Important Notice

UNIVERSITY OF CALGARY

Time-Lapse Seismic Imaging, Full-waveform Inversion, and Uncertainty Quantification

by

Xin Fu

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN GEOLOGY AND GEOPHYSICS

CALGARY, ALBERTA

SEPTEMBER, 2023

Abstract

Time-lapse seismic, also known as 4D seismic, is a powerful tool for monitoring subsurface changes over time. By comparing seismic data acquired at different intervals, it enables the detection and characterization of dynamic reservoir processes, aiding in reservoir management, production optimization, and enhanced oil recovery. It has applications in geothermal energy, CO2 storage monitoring, and environmental impact assessment. However, accurate analysis of time-lapse seismic data remains a challenging task. It requires well-repeated time-lapse seismic surveys, including well-repeated acquisition geometry and equipment as well as well-repeated ambient noise. This thesis is to alleviate the non-repeatability issues in time-lapse seismic imaging and full-waveform inversion (FWI), and to realize the uncertain quantification for time-lapse seismic waveform inversion. A time-lapse imaging approach that involves two new frequency-domain matching filters is developed. The first filter requires source wavelet estimates from both baseline and monitoring data, while the second filter is source-independent but more sensitive to data noise. By applying these filters, we successfully reduce source wavelet non-repeatability, and the new approach improves the accuracy of time-lapse imaging. Furthermore, a stepsize-sharing time-lapse FWI strategy is designed to reduce artifacts caused by the variability of convergence in conventional strategies. The strategy demonstrates good adaptivity in different tested realistic scenarios using synthetic data. It is stable for scenarios using biased starting models, while the conventional strategies fail in this regard. Moreover, to realize the uncertain quantification, a Bayesian time-lapse FWI procedure, based on a Markov chain Monte Carlo (MCMC) algorithm, is formulated. The formulation employs several existing strategies, including the use of a double-difference time-lapse FWI, incorporation of time-domain multi-source data, and application of a local-updating target-oriented inversion. It incorporates these within a stochastic framework, involving the computation of model covariance with an adaptive Metropolis algorithm, and a method to estimate data error statistics based on the features of time-lapse difference data is incorporated. A random walk Metropolis-Hastings MCMC is adopted for optimization.

Preface

Chapter 2 of this thesis has been previously published:

X. Fu and K. A. Innanen."Time-lapse seismic imaging using shot gathers with nonrepeatable source wavelets." Geophysics 88(1):M17-M30, 2023. doi: 10.1190/geo2022-0072.1.

Chapter 4 of this thesis has been previously published:

X. Fu and K. A. Innanen."Stepsize sharing in time-lapse full-waveform inversion" Geophysics 88(2):M59-M70, 2023. doi: 10.1190/geo2022-0094.1.

Chapter 5 of this thesis has been previously published:

X. Fu and K. A. Innanen."A time-domain, multi-source Bayesian / Markov Chain Monte Carlo formulation of time-lapse seismic waveform inversion" Geophysics 87(4):R349-R361, 2022. doi: 10.1190/geo2021-0443.1.

Acknowledgements

I thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grants CRDPJ 461179-13 and CRDPJ 543578-19.

I would like to express my deepest gratitude and appreciation to my supervisor, Dr. Kristopher Innanen, for his invaluable guidance, unwavering support, and continuous encouragement throughout the duration of my doctoral journey. His expertise, patience, and commitment to my academic development have been instrumental in shaping this thesis.

I am also immensely grateful to the members of my thesis committee, Dr. Daniel Trad, Dr. Wenyuan Liao, Dr. Brandon Karchewski and Dr. Jan Dettmer, for their insightful feedback, constructive criticism, and valuable suggestions. Their expertise and scholarly contributions have significantly enriched the quality of this research.

I would like to extend my sincere appreciation to the faculty and staff of the Depart of Geoscience at University of Calgary for creating a conducive academic environment and providing the necessary resources for my research. Their dedication and support have been pivotal in my academic growth.

I am indebted to my colleagues and fellow researchers in CREWES project for their collaboration, intellectual discussions, and shared experiences. Their diverse perspectives and contributions have broadened my understanding and enriched the research process.

I am grateful to my family and friends for their unwavering belief in me and their unconditional support throughout this journey. Their love, encouragement, and understanding have

been a constant source of motivation, enabling me to overcome challenges and persevere.

To everyone who has played a role, big or small, in the completion of this thesis, I extend my heartfelt appreciation. Your contributions and support have been invaluable, and I am truly grateful for the collective effort that has gone into this work.

Dedication

To all the beauty in the universe, and also to my lovely wife, Di.

Table of Contents

List of Figures and Illustrations

Chapter 1

Introduction

1.1 Challenges

At present, seismic monitoring stands out as the most effective technique for long-term subsurface monitoring in the context of CO2 sequestration. The use of time-lapse (or 4D) seismic methods for reservoir monitoring and characterization has been developed for several decades since the mid to late 1980s (Greaves and Fulp, 1987; Lumley, 2001; Landrø, 2001; Calvert, 2005; Hicks et al., 2016; Jack, 2017; Cho and Jun, 2021). Time-lapse seismic technology is particularly useful for tracking changes in reservoirs caused by hydrocarbon production (e.g., enhanced oil recovery) and underground CO_2 storage. As the demand for technologies to control greenhouse gas emissions increases, many researchers have been developing CO_2 storage in the subsurface, and time-lapse seismic methods are used to monitor these subsurface CO_2 storages (Egorov et al., 2017; Cho and Jun, 2021; Ajo-Franklin et al., 2013; Macquet et al., 2019). However, successful seismic monitoring relies heavily on the repeatability of baseline and monitor surveys, which can be affected by various factors such as weather conditions, source and receiver positions, environmental noises, source wavelets, seawater, and near-surface properties.

To mitigate the impact of the aforementioned factors, it is essential to have good ac-

quisition plans and proper data processing methods. These include repeatable acquisition geometries and matching filter designs. For instance, in the Foinhaven and Valhall fields, permanent ocean-bottom-cable installations are used to obtain good repeatable data during time-lapse seismic surveys (Calvert, 2005; Yang et al., 2016). In the Aneth oil field in Utah, receivers are cemented in the monitor well to acquire time-lapse VSP (vertical seismic profile) data (Cheng et al., 2010). In the CO2CRC Otway field experiment, Shulakova et al. (2015) improved the repeatability of land seismic data by burying receivers to reduce noise caused by poor weather conditions, non-repeatable receiver positions, near-surface changes, and non-repeatable survey environments. During time-lapse data processing, a cross-equalization method is often used to enhance the repeatability between baseline and monitor data (Rickett and Lumley, 2001). Fu et al. (2020) proposed a double-wavelet method to eliminate source wavelet non-repeatability, but it requires source wavelets information of baseline and monitor data. To avoid this requirement, a source-independent matching filter will be designed in Chapter 2 in this thesis. In recent years, time-lapse seismic surveys based on fiber-optic distributed acoustic sensing (DAS) systems have become increasingly popular. This is because downhole DAS arrays can be permanently installed, have lower monitoring costs, and provide finer spatial sampling (Zwartjes et al., 2018; Byerley et al., 2018; Wilson et al., 2021).

As a powerful tool to monitor subsurface reservoir changes and/or CO_2 storages, time-lapse seismic full-waveform inversion (FWI) (Lailly et al., 1983; Tarantola, 1984; Virieux and Operto, 2009) can provide high-resolution imaging of the physical properties of subsurface media, and it can solve the problem of non-repeatable receiver/source positions in time-lapse seismic surveys (Zhou and Lumley, 2021c; Fu and Innanen, 2022a). In the past decade, many time-lapse FWI methods have been developed. The most conventional time-lapse FWI strategy is the parallel strategy (PRS) (Lumley et al., 2003; Plessix et al., 2010). However, its result is prone to be affected by the convergence difference (Yang et al., 2015a) between baseline and monitor inversions as well as the non-repeatable receiver/source positions (Zhou

and Lumley, 2021c; Fu and Innanen, 2022a). Routh et al. (2012) present the sequential strategy (SQS), using the inverted baseline model as a starting model for monitor inversion, which can help to save computational cost and has been justified in the case of field VSP data (Egorov et al., 2017). However, this strategy often generates strong artifacts since it enhances the convergence difference between baseline and monitor inversions (Yang et al., 2015a; Zhou and Lumley, 2021c). But a local-updating SQS can efficiently reduce the artifacts and perform well in both synthetic and field time-lapse data (Raknes and Arntsen, 2014; Asnaashari et al., 2015). Also, the local-updating method can be incorporated with the double-difference strategy (DDS), which will be introduced later, to improve the time-lapse results (Zhang and Huang, 2013; Li et al., 2021), or to alleviate the impact of taking an acoustic approximation to elastic subsurface rocks (Willemsen et al., 2016). Moreover, the local-solver-based local-updating method can significantly decrease the computational cost of time-lapse FWI (Willemsen, 2017; Huang et al., 2018; Kotsi et al., 2020). Even so, the local-updating method needs prior location information about reservoir change, which may be challenging to be sure of in some cases of non-repeatable time-lapse surveys.

The DDS, directly minimizing residuals between synthetic difference data (synthetic monitor data minus synthetic baseline data) and observed difference data (observed monitor data minus observed baseline data), applied in time-lapse FWI first by Zheng et al. (2011), has been adopted by several researchers (Zhang and Huang, 2013; Raknes and Arntsen, 2014; Yang et al., 2015a; Willemsen et al., 2016; Fu and Innanen, 2021) including a real data case in Yang et al. (2016). It can focus on reservoir changes and reduce artifacts outside the reservoir, and its result is not sensitive to the convergence degree of the inverted baseline model. Nevertheless, the DDS requires well-repeated time-lapse surveys, for instance, it is vulnerable to the non-repeatability of receiver/source positions. The common-model strategy (CMS), presented by Hicks et al. (2016), can also decay the artifacts caused by the divergence difference between baseline and monitor inversions (Fu and Innanen, 2022a). Its philosophy is to employ the same starting model, which has partially converged, for baseline

and monitoring inversions to guide them into close local minima. It has been applied in field cases in Hicks et al. (2016) and Bortoni et al. (2021). Moreover, Maharramov et al. (2016) present a joint method in which baseline and monitor models are simultaneously inverted; Zhou and Lumley (2021a) propose a central-difference strategy containing two sequential strategies, which assumes that the artifacts in the time-lapse changes, inverted from the two sequential strategies, have opposite polarities. However, of all the methods mentioned above, none has demonstrated the ability to solve the non-repeatability issue of seawater or near-surface property changes during baseline and monitor seismic surveys. To date, the capacity of time-lapse FWI technology to solve this issue has been demonstrated rarely.

Overall, accurate analysis of time-lapse seismic data remains a challenging task. It requires well-repeated time-lapse seismic surveys, including well-repeated acquisition geometry and equipment as well as well-repeated ambient noise. However, certain acquisition parameters are difficult to replicate, such as ambient noise, source parameters, and properties of the near-surface or seawater. In fact, it is hardly realistic to artificially control them. Therefore, it is necessary to improve the repeatability of time-lapse seismic data through appropriate processing. FWI, an advanced tool for time-lapse data analysis, still faces challenges regarding the convergence difference between baseline and monitor inversions. This discrepancy can significantly impact the effectiveness of time-lapse FWI when the repeatability of the data is inadequate. Factors contributing to the convergence difference in time-lapse FWI include nonrepeatable acquisition parameters, nonrepeatable near-surface or seawater properties, nonrepeatable inversion parameters, and biased starting models, among others. Additionally, enhancing the signal-to-noise ratio of time-lapse data poses another challenge. This is because seismic responses to subsurface changes over time are typically weak and more susceptible to noise interference compared to the baseline and monitor seismic data. Indeed, the subtle medium variations associated with enhanced oil recovery and/or CO_2 storage problems give uncertainty quantification within the inverse approaches height importance. Nevertheless, quantifying the uncertainty specifically for time-lapse FWI presents

an even greater challenge compared to the aforementioned difficulties.

1.2 Contributions

Solving all problems in time-lapse seismic analysis within a single thesis is not what we expect, and the contributions of this thesis will lay in the following points:

1 In Chapter 2, a time-lapse imaging approach that involves two frequency-domain matching filters as well as a new workflow is developed. The first filter requires source wavelet estimates from both baseline and monitoring data, while the second filter is source-independent but more sensitive to data noise. By applying these filters, we successfully reduce source wavelet non-repeatability. Additionally, we use a time-shift correction technique, utilizing a published fast local cross-correlation algorithm, to further minimize non-repeatability. Finally, we employ a reverse time migration algorithm with a Poynting vector imaging condition to generate depth images and address any remaining errors caused by the inaccuracy of the source-independent matching filter.

2 In Chapter 4, a stepsize-sharing time-lapse FWI strategy is designed to reduce artifacts caused by the variability of convergence in conventional strategies. The strategy is tested in five scenarios, including noise-free data, non-repeated noises, non-repeatable source positions, biased starting models, and a combination of the latter three. The comparisons between the new strategy and the conventional strategies show that the new method can adapt to all tested scenarios well. Especially, only the new strategy can provide meaningful results in the latter two scenarios when compared with others.

3 In Chapter 5, a Bayesian time-lapse FWI procedure, based on a Markov chain Monte Carlo (MCMC) algorithm, is formulated. The formulation employs several existing strategies, including the use of a double-difference time-lapse FWI, incorporation of time-domain multi-source data, and application of a local-updating target-oriented

inversion. However, it incorporates these within a stochastic framework, involving the computation of model covariance with an adaptive Metropolis algorithm, and a method to estimate data error statistics based on the features of time-lapse difference data is incorporated. A random walk Metropolis-Hastings MCMC is adopted for optimization.

1.3 Thesis overview

This thesis contains six chapters:

1 Chapter 1 is the introduction, which is to introduce the research status and challenges in time-lapse data analysis as well as the contributions and outlines of this thesis.

2 Chapter 2 is titled "Time-lapse seismic imaging using shot gathers with non-repeatable source wavelets". In which the research status and challenges in time-lapse prestack imaging are introduced, and a new time-lapse imaging approach is proposed and verified by numerical examples.

3 Chapter 3, entitled "Full-waveform Inversion", provides a comprehensive review of the theory behind FWI. The chapter covers key aspects such as "Objective Function Optimization", "Gradient Calculation", and "Line Search Technology".

4 Chapter 4, titled "Stepsize Sharing in Time-lapse Full-waveform Inversion", introduces the current research status and challenges in time-lapse FWI. The chapter presents a novel strategy, namely the stepsize-sharing approach, designed to mitigate artifacts resulting from convergence variability in conventional strategies.

5 Chapter 5, titled "A time-domain, multi-source Bayesian/Markov Chain Monte Carlo formulation of time-lapse seismic waveform inversion", introduces the current research status and challenges in FWI using stochastic global optimizations. The chapter presents a novel time-lapse strategy, namely a Bayesian FWI procedure for time-lapse data, based on a MCMC algorithm.

6 Chapter 6 includes conclusions for the whole thesis and future work.

Chapter 2

Time-lapse seismic imaging using shot gathers with non-repeatable source wavelets

2.1 Abstract

In time-lapse seismic applications, the signal produced by changes in the properties of subsurface rocks is generally obscured by noise associated with imperfect repeatability between surveys. A particularly important obstacle in the formation of time-lapse difference images is variation in the effective source wavelet between baseline and monitoring datasets. However, the partially separable influence of the wavelet within the Green's function model of seismic data permits two frequency-domain matching filters to be designed, which act to reduce source wavelet non-repeatability. One is based on the spectral ratio of the baseline and monitoring wavelets, and can be applied when prior estimates of the wavelets are available; the other is the average spectral ratio of the baseline and monitoring traces, and can be applied when prior estimates are unavailable. After balancing the datasets with one or other of these filters, we further prepare for the imaging step with time-shift corrections, using a published

fast local cross-correlations algorithm, preparing the difference data for use in an imaging algorithm. Reverse-time migration (RTM) is engaged for the imaging task, but we observe that residual repeatability errors tend to be magnified at this stage when source-normalized cross-correlation imaging conditions are used. Testing indicates that replacing this with a Poynting vector imaging condition strongly suppresses remaining errors in a robust manner. This includes stability within simulated data environments to both noise-free data and data with random noise, up to signal-to-noise ratios of roughly 1. Furthermore, our method illustrates better performance when compared with the conventional least-squares matching filter and common-depth-point-domain warping. At present, there is no common workflow for seismic imaging directly using time-lapse shot gathers. Our research contribution lies not only in the two matching filters, but also in a novel workflow for time-lapse seismic imaging.

2.2 Introduction

Time-lapse or 4D seismic technology is now a regular part of reservoir monitoring (e.g., enhanced oil recovery and/or CO_2 storage problems) and characterization for decades (Greaves and Fulp, 1987; Ross and Altan, 1997; Wang et al., 1998; Barkved et al., 2003; Arts et al., 2003; Barkved et al., 2005; Chadwick et al., 2009; Kazemeini et al., 2010; Pevzner et al., 2017), and has recently begun to be incorporated into reservoir development plans (Jack, 2017). If its requirements in terms of acquisition and processing can be met, time-lapse seismic can significantly impact the interpretability of detailed seismic information. In 3D or single-time seismic data, static geological information and dynamic fluid-flow information are mixed together, but in a time-lapse setting, which include two or more datasets obtained at different times, time-variant and time-invariant signal can be effectively separated, allowing fluid-flow changes to be directly sensed (Lumley, 2001). However, repeatability requirements are difficult to meet, and issues arising from data variations due to changes unrelated to the geological target zone remain subjects of research and acquisition/processing

workflow development.

Changes in the non-target geology between the times of a baseline and a monitoring seismic survey can be divided into roughly three categories. The first encompasses geological changes external to the target, for instance, reservoir subsidence, compaction, and porosity, or/and overburden pressure changes during hydrocarbon production and/or fluid injection (water, gas, steam, CO_2, etc.). Such changes have been reported, for instance, in the North Sea (Hicks et al., 2016) and in the Valhall Field (Hall et al., 2005). The second is the suite of differences due to experimental non-repeatability, i.e., variation in acquisition and processing between baseline and monitoring datasets. Acquisition differences can be environmental in origin (noise conditions, near-surface properties, water table, seawater properties, seal level, weather, etc.), geometrical (source and receiver positions, etc.), or be due to differences in acquisition equipment, affecting receiver response, and source wavelet. Processing non-repeatability can be caused by variation amongst processors, processing parameters, and workflows. The third involves changes that are due to the target medium property variations we are interested in, but which appear as data variations outside of our scheme for analysis. Most time-lapse monitoring estimation schemes are designed to make use of seismic amplitude changes in the reservoir zone, with some limited use of time shifts (travel time differences) between monitoring and baseline data (Hatchell and Bourne, 2005). Large time changes, and variations that, though due to local geological changes, are distributed globally throughout the data, e.g., via multiple scattering, are rarely accommodated.

The category of acquisition non-repeatability issues can be addressed to a degree by carrying out the acquisitions in the same season, and permanently installing receivers. However, processing and analysis methods that are specially designed to mitigate residual non-repeatability are nevertheless required. The most common approach is cross-equalization (Ross et al., 1996). For instance, Tucker et al. (2000) used a time-domain least-squares or optimum Wiener matching filter to implement cross-equalization; Al-Ismaili and Warner (2002) made use of artificial neural networks; Gallop (2011) apply midpoint matched filters.

Almutlaq and Margrave (2013) evaluated surface-consistent matching filters as a means to handle time-lapse seismic data with non-repeatable acquisition parameters. Bergmann et al. (2014) instead introduced a static correction method to reduce differences in reflection travel times. Hatchell and Tatanova (2019) concluded that a spectral balancing filter had an improved response to noise than a standard time-domain least-squares filter. Warping has been introduced (Rickett and Lumley, 2001) to colocate reflectors imaged at different positions, alleviating residual errors after cross-equalization. Hale (2006) and Hale (2013) developed fast local cross-correlations and dynamic warping for seismic image registration. Three-component apparent displacement vectors can also be estimated from time-lapse seismic images, using local phase correlations and a cyclic sequence of searches for correlation peaks (Hale, 2009). Fomel and Jin (2009) applied the local similarity attribute (Fomel, 2007) for time-lapse seismic image registration, an approach that has subsequently been built upon (Liu et al., 2021). Other warping approaches have been set out by Williamson et al. (2007), Phillips and Fomel (2016), Karimi et al. (2016), and Dramsch et al. (2019). In this paper, we adopt the fast local cross-correlation warping approach of Hale (2006) to correct the errors caused by time shifts in the time-lapse image.

Currently, the most common approach to reduce source wavelet non-repeatability is through cross-equalization. It comprises (Rickett and Lumley, 2001): (1) picking a window above the region of time-lapse change, (2) calculating a spectral balancing filter, or a least-squares time-domain filter, using the data in the window for each trace, and (3) applying each filter to the corresponding trace as a whole. Such processing can be applied to data in either shot gather or common depth point (CDP) gather form. Our approach has been designed in part to address issues with this approach, which include: the challenges of positioning the reservoir within the shot gathers; the fact that the window-picked signals can also provide information about the time-lapse change, for instance, the property change of overburden strata; the difficulty of removing residual wavelet non-repeatability after cross-equalization in the CDP domain, in which the wavelet is variable vertically (temporally) and

horizontally (spatially), and the problems of de-noising through trace-by-trace filtering.

Building on this suite of methods, we examine two new frequency-domain matching filters designed to effectively reduce source wavelet non-repeatability in the analysis of time-lapse shot gathers. The approaches are wave-equation-based, and build on a Green's function description of the seismic data. Having in this way reduced problems of source wavelet non-repeatability, we carry out a time-shift correction of remaining data non-repeatability, making use of a fast local cross-correlation algorithm (Hale, 2006). Finally, a reverse time migration (RTM) with a Poynting-vector imaging condition is examined as a means to further remove errors.

2.3 Theory

The aim of time-lapse seismic data processing is to suppress errors due to non-repeatability between baseline and monitoring surveys. We examine two processing approaches based on matched filters, which are specially designed to reduce errors due to source wavelet non-repeatability. After applying the filters, we correct for local time shifts between two datasets using a fast local cross-correlation algorithm. Finally, after decreasing the time shifts between two datasets, an RTM in depth with a Poynting vector imaging condition is applied to image the difference data.

2.3.1 Filter design for suppression of source wavelet non-repeatability

We consider the 2D constant-density acoustic wave equation:

$$\frac{\partial^2 P(x,z,t)}{\partial x^2} + \frac{\partial^2 P(x,z,t)}{\partial z^2} - \frac{1}{c^2(x,z)} \frac{\partial^2 P(x,z,t)}{\partial t^2} = w(t)\delta(x-x_0)\delta(z-z_0), \tag{2.1}$$

where $P(x,z,t)$ is the wavefield depending on coordinates (x,z) and time t, $c(x,z)$ is the P-wave velocity field, and $w(t)$ is the time-dependent source wavelet. The corresponding

Green's function $G(x, z, t)$ satisfies:

$$\frac{\partial^2 G(x,z,t)}{\partial x^2} + \frac{\partial^2 G(x,z,t)}{\partial z^2} - \frac{1}{c^2(x,z)} \frac{\partial^2 G(x,z,t)}{\partial t^2} = \delta(t)\delta(x - x_0)\delta(z - z_0), \tag{2.2}$$

which we estimate through a time-domain finite-difference method in this study. Seismic data (i.e., evaluation of the wavefield P over a set of receivers) can be expressed as convolutions of the Green's function and the source wavelet:

$$d_{i,j}(t) = w_i(t) * G_{i,j}(t), \tag{2.3}$$

where $*$ denotes the convolution operator, subscripts i and j are shot number and receiver number, respectively, $d_{i,j}(t)$ is the observed seismic trace associated with shot i at receiver j, $G_{i,j}(t)$ is the corresponding Green's function, and $w_i(t)$ is the source wavelet associated with shot i, which is assumed to be the same for all traces in the same shot. For a time-lapse survey, let the additional subscripts 1 and 2 represent baseline and monitoring surveys, respectively. The baseline seismic trace can then be represented as:

$$d_{1i,j}(t) = w_{1i}(t) * G_{1i,j}(t), \tag{2.4}$$

and the monitoring seismic trace:

$$d_{2i,j}(t) = w_{2i}(t) * G_{2i,j}(t). \tag{2.5}$$

If the baseline wavelet $w_{1i}(t)$ is equal to the monitoring wavelet w_{2i}, differences between baseline and monitoring datasets arise due to differences within the Green's functions only, which we assume means that they are caused by changes to the physical properties of the subsurface. If $w_{1i}(t)$ is different from w_{2i}, data differences will then be mixtures of those caused by both source wavelet non-repeatibility and those caused by subsurface property

changes. The double-wavelet method approach described by Fu et al. (2020, 2023), in which effective baseline and monitoring datasets, so constructed to have a common wavelet, $w_{1i}(t) * w_{2i}(t)$, is implemented by convolving the baseline wavelet with the monitoring trace and vice versa. However, this method requires two wavelet estimates as input. Although such estimates can usually be constructed, the uncertainty caused by inaccurate estimates is difficult to quantify.

Here we seek to construct an alternative approach with minimal requirements for a priori source wavelet knowledge. The new method is formulated in the frequency domain. We Fourier transform equations 2.4 and 2.5 after which for the baseline seismic trace we have:

$$\hat{d}_{1i,j}(\omega) = \hat{w}_{1i}(\omega)\hat{G}_{1i,j}(\omega), \tag{2.6}$$

and for the monitoring seismic trace:

$$\hat{d}_{2i,j}(\omega) = \hat{w}_{2i}(\omega)\hat{G}_{2i,j}(\omega), \tag{2.7}$$

where ω is the angular frequency and the $\hat{\cdot}$ denotes the Fourier transform. If we had in hand the filter

$$f_w(\omega) = \hat{w}_{1i}(\omega)/\hat{w}_{2i}(\omega), \tag{2.8}$$

then after application to $\hat{d}_{2i,j}(\omega)$, the filtered monitoring trace would be

$$\hat{d}'_{2i,j}(\omega) = f_w(\omega)\hat{d}_{2i,j}(\omega) = \hat{w}_{1i}(\omega)\hat{G}_{2i,j}(\omega), \tag{2.9}$$

i.e., it would share the same medium information as the original monitoring trace but it would have the same wavelet as the baseline trace. The construction of a matching filter that has this effect without requiring the direct use of the two source wavelet spectra is the goal. We design this source-independent matching filter, $f_s(\omega)$, from the average spectra of

the ratios of the baseline and monitoring traces:

$$f_s(\omega) = \frac{1}{N} \sum_{i=1}^{ns} \sum_{j=1}^{nr} \frac{\hat{d}_{1i,j}(\omega)}{\hat{d}_{2i,j}(\omega)}, \tag{2.10}$$

where $N = ns \times nr$ is the total trace number of all shots in each dataset (baseline or monitoring), and ns and nr is the numbers of shots and receivers, respectively. Assuming that traces from the same shot record have the same source wavelet, we replace the subscript pair (i, j) with the single index k and suppress the shot subscripts, giving

$$f_s(\omega) = \frac{1}{N} \sum_{k=1}^{N} \frac{\hat{d}_{1k}(\omega)}{\hat{d}_{2k}(\omega)} = \frac{1}{N} \sum_{k=1}^{N} \frac{\hat{w}_1(\omega)\hat{G}_{1k}(\omega)}{\hat{w}_2(\omega)\hat{G}_{2k}(\omega)}. \tag{2.11}$$

We argue that $f_s(\omega) \approx f_w(\omega)$. To make this case, let $\delta\hat{G}_k = \hat{G}_{1k} - \hat{G}_{2k}$ represent time-lapse changes within the Green's function (i.e., to parts of the wavefield connected to medium property changes). Equation 2.11 can be re-written

$$f_s(\omega) = \frac{1}{N} \sum_{k=1}^{N} \frac{\hat{w}_1(\omega)\left[\hat{G}_{2k}(\omega) + \delta\hat{G}_k(\omega)\right]}{\hat{w}_2(\omega)\hat{G}_{2k}(\omega)}. \tag{2.12}$$

Re-arranging, we obtain

$$f_s(\omega) = \frac{\hat{w}_1(\omega)}{\hat{w}_2(\omega)} \left(1 + \frac{1}{N} \sum_{k=1}^{N} \frac{\delta\hat{G}_k(\omega)}{\hat{G}_{2k}(\omega)}\right), \tag{2.13}$$

where $\frac{\delta\hat{G}_k(\omega)}{\hat{G}_{2k}(\omega)}$ is the relative, or fractional change in the Green's function, and $\frac{1}{N}\sum_{k=1}^{N} \frac{\delta\hat{G}_k(\omega)}{\hat{G}_{2k}(\omega)}$ is the average of all of these changes. At this point it is clear that if the average of the relative change in the Green's function is small,

$$f_s(\omega) \approx \frac{\hat{w}_1(\omega)}{\hat{w}_2(\omega)} = f_w(\omega), \tag{2.14}$$

and our case is made. We examine this quantity numerically. In Figure 2.1, the monitoring

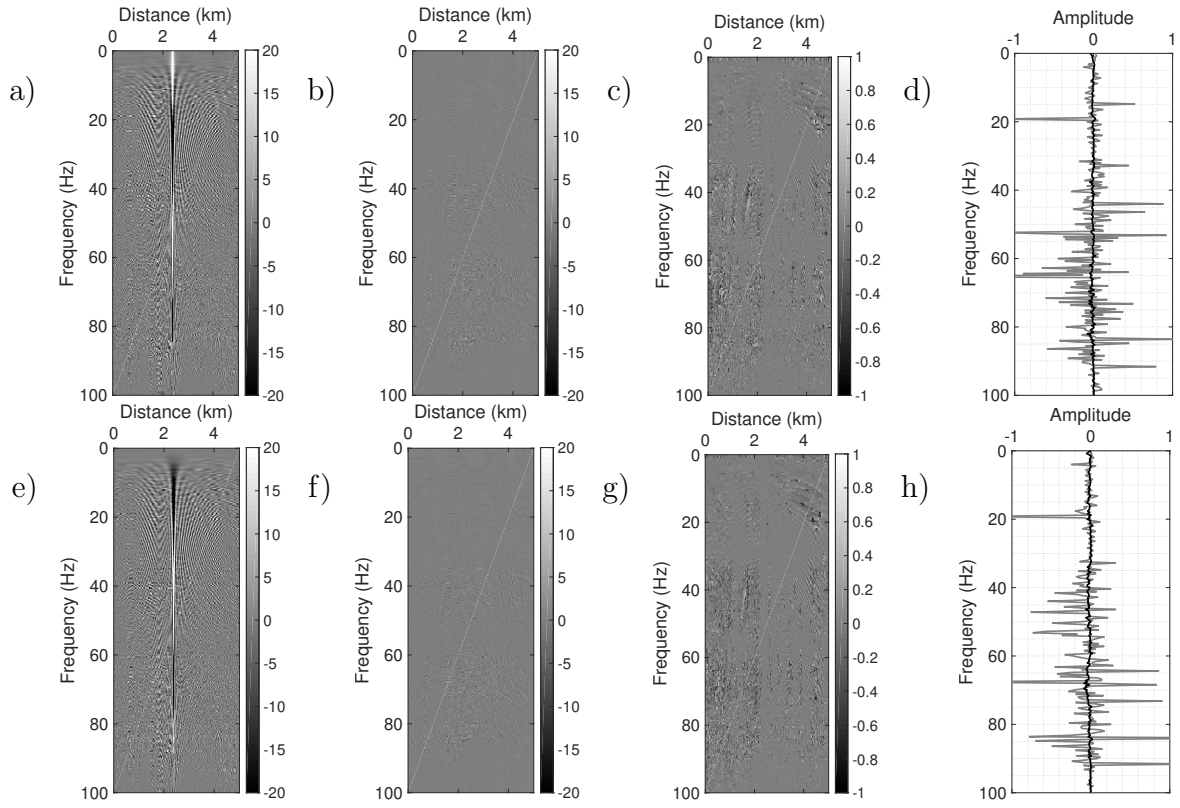

Figure 2.1: The real parts of the monitoring Green's function (a), the Green's function difference (b), the relative Green's function difference (c), and the average of relative Green's function differences (the black line in d). Correspondingly, the imaginary parts are plotted in e, f, g, and h. The gray lines in d and h are extracted from c and g, respectively, at the distance of 2km.

Green's function $\hat{G}_{2k}(\omega)$, the change in the Green's function, the relative change in the Green's function, and its average, are plotted. Models and acquisition parameters used are the same as those in Figure 2.3 and 2.5. We observe in Figure 2.1c and d (or g and h) that the relative change cannot be ignored within individual traces, as its amplitudes can exceed 1; however, the average values (Figure 2.1d and h) are all significantly smaller, and in fact tend very close to zero, as the variations tend to be incoherent. Provided this incoherency dominates, which we will now assume, the second term in equation 2.13 can be safely neglected.

An updated monitoring dataset is obtained by multiplying $f_s(\omega)$ with $\hat{d}_{2k}(\omega)$, frequency by frequency, and trace by trace (the filter is the same for all traces). Here we are treating

the baseline data as the reference data, but this could equally well have been the monitoring data; if either dataset has a narrower frequency bandwidth, selecting that as the reference will tend to avoid errors arising from frequency expansion. During the calculation of the filter, in practice, a damping factor is generally needed for stability, i.e., to avoid situations where the denominator in equation 2.10 becomes zero. The adapted monitoring data are finally transformed back to the time domain to complete the process.

2.3.2 Fast local cross-correlation for time-shift correction

Supposing at this stage that source wavelet non-repeatability has been adequately suppressed, we next add a component to processing designed to eliminate remaining time-shift errors between baseline and monitoring data. This is based on local cross-correlations, which extract the time shift for each sample.

Let $d_{1k}[i]$ be a discrete sample of the time domain baseline trace $d_{1k}(t)$ at time $t = (i-1)\Delta t$, with i running from 1 to Nt and Δt being the sample interval. With the processed monitoring trace likewise discretized, the cross-correlation of the two traces is

$$c_k[l] = \sum_{i=1}^{Nt} d_{1k}[i] d'_{2k}[i+l]. \tag{2.15}$$

By calculating within Gaussian windows Hale (2006) produced a local cross-correlation for each sample:

$$C_k[i,l] = \sum_{j=i-Nw/2}^{i+Nw/2} d_{1k}[j] f(j-i) d'_{2k}[j+l] f(j-i+l), \tag{2.16}$$

where $f(x) \equiv e^{-x^2/2\sigma^2}$, and Nw denotes the number of non-zero samples in a truncated Gaussian window. Hale (2006) carried out the calculation with a fast algorithm which we adopt for its significant reduction of computation time. The algorithm is set out in

Algorithm 1 (for details, see Hale, 2006).

