Quantitative 4D seismic imaging in complex media using 2D full waveform inversion

Amir Asnaashari

▶ **To cite this version:**

Amir Asnaashari. Quantitative 4D seismic imaging in complex media using 2D full waveform inversion. Geophysics [physics.geo-ph]. Université de Grenoble, 2013. English. NNT : . tel-00932597v1

HAL Id: tel-00932597

https://theses.hal.science/tel-00932597v1

Submitted on 17 Jan 2014 (v1), last revised 23 Sep 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Sciences de la Terre, l'Univers et l'Environnement**

Arrêté ministériel : 7 août 2006

Présentée par

Amir ASNAASHARI

Thèse dirigée par **Jean VIRIEUX, Stéphane GARAMBOIS et François AUDEBERT**

préparée au sein de l' **Institut des Sciences de la Terre**
et de **l'école doctorale Terre Univers Environnement**

Quantitative 4D seismic imaging in complex media using 2D full waveform inversion

14 octobre 2013 ,
devant le jury composé de :

François AUDEBERT
Coordinateur R&D, TOTAL CSTJF, France, Tuteur Entreprise
Hervé CHAURIS
Professeur Ecole des Mines de Paris, France, Rapporteur
Stéphane GARAMBOIS
Maitre de conference Université Joseph Fourier, ISTerre, France, Directeur de thèse
René-Édouard PLESSIX
Chercheur Principal Shell, Pays-bas, Rapporteur
Satish SINGH
Professeur Institut de Physique du Globe de Paris, France, Examinateur
Jean VIRIEUX
Professeur Université Joseph Fourier, ISTerre, France, Directeur de thèse

Acknowledgement

I would like, first of all, to thank my supervisors: Jean Virieux, Stéphane Garambois and François Audebert. I thank their valuable and constructive suggestions during the planning and development of this research work. I thank their support, their advice, their sympathy, their monitoring and involvement in my work and their availability even during week-ends. I appreciate them for their helpful discussions, even during breakfast or dinner time in the period of conferences. I thank them for giving me this opportunity to learn several things from them.

I would like to thank TOTAL for supporting my thesis and providing me excellent conditions for my studies, which allowed me to participate in several international conferences, workshops and a summer school, where I acquired important technical knowledge which helped me for developing my work.

I appreciate René-Édouard Plessix and Hervé Chauris for accepting to review my work. I also thank Satish Singh for accepting to be in my jury as examiner.

I would like to specially thank Romain Brossier for his constructive suggestions, for his availability, his contribution and his friendship during my thesis. Also, my great appreciation to Romain for his knowledge and his new ideas. I should also thank Pierre Thore who was first my official supervisor from TOTAL at the begining of my thesis. I thank him for his contribution and his suggestions particularly for time-lapse aspects.

I thank the permenant members of SEISCOPE consortium: Stéphane Operto, Alessandra Ribodetti, Michel Dietrich and Ludovic Métivier for sharing their knowledge during video-conferences. I specially appereciate Stéphane and Michel for their constructive comments and suggestions for my thesis, particularly for real data application. I thank Alessandra for her encouragement, her kindness, and her helpful advices. My thanks to Ludovic for his helpful mathematical advices and his friendship.

I would like to thank my colleagues, François Lavoué and Isabella Masoni. My special thanks to François for sharing a friendly office at last year of my thesis and for his support, his friendship and his helpful advices for my work. I am very thankful to Isabella and appreciate her friendship and her kindness for reviewing my papers in terms of English corrections. I thank Bastien Dupuy for the scientific discussion on downscaling methods and his advices. I also thank Yaser Gholami, Clarita Castellanos, Aurelien Roques and Cécile Bruyant for their great friendship, their support, their suggestions and discussions. I enjoyed being with them.

I would like to thank the other members of SEISCOPE consortium (ancient and new members): Vincent Prieux, Guanghui Hu, Vincent Etienne, Damien Pageot, Yuelian Jia, Wei Zhou, Yang Li, Boussad Beldjenna, François Bretaudeau, Benjamin Pajot, Jose Tago, Paul Wellington for their friendship and for sharing their knowledge during video-conferences and conferences.

I thank different persons in TOTAL Center in Pau where I spent partial of my thesis period. I would like to thank Ramin Nawab and Kaveh Dehghan who proposed me this interesting PhD subject after my master internship. I also thank Cyril Agut, Paul Williamson, Paul Sexton, Victor Martin, Jean-luc Boelle and Henri Calandra for their helpful advices, our scientific discussions and their involvement in my thesis. My special thanks to Cyril, Jean-Rèmi Pontvianne and Hafedh Ben-Hadj-Ali for solving the informatique problems in Total site and their friendship. I appreciate Christian Hubans and Maud Stankoff-Godart for their constructive comments and suggestions for time-lapse aspects of the real data application.

I thank all the members of ISTerre for providing such a comfortable environment to work and their scientific contributions. I thank all PhD students of ISTerre for their friendship. I enjoyed the every moment we spent together. Special thanks to Jing, Yann, Sophie, Johanes, Ainhoa, Anne, Hilal, Afifa, Anne, Alex, Stefano, François, Olga and Virginie for their wonderful friendship. I spent great time with them. Also I should not forget all of my office mates during this thesis: Aurore, Benoit, Boumédiène, Piero, Florent, Marie, Simon, and Bertrand.

Many thanks to my other friends: Kaveh, Maryam, Behrooz, Shadi, Azadeh, Elahe, Pania, Farzad, Somi, Alireza, Amir, Zahra, Amirhossein and Mehdi for their wonderful friendship. I thank them for their time and outstanding support from the beginning to the end. I also thank all other friends all over the world for every happy time we spent together.

My very special thanks to Zahra, Mohammadreza and Hiva for their great friendship, our scientific discussion and their kindness from the beginning to the end of my thesis.

I thank my family: my parents, my sisters, my brother-in-law, my lovely niece, my grandparents, and my aunts and uncles for their long distance but limitless support and encouragement that yield a great motivation during all of my long term educations.

At the end, please accept my apologizes if I forget someone's name.

Résumé

Le suivi temporel est un processus d'acquisition et d'analyse d'acquisitions multiples répétées au même endroit sur la même cible à différentes périodes de temps. Cela s'applique bien à l'exploration sismique quand les propriétés de la cible varient au cours du temps comme pour les réservoirs pétroliers. Cette technique de sismique, dite 4D en raison de l'intégration du temps dans la construction des images, permet une détection et une estimation des variations du sous-sol survenues lors de l'évolution en temps du milieu. En particulier, dans l'industrie, le suivi et la surveillance peuvent améliorer notre compréhension d'un réservoir de pétrole/gaz ou d'un site de stockage de CO_2. Analyser la sismique 4D peut aider à mieux gérer les programmes de production des réservoirs. Ainsi, des acquisitions répétées permettent de suivre l'évolution des fronts de fluide injecté: on peut optimiser les programmes d'injection de fluides pour une récupération améliorée des hydrocarbures (enhanced oil recovery).

Plusieurs méthodes ont été développées pour l'imagerie variable dans le temps en utilisant les informations des ondes sismiques. Dans ma thèse, je montre que l'inversion de forme d'onde complété (FWI) peut être utilisée pour cette imagerie. Cette méthode offre des images sismiques quantitatives haute résolution. Elle est une technique prometteuse pour reconstruire les petites variations de propriétés physiques macro-échelle du sous-sol. Sur une cible identifiée pour ces imageries 4D, plusieurs informations a priori sont souvent disponibles et peuvent être utilisées pour augmenter la résolution de l'image. J'ai introduit ces informations grâce à la définition d'un modèle a priori dans une approche classique FWI en l'accompagnant de la construction d'un modèle d'incertitudes a priori. De plus, j'ai introduit une pondération dynamique de manière à réduire l'importance de ces modèles a priori lors de la convergence finale. Sur des exemples synthéques réalistes, j'ai montré que l'inversion FWI est moins sensible au modèle initial (qui peut donc être moins précis) grâce à cette utilisation de l'information a priori. Il est donc possible d'obtenir un modèle très précis comme modèle de base pour l'imagerie 4D.

Une fois la reconstruction d'un tel modèle atteinte, plusieurs stratégies peuvent être utilisées pour évaluer les changements de paramètres physiques. On peut réaliser deux reconstructions indépendantes et faire la différence des deux modèles reconstruits: on parle de différence parallèle. On peut aussi effectuer une différence séquentielle où l'inversion de l'ensemble de données de la second acqusition, dite moniteur, se fait à partir du modèle de base et non plus à partir du modèle utilisé initialement. Enfin, l'approche double-différence conduit à l'inversion des différences entre les deux jeux de données que l'on rajoute aux données synthétiques du modèle de base reconstruit. J'étudie quelle stratégie est à adopter pour obtenir des changements vitesse plus précis et plus robustes. En plus, je propose une imagerie 4D ciblée en construisant un modèle d'incertitude a priori grâce à une information (si elle existe) sur la locatisation potentielle des variations attendues. Il est démontré que l'inversion 4D ciblée empêche l'apparition d'artéfacts en dehors des zones cibles: on évite la contamination des zones extérieures qui

pourrait compromettre la reconstruction des changements 4D réels.

Une étude de sensibilité, concernant l'échantillonnage en fréquence pour cette imagerie 4D, montre qu'il est nécessaire de faire agir simultanément un grand nombre de fréquences au cours d'un cycle d'inversion. Ce faisant, l'inversion fournit un modèle de base plus précis que l'approche temporelle, ainsi qu'un modèle des variations 4D plus robuste avec moins d'artéfacts. Toutefois, la FWI effectuée dans le domaine temporel semble être une approche plus intéressante pour l'imagerie 4D. Enfin, l'approche d'inversion 4D régularisée avec un modèle a priori est appliquée sur des ensembles de données réelles d'acquisitions sismiques répétées fournis par TOTAL. Cette reconstruction des variations locales s'inscrit dans un projet d'injection de vapeur pour améliorer la récupération des hydro-carbures: Il est possible de reconstituer des variations de vitesse fines causées par la vapeur injectée.

