

ANÁLISE DO PROBLEMA DE SALTO DE CICLO EM INVERSÃO DE
DADOS SÍSMICOS UTILIZANDO A EQUAÇÃO ACÚSTICA DA ONDA E SUA
DEPENDÊNCIA EM RELAÇÃO AO CONTEÚDO DE FREQUÊNCIA

Jorge Luiz de Lima Matias

Dissertação de Mestrado apresentada ao
Programa de Pós-graduação em Engenharia
Civil, COPPE, da Universidade Federal do
Rio de Janeiro, como parte dos requisitos
necessários à obtenção do título de Mestre em
Engenharia Civil.

Orientador: Franciane Conceição Peters

Rio de Janeiro
Outubro de 2016

ANÁLISE DO PROBLEMA DE SALTO DE CICLO EM INVERSÃO DE
DADOS SÍSMICOS UTILIZANDO A EQUAÇÃO ACÚSTICA DA ONDA E SUA
DEPENDÊNCIA EM RELAÇÃO AO CONTEÚDO DE FREQUÊNCIA

Jorge Luiz de Lima Matias

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

Examinada por:

Prof. Webe João Mansur, Ph.D.

Prof. José Antonio Fontes Santiago, D.Sc..

Prof. Carlos Friedrich Loeffler Neto, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
OUTUBRO DE 2016

Matias, Jorge Luiz de Lima

Análise do problema de salto de ciclo em inversão de dados sísmicos utilizando a equação acústica da onda e sua dependência em relação ao conteúdo de frequência/Jorge Luiz de Lima Matias. – Rio de Janeiro: UFRJ/COPPE, 2016.

XIII, 53 p.: il.; 29, 7cm.

Orientador: Franciane Conceição Peters

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Civil, 2016.

Referências Bibliográficas: p. 45 – 47.

1. Full Waveform Inversion. 2. Salto de Ciclo. 3. Cycle Skipping. 4. Equação Completa da Onda. I. Peters, Franciane Conceição. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Civil. III. Título.

À Livia, com muito amor.

Agradecimentos

Gostaria de agradecer a todos que me ajudarem nesse caminho. Primeiramente a minha família, a quem eu amo muito e que sempre me deram apoio em todos os momentos. Sou eternamente grato.

Agradeço a minha namorada, companheira e amiga, Livia Piovesana, por todo o amor, carinho, compreensão e apoio. E também por todas as discussões sobre os mais variados temas técnicos, incluindo o deste trabalho.

Agradeço a minha orientadora, Franciane Conceição Peters, por todo o apoio e ajuda ao longo deste programa de mestrado. Pelo conhecimento, pelo incentivo e pelas discussões elucidativas que ajudaram muito na realização desta dissertação.

Ao professor Webe João Mansur, pelas palavras de incentivo, pelo apoio, e pelos conhecimentos compartilhados que me ajudaram muito nesta caminhada.

A todos do corpo técnico-administrativo, aos professores, e amigos, do Laboratório de Métodos de Modelagem e Geofísica Computacional - LAMEMO, da COPPE/UFRJ.

Aos examinadores da banca por suas importantes correções, revisões e contribuições para a elaboração do documento final.

Ao colega de trabalho e amigo, Frank Agnar Maaø, por todas as discussões valiosas sobre *Full Waveform Inversion*.

E ao consórcio SEISCOPE, cujo software de código aberto TOY2DAC foi utilizado nesta dissertação.

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

ANÁLISE DO PROBLEMA DE SALTO DE CICLO EM INVERSÃO DE
DADOS SÍSMICOS UTILIZANDO A EQUAÇÃO ACÚSTICA DA ONDA E SUA
DEPENDÊNCIA EM RELAÇÃO AO CONTEÚDO DE FREQUÊNCIA

Jorge Luiz de Lima Matias

Outubro/2016

Orientador: Franciane Conceição Peters

Programa: Engenharia Civil

Inversão sísmica da equação completa da onda, mais conhecida na literatura como *FWI* (*Full Waveform Inversion*), é uma metodologia amplamente utilizada para estimar propriedades físicas da Terra a partir de sismogramas, dados provenientes de levantamentos sísmicos. Em resumo, esta metodologia é um problema inverso que busca encontrar um modelo de parâmetros que minimize a função objetivo do problema inverso. Usualmente esta função objetivo é uma norma dos resíduos entre os dados sísmicos medidos e dados modelados a partir de um modelo. Para isto, o algoritmo inicia com um modelo inicial e o modifica a cada iteração, até à convergência do método. Este é um problema mal-posto e não-linear, usualmente resolvido como um problema de otimização local.

A característica oscilatória das ondas sísmicas torna a função objetivo de *FWI* também oscilatória, com a presença de diversos mínimos locais diferentes do mínimo global da solução. Dependendo do modelo inicial, o algoritmo de *FWI* pode convergir erroneamente para um destes mínimos locais em diferentes regiões do modelo, problema conhecido na literatura como Salto de Ciclo, ou *Cycle Skipping*. Este problema será o objeto de estudo desta dissertação, com o objetivo de mostrar a dependência do algoritmo de *FWI* às baixas frequências e de um bom modelo inicial.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

ANALYSIS OF THE CYCLE SKIPPING PROBLEM IN SEISMIC FULL
WAVEFORM INVERSION AND ITS DEPENDENCY ON FREQUENCY
CONTENT

Jorge Luiz de Lima Matias

October/2016

Advisor: Franciane Conceição Peters

Department: Civil Engineering

Full Waveform Inversion, also known as *FWI*, is the seismic inversion of the complete wave equation. It is a methodology broadly used to estimate physical properties of the Earth from shot seismograms of seismic surveys. In short, this method aims to estimate a model of physical properties which minimizes the cost function of the inverse problem. This cost function usually is the residual norm between real seismic measurements and modelled seismic data. The method algorithm begins with an initial model which is modified at each iteration until convergence. That is an ill-posed, non-linear inverse problem which is usually solved with a local optimization algorithm.

The oscillatory characteristic of the seismic waves makes the *FWI* cost function also oscillatory, thus with several points of local minima different from the global solution. Depending on the initial model the FWI solution may converge erroneously into a local minima in some or several areas of the model, and this problem is known as *Cycle Skipping*. This study aims to analyze the Cycle Skipping issue and study how *FWI* results are dependent on seismic low frequency content and a proper initial model.

Sumário

Lista de Figuras	x
Lista de Tabelas	xiii
1 Introdução	1
1.1 Contextualização do trabalho	3
1.2 Estrutura da Dissertação	5
1.3 Revisão Bibliográfica	7
1.4 Objetivos do trabalho	8
2 <i>Full Waveform Inversion</i>	9
2.1 O Problema Direto	9
2.1.1 A equação completa da onda	9
2.1.2 Domínio do tempo e da frequência	11
2.1.3 Solução numérica ao problema direto	13
2.2 Formulação	14
2.3 Função Objetivo	15
2.4 Gradiente - Método Adjunto	17
2.4.1 Método Adjunto	18
2.4.2 Gradiente da onda acústica	21
2.5 Principais algoritmos de otimização local	22
2.5.1 Método da máxima descida	22
2.5.2 Gradiente Conjugado	22
2.5.3 L-BFGS	22
3 O problema do Salto de Ciclo - <i>Cycle Skipping</i>	24
3.1 Cycle Skipping	24
3.2 Estratégia multi-escala	27
3.3 Outras soluções ao <i>Cycle Skipping</i>	29
4 Experimentos e Discussões	31
4.1 Modelo de perturbação gaussiana	31

4.2	Modelo Marmousi	34
4.2.1	O problema	34
4.2.2	Frequências discretas	35
4.2.3	Estratégia multi-escala	37
5	Conclusões	44
	Referências Bibliográficas	45
A	<i>Adaptive Waveform Inversion</i>	48
A.1	Formulação de <i>AWI</i>	50
A.2	Gradiente de <i>AWI</i>	50
A.3	Convergência	51

Lista de Figuras

1.1	Ilustração de um tiro de um levantamento sísmico.	2
1.2	Ilustração da diferença entre os dados sísmicos organizados em sísmograma de tiro (como são medidos, na imagem superior) e os dados organizados em <i>CMP</i> (imagem da esquerda) e <i>CDP</i> (imagem da direita).	5
2.1	Fluxograma básico do método <i>Full Waveform Inversion</i>	15
3.1	Ilustração do problema de Cycle Skipping em um traço. Figura adaptada de VIRIEUX & OPERTO [1]	25
3.2	Exemplo ilustrativo da função objetivo calculada a partir de um traço sísmico medido, e o	26
3.3	Ilustração da divisão de um traço sísmico (ilustrado em preto da figura) em componentes de várias frequências, das mais baixas (mínima 1Hz) até as mais altas (máxima 30Hz). Essa decomposição de frequências é importante em FWI pois o período da onda é maior nos componentes do traço em frequências mais baixas (VIRIEUX & OPERTO [1])	28
3.4	Ilustração da estratégia multi-escala, começando no primeiro estágio com uma frequência mais baixa (imagem do topo), aumentando a frequência até o último estágio (imagem da base). Figura adaptada de FICHTNER[2]	29
4.1	Modelo com uma perturbação de alta velocidade no centro (vermelho). O meio tem velocidade V_p de 1500m/s e a perturbação no centro tem velocidade de 1800m/s.	31
4.2	Resposta do algoritmo de <i>FWI</i> utilizando-se um dado com frequência de 2Hz	32
4.3	Resposta do algoritmo de <i>FWI</i> utilizando-se um dado com frequência de 4Hz	32
4.4	Resposta do algoritmo de <i>FWI</i> utilizando-se dados com frequência discretas de 6Hz, 8Hz, 10Hz e 20Hz	33

4.5	Modelo de velocidade V_p Marmousi. Escala em m/s	34
4.6	Modelo inicial de velocidade V_p , uma suavização do modelo Marmousi da Figura 4.5.	35
4.7	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos de frequência 2Hz.	35
4.8	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos de frequência 4Hz.	36
4.9	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos de frequência 8Hz.	36
4.10	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos de frequência 16Hz.	37
4.11	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos com seis frequências discretas, começando pela mais baixa e sucessivamente aumentando a frequência. Foram utilizadas as frequências 1,5Hz, 2Hz, 3Hz, 4Hz, 6Hz e 8Hz.	38
4.12	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos com seis frequências discretas, começando pela mais baixa e sucessivamente aumentando a frequência. Foram utilizadas as frequências 3Hz, 4Hz, 6Hz, 8Hz, 12Hz e 16Hz.	39
4.13	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos com 4 grupos de frequências descritos na Tabela 4.1.	40
4.14	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos do Grupo 1 de frequências descrito na Tabela 4.1.	40
4.15	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos do Grupo 1 e Grupos 2 de frequências descritos na Tabela 4.1, sucessivamente.	41
4.16	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos do Grupo 1, Grupos 2 e Grupo 3 de frequências descritos na Tabela 4.1, sucessivamente.	41
4.17	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos do Grupo 3, considerando ausentes os dados dos Grupos 1 e Grupo 2 de frequências descritos na Tabela 4.1, sucessivamente.	42
4.18	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos do Grupo 2, considerando ausentes os dados dos Grupos 1 de frequências	42

4.19	Solução do problema inverso fornecido pelo algoritmo de <i>FWI</i> utilizando dados sísmicos do Grupo 2 e 3, sucessivamente, considerando ausentes os dados do Grupos 1 de frequências	43
A.1	Comparação entre a função objetivo da <i>FWI</i> (a) e da <i>AWI</i> (b) Warner & Guasch 2014	50
A.2	Fluxograma básico da metodologia de <i>AWI</i>), alterado de WARNER & GUASCH [3].	53

Lista de Tabelas

4.1	Tabela descrevendo os grupos de frequências discretas utilizados no problema inverso, em uma estratégia multi-escala	39
-----	--	----

Capítulo 1

Introdução

Entre todos os métodos geofísicos, o método sísmico se destaca em seus usos na indústria de exploração e produção de óleo e gás, que estão presentes desde a etapa de estudo geológico regional em subsuperfície, definição estrutural de reservatórios até estudos mais detalhados de delineamento das geometrias e propriedades das rochas e fluidos nos reservatórios. Ele é predominante devido a sua acurácia, alta resolução e grande penetração na subsuperfície. Em TELFORD[4], livro primeiramente escrito em 1976 e cuja mais recente revisão foi em 2003, pode-se encontrar uma ampla revisão de todos os métodos geofísicos, incluindo o objeto de estudo desta tese, os métodos sísmicos.

O método sísmico é em si um problema inverso, onde as propriedades em subsuperfície são estimadas a partir dos tempos de chegada e amplitudes de ondas medidas em receptores após terem se propagado pelo meio em estudo. Em sismologia estas propriedades são as características do interior da Terra, e em Sísmica de Exploração os estudos focam nas estruturas mais rasas. É possível através do método sísmico estimar a localização, direção e inclinação de estruturas geológicas, as variações e contrastes de impedância acústica, além das propriedades elásticas no meio.

As propriedades físicas que são estudadas na sísmica são a velocidade que as ondas elásticas se propagam no meio e a densidade deste meio. Tradicionalmente, consideram-se as propriedades chamadas de impedância acústica e impedância cisalhante, sendo estas o resultado da multiplicação da densidade pela velocidade da onda P e da onda S do meio, respectivamente. Além disso, estas propriedades podem variar de acordo com a direção de propagação, fenômeno conhecido por anisotropia.

As ondas P são as primárias, suas velocidades são maiores comparando-se com as ondas S e elas propagam-se pela Terra de forma longitudinal, ou seja, comprimem e estendem as rochas em direção paralela à propagação da onda. As ondas S ou secundárias são, por sua vez, ondas transversais ou de cisalhamento. As ondas S não se propagam em meios líquidos ou gasosos, sendo, portanto, impossíveis de serem medidas em tradicionais levantamentos sísmicos marinhos feitos com cabos

de receptores sendo puxados por um navio sísmico e uma fonte de compressão de ar.

Em Sísmica de Exploração, o problema das ondas sísmicas é normalmente simplificado como uma propagação de ondas acústicas, desprezando o fator do cisalhamento na modelagem física-matemática do problema principalmente no caso de levantamentos marinhos. A equação da onda acústica será utilizada neste trabalho. A mesma e sua modelagem serão explicados na sequência.

Há diversas técnicas de processamento, interpretação e inversão sísmica para estimar o meio geológico em subsuperfície a partir do dado sísmico, e este trabalho irá focar principalmente no problema inverso conhecido como *Full Waveform Inversion* (*FWI*), método de inversão de dados sísmicos utilizando a equação completa da onda. Nesta metodologia a equação completa da onda é utilizada para simular a propagação da onda mecânica acústica gerada por tiros de uma aquisição sísmica, buscando resolver o problema inverso a partir dos dados de sismogramas de tiro adquiridos no levantamento sísmico.

Na Figura 1.1 está ilustrado um tiro de um levantamento sísmico, onde uma onda propaga-se pelo meio a partir de uma fonte, e ao interagir com o meio esta frente de onda é medida pelos receptores na forma de energia refletida, refratada, entre outras. A partir destes dados medidos em superfície que busca-se estimar as propriedades físicas do meio. Este é o problema inverso que *FWI* e outras metodologias em Sísmica de Exploração buscam resolver.

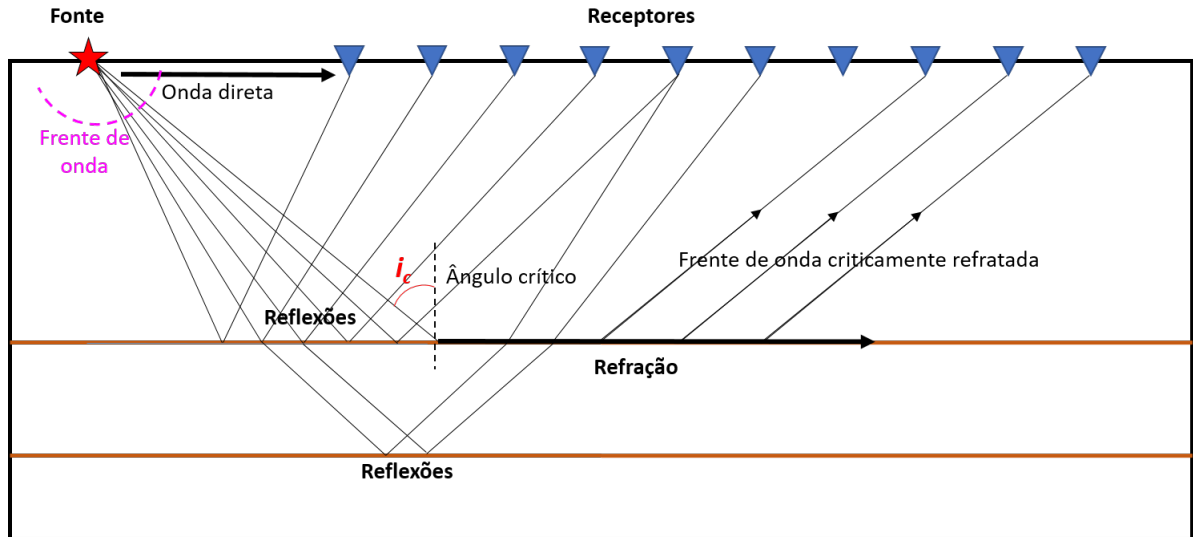


Figura 1.1: Ilustração de um tiro de um levantamento sísmico.

O problema inverso de *FWI* é um problema mal-posto, mal-condicionado, não linear, não há unicidade de solução. Os dados são finitos enquanto o espaço geológico é contínuo, partes da geologia não são identificáveis pelos dados sísmicos, e há pre-

sença de ruídos importantes nos dados. Mesmo o caso acústico e isotrópico, onde a impedância acústica é assumida como dominante e há apenas uma variável para inversão, estas limitações continuam importantes.

Entretanto, *FWI* tem o potencial de estimar uma solução global para os parâmetros físicos em subsuperfície em alta-resolução, ou seja, com um alto nível de detalhamento. *FWI* utilizando a equação acústica da onda, por exemplo, é usualmente utilizada para estimar uma solução global para o campo de velocidades com uma resolução não possível de ser alcançada em outros métodos tradicionalmente utilizados para este propósito, como Análise de Velocidades ou Tomografia sísmica.

Na presente dissertação, as limitações do método de Full Waveform Inversion serão estudadas, principalmente sua dificuldade em chegar a um mínimo global dependendo do modelo inicial.

Neste capítulo introdutório, será apresentada primeiramente uma contextualização do trabalho explanando os principais desafios na metodologia de *Full Waveform Inversion*, assim como as principais metodologias utilizadas para contorná-los. Na sequência serão apresentadas a estrutura da dissertação, uma revisão bibliográfica sobre o tema e os objetivos do trabalho.

1.1 Contextualização do trabalho

Full Waveform Inversion, ou *FWI* é uma tecnologia promissora para a extração de parâmetros físicos diretamente dos dados de sismogramas de tiros. A técnica pode buscar resolver apenas a velocidade da onda P, bem como também parâmetros de anisotropia ou até mesmo parâmetros elásticos (a velocidade da onda P pode ser estimada utilizando a equação completa da onda acústica, onde parâmetros elásticos são ignorados). No entanto, esta tecnologia possui diversos desafios que a impedem de obter resultados em alta resolução, considerando que sua resolução potencial seria de metade do comprimento de onda.

