Caroline Win, Graham Conroy & Alexandre Bertrand 2013, ‘Anisotropic 3d full-waveform inversion’, *GEOPHYSICS* **78**(2), R59–R80.
- Warner, Michael & Lluís Guasch 2016, ‘Adaptive waveform inversion: Theory’, *GEOPHYSICS* **81**(6), R429–R445.
- Williamson, P. R. 1991, ‘A guide to the limits of resolution imposed by scattering in ray tomography’, *GEOPHYSICS* **56**(2), 202–207.
- Yang, Pengliang 2014, ‘A numerical tour of wave propagation’.
URL: https://yangpl.files.wordpress.com/2019/11/yang_2014_primer.pdf
- Yang, Pengliang, Jinghuai Gao & Baoli Wang 2014, ‘RTM using effective boundary saving: A staggered grid GPU implementation’, *Computers & Geosciences* **68**, 64–72.
- Yang, Pengliang, Romain Brossier & Jean Virieux 2016, ‘Wavefield reconstruction by interpolating significantly decimated boundaries’, *GEOPHYSICS* **81**(5), T197–T209.
- Yilmaz, Özdogan 2001, *Seismic data analysis*, Society of exploration geophysicists Tulsa, OK.

Apêndice A

Produção bibliográfica

Abaixo, esta a lista de referências dos trabalhos nos quais participei desde o início do doutorado. A lista inclui os artigos em revistas científicas, resumos expandidos publicados e, também, artigos em processo de revisão.

Lista atualizada em 30 de setembro de 2021.

A.1 Artigos completos publicados em periódicos

1. [Edwin Fagua Duarte](#), Carlos A N da Costa, João M de Araújo, Yanghua Wang, Ying Rao, *Seismic shot-encoding schemes for waveform inversion*, Journal of Geophysics and Engineering, Volume 17, Issue 5, October 2020, Pages 906–913, DOI: 10.1093/jge/gxaa051.
2. Francisco A Moura, Wagner A Barbosa, [Edwin F Duarte](#), Danyelle P Silva, Mauro S Ferreira, Liacir S Lucena, João M de Araújo, *Patched Green's function method applied to acoustic wave propagation in disordered media: an interdisciplinary approach*, Journal of Geophysics and Engineering, Volume 17, Issue 5, October 2020, Pages 914–922, DOI:10.1093/jge/gxaa052.
3. Danyelle da Silva, [Edwin Fagua Duarte](#), Wagner Almeida, Mauro Ferreira, Francisco Alirio Moura, Joao Medeiros de Araújo, *Target-oriented inversion using the Patched Green function method*, Geophysics, DOI: 10.1190/geo2020-0640.1.

A.2 Artigos submetidos

1. P.T. Carvalho, S.L.E.F. da Silva, [E.F. Duarte](#), R. Brossier, G. Corso, J.M. de Araujo. *Full waveform inversion based on the non-parametric estimate of the probability distribution of the residuals*, Geophysical Journal International.

A.3 Artigos publicados em anais de conferências

1. [Edwin Fagua Duarte](#), Carlos A. N. da Costa, and Joao Medeiros de Araujo, (2019), "Alternative encoding strategies for encoded simultaneous sources full waveform inversion in the time domain,"SEG Technical Program Expanded Abstracts : 1670-1674. DOI: 10.1190/segam2019-3215402.1
2. W. Almeida, [E. Fagua Duarte](#), A. Moura, M. Ferreira and J. Medeiros de Araujo, *Target-Oriented Wave Propagator Using the Patched Green Function Method*, EAGE 2020 Annual Conference & Exhibition Online. DOI: 10.3997/2214-4609.202010764.

Também esta sendo preparado um artigo relacionado aos capítulos 4 e 5, o qual não esta nesta lista.