Abstract

Time-lapse monitoring is the process of acquiring and analysing multiple seismic surveys, repeated at the same place at different time periods. This seismic technique, called 4D because of the integration time in the construction of images, allows detection and estimation of the subsurface parameter variations occured through a time evolution. Particularly, in industries, the monitoring can improve our understanding of a producing oil/gas reservoir and CO_2 storage site. Analyzing the time-lapse seismics can help to better manage production programs of reservoirs. In addition, repeated surveys can monitor the evolution of injected fluid fronts and can permit to optimize injection programs which are considered for enhanced oil recovery (EOR) techniques.

Several methods have been developed for time-lapse imaging using seismic wave information. In my thesis, I show that full waveform inversion (FWI) can be used for time-lapse imaging, since this method delivers high-resolution quantitative seismic images. It is a promising technique to recover small variations of macro-scale physical properties of the subsurface. In time-lapse applications, several sources of prior information are often available and should be used to increase the image reliability and its resolution. I have introduced this information through a definition of a prior model in a classical FWI approach by also considering a prior uncertainty model. In addition, I have suggested a dynamic weighting to reduce the importance of these prior models in the final convergence. In realistic synthetic cases, I have shown how the prior model can reduce the sensitivity of FWI to a less accurate initial model. It is therefore possible to obtain a highly accurate baseline model for 4D imaging.

Once the baseline reconstruction is achieved, several strategies can be used to assess the physical parameter changes. We can make two independent reconstructions of baseline and monitor models and make subtraction of the two reconstructed models. This strategy is called parallel difference. The sequential difference strategy inverts the monitor dataset starting from the recovered baseline model, and not from the model used initially. Finally, the double-difference strategy inverts the difference data between two datasets which are added to the calculated baseline data computed in the recovered baseline model. I investigate which strategy should be adopted to get more robust and more accurate time-lapse velocity changes. In addition, I propose a target-oriented time-lapse imaging using regularized FWI including prior model and model weighting, if the prior information exists on the location of expected variations. It is shown that the target-oriented inversion prevents the occurrence of artifacts outside the target areas, which could contaminate and compromise the reconstruction of the effective time-lapse changes.

A sensitivity study, concerning several frequency decimations for time-lapse imaging, shows that the frequency-domain FWI requires a large number of frequencies inverting simultaneously.

By doing so, the inversion provides a more precise baseline model and more robust time-lapse variation model with less artifacts. However, the FWI performed in the time domain appears to be a more interesting approach for time-lapse imaging considering all frequency content. Finally, the regularized time-lapse FWI with prior model is applied to the real field time-lapse datasets provided by TOTAL. The reconstruction of local variations is part of a steam injection project to improve the recovery of hydrocarbons: it is possible to reconstruct the velocity variations caused by the injected steam.

Contents

Introduction

Time-lapse seismic reservoir monitoring has received much attention over the past decade for improving our understanding of complex phenomena in the upper crust from natural geological variations as well as from anthropagenic influences. Time-lapse monitoring is the process of acquiring and analysing multiple seismic surveys, repeated at the same place at different time periods, in order to image potential fluid changes in a producing reservoir and/or CO_2 storage site (Landrø, 2001; Lumley, 2001; Calvert, 2005). Due to the extra fourth dimension, which is time, the repeated 3D seismic survey is often called *4D seismics*.

The final goal of seismic monitoring is the estimation of fluid parameter variations in a reservoir during production time and/or CO_2 storage during stockage time. While fluid saturations, pressures and temperatures in the reservoir change, seismic reflection properties change according to their sensitivity to each parameter. In time-lapse applications, by assuming the solid skeleton is time-invariant during production time (sometimes this may not be the case, due to physical changes and/or chemical reactions), it is possible to separate the dynamic fluid flow properties from the static geology footprint in seismic data and to produce images of the time-variant fluid changes (Lumley, 1995b, 2001). Figure 1 shows two seismic attributes acquired at two different times. By interpreting them, it is possible to detect the changes of oil water contact (OWC) over time. Amplitude and phase of seismic data change because of substitution of oil by water.

Compared to the 3D seismic, which is an exploration tool, 4D seismics is an important reservoir engineering management tool. Time-lapse images can identify bypassed hydrocarbon and show potential fluid migration paths. Therefore, it can be used for making a decision to drill or not to drill new wells in the field. Another example of the use of time-lapse monitoring concerns one of the EOR (Enhanced Oil Recovery) techniques, related to injection of fluid (water, steam, gas, etc.) into the reservoir. Repeated surveys can monitor the evolution of injected fluid fronts and can permit to optimize injection programs (Lumley, 2001). In addition, time-lapse models can distinguish the fluid-flow properties of sealing or non-sealing faults in complex reservoir.

Acquisition & processing

Repeatability in seismic acquisition is a crucial issue for time-lapse monitoring. Advances in repeatability of acquisitions can result in a major enhancement in time-lapse signal-to-noise ratio (Zabihi Naeini, 2012). Several studies have been shown that small changes in tides, water velocity, near surface properties (due to temperature changes), ambient noise, source and receiver positioning, for example, can produce significant non-repeatability effects on time-lapse

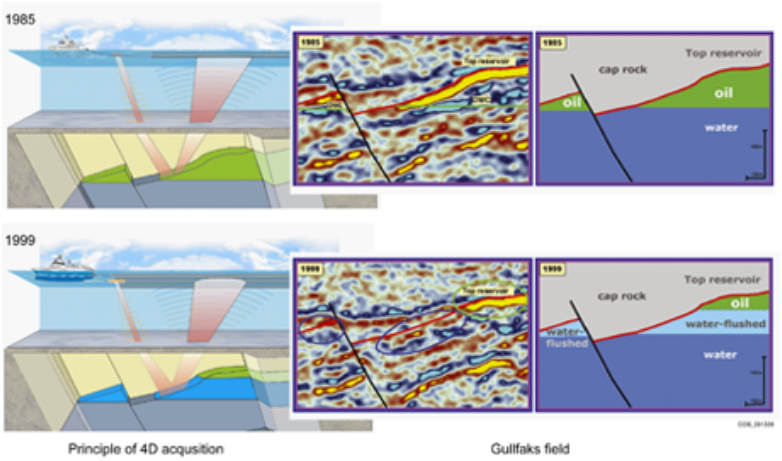


Figure 1: A repeated survey allows to monitor fluid variation in the subsurface (from Statoil).

data (Moldoveanu et al., 1996; Beasley et al., 1997; Rennie et al., 1997; Porter-Hirsche and Hirsche, 1998; Ronen et al., 1999; MacKay et al., 2003; Bertrand and MacBeth, 2003). Some of these acquisition effects can be reduced by performing more repeatable acquisition geometries, using minimized streamer cable feather, using ocean-bottom cable (OBC), and with the installation of permanent sensors on the site (Lumley, 2001), as shown by two examples: SeisMovie technology suggested by CGG (`http://www.cgg.com/default.aspx?cid=5899`) and OptoSeis system recently installed in deepwater offshore Brazil by PGS (`http://www.pgs.com/en/Geophysical-Services/4D-Seismic/Permanent-Monitoring/`). Indeed, permanent receivers are widely used in microseismic projects and could also be used for permanent reservoir monitoring. Some other effects such as ocean velocity variation and even changes in source and receiver locations could be reduced during the time-lapse data processing steps, using statics correction and 4D binning (Zabihi Naeini et al., 2009; Zabihi Naeini, 2012; Zabihi Naeini et al., 2012).

The main aim of 4D processing is the mitigation of the 4D noise caused by changes in acquisition parameters or environmental conditions, and the improvement of the 4D signature of the reservoir caused by changes in fluid, pressure and stress. For this reason, there is a need to obtain excellent 3D or 2D seismic images for each dataset, and simultaneously optimize time-lapse repeatability in areas without any changes. This processing workflow is known as *cross equalization* in the industry (Lumley, 2001). The main purpose of 4D cross-equalization tools is to process repeated vintage data together in order to equalize spectral bandwidth, amplitude gain variations, and event positioning to optimize 4D seismic difference anomalies (Harris and Henry, 1998; Rickett and Lumley, 1998). Figure 2 shows an example of the cross-equalization process (before and after). We can see that, after applying the cross-equalization process, the spectral bandwidth of both surveys become almost similar.

Several processing steps should be performed for time-lapse projects. 4D binning is an essential step to improve the repeatability of time-lapse (4D) data. Binning is a strategy to select the most compatible traces which are best matched between two surveys in terms of source, receiver and midpoint position, offset and azimuth. In other words, the 3D image quality is reduced in order to gain better subtraction between repeated vintages (Meunier and Herculin, 2003). The binning criteria corresponding to the selection of the most compatible

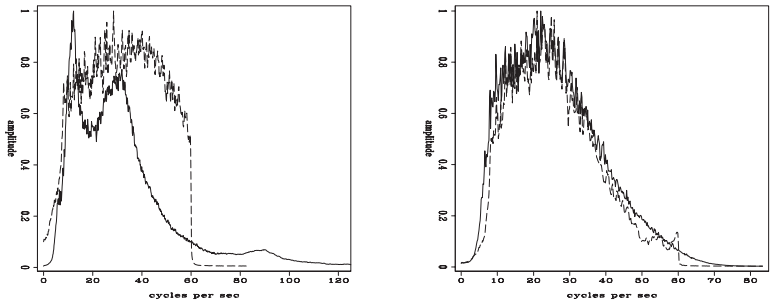


Figure 2: An example of the cross-equalization process. Amplitude spectra before (left) and after (right) cross-equalization. The solid line corresponds to the baseline survey and the dashed line to the monitor survey acquired after 12 years later (from Rickett and Lumley (1998)).

traces can be extended to include statistical measurements of 4D data quality based on cross-correlation and normalized RMS (NRMS) or predictability (`http://www.cgg.com`). Recently, Zabihi Naeini et al. (2009) have proposed a simultaneous multi-vintage (SMV) 4D binning algorithm which provides the best possible repeatability across all vintages and in each bin, instead of cascaded (pair-wise) processing (Figure 3).