Algorithm 1: Fast local cross-correlations (Hale, 2006)

```
for k ← 1, 2, ..., N do
    for l ← −Nl/2, ..., Nl/2 do                    /* Nl is the number of lags */
        for i ← 1, 2, ..., Nt do
            h[i] ← d_1k[i] × d'_2k[i + l]
        end
        for i ← 1, 2, ..., Nt do                                  /* begin shift */
            h̃[i] ← h[i − l/2]                          /* interpolate for odd l */
        end                                                        /* end shift */
        for i ← 1, 2, ..., Nt do                      /* begin Gaussian filter */
            C_k[i; l] ← 0
            for j ← i − Nv/2, ..., i + Nv/2 do                  /* Nv ≈ Nw/√2 */
                C_k[i; l] ← C_k[i; l] + h̃[j] × f(j − i − l/2) × f(j − i + l/2)
            end                                          /* end Gaussian filter */
        end
    end
end
```

After obtaining local cross-correlations for each sample in each trace, we extract the maximum local cross-correlation in each sample, via:

$$C_{max}[i,k] = \max_{l \in [-Nl/2 \;\; Nl/2]} C_k[i,l], \tag{2.17}$$

and pick the corresponding lag l_{max} (Nl is the total number of lags). The adjusted monitoring trace, after time-shift correction, is then

$$d''_{2k}[i] = d'_{2k}[i + l_{max}]. \tag{2.18}$$

2.3.3 Reverse time migration with a Poynting vector imaging condition

To create interpretable outputs from the time lapse data, we apply a reverse-time migration (RTM) algorithm to map the shot gathers into a reflector image in depth. In RTM, imaging conditions based on crosscorrelation, or source-normalized cross-correlation (Claerbout, 1971; Lee et al., 1991; Kaelin and Guitton, 2006), the latter of which produces images with the same (dimensionless) unit, scaling, and sign as the reflection coefficient (Chattopadhyay and McMechan, 2008), are most often selected. To within a linear approximation, difference data can be used as input to RTM in order to produce difference images (Innanen et al., 2014). However, RTM image artifacts caused by events which are improperly managed with the imaging condition, for instance refracted arrivals, are magnified in difference data. Although in principle these arrivals are suppressed in differencing when time-lapse repeatability is improved, because they tend to be of large amplitude, residual amplitudes tend to produce large artifacts, especially when compared with the small-amplitude reservoir changes of interest. To mitigate these issues, we employ a Poynting vector imaging condition (Yoon and Marfurt, 2006):

$$\text{image}(x,z) = \frac{\sum_t P_f(x,z,t) P_b(x,z,t) W(\cos\theta(x,z,t))}{\sum_t P_f(x,z,t) P_f(x,z,t)}, \tag{2.19}$$

where x and z are horizontal and depth coordinates respectively, $P_f(x,z,t)$ is the forward-propagated wavefield from the source location, $P_b(x,z,t)$ is the backward-propagated (difference) wavefield in reverse time from receiver locations, and where

$$\cos\theta(x,z,t) = \frac{\mathbf{v}_f P_f(x,z,t) \cdot \mathbf{v}_b P_b(x,z,t)}{|\mathbf{v}_f P_f(x,z,t)||\mathbf{v}_b P_b(x,z,t)|}, \tag{2.20}$$

with

$$\mathbf{v}_f = -\left(\frac{\partial P_f}{\partial x}, \frac{\partial P_f}{\partial z}\right)\frac{\partial P_f}{\partial t},\ \mathbf{v}_b = -\left(\frac{\partial P_b}{\partial x}, \frac{\partial P_b}{\partial z}\right)\frac{\partial P_b}{\partial t}, \tag{2.21}$$

in which $\mathbf{v}_f P_f$ and $\mathbf{v}_b P_b$ are the Poynting vectors associated with forward- and backward-propagating wavefields, respectively, $\theta(x,z,t)$ is the opening angle, $W(\cos\,\theta(x,z,t))$ is a weighting factor to suppress image contributions within a specified angle range. In the latter case, for example, if we want to keep contributions from incident angles below 50° (i.e., 100° opening angle), the weighting $W(\cos\theta(x,z,t))$ can be set to 1 for $\cos\theta(x,z,t) < \cos(100°)$, and 0 otherwise. We will compare result of this imaging condition to that of the source-normalized condition, which is of the form

$$\text{image}(x,z) = \frac{\sum_t P_f(x,z,t) P_b(x,z,t)}{\sum_t P_f(x,z,t) P_f(x,z,t)}. \tag{2.22}$$

In equations 2.20 and 2.21, we use the baseline or monitoring data to obtain the backward-propagating wavefield, instead of the difference data used in equations 2.19 and 2.22, because the difference data have much lower SNR (signal-to-noise ratio) than baseline or monitoring data, which will lower the quality of calculated incident angles.

The theory section has been summarized as a workflow, illustrated in Figure 2.2.

2.4 Numerical examples

In this section, a modified P-wave Marmousi model and a 2D constant-density acoustic wave equation solved by a finite difference method are used to evaluate the approach numerically. A perfectly matched layer (PML) is employed on each side of the model, except the top side on which a free surface is set up. In Figure 2.3, the baseline model, time-lapse model (monitoring model minus baseline model), smoothing migrated velocity, and geometry are plotted. The model size is 257-by-502, with a 10m grid spacing, with 25 sources evenly distributed across the top of the velocity model (at 30m depth), and 502 receivers are located at the grid points on the surface. A time-lapse velocity change, 4% of the corresponding reservoir velocity, is placed at the center in the baseline model.

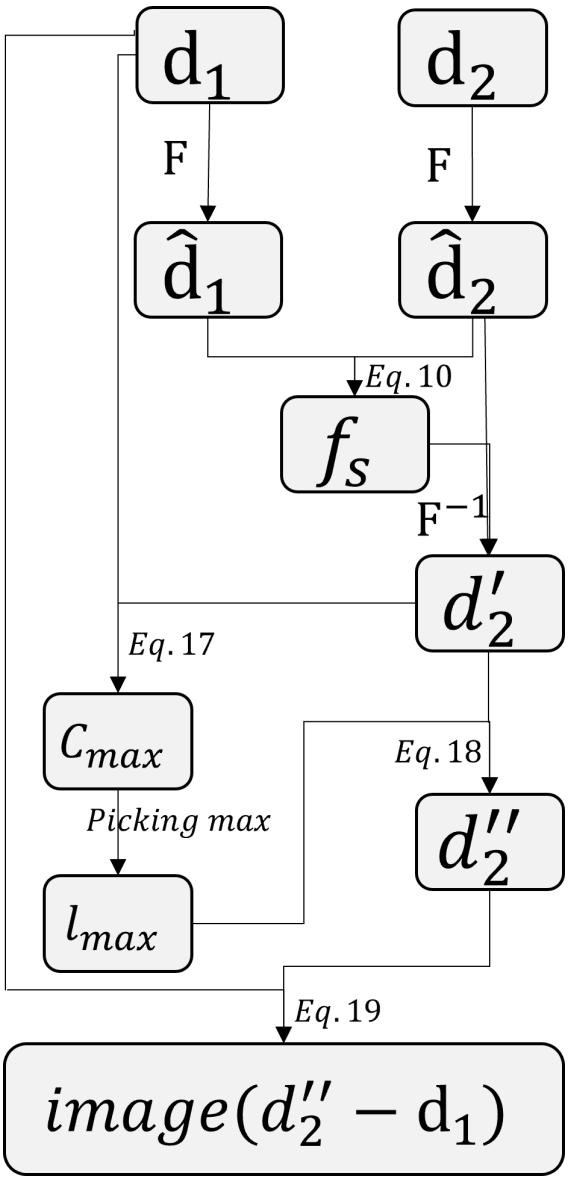

Figure 2.2: Time-lapse seismic imaging workflow for shot gathers with non-repeatable source wavelets. Note that in the workflow, F denotes Fourier transform, and F^{-1} denotes inverse Fourier transform.

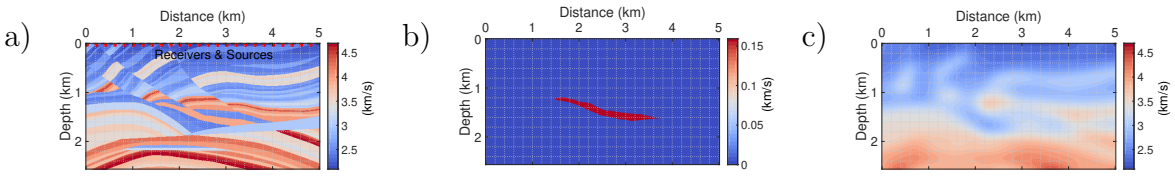

Figure 2.3: (a) Baseline model (P-wave velocity); (b) time-lapse model, monitor model minus baseline model; (c) smoothing migrated velocity model. The dash lines and asterisks in (a) are locations of receivers and sources, respectively.

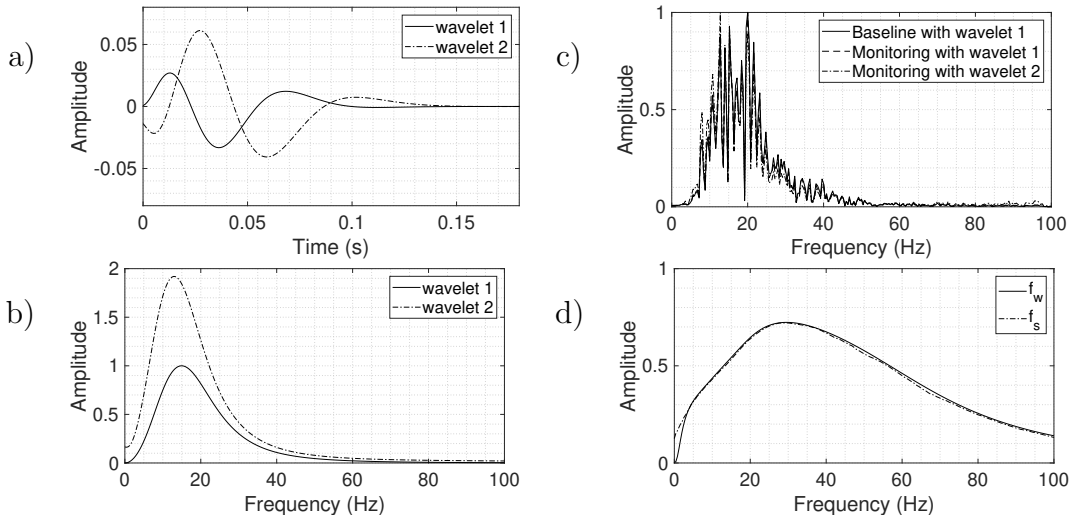

Figure 2.4: Two time-domain wavelets (a) and their frequency spectra (b). (c) Normalized frequency spectra of a certain baseline and monitoring traces. (d) Calculated frequency-domain filters, $f_w(\omega)$ and $f_s(\omega)$.

2.4.1 Single-shot gather example

A single shot gather associated with a source at the center-top of the model is used first Figure 2.3a. Two different minimum-phase wavelets are introduced to simulate non-repeatability (see Figures 2.4a-b). Wavelet 1 has a 15Hz peak frequency; wavelet 2 has a 13Hz peak frequency, and, relative to wavelet 1, has double the amplitude and a -90° phase rotation. The two wavelets are therefore different in peak frequency, maximum amplitude, and phase. For benchmarking, we will generate images derived using identical baseline and monitoring wavelets. For these, we use wavelet 1 for both baseline and monitoring surveys, with wavelet 2 introduced subsequently to test our approach.

Normalized frequency spectra from selected baseline and monitoring traces are plotted in Figure 2.4c. We observe that the discrepancies between these spectra are strongly controlled by the repeatability of the source wavelets. The filters $f_w(\omega)$ and $f_s(\omega)$ calculated from equation 2.10 are plotted in Figure 2.4d. They are similar, except for frequencies below about 3Hz, where our neglect of the average term in equation 2.13 is not justified. This can also be observed in Figure 2.1d-h. This low-frequency error has minimal impact, as

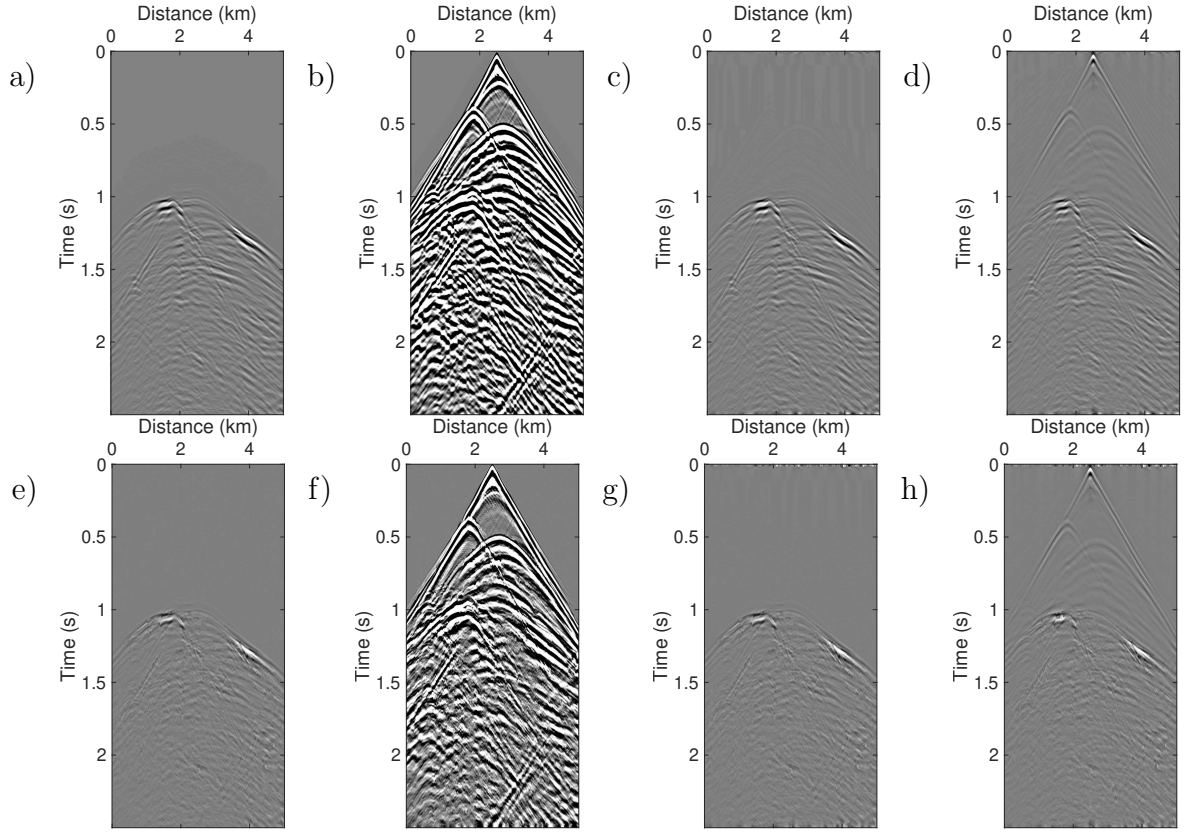

Figure 2.5: The difference data before (a,b,c,d) and after (e, f, g, h) the time-shift correction. (a) Baseline and monitoring wavelets (wavelet 1) are identical. (b) Baseline and monitoring wavelets are different (wavelet 1 for baseline data, wavelet 2 for monitoring data), and there is no processing to both datasets. (c) Baseline and monitoring wavelets are different, but the monitoring data are filtered by $f_w(\omega)$. (d) Baseline and monitoring wavelets are different, but the monitoring data are filtered by $f_s(\omega)$. (e)-(h) are the same as (a)-(d) but having implemented time-shift corrections to monitoring data. All panels are clipped identically.

frequencies below 3Hz are typically weak in seismic data.

In Figure 2.5, the difference data are plotted. In Figure 2.6 and 2.7, traces at a lateral position of 3km, extracted from each panel of Figure 2.5, are plotted. From the data plotted in Figures 2.5-2.7, we observe that $f_w(\omega)$ and $f_s(\omega)$ both effectively eliminate differences caused by non-repeatable wavelets, and time-shift corrections based on the fast local cross-correlations can effectively attenuate errors resulting from the time shifts between baseline and monitoring data, but no particular difference is detected when the wavelets are non-repeatable. Similar conclusions are arrived at by analysis of Figures 2.8 and 2.9, in which the

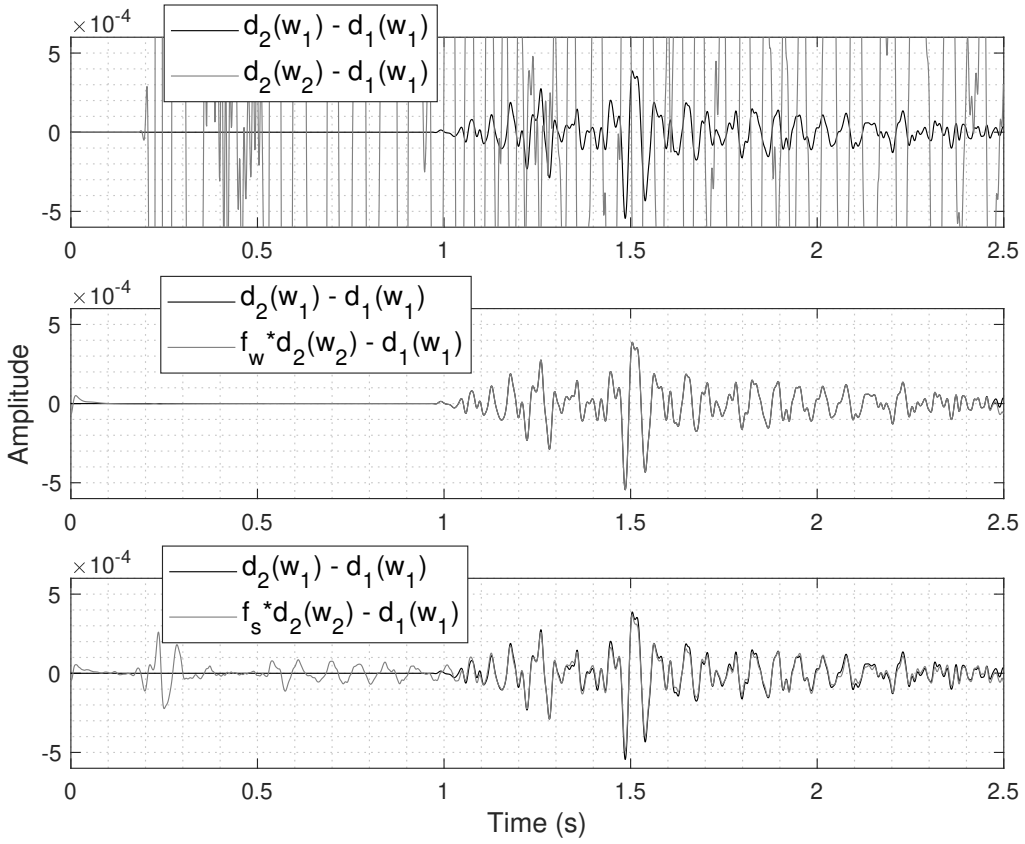

Figure 2.6: Traces extracted from Figure 2.5a-d at the distance of 3km. All black lines are identical, which are extracted from Figure 2.5a. The gray lines in the first to the third panels are, respectively, extracted from Figure 2.5b, c, and d.

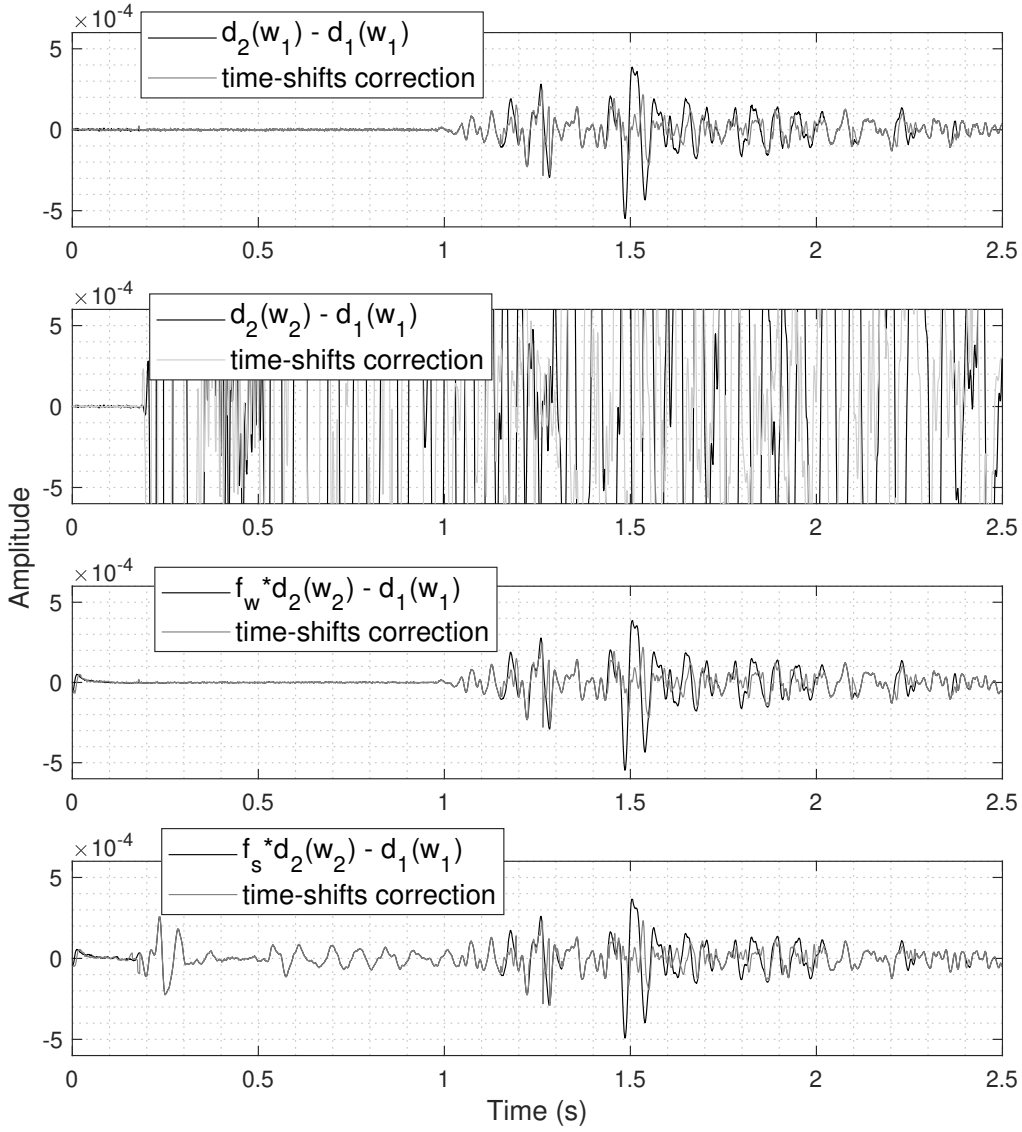

Figure 2.7: Traces extracted from Figure 2.5a-h at the distance of 3km. The black lines in the first to the fourth panels are, respectively, abstracted from Figure 2.5a-d. The gray lines in the first to the fourth panels are, respectively, abstracted from Figure 2.5e-h.

difference data in Figures 2.5a and h are imaged with the source-normalized RTM algorithm.

In Figures 2.8 and 2.9, we furthermore observe that the application of $f_w(\omega)$ leads to images that are almost the same as those derived from perfectly repeatable wavelets. The application of $f_s(\omega)$, however, leads to images with noticeable errors, mainly above the region of time-lapse change (see Figure 2.5d and the third panel of Figure 2.6). However, images generated using the Poynting vector imaging condition (Figure 2.8e) is effective at suppressing these types of error. RTM images after time-shift corrections are plotted in Figure 2.9. Here we observe that time-shift corrections can reduce artifacts, especially those appearing immediately beneath the region of time-lapse change. The three interface images in Figure 2.8 drop to two in Figure 2.9, and these exhibit the expected opposing polarities and similar amplitudes of the time lapse reflection coefficients.

Remaining errors are still visible outside of the region of time-lapse change in 2.9c and e, due to (1) remaining time shifts, (2) linearization errors associated with the imaging of difference data (e.g., Innanen et al., 2014), (3) inaccuracies in the velocity model used for RTM imaging, and (4) inaccuracies in $f_s(\omega)$, in the case of the result in Figure 2.9e.

2.4.2 Multiple shot gather example

In this section we generate similarly-designed examples using all shots, including 25 baseline shots and 25 monitoring shots, and include both noisy and noise-free data.

Noise-free data tests

In Figure 2.10, the source-independent matching filters $f_s(\omega)$, calculated twice, once from the single central shot record and once from the full suite of noise-free shot records, are plotted in comparison with $f_w(\omega)$. We observe that the two $f_s(\omega)$ are very similar, both close to $f_w(\omega)$ but with a bias when the frequency drops below 3Hz.

In Figure 2.11 and 2.12, images formed from the difference data before (Figure 2.11) and after (Figure 2.12) time-shift corrections, using all shots, both with the source-normalized

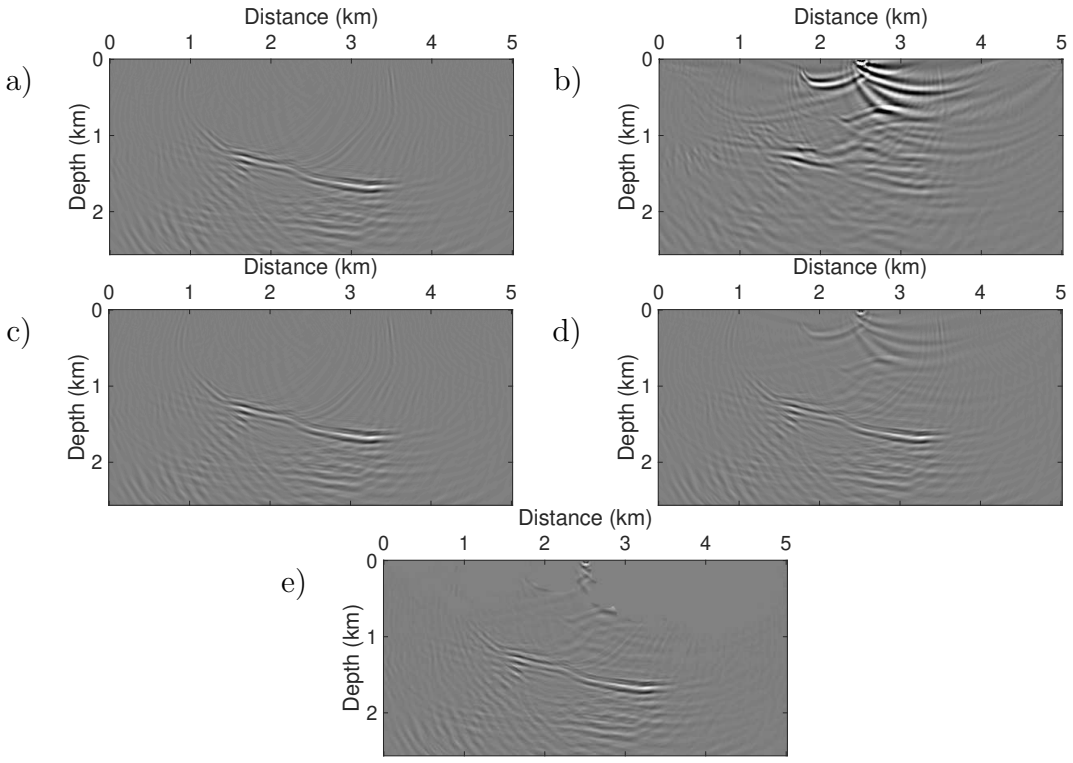

Figure 2.8: RTM images produced using the difference data in Figure 2.5a-d, prior to time-shift correction. (a)-(d) Images formed from the data in Figures 2.5a-d, respectively, using the source-normalized imaging condition (equation 2.22); (e) image formed from the data in Figure 2.5d, using the Poynting vector imaging condition (equation 2.19), with incident angles limited to within 50°. The correct image should contain two reflection events with opposite polarities and similar reflection coefficients.

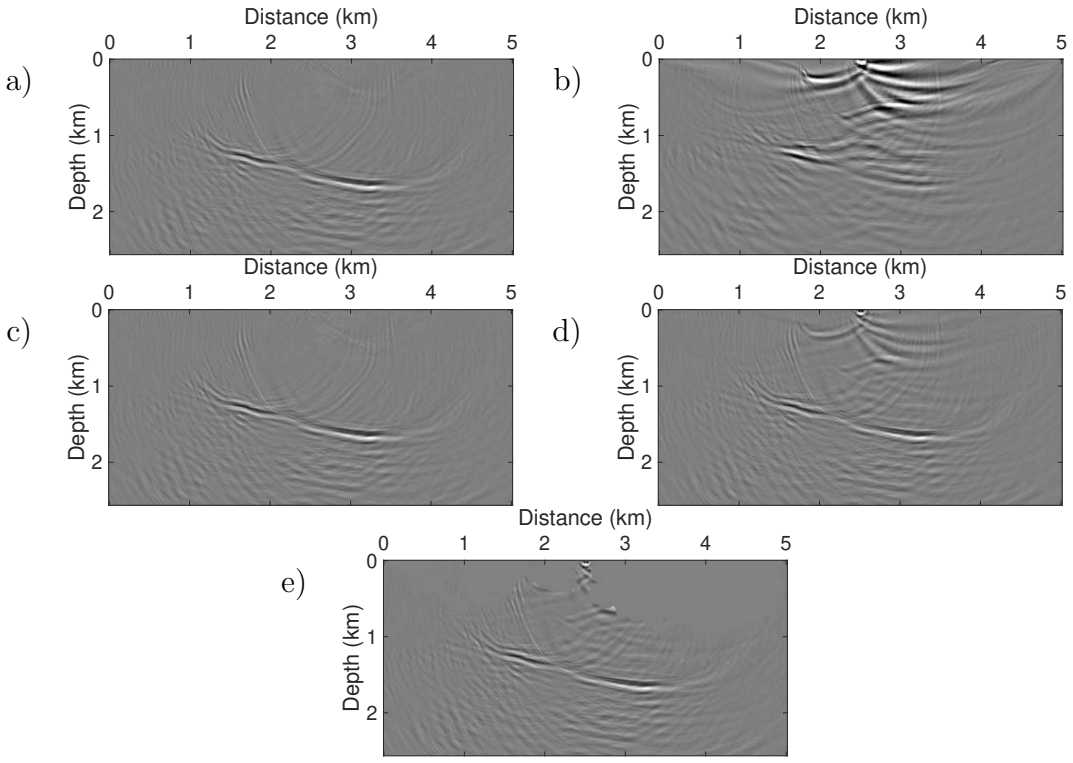

Figure 2.9: RTM images produced using the difference data in Figure 2.5e-h, after time-shift correction. (a)-(d) Images formed from the data in Figures 2.5e-h, respectively, using the source-normalized imaging condition (equation 2.22); (e) image formed from the data in Figure 2.5h, using the Poynting vector imaging condition (equation 2.19), with incident angles limited to within 50°. Except (b), all images contain two reflection events with opposite polarities and similar reflection coefficients, which are close to the real situation.

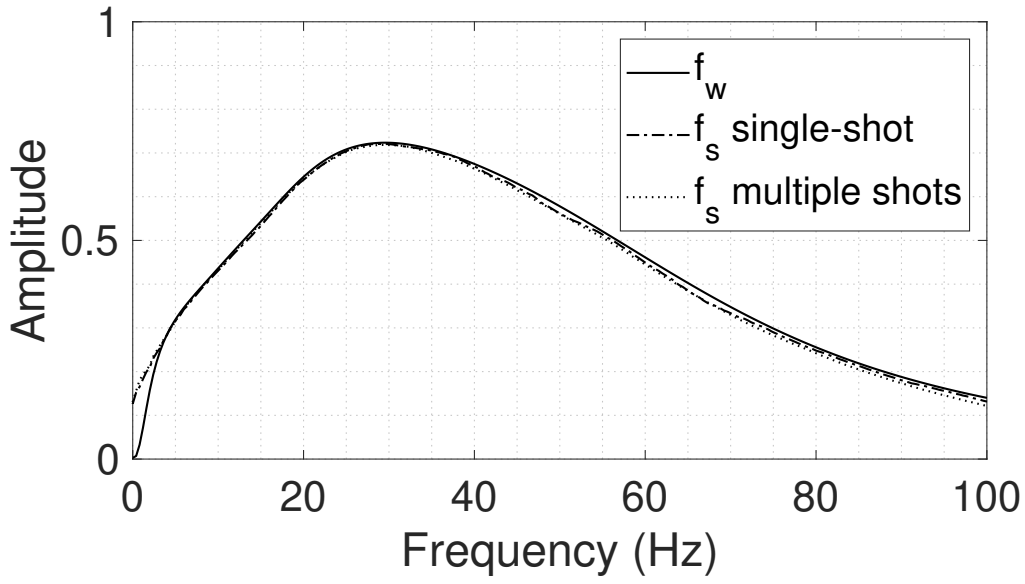

Figure 2.10: The solid and dash lines are the same as that in Figure 2.4d, the former is calculated from the wavelets spectra and the latter is calculated from the single-shot gathers. The dot line is $f_s(\omega)$ calculated from all shots.

imaging condition and with the Poynting vector imaging condition, are plotted. Similarly our observations regarding Figures 2.8 and 2.9, comparable improvements are produced by application of $f_w(\omega)$ and $f_s(\omega)$. We further observe that time-shift corrections again produce images with the expected two polarity-reversed interfaces in the region of time-lapse change. The Poynting vector imaging condition suppresses artifacts above the time-lapse area, especially up shallow.

Noisy data tests

We next examine the response of the time lapse imaging test with Gaussian random noise selected such that the signal-to-noise ratio (SNR) is 2, and then again with SNR=1. Different noise traces are added to the noise-free baseline and monitoring data. The resulting baseline shot records (for the central shot point) and corresponding difference data are plotted in Figure 2.13.

In Figure 2.14, average frequency spectra from the central shot record are plotted for the

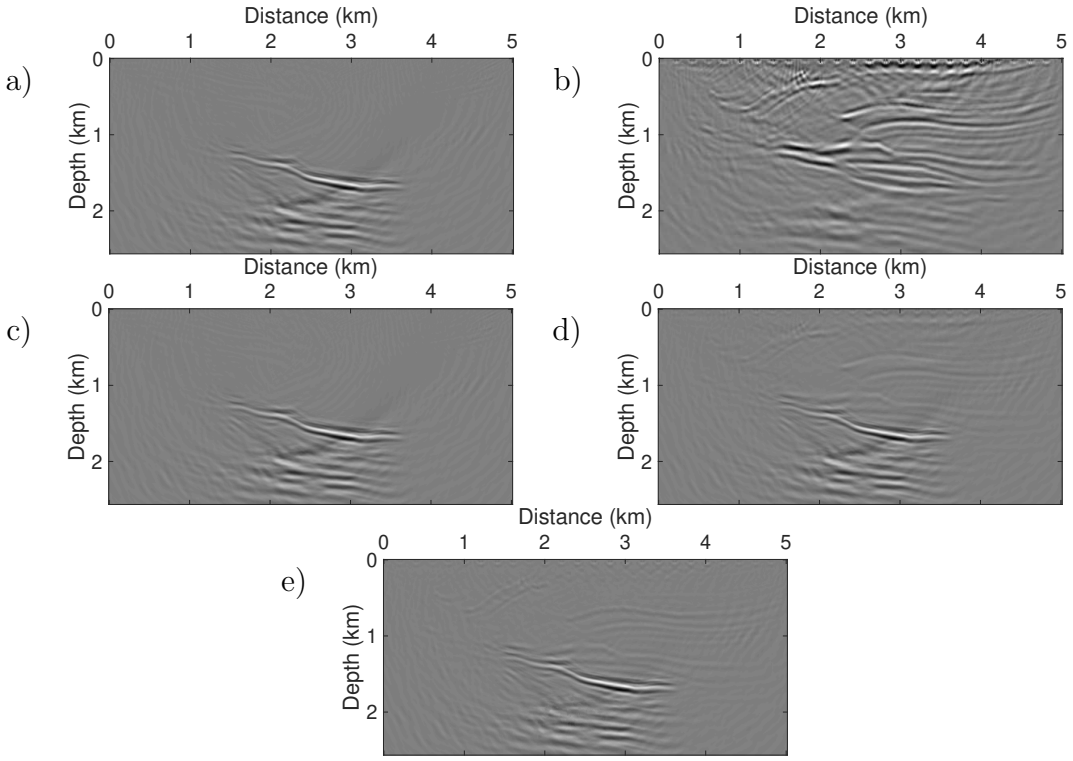

Figure 2.11: RTM images produced using difference data of all 25 shots, prior to time-shift correction. (a)-(d) Images formed from the difference data by using the source-normalized imaging condition (equation 2.22); (e) image formed from the data in Figure 2.5h, using the Poynting vector imaging condition (equation 2.19), with incident angles limited to within 50°. (a) Baseline and monitoring wavelets are identical. (b) Baseline and monitoring wavelets are different, and monitor data are not processed. (c) Baseline and monitoring wavelets are different, and monitor data are filtered by f_w. (d) and (e) Baseline and monitoring wavelets are different, and monitor data are filtered by f_s. The correct image should contain two reflection events with opposite polarities and similar reflection coefficients.