A técnica de *FWI*, explicada em detalhe no Capítulo 2, resume-se na minimização de uma função objetivo, que está relacionada com os resíduos entre os dados medidos e os dados simulados. Este resíduo tem a mesma característica oscilatória dos dados sísmicos, criando, portanto, mínimos locais na função objetivo do problema inverso. Este é um problema não-linear, e os métodos de otimização utilizados para determinar a solução são geralmente do tipo otimização por descida de gradiente. Com este tipo de implementação, a técnica *FWI* encontrará invariavelmente um mínimo local a partir de um modelo inicial fornecido ao algoritmo, e este mínimo local pode eventualmente não coincidir com o mínimo global. Ao fenômeno onde a solução fica presa em um mínimo local diferente do mínimo global dá-se o nome de Salto de Ciclo, encontrado mais frequentemente na literatura como *Cycle Skipping*,

pois a solução encontrada pelo algoritmo diverge em um ou mais ciclos de onda da resposta original.

Muitas técnicas são utilizadas para mitigar o problema dos mínimos locais, sendo a estratégia de multi-escala no domínio da frequência a mais comumente utilizada. Entretanto, mesmo esta tem importantes limitações que a impedem, em alguns casos, de encontrar um resultado satisfatório. Esta técnica depende do sucesso da solução do problema inverso nas primeiras frequências utilizadas, as mais baixas. Se mesmo a solução nas primeiras frequências ficarem presas em um mínimo local, toda a solução ficará comprometida. Logo a estratégia multi-escala depende da qualidade dos dados medidos nas mais baixas frequências, o que nem sempre é possível de se obter. Nos levantamentos sísmicos marinhos, por exemplo, o conteúdo de frequência tem uma banda limitada e está diretamente relacionado com a profundidade dos cabos, onde o posicionamento mais raso prioriza o conteúdo de mais altas frequências, e o posicionamento mais profundo prioriza o conteúdo de mais baixas frequências.

Como o dado sísmico tem notoriamente um conteúdo de frequências limitado, há restrições importantes para a técnica de *FWI*. A solução mais comumente utilizada na indústria é buscar em outras metodologias um modelo de velocidade inicial próximo suficiente da solução final para que a aplicação da técnica de *FWI* seja possível.

A mais comum dessas técnicas utilizadas para determinar um modelo inicial de velocidades é a Tomografia, e para tanto o dado sísmico precisa estar já processado e organizado em famílias CMP (Common Mid Point) ou CDP (*Common Depth Point*), sendo estas famílias uma forma de organizar as reflexões geradas em um mesmo ponto em subsuperfície a partir de fontes em posições distintas e registradas em receptores também em posições distintas. A Figura 1.2 ilustra a diferença entre o sísmograma de tiro e os dados organizados em família *Common mid point* e *Common depth point*.

É importante ressaltar que estas técnicas de processamento sísmico são complexas, e demanda tempo, recursos computacionais e humanos. Portanto, por mais que potencialmente *FWI* possa ser diretamente aplicado utilizando os sísmogramas de tiro, precisa-se esperar todo um projeto de processamento e criação de um modelo de velocidades para de fato ter em mãos um modelo inicial para começar *FWI*, que normalmente só é obtido depois de algumas iterações de tomografia. Atualmente na indústria é comum este processo entre a aquisição e a obtenção de um bom modelo de velocidade de tomografia demorar entre 6 meses e 12 meses, ou mais, dependendo da complexidade da área e qualidade dos dados.

Neste contexto, o problema de Salto de Ciclo (*Cycle Skipping*) ainda é um dos maiores limitadores para a aplicação de *FWI*. Esta é a principal questão a ser estudada nesta dissertação, mostrando o impacto de uma estratégia utilizando vários

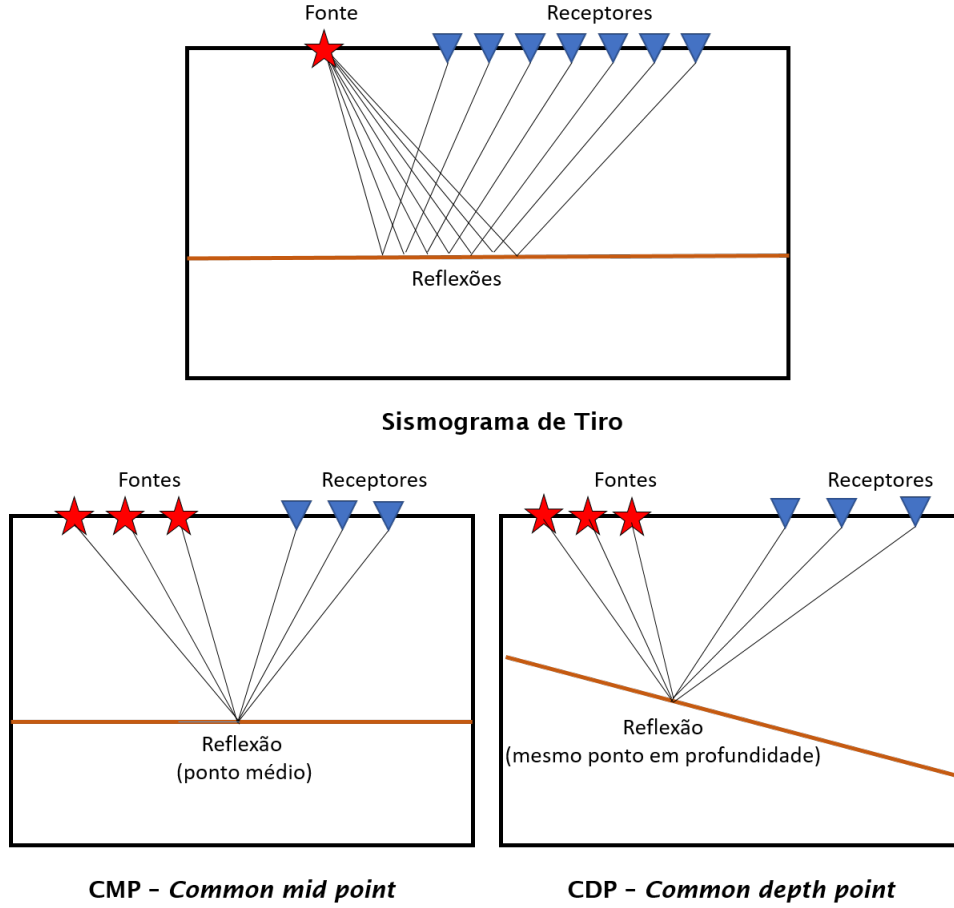


Figura 1.2: Ilustração da diferença entre os dados sísmicos organizados em sismo-grama de tiro (como são medidos, na imagem superior) e os dados organizados em *CMP* (imagem da esquerda) e *CDP* (imagem da direita).

intervalos de frequência na solução do problema e a influência da presença de baixas frequências para o resultado final.

É importante mencionar que outras metodologias mais robustas em relação a mínimos locais estão em desenvolvimento, como *Adaptive Waveform Inversion*, uma nova metodologia desenvolvida por WARNER & GUASCH [3] que está descrita no Apêndice B dessa dissertação.

1.2 Estrutura da Dissertação

Após o contexto do trabalho ser descrito na sessão anterior, na próxima seção desta introdução (1.3) será apresentado um conciso levantamento da literatura sobre o tema. Em seguida, na seção 1.4, serão explanados todos os objetivos desta dissertação. No Capítulo 2 a metodologia de *Full Waveform Inversion (FWI)* será descrita em conjunto com suas implementações mais comumente usadas. Este capítulo tem por objetivo introduzir todos os principais termos sobre o tema. Serão expla-

nados a formulação e o algoritmo da metodologia, a função objetivo do problema inverso, assim como o cálculo do seu gradiente que será utilizado na solução do problema inverso e os principais algoritmos de otimização local.

Além disso, serão descritas no Capítulo 2 as principais diferenças na implementação deste problema inverso no domínio do tempo e no domínio da frequência. A decomposição do dado sísmico no domínio da frequência é comumente utilizado em métodos sísmicos, embora o dado sísmico seja registrado no domínio do tempo e usualmente visualizado e analisado neste domínio. A implementação do método *FWI* em cada domínio trás importantes vantagens e desvantagens para a solução do problema e para o custo computacional que devem ser consideradas dependendo do tamanho e complexidade do problema.

Neste capítulo também há uma seção reservada ao tema da propagação de ondas sísmicas incluindo a descrição da sua modelagem como foi utilizada nesta dissertação. *FWI* é uma técnica que busca resolver um problema inverso da equação completa da onda, que precisa em sua implementação utilizar algum tipo de modelagem desta equação diferencial, sendo que neste trabalho foi utilizado uma simples modelagem por diferenças finitas da equação da onda acústica em um ambiente isotrópico.

No próximo capítulo, Capítulo 3, será explanado o problema de Salto de Ciclo, principal tema desta dissertação. Este capítulo tem por objetivo descrever este problema prático da metodologia de *FWI* implementada como um problema de otimização local, e sua dependência em relação ao conteúdo de frequência do dado sísmico. A estratégia multi-escala, a mais utilizada técnica para mitigar o problema de Salto de Ciclo em *FWI*, também será descrita neste capítulo.

Na sequência, Capítulo 4, serão expostos todos os experimentos, resultados e discussões neste trabalho. Estes são experimentos utilizando todas as técnicas aqui descritas, e ao longo do texto estarão presente discussões relacionadas com cada experimento. Serão explorados principalmente o uso das estratégias multi-escala e a dependência dos resultados de *FWI* em relação a presença de baixas frequências nos dados sísmicos.

Finalmente, no Capítulo 5 serão expostas as principais conclusões neste trabalho e discussões mais gerais sobre as metodologias em estudos, além de sugestões de trabalhos futuros.

E por fim, após as conclusões do Capítulo 5, estão dispostas toda as referências bibliográficas deste trabalho.

No Apêndice A está descrito uma técnica chamada *AWI - Adaptive Waveform Inversion*, que é relevante para a discussão desta dissertação pois busca uma solução ao problema de Salto de Ciclo modificando a função objetivo do problema inverso de *FWI*. Com uma nova função objetivo, a técnica busca não produzir mínimos locais, e portanto eliminar o problema de Salto de Ciclo.

1.3 Revisão Bibliográfica

O problema da inversão da onda sísmica utilizando a equação completa da onda já vem sendo estudado e desenvolvido desde a década de 80, quando um dos primeiros trabalhos desenvolvendo a formulação clássica de *Full Waveform Inversion*, *FWI*, foi apresentado por TARANTOLA[5] em 1984. Outros trabalhos da mesma época buscavam solucionar o problema inverso dos dados sísmicos, como por exemplo o trabalho de LAILLY[6] de 1983 sobre o problema inverso da sísmica como uma sequência de migrações pré estaqueamento, entre outros.

PRATT[7], em 1999, apresentou a formulação de *FWI* no domínio da frequência, utilizando equação da onda acústica, que ainda hoje é uma das mais utilizadas para resolver o problema inverso de dados sísmicos de sismogramas de tiros. A implementação no domínio da frequência possui diversas vantagens em comparação com a metodologia original no domínio do tempo proposta por TARANTOLA[5], como por exemplo a estratégia multiescala que será discutida em detalhes nesta dissertação. Estes desenvolvimentos juntos com as ideias de BUNKS[8] de uma inversão sísmica multi-escala, de 1995, são a origem dos métodos atuais de solução do problema inverso no domínio da frequência em multi-escala com diversos grupos de frequências, das mais baixas às mais altas.

SIRGUE & PRATT[9] completam os estudos para a inversão no domínio da frequência apresentando uma estratégia para as escolhas das frequências a serem utilizadas na inversão em seu trabalho de 2004. Fica demonstrado como um dado sísmico pode ser adequadamente representado por poucas frequências discretas no problema inverso. Este ponto é importante para a aplicação das estratégias conhecidas como multi-escala, onde busca-se minimizar os problemas com Salto de Ciclo (*Cycle Skipping*). Os problemas com Salto de Ciclo serão detalhados no Capítulo 3.

PLESSIX[10] em seu trabalho de 2006 apresenta uma abrangente revisão do método adjunto (adjoint state method) para computação do gradiente de um funcional, explanando a metodologia aplicável às técnicas geofísicas como o *FWI*. O método adjunto é a forma mais eficiente e a mais utilizada para o cálculo do gradiente do problema, e é esta a metodologia utilizada também nessa dissertação.

VIRIEUX & OPERTO[1] apresentam em 2009 uma abrangente revisão bibliográfica sobre *FWI*. Esta é uma referência importante que explica o estado-da-arte das soluções de *FWI*, e também os principais desafios remanescentes da metodologia.

FICHTNER[2], em seu livro de 2011, apresenta uma ainda mais ampla abordagem sobre a metodologia *FWI*. Nesta referência é possível encontrar um apanhado completo dos tópicos relevantes em modelagem sísmica da equação completa da onda e seu problema inverso.

WARNER[11], 2013, apresenta uma abrangente revisão sobre a metodologia *FWI*

com especial ênfase a suas aplicações 3D, e considerando anisotropia no problema inverso.

Recentes desenvolvimentos buscam formas alternativas à estratégia multi-escala na frequência para suplantare o problema de Salto de Ciclo. Alguns destes desenvolvimentos vão na direção de outras formas de função objetivo, como o método de GUASCH & WARNER [3], *AWI - Adaptive Waveform Inversion*, apresentado em 2014. A busca por uma solução definitiva ao problema de Salto de Ciclo ainda é um tema de pesquisa promissor e sem solução simples.

1.4 Objetivos do trabalho

O principal objetivo é uma análise abrangente sobre o problema dos funcionais oscilatórios e seus mínimos locais e seu impacto na solução do problema inverso e na aplicabilidade da metodologia de *FWI* em casos reais.

Serão objetos de estudo nesta dissertação o impacto do conteúdo de frequência do dado sísmico na solução do problema inverso, a estratégia multi-escala no domínio da frequência para evitar problemas com Salto de Ciclo ou *Cycle Skipping* e a dependência dos resultados a um modelo inicial adequado.

A partir de análise de dados sintéticos, serão avaliadas situações realísticas como ausência de um conteúdo amplo de baixas frequências ou um pobre modelo inicial.

Capítulo 2

Full Waveform Inversion

Na primeira seção deste capítulo será discutido o problema direto de *FWI*, ou seja, a equação da onda e sua solução numérica. Esta é a equação governante do problema. Nas seções posteriores será discutido o problema inverso de *FWI*, desde a formulação do problema e sua função objetivo, até mais detalhes sobre o cálculo do gradiente desta função objetivo.

2.1 O Problema Direto

2.1.1 A equação completa da onda

O problema direto da equação completa da onda pode ser resolvido tanto no domínio do tempo quanto da frequência. A equação completa da onda elástica e acústica pode ser descrita de forma compacta como na Equação 2.1 (VIRIEUX et. al. [12], BROSSIER & ETIENNE [13] e ETIENNE et. al. [14]). Este é um sistema geral montado pela discretização por diferenças-finitas ou elementos-finitos, assim como descrito em MARFURT [15]:

$$\mathbf{M}(\mathbf{x}) \frac{d^2 u(\mathbf{x}, t)}{dt^2} = \mathbf{S}(\mathbf{x}) u(\mathbf{x}, t) + f(\mathbf{x}, t) \quad (2.1)$$

onde $\mathbf{M}$ e $\mathbf{S}$ são as representações matriciais de massa e de rigidez, respectivamente, u representa o campo de onda sísmico e f a fonte da onda sísmica. O tempo é representado por t e a posição no espaço é representado por $\mathbf{x}$, sendo $\mathbf{x}$ multi-dimensional, dependendo do problema que se quer resolver (uma, duas ou três dimensões no espaço). O campo de onda sísmico elástico representa os deslocamentos verticais e horizontais da partícula. No caso da simplificação para onda acústica, usualmente u representa pressão.

A representação matricial discretizada da Equação 2.1 é usualmente utilizada na literatura para ilustrar de forma generalizada o caso acústico e elástico no domínio

do tempo. Esta forma também ajuda a comparar a solução da onda no domínio do tempo com a solução no domínio da frequência, o que será discutido posteriormente na Seção 2.1.2.

O problema da propagação da onda pode ser estudado de forma crescentemente complexa, começando pela equação da propagação da onda acústica em um meio homogêneo, considerando a velocidade de propagação da onda e a densidade constantes:

$$\nabla^2 u(\mathbf{x}, t) - \frac{1}{v^2} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = 0 \quad (2.2)$$

onde u representa o campo de pressão da onda acústica, v representa a velocidade da onda neste meio homogêneo.

Para o caso da propagação de uma onda acústica em um meio heterogêneo e isotrópico, onde a velocidade de propagação da onda acústica e a densidade podem variar no espaço mas a velocidade da onda não varia de acordo com a direção de propagação, a equação da onda fica da seguinte forma:

$$\frac{1}{\rho(\mathbf{x})V_p^2(\mathbf{x})} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} - \nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \vec{\nabla} u(\mathbf{x}, t) \right) = f(\mathbf{x}, t) \quad (2.3)$$

Onde foi também incluído as perturbações externas no campo de pressão $f(\mathbf{x}, t)$, que no caso de um levantamento sísmico representa uma fonte sísmica, e $\rho(\mathbf{x})$ e $V_p^2(\mathbf{x})$ variam no meio.

O problema da propagação da onda pode ser descrito de forma ainda mais completa. Falta à Equação 2.3 a inclusão de anisotropia, efeitos de absorção da energia em subsuperfície (caso visco-acústico), e principalmente a inclusão de ondas cisalhantes. O caso acústico compreende apenas alterações no campo de pressão por ondas compressoriais, mas em um meio rochoso sólido, que é o objeto de estudo do problema de *FWI* em Sísmica de Exploração, também há a propagação de ondas cisalhantes a partir da fonte (no caso de aquisições sísmicas terrestres) e a conversão de parte da energia de ondas compressoriais em ondas cisalhantes nas heterogeneidades do meio. Por exemplo, o operador do problema do campo de onda $L(\mathbf{u})$ para a deformação elástica plana, na direção de propagação, é definido por:

$$\mathbf{L}(u) = \nabla [(\lambda + \mu) \nabla \cdot u] + \nabla \cdot (\mu \nabla u) - \frac{\partial}{\partial t} \left(\rho \frac{\partial u}{\partial t} \right) \quad (2.4)$$

onde λ representa o módulo de incompressibilidade e μ o módulo de cisalhamento.

Todas essas possibilidades de descrever o problema direto da equação da onda, dependendo da complexidade do fenômeno objeto de estudo, estão representadas de forma simplificada pela equação da onda discretizada (2.1).

2.1.2 Domínio do tempo e da frequência

O dado sísmico é medido no domínio do tempo. Entretanto, é comum passar este dado para o domínio da frequência em diversas aplicações, com o objetivo de dividir o traço sísmico em harmônicos. Para isso, utiliza-se a Transformada de Fourier, que é definida pela seguinte convenção

$$f(\omega) = \int_{-\infty}^{+\infty} f(t)e^{i\omega t} dt. \quad (2.5)$$

A equação da onda no domínio do tempo, representada pela Equação 2.1, é resolvida como um problema de evolução no tempo, ou marcha no tempo, onde a partir de um tempo inicial o resultado da equação nos tempos posteriores são calculadas sequencialmente. Este processo é calculado para cada tiro do levantamento sísmico, representado por $f(\mathbf{x}, t)$ na Equação 2.1. Esta implementação do problema tem algumas vantagens como sua versatilidade de resolver uma grande variedade de aplicações, tamanho de problema, parametrização, e pré-processamento de dados.

Entretanto, como cada tiro deve ser calculado separadamente no domínio do tempo, o custo computacional é proporcional ao número de tiros. Isto pode demandar um grande esforço computacional para um levantamento sísmico com uma densidade de tiros elevada. Mesmo que o tempo computacional possa ser mitigado paralelizando a modelagem de cada tiro em diferentes processadores, isto exige o uso de um grande cluster de alta performance.