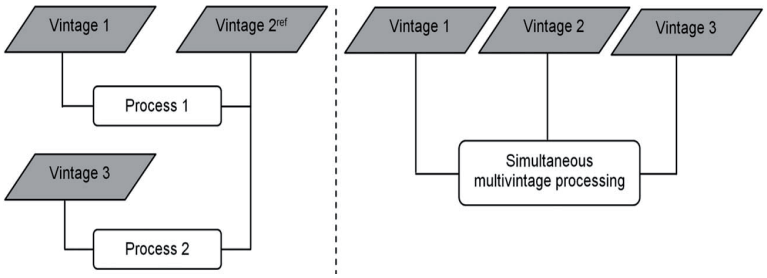


Figure 3: Comparison of cascaded (left) and simultaneous (right) multi-vintage processing approaches for a time-lapse study with three vintages (from Zabihi Naeini et al. (2010)). In cascaded processing, the vintage 1 and 2 are compared as well as the vintage 2 and 3, but comparison of vintage 1 and 3 is not considered. While in simultaneous approaches, all the possible comparisons are considered at the same time.

Rock physics and time-lapse data

Fluid parameter variations, pressure and temperature changes have a direct effect on subsurface macroscale parameters such as the P-wave velocity V_p, the S-wave velocity V_s, the density, etc.. In the 1980's, it has been shown that thermal effects and/or presence of free gas (steam injection) on heavy-oil saturated field lead to large decreases in seismic rock velocity (Nur et al., 1984; Pullin et al., 1987; Greaves and Fulp, 1987; Wang and Nur, 1990). The significant decrease of rock seismic impedance due to the presence of free gas allows to track the motion of gas-fluid contact and detect chambers of injected gases (Johnstad et al., 1995; Harris et al., 1996; QueiBer and Singh, 2013).

Numerous models aim to link the P- and S-wave velocities (V_p, V_s) to the porosity (ϕ) in a downscaling procedure. Mainly, when porosity increases, P- or S-wave velocity decreases, as well as pressure and shear moduli. Empirical relations try to make a link between quantities V_p and ϕ, but they are dependent on the rock type because they are based on experimental measurements. For example, Han et al. (1986) have established linear relations between V_p, V_s, porosity and the clay content for clastic sediments during their full diagenetic evolution (from unconsolidated sands to sandstones). Other empirical relations exist, as those from Wyllie et al. (1956), Raymer et al. (1980) or Raiga-Clemenceau et al. (1988). However, in these relations, the P-wave velocity is only linked to the porosity while, in real media, the drained medium (skeleton) rigidity plays an important role. This skeleton rigidity depends also on the mineral grains rigidity and the grain arrangement (geometry, compaction, etc.). Moreover, we need to consider the rigidity of fluids inside pores. On the other hand, numerous empirical relations between V_p and V_s have been established for various rocks (as an example, summarized by Castagna et al. (1993) for limestones, sandstones, shales and dolomites).

Therefore, if it is possible to recover the time-lapse variations related to macroscale parameters using time-lapse seismic inversions, we can hope to get fluid and poroelastic parameters variations by downscaling approaches (Gassmann, 1951; Berryman et al., 2002; Avseth et al., 2005; Rubino and Velis, 2011; Dupuy, 2011; De Barros et al., 2012).

Analysis of time-lapse data: how to get maps of changes

After processing datasets and obtaining optimal 4D seismic signal anomalies (with less 4D noise), we can analyse the time-lapse signals to get qualitative and quantitative interpretations. The first qualitative interpretation is to calibrate the time-lapse changes in the seismic data with changes observed in other reservoir data such as pressure and temperature data in well logs, production and injection history data, etc. This step can show that time-lapse seismic changes are real and are not due to artifacts related to acquisition and processing. At this step, we can have a qualitative idea about fluid saturation changes and/or pressure and temperature variations inside reservoirs (Lumley, 1995a; Jenkins et al., 1997; Anderson et al., 1997; He et al., 1998).

To make a better reservoir management recommendation, it is preferred to make a quantitative time-lapse analysis. This can be done by inverting time-lapse seismic data to produce dynamic-property variation maps and this is an active research topic (Lumley, 2001). An attribute clustering technique has been used by Sonneland et al. (1997) to obtain oil and gas saturation maps from Gullfaks field. Later, time-lapse amplitude versus offset (AVO) inversion of reflectivity has been presented to estimate pressure and fluid saturation changes (Tura and Lumley, 1999; Landrø, 2001).

Another technique, widely used for quantitative time-lapse analyses, is a warping method to improve an alignment of seismic volume or image with the other one (align baseline and monitor surveys) (Hall et al., 2002, 2005; Hall, 2006; Williamson et al., 2007). Standard warping algorithms estimate time-shifts by selecting windows of data from one image and searching over time displacements to maximize the correlation (trace by trace). This method depends on the length of correlation window and may induce artifacts in the difference volume. Therefore, the correlation-based warping technique has shown insufficient resolution, lower stability and less accuracy (Williamson et al., 2007). The main limitation of this approach is the poor-sensitivity

to density variations (reflectivity), since it estimates the time-shifts which are sensitive to the velocity variations. However, Williamson et al. (2007) have proposed a new warping method to explain both time-shifts and amplitude changes and to provide stable estimation of velocity-change attributes. This new approach is limited to flat reflectors, for near-offsets and the velocity should vary smoothly laterally. Recently, (Hale, 2013) proposed a dynamic warping method which is able to align both time-shifts and amplitude between two traces. It appears that this dynamic warping could also be used for time-lapse applications in the future.

During the past decade, full waveform inversion (FWI) has become a promising technique for seismic inversion, especially in exploration projects (Virieux and Operto, 2009). This approach can deal with the complex datasets and the complex physical model parameters and it delivers high resolution and quantitative interpretation of macroscale physical parameters (Brossier, 2009; Brossier et al., 2009). This framework can provide P-wave, S-wave velocities, density, attenuation factor and even the anisotropy parameters through multiparameter inversion (Virieux and Operto, 2009; Brossier et al., 2009; Malinowski et al., 2011; Plessix, 2012; Prieux et al., 2013a,b; Gholami et al., 2013a). Moreover, difference-based FWI takes into account both phase and amplitude information of seismic data. FWI has been promoted as high resolution seismic imaging with a theoretical resolution of half the minimal wavelength. Moreover FWI aims for an improved automated workflow with less inputs from interpreters. The initial model design still remains an issue and if its building relies on travel-time tomography, some pickings, more or less automatic, will be required: an open question for the moment.

Considering these advantages, FWI could be an interesting technique to be used for inverting the time-lapse datasets, even if it is not yet widely applied (Gosselet and Singh, 2008; Abubakar et al., 2009; Plessix et al., 2010; QueiBer and Singh, 2013; Zheng et al., 2013).

FWI in brief

Today, one of the interesting imaging techniques is known as full waveform inversion (FWI). The purpose is the interpretation of the entire seismograms using complete wave propagation features. The seismic data-fitting procedure was presented as a local optimization method by Lailly (1983b) and Tarantola (1984a). It is shown to be a least-squares minimization of the misfit between observed and calculated data. Figure 4 shows an example of elastic FWI on a synthetic wide-aperture marine dataset (after Shipp and Singh (2002)). It shows the observed and final calculated data and final residuals. The relationship between the seismic data and the model parameters is nonlinear in the forward modeling. Therefore, the inversion appears to be nonlinear and hence this method is reduced to an iterative linearized approach. This seismic inversion method is called the full waveform inversion, because the full information content of the seismogram is considered in the optimization. The waveform inversion has been developed in time-domain inversion in 1980s (Tarantola, 1984a, 1987; Gauthier et al., 1986; Mora, 1989), and in 1990s in frequency domain (Pratt and Worthington, 1990; Pratt, 1990a). Later, the frequency-domain FWI became a widely applied seismic inversion method (see e.g. Pratt and Shipp (1999); Pratt and Symes (2002); Ravaut et al. (2004); Gelis et al. (2004); Operto and Virieux (2006); Sirgue et al. (2007); Malinowski and Operto (2008); Ben Hadj Ali (2009) and Plessix (2009)), however several nice applications have been also shown for time-domain FWI (Shipp and Singh, 2002; Sears et al., 2008, 2010).

In full waveform inversion, the initial model (often rather smooth but kinematically accu-

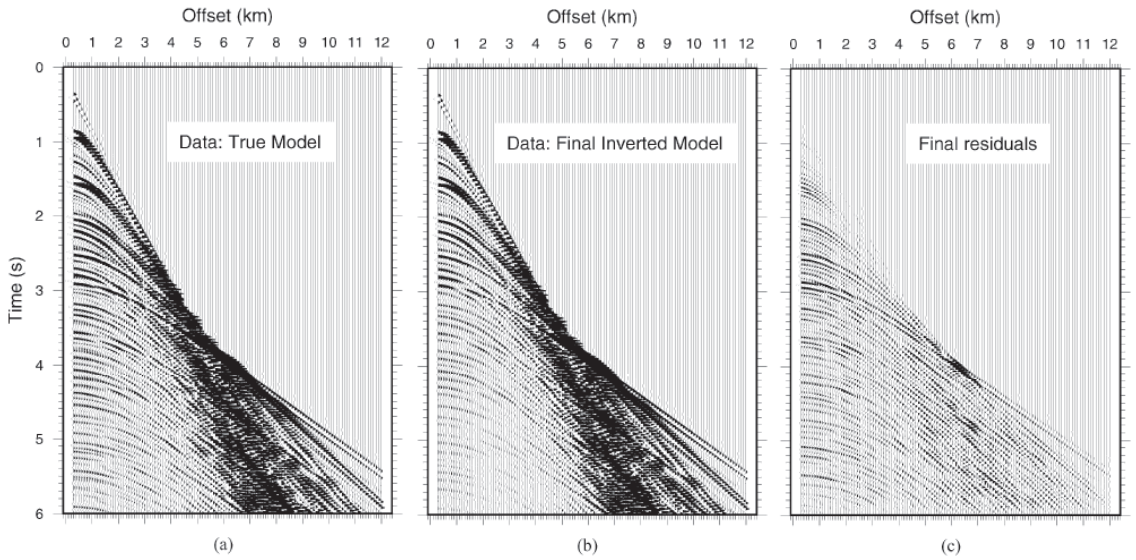


Figure 4: The wavefield inversion experiment on a synthetic wide-aperture marine streamer data with 12 km offset, after Shipp and Singh (2002). (a) The observed data, (b) the final data computed by elastic FWI, and (c) the residual between observed data and the FWI computed data.

rate) is updated with the perturbation obtained by an optimization algorithm. The updating model continues toward minimization of the misfit function until reaching the convergence. In general, the local optimization method has difficulties to converge toward the global minimum of the misfit function, if the initial point is far away from the global minimum. Due to the ill-posedness of the inverse problem, there is a possibility to get trapped in existing local minima. The main challenges for FWI concern the accuracy of the initial model and also the lack of low frequency content in the observed data as we relax conditions on the initial model with lower and lower frequency content. Therefore, in order to retrieve low spatial wavenumbers, it is essential that the observed data contain the transmitted and diving waves (wide-aperture) (Pratt et al., 1996; Sirgue and Pratt, 2004). These arrivals are useful for the reconstruction of the large wavelengths of the subsurface. Less accurate initial models lead to cycle-skipping problems. The initial model should predict arrival times with errors less than half of the period to avoid the cycle-skipping ambiguity (Virieux and Operto, 2009).