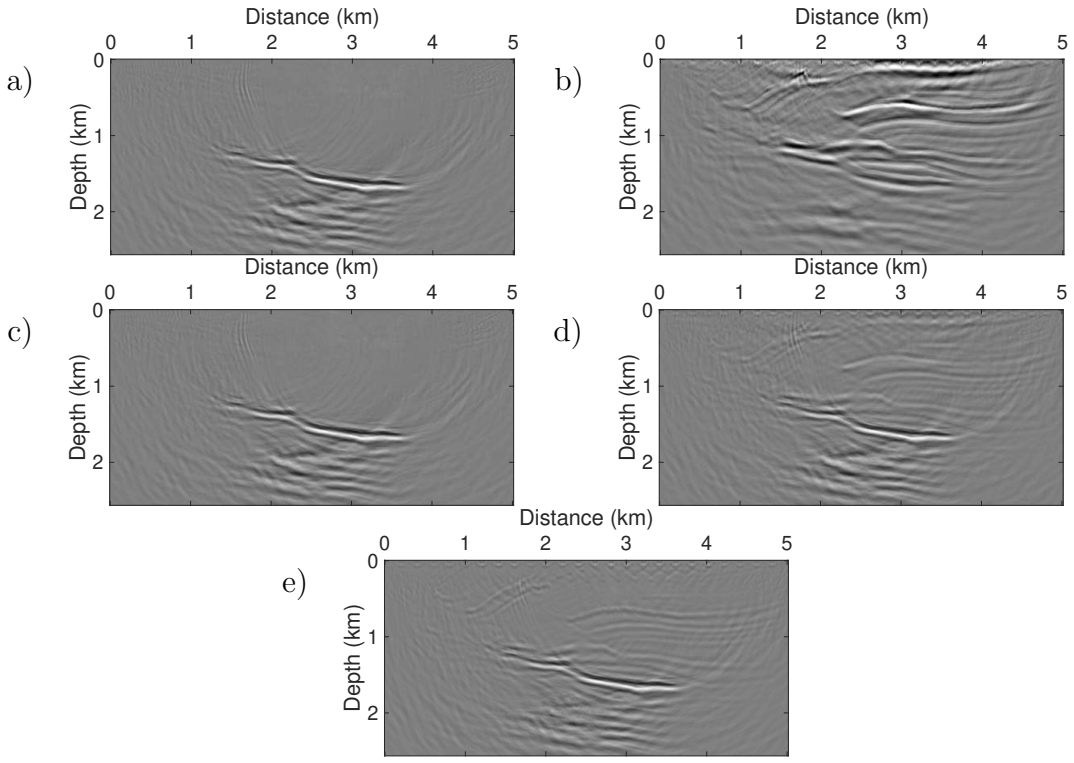

Figure 2.12: RTM images produced using difference data of all 25 shots, after time-shift correction. (a)-(d) Images formed from the difference data by using the source-normalized imaging condition (equation 2.22); (e) image formed from the data in Figure 2.5h, using the Poynting vector imaging condition (equation 2.19), with incident angles limited to within 50°. (a) Baseline and monitoring wavelets are identical. (b) Baseline and monitoring wavelets are different, and monitor data are not processed. (c) Baseline and monitoring wavelets are different, and monitor data are filtered by f_w. (d) and (e) Baseline and monitoring wavelets are different, and monitor data are filtered by f_s. Except (b), all images contain two reflection events with opposite polarities and similar reflection coefficients, which are close to the real situation.

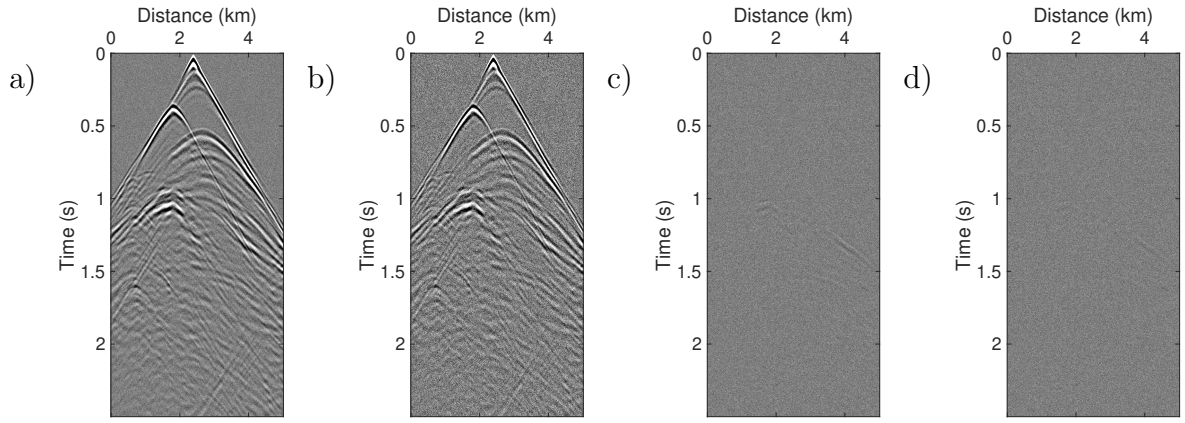

Figure 2.13: The central shots of baseline data with SNR=2 (a) and SNR=1 (b). The difference data of the central shots with SNR=2 (c) and SNR=1 (d).

noise-free and noisy baseline cases respectively, as well as the corresponding matching filters ($f_w(\omega)$ and $f_s(\omega)$). As the noise level increases, the accuracy of the source-independent matching filter ($f_s(\omega)$) degrades. It is most accurate near the peak frequency (10-20Hz), where the signal amplitude most strongly outweighs that of the noise.

RTM images formed from the noisy difference data, including all 25 shots after time-shift correction, using first the source-normalized imaging condition (equation 2.22) and second the Poynting vector imaging condition (equation 2.19) are plotted in Figures 2.15 and 2.16. For the dataset with SNR=2, we still clearly recognize the time-lapse region from the images constructed using $f_s(\omega)$. This is no longer true for the data with SNR=1. In comparison, $f_w(\omega)$ remains stable for across both noise levels. In fact, $f_w(\omega)$ functions as a band-pass filter, to some degree reducing the noise in the final image. The Poynting vector imaging condition suppresses coherent errors in the up-shallow regions within the images formed with noisy data.

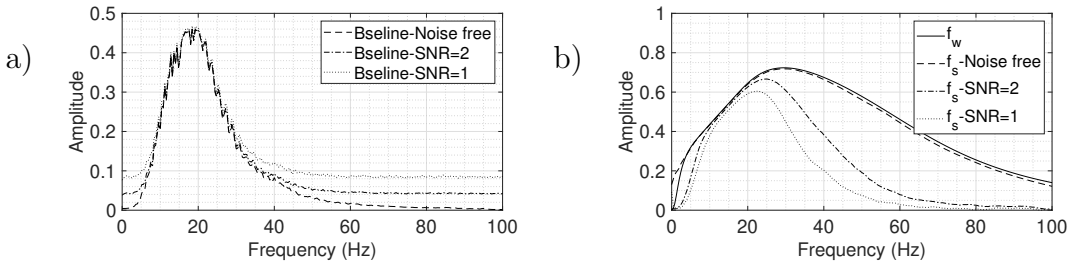

Figure 2.14: (a) Average frequency spectra of the central shot of noise-free and noisy baseline data.(b) Calculated matching filters for noise-free and noisy data.

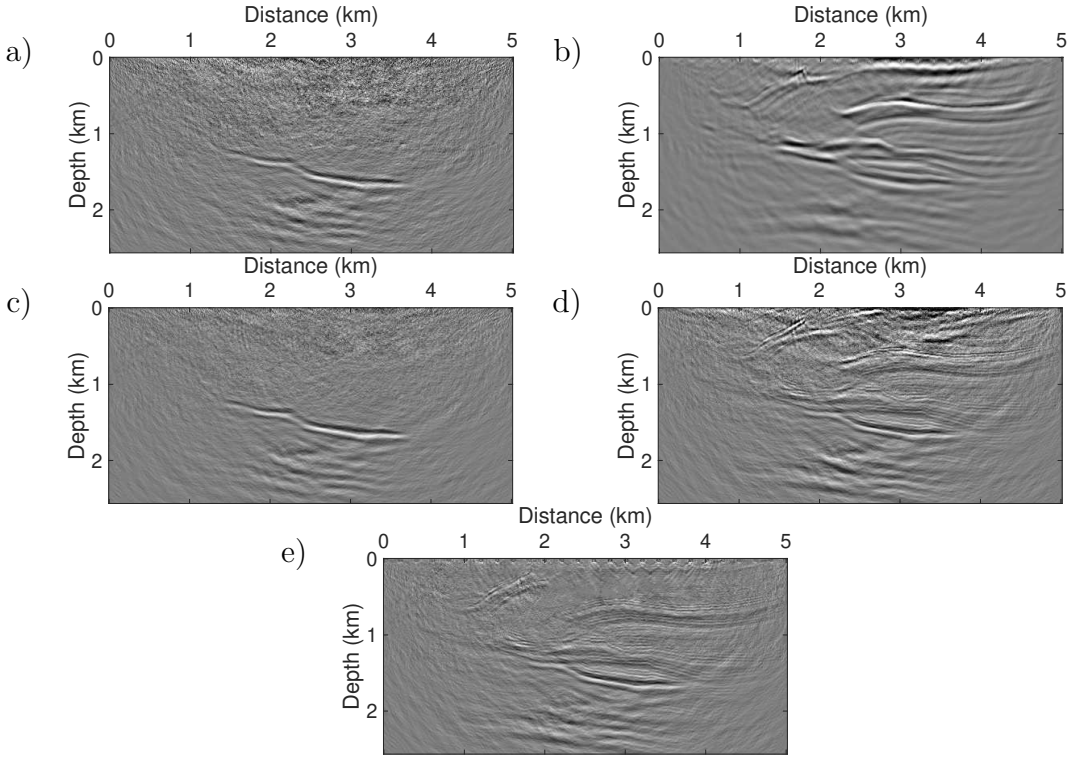

Figure 2.15: RTM images produced using difference data of all 25 shots (SNR=2), after time-shift correction. (a)-(d) Images formed from the difference data by using the source-normalized imaging condition (equation 2.22); (e) image formed from the data in Figure 2.5h, using the Poynting vector imaging condition (equation 2.19), with incident angles limited to within 50°. (a) Baseline and monitoring wavelets are identical. (b) Baseline and monitoring wavelets are different, and monitor data are not processed. (c) Baseline and monitoring wavelets are different, and monitor data are filtered by f_w. (d) and (e) Baseline and monitoring wavelets are different, and monitor data are filtered by f_s. Except (b), all images contain two reflection events with opposite polarities and similar reflection coefficients, which are close to the real situation.

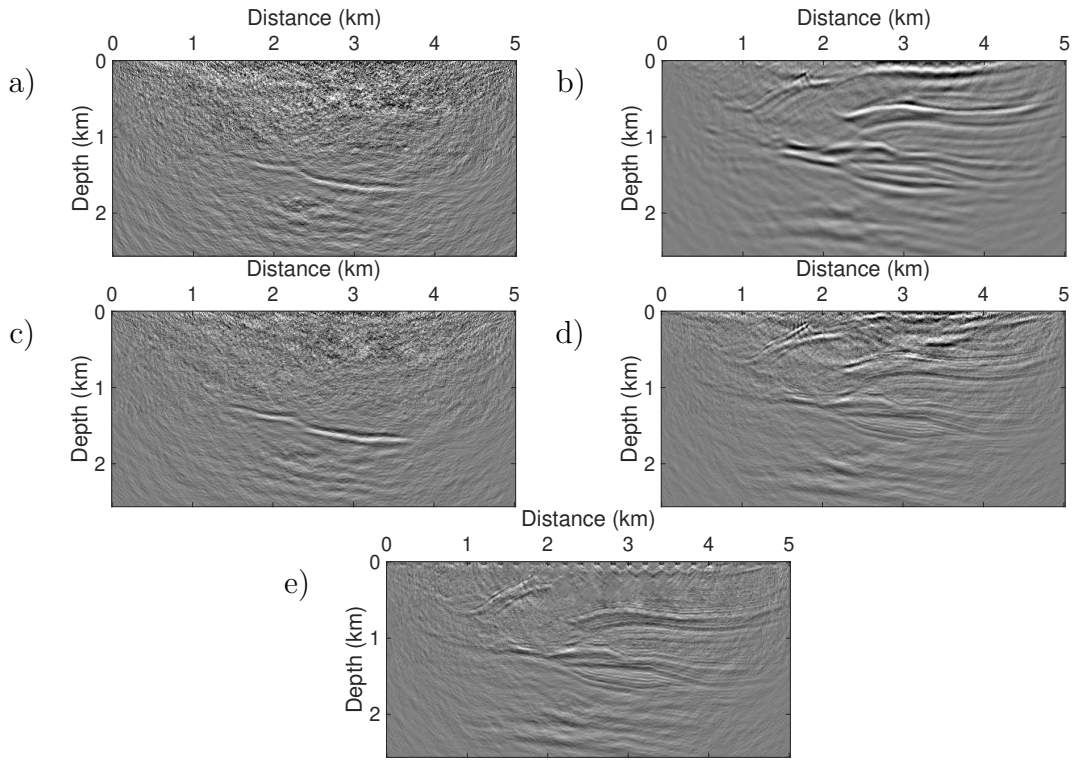

Figure 2.16: RTM images produced using difference data of all 25 shots (SNR=1), after time-shift correction. (a)-(d) Images formed from the difference data by using the source-normalized imaging condition (equation 2.22); (e) image formed from the data in Figure 2.5h, using the Poynting vector imaging condition (equation 2.19), with incident angles limited to within 50°. (a) Baseline and monitoring wavelets are identical. (b) Baseline and monitoring wavelets are different, and monitor data are not processed. (c) Baseline and monitoring wavelets are different, and monitor data are filtered by f_w. (d) and (e) Baseline and monitoring wavelets are different, and monitor data are filtered by f_s.

2.4.3 Comparisons with the least-squares matching filter

A conventional matching filter, f_{LSQ}, obtained from a least-squares method, can be expressed in frequency domain as (Ayeni and Nasser, 2009):

$$f_{LSQ}(\omega, k) = \frac{conj(\hat{d''_{2k}}(\omega))\hat{d''_{1k}}(\omega)}{conj(\hat{d''_{2k}}(\omega))\hat{d''_{2k}}(\omega) + \epsilon^2}, \tag{2.23}$$

where $conj(\cdot)$ denotes the conjugate operator; $\hat{d''_{1k}}$ and $\hat{d''_{2k}}$ are the frequency spectra of number k baseline and monitor traces, respectively, in a chosen time window; and ϵ^2 is a damp factor to avoid the case of a zero-value denominator. f_{LSQ} is a filter to match monitor trace to baseline trace, which needs to be calculated for each trace.

To compare f_{LSQ} with f_s, both do not need wavelet information, we employ two numerical examples in this section. One is a layer model plotted in Figure 2.17, another is the same as that in Figure 2.3 which has higher geological complexity. Two models are identical in acquisition geometry and grid spacing. The extra model used in Figure 2.17 is also to further prove the feasibility of f_s.

Traces, at the distance of 2.5km, of shot gathers with sources located at the centers of the models, are plotted in Figures 2.18-2.21. The difference traces, obtained by using the filtered monitor traces minus the baseline traces, are plotted in Figures 2.19 and 2.21. They show f_{LSQ} can perform well on the layer model, but for the geologically complex model (Figure 2.3), f_{LSQ} can not effectively eliminate the impact of wavelet non-repeatability, and f_s are able to reduce the wavelet non-repeatability for both models and are obviously better than f_{LSQ}.

2.4.4 Comparisons with the CDP-domain warping

In this study, a 1D (one-dimentional) warping algorithm, i.e., the algorithm of fast local cross-correlations in algorithm 1, is employed in shot gathers to decay the kinematical difference between baseline and monitor data. In contrast, a conventional time-lapse imaging

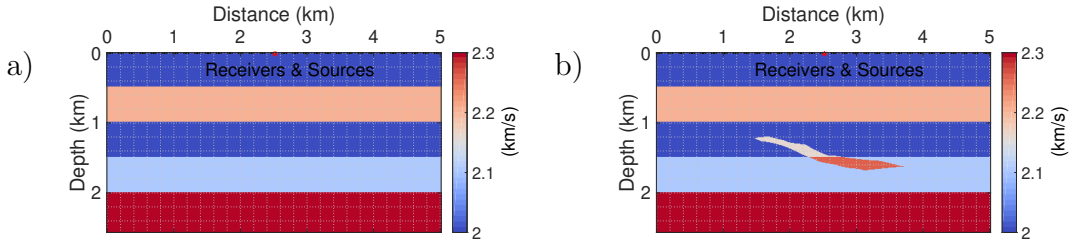

Figure 2.17: Layer models and acquisition geometries. (a) Baseline model; (b) monitor model. The time-lapse model (monitor model minus baseline model) is the same as that in Figure 2.3b. The dash lines and asterisks are the locations of receivers and sources, respectively, which are the same for the baseline model and the monitor model.

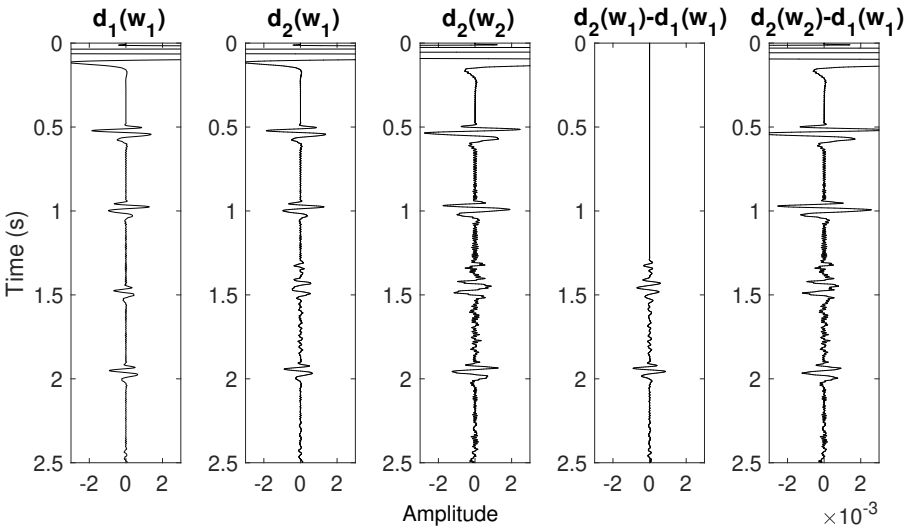

Figure 2.18: Traces of shot gathers, for the models in Figure 2.17, at the distance of 2.5km. $d_1(w_1)$ is the baseline trace obtained with wavelet 1 (Figure 2.4a). $d_2(w_1)$ is the monitor trace obtained with wavelet 1 (Figure 2.4a). $d_2(w_2)$ is the monitor trace obtained with wavelet 2 (Figure 2.4a). $d_2(w_1) - d_1(w_1)$ is the difference trace in the case of baseline and monitor wavelets (wavelet 1) are the same, which is the response of the true time-lapse model and is used as a reference trace in Figure 2.19. $d_2(w_2) - d_1(w_1)$ is the difference trace in the case of baseline and monitor wavelets are different, in which the response of the true time-lapse model is fully submerged by the wavelet non-repeatability.

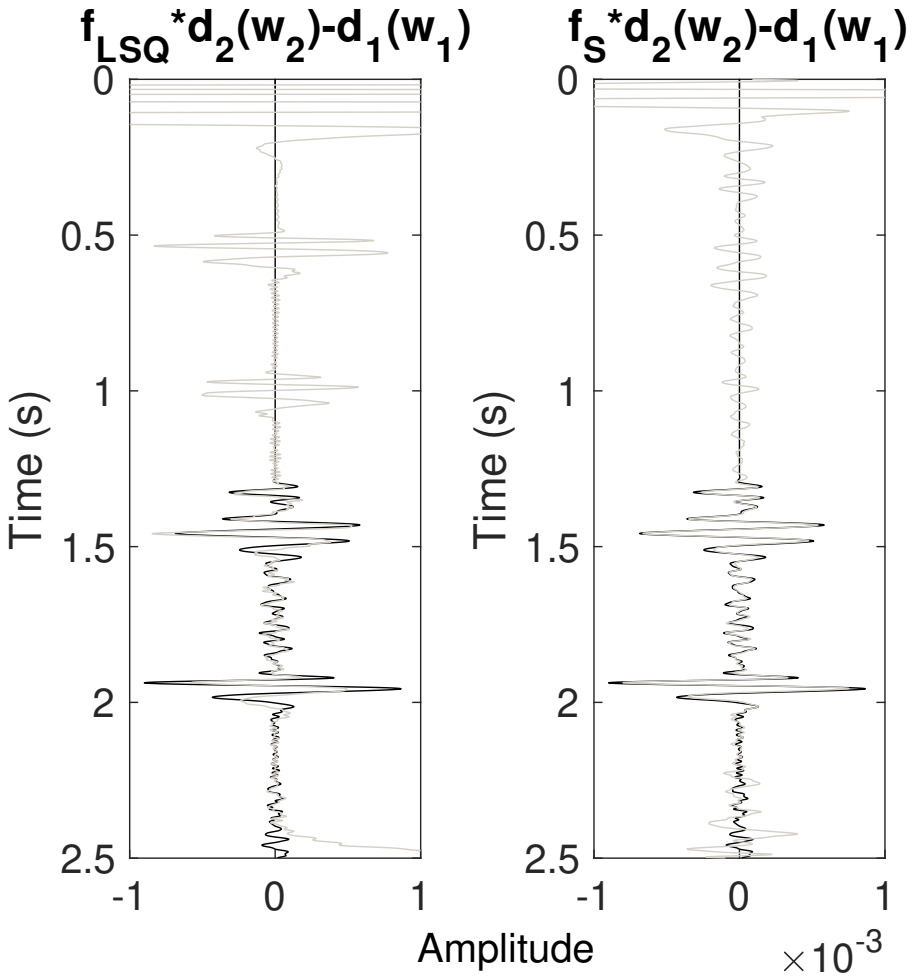

Figure 2.19: The black line in each panel is the reference trace in Figure 2.18. The gray line in the left panel is the difference trace obtained by using the conventional matching filter f_{LSQ} to trace $d_2(w_2)$ in Figure 2.18. The gray line in the right panel is the difference trace obtained by using f_s to trace $d_2(w_2)$ in Figure 2.18. Time window chosen to calculate f_{LSQ} is from 0 to $1.2s$.

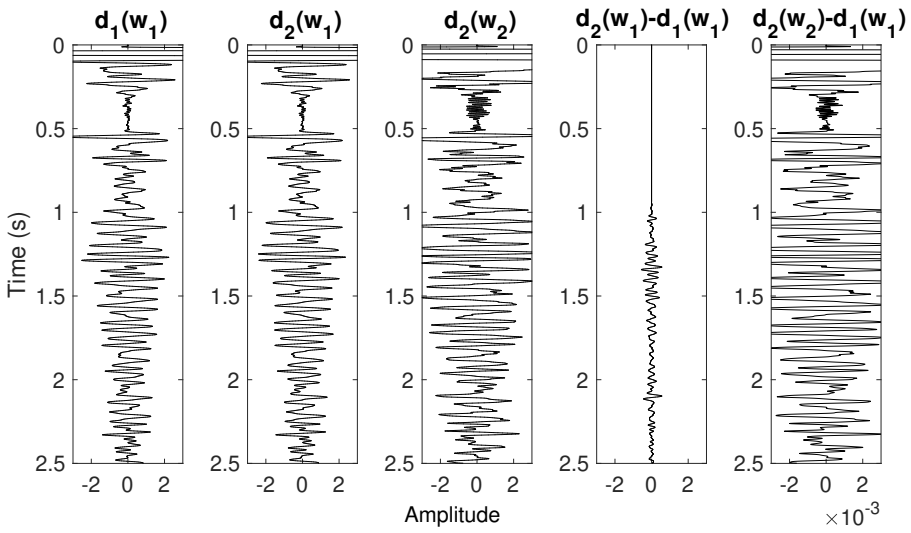

Figure 2.20: Traces of shot gathers, for the models in Figure 2.3, at the distance of 2.5km. $d_1(w_1)$ is the baseline trace obtained with wavelet 1 (Figure 2.4a). $d_2(w_1)$ is the monitor trace obtained with wavelet 1 (Figure 2.4a). $d_2(w_2)$ is the monitor trace obtained with wavelet 2 (Figure 2.4a). $d_2(w_1) - d_1(w_1)$ is the difference trace in the case of baseline and monitor wavelets (wavelet 1) are the same, which is the response of the true time-lapse model and is used as a reference trace in Figure 2.21. $d_2(w_2) - d_1(w_1)$ is the difference trace in the case of baseline and monitor wavelets are different, in which the response of the true time-lapse model is fully submerged by the wavelet non-repeatability.

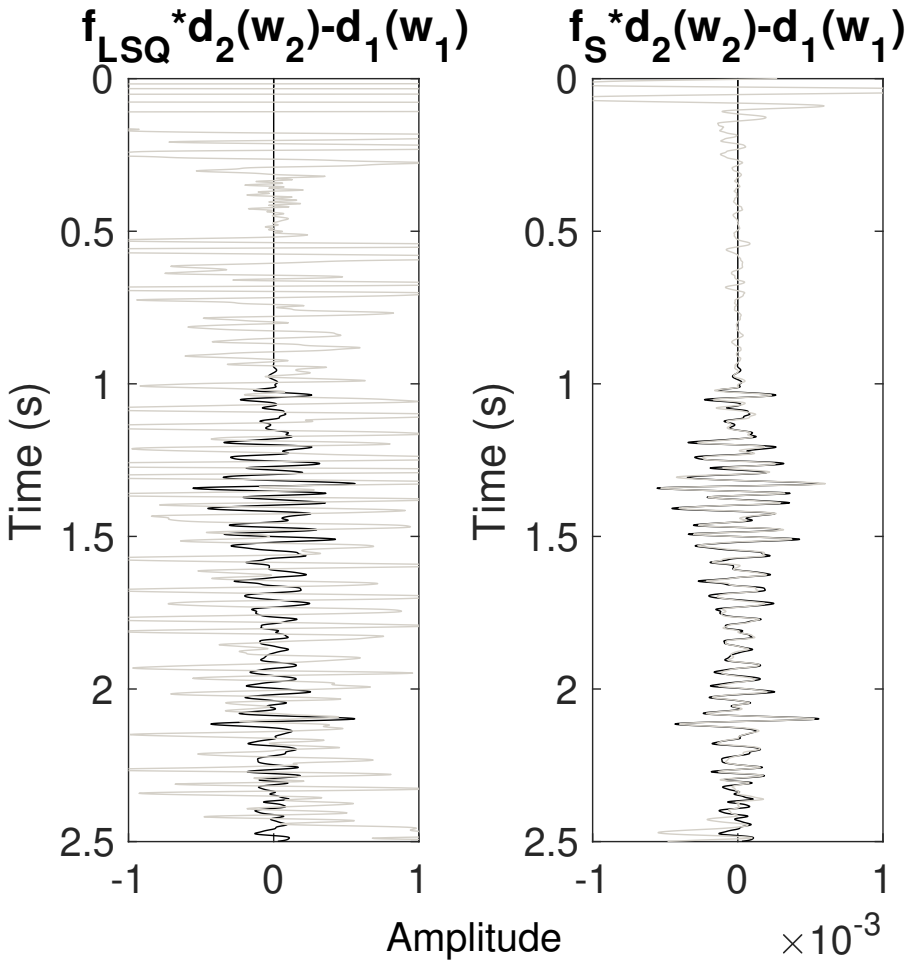

Figure 2.21: The black line in each panel is the reference trace in Figure 2.20. The gray line in the left panel is the difference trace obtained by using the conventional matching filter f_{LSQ} to trace $d_2(w_2)$ in Figure 2.20. The gray line in the right panel is the difference trace obtained by using f_s to trace $d_2(w_2)$ in Figure 2.20. Time window chosen to calculate f_{LSQ} is from 0 to $0.9s$.

workflow usually implements the warping algorithm on common-depth-point (CDP) gathers to eliminate this difference. To have comparisons between the shot-gather-domain warping used in this study and the conventional CDP-domain warping, we have shown the time-lapse images of them in Figure 2.22. And all data used are the same as that used in Figure 2.8, in which the wavelets for baseline and monitor shot gathers are identical to avoid the wavelet non-repeatability interference.

Figure 2.22a (same as Figure 2.8a) is the time-lapse image before warping, in which the affect of kinematical differences, i.e., the time shifts, are comtained, enhence, the correct relative amplitudes of events cannot be reflected in this image. In Figure 2.22b and 2.22c, we plot the time-lapse image after 1D (vertical) and 2D (both vertical and horizontal) CDP-domain warping, respectively. The warping algorithm applyed is the dynamic time warping (DTW) algorithm, which is more advanced than algorithm 1, but requires more computational cost. To obtain Figure 2.22b and 2.22c, first, we implemet RTM to baseline and monitor shot gathers, individually, then, apply 1D or 2D DTW to warp the CDP-domain baseline and monitor RTM images, and the final time-lapse images are the differences between the warped baseline and monitor images. And Figure 2.22d (same as Figure 2.9a) is time-lapse image produced using shot gathers after 1D warping or time-shift correction by our method. We observe the conventional CDP-domain warping can reduce the artifacts outside the reservoir area, however, it over-warps and destroys events corresponding to the time-lapse change, especially, in Figure 2.22c, obtained by 2D DTW, the events are nearly unrecognizable. And in the image produced by shot-gather-domain warping (Figure 2.22d) the reflections linking with the time-lapse change can be clearly distinguished.

2.5 Discussion

We have focus on the source wavelet non-repeatability issue. To build on the current approach, additional non-repeatability issues should also be considered. For example, before

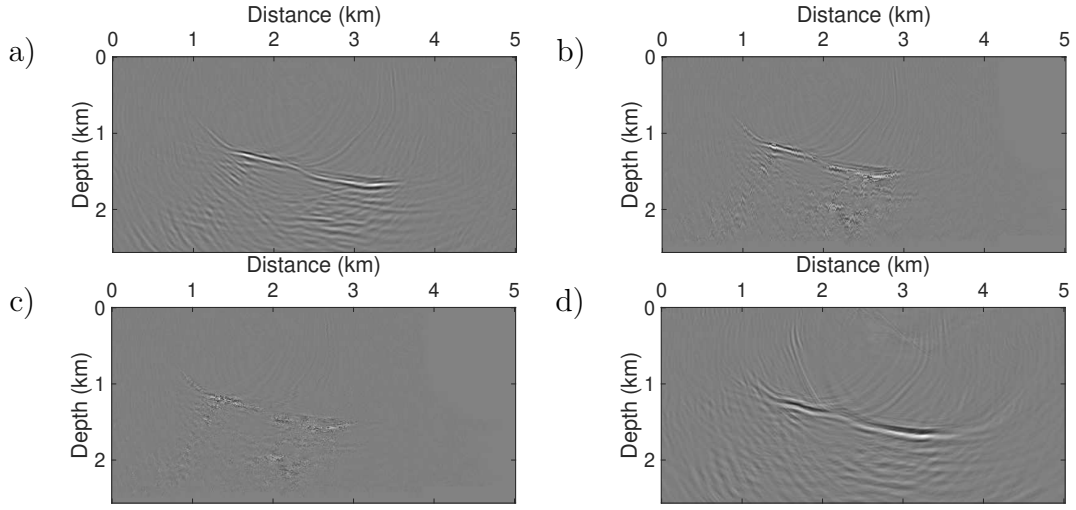

Figure 2.22: (a) Time-lapse image before warping (same as Figure 2.8a). (b) Time-lapse image after 1D (vertical) CDP-domain warping. (c) Time-lapse image after 2D (both vertical and horizontal) CDP-domain warping. (d) Time-lapse image produced using shot gathers after 1D warping or time-shift correction (same as Figure 2.9a). The correct image should contain two reflection events with opposite polarities and similar reflection coefficients.

applying the matching filters we have made use of, baseline and monitoring shot gathers could be interpolated and resampled into the same grids, and denoising and low-pass filtering should also be carried out, without changing the relationship in equation 2.3. The methods we present are also in principle applicable for elastic waves, and, since raw shot gathers are used, for 3D acquisitions, to enable full 4D imaging.

The workflow in this study is using a Poynting vector imaging condition to reduce the remaining errors arising from the inaccuracy of the source-independent matching filter. The errors are mainly from the remaining direct and refraction waves, considering the high computational cost of calculating Poynting vectors, an alternative of the Poynting vector imaging condition can be manually muting signals near the first arrivals, i.e., direct and refraction waves, in shot gathers first, and then applying the source-normalized condition to image.

We have proposed a solution for the non-repeatability issue on the source wavelet, but in which the derivation of f_s is based on the assumption that source locations for baseline and monitor surveys are identical, and the time-shift correction is also based on the same

assumption. Hence, there may still be at least two non-repeatability issues waiting for solutions before imaging for time-lapse raw field shot gathers, which are issues of non-repeatable source locations and seasonal variations (e.g., seawater velocity changes for marine data and near-surface property changes for land data) between baseline and monitor surveys. Even for the more advanced time-lapse full-waveform inversion that costs much more computational time than RTM, the seasonal variations can often vitally destroy the time-lapse imaging. Some additional efforts to address these will likely be required in practice.

2.6 Conclusions

We have developed a time-lapse imaging approach involving the design of two frequency-domain matching filters, to reduce source wavelet non-repeatability; the first requires source wavelet estimates from both of baseline and monitoring data; the second is source-independent but more sensitive to data noise. After suppressing source wavelet non-repeatability, we employ a time-shift correction, using a published fast local cross-correlations algorithm to further reduce the non-repeatability. Finally, depth images using an RTM algorithm with a Poynting vector imaging condition reduces remaining errors arising from the inaccuracy of the source-independent matching filter. The feasibility of our methods is demonstrated with synthetic noise-free and noisy data tests. Spectral ratios of the baseline and monitoring wavelets appear to be a robust and accurate means of suppressing source wavelet nonrepeatability, but high-quality (low noise) data are nevertheless important to the success of the filter design, and consequently the source-independence of the approach. Furthermore, our method has been demonstrated to have better performance when compared with the conventional least-squares matching filter and CDP-domain warping. At present, there is no common workflow for seismic imaging directly using time-lapse shot gathers. Our research contribution lies not only in the two matching filters, but also in a novel workflow for time-lapse seismic imaging.

Chapter 3

Full-waveform inversion

Full-waveform inversion (FWI) is a data-driven computational technique used in seismic exploration and imaging. It aims to estimate subsurface properties by iteratively minimizing the misfit between observed and modeled seismic waveforms.