A representação compacta da equação da onda discretizada no espaço e no tempo dada pela Equação 2.1 ajuda a enfatizar as diferenças entre a solução da onda no domínio do tempo e no domínio da frequência. A mesma equação no domínio da frequência pode ser representada como na Equação 2.6, onde B é conhecida por Matriz de Impedância (VIRIEUX et. al. [12], BROSSIER & ETIENNE [13], MARFURT [15]).

$$\mathbf{B}(\mathbf{x}, \omega)u(\mathbf{x}, \omega) = F(\mathbf{x}, \omega) \quad (2.6)$$

Na Equação 2.6, ω representa a frequência. Percebe-se que esta equação é muito mais simplificada que sua versão no domínio do tempo. A equação da onda no domínio da frequência pode ser resolvida como um sistema de equações do tipo $\mathbf{Ax} = \mathbf{b}$, para cada frequência. Ou seja, no caso do domínio da frequência o problema pode ser resolvido de forma eficiente com uma eliminação gaussiana, ou Decomposição LU. E uma vez que a matriz B da Equação 2.6 for decomposta, torna-se simples resolver o problema para múltiplas fontes (múltiplos lados direitos da equação) de forma bastante eficiente na fase de substituição da solução do sistema. Basta, para isto, calcular a solução por substituição para cada fonte, utilizando a

mesma decomposição LU já calculada.

Portanto, a solução do problema no domínio da frequência se difere muito daquela no domínio do tempo. O custo computacional do problema não mais depende do número de tiros, uma vez que a solução para cada fonte é computacionalmente eficiente, e a operação mais dispendiosa computacionalmente é a decomposição LU para cada frequência. Portanto, o problema passa a ser proporcional ao número de frequências utilizadas e do tamanho do problema.

O principal problema do método na frequência é a demanda de memória para armazenar e resolver a decomposição LU e a escalabilidade da fatoração LU em casos de larga escala. Em levantamentos sísmicos 3D de grande escala, a implementação do problema no domínio do tempo pode ser mais flexível e acessível do que no domínio da frequência. Recentes desenvolvimentos na solução deste tipo de sistema já tem mostrado avanços na utilização de *FWI* no domínio da frequência em casos visco-acústicos em levantamento sísmico 3D (AMESTOY et. al. [18]).

Nos experimentos desta dissertação (Capítulo 4), foi utilizada a equação da onda acústica isotrópica para a modelagem direta dos dados e para a solução do problema de *FWI*. O caso acústico isotrópico no domínio do tempo é dado por

$$\frac{1}{\rho(\mathbf{x})V_p^2(\mathbf{x})} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} - \nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \vec{\nabla} u(\mathbf{x}, t) \right) = f(\mathbf{x}, t). \quad (2.7)$$

E a equação da onda acústica isotrópica no domínio da frequência é dada por

$$\omega^2 u(\mathbf{x}, \omega) + \rho(\mathbf{x})V_p^2(\mathbf{x}) \operatorname{div} \left(\frac{1}{\rho(\mathbf{x})} \nabla u(\mathbf{x}, \omega) \right) = F(\mathbf{x}, \omega). \quad (2.8)$$

Devido a redundância no espectro de número de onda, ou ângulos de difração, é possível resolver o problema de *FWI* utilizando um número reduzido de frequências. Cada frequência tem uma contribuição limitada e de banda finita no espectro de número de onda, e a estratégia para imagear propriamente um alvo em uma ampla extensão de banda de número de ondas é escolher as frequências tal que a frequência maior (posterior) tenha o número de onda mínimo igual ao máximo número de onda da frequência imediatamente menor (anterior). Desta forma o incremento das frequências não é constante e aumenta linearmente com a frequência (SIRGUE & PRATT [9]). O número de onda vertical máximo e mínimo de uma determinada frequência depende da geometria do levantamento sísmico (máximo e mínimo offset).

Concluindo, o problema de *FWI* pode ser resolvido tanto no domínio do tempo quanto no domínio da frequência. Há vantagens e desvantagens computacionais em ambos os métodos, dependendo do tamanho e complexidade do problema. A escolha do domínio dependerá da aplicação do problema. Como será mostrado no Capítulo 3, a implementação no domínio da frequência tornou-se popular na literatura não apenas pela facilidade de solução por um sistema de equações, mas também por

possibilitar uma estratégia multi-escala. Esta estratégia multi-escala é uma das principais formas de evitar problemas com *Cycle Skipping*.

2.1.3 Solução numérica ao problema direto

Modelar as equações de onda num meio contínuo exigem soluções para a equação diferencial parcial (EDP) da onda. Como o meio geológico onde busca-se a solução é altamente heterogêneo, onde os parâmetros elásticos da subsuperfície podem variar em diversas escalas, as interações entre as ondas e o meio tendem a ser complexas, complicando o uso de métodos numéricos de aproximações analíticas para a EDP.

Para simular a onda num meio complexo é, portanto, necessária a aplicação de soluções numéricas das EDPs de forma completa, ou de sua equação integral correspondente. Estas soluções numéricas tem limitações intrínsecas relacionadas a discretização do meio (espaço) e do tempo que devem ser balanceadas de acordo com os objetivos do imageamento.

Há diversas técnicas para a modelagem direta na literatura. Entre as mais utilizadas para a equação da onda, estão:

- método de diferenças finitas (VIRIEUX [24], LEVANDER [25]),
- métodos de elementos finitos (MARFURT [15], MIN et. al. [26]),
- método de volumes finitos (BROSSIER et. al.[27]),
- métodos pseudo-espectrais (DANECEK & SERIANI [28]),
- entre outros.

Para resolver o problema direto da Equação 2.8 foi utilizado o método de solução numérica de diferenças finitas. Existem várias formulações possíveis para o Método de Diferenças Finitas (MDF), e no caso desta dissertação utilizou-se uma estratégia de discretização de quarta ordem, com um Estêncil otimizado de 13 pontos numa malha intercalada. Para mais detalhes dessa solução numérica, verificar HUSTEDT et. al [29].

A solução numérica utilizada aqui está implementada no código aberto TOY2DAC criado pelo consórcio SEISCOPE, para mais detalhes ver em OPERTO et. al [30].

A solução da equação da onda no domínio da frequência assemelha-se a uma solução de um sistema linear esparsa de larga escala. Para resolver este sistema foi utilizado o solver conhecido como MUMPS, que implementa uma decomposição LU multi-frontal. Para resolver o problema direto da Equação 2.8 apenas uma decomposição LU é necessária para uma determinada frequência, sendo que o restante do sistema é resolvido por um simples processo de retro-substituição.

2.2 Formulação

Full Waveform Inversion é o método de problema inverso que busca estimar parâmetros físicos de subsuperfície a partir diretamente dos sismogramas de tiro. Para isto o método busca um modelo de parâmetros tal que os dados sísmicos simulados utilizando este modelo sejam igual ou o mais próximo possível dos dados reais medidos no levantamento geofísico. Para tal, o método parte do princípio que é possível simular os dados sísmicos a partir da equação completa da onda.

Em linhas gerais, o método de *FWI* exige um modelo inicial de parâmetros físicos, os mesmo parâmetros aos quais busca-se estimar uma solução. Estes parâmetros físicos dependem do problema que se está buscando resolver:

- Onda Acústica Isotrópica: busca-se identificar um único parâmetro, a velocidade da onda p (V_p).
- Onda Elástica Isotrópica: busca-se resolver três parâmetros, velocidade da onda P (V_p), da onda S (V_s) e a densidade (ρ).
- Onda Elástica Anisotrópica: o número de parâmetros depende do modelo de anisotropia a ser utilizado. Pela Lei de Hooke, um caso elástico anisotrópico pode ter até 21 constantes elásticas além da densidade. Entretanto, normalmente a anisotropia sísmica é simplificada utilizando os parâmetros de Thomsen ([16]).

A partir deste modelo inicial de parâmetros elásticos, o algoritmo de *FWI* irá atualizar o modelo de tal forma a minimizar a função objetivo. Ao buscar-se a convergência da função objetivo para o mínimo, a metodologia estará fazendo uma estimativa dos parâmetros elásticos reais do meio onde os dados sísmicos foram medidos. Como a relação entre estes parâmetros elásticos e os dados sísmicos regidos pela equação da onda não é linear, o problema inverso de *FWI* também não é linear.

O problema não-linear de *FWI* é usualmente resolvido como um algoritmo de otimização local de forma iterativa, onde uma direção de busca é calculada para cada posição do modelo em cada iteração, visando minimizar a função objetivo nesta posição. Esta direção é baseada no inverso do gradiente em relação ao modelo, que aponta para a direção de maior crescimento da função objetivo. Um esquema simplificado do algoritmo de *FWI* está exposto na Figura 2.2, com uma visão geral das principais etapas da metodologia. O algoritmo é baseado em atualização do modelo na direção de busca calculada até que o mínimo da função objetivo é atingido, ou seja, o modelo obtido é tal que os dados simulados com este modelo são similares o suficiente dos dados sísmicos medidos no levantamento sísmico (VIRIEUX & OPERTO [1]).

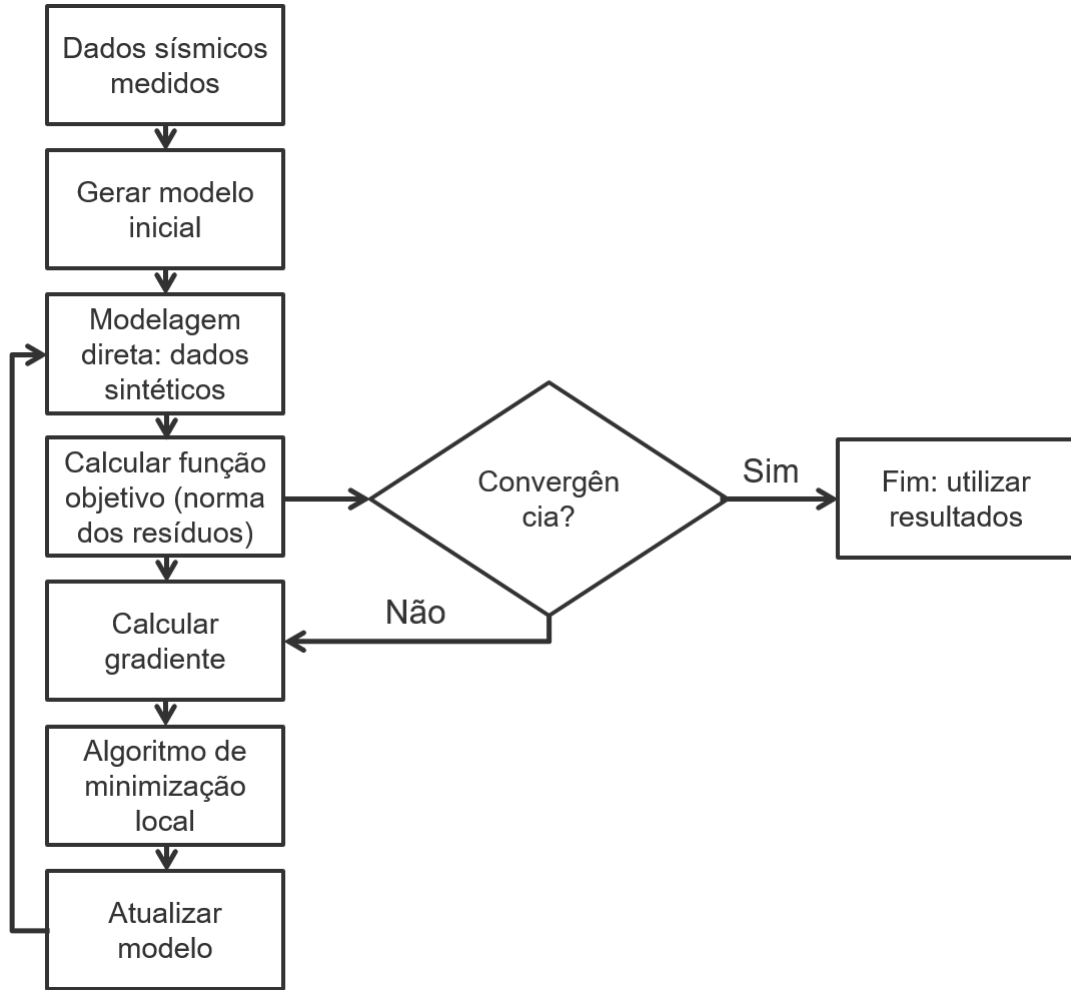


Figura 2.1: Fluxograma básico do método *Full Waveform Inversion*

2.3 Função Objetivo

Em Full Waveform Inversion a função objetivo tem por base os resíduos calculados entre dados sísmicos simulados e os dados sísmicos medidos, representados pela equação

$$\Delta d = d - p(\mathbf{m}). \quad (2.9)$$

Onde Δd são os resíduos, d são os dados sísmicos medidos e $p(\mathbf{m})$ representa os dados modelados (por exemplo pelo MDF brevemente descrito na Seção 2.1.3) utilizando o modelo de propriedades físicas $\mathbf{m}$. Lembrando que estes resíduos não são contínuos, os dados modelados estão limitados por uma discretização no tempo e no espaço, e os dados medidos estão limitados pela amostragem do equipamento utilizado no levantamento sísmico. No caso da onda acústica, isotrópica, este modelo $\mathbf{m}$ é formado pelo modelo de velocidades da Onda P (V_p). Em casos mais completos, são adicionados parâmetros a $\mathbf{m}$, e.g. parâmetros de anisotropia, parâmetros de

absorção (visco-acústico), parâmetros elásticos (densidade, velocidade da onda S).

Os dados observados nos receptores podem ser a diferença de pressão ao longo do tempo medidos em hidrofones, ou a movimentação/aceleração observada por um geofone. Os hidrofones podem ser descritos brevemente como sensores utilizados usualmente em aquisições marinhas combinados em cabos emersos na água, que são sensíveis a mudanças de pressão ao seu redor. Os hidrofones tem apenas a capacidade de medir a diferença de pressão na água, sem reconhecer a direção de propagação da onda, e devido a isso são classificados como um tipo de sensor escalar. O geofone por sua vez é o sensor que tem sensibilidade para detectar movimento. Ele é usualmente utilizado acoplado a superfície em levantamentos sísmicos terrestres ou no fundo do mar, e detecta a movimentação do solo ao qual está acoplado (vibrações). O geofone tem a capacidade de medir a magnitude desta vibração e a sua direção, ou seja, é um sensor vetorial.

O dado gravado por cada receptor é tratado como um traço sísmico, e a função objetivo χ mais comumente utilizada em FWI é o somatório da Norma dos resíduos de todos os traços, dado por

$$\chi(\mathbf{m}) = \|d - p(\mathbf{m})\|^2 = \sum_{s,r} \int_0^T (p(t, \mathbf{m}) - d(t))^2 dt. \quad (2.10)$$

Na Equação 2.10 foi utilizada a Norma 2 dos resíduos, s representa as fontes e r os receptores. Outros estudos mostram alternativas à norma 2 dos resíduos, como por exemplo a norma de Huber apresentada em [17]. Entretanto, nesta dissertação apenas a norma 2 será utilizada por ser a mais comum na literatura em problemas de *FWI*.

Devido a natureza oscilatória dos dados sísmicos, a subtração de dois traços oscilatórios também será oscilatório a menos que estes traços sejam idênticos. Esta função objetivo χ também é não-linear e de comportamento oscilatório, portanto é uma função objetivo com diversos pontos de mínimo além do mínimo local.

Além disso, o problema de *FWI* é mal posto em relação ao seu domínio. O meio geológico no qual levantamentos sísmicos são realizados é contínuo; entretanto, os dados sísmicos adquiridos são apenas correspondentes às posições dos receptores. Com isto, o problema inverso posto tenta estimar os parâmetros físicos elásticos de um meio contínuo com uma amostragem limitada de dados. Logo, o problema de *FWI* é limitado pela geometria de aquisição do dado sísmico. Podemos escrever então que o método de *FWI* tenta resolver a Equação 2.11, onde m é o modelo dos parâmetros elásticos que busca-se estimar, χ é a função objetivo, u_s é o campo de onda calculado como solução do problema direto para a fonte f_s , N_s representa o número de fontes, R_s representa o operador de restrição de u_s para a posição dos receptores no meio, e d_s são os dados reais registrados no levantamento sísmico

(sismograma).

$$\min_{\mathbf{m}} \chi(\mathbf{m}) = \frac{1}{2} \sum_{s=1}^{N_s} \|R_s u_s(\mathbf{m}) - d_s\|^2 \quad (2.11)$$

2.4 Gradiente - Método Adjunto

O tamanho dos problemas de levantamentos sísmicos 2D e 3D impedem a utilização de métodos de otimização global para a solução do método de *FWI*. Devido a isto, para otimizar a função objetivo se faz necessária a aplicação de algoritmos de otimização local, ou seja, métodos baseados em uma estimativa de gradiente da função objetivo dado um modelo. *FWI* é um problema não linear e mal posto, então várias iterações são necessárias para atingir a convergência da função objetivo, e com isto uma estimativa dos parâmetros que deseja-se obter. Em cada iteração a função objetivo é analisada com o modelo de parâmetros obtido na iteração anterior (ou o modelo inicial no caso da primeira iteração); calcula-se o gradiente da função objetivo e ambos, função objetivo e gradiente, são fornecidos para um algoritmo de otimização local que irá calcular mudanças no modelo de parâmetros. Após atualizar o modelo, começa uma nova iteração. Este fluxograma está representado na Figura 2.2, e essa seção irá focar na estimativa do gradiente da função objetivo.

Para os algoritmos de minimização, que serão descritos na próxima seção, é necessário o cálculo do gradiente da função objetivo em relação aos parâmetros para um dado modelo. O método mais utilizado em *FWI* para a estimativa do gradiente é o Método Adjunto (BUNKS et. al. [8], PLESSIX [10]).

O gradiente da função objetivo é definido como

$$(\nabla_m \chi(\mathbf{m}))_i = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [\chi(\mathbf{m} + \varepsilon \mathbf{m}_i) - \chi(\mathbf{m})]. \quad (2.12)$$

onde $\mathbf{m}$ é o modelo de parâmetros, χ é a função objetivo, e $(\nabla_m \chi(\mathbf{m}))_i$ é o gradiente calculado na posição i do modelo. Aqui foi usado a notação ∇_m para indicar o gradiente em relação ao modelo de parâmetros $\mathbf{m}$.

Entretanto, calcular o gradiente utilizando esta formulação seria computacionalmente proibitivo, pois exige a avaliação da função objetivo para cada ponto do modelo de parâmetros.

O Método Adjunto, ou também conhecido na literatura como Método do Estado Adjunto, permite o cálculo do gradiente de $\chi(m)$ de uma forma mais eficiente. Isto tanto para a formulação no domínio do tempo quanto no domínio da frequência.

2.4.1 Método Adjunto

O gradiente da função objetivo de FWI pode ser calculado em relação ao modelo m ($\nabla_m \chi(m)$) utilizando o Método Adjunto. Este método busca a solução do gradiente através da utilização de Operadores Adjuntos.