The starting model of FWI can be obtained by ray-tracing based methods (Červený, 2001; Virieux and Lambaré, 2007), reflection/refraction traveltime tomography (Taillandier et al., 2009; Roux et al., 2011; Prieux et al., 2013c) and by migration velocity analysis (MVA) techniques (Chauris et al., 2002a,b; Sava and Biondi, 2004; Symes, 2008), which all can usually provide a smooth initial model. The MVA procedure relies on iterative applications of prestack depth-migration and velocity analysis. The velocity updating can be performed through minimization of a differential semblance functional in the image domain (Symes and Carazzone, 1991; Chauris and Noble, 2001; Shen and Symes, 2008). The drawback of these MVA approaches is the computational cost of the iterative application of prestack depth migration and velocity analysis. Recently, Wang et al. (2013) have proposed an integrated workflow of

combination wave-equation tomography (WET) and full waveform inversion (FWI) through a hybrid misfit function. This workflow allows inverting for high-resolution subsurface velocity structure from a poor constrained initial model. Building accurate initial models for full waveform inversion still remains an important research topic (Chauris et al., 2008).

Objectives

In this thesis dissertation, I want to answer these following questions. Can the FWI method be used for the reconstruction of small time-lapse variations or not? Can this method deliver a robust time-lapse result for macroscale parameter changes? How to obtain accurate and robust time-lapse models using FWI approach? I mainly focus on inversion of time-lapse datasets in depth domain to get the map of P-wave velocity variations using 2D acoustic full waveform inversion (FWI). As a first step, time-lapse analysis is considered as a single parameter inversion. Multi-parameter FWI is still an issue for general exploration projects to accurately estimate several macroscale parameters. Therefore, in this thesis, the time-lapse inversion is restricted to acoustic approximation and single parameter inversion in order not to have influence of multi-parameter inversion problems on the estimation of small time-lapse variations. In this case, our conclusion will not be affected by the multi-parameter issue.

The goal of this thesis is to focus on the analysis of time-lapse data to get macroscale parameters in more complex 2D models. I will show how the time-lapse velocity can be obtained in a robust way with less time-lapse noise artifacts. However, the suggested approaches in this thesis can be extended to elastic and multi-parameter inversion in the future, which is an important issue to get more accurate and more robust fluid and microscale parameter changes by downscaling. Indeed, in addition to V_p, Q_p, the knowledge of V_s and Q_s changes may contribute to image fluid changes (Dupuy, 2011). Before transition to multi-parameter time-lapse inversion, it is crucial to develop the current FWI algorithms to be adapted for more precise multi-parameter estimation, an open active research line.

This research study includes four main objectives: (a) taking into account the available non-seismic prior data into the inversion scheme by adding prior model misfit term into data-driven FWI framework and showing that the prior model information is essential for time-lapse FWI , (b) better understanding of which time-lapse strategy can deliver more robust and more accurate time-lapse velocity variations. I also propose a target-oriented time-lapse inversion in order to make inversion focus only on the expected area of changes. Sensitivity analysis on synthetic datasets are performed as well, (c) studying a sensitivity analysis of reconstruction time-lapse variation models using different frequency decimations for the baseline (time zero) and monitor (greater times) inversions. Which sampling frequency can be suitable to recover baseline accurately, and consequently to reconstruct small time-lapse perturbations? I show that it may be better to use time-domain FWI for time-lapse applications, (d) applying the suggested time-lapse strategies to realistic synthetic and real field datasets (steam injection case).

In **Chapter 1**, I discuss how to include the prior model information into standard FWI, which is necessary for time-lapse applications. It is illustrated that the prior model term may significantly reduce the inversion sensitivity to less accurate initial model conditions. It is highlighted how the limited range of spatial wavenumber sampling due to the acquisition may

be compensated with the prior model information, in order to get a more accurate baseline model.

Chapter 2 explores different time-lapse FWI strategies and their sensitivities. After baseline reconstruction, several strategies can be used to recover the physical parameter changes, such as parallel difference (two separate inversions of baseline and monitor datasets), sequential difference (inversion of the monitor dataset starting from the recovered baseline model), and double-difference (inversion of the difference data starting from the recovered baseline model) strategies. I compare the robustness of these strategies on two synthetic datasets, on famous and complex Marmousi2 model (Martin et al., 2006) and a modified steam injection model originally introduced by Dai et al. (1995). In addition, a target-oriented time-lapse inversion is proposed based on the prior model and the prior weighting, in order to focus inversion on the expected area of time-lapse changes.

In **Chapter 3**, several studies are investigated in order to find which domain, time or frequency, would be more suitable for a time-lapse application. I propose to compare the results of an inversion approach in frequency domain inverting all the frequencies simultaneously (similar to the time-domain inversion) and several decimations on the selected frequencies on the Marmousi synthetic data. It is shown that to get more accurate time-lapse model with less 4D artifacts, it is necessary to perform frequency-domain FWI with dense frequency samples.

Chapter 4 contains the application of time-lapse FWI on a real field steam injection dataset (Canada). The presented workflow for synthetic cases at early chapters are also taken into account for this real field application.

Chapter 1

Regularized Full Waveform Inversion

Contents

In this chapter, we first discuss the forward modeling and standard full waveform inversion. We then present an article, published in *Geophysics* journal, which discusses how to include the prior model information into the standard procedure of FWI. For time-lapse inversion, it is necessary to have high accurate and high precise baseline model in order to reconstruct small time-lapse variations (Asnaashari et al., 2011). The issue with this result is that standard FWI, which generally only uses information from surface seismic data, cannot deliver the proper accuracy in baseline model estimation. Figure 1.1 shows the sensitivity of reconstructed time-lapse models by differential waveform inversion with respect to reference models. Figure 1.1j shows the recovered time-lapse variation which is similar to the true one. In these simple synthetic examples, we can clearly see that an improvement of the recovered baseline model leads to an improvement of the time-lapse model. For more details, please refer to Appendix A.1.

In order to increase the precision of the recovered baseline model, it may be interesting to add more information into the inversion scheme. These supplementary information addition to seismic data, called prior information, could be integrated inside additional terms of the objective function. This is the main goal of the article presented hereafter (section 1.2). It must be noted that, in time-lapse applications, there are usually several sources of available prior information, which may be crucial in term of final precision. This leads to end up with more accurate baseline and more robust time-lapse variation models.

1.1 Full waveform inversion

Generally, the full waveform inversion (FWI) is represented as data-driven approach based on a least-squares local optimization problem. A representation of inverse problem is performed on probabilistic maximum likelihood or generalized inverse formulation (Menke, 1984; Tarantola, 1984a). Lailly (1983a) and Tarantola (1984a) represented the seismic inversion problem by recasting the migration imaging principle (Claerbout, 1971, 1976), as a local optimization problem. The reader is referred to Virieux and Operto (2009) for an overview on the FWI problem in exploration geophysics. The generalized inverse problem is based on minimizing the difference between the observed data $\mathbf{d}_{obs}$, and the data calculated in an estimated model $\mathbf{d}_{cal}(\mathbf{m})$. Before going to inverse problem, we need to have an accurate forward modeling engine to simulate the calculated data.

1.1.1 Forward modeling

The modeling of the synthetic data is performed using full wave equation, which provides a data vector with complete waveforms sampling the subsurface at receiver positions. Several techniques of discretization of wave equation have been studied, such as finite-difference (FD) or finite-element (FE) methods. Among them, one of the possible methods is the finite element discontinuous Galerkin (DG) method (Käser and Dumbser, 2006; Dumbser and Käser, 2006; Brossier et al., 2008). This DG method allows the use of triangular/tetrahedral meshes, which is suitable for the handling of strong physical contrasts in the medium, including liquid/solid contact. However this method demands higher computational costs compared to the finite-difference method and is more difficult to be implemented. Therefore in this study, we use the finite-difference method to solve the 2D acoustic wave equation in the time domain,