FWI utilizes the complete waveform information recorded by seismic sensors to extract detailed information about the subsurface. It involves solving the wave equation iteratively to generate synthetic waveforms that closely match the observed data. By comparing these synthetic waveforms with the recorded data, FWI seeks to update the subsurface model parameters, such as velocity or density, to minimize the misfit.

FWI is known for its ability to provide high-resolution subsurface images and accurately estimate subsurface properties. However, it is a computationally intensive process that requires accurate modeling of wave propagation and careful handling of various challenges, such as nonlinearity, frequency content, and noise.

Despite its challenges, FWI has become a powerful tool in seismic imaging and is widely used in the oil and gas industry, as well as in geotechnical and environmental studies, for mapping subsurface structures and characterizing reservoirs or geological formations.

In this chapter, we will introduce the theory of FWI based on the adjoint method proposed by Tarantola (1984), and the derivation in Yang et al. (2015b) is also referred to.

3.1 Objective function optimization

Standard FWI starts from a given model $\mathbf{m}_0$ and seeks through local gradient information a model $\mathbf{m}$ that matches simulated data with observed data $\mathbf{d_{obs}}$. This is most often achieved by minimizing the $L2$ norm of the data residual $(\mathbf{d_{obs}} - \mathbf{F}(\mathbf{m}))$

$$E(\mathbf{m}) = \frac{1}{2}(\mathbf{d_{obs}} - \mathbf{F}(\mathbf{m}))^T(\mathbf{d_{obs}} - \mathbf{F}(\mathbf{m})). \tag{3.1}$$

Setting $\mathbf{m} = \mathbf{m}_0 + \Delta\mathbf{m}$ where $\Delta\mathbf{m}$ is the model perturbation, and taking Taylor expansion to $E(\mathbf{m})$, then we have

$$E(\mathbf{m}) = E(\mathbf{m}_0) + \frac{\partial E(\mathbf{m})}{\partial \mathbf{m}}\Delta\mathbf{m}. \tag{3.2}$$

Taking the first derivative with respect to $\mathbf{m}$ on both left and right sides, and it gives

$$\frac{\partial E(\mathbf{m})}{\partial \mathbf{m}} = \frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} + \frac{\partial E(\mathbf{m})}{\partial^2 \mathbf{m}^2}\Delta\mathbf{m}. \tag{3.3}$$

At the minimum of the objective function, the gradient should be zero. Hence, by forcing $\frac{\partial E(\mathbf{m})}{\partial \mathbf{m}} = 0$, we have

$$\Delta\mathbf{m} = -\mathbf{H}^{-1}\mathbf{g}, \tag{3.4}$$

where

$$\mathbf{H} = \frac{\partial E(\mathbf{m})}{\partial^2 \mathbf{m}^2}, \mathbf{g} = \frac{\partial E(\mathbf{m})}{\partial \mathbf{m}}, \tag{3.5}$$

in which, $\mathbf{H}$ and $\mathbf{g}$ are the Hessian matrix and the gradient, respectively. To avoid the expensive calculation cost of the Hessian matrix, it can be replaced by the diagonal approximation

$$\mathbf{H}_a = \mathbf{g}^T\mathbf{g}. \tag{3.6}$$

This compromises the capacity of the Hessian matrix to accurately reflect the relationship between different parameters. However, it still serves the purpose of compensating for the

sphere-spreading effect (Shin et al., 2001).

Putting equation 3.7 into equation 3.4, then we have the model perturbation

$$\Delta \mathbf{m} = -\mathbf{H}_a^{-1}\mathbf{g} = -(\mathbf{g}^T\mathbf{g})^{-1}\mathbf{g}, \tag{3.7}$$

then the model is iteratively updated by

$$\mathbf{m} = \mathbf{m} + \mu\Delta\mathbf{m}, \tag{3.8}$$

where μ is the stepsize and will be further introduced later.

3.2 Gradient calculation

We consider the 2D constant-density acoustic wave equation:

$$\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 P(\mathbf{x},t;\mathbf{x}_s) = f_s(\mathbf{x}_s,t), \tag{3.9}$$

where x is location vector, and $\mathbf{x}_s$ the source location.

The corresponding Green's function $G(\mathbf{x},t;\mathbf{x}_s,t')$ can be expressed as

$$\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 G(\mathbf{x},t;\mathbf{x}_s,t')}{\partial t^2} - \nabla^2 G(\mathbf{x},t;\mathbf{x}_s,t') = \delta(\mathbf{x}-\mathbf{x}_s)\delta(t-t'), \tag{3.10}$$

and, the integral representation of the solution can be given by

$$P(\mathbf{x},t;\mathbf{x}_s) = \int d\mathbf{x} \int dt' G(\mathbf{x},t;\mathbf{x}_s,t') f_s(\mathbf{x}_s,t'). \tag{3.11}$$

Since $v(\mathbf{x})$ is independent of time, Green's function is invariant with respect to time translation, i.e.

$$G(\mathbf{x},t;\mathbf{x}_s,t') = G(\mathbf{x},t-t';\mathbf{x}_s,0) = G(\mathbf{x},0;\mathbf{x}_s,t'-t). \tag{3.12}$$

Putting equation 3.12 into equation 3.11 and using the reciprocity theorem for the Green's function, we have

$$\begin{aligned} P(\mathbf{x},t;\mathbf{x}_s) &= \int d\mathbf{x} \int dt' G(\mathbf{x},t;\mathbf{x}_s,t') f_s(\mathbf{x}_s,t') \\ &= \int d\mathbf{x} \int dt' G(\mathbf{x},t-t';\mathbf{x}_s,0) f_s(\mathbf{x}_s,t') \\ &= \int d\mathbf{x} G(\mathbf{x},t;\mathbf{x}_s,0) * f_s(\mathbf{x}_s,t) \end{aligned} \tag{3.13}$$

Then, for the forward modeling $F(\mathbf{m}) = P(\mathbf{x}_r,t;\mathbf{x}_s)$ where $P(\mathbf{x}_r,t;\mathbf{x}_s)$ is the forward modeling wavefields at receiver locations $\mathbf{x}_r$, we have

$$P(\mathbf{x}_r,t;\mathbf{x}_s) = \int d\mathbf{x} G(\mathbf{x}_r,t;\mathbf{x}_s,0) * f_s(\mathbf{x}_s,t). \tag{3.14}$$

For velocity $v(\mathbf{x})$ added a small perturbation $\delta v(\mathbf{x})$, the corresponding wave equation is given by

$$\frac{1}{(v(\mathbf{x})+\delta v(\mathbf{x}))^2} \frac{\partial^2 (P(\mathbf{x},t;\mathbf{x}_s)+\delta P(\mathbf{x},t;\mathbf{x}_s))}{\partial t^2} - \nabla^2 (P(\mathbf{x},t;\mathbf{x}_s)+\delta P(\mathbf{x},t;\mathbf{x}_s)) = f_s(\mathbf{x}_s,t), \tag{3.15}$$

where

$$\frac{1}{(v(\mathbf{x})+\delta v(\mathbf{x}))^2} = \frac{1}{v^2(\mathbf{x})} - \frac{2\delta v(\mathbf{x})}{v^3(\mathbf{x})} + O((\delta v)^2). \tag{3.16}$$

Keeping the first two terms of equation 3.16 and putting them into equation 3.15, then after comparing with equation 3.9, we have

$$\frac{1}{v^2(\mathbf{x})} \frac{\partial^2 \delta P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 \delta P(\mathbf{x},t;\mathbf{x}_s) = \frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} \frac{2\delta v(\mathbf{x})}{v^3(\mathbf{x})}, \tag{3.17}$$

According to equation 3.14, we obtain the integral representation of equation 3.17

$$\delta P(\mathbf{x}_r,t;\mathbf{x}_s) = \int d\mathbf{x} G(\mathbf{x}_r,t;\mathbf{x},0) * \frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} \frac{2\delta v(\mathbf{x})}{v^3(\mathbf{x})}, \tag{3.18}$$

and then the gradient of the forward modeling function with respect to the velocity model is given by

$$\frac{\partial P(\mathbf{x}_r,t;\mathbf{x}_s)}{\partial v(\mathbf{x})} \approx \frac{\delta P(\mathbf{x}_r,t;\mathbf{x}_s)}{\delta v} = \int d\mathbf{x} G(\mathbf{x}_r,t;\mathbf{x},0) * \frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}\frac{2}{v^3(\mathbf{x})}. \tag{3.19}$$

We re-write the objective function equation 3.1 as following format

$$E(\mathbf{m}) = \frac{1}{2}\sum_{r=1}^{nr}\sum_{s=1}^{ns}\int dt \left(d_{obs} - P(\mathbf{x}_r,t;\mathbf{x}_s)\right)^2, \tag{3.20}$$

then the gradient of the objective function with respect to the velocity model $v(\mathbf{x})$ is given by

$$\begin{aligned}\frac{\partial E(\mathbf{m})}{\partial v(\mathbf{x})} &= -\sum_{r=1}^{nr}\sum_{s=1}^{ns}\int dt \frac{\partial P(\mathbf{x}_r,t;\mathbf{x}_s)}{\partial v(\mathbf{x})}\delta P(\mathbf{x}_r,t;\mathbf{x}_s) \;\; (\delta P(\mathbf{x}_r,t;\mathbf{x}_s) = d_{obs} - P(\mathbf{x}_r,t;\mathbf{x}_s)) \\ &= -\frac{2}{v^3(\mathbf{x})}\sum_{r=1}^{nr}\sum_{s=1}^{ns}\int dt \int d\mathbf{x} G(\mathbf{x}_r,t;\mathbf{x},0) * \frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}\delta P(\mathbf{x}_r,t;\mathbf{x}_s).\end{aligned} \tag{3.21}$$

According to the convolution rule below

$$\int dt \left[g(t) * f(t)\right] h(t) = \int dt f(t)[g(-t) * h(t)], \tag{3.22}$$

we re-write the gradient equation 3.21 as

$$\begin{aligned}\frac{\partial E(\mathbf{m})}{\partial v(\mathbf{x})} &= -\frac{2}{v^3(\mathbf{x})}\sum_{r=1}^{nr}\sum_{s=1}^{ns}\int dt \frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}\int d\mathbf{x} G(\mathbf{x}_r,-t;\mathbf{x},0) * \delta P(\mathbf{x}_r,t;\mathbf{x}_s) \\ &= -\frac{2}{v^3(\mathbf{x})}\sum_{r=1}^{nr}\sum_{s=1}^{ns}\int dt \frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}\int d\mathbf{x} G(\mathbf{x}_r,0;\mathbf{x},t) * \delta P(\mathbf{x}_r,t;\mathbf{x}_s).\end{aligned} \tag{3.23}$$

Furthermore, using the reciprocity theorem for the Green's function we have

$$\frac{\partial E(\mathbf{m})}{\partial v(\mathbf{x})} = -\frac{2}{v^3(\mathbf{x})}\sum_{r=1}^{nr}\sum_{s=1}^{ns}\int dt \frac{\partial^2 P(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}\int d\mathbf{x} G(\mathbf{x},0;\mathbf{x}_r,t) * \delta P(\mathbf{x}_r,t;\mathbf{x}_s). \tag{3.24}$$

Set

$$P_b(\mathbf{x}, 0; \mathbf{x}_r, t) = \int d\mathbf{x} G(\mathbf{x}, 0; \mathbf{x}_r, t) * \delta P(\mathbf{x}_r, t; \mathbf{x}_s) \tag{3.25}$$

which is called the back propagation wave field of the data residual, and according to equation 3.25, it can be calculated by solving

$$\frac{1}{v^2(\mathbf{x})} \frac{\partial^2 P_b(\mathbf{x}, 0; \mathbf{x}_r, t)}{\partial t^2} - \nabla^2 P_b(\mathbf{x}, 0; \mathbf{x}_r, t) = \delta P(\mathbf{x}_r, t; \mathbf{x}_s). \tag{3.26}$$

Finally, the element, in gradient $\mathbf{g}$ in equation 3.7, corresponding to the location $\mathbf{x}$, is summarized as

$$\frac{\partial E(\mathbf{m})}{\partial v(\mathbf{x})} = -\frac{2}{v^3(\mathbf{x})} \sum_{r=1}^{nr} \sum_{s=1}^{ns} \int dt \frac{\partial^2 P(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} P_b(\mathbf{x}, 0; \mathbf{x}_r, t). \tag{3.27}$$

3.3 Line search technology

Stepsize calculation is a key factor in ensuring inversion efficiency and accuracy. In this study, a parabolic line search method (Vigh and Starr, 2008) is adopted. For each iteration during the inversion, after the model perturbation $\delta\mathbf{m}$ (equation (3.7)) is calculated, two different stepsizes μ_1 and μ_2 are properly chosen, then two corresponding data misfits E_1 and E_2 are calculated. If we assume the curve of data misfit versus the stepsize is parabolic, then, including the data misfit E_0 for the starting model of each iteration, we have three points $(0, E_0)$, (μ_1, E_1), and (μ_2, E_2) on the curve to figure out the exact parameters about the parabolic curve, see Figure 3.1. Finally, we can work out the best stepsize giving the minimum data misfit by

$$\mu_{best} = \frac{(E_1 - E_0)\mu_2^2 - (E_2 - E_0)\mu_1^2}{2(E_1 - E_0)\mu_2^2 - (E_2 - E_0)\mu_1^2}. \tag{3.28}$$

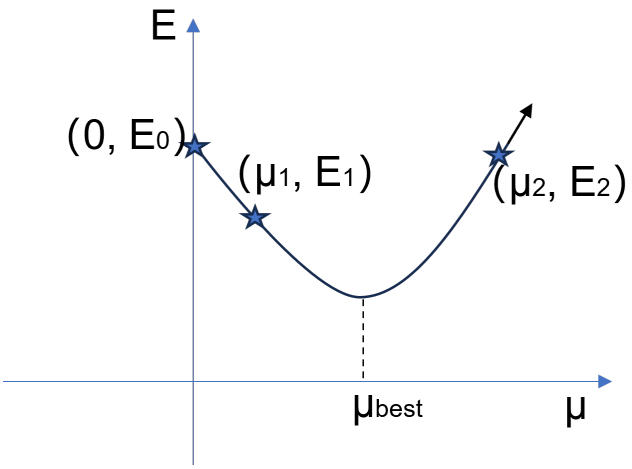

Figure 3.1: Stepsize calculation using the parabolic line search method by three points.

Chapter 4

Stepsize sharing in time-lapse full-waveform inversion

4.1 Abstract

Full waveform inversion (FWI) methods can produce high-resolution images of the physical properties of the subsurface. FWI has become a powerful tool for time-lapse or 4D seismic inversion, with applications in the monitoring of reservoir changes with injection and production, and potentially long term storage of carbon. Current time-lapse FWI strategies include the parallel strategy (PRS), the sequential strategy (SQS), the double-difference strategy (DDS), the common-model strategy (CMS), and the central-difference strategy (CDS). PRS time-lapse inversion is affected by convergence differences between the baseline and monitoring inversions, as well as non-repeatable noise and non-repeatable acquisition geometries between surveys. The other strategies are largely efforts to fix the sensitivities of PRS, but robust solutions are still sought. We hypothesize that several problems in time-lapse FWI arise from the independence of step lengths during updating. This is supported by synthetic data tests, which indicate that stepsize-sharing reduces artifacts caused by the variability in PRS convergence. Two strategies, which we refer to as stepsize-sharing PRS (SSPRS) and

stepsize-sharing CMS (SSCMS), are then designed to address these remaining issues. In this paper, we have tested our methods in five scenarios, including noise-free data, non-repeated noises, non-repeatable source positions, biased starting models, and a combination of the latter three. The comparisons between the SSPRS and other strategies show that the SSPRS can adapt to all tested scenarios well. Especially, except for the DDS which is extremely sensitive to the non-repeatable source positions, only the SSPRS can provide meaningful results in the latter two scenarios when compared with others. Furthermore given that SSPRS through its sharing incurs half of the time cost of seeking stepsizes compared with the PRS and DDS, and the total computational cost of SSPRS is less than half of that of the CMS and CDS.

4.2 Introduction

Time-lapse or 4D seismic analysis is a crucial technology for reservoir monitoring problems such as enhanced oil recovery and CO_2 storage (Greaves and Fulp, 1987; Ross and Altan, 1997; Wang et al., 1998; Lumley, 2001; Barkved et al., 2003; Arts et al., 2003; Barkved et al., 2005; Chadwick et al., 2009; Kazemeini et al., 2010; Pevzner et al., 2017). It has begun to be incorporated as a matter of course into reservoir development plans (Jack, 2017). Full waveform inversion (FWI) (Lailly et al., 1983; Tarantola, 1984; Virieux and Operto, 2009), a technology with the capacity to create high-resolution images of physical properties of subsurface media, has become a powerful tool for time-lapse inversion. Real field data applications have been reported with increased frequency (Raknes and Arntsen, 2014; Hicks et al., 2016; Yang et al., 2016; Kamei et al., 2017; Bortoni et al., 2021), but, as we will review next, challenges remain.

The time-lapse FWI approach we will refer to as "conventional" is known as the parallel strategy (PRS) (Lumley, 2001). In PRS, baseline and monitor inversions are carried out independently but using the same starting model. Its challenges derive from detailed differ-

ences within the the two independent inversions within it. They commonly exhibit different convergence properties, and these induce artifacts in the time-lapse inversion. To avoid some of these issues, Routh et al. (2012) introduced sequential strategy (SQS), which uses the inverted baseline as the starting model for the monitor inversion. However, examples presented by Yang et al. (2015a) and Zhou and Lumley (2021a) indicate that it can cause strong artifacts in the time-lapse inversion by amplifying convergence differences between the baseline and monitor inversions. Target-oriented SQS (Raknes and Arntsen, 2014; Asnaashari et al., 2015) has been shown to effectively constrain artifacts in the target zone, this mitigation strategy requires significant prior information, especially about the location of time-lapse change. Probably the most widely-adopted strategy at the moment is the double-difference strategy (DDS), proposed by Zheng et al. (2011), which has been used and adapted by many researchers (Zhang and Huang, 2013; Raknes et al., 2013; Fu and Innanen, 2021, 2022b), and has been vetted with real data case in Yang et al. (2016). It has several points of weakness, some of which are addressable. For instance, Fu et al. (2020) introduced a double-wavelet DDS to mitigate the impact of non-repeatability of baseline and monitor source wavelets to the inversion. However, DDS remains very sensitive to non-repeatability of source/receiver locations (Yang et al., 2015a; Zhou and Lumley, 2021b). A different approach was taken by Hicks et al. (2016), who introduced the common-model strategy (CMS) and applied it to a North Sea field case study; it was adopted by Bortoni et al. (2021) in real data of a post-salt field in the Campos Basin. A different variant, introduced by (Zhou and Lumley, 2021a) and called the central-difference strategy (CDS), has recently been shown to be robust to non-repeatable noise in time-lapse FWI (Zhou and Lumley, 2021b). As they are amongst the most recent and robust time-lapse methods in the literature, we will examine CMS and CDS under different conditions in this paper.

The main results of this paper are two new candidate time-lapse FWI strategies, which aim to grow and extend the robustnesses sought in the above references, and simultaneously address computational burden. The main concept we will leverage is the sharing of waveform

inversion stepsizes across baseline and monitoring inversions. Conclusions are based on benchmark synthetic data and comparative inversions with the new and recent methods.

4.3 Time-lapse FWI methods

In standard FWI (Lailly et al., 1983; Tarantola, 1984; Virieux and Operto, 2009) we minimize the L2 norm misfit function:

$$E(\mathbf{m}) = \frac{1}{2}||\mathbf{d}_{obs} - \mathbf{F}(\mathbf{m})||^2, \tag{4.1}$$

where $\mathbf{d}_{obs}$ is the observed data or recorded wavefields, $\mathbf{F}(\cdot)$ is a forward modeling operator based on the wave equation, and $\mathbf{m}$ is the updating model (e.g., P-wave velocity). Via some appropriate optimization approaches, based on steepest descents, conjugate gradients, etc., the model is updated iteratively as:

$$\mathbf{m}^k = \mathbf{m}^{k-1} + \delta\mathbf{m}^k, \tag{4.2}$$

where k is the iteration number, and

$$\delta\mathbf{m}^k = \mu^k \mathbf{g}(\mathbf{m}^{k-1}, \mathbf{d}_{res}^{k-1}), \tag{4.3}$$

where

$$\mathbf{d}_{res}^{k-1} = \mathbf{d}_{obs} - \mathbf{F}(\mathbf{m}^{k-1}), \tag{4.4}$$

in which $\mathbf{g}(\mathbf{m}^{k-1}, \mathbf{d}_{res}^{k-1})$ is the updating direction of model in iteration k, which depends on the updated model $\mathbf{m}^{k-1}$ and data residual $\mathbf{d}_{res}^{k-1}$ in iteration $k-1$, and μ^k, obtained from a line search in this study, is the stepsize for iteration k. In the steepest descent method, adopted in this paper, $\mathbf{g}$ represents the gradient of the misfit function (equation 4.1) with respect to $\mathbf{m}$, which is the zero-lag cross-correlation between forward wavefields

and backward wavefields of data residuals. For the first iteration, a starting model $\mathbf{m}^0$ have to be prepared, which can be obtained by velocity analysis or tomography. Furthermore, combining equation 4.3 and 4.4, we have

$$\delta\mathbf{m}^k = \mu^k \mathbf{g}(\mathbf{m}^{k-1}, \mathbf{d}_{obs}), \tag{4.5}$$

where the updating direction $\mathbf{g}$ depends on observed data and the updated model $\mathbf{m}^{k-1}$ in iteration $k-1$. In this study, we use a time-domain constant-density acoustic finite-difference method as the forward modeling operator, the steepest descent method as the optimization, and we precondition the gradient with the diagonal approximation of the Hessian matrix (or source illumination) (Shin et al., 2001).

4.3.1 Common time-lapse inversion strategies

Parallel strategy

The parallel strategy (PRS) follows the workflow in Figure 5.1a. It includes two independent FWI processes: baseline model inversion, with baseline data and a starting model as inputs, and monitor model inversion, with monitor data and the same (baseline) starting model as inputs. The estimated time-lapse model is the difference between the inverted monitor model and the inverted baseline model. Since FWI is highly non-linear, with local minima as common traps, the two FWI processes mentioned above often have different convergence properties, with each difference tending to produce artifacts upon subtraction.

Sequential strategy

The sequential strategy (SQS) is summarized in the workflow in Figure 5.1b. It involves having the same baseline inversion as PRS, and takes baseline data and a starting model as inputs. The monitor inversion is different. In it, the inverted baseline model is used as the starting model for the monitor inversion, and the difference of the two final inversions is the

time-lapse model. Inversions with different starting models also tend to produce different convergence histories, however, and again strong artifacts are the result (Yang et al., 2015a; Fu et al., 2020; Zhou and Lumley, 2021a). We do not pursue SQS further in this study.

Double-difference strategy

The double-difference strategy (DDS), with workflow illustrated in Figure 4.1c, also contains two FWI steps. The first is again the baseline model inversion. In the second, the starting model is the inverted baseline model, as with SQS, but the input monitor data are not the observed monitor data. Instead, a composited data set is introduced:

$$\mathbf{d}_{DD} = \mathbf{F}(\mathbf{m}_{bas}) + (\mathbf{d}_{mon} - \mathbf{d}_{bas}), \tag{4.6}$$

where $\mathbf{F}(\mathbf{m}_{bas})$ are synthetic data predicted from the inverted baseline model $\mathbf{m}_{bas}$, $(\mathbf{d}_{mon} - \mathbf{d}_{bas})$ are difference data (observed monitor data $\mathbf{d}_{mon}$ minus the observed baseline data $\mathbf{d}_{bas}$). Accordingly, the misfit function for the monitor inversion becomes:

$$\mathbf{E}_{DD}(\mathbf{m}_{mon}) = \frac{1}{2}||\mathbf{d}_{DD} - \mathbf{F}(\mathbf{m}_{mon})||^2, \tag{4.7}$$

where $\mathbf{F}(\mathbf{m}_{mon})$ are the synthetic data predicted from the inverted monitor model $\mathbf{m}_{mon}$. Rewriting equation 4.7 as:

$$\mathbf{E}_{DD}(\mathbf{m}_{mon}) = \frac{1}{2}||(\mathbf{d}_{mon} - \mathbf{F}(\mathbf{m}_{mon})) - (\mathbf{d}_{bas} - \mathbf{F}(\mathbf{m}_{bas}))||^2, \tag{4.8}$$

we observe that minimizing the misfit function tends to equalize the baseline data residual and the monitor data residual. DDS can reduce convergence differences between baseline and monitor inversions, tending to suppress artifacts outside of the time-lapse change zone. It is, in other words, a kind of target-oriented strategy. Its disadvantages arise largely in difference data in equation 4.6, which are often weak and are easy influenced by non-repeatability in

time-lapse surveys. Those influences can leak strongly into the final inverted time-lapse model, and cause heavy artifacts.

Common-model strategy

The common-model strategy (CMS), with workflow illustrated in Figure 4.1d, can be seen as an upgraded version of the PRS. Essentially, it contains two instances of PRS. First, the baseline and monitor inversions are performed independently with the same starting model. Then a new starting model is taken from the average of the baseline and monitor models; with this, the baseline and monitor inversions are performed independently again, still with the original data sets. The final time-lapse model is obtained from a difference of the baseline and monitor models derived during the second PRS. Although essentially an upgraded PRS, the CMS strategy has proved in both synthetic and field case studies to outperform the PRS on account of the changes in the effective baseline model (Hicks et al., 2016).

Central-difference strategy

The central-difference strategy (CDS), with workflow illustrated in Figure 4.1e comprises two instances of SQS, called the forward bootstrap FWI and the reverse bootstrap FWI. The forward bootstrap FWI uses the baseline data and a starting model to invert for a baseline model; then it uses the inverted baseline model as the starting model for monitor inversion. The first time-lapse model is obtained by subtracting the baseline model from the monitor model, and the second (i.e., the reverse bootstrap FWI) uses the monitor data and the same starting model as in the first FWI to invert for the monitor model. Then the inverted monitor model is used as the starting model for baseline model inversion. The second time-lapse model is obtained by subtracting the baseline model from the monitor model. The final time-lapse model is the average of these two time-lapse models. Zhou and Lumley (2021a) find that artifacts in the first and second time-lapse models have opposite phases, whereas the correct time-lapse estimations in the two models have identical phases,

hence the artifacts are attenuated after averaging the two models.

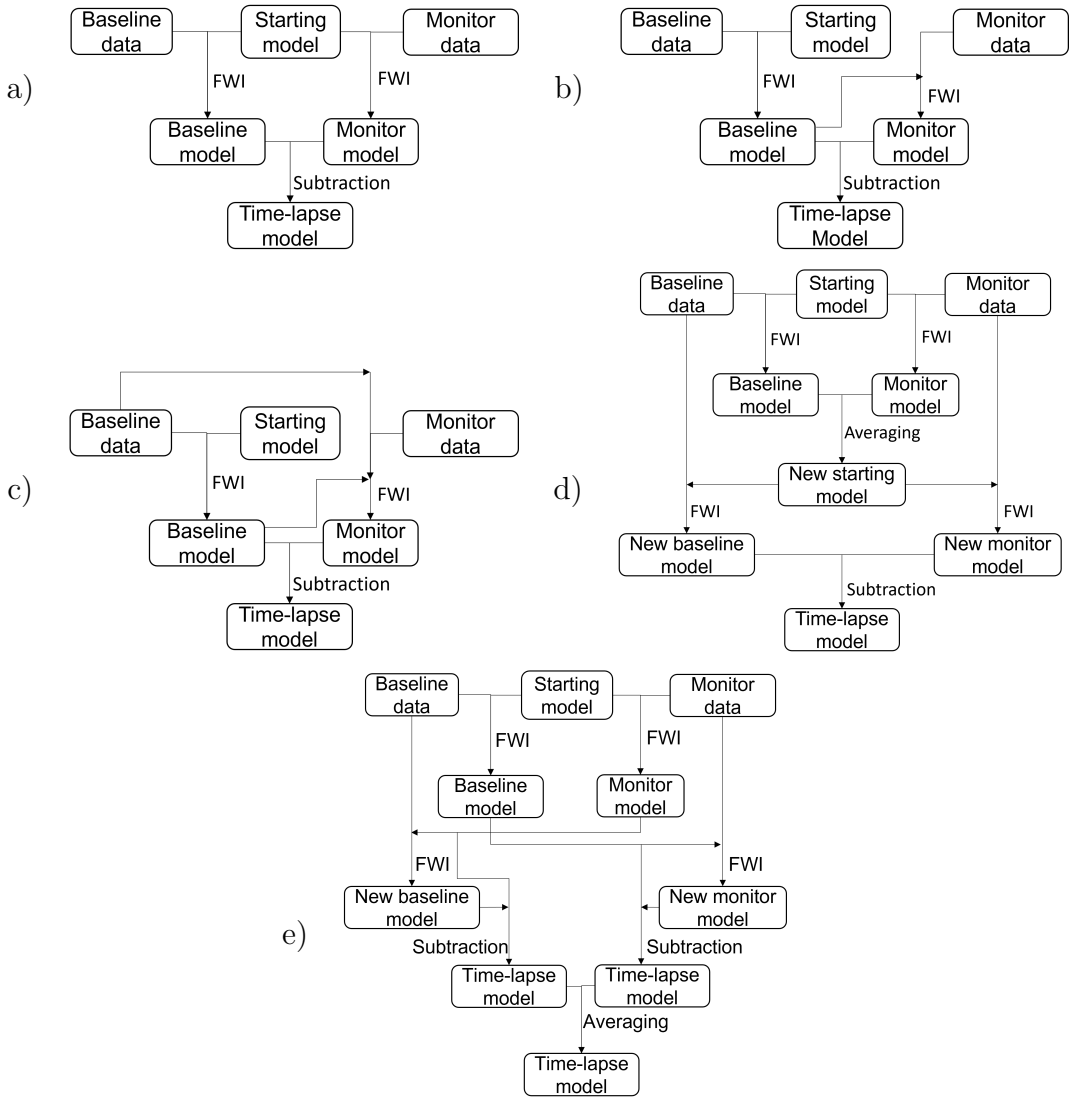

Figure 4.1: Flowcharts of present time-lapse FWI strategies. (a) Parallel strategy (PRS). (b) Sequential strategy (SQS). (c) Double-difference strategy (DDS). (d) Common-model strategy (CMS). (e) Central-difference strategy (CDS).

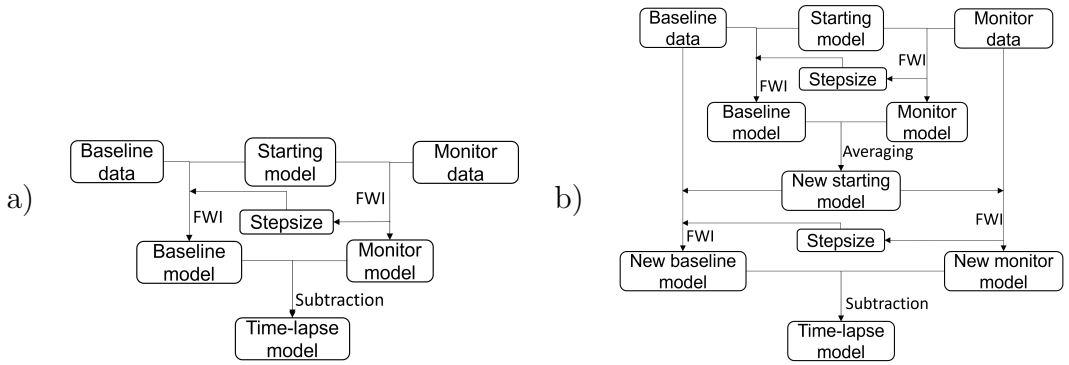

Figure 4.2: Flowcharts of stepsizes-sharing time-lapse FWI strategies. (a) Stepsizes-sharing parallel strategy (SSPRS). (b) Stepsizes-sharing common-model strategy (SSCMS).

4.4 Stepsize sharing time-lapse inversion strategies

In this section, we propose two new strategies for time-lapse FWI that add to these mitigating efforts the idea of sharing of stepsizes. Returning to the sequence in equations 4.2 to 4.5, if we substitute equation 4.5 into 4.2, we produce an expression for an updated baseline model at iteration k:

$$\mathbf{m}_{bas}^{k} = \mathbf{m}_{bas}^{k-1} + \mu_{bas}^{k}\mathbf{g}(\mathbf{m}_{bas}^{k-1}, \mathbf{d}_{bas,obs}), \tag{4.9}$$

and an updated monitor model at iteration k expressed as:

$$\mathbf{m}_{mon}^{k} = \mathbf{m}_{mon}^{k-1} + \mu_{mon}^{k}\mathbf{g}(\mathbf{m}_{mon}^{k-1}, \mathbf{d}_{mon,obs}), \tag{4.10}$$

where the subscript *bas* or *mon* denotes the parameter is in the baseline or monitor inversion. At the end of two FWI procedures, the inverted baseline and monitor models are, respectively:

$$\mathbf{m}_{bas} = \mathbf{m}_{bas}^{0} + \sum_{k=1}^{m} \mu_{bas}^{k}\mathbf{g}(\mathbf{m}_{bas}^{k-1}, \mathbf{d}_{bas,obs}), \tag{4.11}$$

and

$$\mathbf{m}_{mon} = \mathbf{m}_{mon}^{0} + \sum_{k=1}^{n} \mu_{mon}^{k}\mathbf{g}(\mathbf{m}_{mon}^{k-1}, \mathbf{d}_{mon,obs}), \tag{4.12}$$

where $\mathbf{m}_{bas}^0$ and $\mathbf{m}_{mon}^0$ are, respectively, the starting models for baseline and monitor inversions, and m and n are the maximum baseline and monitor iteration numbers, respectively. Let

$$\mathbf{d}_{dif} = \mathbf{d}_{mon,obs} - \mathbf{d}_{bas,obs}, \tag{4.13}$$

be the difference data, and

$$\mathbf{m}_{mon}^{k-1} = \mathbf{m}_{mon,bas}^{k-1} + \mathbf{m}_{tl}^{k-1}, \tag{4.14}$$

in which $\mathbf{m}_{mon,bas}^{k-1}$ is the baseline model implied by the monitor model $\mathbf{m}_{mon}^{k-1}$ and $\mathbf{m}_{tl}^{k-1}$ is the time-lapse model in iteration $k-1$. With these in place in equation 4.12, we have

$$\mathbf{m}_{mon} = \mathbf{m}_{mon}^0 + \sum_{k=1}^{n} \mu_{mon}^k \mathbf{g}(\mathbf{m}_{mon,bas}^{k-1} + \mathbf{m}_{tl}^{k-1}, \mathbf{d}_{bas,obs} + \mathbf{d}_{dif}). \tag{4.15}$$

According to the adjoint state method, the gradient $\mathbf{g}$ is calculated by applying a zero-lag cross-correlation between the forward modeling wavefields and backpropagated data residuals (Lailly et al., 1983; Tarantola, 1984), hence, $\mathbf{g}$ is linear with respect to the observed data, equation 4.15 can be rewritten as:

$$\mathbf{m}_{mon} = \mathbf{m}_{mon}^0 + \sum_{k=1}^{n} \mu_{mon}^k \mathbf{g}(\mathbf{m}_{mon,bas}^{k-1} + \mathbf{m}_{tl}^{k-1}, \mathbf{d}_{bas,obs}) + \sum_{k=1}^{n} \mu_{mon}^k \mathbf{g}(\mathbf{m}_{mon,bas}^{k-1} + \mathbf{m}_{tl}^{k-1}, \mathbf{d}_{dif}). \tag{4.16}$$

And the quantity $\mathbf{g}$ is nonlinear with respect to the model, so we approximate the updating direction by Taylor expansion as:

$$\mathbf{g}(\mathbf{m}_{mon,bas}^{k-1} + \mathbf{m}_{tl}^{k-1}, \mathbf{d}_{bas,obs}) \approx \mathbf{g}(\mathbf{m}_{mon,bas}^{k-1}, \mathbf{d}_{bas,obs}) + \mathbf{g}'(\mathbf{m}_{mon,bas}^{k-1}, \mathbf{d}_{bas,obs})\mathbf{m}_{tl}^{k-1}, \tag{4.17}$$

where $\mathbf{g}^{'}$ is the derivative of the updating direction $\mathbf{g}$ with respect to model $\mathbf{m}^{k-1}_{mon,bas}$. Putting equation 4.17 into equation 4.16, we have

$$\begin{aligned}\mathbf{m}_{mon} = \mathbf{m}^0_{mon} &+ \sum_{k=1}^{n} \mu^k_{mon}\mathbf{g}(\mathbf{m}^{k-1}_{mon,bas}, \mathbf{d}_{bas,obs}) + \sum_{k=1}^{n} \mu^k_{mon}\mathbf{g}^{'}(\mathbf{m}^{k-1}_{mon,bas}, \mathbf{d}_{bas,obs})\mathbf{m}^{k-1}_{tl} \\ &+ \sum_{k=1}^{n} \mu^k_{mon}\mathbf{g}(\mathbf{m}^{k-1}_{mon}, \mathbf{d}_{dif}).\end{aligned} \tag{4.18}$$

or

$$\mathbf{m}_{mon} = \mathbf{m}_{mon,bas} + \mathbf{m}_{tl}, \tag{4.19}$$

where

$$\mathbf{m}_{mon,bas} = \mathbf{m}^0_{mon} + \sum_{k=1}^{n} \mu^k_{mon}\mathbf{g}(\mathbf{m}^{k-1}_{mon,bas}, \mathbf{d}_{bas,obs}), \tag{4.20}$$

$$\mathbf{m}_{tl} = \sum_{k=1}^{n} \mu^k_{mon}\mathbf{g}^{'}(\mathbf{m}^{k-1}_{mon,bas}, \mathbf{d}_{bas,obs})\mathbf{m}^{k-1}_{tl} + \sum_{k=1}^{n} \mu^k_{mon}\mathbf{g}(\mathbf{m}^{k-1}_{mon}, \mathbf{d}_{dif}). \tag{4.21}$$

Comparing equation 4.11 with equation 4.20, we observe that the implicit baseline model can be eliminated from the inverted monitor model under the conditions:

(1) We employ the same starting model, i.e. $\mathbf{m}^0_{bas} = \mathbf{m}^0_{mon}$;

(2) The iteration number is the same, i.e. $m = n$;

(3) The stepsizes are the same, i.e. $\mu^k_{mon} = \mu^k_{bas}$;

(4) The updated baseline models are the same, i.e. $\mathbf{m}^{k-1}_{mon,bas} = \mathbf{m}^{k-1}_{bas}$.