Considerando-se um produto interno dado por

$$\langle \mathbf{a}, \mathbf{b} \rangle = \langle \overline{\mathbf{a}}, \mathbf{b} \rangle \quad (2.13)$$

Sendo $\mathbf{a}$ e $\mathbf{b}$ vetores pertencentes a um espaço de Hilbert ($\mathbf{a}, \mathbf{b} \in H$). Um Operador Adjunto de um operador linear $\mathbf{A}$ é aquele que respeita a relação

$$\langle \mathbf{a}, \mathbf{A}\mathbf{b} \rangle = \langle \mathbf{A}^\dagger \mathbf{a}, \mathbf{b} \rangle \quad (2.14)$$

Portanto, definindo-se o operador de modelagem do problema direto da equação da onda como $\mathbf{L}$, que atua em um conjunto de campos u que dependem do modelo de parâmetros $\mathbf{m}$, e cujas fontes são representadas por f , por

$$\mathbf{L}(u, \mathbf{m}) = f \quad (2.15)$$

Seu operador adjunto $\mathbf{L}^\dagger$ deve satisfazer a seguinte equação:

$$\langle \mathbf{a}, \mathbf{L}\mathbf{b} \rangle = \langle \mathbf{L}^\dagger \mathbf{a}, \mathbf{b} \rangle \quad (2.16)$$

O primeiro passo para obter o gradiente pelo método adjunto é utilizar a regra da cadeia no gradiente da função objetivo em relação ao modelo. Primeiramente, a notação simplificada da função objetivo utilizada aqui como $\chi(\mathbf{m})$ subentende que a função objetivo depende do campo de onda u , e este por sua vez depende do modelo m , ou seja, a função objetivo é dada por

$$\chi(\mathbf{m}) = \chi(u(\mathbf{m})) \quad (2.17)$$

E a partir desta relação, a derivada da função objetivo pelo modelo de parâmetros poder ser determinada através da seguinte regra da cadeia

$$\frac{\partial}{\partial m} = \frac{\partial}{\partial u} \frac{\partial u}{\partial m}. \quad (2.18)$$

Seguindo esta regra da cadeia, o gradiente da função objetivo em relação ao modelo é dado por

$$\nabla_{\mathbf{m}} \chi(\mathbf{m}) = \nabla_u \chi(\mathbf{m}) \nabla_{\mathbf{m}} u. \quad (2.19)$$

A escolha feita nos experimentos desta dissertação para a função objetivo foi uma

norma L_2 dos resíduos entre dados modelados e medidos, sendo esta uma escolha bastante popular em problemas de FWI. Entretanto, para tornar a formulação do método adjunto genérica, pode-se denotar a função objetivo como

$$\chi(\mathbf{m}) = \int_T \int_G d^3\mathbf{x} \chi_1[u(\mathbf{m}; \mathbf{x}, t)] = \langle \chi_1(\mathbf{m}) \rangle \quad (2.20)$$

onde χ_1 é uma função objetivo a ser especificada, e $\langle \cdot \rangle$ é uma representação para a integração no tempo T e no espaço G .

A partir da relação da Equação 2.19, incluindo o termo $\delta\mathbf{m}$ em ambos os lados, obtém-se a relação

$$\nabla_{\mathbf{m}}\chi(\mathbf{m})\delta\mathbf{m} = \nabla_u\chi(\mathbf{m})\nabla_{\mathbf{m}}u\delta\mathbf{m} = \nabla_u\chi(\mathbf{m})\delta\mathbf{u} \quad (2.21)$$

Onde $\delta\mathbf{u} = \nabla_{\mathbf{m}}u\delta\mathbf{m}$ é a derivada de u com relação a $\mathbf{m}$ na direção de $\delta\mathbf{m}$. $\nabla_{\mathbf{u}}\chi(\mathbf{m})$ representa o gradiente da função objetivo em relação ao campo de onda u . Desenvolvendo a Equação 2.21, substituindo $\chi(\mathbf{m})$ pela Equação 2.20, chega-se em

$$\nabla_{\mathbf{m}}\chi(\mathbf{m})\delta\mathbf{m} = \int_T \int_G d^3\mathbf{x} \nabla_{\mathbf{u}}\chi_1[u(\mathbf{m}; \mathbf{x}, t)]\delta\mathbf{u} = \langle \nabla_{\mathbf{u}}\chi_1\delta\mathbf{u} \rangle. \quad (2.22)$$

Avaliar $\delta\mathbf{u}$ diretamente seria computacionalmente custoso, por compreender a modelagem de $u(m + \varepsilon\delta m)$ para todas as direções $\delta\mathbf{m}$ possíveis. Para remover a dependência de $\delta\mathbf{u}$, primeiramente aplica-se o mesmo procedimento das equações 2.21 e 2.22 na equação do operador do problema direto 2.15, obtendo-se a relação

$$\nabla_{\mathbf{m}}L\delta\mathbf{m} + \nabla_{\mathbf{u}}L\delta\mathbf{u} = 0 \quad (2.23)$$

Onde o lado direito da Equação 2.23 é nulo pois as fontes são independentes do modelo $\mathbf{m}$. A Equação 2.23 pode ser modificada multiplicando-se uma função teste $u^\dagger$ a ser definida, e aplicando a integração sobre o tempo e espaço, denotada por $\langle \cdot \rangle$.

$$\langle u^\dagger \nabla_{\mathbf{m}}L\delta\mathbf{m} \rangle + \langle u^\dagger \nabla_{\mathbf{u}}L\delta\mathbf{u} \rangle = 0 \quad (2.24)$$

Adicionando a Equação 2.24 na Equação 2.22, obtém-se

$$\nabla_{\mathbf{m}}\chi(\mathbf{m})\delta\mathbf{m} = \langle \nabla_{\mathbf{u}}\chi_1\delta\mathbf{u} \rangle + \langle u^\dagger \nabla_{\mathbf{m}}L\delta\mathbf{m} \rangle + \langle u^\dagger \nabla_{\mathbf{u}}L\delta\mathbf{u} \rangle \quad (2.25)$$

Na Equação 2.25 será utilizado o método adjunto, aplicando-se os operadores adjuntos $\nabla_u\chi_1^\dagger$ e $\nabla_uL^\dagger$, que satisfazem as equações

$$\langle \nabla_{\mathbf{u}}\chi_1\delta\mathbf{u} \rangle = \left\langle \delta\mathbf{u} \cdot \nabla_{\mathbf{u}}\chi_1^\dagger \right\rangle \quad (2.26)$$

$$\langle u^\dagger \nabla_{\mathbf{u}} L \delta \mathbf{u} \rangle = \langle \delta \mathbf{u} \cdot \nabla_{\mathbf{u}} L^\dagger u^\dagger \rangle \quad (2.27)$$

Substituindo as equações 2.26 e 2.27, que representam os operadores adjuntos, na Equação do gradiente (2.25), obtém-se a equação para o gradiente isolando $\delta \mathbf{u}$ dada por

$$\nabla_{\mathbf{m}} \chi(\mathbf{m}) \delta \mathbf{m} = \langle \delta \mathbf{u} \cdot \nabla_{\mathbf{u}} \chi_1^\dagger \rangle + \langle \delta \mathbf{u} \cdot \nabla_{\mathbf{u}} L^\dagger u^\dagger \rangle + \langle u^\dagger \nabla_{\mathbf{m}} L \delta \mathbf{m} \rangle \quad (2.28)$$

$$\nabla_{\mathbf{m}} \chi(\mathbf{m}) \delta \mathbf{m} = \langle \delta \mathbf{u} \cdot (\nabla_{\mathbf{u}} L^\dagger u^\dagger + \nabla_{\mathbf{u}} \chi_1^\dagger) \rangle + \langle u^\dagger \nabla_{\mathbf{m}} L \delta \mathbf{m} \rangle \quad (2.29)$$

E finalmente, para eliminar a dependência da Equação 2.29 de δu , deve-se exigir que o campo auxiliar $u^\dagger$ satisfaça a condição

$$\nabla_{\mathbf{u}} L^\dagger u^\dagger = -\nabla_{\mathbf{u}} \chi_1^\dagger \quad (2.30)$$

A condição da Equação 2.30 é conhecida como Equação Adjunta, e $u^\dagger$ e $-\nabla_{\mathbf{u}} \chi_1^\dagger$ são conhecidos como campo adjunto e fonte adjunta, respectivamente (FICHTNER [2]).

Dessa forma, encontrada a solução para a equação adjunta, o gradiente utilizando o método adjunto é dado pela seguinte equação

$$\nabla_{\mathbf{m}} \chi(\mathbf{m}) \delta \mathbf{m} = \langle u^\dagger \nabla_{\mathbf{m}} L \delta \mathbf{m} \rangle \quad (2.31)$$

Um paralelo pode ser feito para o domínio da frequência para encontrar o gradiente da função objetivo utilizando o método adjunto. Neste domínio, o campo de onda adjunto discreto no domínio da frequência $\bar{u}^\dagger$ é definido em FICHTNER [2] como

$$L^T \cdot \bar{u}^\dagger = -\nabla_{\mathbf{u}} \chi \quad (2.32)$$

Um desenvolvimento detalhado do método adjunto no domínio da frequência pode ser encontrado, por exemplo, em PRATT [7] e em PLESSIX [10]. O cálculo do campo adjunto no domínio da frequência é computacionalmente bastante eficiente, uma vez que pode-se utilizar a mesma fatoração LU na matriz de impedâncias B da equação da onda (Equação 2.6).

A expressão para o gradiente da função objetivo em relação ao modelo no domínio da frequência fica

$$\nabla_{\mathbf{m}} \chi = \text{Real}(F^t \bar{u}^\dagger) \quad (2.33)$$

E o campo adjunto discreto é dado por

$$\bar{u}^\dagger = B^{-1} \delta \hat{d}^* \quad (2.34)$$

2.4.2 Gradiente da onda acústica

Considerando o modelo $\mathbf{m} = V_p(\mathbf{x})$, o operador L da equação da onda acústica no domínio do tempo (Equação 2.3) é dado por

$$L(u, V_p, \rho) = \frac{1}{\rho(\mathbf{x})V_p^2(\mathbf{x})} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} - \nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \vec{\nabla} u(\mathbf{x}, t) \right) \quad (2.35)$$

Substituindo a Equação 2.35, do operador da onda acústica no domínio do tempo, na Equação 2.31, a equação do gradiente pelo método adjunto, e , obtêm-se a seguinte equação do gradiente da onda acústica

$$\begin{aligned} \nabla_{\mathbf{V}_p} \chi(V_p) \delta \mathbf{V}_p &= \langle u^\dagger \cdot \nabla V_p L \delta \mathbf{V}_p \rangle = \int_T dt \int_G d^3 \mathbf{x} u^\dagger \nabla V_p L \delta \mathbf{V}_p \\ &= \int_T dt \int_G d^3 \mathbf{x} u^\dagger \left(\frac{-2}{\rho(\mathbf{x})V_p^3(\mathbf{x})} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} \right) \delta \mathbf{V}_p \\ &\Leftrightarrow \nabla_{\mathbf{V}_p} \chi(V_p) = \frac{-2}{\rho(\mathbf{x})V_p^3(\mathbf{x})} \int_T dt u^\dagger \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} \end{aligned} \quad (2.36)$$

Onde o campo adjunto $u^\dagger$ é dado pela Equação Adjunta 2.30. A equação adjunta pode ser interpretada como uma propagação de onda reversa no tempo, semelhante a uma migração *RTM* (*Reverse Time Migration*), da fonte adjunta $-\nabla_{\mathbf{u}} \chi_1^\dagger$. A função objetivo χ utilizada nesta dissertação está descrita na Equação 2.11, e portanto a fonte adjunta é dada pela equação

$$f^\dagger(\mathbf{x}, t) = -\nabla_{\mathbf{u}} \chi_1^\dagger = -[u(V_p, \mathbf{x}^r, t) - d(\mathbf{x}^r, t)] \delta(\mathbf{x} - \mathbf{x}^r) \quad (2.37)$$

Onde $\mathbf{x}^r$ indica as posições dos receptores no espaço do problema.

Portanto o campo adjunto é calculado a partir dos resíduos simples, os mesmos utilizados para calcular a função objetivo, que é a Norma 2 destes mesmos resíduos. Esta solução no domínio do tempo dada pelas equações 2.36 e 2.37 significa a correlação temporal entre este campo adjunto e a derivada segunda dos dados modelados, sendo que o campo adjunto é calculado de forma reversa no tempo utilizando os resíduos como fonte. Esta correlação é calculada para todas as fontes, e a contribuição de cada fonte é somada para a construção do gradiente da função objetivo.

Propagação reversa no tempo aparece em diversas aplicações, incluindo migração *RTM* (*Reverse Time Migration*) e *time-reverse imaging of seismic sources*. A maioria destas aplicações tem uma relação próxima com o método adjunto (FICHTNER

[2]).

2.5 Principais algoritmos de otimização local

Uma vez calculado o gradiente $\nabla \chi_{\mathbf{m}}$ da função objetivo de *FWI*, é necessário calcular a atualização do modelo Δm , e o novo modelo que irá para a próxima iteração do algoritmo de *FWI* será $m^{k+1} = m^k + \Delta m^k$.

Para isto, é necessário utilizar um algoritmo de otimização local. Nesta seção estão escritos alguns dos mais utilizados em *FWI*.

2.5.1 Método da máxima descida

O método de máxima descida ou declividade é um dos mais simples algoritmos de otimização local, sendo definido pela equação

$$\Delta x_k = -\alpha_k \nabla f(x_k) \quad (2.38)$$

onde $\nabla f(x_k)$ é o gradiente da função f em relação a x_k que deseja-se otimizar, α_k representa o tamanho do passo, e Δx_k é a direção de máxima declividade.

Usualmente este método é utilizado com algum tipo de preconditionamento. Se uma matriz de preconditionamento P_k estiver disponível, o método fica definido por

$$\Delta x_k = -\alpha_k P_k \nabla f(x_k) \quad (2.39)$$

Para o problema de *FWI* descrito neste trabalho, x_k é o modelo de parâmetros anteriormente descrito como m .

2.5.2 Gradiente Conjugado

O método de Gradiente Conjugado não-linear e preconditionado é definido como

$$\Delta x_k = -P_k \nabla f(x_k) + \beta_k \Delta x_{k-1} \quad (2.40)$$

Onde β_k é um fator calculado seguindo a fórmula de Dai & Yuan [19].

2.5.3 L-BFGS

O método *l-BFGS* utiliza a seguinte direção de descida:

$$\Delta x_k = -Q_k \nabla f(x_k) \quad (2.41)$$

Onde Q_k é a aproximação l-BFGS da inversa da matriz Hessiana $H(x_k)^{-1} = \nabla^2 f(x_k)^{-1}$ (BYRD et. al. [20], NOCEDAL & WRIGHT [21]).

Este método foi utilizado nesta dissertação, utilizando a implementação da *Optimization Toolbox* do consórcio *SEISCOPE*. Mais detalhes em METIVIER & BROSSIER [22].

Capítulo 3

O problema do Salto de Ciclo - *Cycle Skipping*

3.1 Cycle Skipping

A relação entre os dados sintéticos modelados e os dados sísmicos medidos é não-linear. *Full Waveform Inversion* utiliza essa relação (o resíduo entre dados sintéticos e medidos) como função objetivo, e por consequência esta função objetivo também é não-linear. Como descrito no Capítulo 2, o problema inverso de FWI é usualmente resolvido como um problema de otimização local, uma vez que uma otimização global seria muito dispendiosa computacionalmente.

Embora os algoritmos de otimização local permitam que a função objetivo da *FWI* seja minimizada de forma computacionalmente eficiente, existe a possibilidade indesejada de convergência para um mínimo local diferente do mínimo global. Pela natureza oscilatória dos dados sísmicos, o resíduo entre o dado sintético modelado e o dado real medido só poderá ser minimizado se a diferença entre ambos for menor que a metade do período da oscilação ([1]).

A Figura 3.1 ilustra o problema dos mínimos locais para um traço sísmico, sendo um traço sísmico o dado sísmico gerado por uma única fonte e medido em um único receptor, ou seja, um traço do sismograma de tiro. O traço preto em destaque no centro da figura representa o traço sísmico medido. O traço em azul na parte inferior da figura é um traço modelado cuja defasagem para o dado real é menor que metade do período de oscilação. Observa-se que neste caso a direção de minimização da função objetivo de *FWI*, representado pela seta, aponta na direção correta para a convergência entre dados sintético e real, e neste caso a solução de otimização local da *FWI* convergirá para o resultado correto. Na mesma figura, a linha azul na parte superior representa o caso onde o dado modelado inicial tem uma defasagem para o traço sísmico medido maior que metade do período, e neste caso o gradiente para

a minimização local do resíduo aponta para a direção oposta, buscando aproximar o pico $n + 1$ do dado sintético para a posição do pico n do dado medido. Neste caso a solução da FWI ficará presa em um mínimo local, e portanto irá resultar em um modelo de subsuperfície equivocado (por exemplo, a velocidade da onda estará errada). Na literatura este problema é conhecido como Salto de Ciclo, ou *Cycle Skipping*, pois ao aproximar o ciclo representado por $n + 1$ ao ciclo n , a resposta de FWI está equivocada em 1 ciclo.

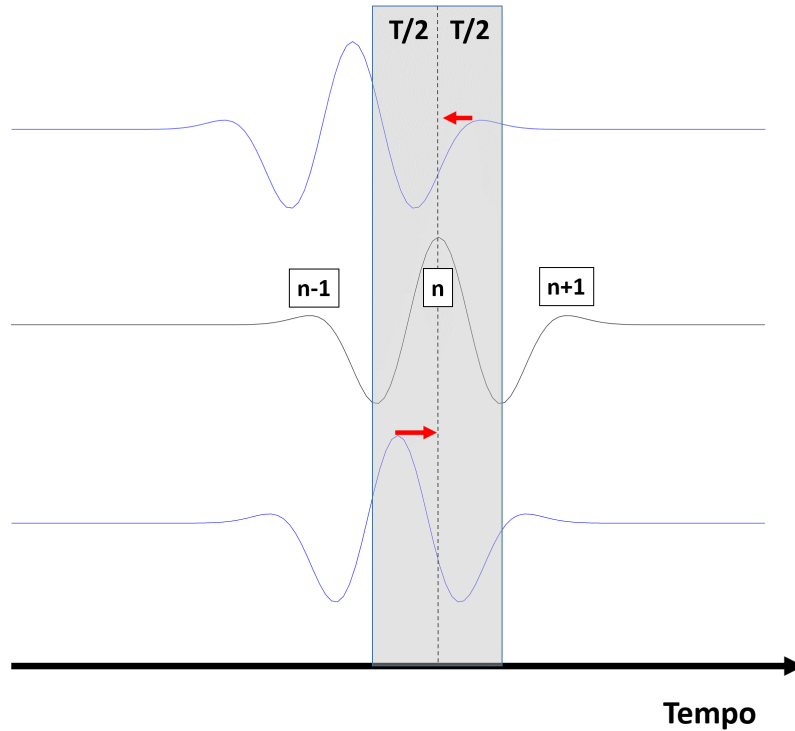


Figura 3.1: Ilustração do problema de Cycle Skipping em um traço. Figura adaptada de VIRIEUX & OPERTO [1]

O caso da Figura 3.1 é apenas 1D, para um traço simples. Porém a metodologia de FWI é utilizada para tentar resolver parâmetros elásticos da Terra em problemas 2D e 3D. E como o método de otimização é local, com o gradiente da função objetivo sendo calculado para cada ponto do modelo, o problema de Salto de Ciclo pode ocorrer em qualquer ponto ou área do modelo, comprometendo o restante da resposta.

A função objetivo para um traço sísmico, aqui representada pela norma L2 dos resíduos entre o traço sísmico medido e o traços sísmicos modelados, a partir de um modelo de velocidades, com diferentes defasagens em relação ao dado medido, está ilustrada no exemplo da Figura 3.2. A figura mostra que a função objetivo, neste caso simples, tem um valor de Mínimo Global no centro da figura. Além deste, há Mínimos locais em ambos os lados das figuras, dependendo da defasagem que o

traço sísmico modelado tem em relação ao traço medido. Como o problema de FWI é resolvido por algoritmos de otimização local, a solução só irá convergir se o traço sísmico simulado por um modelo de velocidades inicial for próximo o suficiente do mínimo global. Ou seja, só haverá convergência para o mínimo global se direção do gradiente da função objetivo a partir de um dado modelado estiver apontando na direção deste mínimo global.