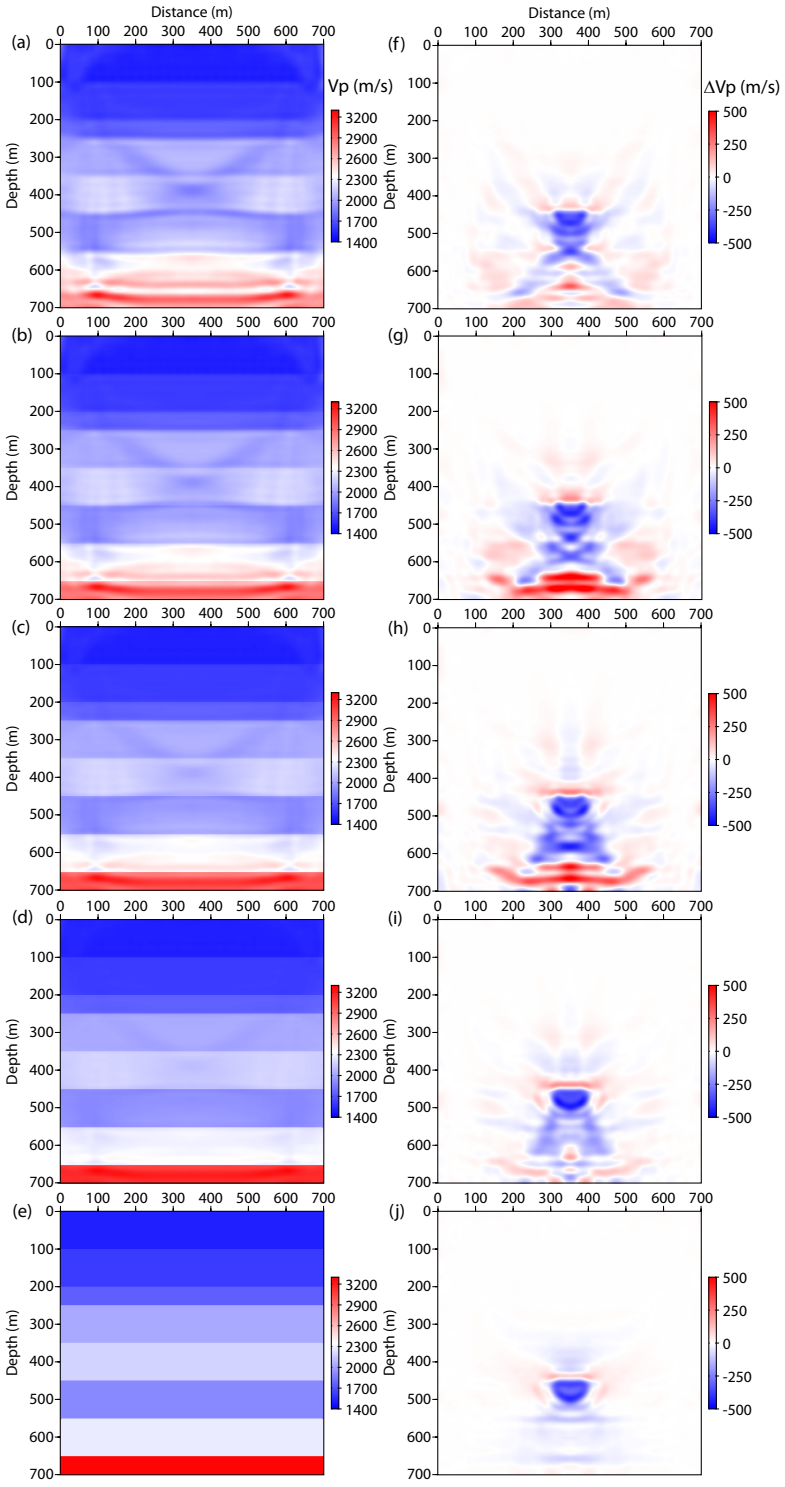


Figure 1.1: Reference and time-lapse V_p models: left panel shows the evolution of reference models from (a) reconstructed baseline image to (e) the true baseline model, and right panel illustrates each time-lapse image corresponding to the reference model at the left part.

$$\nabla.\Big(\frac{1}{\rho(x,z)}\nabla u(x,z,t)\Big) - \frac{1}{\rho(x,z)V_p^2(x,z)}\frac{\partial^2 u(x,z,t)}{\partial t^2} = -s(x,z,t), \tag{1.1}$$

where quantities ρ and V_p represent the density and velocity of the medium. u and s denote the pressure wavefield and an acoustic source, respectively. The finite-difference scheme suggested by Levander (1988) adapted for the acoustic case is used. The velocity-pressure formulation of the 2D acoustic wave equation, which is reduced from 2D P-SV formulation, is given by

$$\begin{aligned}
\frac{\partial u_x(x,z,t)}{\partial t} &= \kappa(x,z)\Big(\frac{\partial v_x(x,z,t)}{\partial x} + s(x,z,t)\Big)\\
\frac{\partial u_z(x,z,t)}{\partial t} &= \kappa(x,z)\frac{\partial v_z(x,z,t)}{\partial z}\\
\frac{\partial v_x(x,z,t)}{\partial t} &= b(x,z)\frac{\partial u(x,z,t)}{\partial x}\\
\frac{\partial v_z(x,z,t)}{\partial t} &= b(x,z)\frac{\partial u(x,z,t)}{\partial z},
\end{aligned} \tag{1.2}$$

where $v_x(x,z,t)$ and $v_z(x,z,t)$ are the horizontal and vertical particle velocities. The pressure components $u_x(x,z,t)$ and $u_z(x,z,t)$ are used to separate the horizontal and vertical derivatives and also to take into account for the perfectly matched layers (PML) absorbing condition (Berenger, 1994; Operto et al., 2002). The real pressure wavefield can be computed by $u(x,z,t) = u_x(x,z,t) + u_z(x,z,t)$. The quantities $b(x,z)$ and $\kappa(x,z)$ represent the inverse of density (buoyancy) and bulk modulus (ρV_p^2).

This system can be easily discretized using Virieux-Levander staggered-grid stencils (Virieux, 1986; Levander, 1988). Geometry of the staggered-grid is shown in Figure 1.2. The FD classic staggered-grid stencils provide calculated pressure wavefields with an accuracy of the fourth-order in space and second-order in time $\mathcal{O}(h^4, \Delta t^2)$. Spatial derivatives are discretized along the Cartesian directions. The FD grids are uniform in time and space domains.

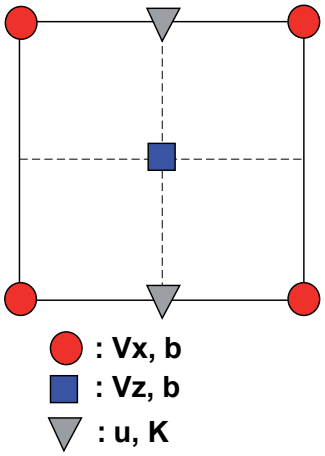


Figure 1.2: Geometry of the staggered grid for Virieux-Levander stencils adapted to the acoustic case (after Operto et al. (2007)).

1.1.1.1 Spatial discretization rule: numerical dispersion

The usual discretization rule is 10 and 5 grid points per minimum wavelength for the classic $\mathcal{O}(h^2)$ and $\mathcal{O}(h^4)$ staggered-grid stencils, respectively (Virieux, 1986; Levander, 1988). For our case, since we use the fourth-order stencils, the spatial discretization h should satisfy the condition 1.3,

$$\begin{aligned} h &< \frac{\lambda_{min}}{5} \\ h &< \frac{V_{p_{min}}}{5 f_{max}}, \end{aligned} \tag{1.3}$$

where $V_{p_{min}}$ and f_{max} denote minimum P-wave velocity of medium and maximum frequency content of source wavelet.

1.1.1.2 Time discretization rule: stability conditions

To keep numerical calculation stable, the time increment Δt must be less than or equal to the Courant limit (Richtmyer and Morton, 1967) known as the CFL condition. The stability condition in 2D case for $\mathcal{O}(h^4, \Delta t^2)$ is given by,

$$\Delta t \leq 0.606 \frac{h}{V_{p_{max}}}. \tag{1.4}$$

To have an accurate modeling of seismic waves, both conditions 1.3 and 1.4 must be satisfied.

1.1.2 Inverse problem

In general, the inverse problem involves optimization of a misfit functional, which measures the distance between observed and computed data. The misfit functional is defined based on the considered ℓ_2 or ℓ_1 norms in the data space. The theory of the generalized inverse problem (as a least-squares optimization), and different methods for model parameter estimation were established by Tarantola (1987). One of the main principle of the least-squares criterion of the misfit function underlies on the hypothesis that all uncertainties in the inverse problem are modelled with Gaussian distributions (Tarantola, 1987). In FWI, the residual or misfit vector $\Delta \mathbf{d}$, is defined as the difference between observed and computed data vectors at the receiver positions for each source. The residual vector is the criterion showing the closeness of the computed data vector to the observed data vector:

$$\Delta d = d_{obs} - d_{cal}(m). \tag{1.5}$$

In FWI, the inverse problem represents a nonlinear relation between data and model spaces. Hence, the solution of the inversion is computed with linearized iterative methods to minimize the misfit function. The seismic wavefield data in the time domain are represented by real-valued seismograms, and in the frequency domain by complex-valued data for each frequency component. In this chapter, we focus on inversion in time domain, hence the formulation is

shown in this domain. In Chapter 3, a complete study for the frequency-domain inversion for a time-lapse application will be discussed.

The solution of the forward problem is the wavefield of each seismic source in the model domain. Mathematically, the data is obtained by applying a sampling operator $\mathcal{P}_{data}$ (corresponding to the receivers) on the solution u in the model domain ($d_{cal} = \mathcal{P}_{data}u$). The objective of full waveform inversion is to seek the minimum of the misfit function starting from the initial model m_0 The ℓ_2 norm misfit function of data is defined by the least-squares equation as (Tarantola, 1987):

$$\begin{aligned} \mathcal{C}(m) &= \frac{1}{2}\int_0^T (d_{obs} - d_{cal}(m))^t (d_{obs} - d_{cal}(m))\, d\tau \\ &= \frac{1}{2}\int_0^T \Delta d^t \Delta d\, d\tau, \end{aligned} \tag{1.6}$$

where ^{t} is the transpose operator (reminding that the data in frequency domain is complex and the complex conjugate operator must be applied as well) and T is the trace length (recording time). Some weighting (more desirable to be related to the variance of data) may be applied on the residual vector to equalize the role of each residual vector components in the misfit function. This data weighting matrix (W_d) can be seen as an inverse of the data covariance matrix (Tarantola, 2005). When the weighting matrix is equal to identity matrix, it means same weight is applied to all data misfit components. For a seismic wide-aperture/wide-azimuth dataset, the weighting matrix can be applied based on the offset distance and/or aperture angle. The weighting depends on the distribution of the misfit between the observed and computed data vectors. The weighting coefficient on each components of misfit vector can penalise the data vector components. By considering the prior information on the data using weighting matrix, the misfit function (equation 1.6) would appear as a weighted misfit function and leads to a weighted least-squares problem:

$$\mathcal{C}(m) = \frac{1}{2}\int_0^T \Delta d^t W_d \Delta d\, d\tau. \tag{1.7}$$

In FWI, because of the huge number of estimated parameters and of the importance of computational cost for solving the direct problem, FWI method generally uses linearized and iterative local optimization scheme.