Conditions (1) and (3) in fact enforce condition (4), since equations 4.11 and 4.20 are recursive. Hence, we only need to meet the first three conditions to eliminate the implicit baseline model.

In PRS, two FWI procedures are enacted, starting from the same model, $\mathbf{m}^0_{bas} = \mathbf{m}^0_{mon}$, and having the same iteration number, $m = n$ (the case of different iteration numbers normally deteriorates artifacts on the final time-lapse model). The implicit baseline model

$\mathbf{m}_{mon,bas}$ cannot be eliminated completely in the PRS, since the condition (3) is not met. The remaining baseline model can be considered the source of the coherent artifacts in the final time-lapse model.

PRS can be adapted to produce the stepsize-sharing parallel strategy (SSPRS), to meet the condition (3), wherein $\mu^k_{mon} = \mu^k_{bas}$, such that the implicit baseline model is eliminated. The workflow is illustrated in Figure 4.2a, in which we perform the monitor inversion first with the inputs of a starting model and observed monitor data, and the outputs are not only the monitor model but also the stepsizes for each iteration. Then, in the second FWI, the baseline inversion, the inputs are not only the observed baseline data and the same starting model, but also the stepsizes from the first monitor inversion. The final time-lapse model is obtained by subtracting the baseline model from the monitor model. In addition to the elimination of what is largely a source of artifacts, an advantage of this SSPRS approach is that it saves on the cost of seeking the stepsizes during the second FWI. This is the first of the two proposed methods.

The CMS approach can also be upgraded to incorporate this idea, leading to the stepsize-sharing common-model strategy (SSCMS), the second of the two proposed methods. The workflow is illustrated in Figure 4.2b. It involves four FWI steps. The first two are the same as in SSPRS. In the second two, we repeat SSPRS but with a starting model comprising the average of the inverted baseline and monitor model in the first two. The final time-lapse model is the difference between the inverted monitor and baseline models in the second two FWI outputs.

The baseline updating directions ($\mathbf{g}(\mathbf{m}^{k-1}_{bas}, \mathbf{d}_{bas,obs})$) and the monitor updating directions ($\mathbf{g}(\mathbf{m}^{k-1}_{mon}, \mathbf{d}_{mon,obs})$) may not be produced at same scale, depending on the character of data noise and differences in baseline and monitor source/receiver locations. So, in both SSPRS and SSCMS, we recommend that the baseline and monitor update directions be calibrated to a common root-mean-square before multiplying the shared stepsizes from monitor FWI. Actually, the gradients involving time-lapse change are normally very small when compared

with the background (or baseline) gradients, and the calibration has little impact on the gradients when baseline and monitoring wavelets are identical. The gradient calibration is to avoid the case that baseline and monitoring wavelets are different and the effect of the wavelet is not canceled in the gradient (i.e., a gradient without the source illumination).

4.5 Numerical examples

In this section, we use a land model to test our methods and have comparisons with present methods. The true baseline model is displayed in Figure 4.3a, two reservoirs are located at the left below corner and near the center, respectively. To mimic the fluid change, 4% or $49m/s$ velocity changes, displayed in Figure 4.3b, are added at the two reservoirs to obtain the monitor model. A smooth starting model is displayed in Figure 4.3c, which is employed in the first FWI of all time-lapse strategies mentioned above. The model size is 101-by-208 with 10m spacing. On the top of the model, seven sources are evenly spread at the depth of 10m and each surface cell grid is located a receiver. The source wavelet used for baseline and monitor data sets is identical, which is a minimum phase wavelet with a dominant frequency of $10Hz$. The time sampling interval is 2 milliseconds and the maximum recording time is 1.2 seconds. To show the capacity of the FWI program, a time-domain constant-density acoustic FWI, used in this study, we display the inverted baseline model in Figure 4.4a, two traces crossing the two reservoirs at distances of 970m and 1700m are abstracted and plotted in Figure 4.4b, and the curve of misfits versus iteration numbers is plotted in Figure 4.4c. The results show a good performance of the used FWI program. These are to avoid the interferences, caused by the use of an incompetent FWI program, to our following tests. In the next subsections, we will implement the PRS, SSPRS, DDS, CMS, CDS, and SSCMS with noise-free data sets, noisy data sets, data sets with non-repeatable source locations, and biased starting models, and have comparisons between different strategies. All data sets are acoustic and synthetic, and all inversions have the same iteration number.

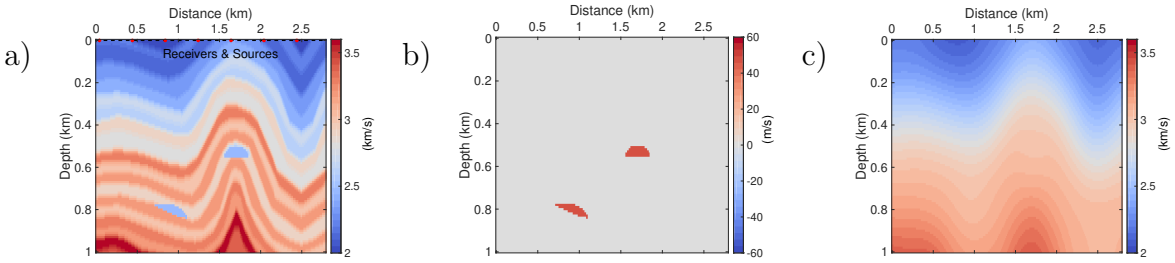

Figure 4.3: (a) True baseline model and acquisition geometry, the dashed line denotes positions of receivers, and the red asterisks denote positions of sources. (b) True time-lapse model, all non-zero values are identical, $49m/s$, and clipped in $[-60, 60]$. (c) Starting model.

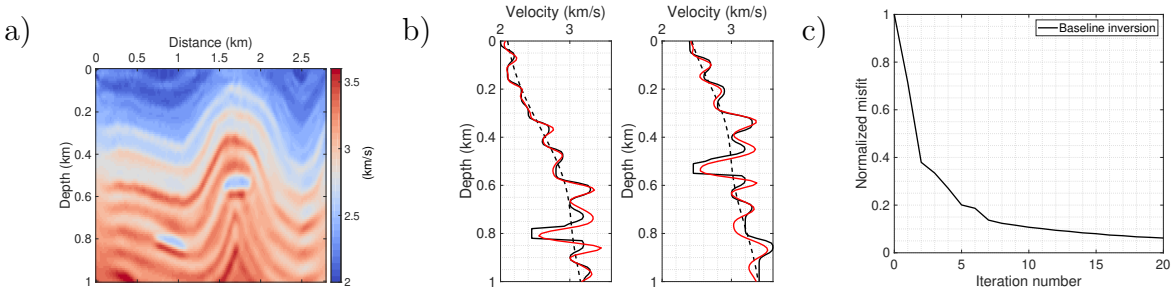

Figure 4.4: (a) Inverted baseline model. (b) The solid black lines are the true model, the dash black lines are starting models, and the red lines are inverted baseline models at distance 970m (left) and 1700m (right). (c) Misfits versus iteration numbers for baseline inversion.

4.5.1 Noise-free data tests

In this subsection, noise-free data sets with perfectly repeatable acquisition geometries are employed. The stepsizes for baseline and monitor inversions in the PRS are plotted in Figure 4.5, which are different in twice inversions. The results of different strategies are displayed in Figure 4.6. We observe, compared with the PRS, the SSPRS, DDS, and SSCMS can significantly eliminate artifacts caused by the convergence difference, and the CMS and CDS can also reduce the artifacts to some extent. The best result is produced by the SSPRS (Figure 4.6b). It contains fewer coherent artifacts as compared with others including the DDS (Figure 4.6c), which is target-oriented.

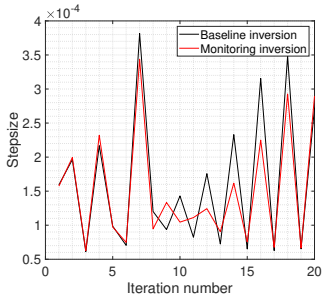

Figure 4.5: Stepsizes for baseline and monitor inversions in the parallel strategy (PRS).

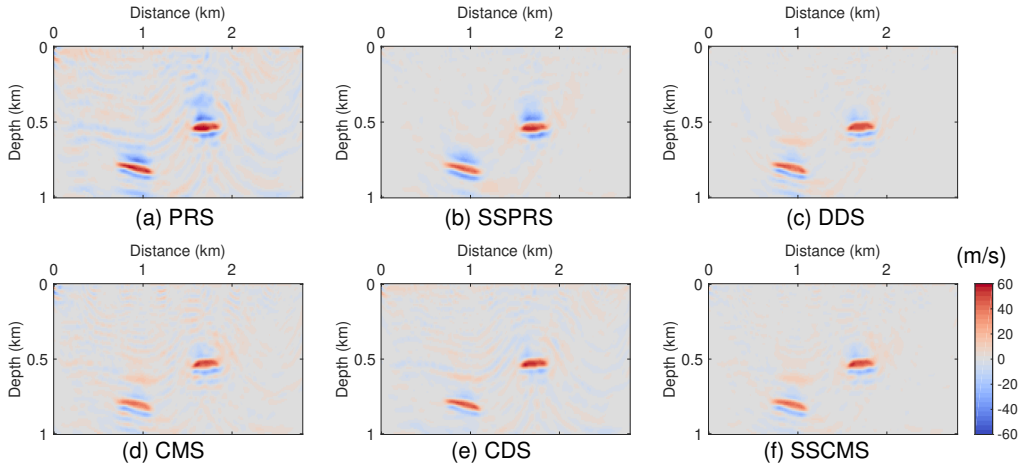

Figure 4.6: Inverted time-lapse results using noise-free data for different strategies: (a) parallel strategy (PRS), (b) stepsize-sharing parallel strategy (SSPRS), (c) double-difference strategy (DDS), (d) common-model strategy (CMS), (e) central-difference strategy (CDS), (f) stepsize-sharing common-model strategy (SSCMS).

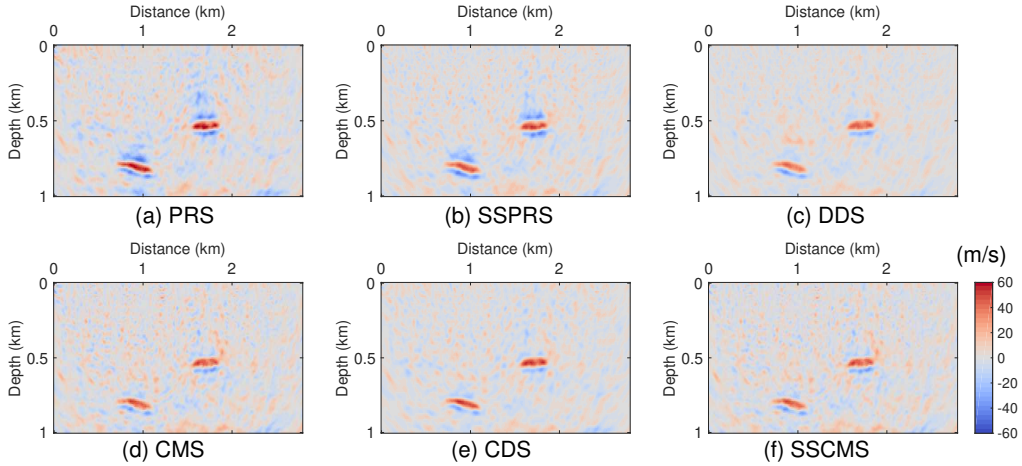

Figure 4.7: Inverted time-lapse results using data with SNR=20dB for different strategies. All figures are clipped in the same color bar.

4.5.2 Non-repeatable random noise

In this subsection, noisy data sets with a perfectly repeatable acquisition geometries are considered. We incorporate into the simulated data Gaussian random noise, with different realizations of the same noise level for both baseline and monitor data. In Figure 4.7-4.9, we display the results of different strategies using data sets with SNR (signal-to-noise ratio) 20dB, 10dB, and 5dB. We observe all six strategies have a similar anti-noise response. For the cases of SNR 20dB and 10dB, the time-lapse velocity changes can still be clearly recognized, nevertheless, in the case of SNR 5dB, the artifacts almost submerge the time-lapse changes. In practice, the real field data often have a lower SNR which would further hurt the quality of inverted time-lapse changes.

4.5.3 Non-repeatable source positions

In this subsection, we analyze noise-free data sets generated with acquisition geometries that are perturbed between baseline and monitor surveys. We assume sparse source positions and dense receiver positions, and model non-repeatability through the positions of the sources. Receivers are located in each surface cell grid of the model, and the seven monitor source

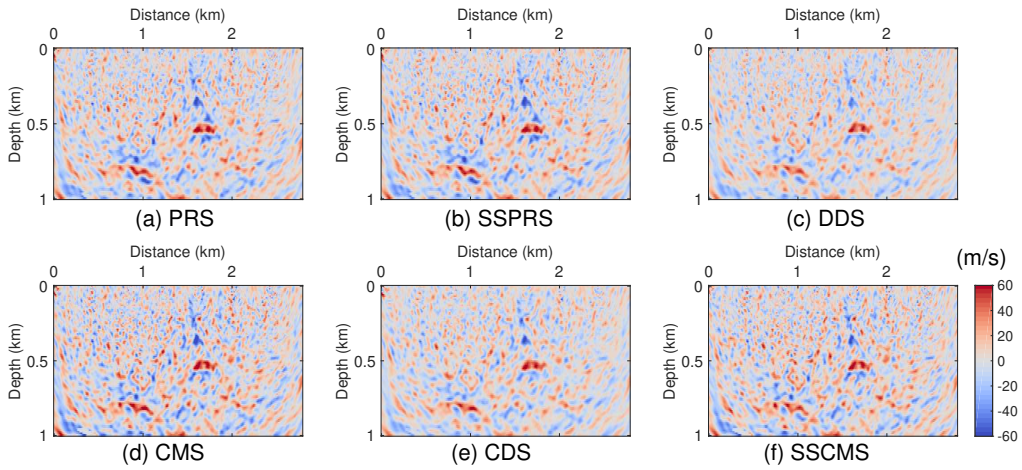

Figure 4.8: Inverted time-lapse results using data with SNR=10dB for different strategies. All figures are clipped in the same color bar.

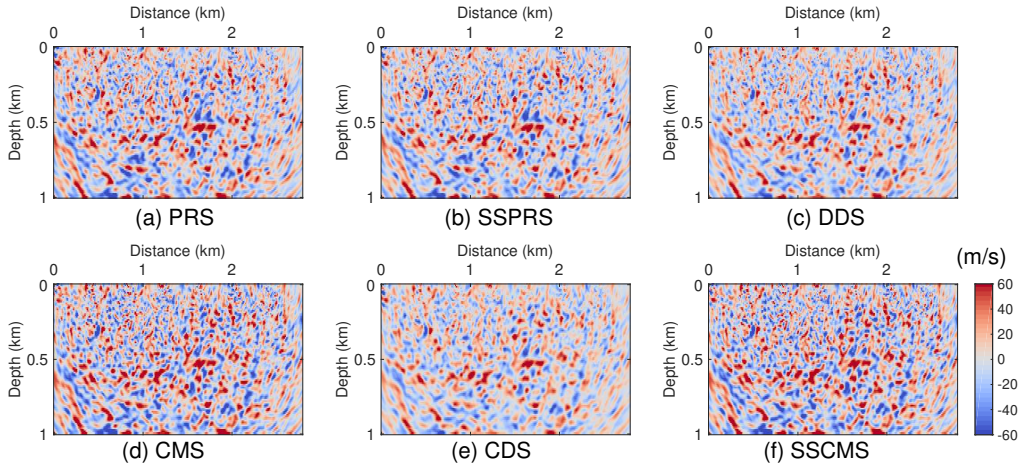

Figure 4.9: Inverted time-lapse results using data with SNR=5dB for different strategies.

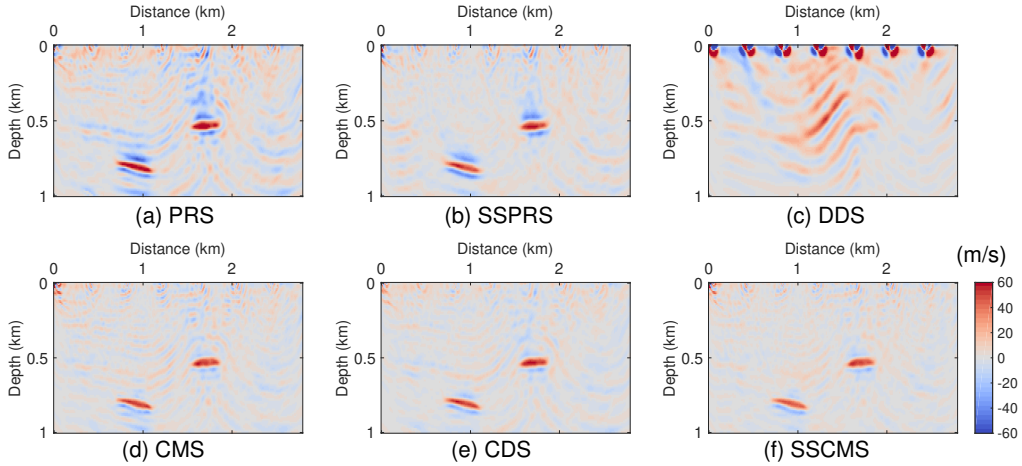

Figure 4.10: Inverted time-lapse results of different strategies in the case that monitor source locations as a whole have been moved to the right of baseline source locations by 10m. All figures are clipped in the same color bar.

positions will move the same distance to the right from the baseline source position.

In Figure through 4.10 to 4.12, we plot the time-lapse results of different strategies, and the source position differences between twice surveys are, respectively, 10m, 20m, and 40m, that is, monitor source locations as a whole have been moved to the right by 10m, 20m, and 40m, respectively, from the original positions which are the same as that of baseline sources. We observe that the DDS is sensitive to the source position non-repeatability. And the others can still effectively reflect the velocity changes, although the inversion quality has a certain decrease with the source position difference increasing. In the near-surface area, we observe obvious acquisition footprints, and the footprints in the CMS, CDS, and SSCMS are slighter than that in the PRS, SSPRS, and DDS.

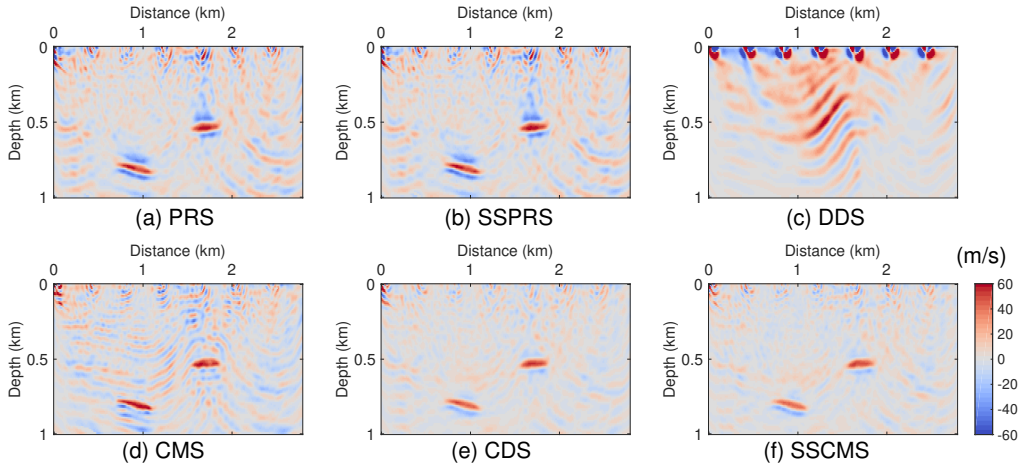

Figure 4.11: Inverted time-lapse results of different strategies in the case that monitor source locations as a whole have been moved to the right of baseline source locations by 20m.

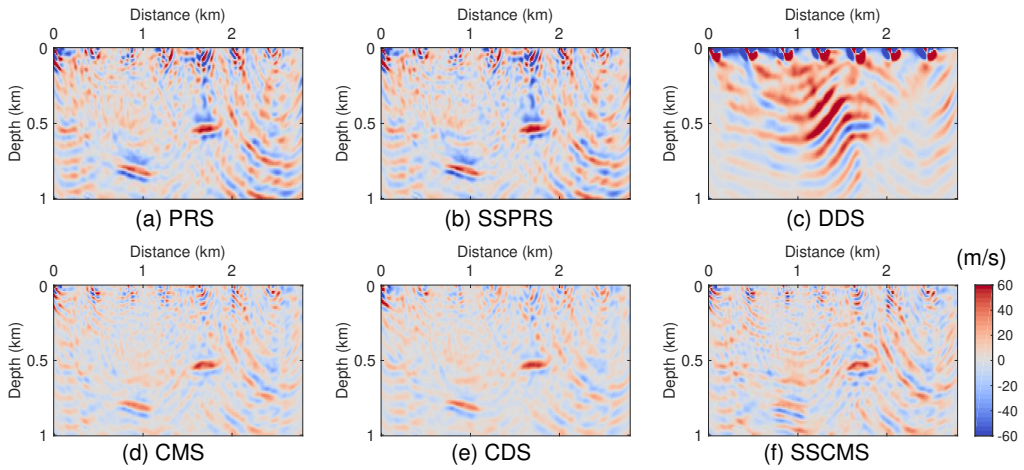

Figure 4.12: Inverted time-lapse results of different strategies in the case that monitor source locations as a whole have been moved to the right of baseline source locations by 40m.

4.5.4 Biased starting models

In the previous tests, the starting model (Figure 4.3c) is unbiased, being a smoothed version of the true baseline model (Figure 4.3a). In this subsection, we employ two biased starting models (Figure 4.13a and 4.14a) to test the resilience of the different strategies to bias, using baseline and monitor data sets that are noise-free and are of identical acquisition geometries. The first biased starting model, plotted in Figure 4.13a, is equal to the unbiased model (Figure 4.3c) plus $100m/s$. The inverted model is plotted in Figures 4.13b and c. The second biased starting model, plotted in Figure 4.14a, is equal to the unbiased model minus $100m/s$. The corresponding inverted baseline model is plotted in Figures 4.14b and c. In Figures 4.15a-c, the observed and predicted baseline data are plotted. We observe that the biased starting models tend to produce serious deviations in the inversion results, but the main geological structures are still visible; all of the results tend to produce a good data fit, suggesting that a local minimum has been identified.

In Figures 4.16 and 4.17, the inverted time-lapse models derived using the different strategies are plotted. We observe the CMS, CDS, and SSCMS, which involve 4 of FWI calculations, fail to give meaningful results. However, the reservoir changes are still recognizable in the results of the PRS, SSPRS, and DDS, with the SSPRS and DDS producing fewer artifacts than the PRS. Since the regions above the two reservoirs in Figure 4.14 are more accurately estimated than in Figure 4.13, the results in Figure 4.17 have in general fewer artifacts than those in Figure 4.16. Here the true reservoir change is visible. However, errors in the reservoir position and the strong erroneous negative velocity change near the reservoirs would likely mislead a reservoir interpretation.

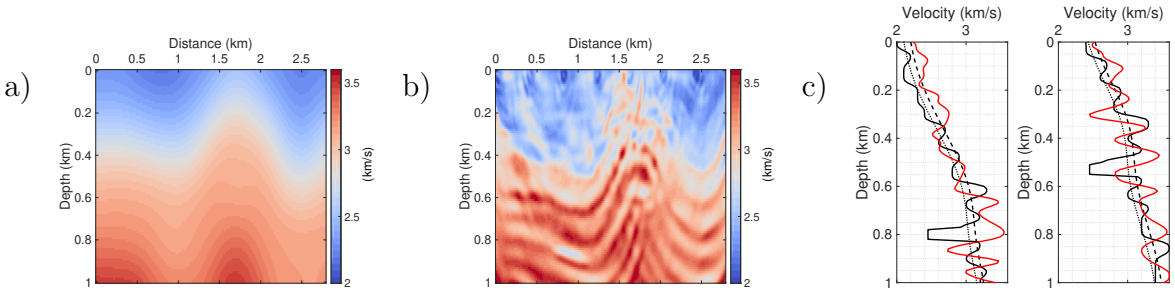

Figure 4.13: (a) The biased starting model which equals the unbiased model (Figure 4.3b) plus $100m/s$. (b) The corresponding inverted baseline model. (c) Traces abstracted at distances 970m (left) and 1700m (right), the solid black lines are the true model, the dash black lines are the biased starting models, the dot black lines are the unbiased starting models, and the red lines are inverted baseline models.

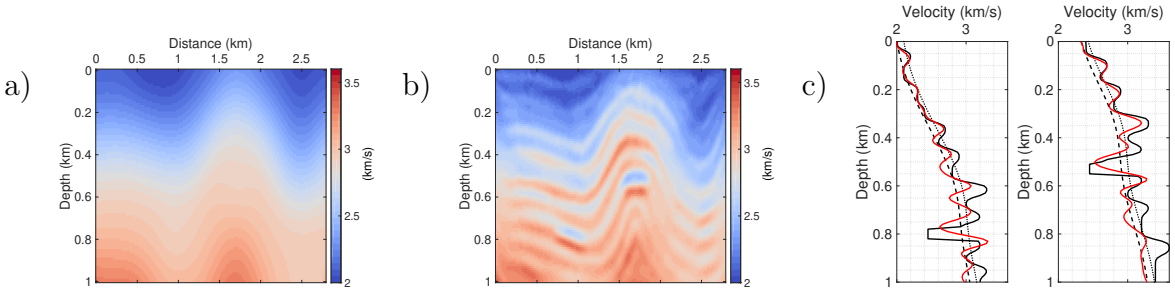

Figure 4.14: (a) The biased starting model which equals the unbiased model (Figure 4.3b) minus $100m/s$. (b) The corresponding inverted baseline model. (c) The traces abstracted at the distances 970m (left) and 1700m (right), and legends are the same as that in Figure 4.13.

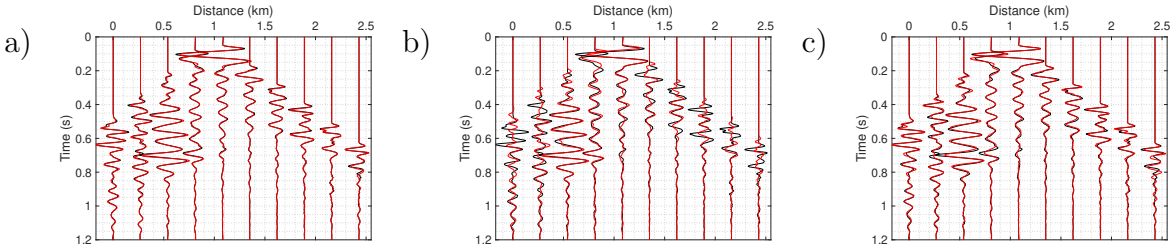

Figure 4.15: Black curves are observed baseline data, and red curves are synthetic data of the inverted baseline models using (a) the unbiased starting model in Figure 4.3b, (b) the biased starting model in Figure 4.13a, and (c) the biased starting model in Figure 4.14a. All synthetic data can fit the observed data well.

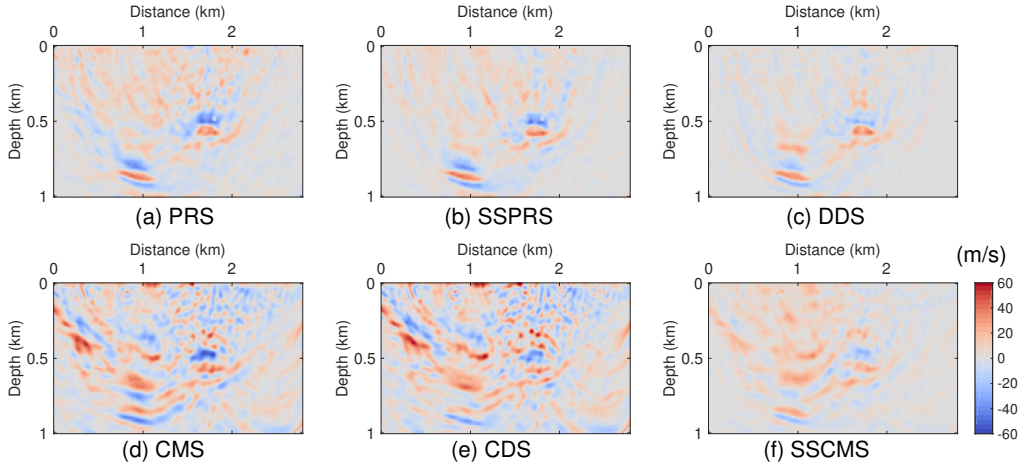

Figure 4.16: Inverted time-lapse results of different strategies in the case of starting model (Figure 4.13a) is $100m/s$ larger than the unbiased one.

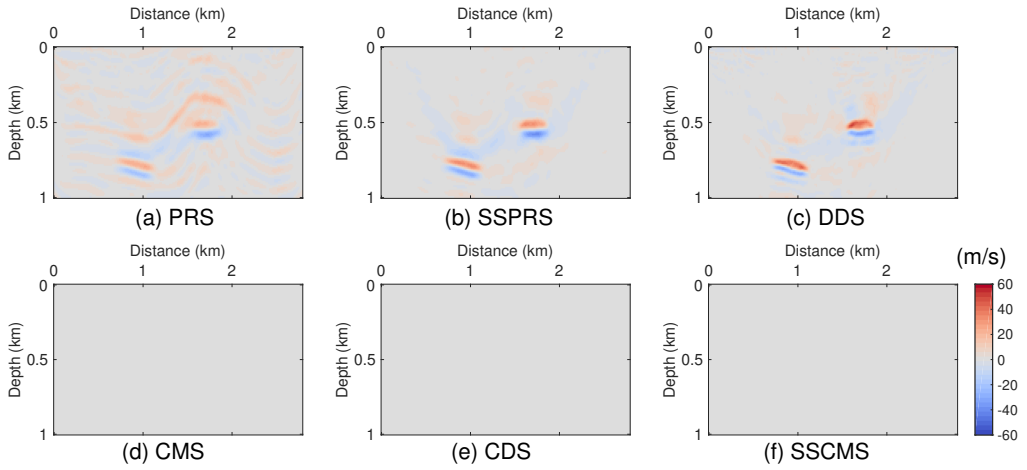

Figure 4.17: Inverted time-lapse results of different strategies in the case of starting model (Figure 4.14a) is $100m/s$ smaller than the unbiased one.

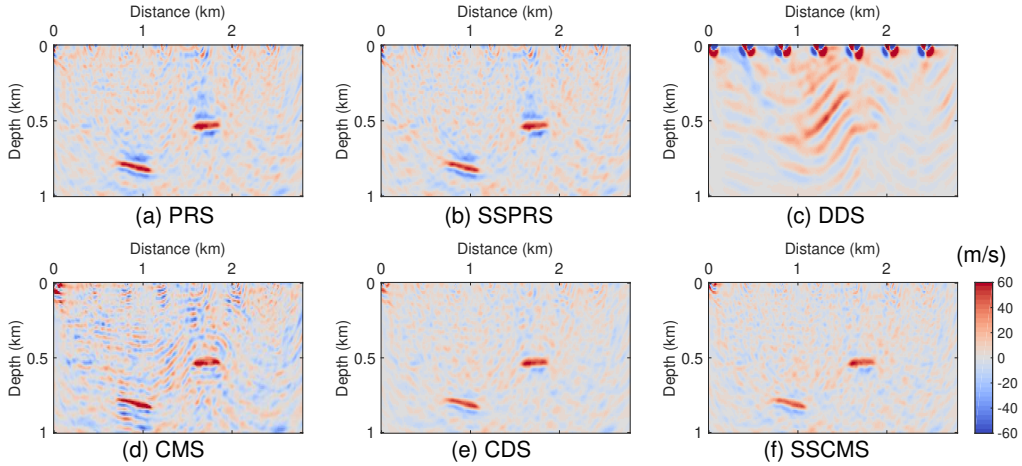

Figure 4.18: Inverted time-lapse results of different strategies in the case of SNRs for both baseline and monitor data sets are 20dB, the distances of monitor source locations are $10m$ larger than that of baseline source locations, and starting model (Figure 4.3c) is the unbiased one. All figures are clipped in the same color bar.

4.5.5 Combined random noise, non-repeatable source locations, and biased starting models

In Figures 4.18, 4.19, and 4.20, we test the response of different strategies to combinations of all of the perturbing conditions tested previously. A SNR=20dB is selected for both baseline and monitor data sets; the monitor source locations are $10m$ larger than baseline source locations, and all unbiased and biased starting models are employed (Figure 4.3c for Figure 4.18, Figure 4.13a for Figure 4.19, and Figure 4.14a for Figure 4.20). In Figure 4.18, we observe, except for the DDS, all strategies produce interpretable results, with artifacts caused by CMS relatively stronger than those of others. Comparing the results in Figure 4.19 and 4.20 with those in Figure 4.18, we find the impact on time-lapse inversion caused by the biased starting model is much heavier than those caused by the other factors. In Figure 4.19, we can only recognize the time-lapse changes in the results of SSPRS and SSCMS. Only the SSPRS gives a meaningful result in Figure 4.20; this result appears to be noticeably different from the others tested.