Como este caso sintético da Figura 3.2 foi calculado sem a consideração de ruídos, e para um único traço, na defasagem zero entre os dados modelados e medidos a função objetivo é igual a zero. Mas em casos reais, sempre há presença de ruídos nos dados medidos que não podem ser modelados, logo o mínimo global não terá amplitude zero nesses casos.

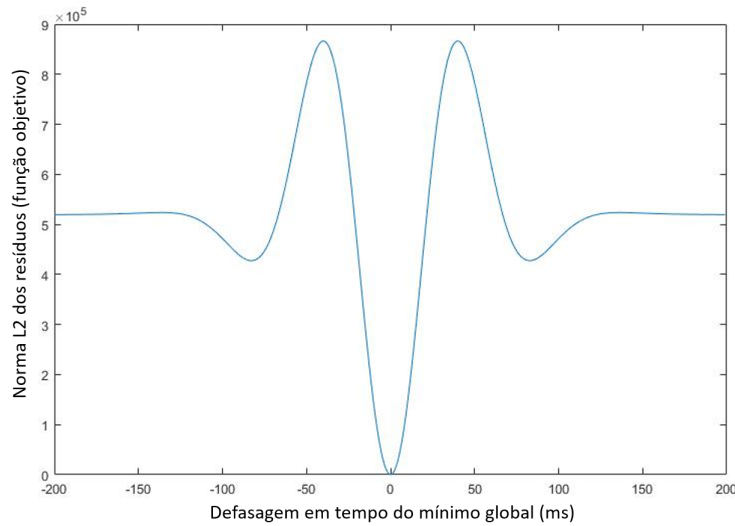


Figura 3.2: Exemplo ilustrativo da função objetivo calculada a partir de um traço sísmico medido, e o

Por este motivo, o resultado da *FWI* é dependente da qualidade do modelo inicial. O modelo inicial deve ser tal que a diferença entre dados sintéticos e reais não seja maior que um “ciclo”, ou meio período da oscilação local. E garantir a qualidade do modelo inicial em todo o espaço da solução é uma tarefa difícil em casos de estudos em grandes áreas de levantamento sísmico 2D e 3D. Na maioria dos casos é necessário ser gasto um longo tempo em processamento de dados e construção de modelo de velocidades utilizando técnicas como análise de velocidades e tomografia para obter-se um modelo de velocidades que possa ser usado como modelo inicial no processo de *FWI*. A mitigação deste problema com *Cycle Skipping*, ou seja, a redução da dependência ao modelo inicial, é um dos maiores desafios atuais para a metodologia de *FWI*.

3.2 Estratégia multi-escala

Problemas inversos realísticos são caracterizados pela presença de múltiplos mínimos locais e de não-unicidade de mínimo global. Para evitar a convergência do método para potenciais soluções em mínimos locais sem significado, devem ser adotadas algumas estratégia especiais de inversão, como a estratégia multi-escala descrita nessa seção (FICHTNER [2]).

A chamada estratégia multi-escala consiste em subdividir o problema em frequências discretas e resolvê-lo a partir das frequências mais baixas até as frequências mais altas progressivamente, vide Figura 3.3. Com isto procura-se reduzir a probabilidade de o problema ficar preso em um mínimo local diferente do mínimo global (*Cycle Skipping*) dado um modelo inicial. Como o período (ou ciclo) é muito maior na onda de baixa frequência, menores são as exigências ao modelo inicial, que precisa ser próximo o suficiente da resposta real para que esteja dentro do ciclo do mínimo global. Após o problema inverso ser resolvido para a mais baixa frequência, o modelo resultante é utilizado como modelo inicial para a próxima frequência discreta, onde o comprimento de onda é um pouco menor, mas provavelmente não haverá problemas com *Cycle Skipping* já que o modelo foi refinado utilizando os dados na frequência anterior.

Outra possibilidade é a implementação simultânea da estratégia multi-escala, onde o processo tenta minimizar a função objetivo em todas as frequências escolhidas simultaneamente.

Portanto, um dos principais fatores para o sucesso da estratégia multi-escala é a escolha das frequências a serem utilizadas no processo de inversão. A frequência mínima deve ser baixa o suficiente para garantir proximidade entre o modelo inicial e o resultado final neste nível de detalhamento. Usualmente as frequências são escolhidas de forma que o número de onda máximo imageado por uma frequência seja semelhante ao mínimo número de onda vertical imageado pela próxima frequência. Como os número de ondas imageados dependem do máximo e mínimo *offset* do levantamento sísmico (SIRGUE & PRATT [9]) é importante levar a geometria de aquisição em consideração ao fazer-se a discretização das frequências. Em um levantamento sísmico com longos *offsets* disponíveis, um número reduzido de frequências é suficiente para representar um dado sísmico como todo e de resolver o problema de *FWI* de forma robusta (AMESTOY et. al. [23]). Lembrando que *offset* é a distância entre as fontes e os receptores do levantamento sísmico.

Esta estratégia de discretização das frequências suprime a redundância na cobertura dos números de onda do dado sísmico, o que é desejável uma vez que o custo computacional é proporcional ao número de frequências utilizadas no processo. Entretanto, uma parte desta redundância pode ser desejada na presença de ruído para

garantir a redução deste ruído pela soma dos ruídos não coerentes (SIRGUE & PRATT [9]).

A Figura 3.3 mostra a decomposição de um traço sísmico no domínio da frequência, ou seja, em traços de diferentes frequências. O traço sísmico pode ser decomposto em baixas e altas frequências, e nas mais baixas frequências o período de onda maior permite uma maior defasagem entre o dado simulado por um modelo inicial e um traço sísmico medido. Ou seja, a exigência de detalhamento do modelo inicial é menor em frequências mais baixas. *FWI* é, portanto, menos susceptível a *Cycle Skipping* se o processo começar com frequências suficientemente baixas, se comparado com a aplicação direta do algoritmo no dado original, não decomposto.

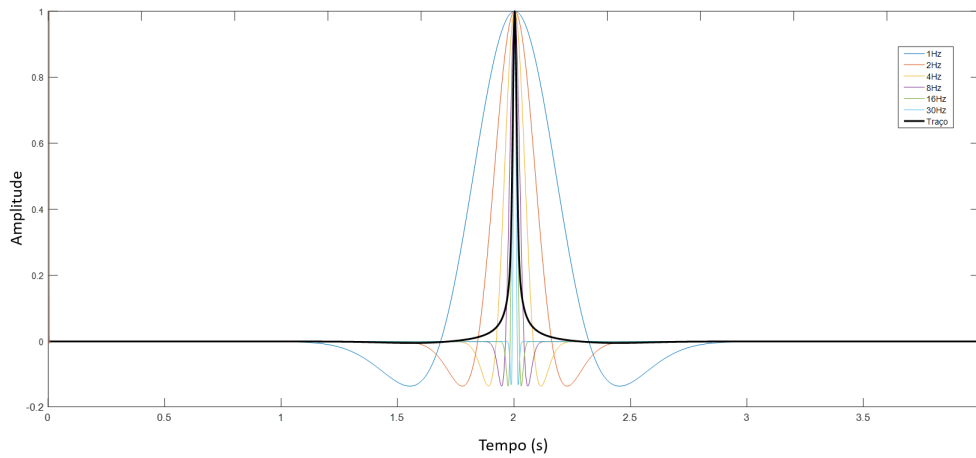


Figura 3.3: Ilustração da divisão de um traço sísmico (ilustrado em preto da figura) em componentes de várias frequências, das mais baixas (mínima 1Hz) até as mais altas (máxima 30Hz). Essa decomposição de frequências é importante em *FWI* pois o período da onda é maior nos componentes do traço em frequências mais baixas (VIRIEUX & OPERTO [1])

A Figura 3.4 é um exemplo ilustrativo adaptado de FICHTNER [2] da estratégia multiescala, começando de uma frequência mais baixa e aumentando sucessivamente para frequências mais altas. Ao aumentar a frequência, o detalhamento aumenta proporcionalmente, e o tamanho dos ciclos da função objetivo representada por $\chi(m)$ diminuem. Sendo assim, neste exemplo, se a estratégia multiescala não fosse aplicada e o processo de *FWI* começasse com o modelo m_0 usando o dado de período intermediário ou o ainda mais curto, provavelmente a solução iria convergir para um mínimo local sem significado, diferente no mínimo global.

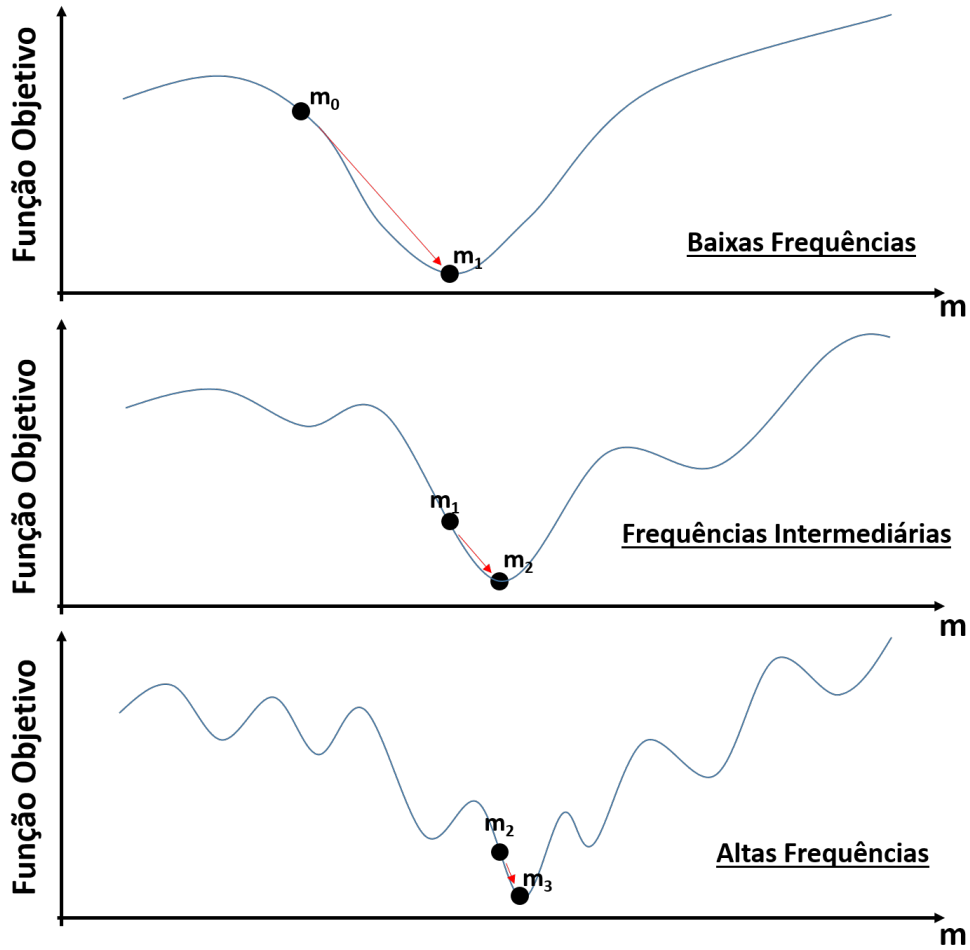


Figura 3.4: Ilustração da estratégia multi-escala, começando no primeiro estágio com uma frequência mais baixa (imagem do topo), aumentando a frequência até o último estágio (imagem da base). Figura adaptada de FICHTNER[2]

3.3 Outras soluções ao *Cycle Skipping*

Com o problema inverso de *FWI* posto como a minimização da norma dos resíduos entre dados sísmicos medidos e modelado, ainda é possível outras soluções para o problema de Salto de Ciclo (*Cycle Skipping*). A mais direta seria buscar a utilização de algoritmos de Minimização Global, no lugar dos algoritmos de minimização local aqui utilizados. Esta alternativa esbarra no alto custo computacional devido à busca intensiva do espaço dos parâmetros.

Outra alternativa são desenvolvimentos de outras funções objetivo, que sejam mais robustas em relação ao problema de Salto de Ciclo. Uma das mais recentes está descrita no Apêndice A, onde na função objetivo os dados sísmicos medidos e modelados são comparados utilizando filtros convolucionais, no lugar da comparação por resíduos. Como o filtro convolucional entre dois dados idênticos é um Delta de

Dirac, então o processo de FWI , neste caso, tentará aproximar o filtro convolucional a um Delta de Dirac, em vez de tentar minimizar a norma dos resíduos.

Capítulo 4

Experimentos e Discussões

Nesta dissertação uma série de experimentos foram realizados para estudar o problema de salto de ciclo, também conhecido na literatura como *Cycle Skipping*.

4.1 Modelo de perturbação gaussiana

Este é um modelo de tamanho pequeno bastante simples. Há apenas uma perturbação no centro do modelo. O modelo está ilustrado na Figura 4.1, onde a perturbação no centro do modelo tem uma velocidade maior que o meio no entorno.

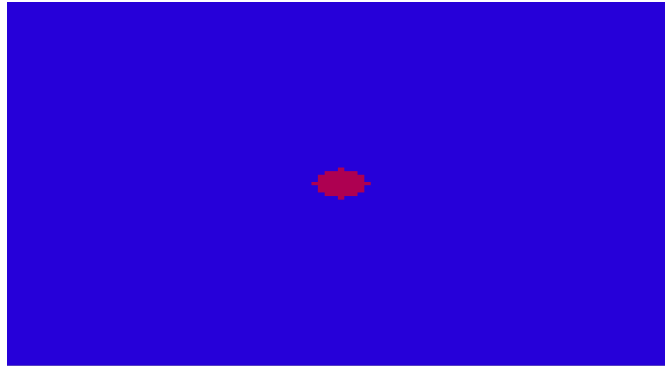


Figura 4.1: Modelo com uma perturbação de alta velocidade no centro (vermelho). O meio tem velocidade V_p de 1500m/s e a perturbação no centro tem velocidade de 1800m/s.

O modelo inicial dado como entrada para o algoritmo de *FWI* é um modelo homogêneo, apenas com a velocidade do meio sem a perturbação de velocidade no centro.

Este primeiro experimento busca simplesmente ilustrar o aumento do detalhamento da resposta do método ao se utilizar de frequências mais altas. A motivação é mostrar que para obter uma resposta detalhada do método *FWI*, faz-se necessário o uso de altas frequências. Como explicitado no Capítulo 3, quanto maior a frequência

maior a chance de Salto de Ciclo, ou seja, da solução potencialmente parar em mínimos locais em algumas regiões do modelo. Entretanto, sem a utilização de altas frequências não é possível a obtenção de um resultado com o detalhamento desejado.

Começamos por uma frequência baixa. A Figura 4.2 é a resposta do algoritmo de *FWI* utilizando um dado sísmico com frequência de 2Hz. Embora o resultado tenha capturado alguma forma de perturbação no centro do modelo, nota-se que com esta frequência a resposta tem uma resolução muito baixa.

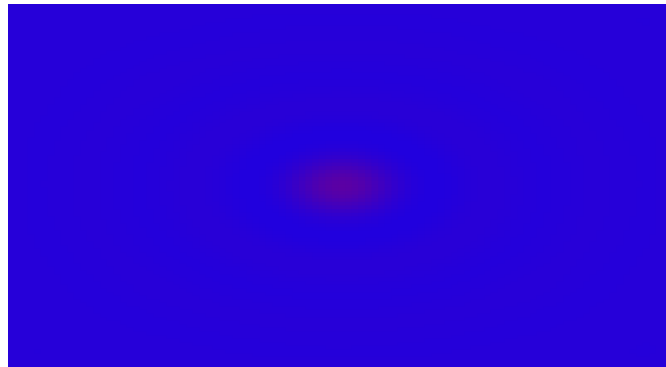


Figura 4.2: Resposta do algoritmo de *FWI* utilizando-se um dado com frequência de 2Hz

Com uma frequência maior, 4Hz, a perturbação de velocidade no centro do modelo passa a ficar um pouco mais nítida, como pode-se observar na Figura 4.3

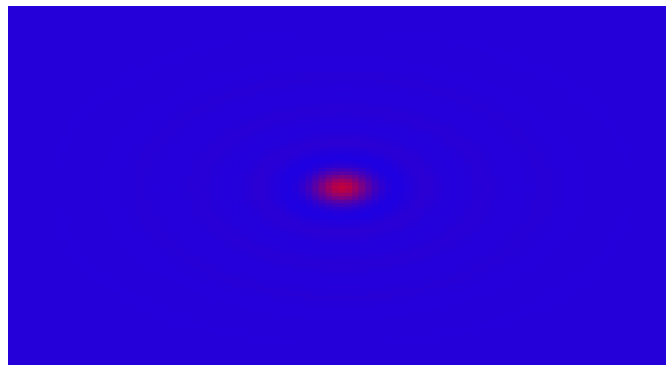


Figura 4.3: Resposta do algoritmo de *FWI* utilizando-se um dado com frequência de 4Hz

Aumentando mais a frequência, o detalhamento melhora consequentemente. Na Figura 4.4 estão ilustrados alguns exemplos com frequências sequencialmente maiores, mostrando a melhora no detalhamento da anomalia no centro do modelo. Este exercício demonstra a importância de tentar utilizar frequências altas, e como o modelo é simples, não houve problemas com salto de ciclo.

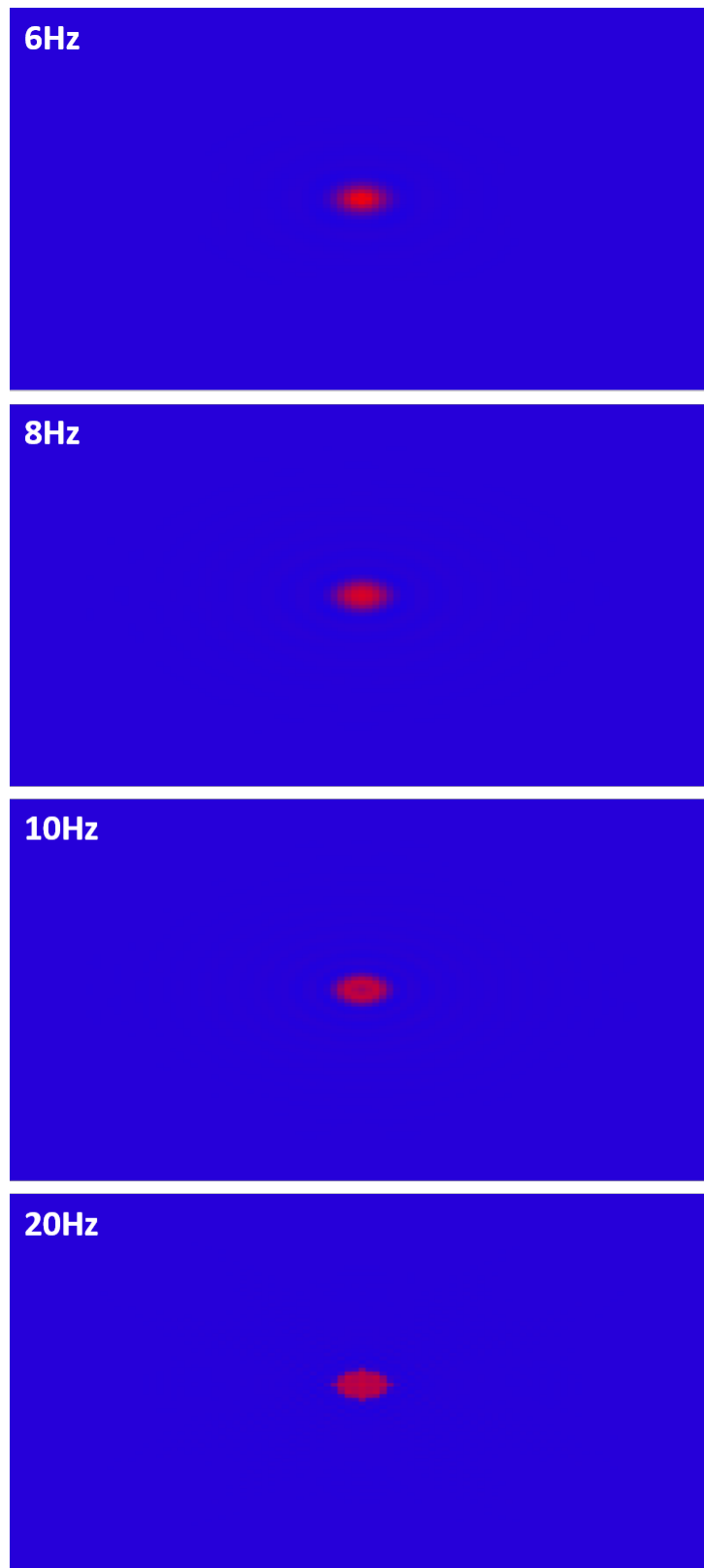


Figura 4.4: Resposta do algoritmo de *FWI* utilizando-se dados com frequência discretas de 6Hz, 8Hz, 10Hz e 20Hz

4.2 Modelo Marmousi

4.2.1 O problema

O modelo de teste Marmousi é um problema amplamente utilizado na geofísica para testar soluções de problemas inversos. O modelo Marmousi inclui feições complexas no meio tanto na sua parte rasa quanto na mais profunda do meio. Este modelo de velocidades está ilustrado na Figura 4.5.