Let us consider the framework of small perturbation or Born approximation in the scattering theory (see e.g. Born and Wolf (1980); Hudson and Heritage (1981); Beydoun and Tarantola (1988); Beydoun and Mendes (1989); Coates and Chapman (1990); Born and Wolf (1993); Forgues and Lambaré (1997)). At each iteration k, we can express the updated model based on the starting model m^{k-1} and the perturbation vector Δm^k: $m^k = m^{k-1} + \Delta m^k$. We search for the local minimum of the misfit function $\mathcal{C}(m^k)$ departing from the model m^{k-1} of previous iteration:

$$\mathcal{C}(m^k) = \mathcal{C}(m^{k-1} + \Delta m^k). \tag{1.8}$$

By assuming that the model perturbation vector is small enough with respect to model vector, the second-order Taylor expansion of misfit function gives:

$$\mathcal{C}(m^{k-1} + \Delta m^k) = \mathcal{C}(m^{k-1}) + \sum_{i=1}^{n} \frac{\partial \mathcal{C}(m^{k-1})}{\partial m_i} \Delta m_i^k$$

$$+ \quad \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2\mathcal{C}(m^{k-1})}{\partial m_i \partial m_j}\Delta m_i^k \Delta m_j^k + \mathcal{O}(m^3), \tag{1.9}$$

where n is the size of model parameter vector. Taking the derivative of equation 1.9 with respect to the model parameter m_l leads to:

$$\frac{\partial\mathcal{C}(m^k)}{\partial m_l} = \frac{\partial\mathcal{C}(m^{k-1})}{\partial m_l} + \sum_{i=1}^{n}\frac{\partial^2\mathcal{C}(m^{k-1})}{\partial m_i \partial m_l}\Delta m_i^k. \tag{1.10}$$

The optimum solution is obtained, when the quadratic ($\mathcal{O}(m^3) = 0$) misfit function is tracked in minimum ($\partial\mathcal{C}(m^k)/\partial m_l = 0$):

$$\frac{\partial^2\mathcal{C}(m^{k-1})}{\partial m^2}\Delta m^k = -\frac{\partial\mathcal{C}(m^{k-1})}{\partial m}. \tag{1.11}$$

The first-order term $\partial\mathcal{C}(\mathbf{m}^{k-1})/\partial\mathbf{m}$ in equation 1.11 is the gradient vector of the misfit function and the second order term $\partial^2\mathcal{C}(\mathbf{m}^{k-1})/\partial\mathbf{m}^2$ is the Hessian matrix. Equation 1.11 can be expressed in linear form as

$$H^{k-1}\Delta m^k = -\mathcal{G}^{k-1}, \tag{1.12}$$

the so-called normal equation. We may solve this equation through a Newton approach delivering the minimum of misfit function in one iteration in case of a inverse problem with a linear forward problem. For the FWI, since the inverse problem is highly non-linear, it must be solved iteratively with a linearized problem at each iteration. The complete forward problem is performed at each update of the model. The forward problem is not approached by any means for a linear formulation. The gradient of misfit function $\mathcal{G}$, is the vector whose components are the partial derivatives of misfit function $\mathcal{C}$ with respect to the model parameters:

$$\mathcal{G} = (\frac{\partial\mathcal{C}}{\partial m_1}, \frac{\partial\mathcal{C}}{\partial m_2}, \cdots, \frac{\partial\mathcal{C}}{\partial m_n})^t. \tag{1.13}$$

In other words, the negative of the gradient of misfit function ($\Delta\mathcal{C}$) represents the direction of descending toward the minimum. The Hessian matrix is the square matrix of second-order partial derivatives of misfit function:

$$H = \begin{bmatrix} \frac{\partial^2\mathcal{C}}{\partial m_1^2} & \frac{\partial^2\mathcal{C}}{\partial m_1 \partial m_2} & \cdots & \frac{\partial^2\mathcal{C}}{\partial m_1 \partial m_n} \\ \frac{\partial^2\mathcal{C}}{\partial m_2 \partial m_1} & \frac{\partial^2\mathcal{C}}{\partial m_2^2} & \cdots & \frac{\partial^2\mathcal{C}}{\partial m_2 \partial m_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2\mathcal{C}}{\partial m_n \partial m_1} & \frac{\partial^2\mathcal{C}}{\partial m_n \partial m_2} & \cdots & \frac{\partial^2\mathcal{C}}{\partial m_n^2} \end{bmatrix}. \tag{1.14}$$

When the misfit function is continuous, the Hessian matrix is a symmetric square matrix (*i.e.* $\frac{\partial^2\mathcal{C}}{\partial m_i \partial m_j} = \frac{\partial^2\mathcal{C}}{\partial m_j \partial m_i}$). The Hessian matrix represents the curvature trend of the quadratic misfit function. The model perturbation vector is the solution of the linear system 1.12, which is solved by iterative optimization methods.

The gradient and Hessian expressions of the misfit function (equation 1.7) should be detailed here. Taking the first-order derivative of the misfit function with respect to model parameter leads to an expression for the gradient as

$$\begin{aligned} \mathcal{G}(m) &= -\frac{\partial \mathcal{C}(m)}{\partial m} \\ &= -\int_0^T (\frac{\partial d_{cal}(m)}{\partial m})^t W_d (d_{obs} - d_{cal}(m))\, d\tau \\ &= -\int_0^T J^t W_d \Delta d\, d\tau. \end{aligned} \tag{1.15}$$

The operator **J** is the Jacobian or Fréchet derivative matrix. The Jacobian matrix is the first-order derivative of calculated data vector with respect to model vector. Note that the operator **J** represents the discrete first-order Born operator if the linearized direct problem is considered $\Delta d = J\Delta m$ (Tarantola, 1987; Pratt et al., 1996; Brossier, 2009). By taking the second-order derivative of misfit function or the first-order derivative of the gradient with respect to the model vector, the Hessian operator is expressed as

$$H = -\frac{\partial^2 \mathcal{C}(m)}{\partial m^2} = \int_0^T \left[J^t W_d J + (\frac{\partial J}{\partial m})^t W_d (\Delta d \cdots \Delta d) \right] d\tau. \tag{1.16}$$

Replacing the gradient (equation 1.15) and the Hessian (equation 1.16) into the normal equation 1.12, the linear system would be expressed as

$$\Big(\int_0^T J^t W_d J + (\frac{\partial J}{\partial m})^t W_d (\Delta d \cdots \Delta d)\, d\tau \Big)^{k-1} \Delta m^{(k)} = \Big(\int_0^T J^t W_d \Delta d\, d\tau \Big)^{k-1}. \tag{1.17}$$

One should note that, for a linear direct problems where **d=G.m**, the second order derivative of the data with respect to model parameters is zero. Hence, the second term of equation 1.16 on the left-hand side would disappear. The first term of the Hessian is referred to as approximated Hessian $H_a = \int_0^T J^\dagger W_d J\, d\tau$. Methods which solve model perturbation using the approximated Hessian H_a, is referred to as Gauss-Newton methods. An estimation of the approximated Hessian, called pseudo-Hessian, is proposed by Shin et al. (2001). Full Newton method used the complete Hessian (equation 1.16) while the Gauss-Newton method considers the approximated Hessian. These methods have theoretically the quadratic convergences, but they require to perform supplementary forward problems. Therefore, these methods are not generally considered for large size problems, because of their computational costs.

Quasi-Newton optimization considers a cheap estimation of inverse of Hessian operator. This method provides a less expensive way to take into account the benefits of information contains in the Hessian, without paying the price of expensive full Newton or Gauss-Newton methods.

In our implementations, the Hessian is not computed. Instead, we minimize our problem with a bounded quasi-Newton method using the L-BFGS-B (limited-memory Broyden-Fletcher-Goldfarb-Shanno) routine (Nocedal, 1980; Byrd et al., 1995; Nocedal and Wright, 1999). This method provides an approximation of the product of the inverse of the Hessian $(H^{k-1})^{-1}$ by finite differences of a limited number (l) of gradients $(\mathcal{G}^{k-l}, ..., \mathcal{G}^{k-1})$ and model difference vectors coming from previous iterations. The double-loop recursive algorithm designed by Nocedal (1980) does not explicitly build and store $(H^{k-1})^{-1}$, but directly computes the perturbation

vector $\Delta m^k = -(H^{k-1})^{-1}\mathcal{G}^{k-1}$ with additions, differences and inner products of vectors. A diagonal approximation of the inverse Hessian computed from the diagonal terms of the approximated Hessian (Pratt et al., 1998; Operto et al., 2006) or pseudo-Hessian (Shin et al., 2001), can be provided to the L-BFGS algorithm for a better and faster estimation of $(H^{k-1})^{-1}\mathcal{G}^{k-1}$ (Brossier, 2011).

This bounded limited-memory quasi-Newton method, by considering an approximation of the non-diagonal Hessian terms, is an efficient alternative to preconditioned steepest-descent based only on gradients and/or approximate diagonal Hessian approaches. This cheap and efficient estimation of the influence of the inverse Hessian in the optimization improves focusing, partially corrects the descent direction from effects due to limited aperture illumination and frequency bandwidth and respects dimensionalities of the different parameter values (Brossier et al., 2009).

1.1.2.1 Gradient computation

The adjoint-state method is a general method to compute the gradient of a functional that depends on a set of state variables, which are solutions of forward equations (Plessix, 2006). For computing directly the Fréchet derivatives by finite-difference method, we need to perform the forward modeling for each model parameter, which is so costly. The adjoint-state method can deliver an efficient computation of the gradient without implicitly performing the Fréchet derivatives. The theory of adjoint-state method in inverse problems was introduced by Chavent (1974) in order to efficiently compute the gradient of the misfit function without computing the Fréchet derivatives. This method was widely used within numerical community. In geophysics, the gradient of the misfit function was also computed by adjoint-state method (Lailly, 1983a; Chavent and Jacewitz, 1995; Tromp et al., 2005; Plessix, 2006; Chavent, 2009). Developing the adjoint-state method in time domain is slightly more complicated than the frequency domain because of the initial boundary conditions (Plessix, 2006). However, at the end, interpretations of the gradient in both domains are similar.

The pressure wavefield $u(\mathbf{x},t)$ satisfies equation 1.1 with the initial boundary conditions $u(\mathbf{x},t=0)=0$ and $\partial_t u(\mathbf{x},t=0)=0$. Equation 1.1 is used for the imaging formulation and optimisation framework.

Before deriving the adjoint problem, we should define a real scalar product of two real functions f and g, over the Ω spatial domain

$$< f|g >_D = \int_T \int_\Omega dt d\mathbf{x}\, f(\mathbf{x},t) g(\mathbf{x},t), \tag{1.18}$$

for a time window of observation going upto time T (recording time). Let us repeat that the pressure wavefield in time domain has real values.