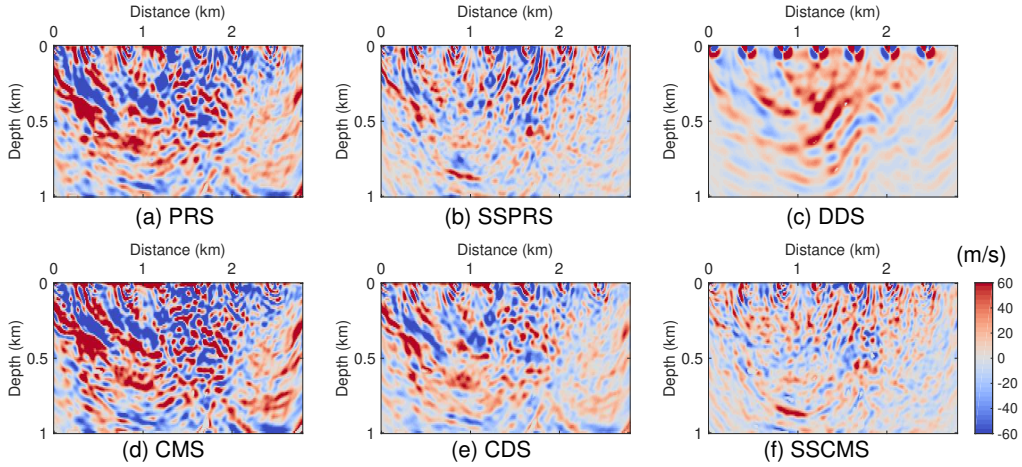

Figure 4.19: Inverted time-lapse results of different strategies in the case of SNRs for both baseline and monitor data sets are 20dB, the distances of monitor source locations are $10m$ larger than that of baseline source locations, and starting model (Figure 4.13a) is $100m/s$ larger than the unbiased one.

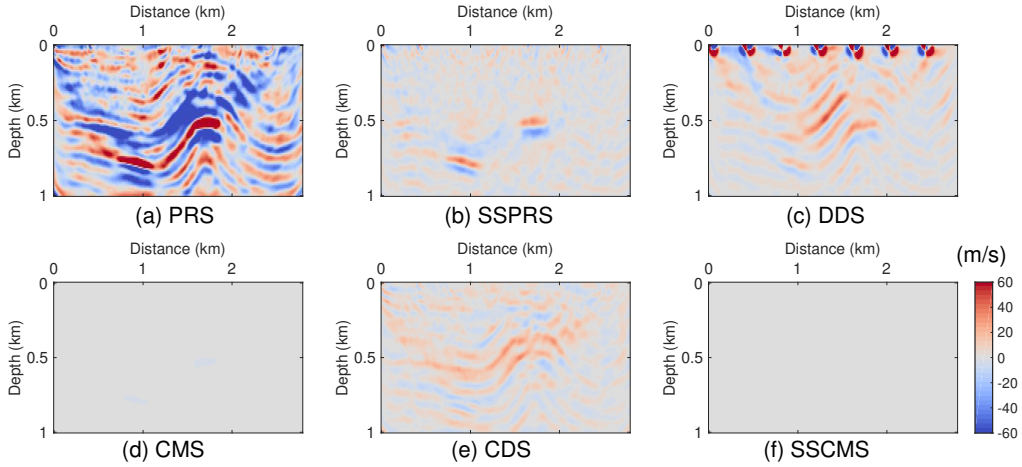

Figure 4.20: Inverted time-lapse results of different strategies in the case of SNRs for both baseline and monitor data sets are 20dB, the distances of monitor source locations are $10m$ larger than that of baseline source locations, and starting model (Figure 4.14a) is $100m/s$ smaller than the unbiased one.

4.6 Discussion

The stepsize-sharing methods are developed based on the steepest descent optimization, their adaption to other optimizations should be further proved, since a key assumption in them is the updating direction is linear with respect to the observed data, which could not be satisfied in other optimizations. Furthermore, the stepsize-sharing methods do not take the source wavelet non-repeatability into account, hence, eliminating source wavelet non-repeatability between baseline and monitor shot gathers is recommended before performing them. An available means of doing it has been proposed by (Fu and Innanen, 2023).

Our results are confined to the acoustic approximation, and single land synthetic model. Practical time-lapse issues such as seasonal change in the near surface, and overburden variations have not been included. Extensions to elastic, whether full elastic FWI or acoustic FWI with elastic data (Zhou and Lumley, 2021a), and both land and marine models, and systematic study of geological variations not related to the target, are important next steps. However, the numerical responses we observe in our study, and the differential sensitivities, appear to be quite general, and both (1) suggest the value of stepsize sharing to time-lapse inversions in general, and (2) reflect fundamental differences in the sensitivities of the various strategies.

During the derivation of the stepsize-sharing methods, we assume that the gradient (e.g. in equation 4.15) is only in terms of parameter model and observed data, and receiver/source positions and wavelets for baseline and monitor survey are perfectly repeatable. In fact, the gradient also depends on receiver/source positions and wavelet, and their non-repeatability can affect the performance of stepsize-sharing methods. According to the tests in our study, we observe that a certain difference between receiver/source positions is acceptable. As for the difference in wavelets, we believe the impact is also limited since the gradient is preconditioned by the diagonal approximation of the Hessian matrix, in which the wavelet effect can be eliminated.

In this study, we do not incorporate any constraint to enhance the inversion results.

We believe efforts of other researchers can be incorporated in present strategies, such as model-difference regularizations (Willemsen and Malcolm, 2015; Maharramov et al., 2016), target-oriented schemes (Raknes and Arntsen, 2014; Asnaashari et al., 2015; Willemsen, 2017; Huang et al., 2018). Moreover, a regularized FWI algorithm could be incorporated into our work to improve the anti-noise property of our method.

4.7 Conclusions

There are no significant differences in the responses of the different strategies to random noise. And there is a general sensitivity of all methods to an increase in the strength of random noise. Source positioning non-repeatability produces artifacts in time-lapse FWI inversion results, but its differential influence on the various time-lapse inversion strategies we tested is limited, except for the double-difference strategy (DDS). DDS is very sensitive to acquisition non-repeatability. The impact on time-lapse inversion of bias in the starting model is much stronger than that of the other factors. The parallel strategy (PRS) exhibits strong artifacts caused by differences in the convergence properties of the two inversion procedures, and is sensitive to biased starting models. The double-difference strategy (DDS) is robust to bias in the starting models, and is able to eliminate artifacts caused by variation in convergence properties of the two basic inversion steps, but it is strongly sensitive to source position non-repeatability. The common-model strategy (CMS) and central-difference strategy (CDS) both exhibit a capacity to avoid artifacts resulting from differing convergence properties, and are stable with respect to unrepeatable source locations, but are strongly sensitive to bias in the starting model. Stepsize-sharing as a strategy produces results that are robust to convergence differences; the SSPRS in addition is insensitive to non-repeatable source locations and biased starting models. Furthermore given that SSPRS through its sharing incurs half of the time cost of seeking stepsizes compared with the PRS and DDS, and the total computational cost of SSPRS is less than half of that of the CMS and CDS. Although

the stepsize-sharing common-model strategy (SSCMS) also exhibits fewer artifacts caused by the convergence differences, and is robust to non-repeatable source locations, it is sensitive to bias in starting models.

Chapter 5

A time-domain, multi-source Bayesian/Markov Chain Monte Carlo formulation of time-lapse seismic waveform inversion

5.1 Abstract

Seismic waveform inversion has been shown to be well-suited for identification and characterization of time-lapse changes in a reservoir. However, the subtle medium variations associated with enhanced oil recovery and/or CO_2 storage problems give uncertainty quantification within such inverse approaches a heightened importance. To analyze both of these features of the time-lapse inverse problem, we formulate a Bayesian full waveform inversion (FWI) procedure, based on a Markov chain Monte Carlo (MCMC) algorithm. The formulation employs several existing strategies, including use of a double-difference time-lapse FWI (DDFWI), incorporation of time-domain multi-source data, and application of a local-updating target-oriented inversion. However, it incorporates these within a stochastic

framework, involving computation of model covariance with an adaptive Metropolis algorithm, and a method to estimate data error statistics based on the features of time-lapse difference data is incorporated. A random walk Metropolis-Hastings MCMC is adopted for optimization. In conventional, i.e., deterministic, DDFWI, inversions are carried out for the baseline and monitoring models; in the MCMC approach, a deterministic FWI procedure is carried out for the baseline model, and the MCMC algorithm is applied in the monitoring inversion stage; the final time-lapse model is the difference between these. A feasibility study is carried out using synthetic 2D acoustic models and data, including both time-lapse model estimation and uncertainty quantification. We compare the MCMC approach with conventional deterministic optimization DDFWI, and remark on benefits derived, which appear to justify the expanded complexity and cost of a global approach. In addition to the availability of posterior distributions, which are critical for assessment of the estimations, we observe that the MCMC approach tends to produce monitoring images with clearer edges and fewer coherent errors.

5.2 Introduction

Full-wave inversion, or FWI (Lailly et al., 1983; Tarantola, 1984; Virieux and Operto, 2009), has been employed extensively in geophysics. Time-lapse FWI, designed to detect or characterize production, injection, or storage-related property changes within a subsurface volume, in turn has become an important mode of FWI application. Time-lapse FWI normally involves two or more inversions, one baseline inversion and one for each monitoring stage; a time-lapse model is then produced by subtracting the baseline model from the monitoring model. Within this structure, time-lapse FWI can be classified into three basic categories: parallel difference FWI (in which baseline and monitoring data are incorporated independently, both using the same starting model), sequential difference FWI (in which baseline data and monitoring data are inverted independently, but the latter using the inverted base-

line model) (Oldenborger et al., 2007; Routh and Anno, 2008), and the double-difference FWI (DDFWI) (Watanabe et al., 2004; Onishi et al., 2009; Denli and Huang, 2009; Zheng et al., 2011; Asnaashari et al., 2011; Routh et al., 2012; Raknes et al., 2013; Maharramov and Biondi, 2014; Raknes and Arntsen, 2014; Yang et al., 2016) adopted in this work.

In the DDFWI procedure, the baseline inversion is the same as those of the other two strategies. The input elements are the baseline data and a reasonable starting model, using which a standard FWI inversion is carried out. In the monitoring inversion, DDFWI makes use of composited data, built from the difference data (the difference between the monitoring data and the baseline data) and synthetic data computed with the inverted baseline model. The approach is motivated by the exposure of both parallel difference FWI and sequential difference FWI to coherent errors, caused by two different convergence histories (Asnaashari et al., 2014; Yang et al., 2015a), which are reduced in the DDFWI approach. DDFWI requires high repeatability between the baseline and monitoring surveys (Yang et al., 2015a), but aspects of this sensitivity can be addressed. For instance, Fu et al. (2020) developed a double-wavelet DDFWI method to manage wavelet differences between baseline and monitoring data.

Inversion based on deterministic optimization (DO) methods, in particular FWI, is an increasingly common tool for determination of the physical properties of the subsurface media, and a growing number of successful examples have been reported. However, their success depends strongly on the starting model (Virieux and Operto, 2009), especially when low-frequency data are unavailable. Stochastic global optimization methods have the capacity to avoid local minima, and provide uncertainty quantification, but at the cost of a very significant increase in numbers of simulations. Consequently, systematic analyses and applications of global optimization methods to seismic problems are largely restricted to 1D models (Hong and Sen, 2009; Afanasiev et al., 2014; Aleardi and Mazzotti, 2016; Ray et al., 2016). To make these more tractable, and to accommodate multidimensional, and otherwise more complex and realistic models, combinations of fully global approaches with additional

strategies are generally required. Datta and Sen (2016) reduce the number of unknown parameters by sparsely parameterizing the velocity as several interfaces, and then estimate these parameters with a global very fast simulated annealing algorithm. Sajeva et al. (2016) apply a two-grid technique, involving a coarse grid for the subsurface model but a finer grid for the forward modeling, and invoke a genetic algorithm to lower the model dimension (an approach also applied by Mazzotti et al., 2016). Biswas and Sen (2017) parameterize the velocity model with Voronoi cells, represent the parameters with certain nuclei points, and then utilize a custom-developed Reversible Jump Hamiltonian Monte Carlo algorithm for optimization. Ely et al. (2018) employ a fast field expansion method to accelerate wavefield simulations; da Silva et al. (2019) solve for a sparse macro Q-model via a quantum particle-swarm optimization approach; and, Visser et al. (2019) parameterize models in terms of local velocity, thickness, and lower interface dip angles within a Bayesian transdimensional approach, sampled via a Markov chain Monte Carlo (MCMC) approach, then stitch the local models together into a final 2D model.

Globally-optimized full waveform applications have to date largely been aimed at constructing starting models, via a sparse parameterization, which are then subsequently used in a more standard DO FWI. However, there are examples in which global optimization methods have been invoked to invert for the 2D models directly. For example, Stuart et al. (2019) combine a two-stage MCMC with a coarse-grid filter to enhance the acceptance rate of MCMC, solving for layered models. Also, Gebraad et al. (2020) formulated a Bayesian elastic FWI, based on Hamiltonian Monte Carlo for converted wave data. Zhang and Curtis (2021) apply variational inference to solve acoustic FWI problems including uncertainties.

In this paper, we discuss the formulation of a stochastic, global solution to the full waveform seismic time-lapse problem. Our aim is to solve for relatively general 2D velocity models, without applying a sparse parameterization. We set up a Bayesian MCMC problem, and employing a Metropolis-Hastings algorithm, similar to that of Kotsi et al. (2020). However, rather than a single-shot, single-frequency data set, we consider local updating with

time domain multi-source shot gathers as input. Furthermore, we introduce a new estimation method for data error statistics based on the natural features of time-lapse data. The method is exemplified with simulated data, and we compare results of its application with otherwise comparable DO FWI-based time lapse inversion of the same data. This study is an expansion of Fu and Innanen (2021).

5.3 Theory and background

5.3.1 Standard full-waveform inversion with deterministic optimization

Standard FWI, which is an example of inversion based on deterministic optimization (DO), starts from a given model $\mathbf{m}_0$ and seeks through local gradient information a model $\mathbf{m}$ that matches simulated data $\mathbf{d}_{syn}(\mathbf{m})$ to observed data $\mathbf{d}_{obs}$. This is most often achieved by minimizing the L2 norm of the data residual $\delta\mathbf{d} = \mathbf{d}_{syn}(\mathbf{m}) - \mathbf{d}_{obs}$:

$$E(\mathbf{m}) = \frac{1}{2}\delta\mathbf{d}^T\delta\mathbf{d}, \tag{5.1}$$

where T denotes the transpose operation. For constant-density acoustic FWI, the model $\mathbf{m}$ in equation 5.1 is a vector of pressure wave velocity values $v(\mathbf{x})$ in which $\mathbf{x}$ is position vector, and $\mathbf{d}_{syn}$ is a vector of values of the simulated pressure field $P(\mathbf{x}, t)$, evaluated at receiver positions. In this paper, the simulated field $P(\mathbf{x}, t)$ is obtained by solving the time-domain constant-density acoustic wave equation

$$\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 P(\mathbf{x}, t)}{\partial t^2} - \bigtriangledown^2 P(\mathbf{x}, t) = s(t)\delta(\mathbf{x} - \mathbf{x}_s), \tag{5.2}$$

where $\bigtriangledown^2$ is the Laplacian operator, $s(t)$ is the source wavelet, δ is the Dirac delta function, and $\mathbf{x}_s$ is the source position vector. A finite-difference method (eighth order in space and

second order in time), coupled with a perfectly matched layer (PML) boundary condition are used to solve this wave equation throughout this paper.

A wide range of optimization methods can be applied within FWI. These include Newton-type optimizations (e.g., full Newton and Gauss-Newton methods), gradient-based optimizations (e.g., steepest-descent, or SD, and non-linear conjugate-gradient, or NCG, methods), Quasi-Newton optimizations (e.g., BFGS and l-BFGS methods), truncated-Newton optimations, etc. In the instances in this paper in which we carry out a DO FWI, we restrict ourselves to SD optimizations with a simple illumination-compensating preconditioning step. Following the adjoint-state method (Tarantola, 1984; Bunks et al., 1995; Plessix, 2006), and preconditioning the gradient with the approximate inverse Hessian of Shin et al. (2001), avoiding explicit Hessian calculations, a DO FWI model update is computed via

$$\Delta v(\mathbf{x}) = -\mu \sum_{r=1}^{ng} \sum_{i=1}^{ns} \frac{2}{v(\mathbf{x})^3} \frac{\int_0^{t_{max}} dt[\ddot{P}_f(\mathbf{x},t;\mathbf{x}_s) P_b(\mathbf{x},t;\mathbf{x}_r)]}{\int_0^{t_{max}} dt[\ddot{P}_f(\mathbf{x},t;\mathbf{x}_s) \ddot{P}_f(\mathbf{x},t;\mathbf{x}_s) + \lambda I_{max}]}, \tag{5.3}$$

where μ is the step length obtained by the line search technique; ng, ns are the numbers of receivers and shots, respectively; t_{max} is the maximum forward/backward propagating time t of wavefields; $\mathbf{x_r}$ is the receiver position vector; $P_f(\mathbf{x},t;\mathbf{x}_s)$ is the forward wavefield due to the source at $\mathbf{x}_s$ and $\ddot{P}_f(\mathbf{x},t;\mathbf{x}_s)$ is its the second derivative with respect to t; $P_b(\mathbf{x},t;\mathbf{x}_r)$ is the backward/time-reversal wavefield due to the data residual $\delta\mathbf{u}$ at position $\mathbf{x}_r$; $I_{max} = \max_{\mathbf{x},t} \left[\ddot{P}_f(\mathbf{x},t;\mathbf{x}_s)\ddot{P}_f(\mathbf{x},t;\mathbf{x}_s)\right]$ is the square of maximum absolute value in the derivative of forward propagation wavefield; λ is the damp factor.

5.3.2 Time-lapse full waveform inversion

In double-difference FWI, or DDFWI, the first inversion is for the baseline model; the input elements are the baseline data and a starting model. The second is for the monitoring model; the inverted baseline model is used as the starting model, but a composited data set is used

as an alternative to the monitoring data. The vector of composited data set is:

$$\mathbf{d}'_{obs2} = \mathbf{d}_{syn}(\mathbf{m}'_1) + (\mathbf{d}_{obs2} - \mathbf{d}_{obs1}), \tag{5.4}$$

where $\mathbf{d}_{syn}(\mathbf{m}'_1)$ is the vector of synthetic data predicted with the inverted baseline model $\mathbf{m}'_1$, and $\mathbf{d}_{obs2}$ and $\mathbf{d}_{obs1}$ are the observed monitoring and baseline data respectively. The combination $\mathbf{d}_{obs2} - \mathbf{d}_{obs1}$ is the vector of difference data. During the monitoring inversion stage, the mistfit function

$$E_{DDFWI}(\delta\mathbf{m}') = \frac{1}{2}||\mathbf{d}_{syn}(\mathbf{m}'_1 + \delta\mathbf{m}') - \mathbf{d}'_{obs2}||_2^2, \tag{5.5}$$

where $||\ ||_2^2$ denots L2 norm, and $\delta\mathbf{m}'$ is a perturbation of the background model $\mathbf{m}'_1$, is employed. Substituting the relationship $\mathbf{d}_{syn}(\mathbf{m}'_1 + \delta\mathbf{m}') = \delta\mathbf{d}_{syn}(\delta\mathbf{m}') + \mathbf{d}_{syn}(\mathbf{m}'_1)$, where $\delta\mathbf{d}_{syn}(\delta\mathbf{m}')$ is the synthetic difference data vector, into equation 5.5, we obtain

$$E_{DDFWI}(\delta\mathbf{m}') = \frac{1}{2}||\delta\mathbf{d}_{syn}(\delta\mathbf{m}') - \delta\mathbf{d}||_2^2, \tag{5.6}$$

where $\delta\mathbf{d} = \mathbf{d}_{obs2} - \mathbf{d}_{obs1}$ is the observed difference data vector. By minimizing equation 5.5, we are in essence minimizing the residual of two difference data vectors. The inverted time-lapse model is $\delta\mathbf{m}'$.

This inverted time-lapse model $\delta\mathbf{m}'$ is distinct from the "true" time-lapse model $\delta\mathbf{m}$. The perturbation $\delta\mathbf{m}'$ is that by which the difference data $\delta\mathbf{d}_{syn}(\delta\mathbf{m}')$ is linearly developed from $\mathbf{d}_{syn}(\mathbf{m}'_1)$. In contrast, $\delta\mathbf{m}$ relates the difference data $\delta\mathbf{d}$ and the background data $\mathbf{d}_{obs1}$, assuming the baseline and monitoring surveys are identical. When $\mathbf{d}_{syn}(\mathbf{m}'_1)$ is close to $\mathbf{d}_{obs1}$, or equivalently when $\mathbf{m}'$ is close to the true baseline model, $\delta\mathbf{m}'$ will be close to $\delta\mathbf{m}$, and vice versa. Therefore, a relatively accurate inverted baseline model is important for DDFWI (Asnaashari et al., 2011). How the initial monitoring model impacts the final inverted time-lapse model is part of the analysis in this study.

In a conventional DO DDFWI, both the baseline inversion and monitoring inversion are based on the DO FWI, i.e., input a starting model $\mathbf{m}_0$ and the observed data $\mathbf{d}_{obs1}$ into the DO FWI to obtain the inverted baseline model $\mathbf{m}_1'$, and then input the inverted $\mathbf{m}_1'$ and the new observed data $\mathbf{d}_{obs2}'$ into the DO FWI to obtain the inverted monitoring model. In our method, called MCMC DDFWI, we still use the DO FWI with inputs $\mathbf{m}_0$ and $\mathbf{d}_{obs1}$ to obtain $\mathbf{m}_1'$, but we obtain the monitoring model by inputting $\mathbf{m}_1'$ and $\mathbf{d}_{obs2}'$ into the Bayes' frame solved by the MCMC algorithm. The reasons for using DO FWI in baseline inversion are that employing a stochastic global optimization method to invert the whole 2D model is still too computationally expensive and whether the stochastic global optimization method is capable of that mission is still uncertain because of the large parameter number and high complexity of the inverse problem.

5.3.3 Bayesian inference and MCMC

In this section, we formulate an inverse procedure based on the above concepts, but in a Bayesian setting (Tarantola, 2005). This involves combining data with model prior information to infer optimal models; its output, which is in the form of posterior probability distributions, also allows uncertainty estimation. The prior probability density function, or PDF $p(\mathbf{m}_2)$, is called the model prior, and reflects our state of information about the monitoring model $\mathbf{m}_2$ independent of the data. The conditional PDF $p(\mathbf{d}_{obs2}^{'}|\mathbf{m}_2)$, or likelihood function, is the probability of observing a particular set of observed data $\mathbf{d}_{obs2}^{'}$ over a range of models, and is how simulation and synthetic data are incorporated. These two quantities are combined to determine the numerator of the posterior PDF $p(\mathbf{m}_2|\mathbf{d}_{obs2}^{'})$ via Bayes' theorem:

$$p(\mathbf{m}_2|\mathbf{d}_{obs2}^{'}) = \frac{p(\mathbf{d}_{obs2}^{'}|\mathbf{m}_2)p(\mathbf{m}_2)}{p(\mathbf{d}_{obs2}^{'})} \propto p(\mathbf{d}_{obs2}^{'}|\mathbf{m}_2)p(\mathbf{m}_2), \tag{5.7}$$

where $p(\mathbf{d}_{obs2}^{'}) = \int p(\mathbf{d}_{obs2}^{'}|\mathbf{m}_2)p(\mathbf{m}_2)d\mathbf{m}_2$ is the marginal likelihood or model evidence. Because it is independent of the model and thus acts as a scalar normalization factor, it is

neglected here. We will proceed assuming that the prior PDF $p(\mathbf{m}_2)$ and likelihood function $p(\mathbf{d}'_{obs2}|\mathbf{m}_2)$ are the Gaussian or normal distributions

$$p(\mathbf{m}_2) \propto \exp\left\{-\frac{1}{2}(\mathbf{m}_2 - \bar{\mathbf{m}})^T \mathbf{C}_m^{-1}(\mathbf{m}_2 - \bar{\mathbf{m}})\right\}, \tag{5.8}$$

$$p(\mathbf{d}'_{obs2}|\mathbf{m}_2) \propto \exp\left\{-\frac{1}{2}(\mathbf{d}_{syn}(\mathbf{m}_2) - \mathbf{d}'_{obs2})^T \mathbf{C}_d^{-1}(\mathbf{d}_{syn}(\mathbf{m}_2) - \mathbf{d}'_{obs2})\right\}, \tag{5.9}$$

where $\mathbf{C}_m$ and $\mathbf{C}_d$ are, respectively, prior model covariance and data error covariance, and $\bar{\mathbf{m}}$ is the mean model. With equations 5.8 and 5.9, equation 5.7 becomes

$$p(\mathbf{m}_2|\mathbf{d}'_{obs2}) \propto \exp\{-\chi(\mathbf{m}_2)\}, \tag{5.10}$$

where

$$\chi(\mathbf{m}_2) = \frac{1}{2}(\mathbf{m}_2 - \bar{\mathbf{m}})^T \mathbf{C}_m^{-1}(\mathbf{m}_2 - \bar{\mathbf{m}}) + \frac{1}{2}(\mathbf{d}_{syn}(\mathbf{m}_2) - \mathbf{d}'_{obs2})^T \mathbf{C}_d^{-1}(\mathbf{d}_{syn}(\mathbf{m}_2) - \mathbf{d}'_{obs2}). \tag{5.11}$$

The inverse problem is to determine $p(\mathbf{m}_2|\mathbf{d}'_{obs2})$, or its parameters, given $\mathbf{d}'_{obs2}$, so that its form over a range of possible models $\mathbf{m}_2$ can be interpreted.

The inverse problem is therefore a sampling problem, where each sample involves a potentially intensive simulation. Monte-Carlo algorithms (e.g., Mosegaard and Tarantola, 1995) are a natural choice for this, however the number of samples such algorithms require to produce robust estimates of $p(\mathbf{m}_2|\mathbf{d}'_{obs2})$ make them impractical without adaptation. Markov Chain Monte Carlo (MCMC) methods reduce the number of samples required by selecting the current sample based on statistics produced from the previous accepted sample. The details of the acceptance/rejection varies across a range of algorithms. In our approach we employ the Metropolis algorithm (Metropolis et al., 1953), a special case of the Metropolis-Hastings (MH) MCMC algorithm (Hastings, 1970). The sampling procedure is as follows. We move about model space, from one model $\mathbf{m}_2$ to another $\mathbf{m}^*$, according to predefined

transitional probabilities in the proposal distribution $T(\mathbf{m}_2^*|\mathbf{m}_2)$. Each new model $\mathbf{m}_2^*$ is a proposal, to be accepted or rejected as a sample contributing to the posterior. In the Metropolis-Hastings algorithm, a criterion for acceptance or rejection of proposed models of the form

$$a(\mathbf{m}_2^*, \mathbf{m}_2) = \min\left[1, \frac{T(\mathbf{m}_2|\mathbf{m}_2^*)p(\mathbf{m}_2^*|\mathbf{d}'_{obs2})}{T(\mathbf{m}_2^*|\mathbf{m}_2)p(\mathbf{m}_2|\mathbf{d}'_{obs2})}\right], \tag{5.12}$$

is adopted. However, if symmetric distributions (e.g., Gaussian distribution) $T(\mathbf{m}_2|\mathbf{m}_2^*) = T(\mathbf{m}_2^*|\mathbf{m}_2)$ are used, this simplifies to the Metropolos algorithm

$$a(\mathbf{m}_2^*, \mathbf{m}_2) = \min\left[1, \frac{p(\mathbf{m}_2^*|\mathbf{d}'_{obs2})}{p(\mathbf{m}_2|\mathbf{d}'_{obs2})}\right]. \tag{5.13}$$

After selecting a starting model, an iterative procedure begins:

1. A random number u is drawn from a uniform distribution between 0 and 1.

2. The acceptance a is computed, requiring a simulation at the proposed model point $\mathbf{m}_2^*$. If $a > u$, the proposed model $\mathbf{m}_2^*$ is accepted as the current model $\mathbf{m}_2$, and stored. If $a <= u$, $\mathbf{m}_2^*$ is rejected and $\mathbf{m}_2$ is retained as the current model.

Steps 1-2 are repeated until a fixed maximum number of iterations is reached, at which point the accepted models are used to determine features of the expected distribution $p(\mathbf{m}_2|\mathbf{d}'_{obs2})$.

5.4 Methodology

The theoretical approaches reviewed in the previous section are combined in our approach, along with several practical considerations. In this section we review the key simplification strategies to be employed, and end with a description of our algorithm.

5.4.1 Local-updating target-oriented time-lapse inversion

DDFWI generally requires high repeatability between baseline and monitoring surveys. Common differences between the two, e.g., source wavelet differences, acquisition geometry changes, noise level, lead to error, and the appearance of difference model energy outside the target area. This is further exacerbated by inaccuracies in the inverted baseline model. An additional helpful constraint is to set the target area; this is possible in practice by, for instance, combining depth migration of the observed difference data with prior information of the reservoir. Raknes and Arntsen (2014) described the use of a simple edge detector on a depth-migrated section to automatically detect the target area, and bounded the model updating to these regions. Asnaashari et al. (2014) applied a Gaussian weighting constraining updates to a target area, after manually and approximately detecting the center of the target area. We follow this philosophy, and permit updating only in a predetermined target region. Target-oriented inversion is, then, an important feature in any time-lapse FWI methodology, but for MCMC DDFWI, it is vital, because of its reduction in the number of unknowns. In Metropolis-Hastings MCMC, the cost of producing independent samples grows as $O(n^2)$ for a growth in the number of unknowns of n (Creutz, 1988). The target-oriented strategy therefore reduces the computational expense by an order of magnitude. The failure rate of these stochastic algorithms also increases for high-dimension problems (Chib and Greenberg, 1995), meaning that any plausible reduction in the number of unknowns should generally be adopted when possible.

5.4.2 Multisource waveform inversion

Multisource waveform inversion reduces the cost of FWI by merging sources to reduce the number of simulations needed to form the update. In standard (DO) FWI, special treatment (e.g., the encoded multisource method of (Krebs et al., 2009)) of each source is required to suppress crosstalk artifacts in the final inverted model. In MCMC-based FWI, these treatments are not necessary, because the intermediate stage of gradient calculation, wherein

the shot crosstalk affects the model, does not take place. MCMC-based FWI decreases data misfit by directly accepting the stochastic samples, allowing us to more freely employ multisource seismic data.

5.4.3 Random walk sampling and sequential inversion

In random walk sampling (Metropolis et al., 1953) the proposal model $\mathbf{m}_2^*$ in equation 5.12 is given by

$$\mathbf{m}_2^* = \mathbf{m}_2 + \mathbf{u}, \tag{5.14}$$

where $\mathbf{m}_2$ is the present model, $\mathbf{u}$ is a random model drawn from the proposal distribution which, for instance, is a zero-mean Gaussian distribution with a standard deviation of 3 in our numerical tests. Normally, $\mathbf{m}_2$ is randomly initialized. However, for FWI, which is strongly nonlinear and high-dimensional, a random initial model will tend to lead to significant extra computation and increased chance of failure. We adopt the inverted baseline model, which in our approach is determined from DO FWI, as the initial model for monitoring inversion within MCMC DDFWI. In addition, we use a sequential scheme, updating only one unknown grid cell at a time, rather than all at a time, to update the monitoring model, in response to the high dimensionality of the model space. Even with target-oriented methods in place, this space can easily be many hundreds or thousands of dimensions. Generating a perturbation $\mathbf{u}$ suitable for all parameters is rarely possible under such conditions.

5.4.4 Model prior

In the Bayesian inference framework of equations 5.10 and 5.11, data error covariance $\mathbf{C}_d$ and model covariance $\mathbf{C}_m$ are both required as input. In our work, $\mathbf{C}_m$ is taken to be a diagonal matrix corresponding to independently distributed model parameters. $\mathbf{C}_m$ can be obtained via an adaptive Metropolis (AM) algorithm, as developed by Haario et al. (2001), in which the model means and covariances are calculated from previous models. In our work,

we utilize the all-samples-calculated model covariance, and the present model is taken as the mean value. This is similar to the adaptive proposal algorithm (Haario et al., 1999), in which the proposal distribution is centered on the present model, but the covariance is calculated from a fixed finite number of previous models. In the AM algorithm, the covariance for iteration $k+1$ proposal model is given by

$$\mathbf{C}_{k+1} = \frac{k-1}{k}\mathbf{C}_k + \frac{s_m}{k}(k\bar{\mathbf{m}}_{2k-1}\bar{\mathbf{m}}_{2k-1}^T - (k+1)\bar{\mathbf{m}}_{2k}\bar{\mathbf{m}}_{2k}^T + \mathbf{m}_{2k}\mathbf{m}_{2k}^T + \epsilon\mathbf{I}_m), \tag{5.15}$$

where $\bar{\mathbf{m}}_{2k} = (1/(k+1)\sum_{i=0}^{k}\mathbf{m}_{2i})$ is the mean model, s_m is a small constant that depends only on model dimension m, $\mathbf{I}_m$ is a m-dimension identity matrix, ϵ is a positive constant. The constant ϵ is selected to keep $\mathbf{C}_{k+1}$ from being singular, and s_m is typically set to $s_d = 2.4^2/m$ (Gelman et al., 1996). Equation 5.15 is a recursion formula requiring little computation. At early iterations, we use $\mathbf{C}_{k+1} = \sigma_{m0}\mathbf{I}_m$ as an alternative to equation 5.15, in which σ_{m0} is a reasonable constant from some prior information or by trial-and-error method; after enough models are obtained, we switch to equation 5.15. All models are constrained in given bounds, for example, we will constrain the time-lapse change in $[-80m/s, 80m/s]$ in our numerical tests (the true time-lapse change is $40m/s$).

5.4.5 Data prior

The character of the time-lapse problem allows us to incorporate data prior information (i.e., the data error covariance $\mathbf{C}_d$ in equation 5.11) in a manner not described elsewhere in the literature. For independent and identically distributed data, the error covariance can be expressed as $\mathbf{C}_d = \sigma_d\mathbf{I}_d$ where all standard deviations σ_d are equal, and $\mathbf{I}_d$ is the identity matrix. Normally, σ_d must be determined by trial-and-error, or by treating it as an unknown to be updated during the inversion (e.g., Malinverno and Briggs, 2004; Bodin et al., 2012). In DDFWI, all noise in the composited data come from the difference data, in which it is often possible to infer regions of signal from regions of noise. We assume that "pure noise"

regions of the data can be identified, and within these regions data error follows zero-mean Gaussian statistics. The success of this method requires that the baseline and monitoring surveys have good repeatability (which permits pure noise regions to be identified), and the availability of sufficient samples within these regions that the estimated distribution is representative. The first of these requirements is also a premise of DDFWI, so the only new assumption being introduced is the second. Finally, we set out the monitoring inversion of MCMC DDFWI in Algorithm 2.