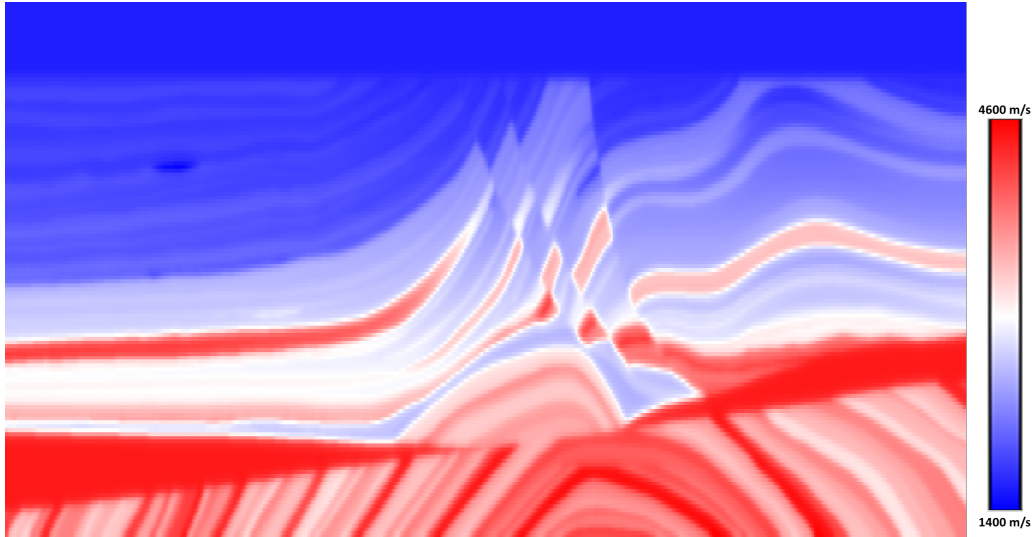


Figura 4.5: Modelo de velocidade V_p Marmousi. Escala em m/s

Da mesma forma como foi feito no experimento anterior, pode-se testar o algoritmo de *FWI* para frequências discretas. No experimento simples da perturbação gaussiana, foi visto que as frequências mais altas eram necessárias para atingir um maior detalhamento da resposta. E no caso deste modelo mais simples, o processo de inversão forneceu resultados satisfatórios para frequências relativamente altas até 20Hz. Entretanto, o modelo Marmousi é maior e muito mais complexo.

O modelo inicial utilizado no problema inverso está ilustrado na Figura 4.6. Ele é uma suavização do modelo inicial. Como pode ser observado, embora suavizado o modelo mantém as feições de escala mais ampla do modelo original de Marmousi. Um modelo inicial próximo do real é necessário uma vez que o problema de *FWI* é não-linear e mal-posto, e um resultado satisfatório dificilmente seria obtido com um modelo inicial homogêneo.

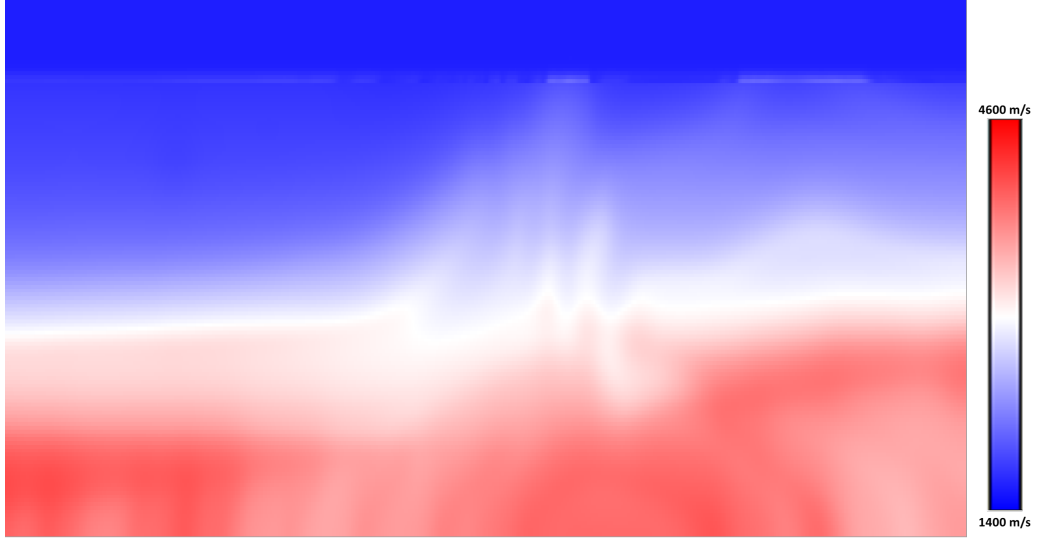


Figura 4.6: Modelo inicial de velocidade V_p , uma suavização do modelo Marmousi da Figura 4.5.

4.2.2 Frequências discretas

No primeiro experimento, o processo de inversão foi testado com a componente do dado de frequência discreta de 2Hz. O resultado está ilustrado na Figura 4.7, e como esperado o modelo de Marmousi não foi totalmente recuperado pelo algoritmo de *FWI*. Assim como no experimento da Figura 4.2, falta resolução ao dado com uma frequência discreta de 2Hz.

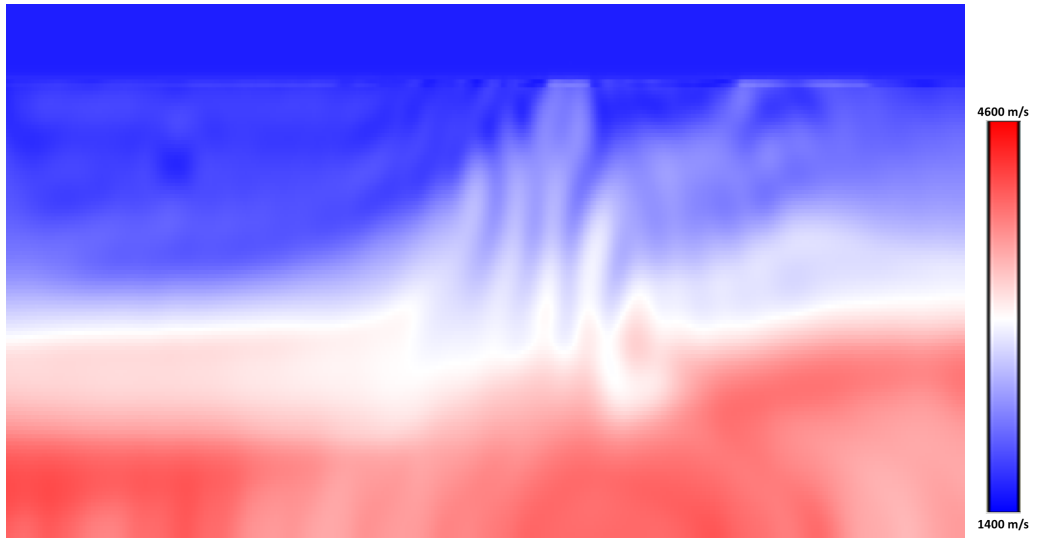


Figura 4.7: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos de frequência 2Hz.

O resultado melhora se um dado com frequência discreta de 4Hz for utilizado (Figura 4.8). Entretanto observa-se alguns artefatos no resultado, e a resposta ainda

está distante do modelo Marmousi. Isso deve-se ao fato de apenas uma frequência discreta não ser suficiente para resolver um modelo tão complexo, e os artefatos podem potencialmente estar relacionados ao problema de Salto de Ciclo.

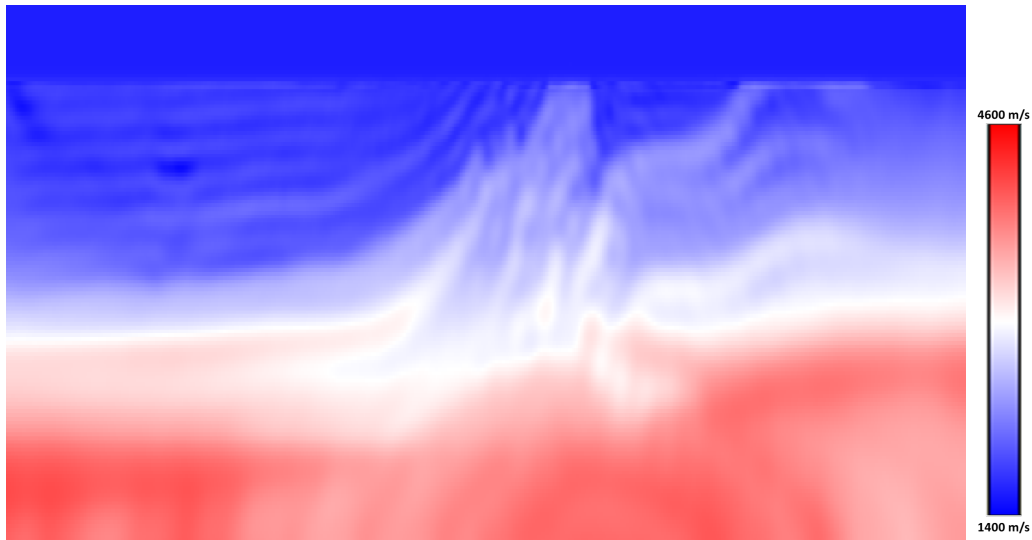


Figura 4.8: Solução do problema inverso fornecido pelo algoritmo de FWI utilizando dados sísmicos de frequência 4Hz.

Ao aumentar ainda mais a frequência, por exemplo para 8Hz (Figura 4.9) e 16Hz (Figura 4.10), o resultado passa a divergir muito do resultado esperado, o modelo de Marmousi da Figura 4.5. O problema de FWI não pode ser resolvido por dados de alta frequência sem informação proveniente de dados de baixa frequência. Todos estes artefatos na imagem mostram onde a solução ficou presa em mínimos locais, pelo motivo do ciclo da função objetivo de um dado de alta frequência ser muito curto.

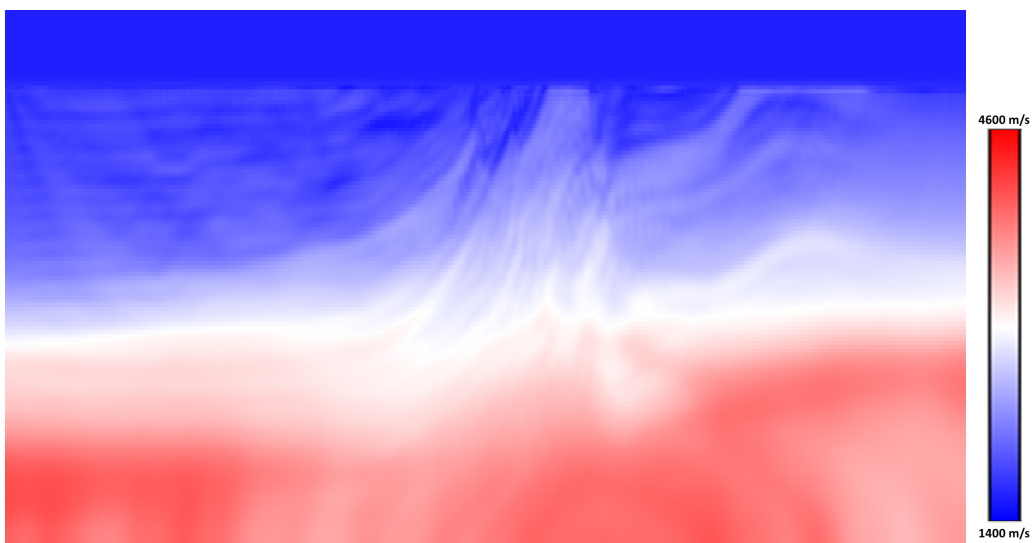


Figura 4.9: Solução do problema inverso fornecido pelo algoritmo de FWI utilizando dados sísmicos de frequência 8Hz.

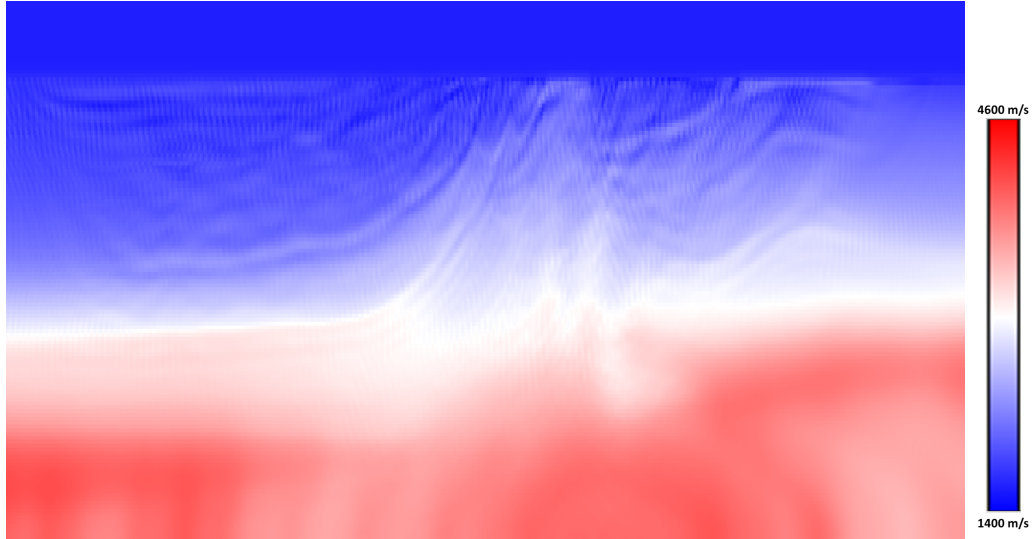


Figura 4.10: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos de frequência 16Hz.

4.2.3 Estratégia multi-escala

Na seção anterior foi mostrado que não é possível recuperar os parâmetros de um modelo complexo como o de Marmousi apenas com uma frequência discreta. As frequências baixas não tem resolução suficiente, e as frequências altas a solução converge para mínimos locais, criando artefatos na solução e divergindo da resposta esperada.

Para solucionar este problema, é necessário aplicar-se a estratégia de multi-escala descrita no Capítulo 3. Um exemplo disto está ilustrado na Figura 4.11, onde foram utilizadas as frequências 1,5Hz, 2Hz, 3Hz, 4Hz, 6Hz e 8Hz de forma sucessiva. O resultado com a estratégia multi-escala é claramente superior aos resultados a partir de frequências discretas. Ao comparar o resultado com a estratégia multi-escala, utilizando seis frequências discretas até 8Hz, com o resultado utilizando dado apenas com a frequência de 8Hz (Figura 4.9), fica evidente que o resultado com a frequência discreta não se aproxima da qualidade da solução alcançada com uma estratégia multi-frequência multi-escala.

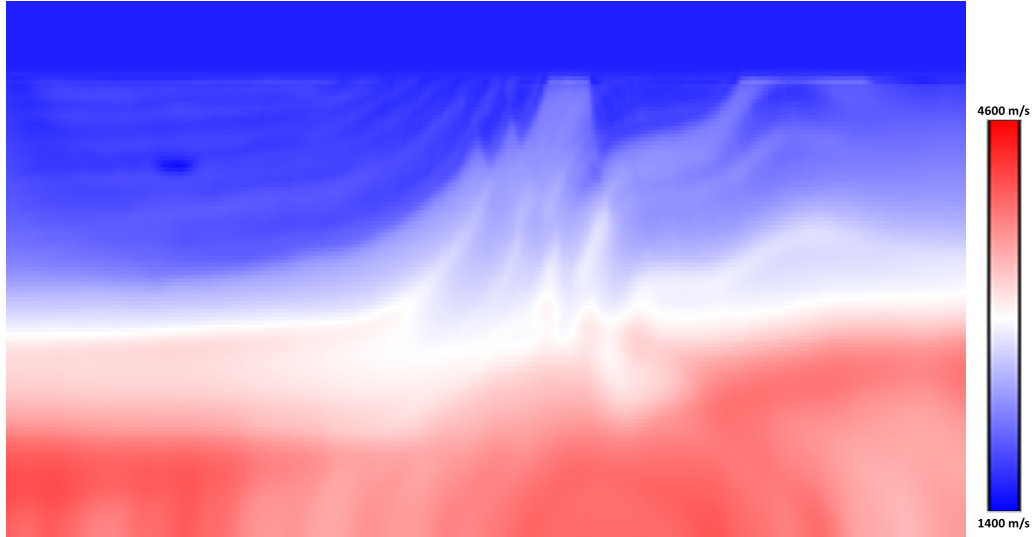


Figura 4.11: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos com seis frequências discretas, começando pela mais baixa e sucessivamente aumentando a frequência. Foram utilizadas as frequências 1,5Hz, 2Hz, 3Hz, 4Hz, 6Hz e 8Hz.

Contudo, a seleção das frequências é importante. Como explicado no Capítulo 3, esta estratégia é dependente do sucesso da inversão nas frequências mais baixas, para então aprimorar o modelo obtido incluindo frequências sucessivamente mais altas. Comparando o resultado com frequências 1,5Hz, 2Hz, 3Hz, 4Hz, 6Hz e 8Hz da Figura 4.11 com o resultado ilustrado na Figura 4.12 fica evidente a importância das baixas frequências. Na Figura 4.12 a menor frequência utilizada foi 3Hz, e mesmo esta sendo uma frequência relativamente baixa para levantamentos de dados sísmicos, observa-se que o resultado se deteriora sem as frequências ainda mais baixas.