A misfit function $\mathcal{C}$ in data space can be defined,

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\sum_r \int_0^T dt\, (d_r - R_r u(\mathbf{x},t))^2, \tag{1.19}$$

where the operator R_r projects the pressure wavefield u onto the receiver positions and d_r denotes the obsereved data at receiver positions. Summation over r denotes a summation over

all the receivers. In our acoustic inversion, the P-wave velocity is reconstructed, therefore the model parameter is considered as P-wave velocity ($m = V_p(\mathbf{x})$).

To compute the gradient of misfit function, a new functional called the Lagrangian $\mathcal{L}$ has to be introduced. The Lagrangian function corresponds to the misfit function, subject to the constraint that the state equation is satisfied. We associate the adjoint state μ^0 and μ^1 with considering the initial boundary conditions $u(t=0) = 0$ and $\partial_t u(t=0) = 0$ (Plessix, 2006), and λ with the wave equation. The Lagrangian function is defined by

$$\begin{aligned}\mathcal{L}(m,u,\lambda,\mu^0,\mu^1) &= \frac{1}{2}\sum_r \int_0^T dt\ (d_r - R_r u(\mathbf{x},t))^2 \\ &+ \int_0^T \int_\Omega dt d\mathbf{x}\ \lambda(\mathbf{x},t)\left(\frac{1}{\rho(\mathbf{x})m^2}\frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} - \nabla.\big(\frac{1}{\rho(\mathbf{x})}\nabla u(\mathbf{x},t)\big) - s(\mathbf{x},t)\right) \\ &+ \int_\Omega d\mathbf{x}\ \mu^0(\mathbf{x})u(\mathbf{x},t=0) + \int_\Omega d\mathbf{x}\ \mu^1(\mathbf{x})\partial_t u(\mathbf{x},t=0). \end{aligned} \tag{1.20}$$

The goal is to minimize the Lagrangian function (equation 1.20). The minimum must satisfy conditions 1.21:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \lambda} &= 0 \\ \frac{\partial \mathcal{L}}{\partial u} &= 0 \\ \frac{\partial \mathcal{L}}{\partial m} &= 0. \end{aligned} \tag{1.21}$$

The derivative of $\mathcal{L}$ with respect to the vector λ should be equal to zero at the minimum and this will provide the wave equation while equating the derivative with respect to the vector u to zero will provide the partial differential equations verified by the vector λ. For computing scalar product, we should consider the initial boundary condition ($u(t=0) = 0$ and $\partial_t u(t=0) = 0$). We can find the derivative with respect to the vector u after two integrations by parts for the second right-hand side term of the Lagrangian equation and by evaluating them at the solution of wave equation

$$\begin{aligned}\partial_u \mathcal{L} &= -\sum_r \int_0^T dt\ (d_r - R_r u(\mathbf{x},t)) \\ &+ \int_0^T \int_\Omega dt d\mathbf{x}\ \left(\frac{1}{\rho(\mathbf{x})m^2}\frac{\partial^2 \lambda(\mathbf{x},t)}{\partial t^2} - \nabla.\big(\frac{1}{\rho(\mathbf{x})}\nabla \lambda(\mathbf{x},t)\big)\right) \\ &+ \int_\Omega d\mathbf{x}\ \lambda(\mathbf{x},T)\frac{1}{\rho(\mathbf{x})m^2}\partial_t u(\mathbf{x},T) - \int_\Omega d\mathbf{x}\ \lambda(\mathbf{x},0)\frac{1}{\rho(\mathbf{x})m^2}\partial_t u(\mathbf{x},0) \\ &- \int_\Omega d\mathbf{x}\ \partial_t\lambda(\mathbf{x},T)\frac{1}{\rho(\mathbf{x})m^2}u(\mathbf{x},T) + \int_\Omega d\mathbf{x}\ \partial_t\lambda(\mathbf{x},0)\frac{1}{\rho(\mathbf{x})m^2}\partial_t u(\mathbf{x},0) \\ &= 0. \end{aligned} \tag{1.22}$$

We ignore the adjoint-state related to μ^0 and μ^1, because they do not play a role in the gradient of misfit function. Four last terms of equation 1.22 can give us the boundary conditions for λ adjoint wavefield. Since $u(t=0)$ and $\partial_t u(t=0)$ are equal to zero and $u(t=T)$, $\partial_t u(t=T)$

are non-zero, therefore $\lambda(t=T)$ and $\partial_t\lambda(t=T)$ must be equal to zero. So, two first terms of equation 1.22 can be written as

$$\begin{aligned}\partial_u\mathcal{L} &= -\int_0^T\int_\Omega dtd\mathbf{x}\sum_r\delta(\mathbf{x}-\mathbf{x}_r)(d_r-R_ru(\mathbf{x},t))\\ &+ \int_0^T\int_\Omega dtd\mathbf{x}\left(\frac{1}{\rho(\mathbf{x})m^2}\frac{\partial^2\lambda(\mathbf{x},t)}{\partial t^2}-\nabla.\big(\frac{1}{\rho(\mathbf{x})}\nabla\lambda(\mathbf{x},t)\big)\right)=0,\end{aligned}\tag{1.23}$$

which gives us the local partial differential equation

$$\frac{1}{\rho(\mathbf{x})m^2}\frac{\partial^2\lambda(\mathbf{x},t)}{\partial t^2}-\nabla.\big(\frac{1}{\rho(\mathbf{x})}\nabla\lambda(\mathbf{x},t)\big)=\sum_r(d_r-R_ru(\mathbf{x},t)).\tag{1.24}$$

The system 1.24 has final boundary conditions $\lambda(t=T)=0$ and $\partial_t\lambda(t=T)=0$. We apply the change of variable in the time axis $t'=T-t$ which gives us

$$\begin{aligned}\frac{1}{\rho(\mathbf{x})m^2}\frac{\partial^2\lambda(\mathbf{x},t')}{\partial t'^2}-\nabla.\big(\frac{1}{\rho(\mathbf{x})}\nabla\lambda(\mathbf{x},t')\big) &= \sum_r(d_r(T-t')-R_ru(\mathbf{x},T-t')),\\ \lambda(t'=0) &= 0,\\ \partial_{t'}\lambda(t'=0) &= 0.\end{aligned}\tag{1.25}$$

λ satisfies the wave equation with a new source term. This system is solved by our forward modeling (system 1.2) to obtain the adjoint wavefield. The source term of the adjoint wavefield is the residual between the observed and the synthetic data, in reverse time. Equation 1.25 back propagates the residual into the earth starting from the final time.

The last derivative of Lagrangian with respect to the model parameter m will provide the gradient of the misfit function. We have the expression

$$\partial_m\mathcal{L}=\partial_mC+\int_0^T\int_\Omega dtd\mathbf{x}\;\lambda(\mathbf{x},t)\left(\frac{-2}{\rho(\mathbf{x})m^3}\right)\frac{\partial^2u(\mathbf{x},t)}{\partial t^2}=0,\tag{1.26}$$

which gives us the gradient $\mathcal{G}$ of the misfit function as

$$\mathcal{G}_m=\partial_mC = -\int_0^T\int_\Omega dtd\mathbf{x}\;\lambda(\mathbf{x},t)\left(\frac{-2}{\rho(\mathbf{x})m^3}\right)\frac{\partial^2u(\mathbf{x},t)}{\partial t^2}.\tag{1.27}$$

The term $-2/(\rho(\mathbf{x})m^3)$ can be written as $\partial M(m)/\partial m_i$ as a matricial form, where matrix $M(m)$ corresponds to the matrix of physical model parameters.

Based on equation 1.27, the gradient can be computed by performing only two forward problems. The gradient will be calculated by correlation between second-time derivative of the incident wavefield u and back-propagated residual wavefield λ, taking into account diffraction kernel or pattern $\partial M/\partial m_i$. The matrix $\partial M/\partial m_i$ is the partial derivative of model matrix with respect to model parameter. The matrix $\partial M/\partial m_i$ is extremely sparse. Physically it is the signature of a diffraction coming from a perturbation point localized on a spatial position in model m_i. The radiation patterns of model parameters generally show the sensitivity of FWI to reconstruct the parameters (especially for better parametrization) (Wu and Aki, 1985; Tarantola, 1986; Ribodetti and Virieux, 1996; Forgues and Lambaré, 1997; Gholami et al., 2013b).

In time domain, the residual wavefield is back-propagated by reversing the time. In the frequency domain, the back propagation is indicated by the conjugate operator. The principle of back propagation is similar to reverse time migration (RTM) method (McMechan, 1989). The equation 1.27 shows that only two forward modeling is necessary for gradient computation: one is the computation of the incident wavefield and another one is for back propagating the residuals.

In our application, we consider the acoustic wave equation and inversion is done in time domain (Lailly, 1983a; Tarantola, 1984a; Mora, 1988). Therefore, the gradient in time domain with multiple sources (ns number) can be sum up over the number of sources.

In the following, we introduce another misfit function by including the prior model information into classical FWI scheme.

1.2 Regularized seismic full waveform inversion with prior model information

Regularized seismic full waveform inversion with prior model information
Amir Asnaashari, Romain Brossier, Stéphane Garambois, François Audebert, Pierre Thore, and Jean Virieux
2013, *Geophysics*, 78(2), R25-R36, *doi: 10.1190/geo2012-0104.1.*

1.2.1 Abstract

Full Waveform Inversion (FWI) delivers high-resolution quantitative images and is a promising technique to obtain macro-scale physical properties model of the subsurface. In most geophysical applications, prior information, as those collected in wells, is available and should be used to increase the image reliability. For this, we propose to introduce three terms in the definition of the FWI misfit function: the data misfit itself, the first-order Tikhonov regularization term acting as a smoothing operator and a prior model norm term. This last term is the way to introduce smoothly prior information into the FWI workflow. On a selected target of the Marmousi synthetic example, we show the significant improvement obtained when using the prior model term for both noise-free and noisy synthetic data. We illustrate that the prior model term may significantly reduce the inversion sensitivity to incorrect initial conditions. It is highlighted how the limited range of spatial wavenumber sampling by the acquisition may be compensated with the prior model information, for both multiple-free and multiple-contaminated data. We also demonstrate that prior and initial models play different roles in the inversion scheme. The starting model is used for wave propagation and therefore drives the data-misfit gradient, while the prior model is never explicitly used for solving the wave equation and only drives the optimization step as an additional constraint to minimize the total objective function. Thus the prior model in not required to follow kinematic properties as precisely as the initial model, except in poor illumination zones. In addition, we investigate the influence of a simple dynamic decreasing weighting of the prior model term. Once the cycle-skipping problem has been solved, the impact of the prior model term is gradually reduced within the misfit function in order to be driven by seismic-data only.