Algorithm 2: Monitoring inversion based on Baysian/MCMC

Input: $\mathbf{m}^{'}$, $\mathbf{d}_{syn}(\mathbf{m}^{'})$, $\mathbf{d}_{obs1}$, $\mathbf{d}_{obs2}$

Output: All samples of $\mathbf{m}_2$

Initialization: $\mathbf{m}_2 = \mathbf{m}^{'}$, $\mathbf{d}^{'}_{obs2} = \mathbf{d}_{syn}(\mathbf{m}^{'}) + (\mathbf{d}_{obs2} - \mathbf{d}_{obs1})$, $\mathbf{C}_m = \sigma_{m0}\mathbf{I}_m$, $\mathbf{C}_d = \sigma_d\mathbf{I}_d$;

for *k=1,2, . . . ,* N **do** `/* N is the maximum sampling number */`

  Generate proposal model $\mathbf{m}_2^*$ by equation 5.14 `/* only sequentially update one grid cell per time */`;

  Calculate the acceptance α by equation 5.13;

  Draw a random number u from a uniform distribution between 0 and 1;

  if $\alpha > u$ **then**

    $\mathbf{m}_2 = \mathbf{m}_2^*$

  else

    $\mathbf{m}_2 = \mathbf{m}_2$

  end

  Store $\mathbf{m}_2$ as a sample;

  Update $\mathbf{C}_m$ by equation 5.15;

end

5.5 Numerical examples

In this section we establish the feasibility this time-lapse approach with a synthetic 2D constant-density acoustic time-lapse model. The models, geometry, and corresponding noise-free multisource shot gather seismic data used are displayed in Figure5.1a-f. The model size is 50-by-50, with a 10m grid spacing. A 40m/s time-lapse velocity change is placed at the center in the time-lapse model (in Figure 5.1c, the monitoring model minus the baseline model is plotted). Ten sources are distributed on the surface, and receivers are placed both on the surface and within two vertical boreholes which define the lateral extent of the volume. The sources are excited simultaneously, each using an identical Ricker wavelet with a 40Hz central frequency. A time-domain finite differencing method, eighth-order in space and second-order in time, is used for simulation. The target area, of size 19-by-22 and containing 418 parameters, is highlighted in Figure 5.1c. Updating is constrained to occur in this region during the monitoring inversion stage, during which we run eight chains simultaneously with the same starting model (i.e., the model inverted during the baseline stage), to obtain samples quickly, and each chain is sampled 83600 times. The burn-in period comprises the first 21900 models; these are not utilized.

5.5.1 DO DDFWI vs MCMC DDFWI

To exemplify the approach, and simultaneously examine the relationship between the recovered model and the starting model, we produce four inversions, each with different starting models (Models 1, 2, 3, and 4). These models, which correspond with true baseline model (Model 1), the inverted baseline model from a standard FWI (Model 2, using noise-free baseline data and a smooth starting Model 3), and two smooth models (Model 3 and 4), are plotted in Figure 5.2a-d. We use noise-free monitoring data, and invert using both conventional DO DDFWI and MCMC DDFWI. The recovered models (all MCMC DDFWI results correspond to the posterior means) are plotted in Figure 5.3. In Figure 5.4a-p, the corre-

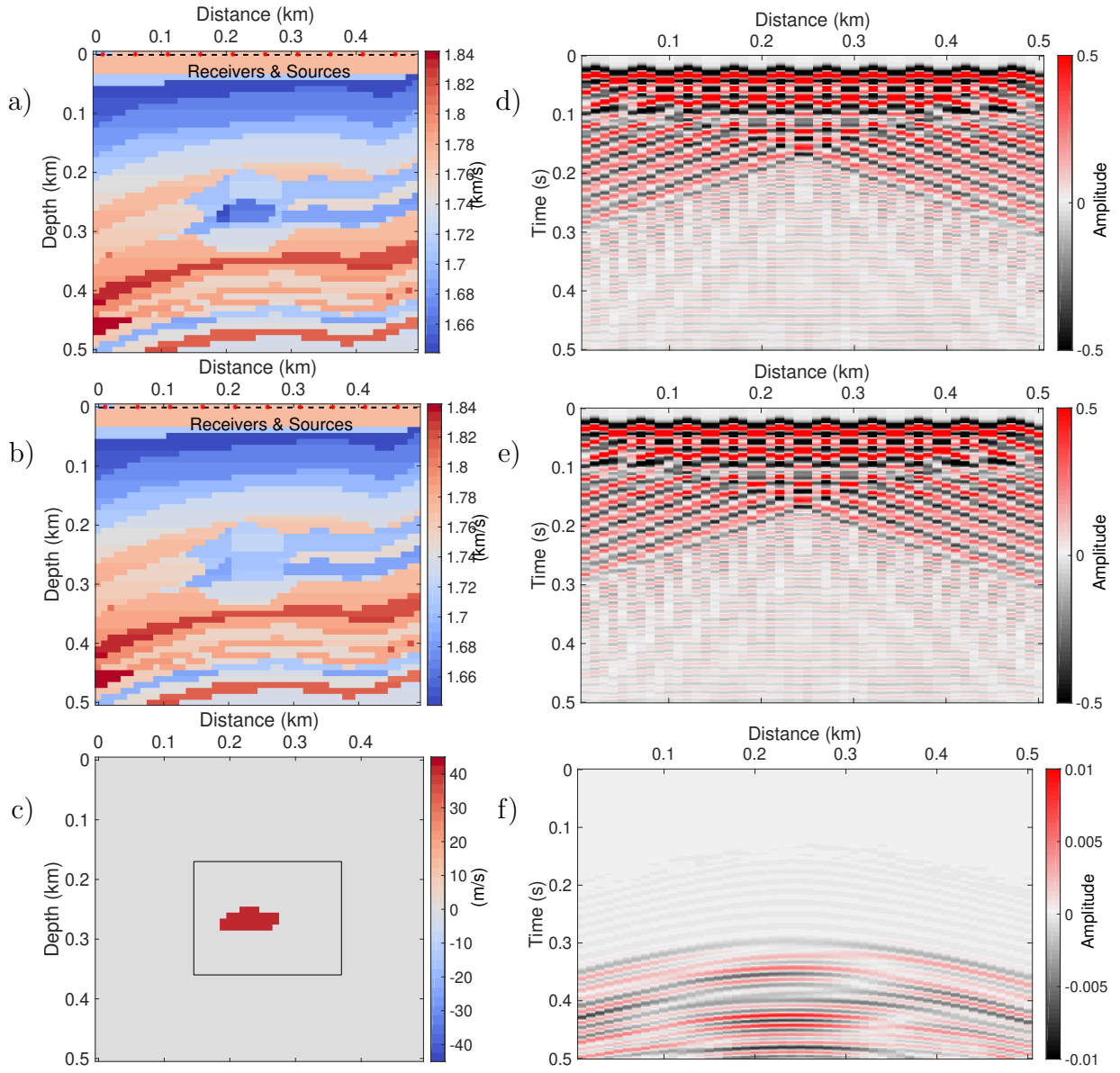

Figure 5.1: In (a), (b), and (c) are the baseline model, monitoring model, and time-lapse model respectively. In (d), (e), and (f) are the multisource baseline data, multisource monitoring data, and difference data, respectively. The acquisition geometries in (a) and (b) are identical. The black dashed lines indicate receivers, and the red stars sources. The black box in (c) is the target area, i.e., the region in which model updating permitted during the monitoring inversion.

sponding absolute model errors of Figure 5.3a-p, are plotted. Compared with conventional DDFWI, the results of MCMC DDFWI exhibit more random variations, which become more significant as the starting model becomes less accurate, but fewer coherent errors. To en-

hance this feature, in Figure 5.3a-p, to reduce the random noise, we use a median filter to the results of MCMC DDFWI and also use the same filter to the results of conventional DDFWI. We observe within the MCMC DDFWI results generally clearer edges of nonzero time-lapse model and less coherent model errors. In Figure 5.5, we plot the curves of the L2 norm of the monitoring data residuals of MCMC DDFWI versus sampling number for different starting models, which shows that all chains converge well, but those associated with the more accurate starting model converge a lower data misfit. The inverted time-lapse models of MCMC DDFWI, using Model 1 as the starting model for monitoring inversion, at different iteration numbers in a certain chain are plotted in Figure 5.6 to show the evolution process of MCMC time-lapse inversion.

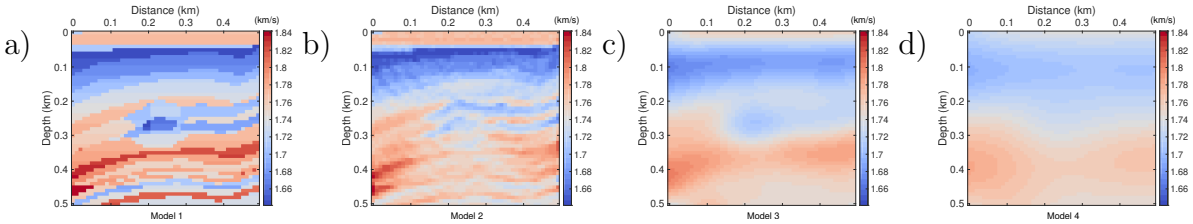

Figure 5.2: Different starting models (Model 1, 2, 3, 4) for monitoring inversion. (a) is the true baseline model. (b) is the inverted baseline model of the DO FWI using (d) as the starting model. (c) and (d) are smooth models from the true baseline model after using a two-dimensional Gaussian filter. The models become worse from left to right.

5.5.2 Noise tests

In Figures 5.7a-c, difference data with different noise levels are plotted. The standard deviation of the noise-free difference data simulations (Figure 5.1f) is 1.2×10^{-3}; we draw noise from zero-mean Gaussian distributions with standard deviations of 2.5×10^{-4}, 2.5×10^{-3}, and 5×10^{-3}, i.e., with signal-to-noise ratios (SNR) of approximately about 23, 0.23, and 0.06, respectively. The SNR calculation does not take the synthetic baseline data into account; they increase significantly after adding the synthetic data to the difference data. Model 2 (Figure 5.2b) is input as the starting model for all monitoring inversions in this subsection.

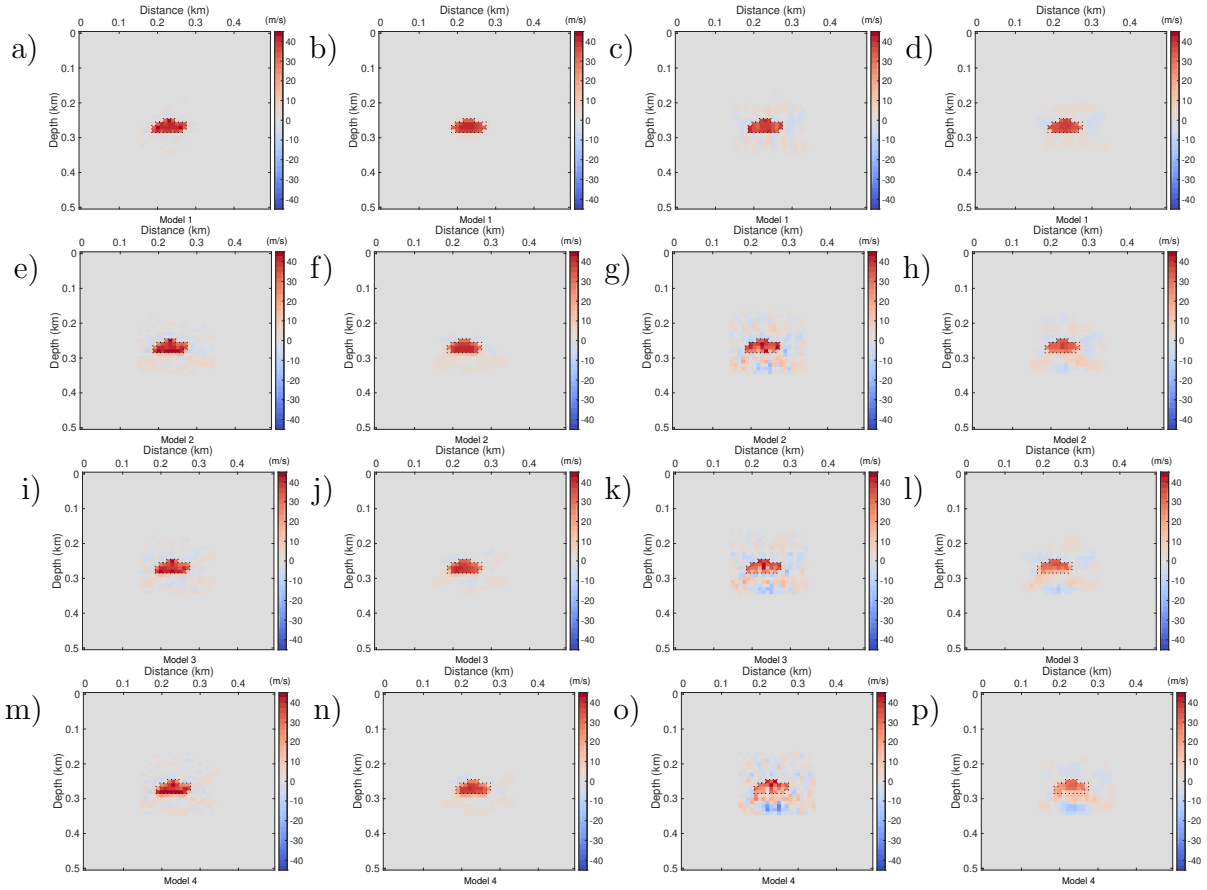

Figure 5.3: Inverted time-lapse models of DO DDFWI and MCMC DDFWI from different starting models in Figure 5.2. (a)-(p) are results for model 1 (a-d), 2 (e-h), 3 (i-l), and 4 (m-p). The results in (a), (e), (i), and (m) are inverted from DO DDFWI, and that in (b), (f), (j), and (n) are obtained by using a median filter to results of (a), (e), (i), and (m), respectively. The results in (c), (g), (k), and (o) are inverted from MCMC DDFWI, and that in (d), (h), (l), and (p) are obtained by using a median filter to results of (c), (g), (k), and (o), respectively. The black dot line box is the edge of the true non-zero model change.

Pure noise regions (above the horizontal lines in Figures 5.7a-c) are used to estimate the data error standard deviation σ_d, and using these values the MCMC DDFWI procedures are carried out for each case. In Figure 5.8a-d, the L2 norm versus sampling number curves are plotted, each containing eight chains. We observe that all chains in each monitoring inversion converge to similar misfit levels, with higher noise levels corresponding to higher misfit levels after convergence, and stronger misfit fluctuations. In Figure 5.9a-i, the inverted time-lapse models including the posterior mean models and the maximum posterior probability (MAP)

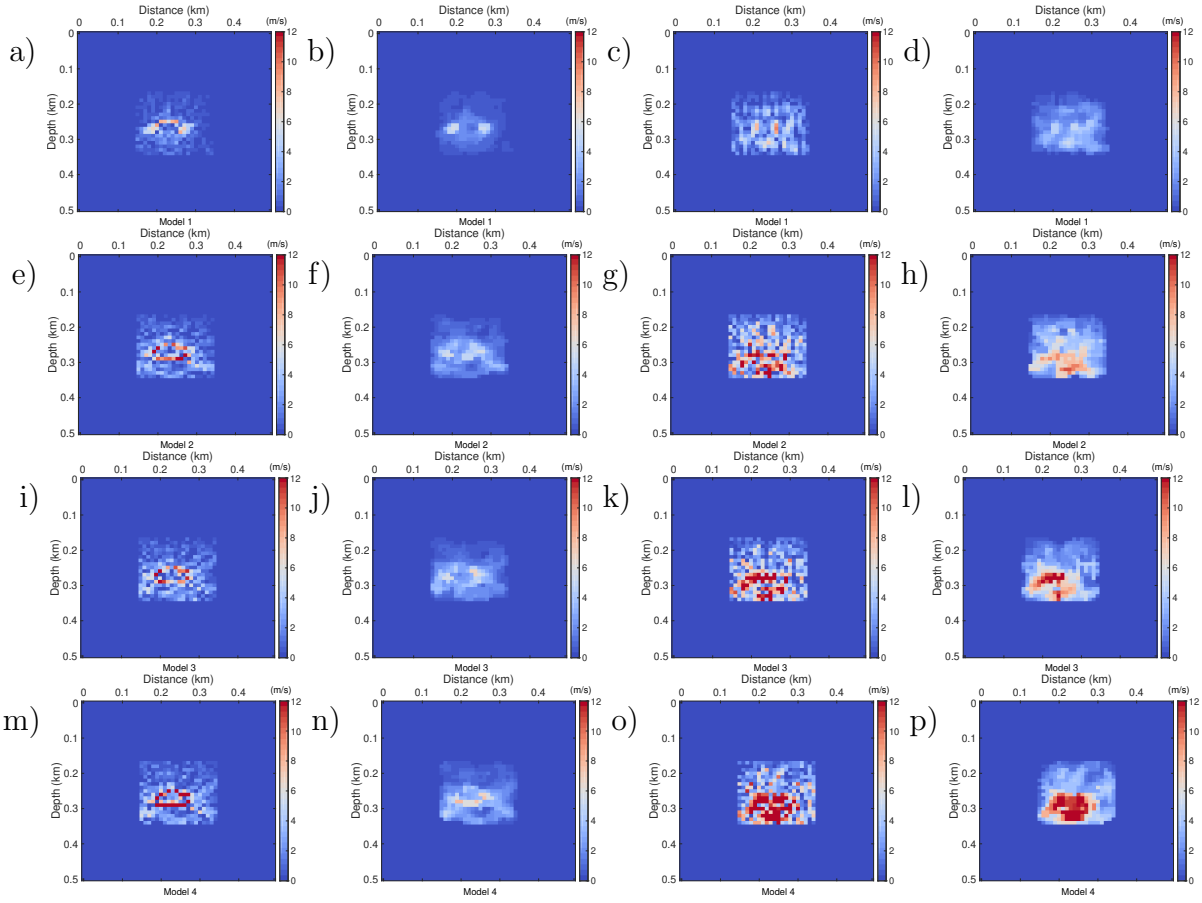

Figure 5.4: (a)-(p) are, respectively, absolute values of model errors of results in Figure 5.3a-p.

models, and the corresponding posterior model standard deviations, are plotted. The posterior mean models are observed to match well with the true time-lapse model. MAP models are worse than the mean models since they may not be at the global minimum, and the mean models can interpret the target better. The standard deviations of the posterior model evidently reflect the noise level of the data. A higher noise level leads to larger posterior model standard deviations. The standard deviations also vary from parameter to parameter and relate to the true value of the time-lapse model. The marginal distributions at the distance of 2.5km are plotted in Figure 5.10a-c, and the corresponding true time-lapse changes and mean time-lapse changes are also plotted on them to have comparisons. From both Figure 5.9 and 5.10, we can see better inversions often correspond to lower standard deviations and

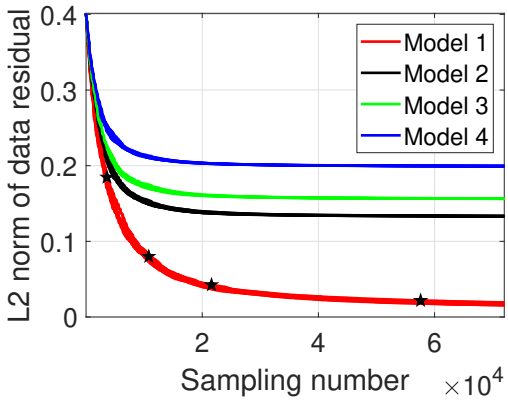

Figure 5.5: Curves of the monitoring data misfit of MCMC DDFWI versus sampling number for different starting models. Each color contains eight curves corresponding to eight Markov chains. The corresponding time-lapse models at iteration numbers marked by black stars are plotted in Figure 5.6.

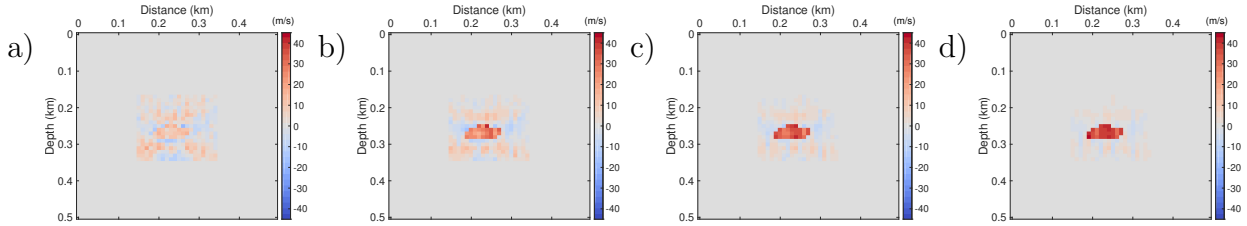

Figure 5.6: The inverted time-lapse models of MCMC DDFWI, using Model 1 as the starting model for monitoring inversion, at the iteration numbers of 3600 (a), 10800 (b), 21600 (c), and 57600 (d) in a certain chain, which are marked by black stars in Figure 5.5.

higher marginal probability. In Figure 5.11a-c, the marginal distributions of six parameters located at the black dots in Figure 5.9a, including three parameters with nonzero true time-lapse change and three parameters with zero true time-lapse change, are plotted. In Figure 5.11a, since the model error standard deviation is too small and the rejection rate is too high, we do not obtain enough effective models to characterize a reasonable distribution for some parameters. But in Figure 5.11b and c, the samples are observed to match the Gaussian distributions closely. In Figure 5.12a-f and 5.13a-f, the predicted noise-free difference data and the noises, are plotted. We observe the predicted values are close to the true values.

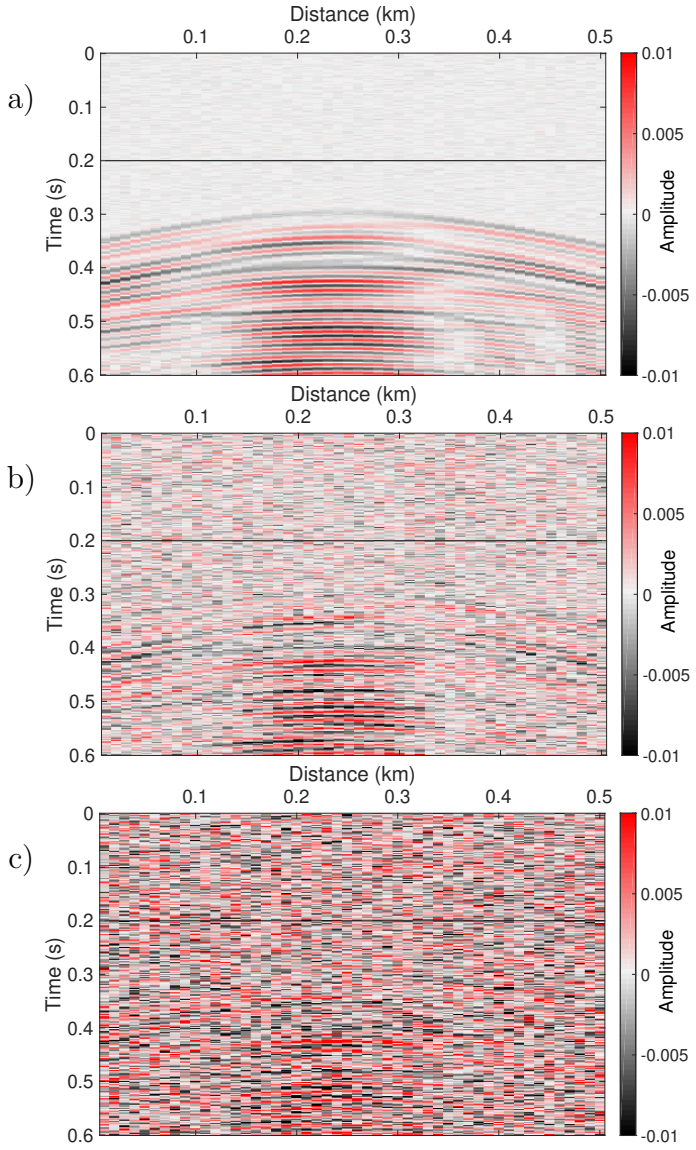

Figure 5.7: The difference data with different noise levels. Noises confirm to zero-mean Gaussian distributions with different standard deviations that are, respectively, 2.5×10^{-4} for (a), 2.5×10^{-3} for (b), and 5×10^{-3} for (c). The data above the black line in each panel is the pure noise region used to estimate the data errors standard deviation.

5.6 Discussion

Global stochastic optimization methods involving wave equation simulations of seismic data avoid several well-known issues experienced by local descent based inversions, and naturally provide uncertainty estimates, but are generally prohibitively expensive. Realistic applica-

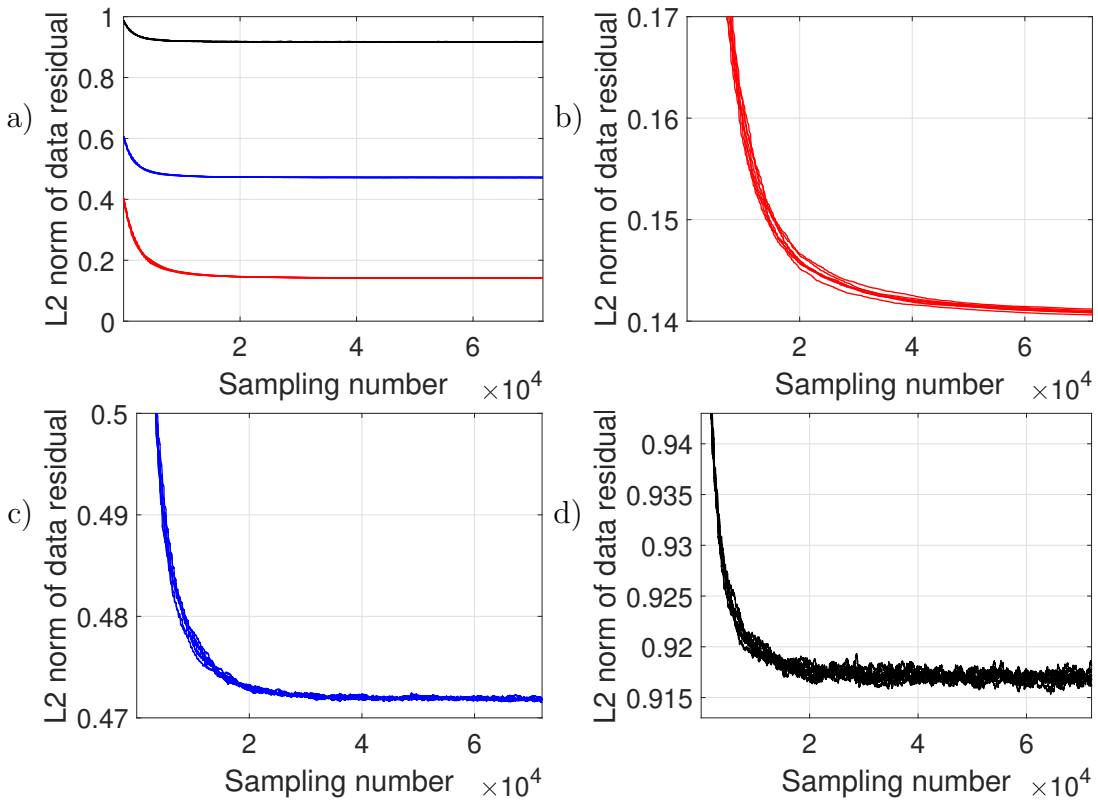

Figure 5.8: Curves of the monitoring data misfit of MCMC DDFWI versus sample number for different level noisy data. The red, blue, and black curves, respectively, correspond to the data in Figure 5.7a, b, and c. All curves are displayed together in (a) and separately displayed in (b)-(d). Each color contains eight curves corresponding to eight Markov chains.

tions of these approaches require effective strategies for fast simulation and dimensionality reduction. In the second category, we have implemented a range of these which are applicable to the time-lapse problem, including target oriented methods, reducing the number of unknown grid cells to update, multi-source methods, reducing the number of simulations required to predict shot records. This, when coupled with small scale 2D models, produces problem which can be run in several hours on a standard desktop computer, and allow us to draw some important feasibility conclusions. As model sizes grow, examples of the kind we present will naturally require more significant high-performance computing solutions. At this stage, invoking additional stochastic algorithms, e.g., Hamiltonian Monte Carlo, will be warranted. In the first category, extending our method to include local solvers (Robertsson

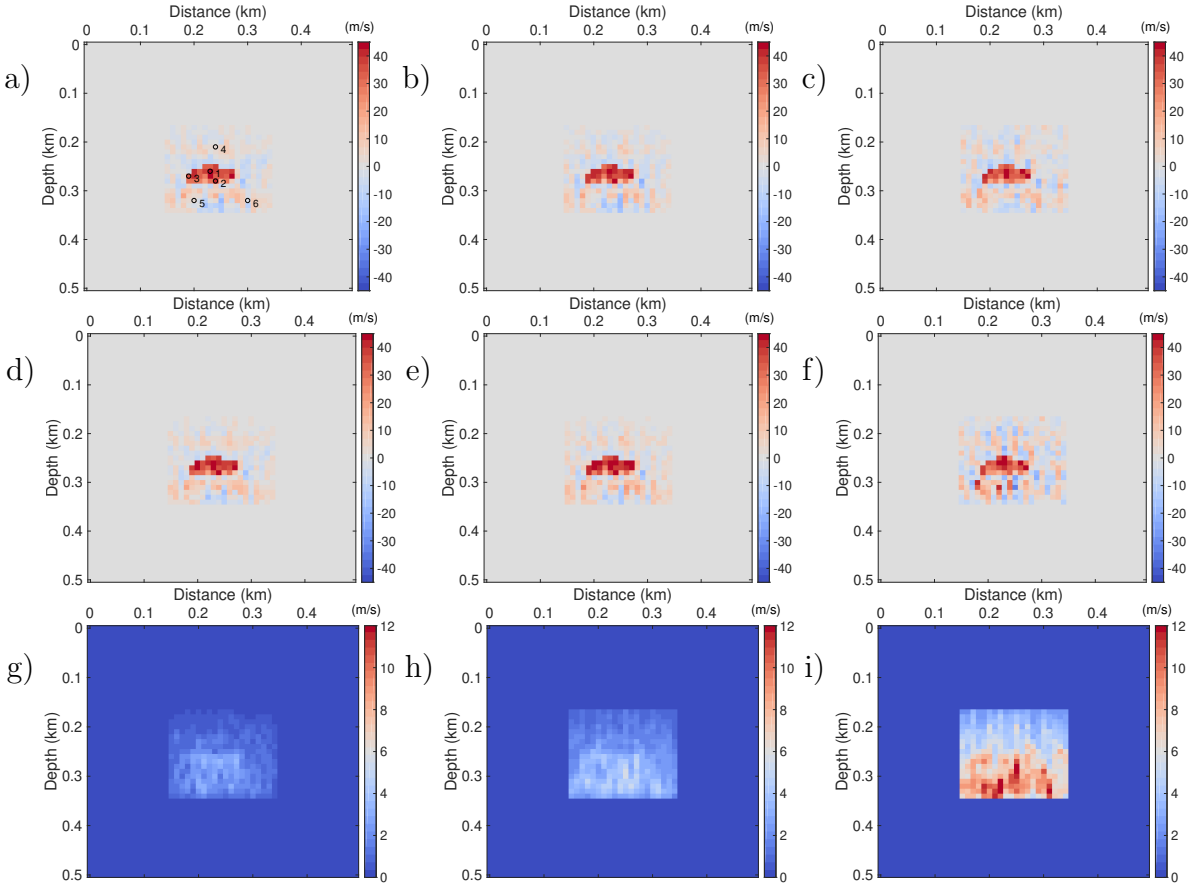

Figure 5.9: Inverted mean time-lapse models and posterior model standard deviations of MCMC DDFWI using Model 2 as the starting model and data with different noise levels. (a), (b), and (c) are mean models using data in Figure 5.7a, b and c, respectively. (d), (e), and (f) are MAP models using data in Figure 5.7a, b and c, respectively. (g), (h), and (i) are posterior model standard deviations using data in Figure 5.7a, b and c, respectively. The six labeled black circles in (a) are the positions where we will appraise the marginal distributions in Figure 5.11.

and Chapman, 2000; Yang et al., 2012; Huang et al., 2018) is a natural next step for solving target-oriented problems of this kind. In our current toolset, the Green's functions necessary for these are not yet in place for the multi-source solver; this inclusion is an important part of ongoing research.

For challenges in real time-lapse data, first, the DDFWI strategy used in this paper requires well-repeated time-lapse surveys, this may make our method unsuitable for some cases, and proper processing to raw shot gathers is necessary to improve the repeatability of

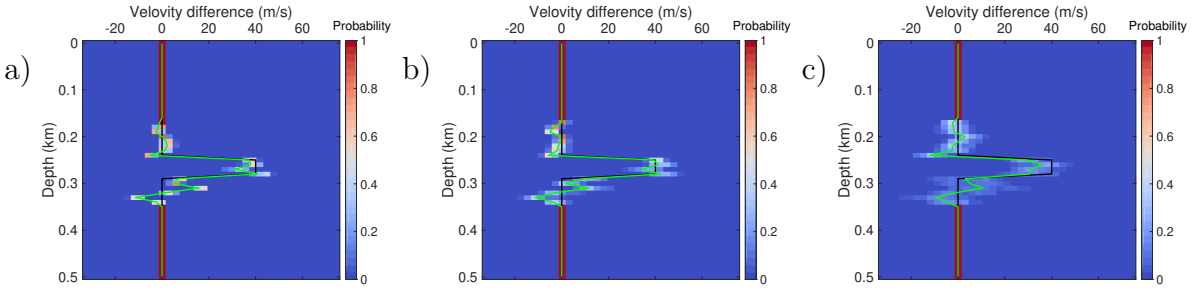

Figure 5.10: The marginal distributions at the distance of 2.5 km for results of MCMC DDFWI using Model 2 as the starting model and data with different noise levels. (a), (b), and (c) are marginal distributions using data in Figure 5.7a, b and c, respectively. Black lines are true time-lapse changes, and green lines are mean time-lapse changes.

time-lapse data. The second is our research is built on an acoustic setting which is different from the elastic property of the earth, hence, seismic signals beyond the acoustic setting will interfere with the inversion to some degree. The third is the realistic noise level may be higher than that in our tests, which could further destroy the small time-lapse signal.

5.7 Conclusions

In this study, have assembled and numerically tested a time-lapse FWI methodology, based on an MH MCMC algorithm, and a method of estimating the data error standard deviation for time-lapse data which makes use of the features of difference data. We have employed DDFWI, multisource data, and local-updating target-oriented inversion in the approach, calculating model covariance with an AM algorithm. Synthetic data tests using a 2D acoustic model establish the feasibility of both model inversion and uncertainty quantification of our method for small synthetic examples, and allow several distinguishing characteristics of the methodology in the presence of noise and relative to standard deterministic FWI. In particular, although both methods tend to introduce errors as the starting model becomes less and less resolved, in the MCMC approach those errors are less spatially correlated, and the anomaly edges are more clearly recovered.