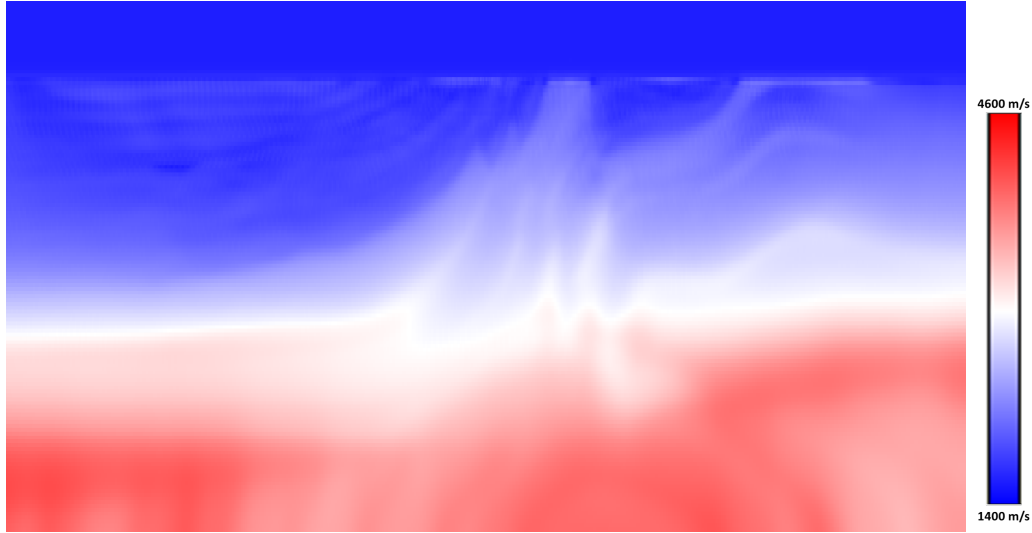


Figura 4.12: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos com seis frequências discretas, começando pela mais baixa e sucessivamente aumentando a frequência. Foram utilizadas as frequências 3Hz, 4Hz, 6Hz, 8Hz, 12Hz e 16Hz.

Até o momento foram apresentados testes relativamente simples, com a utilização de poucas frequências discretas. Entretanto, é comum em problemas de *FWI* dividir a solução em grupos de frequências, com o objetivo de uma discretização menor do dado sísmico e melhorar os resultados, além de tornar a estratégia multi-escala mais robusta. Foi feito um experimento com 4 grupos de frequências, cada um com 5 frequências discretas (Tabela 4.1), começando com o Grupo 1 que contém as frequências mais baixas, de 1,5Hz até 3,5Hz, e o resultado do algoritmo de *FWI* deste grupo é utilizado como entrada para o próximo grupo, o Grupo 2, e assim sucessivamente. O resultado desta estratégia multi-escala utilizando grupos de frequências está ilustrado na Figura 4.13.

Tabela 4.1: Tabela descrevendo os grupos de frequências discretas utilizados no problema inverso, em uma estratégia multi-escala

Grupos	Frequências discretas (Hz)				
Grupo 1	1,5	1,9	2,4	2,9	3,5
Grupo 2	3,5	4,0	4,8	5,8	7,0
Grupo 3	7,0	8,0	9,2	10,6	12,2
Grupo 4	12,2	14,0	16,0	18,4	22,0

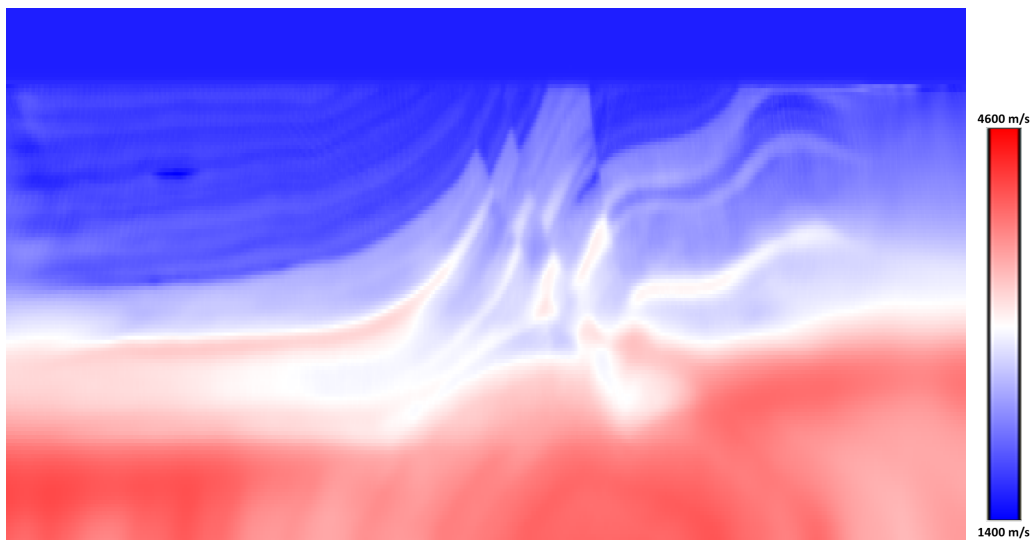


Figura 4.13: Solução do problema inverso fornecido pelo algoritmo de FWI utilizando dados sísmicos com 4 grupos de frequências descritos na Tabela 4.1.

O resultado da Figura 4.13 é superior a todos os anteriores. Neste caso a estratégia multi-escala foi aplicada com sucesso. Vale observar a evolução dos resultados intermediários da estratégia multi-escala, começando pelo Grupo 1 (Figura 4.14), depois utilizando o resultado do Grupo 1 como modelo inicial para o Grupo 2 (Figura 4.15), e utilizando este resultado como modelo inicial para o Grupo 3 (Figura 4.16).

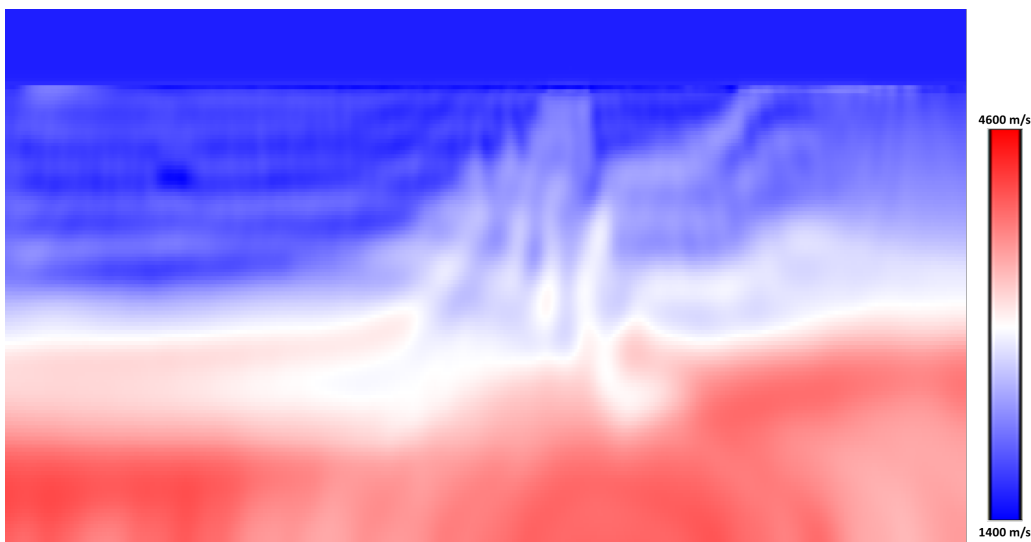


Figura 4.14: Solução do problema inverso fornecido pelo algoritmo de FWI utilizando dados sísmicos do Grupo 1 de frequências descrito na Tabela 4.1.

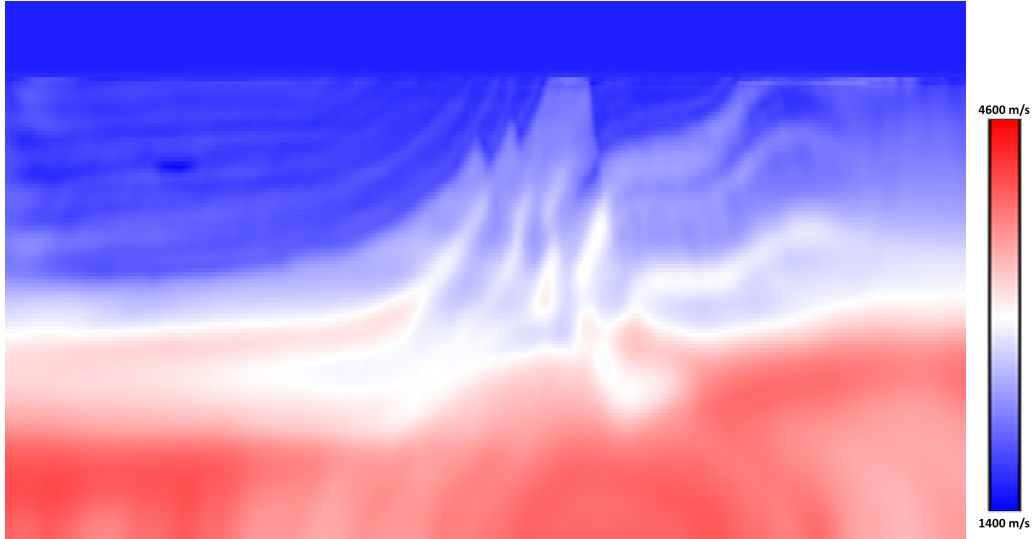


Figura 4.15: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos do Grupo 1 e Grupos 2 de frequências descritos na Tabela 4.1, sucessivamente.

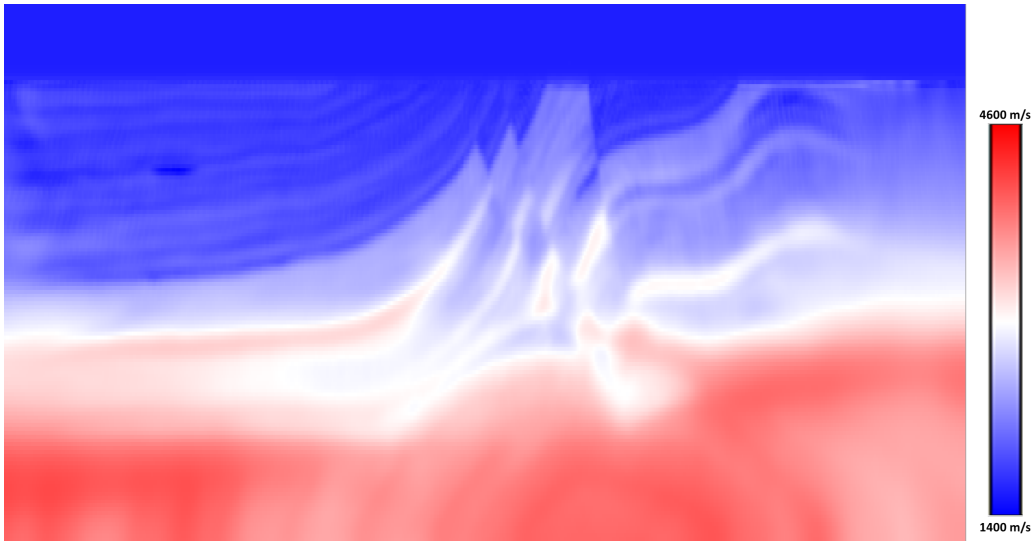


Figura 4.16: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos do Grupo 1, Grupos 2 e Grupo 3 de frequências descritos na Tabela 4.1, sucessivamente.

Nesta estratégia multi-escala a contribuição de todos os grupos de frequências é importante. Caso um destes grupos esteja ausente dos dados sísmicos, o algoritmo de *FWI* pode potencialmente não chegar aos mesmo resultados.

Num experimento, foi considerado que as frequências dos Grupos 1 e 2 estavam ausentes nos dados sísmicos. Neste caso, o resultado de *FWI* para o Grupo 3 seria o ilustrado no Figura 4.17. Observa-se que o resultado tem muito mais artefatos do que se comparado ao resultado do Grupo 3 utilizando os resultados dos Grupos

1 e 2 (Figura 4.16). Isto deve-se ao fato de 7Hz, a menor frequência do Grupo 3, ser uma frequência relativamente alta para resolver este problema inverso, e então o algoritmo de *FWI* encontra problemas de Salto de Ciclo no espaço da solução.

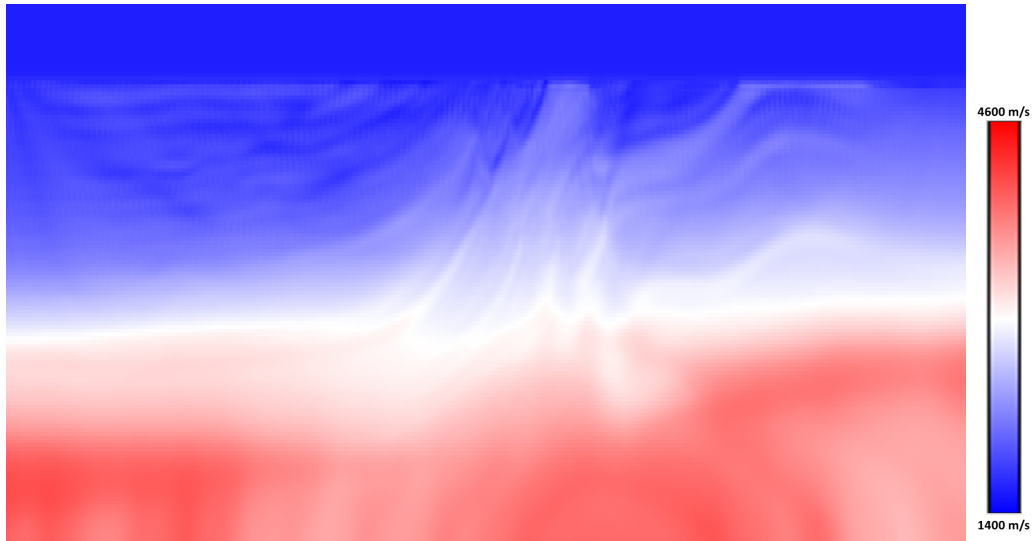


Figura 4.17: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos do Grupo 3, considerando ausentes os dados dos Grupos 1 e Grupo 2 de frequências descritos na Tabela 4.1, sucessivamente.

Mesmo que apenas o Grupo 1 esteja ausente, isto já tem um impacto importante nos resultados (Figuras 4.18 e 4.19).

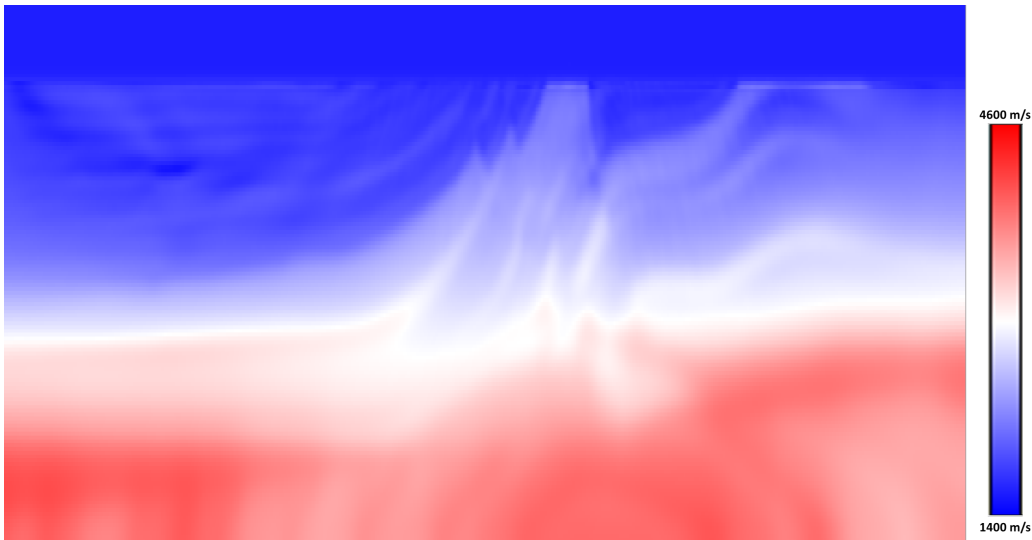


Figura 4.18: Solução do problema inverso fornecido pelo algoritmo de *FWI* utilizando dados sísmicos do Grupo 2, considerando ausentes os dados dos Grupos 1 de frequências

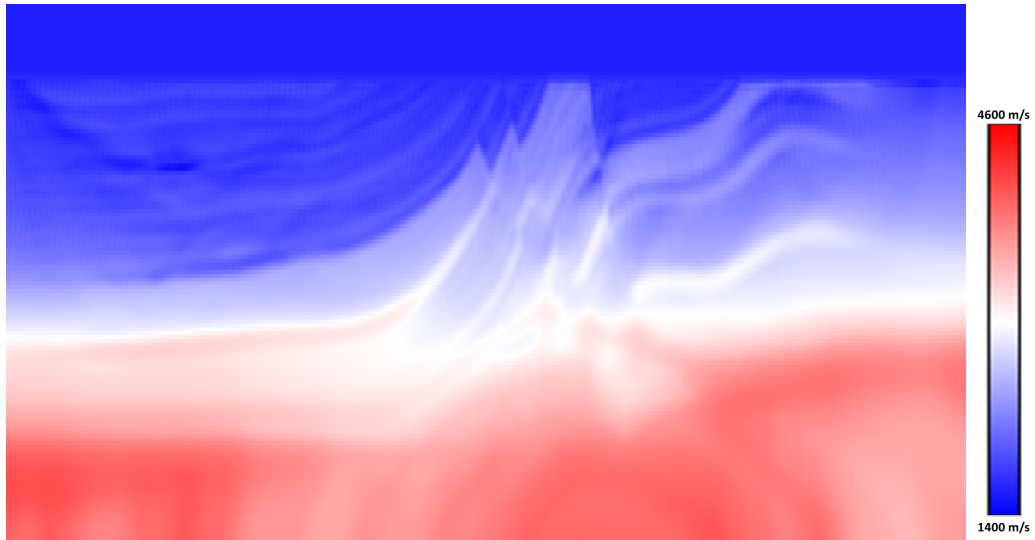


Figura 4.19: Solução do problema inverso fornecido pelo algoritmo de FWI utilizando dados sísmicos do Grupo 2 e 3, sucessivamente, considerando ausentes os dados do Grupos 1 de frequências

Capítulo 5

Conclusões

Pelos experimentos postos nesta dissertação pode-se observar a alta dependência dos resultados do método de *Full Waveform Inversion* ao conteúdo de baixas frequências. Os experimentos demonstram que são necessários dados sísmicos em frequências discretas relativamente baixas, de preferência em torno de 2Hz, mesmo com um modelo inicial razoável.

É possível obter bons resultados apenas com frequências mais altas, desde que o modelo inicial seja próximo suficiente do modelo real para não ocorrer problemas com Salto de Ciclo (*Cycle Skipping*). Quanto mais alta for a frequência mínima disponível nos dados sísmicos medidos, maiores serão as restrições para o modelo inicial. De tal forma que, dependendo do conteúdo de frequência dos dados sísmicos adquiridos, a aplicação de *FWI* pode tornar-se impeditiva, uma vez que não seria possível garantir um modelo inicial tão próximo dos parâmetros físicos reais de subsuperfície. Isto é uma limitação importante para a metodologia, principalmente em sua aplicabilidade em dados sísmicos com conteúdo de baixa frequência restrito, como por exemplo aquisições sísmicas marinhas convencionais.

Concluindo, a estratégia multi-escala no domínio da frequência é uma forma eficiente de resolver o problema inverso de *FWI*, evitando problemas com Salto de Ciclo, com o condicionante de haver um conteúdo abrangente de frequências nos dados sísmicos, principalmente as mais baixas. Caso não haja estas baixas frequências nos dados, o problema com a função objetivo oscilatória de *FWI* não tem solução possível, sendo necessário investigar métodos de otimização global ou modificar a função objetivo.

Referências Bibliográficas

- [1] VIRIEUX, J., OPERTO, S. “An overview of full-waveform inversion in exploration geophysics”, *Geophysics*, v. 74, n. 6, pp. WCC1–WCC26, 2009.
- [2] FICHTNER, A. *Full Seismic Waveform Modelling and Inversion*. Germany, Springer, 2011.
- [3] WARNER, M., GUASCH, L. “Adaptive Waveform Inversion - FWI Without Cycle Skipping - Theory”, *EAGE Extended Abstracts*, 2014.
- [4] TELFORD, W. M., GELDART, L. P., SHERIFF, R. E. *Applied Geophysics*. 2 ed. Cambridge, Cambridge University Press, 1990.
- [5] TARANTOLA, A. “Inversion of seismic reflection data in acoustic approximation”, *Geophysics*, v. 49, n. 8, pp. 1259–1266, 1984.
- [6] LAILLY, P. “The seismic inverse problem as a sequence of before stack migrations”, In *Bednar, R. and Weglein, editors, Conference on Inverse Scattering, Theory and application, Society for Industrial and Applied Mathematics, Philadelphia*, pp. 206–220, 1983.
- [7] PRATT, R. G. “Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model”, *Geophysics*, v. 64, n. 3, pp. 888–901, 1999.
- [8] BUNKS, C., SALECJ, F. M., ZALESKI, S., et al. “Multiscale seismic waveform inversion”, *Geophysics*, v. 60, n. 5, pp. 1457–1473, 1995.
- [9] SIRGUE, L., PRATT, R. G. “Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies”, *Geophysics*, v. 69, n. 1, pp. 231–248, 2004.
- [10] PLESSIX, R.-E. “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications”, *Geophys. J. Int.*, v. 167, pp. 495–503, 2006.

- [11] WARNER, M., RATCLIFFE, A., NANGOO, T., et al. “Anisotropic 3D full-waveform inversion”, *Geophysics*, v. 78, n. 2, pp. R59–R80, 2013.
- [12] VIRIEUX, J., CALANDRA, H., PLESSIX, R.-E. “partial review of modelling methods for geophysical imaging”, *Geophysical Prospecting*, 2011.
- [13] BROSSIER, R., ETIENNE, V., S., O., et al. *Frequency-Domain Numerical Modelling of Visco-Acoustic Waves with Finite-Difference and Finite-Element Discontinuous Galerkin Methods, Acoustic waves*. France, D.W. Dissanayake, 2010.
- [14] ETIENNE, V., CHALJUB, E., VIRIEUX, J., et al. “An hp-adaptive discontinuous Galerkin Finite-element method for 3D elastic wave modelling”, *Geophys. J. Int.*, v. 183, pp. 941–962, 2010.
- [15] MARFURT, K. J. “Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations”, *Geophysics*, v. 49, n. 5, pp. 533–549, 1984.
- [16] THOMSEN, L. “Weak elastic anisotropy”, *Geophysics*, v. 51, n. 10, pp. 1954–1966, 1986.
- [17] GUITTON, A., SYMES, W. W. “Robust inversion of seismic data using the Huber norm”, *Geophysics*, v. 68, n. 4, pp. 1310–1319, 2003.
- [18] AMESTOY, P., BROSSIER, R., L’EXCELLENT, J.-Y., et al. “Efficient 3D frequency-domain full-waveform inversion of ocean-bottom cable data with sparse block low-rank direct solver: a real data case study from the North Sea”, *SEG Technical Program Expanded Abstracts*, pp. 1303–1308, 2015.
- [19] DAY, Y., YUAN, Y. “nonlinear conjugate gradient method with a strong global convergence property”, *SIAM Journal on Optimization*, v. 10, pp. 177–182, 1999.
- [20] BYRD, R. H., LU, P., NOCEDAL, J. “A limited memory algorithm for bound constrained optimization”, *SIAM Journal on Scientific and Statistical Computing*, v. 16, pp. 1190–1208, 1995.
- [21] NOCEDAL, J., WRIGHT, S. J. *Numerical Optimization*. 2 ed. USA, Springer, 2006.
- [22] METIVIER, L., BROSSIER, R. “The seiscopes optimization toolbox: A large-scale nonlinear optimization library based on reverse communication”, *Geophysics*, 2015.

- [23] AMESTOY, P., BROSSIER, R., BUTTARI, A., et al. “Efficient 3D frequency-domain full-waveform inversion of ocean-bottom cable data with sparse block low-rank direct solver: a real data case study from the North Sea”, *SEG Technical Program Expanded Abstracts*, 2015.
- [24] VIRIEUX, J. “P-SV wave propagation in heterogeneous media, velocity-stress finite difference method.” *Geophysics*, v. 51, pp. 889–901, 1986.
- [25] LEVANDER, A. R. “Fourth-order finite-difference P-SV seismograms”, *Geophysics*, v. 53, n. 11, pp. 1425–1436, 1988.
- [26] MIN, D.-J., SHIN, C., PRATT, R. G., et al. “Weighted-averaging finite-element method for 2D elastic wave equations in the frequency domain”, *Bull. Seis. Soc. Am.*, v. 93, n. 2, pp. 904–921, 2003.
- [27] BROSSIER, R., VIRIEUX, J., OPERTO, S. “Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling”, *Geophysical Journal International*, v. 175, n. 2, pp. 541–559, 2008.
- [28] DANECEK, P., SERIANI, G. “An efficient parallel Chebyshev pseudo-spectral method for large-scale 3D seismic forward modelling.” *EAGE Extended Abstract*, 2008.
- [29] HUSTEDT, B., OPERTO, S., VIRIEUX, J. “Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling”, *Geophysical Journal International*, v. 157, pp. 1269–1296, 2004.
- [30] OPERTO, S., VIRIEUX, J., RIBODETTI, A., et al. “Finite-difference frequency-domain modeling of visco-acoustic wave propagation in two-dimensional TTI media”, *Geophysics*, v. 74, n. 5, pp. T75–T95, 2009.
- [31] WARNER, M., GUASCH, L. “Adaptive Waveform Inversion: Theory”, *SEG Extended Abstracts*, 2014.
- [32] GUASCH, L., BURGESS, T., WARNER, M. “Optimised Adaptive Waveform Inversion - Improved Convergence Via Conjugate Gradients and Superior Step-length Calculation”, *EAGE Extended Abstracts*, 2015.

Apêndice A

Adaptive Waveform Inversion

AWI (*Adaptive Waveform Inversion*) é uma metodologia que busca resolver ou mitigar os problemas da FWI convencional com saltos de ciclo (conhecido na literatura como Cycle Skipping). Introduzida por Warner & Guasch em 2014, esta metodologia tem o potencial de lidar com diversas limitações da FWI convencional, como por exemplo reduzir as restrições em relação à qualidade do modelo inicial e à presença de baixas frequências no dado como boa relação sinal/ruído.

O conceito de AWI é utilizar filtros convolucionais como função objetivo no lugar dos resíduos utilizados em FWI (entre dados modelados e dados reais). Este filtro convolucional é construído de forma tal que, ao aplicá-lo aos dados modelados, estes se aproximem dos dados reais, ou vice-versa. Esta nova forma de função objetivo parece não possuir mínimos locais que causem saltos de ciclo, em parte pelo fato do filtro convolucional sempre aproximar os dados reais dos dados modelados, como uma extensão não mínima do modelado. Esta função objetivo está descrita em detalhes a seguir.

Entretanto, a relação dessa nova função objetivo com a diferença entre dados reais e modelados (resíduos) é menos direta do que em *FWI*. Isto gera consequências para a convergência do método descritas na sequência.

FWI convencionalmente busca minimizar a norma quadrada dos resíduos entre dados modelados e dados medidos. O método alternativo proposto na *AWI* utiliza filtros de Wiener no lugar da subtração direta entre dados medidos e dados modelados. Como na equação A.1, onde $p(m)$ são os dados modelados, w é o filtro de Wiener e d são os dados medidos (WARNER & GUASCH [3]).

$$p(m) * w \approx d \tag{A.1}$$

O contrário também pode ser utilizado, onde calculamos o filtro tal que os dados reais se aproximem dos dados medidos (equação).

$$d * w \approx p(m) \quad (\text{A.2})$$

Em suas duas formas, quando o modelo é tal que os dados sintéticos sejam similares aos dados reais, o filtro de Wiener será próximo de um Delta de Dirac no tempo zero. Isso seria o equivalente a minimizar os resíduos no esquema de FWI convencional. Portanto, a primeira função objetivo da metodologia de AWI é encontrar os coeficientes do filtro de Wiener w que minimiza, na forma de mínimos quadrados, A.3, onde P é uma matriz convolucional de Toeplitz que contém o vetor de dados modelados $p(m)$ em cada coluna.

$$f_1 = \|Pw - d\|^2 \quad (\text{A.3})$$

A equação é aplicada traço a traço. Ela é linear, com apenas um mínimo possível, e provê uma solução única para w .

A segunda etapa da inversão, ou segunda função objetivo, é mencionar a amplitude dos coeficientes do filtro de Wiener w . Lembrando que quanto mais semelhantes os dados modelados $p(m)$ forem dos dados reais d , mais próximo o filtro construído será de um Delta de Dirac centralizado no tempo zero. Desta forma, uma possível função objetivo é minimizar as amplitudes dos coeficientes w distantes de zero, como na equação A.4.

$$f_1 = \frac{\|Tw\|^2}{\|w\|^2} \quad (\text{A.4})$$

Onde T é uma matriz diagonal simples que funciona como pesos para os coeficientes do filtro de Wiener. A matriz T deve dar um peso maior aos coeficientes distantes de zero do filtro de Wiener, ao mesmo tempo que permita um peso nulo ou mínimo no tempo zero. Neste caso, o funcional f_2 deve ser minimizado. Já no caso oposto, onde a matriz T dê um peso maior aos coeficientes próximos de zero, a função objetivo f_2 deve ser maximizada.

A normalização dos coeficientes do filtro na equação X é necessária para a inversão não buscar minimizar a função objetivo simplesmente aumentando a amplitude dos dados $p(m)$.

Segundo Warner & Guasch, esta função objetivo f_2 tem a vantagem de não sofrer de saltos de ciclo (Cycle Skipping). A figura ilustra a função objetivo oscilatória da metodologia de FWI convencional em comparação com o funcional adaptativo f_2 da metodologia de AWI.

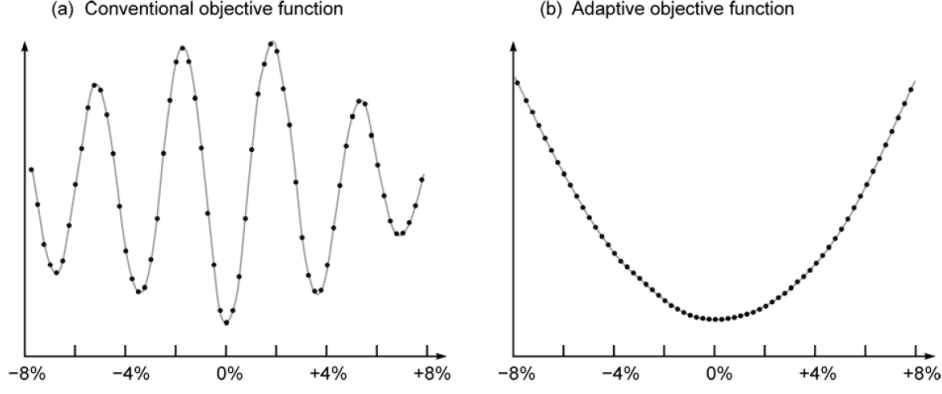


Figura A.1: Comparação entre a função objetivo da FWI (a) e da AWI (b) Warner & Guasch 2014

Concluindo, o uso do filtro convolucional de Wiener é a parte adaptativa da metodologia de AWI, cujo restante do algoritmo é bastante semelhante ao de FWI. Filtros de Wiener são bem conhecidos na literatura, e é a proximidade da comparação entre dados modelados e reais utilizando este filtro que remove problemas com Cycle Skipping da formulação de AWI (Warner & Guasch, 2015).

A.1 Formulação de *AWI*

A formulação de *AWI* é bastante próxima da metodologia de *FWI* convencional. As principais diferenças são na função objetivo, ou funcional, e no cálculo do gradiente que será descrito na sessão seguinte. Enquanto *FWI* busca minimizar os resíduos, *AWI* irá buscar comparar o filtro convolucional com um filtro de referência, no caso um Delta de Dirac centrado no tempo zero. A.1 ilustra um fluxograma completo da metodologia.

A.2 Gradiente de *AWI*

O cálculo do gradiente no *AWI*, resolvido no domínio do tempo, assemelha-se ao cálculo do gradiente do *FWI* convencional. O método adjunto utilizado no *FWI* também é válido para a formulação de *AWI*, mudando apenas a fonte adjunta δS .

A fonte adjunta do método adjunto de cálculo do gradiente é a diferenciação da função objetivo pelos dados modelados p . A fonte adjunta do *FWI* convencional são os resíduos diretamente calculados entre os dados observados reais e os dados modelados (equação). A fonte adjunta é propagada inversamente no tempo e correlacionada com a propagação direta dos dados modelados para estimar o gradiente em relação ao modelo m .

A fonte adjunta para um para fonte-receptor é um pouco mais complicada que no caso de *FWI*. Ela é dada por WARNER & GUASCH [31]:

$$\delta s = \frac{\partial f}{\partial p} = D(w^T w)^{-1} \left(\frac{T^2 - 2f_2 I}{w^T w} \right) w \quad (\text{A.5})$$

Na equação A.5, D representa uma matriz convolucional de Toeplitz dos dados observados, ou seja, representa a convolução com os dados observados. $w^T w$ representa o produto interno dos coeficientes de w , utilizado para normalizar w . $(T^2 - 2f_2 I)$ representam os pesos dos coeficientes de w utilizando uma função temporal. DTD representa a autocorrelação dos dados, que é utilizada para deconvolver o resultado de $(T^2 - 2f I / w^T w) w$ (deconvolução representada pelo -1). Finalmente, o resultado da deconvolução é convolvido com o dado original d .

Esta operação é calculada traço a traço, e seu resultado é propagado inversamente no tempo e correlacionado com os dados reais no esquema do método adjunto para formar o gradiente em relação ao modelo.

Este cálculo traço-a-traço é a maior diferença entre *FWI* e *AWI*, levando mais operações do que o simples cálculo dos resíduos nos receptores. É um cálculo computacionalmente eficiente, e o acréscimo computacional é insignificante comparado ao custo da modelagem direta e da propagação reversa no tempo.

A.3 Convergência

Embora a metodologia de *AWI* seja insensitiva em relação a mínimos locais causados por salto de ciclo (Cycle Skipping), o formato da função objetivo pode não ser ideal para métodos de otimização local, o que pode resultar na necessidade de muitas iterações para a convergência do resultado. Um estudo de convergência foi desenhado em Guasch et. al. 2015, mostrando que uma implementação simples do *AWI* utilizando um método de otimização local steepest-descent com estimativas de step lenght lineares produz uma baixa taxa de convergência em comparação com *FWI* convencional. Sendo a causa disso o fato que a relação entre dado observado e dado modelado é muito mais direta na *FWI*, onde a função objetivo utiliza os resíduos dos dados diretamente e a aproximação de Born justifica a linearização no método de otimização. Como *AWI* consiste na ponderação dos coeficientes de um filtro de Wiener entre dados observados e modelados, a relação deste funcional com os dados é muito menos linear.

Guasch et. al. (2015) também mostra que é possível melhorar as taxas de convergência utilizando um algoritmo melhor para estimativa dos step-lenght combinados com um algoritmo de gradiente conjugado.

Uma vez que a convergência de *AWI* é mais lenta que *FWI*, uma combinação das

duas metodologias também é uma possível solução. Uma vez garantido que o modelo esteja próximo suficiente do modelo real para que FWI não tenha problemas com Cycle Skipping, FWI pode ser aplicada a partir de um modelo gerado por iterações de AWI. Além disso, esta combinação pode ajudar a resolver alguns artefatos na solução combinada de CG e AWI, como mostrado em Guasch et. al. [32].

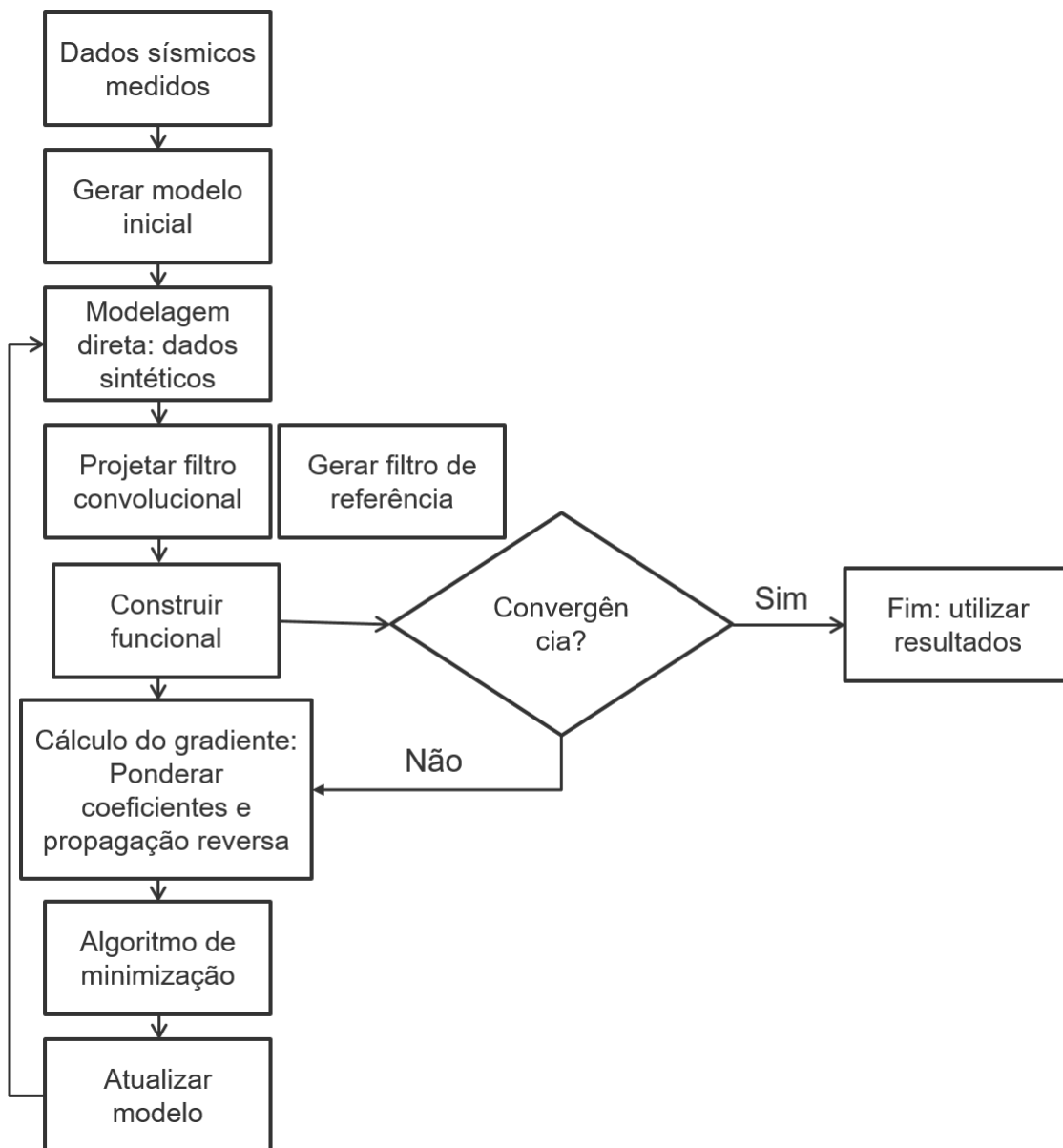


Figura A.2: Fluxograma básico da metodologia de AWI), alterado de WARNER & GUASCH [3].