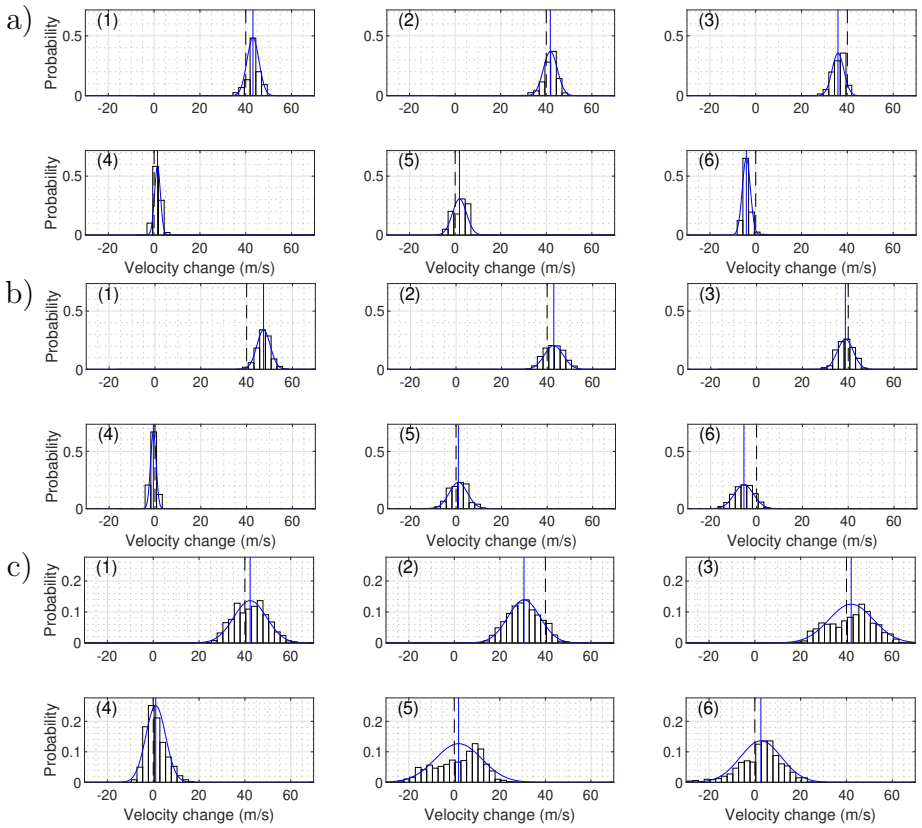

Figure 5.11: The marginal distributions of 6 parameters located at the labeled black circles in Figure 5.9a, including three with nonzero true time-lapse change and three with zero true time-lapse change. (a), (b), and (c), respectively, correspond to the noisy data in Figure 5.7a, b, and c. The histograms denote all samples abstracted from the eight chains excepting the burn-in ones. Blue solid curves are standard Gaussian probability density functions (PDFs) best fitting the histograms. Black dash straight lines and blue solid straight lines denote the true values and inverted mean values, respectively. The labels on the up left corner of each panel correspond to the parameter locations labeled in Figure 5.9a

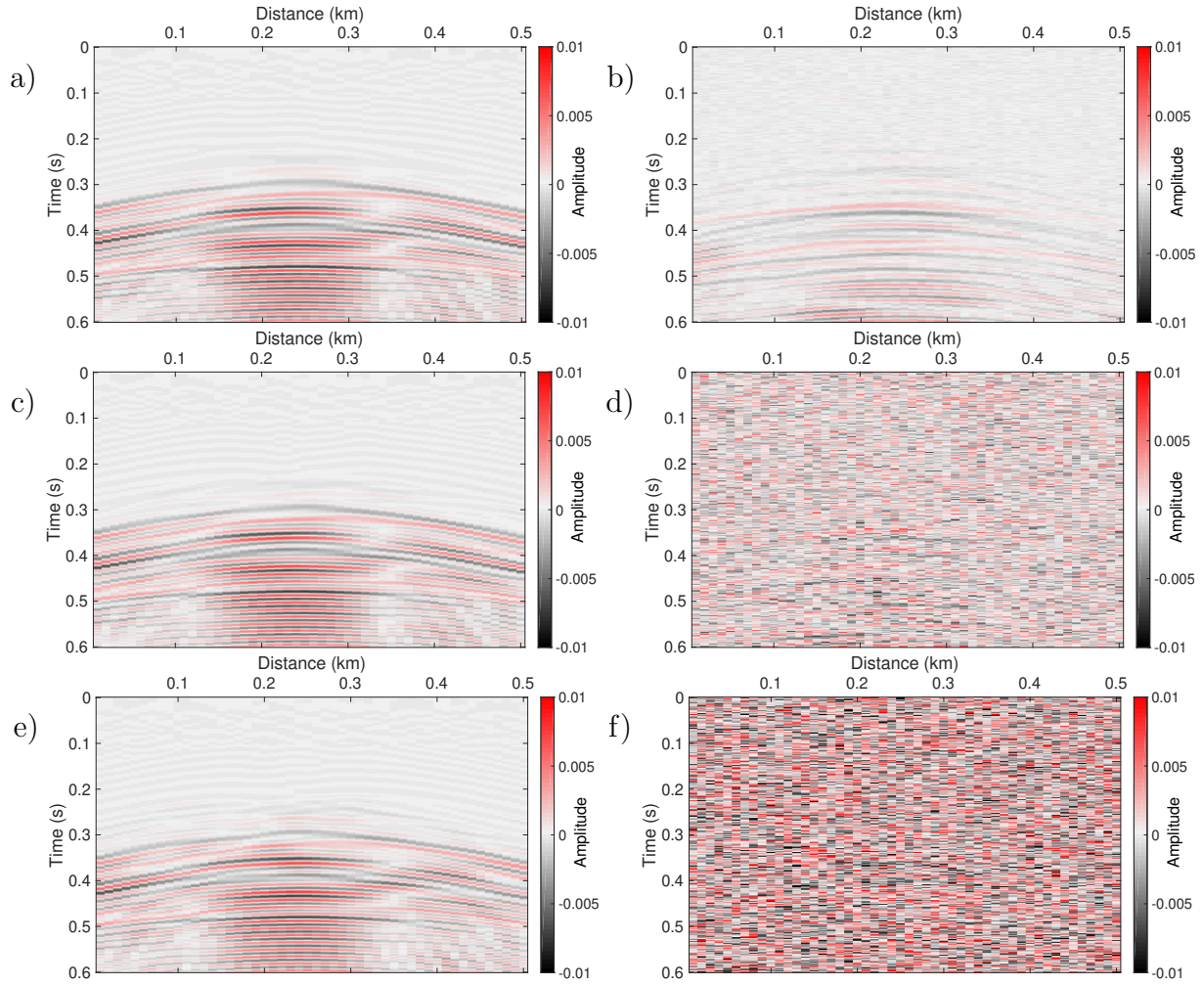

Figure 5.12: The predicted noise-free difference data (a, c, e) and noises (b, d, f) using the inverted mean models in Figure 5.9a, b, and c, respectively. The predicted data in (a) and the predicted noise in (b) correspond to the noisy data in Figure 5.7a. (c) and (d) correspond to the noisy data in Figure 5.7b. And (e) and (f) correspond to the noisy data in Figure 5.7c.

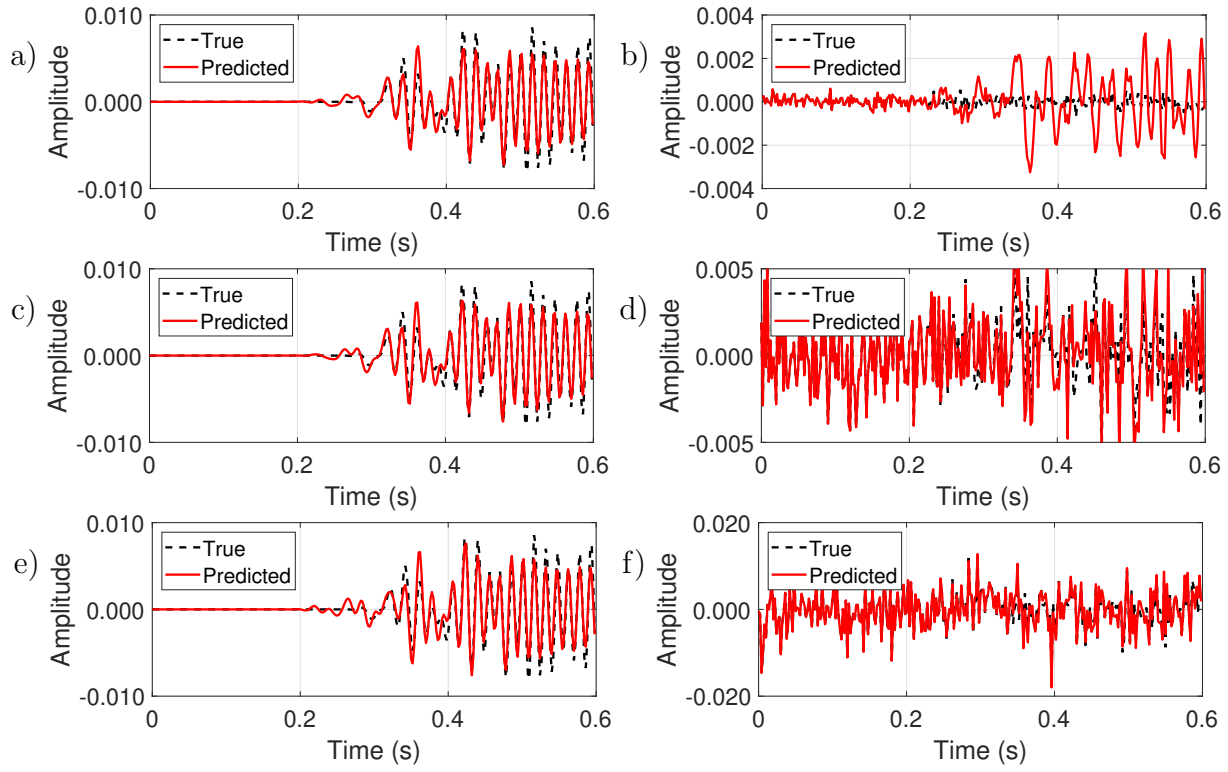

Figure 5.13: Red curves in (a)-(f) are central traces of Figure 5.12a-f,respectively. Black dash curves are the corresponding true values. (a), (c), and (e) display the noise-free data, and (b), (d), and (f) display the corresponding noises.

Chapter 6

Conclusions and future work

6.1 Conclusions

Time-lapse seismic, also known as 4D seismic, is a powerful tool for monitoring subsurface changes over time. By comparing seismic data acquired at different intervals, it enables the detection and characterization of dynamic reservoir processes, aiding in reservoir management, production optimization, and enhanced oil recovery. It has applications in geothermal energy, CO2 storage monitoring, and environmental impact assessment. Despite challenges in data registration, noise mitigation, and inversion uncertainties, time-lapse seismic continues to evolve as an indispensable asset, providing valuable insights into subsurface dynamics and contributing to advancements in geophysics.

In this thesis, novel techniques are developed for time-lapse imaging and inversion. Chapter 2 presents a workflow incorporating frequency-domain matching filters and time-shift correction to reduce source wavelet non-repeatability in imaging. In Chapter 4, a stepsize-sharing strategy is introduced to mitigate artifacts caused by convergence variability in conventional methods. Finally, Chapter 5 formulates a Bayesian time-lapse FWI procedure using a Markov chain Monte Carlo algorithm, incorporating multiple strategies within a stochastic framework. These advancements enhance the accuracy and reliability of time-

lapse imaging and inversion, addressing challenges such as non-repeatability, convergence variability, and uncertainty estimation.

However, there are still challenges in time-lapse seismic analysis that need to be addressed. The time-lapse imaging workflow presented in Chapter 2 focuses solely on the non-repeatability of source waveforms and does not account for non-repeatability in source/receiver positions. While the time-lapse FWI discussed in Chapter 4 demonstrates the ability to handle non-repeatability in source/receiver positions, it does not specifically address non-repeatability in source wavelets. Moreover, the impact of near-surface or seawater property changes during time-lapse surveys on the inversions is not investigated. More effective approaches for overcoming this nonrepeatability should be proposed. Furthermore, the uncertainty quantification approach proposed in Chapter 5 is hindered by the computational cost associated with stochastic optimization methods in waveform inversion. To overcome this limitation, it is necessary to introduce or develop a more efficient optimization approach capable of handling the high-dimensional nature of waveform inversion problems.

6.2 Future work

In order to further advance time-lapse inversion techniques, several avenues for future research can be explored. Firstly, addressing the challenge of non-repeatability in both source wavelets and source/receiver positions as well as the near-surface or seawater nonrepeatable remains a crucial aspect. Developing robust methodologies that can effectively handle these sources of non-repeatability and their combined effects would significantly enhance the accuracy and reliability of time-lapse inversion.

Additionally, the computational efficiency of stochastic optimization approaches for uncertainty quantification needs to be improved. Exploring novel algorithms or optimizing existing ones to reduce the computational burden associated with high-dimensional waveform inversion problems would enable the practical implementation of stochastic optimization

methods in real-world applications.

Furthermore, integrating advanced machine learning and data assimilation techniques into time-lapse inversion workflows holds great promise. Leveraging the power of artificial intelligence and assimilating various types of data, such as production data or well logs, can provide valuable constraints and improve the accuracy of inversion results.

Moreover, considering the challenges posed by complex geological scenarios, such as heterogeneous reservoirs or multiphase flow as well as the elastic, anelastic, or anisotropic media, developing specialized inversion methodologies tailored to these specific situations would be beneficial. These approaches should take into account the inherent complexities and incorporate appropriate physics-based models to ensure accurate and reliable inversion results.

Overall, future work in time-lapse inversion should focus on addressing the challenges of non-repeatability, enhancing computational efficiency, incorporating advanced techniques like machine learning and data assimilation, accommodating complex geological scenarios, and exploring new application domains. By pursuing these avenues, we can continue to advance the state-of-the-art in time-lapse inversion and unlock its full potential for subsurface characterization and monitoring.

Bibliography

Afanasiev, M. V., R. G. Pratt, R. Kamei, and G. McDowell, 2014, Waveform-based simulated annealing of crosshole transmission data: A semi-global method for estimating seismic anisotropy: Geophysical Journal International, **199**, 1586–1607.

Ajo-Franklin, J., J. Peterson, J. Doetsch, and T. Daley, 2013, High-resolution characterization of a co2 plume using crosswell seismic tomography: Cranfield, ms, usa: International Journal of Greenhouse Gas Control, **18**, 497–509.

Al-Ismaili, I. Y. and M. R. Warner, 2002, Non-linear cross-equalization of time-lapse seismic surveys using artificial neural networks: 64th EAGE Conference & Exhibition, cp–5.

Aleardi, M. and A. Mazzotti, 2016, 1d elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm–gibbs sampler approach: Geophysical Prospecting, **65**, 64–85.

Almutlaq, M. H. and G. F. Margrave, 2013, Surface-consistent matching filters for time-lapse seismic processing: Geophysics, **78**, M29–M41.

Arts, R., O. Eiken, A. Chadwick, P. Zweigel, L. Van der Meer, and B. Zinszner, 2003, Monitoring of co2 injected at sleipner using time lapse seismic data: Greenhouse Gas Control Technologies-6th International Conference, 347–352.

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2011, Sensitivity analysis of time-lapse images obtained by differential waveform inversion with respect to reference model, *in* SEG Technical Program Expanded Abstracts 2011, 2482–2486, Society of Exploration Geophysicists.

———, 2014, Time-lapse seismic imaging using regularized full-waveform inversion with a prior model: which strategy?: Geophysical prospecting, **63**, 78–98.

———, 2015, Time-lapse seismic imaging using regularized full-waveform inversion with a prior model: which strategy?: Geophysical prospecting, **63**, 78–98.

Ayeni, G. and M. Nasser, 2009, Optimized local matching of time-lapse seismic data: A case study from the gulf of mexico: Presented at the 2009 SEG Annual Meeting.

Barkved, O., K. Buer, K. Halleland, R. Kjelstadli, T. Kleppan, and T. Kristiansen, 2003, 4d seismic response of primary production and waste injection at the valhall field: 65th EAGE Conference & Exhibition, cp–6.

Barkved, O. I., T. Kristiansen, and E. Fjær, 2005, The 4d seismic response of a compacting reservoir—examples from the valhall field, norway, *in* SEG Technical Program Expanded Abstracts 2005, 2508–2511, Society of Exploration Geophysicists.

Bergmann, P., A. Kashubin, M. Ivandic, S. Lüth, and C. Juhlin, 2014, Time-lapse difference static correction using prestack crosscorrelations: 4d seismic image enhancement case from ketzin: Geophysics, **79**, B243–B252.

Biswas, R. and M. Sen, 2017, 2d full-waveform inversion and uncertainty estimation using the reversible jump hamiltonian monte carlo, *in* SEG Technical Program Expanded Abstracts 2017, 1280–1285, Society of Exploration Geophysicists.

Bodin, T., M. Sambridge, N. Rawlinson, and P. Arroucau, 2012, Transdimensional tomography with unknown data noise: Geophysical Journal International, **189**, 1536–1556.

Bortoni, S., S. Barragan, G. Azevedo, L. Cypriano, A. Ferreira, W. Moreira, P. dos Reis, and W. Filho, 2021, Learnings from an fwi imaging study using 3d and 4d data over a postsalt field in campos basin: First International Meeting for Applied Geoscience & Energy, 587–591.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.

Byerley, G., D. Monk, P. Aaron, and M. Yates, 2018, Time-lapse seismic monitoring of

individual hydraulic frac stages using a downhole das array: The Leading Edge, **37**, 802–810.

Calvert, R., 2005, Insights and methods for 4d reservoir monitoring and characterization: Society of Exploration Geophysicists and European Association of ....

Chadwick, R., D. Noy, R. Arts, and O. Eiken, 2009, Latest time-lapse seismic data from sleipner yield new insights into co2 plume development: Energy Procedia, **1**, 2103–2110.

Chattopadhyay, S. and G. A. McMechan, 2008, Imaging conditions for prestack reverse-time migration: Geophysics, **73**, S81–S89.

Cheng, A., L. Huang, and J. Rutledge, 2010, Time-lapse vsp data processing for monitoring co 2 injection: The Leading Edge, **29**, 196–199.

Chib, S. and E. Greenberg, 1995, Understanding the metropolis-hastings algorithm: The american statistician, **49**, 327–335.

Cho, Y. and H. Jun, 2021, Estimation and uncertainty analysis of the co2 storage volume in the sleipner field via 4d reversible-jump markov-chain monte carlo: Journal of Petroleum Science and Engineering, **200**, 108333.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

Creutz, M., 1988, Global monte carlo algorithms for many-fermion systems: Physical Review D, **38**, 1228.

da Silva, N. V., G. Yao, and M. Warner, 2019, Semiglobal viscoacoustic full-waveform inversionsemiglobal viscoacoustic inversion: Geophysics, **84**, R271–R293.

Datta, D. and M. K. Sen, 2016, Estimating a starting model for full-waveform inversion using a global optimization method: Geophysics, **81**, R211–R223.

Denli, H. and L. Huang, 2009, Double-difference elastic waveform tomography in the time domain, *in* SEG Technical Program Expanded Abstracts 2009, 2302–2306, Society of Exploration Geophysicists.

Dramsch, J. S., A. N. Christensen, C. MacBeth, and M. Lüthje, 2019, Deep unsupervised

4d seismic 3d time-shift estimation with convolutional neural networks.

Egorov, A., R. Pevzner, A. Bóna, S. Glubokovskikh, V. Puzyrev, K. Tertyshnikov, and B. Gurevich, 2017, Time-lapse full waveform inversion of vertical seismic profile data: Workflow and application to the co2crc otway project: Geophysical Research Letters, **44**, 7211–7218.

Ely, G., A. Malcolm, and O. V. Poliannikov, 2018, Assessing uncertainties in velocity models and images with a fast nonlinear uncertainty quantification method: Geophysics, **83**, R63–R75.

Fomel, S., 2007, Local seismic attributes: Geophysics, **72**, A29–A33.

Fomel, S. and L. Jin, 2009, Time-lapse image registration using the local similarity attribute: Geophysics, **74**, A7–A11.

Fu, X. and K. A. Innanen, 2021, An mcmc-based approach to time-lapse full-waveform inversion: First International Meeting for Applied Geoscience & Energy, 3484–3489.

———, 2022a, Stepsize sharing in time-lapse full-waveform inversion: Geophysics, **88**, 1–67.

———, 2022b, A time-domain, multi-source bayesian/markov chain monte carlo formulation of time-lapse seismic waveform inversion: Geophysics, **87**, 1–110.

———, 2023, Time-lapse seismic imaging using shot gathers with nonrepeatable source wavelets: Geophysics, **88**, M17–M30.

Fu, X., S. Romahn, and K. Innanen, 2020, Double-wavelet double-difference time-lapse waveform inversion, *in* SEG Technical Program Expanded Abstracts 2020, 3764–3767, Society of Exploration Geophysicists.

Fu, X., S. Romahn, and K. A. Innanen, 2023, Double-wavelet time-lapse full-waveform inversion: Canadian Journal of Exploration Geophysics, **46**, 1–19.

Gallop, J., 2011, Midpoint match filtering, *in* SEG Technical Program Expanded Abstracts 2011, 4170–4174, Society of Exploration Geophysicists.

Gebraad, L., C. Boehm, and A. Fichtner, 2020, Bayesian elastic full-waveform inversion using hamiltonian monte carlo: Journal of Geophysical Research: Solid Earth, **125**,

e2019JB018428.

Gelman, A., G. Roberts, and W. Gilks, 1996, Efficient metropolis jumping hules: Bayesian statistics.

Greaves, R. J. and T. J. Fulp, 1987, Three-dimensional seismic monitoring of an enhanced oil recovery process: Geophysics, **52**, 1175–1187.

Haario, H., E. Saksman, and J. Tamminen, 1999, Adaptive proposal distribution for random walk metropolis algorithm: Computational Statistics, **14**, 375–396.

Haario, H., E. Saksman, J. Tamminen, et al., 2001, An adaptive metropolis algorithm: Bernoulli, **7**, 223–242.

Hale, D., 2006, Fast local cross-correlations of images, *in* SEG Technical Program Expanded Abstracts 2006, 3160–3164, Society of Exploration Geophysicists.

———, 2009, A method for estimating apparent displacement vectors from time-lapse seismic images: Geophysics, **74**, V99–V107.

———, 2013, Dynamic warping of seismic images: Geophysics, **78**, S105–S115.

Hall, S. A., C. MacBeth, O. I. Barkved, and P. Wild, 2005, Cross-matching with interpreted warping of 3d streamer and 3d ocean-bottom-cable data at valhall for time-lapse assessment: Geophysical Prospecting, **53**, 283–297.

Hastings, W. K., 1970, Monte carlo sampling methods using markov chains and their applications.

Hatchell, P. and S. Bourne, 2005, Measuring reservoir compaction using time-lapse timeshifts, *in* SEG Technical Program Expanded Abstracts 2005, 2500–2503, Society of Exploration Geophysicists.

Hatchell, P. and M. Tatanova, 2019, Matching of 4d seismic data: Spectrum balancing vs conventional least squares filters, *in* SEG Technical Program Expanded Abstracts 2019, 5245–5249, Society of Exploration Geophysicists.

Hicks, E., H. Hoeber, M. Houbiers, S. P. Lescoffit, A. Ratcliffe, and V. Vinje, 2016, Time-lapse full-waveform inversion as a reservoir-monitoring tool—a north sea case study: The

Leading Edge, **35**, 850–858.

Hong, T. and M. K. Sen, 2009, A new mcmc algorithm for seismic waveform inversion and corresponding uncertainty analysis: Geophysical Journal International, **177**, 14–32.

Huang, X., M. Jakobsen, K. S. Eikrem, and G. Nævdal, 2018, A target-oriented scheme for efficient inversion of time-lapse seismic waveform data: Presented at the 2018 SEG International Exposition and Annual Meeting.

Innanen, K. A., M. Naghizadeh, and S. T. Kaplan, 2014, Perturbation methods for two special cases of the time-lapse seismic inverse problem: Geophysical Prospecting, **62**, 453–474.

Jack, I., 2017, 4d seismic—past, present, and future: The Leading Edge, **36**, 386–392.

Kaelin, B. and A. Guitton, 2006, Imaging condition for reverse time migration, *in* SEG Technical Program Expanded Abstracts 2006, 2594–2598, Society of Exploration Geophysicists.

Kamei, R., U. G. Jang, D. Lumley, M. Takanashi, M. Nakatsukasa, T. Mouri, and A. Kato, 2017, Time-lapse full-waveform inversion for cross-well monitoring of microbubble injection: Presented at the 2017 SEG International Exposition and Annual Meeting.

Karimi, P., S. Fomel, and R. Zhang, 2016, Time-lapse image registration using the stratigraphic-coordinate system, *in* SEG Technical Program Expanded Abstracts 2016, 5500–5505, Society of Exploration Geophysicists.

Kazemeini, S. H., C. Juhlin, and S. Fomel, 2010, Monitoring co2 response on surface seismic data; a rock physics and seismic modeling feasibility study at the co2 sequestration site, ketzin, germany: Journal of Applied Geophysics, **71**, 109–124.

Kotsi, M., A. Malcolm, and G. Ely, 2020, Uncertainty quantification in time-lapse seismic imaging: a full-waveform approach: Geophysical Journal International, **222**, 1245–1263.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, **74**, WCC177–WCC188.

Lailly, P., J. Bednar, et al., 1983, The seismic inverse problem as a sequence of before

stack migrations: Conference on Inverse Scattering, Theory and Application, Society for Industrial and Applied Mathematics, Expanded Abstracts, 206–220.

Landrø, M., 2001, Discrimination between pressure and fluid saturation changes from time-lapse seismic data: Geophysics, **66**, 836–844.

Lee, D., I. M. Mason, and G. M. Jackson, 1991, Split-step fourier shot-record migration with deconvolution imaging: Geophysics, **56**, 1786–1793.

Li, Y., T. Alkhalifah, and Q. Guo, 2021, Target-oriented time-lapse waveform inversion using deep learning-assisted regularization: Geophysics, **86**, R485–R495.

Liu, X., X. Chen, M. Bai, and Y. Chen, 2021, Time-lapse image registration by high-resolution time-shift scan: Geophysics, **86**, M49–M58.

Lumley, D., D. C. Adams, M. Meadows, S. Cole, and R. Wright, 2003, 4d seismic data processing issues and examples, *in* SEG Technical Program Expanded Abstracts 2003, 1394–1397, Society of Exploration Geophysicists.

Lumley, D. E., 2001, Time-lapse seismic reservoir monitoring: Geophysics, **66**, 50–53.

Macquet, M., D. C. Lawton, A. Saeedfar, and K. G. Osadetz, 2019, A feasibility study for detection thresholds of co2 at shallow depths at the cami field research station, newell county, alberta, canada: Petroleum Geoscience, **25**, 509–518.

Maharramov, M. and B. Biondi, 2014, Joint full-waveform inversion of time-lapse seismic data sets, *in* SEG Technical Program Expanded Abstracts 2014, 954–959, Society of Exploration Geophysicists.

Maharramov, M., B. L. Biondi, and M. A. Meadows, 2016, Time-lapse inverse theory with applicationstime-lapse inverse theory: Geophysics, **81**, R485–R501.

Malinverno, A. and V. A. Briggs, 2004, Expanded uncertainty quantification in inverse problems: Hierarchical bayes and empirical bayes: Geophysics, **69**, 1005–1016.

Mazzotti, A., N. Bienati, E. Stucchi, A. Tognarelli, M. Aleardi, and A. Sajeva, 2016, Two-grid genetic algorithm full-waveform inversion: The Leading Edge, **35**, 1068–1075.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953, Equa-

tion of state calculations by fast computing machines: The journal of chemical physics, **21**, 1087–1092.

Mosegaard, K. and A. Tarantola, 1995, Monte carlo sampling of solutions to inverse problems: Journal of Geophysical Research: Solid Earth, **100**, 12431–12447.

Oldenborger, G. A., P. S. Routh, and M. D. Knoll, 2007, Model reliability for 3d electrical resistivity tomography: Application of the volume of investigation index to a time-lapse monitoring experiment: Geophysics, **72**, F167–F175.

Onishi, K., T. Ueyama, T. Matsuoka, D. Nobuoka, H. Saito, H. Azuma, and Z. Xue, 2009, Application of crosswell seismic tomography using difference analysis with data normalization to monitor co2 flooding in an aquifer: International Journal of Greenhouse Gas Control, **3**, 311–321.

Pevzner, R., M. Urosevic, D. Popik, V. Shulakova, K. Tertyshnikov, E. Caspari, J. Correa, T. Dance, A. Kepic, S. Glubokovskikh, et al., 2017, 4d surface seismic tracks small supercritical co2 injection into the subsurface: Co2crc otway project: International Journal of Greenhouse Gas Control, **63**, 150–157.

Phillips, M. and S. Fomel, 2016, Seismic time-lapse image registration using amplitude-adjusted plane-wave destruction, *in* SEG Technical Program Expanded Abstracts 2016, 5473–5478, Society of Exploration Geophysicists.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

Plessix, R.-E., S. Michelet, H. Rynja, H. Kuehl, C. Perkins, J. de Maag, and P. Hatchell, 2010, Some 3d applications of full waveform inversion: 72nd EAGE Conference and Exhibition-Workshops and Fieldtrips, cp–162.

Raknes, E. B. and B. Arntsen, 2014, Time-lapse full-waveform inversion of limited-offset seismic data using a local migration regularization: Geophysics, **79**, WA117–WA128.

Raknes, E. B., W. Weibull, and B. Arntsen, 2013, Time-lapse full waveform inversion: Syn-

thetic and real data examples: Presented at the 2013 SEG Annual Meeting.

Ray, A., A. Sekar, G. M. Hoversten, and U. Albertin, 2016, Frequency domain full waveform elastic inversion of marine seismic data from the alba field using a bayesian trans-dimensional algorithm: Geophysical Journal International, **205**, 915–937.

Rickett, J. and D. Lumley, 2001, Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the gulf of mexico: Geophysics, **66**, 1015–1025.

Robertsson, J. O. and C. H. Chapman, 2000, An efficient method for calculating finite-difference seismograms after model alterations: Geophysics, **65**, 907–918.

Ross, C., G. Cunningham, and D. Weber, 1996, Inside the crossequalization black box: The Leading Edge, **15**, 1233–1240.

Ross, C. P. and M. S. Altan, 1997, Time-lapse seismic monitoring: Some shortcomings in nonuniform processing: The Leading Edge, **16**, 931–937.

Routh, P., G. Palacharla, I. Chikichev, and S. Lazaratos, 2012, Full wavefield inversion of time-lapse data for improved imaging and reservoir characterization, *in* SEG Technical Program Expanded Abstracts 2012, 1–6, Society of Exploration Geophysicists.

Routh, P. S. and P. D. Anno, 2008, Time-lapse noise characterization by inversion, *in* SEG Technical Program Expanded Abstracts 2008, 3143–3147, Society of Exploration Geophysicists.

Sajeva, A., M. Aleardi, E. Stucchi, N. Bienati, and A. Mazzotti, 2016, Estimation of acoustic macro models using a genetic full-waveform inversion: Applications to the marmousi modelgenetic fwi for acoustic macro models: Geophysics, **81**, R173–R184.

Shin, C., S. Jang, and D.-J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical prospecting, **49**, 592–606.

Shulakova, V., R. Pevzner, J. C. Dupuis, M. Urosevic, K. Tertyshnikov, D. E. Lumley, and B. Gurevich, 2015, Burying receivers for improved time-lapse seismic repeatability: Co2crc otway field experiment: Geophysical Prospecting, **63**, 55–69.

Stuart, G. K., S. E. Minkoff, and F. Pereira, 2019, A two-stage markov chain monte carlo

method for seismic inversion and uncertainty quantification: Geophysics, **84**, R1003–R1020.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

———, 2005, Inverse problem theory and methods for model parameter estimation: SIAM.

Tucker, B., A. R. Hoover, and P. B. Flemings, 2000, Time-lapse (4-d) seismic monitoring of primary production of turbidite reservoirs at south timbalier block 295, offshore louisiana, gulf of mexico: Geophysics, **65**, 351–367.

Vigh, D. and E. W. Starr, 2008, 3d prestack plane-wave, full-waveform inversion: Geophysics, **73**, VE135–VE144.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Visser, G., P. Guo, and E. Saygin, 2019, Bayesian transdimensional seismic full-waveform inversion with a dipping layer parameterization: Geophysics, **84**, R845–R858.

Wang, Z., M. E. Cates, and R. T. Langan, 1998, Seismic monitoring of a co2 flood in a carbonate reservoir: A rock physics study: Geophysics, **63**, 1604–1617.

Watanabe, T., S. Shimizu, E. Asakawa, and T. Matsuoka, 2004, Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring, *in* SEG Technical Program Expanded Abstracts 2004, 2323–2326, Society of Exploration Geophysicists.

Willemsen, B., J. Cao, and B. Roy, 2016, The impact of the acoustic approximation on time-lapse fwi: Presented at the 2016 SEG International Exposition and Annual Meeting.

Willemsen, B. and A. Malcolm, 2015, Regularizing velocity differences in time-lapse fwi using gradient mismatch information: Presented at the 2015 SEG Annual Meeting.

Willemsen, L. A., 2017, Problems with a localized nature in exploration seismology: PhD thesis, Massachusetts Institute of Technology.

Williamson, P., A. Cherrett, and P. Sexton, 2007, A new approach to warping for quantitative

time–lapse characterisation: 69th EAGE Conference and Exhibition incorporating SPE EUROPEC 2007, cp–27.

Wilson, G. A., M. E. Willis, and A. Ellmauthaler, 2021, Evaluating 3d and 4d das vsp image quality of subsea carbon storage: The Leading Edge, **40**, 261–266.

Yang, D., F. Liu, S. Morton, A. Malcolm, and M. Fehler, 2016, Time-lapse full-waveform inversion with ocean-bottom-cable data: Application on valhall field: Geophysics, **81**, R225–R235.

Yang, D., M. Meadows, P. Inderwiesen, J. Landa, A. Malcolm, and M. Fehler, 2015a, Double-difference waveform inversion: Feasibility and robustness study with pressure data: Geophysics, **80**, M129–M141.

Yang, D., Y. Zheng, M. Fehler, and A. Malcolm, 2012, Target-oriented time-lapse waveform inversion using virtual survey, *in* SEG Technical Program Expanded Abstracts 2012, 1–5, Society of Exploration Geophysicists.

Yang, P., J. Gao, and B. Wang, 2015b, A graphics processing unit implementation of time-domain full-waveform inversiongpu implementation of fwi: Geophysics, **80**, F31–F39.

Yoon, K. and K. J. Marfurt, 2006, Reverse-time migration using the poynting vector: Exploration Geophysics, **37**, 102–107.

Zhang, X. and A. Curtis, 2021, Bayesian full-waveform inversion with realistic priors: Geophysics, **86**, A45–A49.

Zhang, Z. and L. Huang, 2013, Double-difference elastic-waveform inversion with prior information for time-lapse monitoring: Geophysics, **78**, R259–R273.

Zheng, Y., P. Barton, and S. Singh, 2011, Strategies for elastic full waveform inversion of time-lapse ocean bottom cable (obc) seismic data, *in* SEG Technical Program Expanded Abstracts 2011, 4195–4200, Society of Exploration Geophysicists.

Zhou, W. and D. Lumley, 2021a, Central-difference time-lapse 4d seismic full-waveform inversion: Geophysics, **86**, R161–R172.

———, 2021b, Non-repeatability effects on time-lapse 4d seismic full waveform inversion for

ocean-bottom node data: Geophysics, **86**, 1–60.

———, 2021c, Nonrepeatability effects on time-lapse 4d seismic full-waveform inversion for ocean-bottom node data: Geophysics, **86**, R547–R561.

Zwartjes, P., A. Mateeva, D. Chalenski, Y. Duan, D. Kiyashchenko, and J. Lopez, 2018, Frequent, multi-well, stand-alone 3d das vsp for low-cost reservoir monitoring in deepwater, *in* SEG Technical Program Expanded Abstracts 2018, 4948–4952, Society of Exploration Geophysicists.