Keywords

Prior model, Optimization, High resolution, Wells, Illumination

1.2.2 Introduction

Robust reservoir characterization is a key issue for oil and gas exploration and production. The seismic processing workflow can be roughly summarized in three main steps: velocity model building, migration in time or in depth and elastic properties characterization through amplitude variation-with-offset (AVO) or amplitude variation-with-angle (AVA) analysis. The velocity model building remains a key step that can be tackled with reflection/refraction tomography in time and/or depth domain. A recent alternative for velocity model building is the full waveform inversion (FWI) that allows to reconstruct high-resolution velocity models of the subsurface through the extraction of the full information content of the seismic data (Tarantola, 1984b; Virieux and Operto, 2009).

FWI is a multiscale data-fitting method well adapted to wide-angle/wide-azimuth acquisition geometries, as it uses simultaneously diving and reflected waves. FWI is classically solved with local optimization schemes and therefore strongly dependent on the starting model definition. This starting model should predict arrival times with errors less than half of the period to cancel the cycle-skipping ambiguity (Virieux and Operto, 2009). The multiscale strategy performed by moving from low to high frequencies during the inversion allows to reduce the non-linearities and cycle-skipping issues of the inversion and helps convergence towards the global minimum. Recent applications of FWI to real data have shown promising results for exploration projects: see 3D examples in Plessix and Perkins (2010) or Sirgue et al. (2010). Monoparameter reconstruction of the acoustic velocity for exploration is quite impressive even in the anisotropic case (Prieux et al., 2011). Elastic parameters could also be recovered for exploration targets (Brossier et al., 2009; Prieux et al., 2012), but elastic inversion applies rather to seismological scales where phases are nicely separated (Fichtner et al., 2010; Tape et al., 2010).

Preconditioning or regularization techniques may alleviate the non-uniqueness of the ill-posed inverse problem. Tikhonov and Arsenin (1977) have proposed a regularization strategy, within the optimization step, to find the smoothest model that explains the data. Preconditioning techniques acting as a smooth operator on the model update (Operto et al., 2006) may add strong prior features of the expected structure through directive Laplacian preconditioning, such as in (Guitton et al., 2010). Regularization schemes that preserve edges and contrasts have also been developed for specific FWI applications through an ℓ_1 model penalty (Guitton, 2011) or through a multiplicative regularization (Abubakar et al., 2009) that mimics the Total Variation scheme (Rudin et al., 1992). Regularization can also be expressed in the curvelet or wavelet domains (Loris et al., 2010; Herrmann et al., 2009). In such domains, the ℓ_1 norm minimization is generally preferred for the model term penalty as it ensures sparsity in the model space.

All the previous regularization techniques allow to stabilize the inversion scheme by assuming a particular representation or structure of the velocity model (smoothness, sparsity and so on). However, prior model information is generally not used in classical FWI implementation even if Hu et al. (2009) recently suggested to use a prior model in the multiplicative regularization term. Several sources of prior model information are usually available at the exploration

stages, such as sonic logs, exploration well data or geological information of the field. One may want to use such prior information in the FWI scheme as is done in other velocity building techniques. Taking into account the prior information could also be highly important for monitoring purposes, where many different and precise prior data types have been collected for the target zone. Prior information can be introduced through the generalized Tikhonov regularization using the Bayesian formulation (Greenhalgh et al., 2006; Mead and Renaut, 2009) where the prior model is related to the expected model for the Bayesian interpretation. Strict Tikhonov regularization can be recast into this formalism as well. However, combining strict Tikhonov regularization and generalized Tikhonov regularizations may lead to difficulties in defining the respective weights of the different information: prior information and expected smoothness of the model.

Several studies have been done on using two model penalty terms in geophysical electromagnetic applications, such as for the inversion of magnetic stripe data (Farquharson and Oldenburg, 1998) and for the inversion of controlled source audio-frequency magnetotellurics data to recover a 1D conductivity structure (Routh and Oldenburg, 1999). In this study, we investigate the performances of a FWI scheme based on two model penalty terms in the misfit definition in addition to the data term: the Tikhonov term to ensure smoothness, and a prior model term to drive the inversion in a given direction. In the first part, we present the theoretical framework of our study. Then, through a synthetic application on the Marmousi model, we show the critical effect of the prior model penalty term on the FWI results. We shall highlight how the limited range of wavenumber sampling coming from the limited frequency band and the acquisition geometry may be compensated with the prior model information, for both surface multiple-free data and also data containing surface multiples. We shall underline the fundamentally different role of the prior model and of the starting model within the FWI procedure.

1.2.3 Theory

Full Waveform Inversion relies on an iterative local optimization problem that is generally introduced as a linearized least-squares problem. The optimization attempts to minimize the residuals between the observed and the modeled wavefields at the receivers. The linearized inverse problem remains ill-posed, and therefore multiple model solutions can provide a satisfactory fit of the observed data. Prior information is generally introduced through regularization in the inverse formalism. However, for specific applications where other sources of information such as sonic logs, stratigraphic data or geological constraints are available, it is crucial to take these into account in the inversion process and incorporate them into a prior model, to ensure robust and consistent results.

To do so, we briefly introduce the full waveform inversion algorithm including the model norm contribution.

The general definition of the misfit function for solving ill-posed inverse problems could be recast as the Tikhonov function (Tikhonov and Arsenin, 1977):

$$\mathcal{C}(\mathbf{m}) = \mathcal{C}_d(\mathbf{m}) + \lambda \mathcal{C}_m(\mathbf{m}). \tag{1.28}$$

The data misfit $\mathcal{C}_d(\mathbf{m})$ is based on a norm of the residuals between observed and computed data in the data space, and the model norm $\mathcal{C}_m(\mathbf{m})$ term is based on a norm of a model penalty

function in the model space. In the standard Tikhonov approach, this penalty function is based on the first spatial derivative of the current model that should have a minimal norm, thus giving a smooth model. The hyper-parameter λ is the regularization parameter, also called trade-off parameter, that balances contributions between the data and the model terms.

For applications where prior information on the model can be established, we add a second penalty term to the misfit function. This term estimates residuals between the current model at a given iteration and the prior model considered at that same iteration. The objective function can now be written as the following expression,

$$\mathcal{C}(\mathbf{m}) = \mathcal{C}_d(\mathbf{m}) + \lambda_1 \mathcal{C}_{1_m}(\mathbf{m}) + \lambda_2 \mathcal{C}_{2_m}(\mathbf{m}), \tag{1.29}$$

where the Tikhonov term is denoted by $\mathcal{C}_{1_m}(\mathbf{m})$ and the prior model misfit term by $\mathcal{C}_{2_m}(\mathbf{m})$. Two regularization hyper-parameters λ_1 and λ_2 are introduced, to allow weighting of the penalty terms with respect to each other and to the data term.

Let us express these three terms in a more explicit way using ℓ_2 norms. The data term may be written as

$$\mathcal{C}_d(\mathbf{m}) = \sum_{ns} ||\mathbf{W}_d \left(\mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})\right)||^2 = \sum_{ns} \frac{1}{2} \Big\{ \left(\mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})\right)^T \mathbf{W}_d^T \mathbf{W}_d \left(\mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})\right) \Big\}, \tag{1.30}$$

where $\mathbf{d}_{obs}$ and $\mathbf{d}(\mathbf{m})$ are vectors for the observed and computed data respectively. For our specific investigation we consider a time-domain approach, and each component of these vectors are samples of time-domain seismograms recorded at receiver positions for one seismic source. This misfit function results from a sum over the ns sources of the experiment. The matrix $\mathbf{W}_d$ is a weighting operator on the data. This matrix can also be seen as the inverse of the square-root of the covariance matrix of the data, which contains information on data uncertainties. Considering a constant measurement quality and uncorrelated traces, we end up with a diagonal matrix of $\mathbf{W}_d = \sigma_d I$, where σ_d is the standard deviation of the data and I is the identity matrix (Tarantola, 2005). The synthetic data $\mathbf{d}(\mathbf{m})$ non-linearly depend on the model parameters denoted by $\mathbf{m} = \{m_i\}_{i=1,N_m}$, where N_m is the number of unknows. These model parameters should be determined through the inverse procedure by reducing this data term.

The second term of the misfit function is the Tikhonov term and can be written as

$$\mathcal{C}_{1_m}(\mathbf{m}) = ||\mathbf{B}_x \mathbf{m}||^2 + ||\mathbf{B}_z \mathbf{m}||^2 = \frac{1}{2}\{\mathbf{m}^T {\mathbf{B}_x}^T \mathbf{B}_x \mathbf{m} + \mathbf{m}^T {\mathbf{B}_z}^T \mathbf{B}_z \mathbf{m}\} = \frac{1}{2}\{\mathbf{m}^T \mathbf{D} \mathbf{m}\}, \tag{1.31}$$

where $\mathbf{B}_x$ and $\mathbf{B}_z$ are the first-order spatial derivative operator matrices with respect to x and z, respectively. In practice, they can be reduced to the second-order Laplacian operator $\mathbf{D}$. We use a classical five-point finite-difference stencil to implement the operator $\mathbf{D}$.

The third term of the objective function is related to the prior model $\mathbf{m}_p$, which can be designed from different information and could be set prior to the seismic inversion, but which could be also adapted iteratively during the inversion procedure. This so-called prior model norm term is computed using the expression

$$\mathcal{C}_{2_m}(\mathbf{m}) = ||\mathbf{W}_m \left(\mathbf{m} - \mathbf{m}_p\right)||^2 = \frac{1}{2} \Big\{ \left(\mathbf{m} - \mathbf{m}_p\right)^T \mathbf{W}_m^T \mathbf{W}_m \left(\mathbf{m} - \mathbf{m}_p\right) \Big\}, \tag{1.32}$$

where the matrix $\mathbf{W}_m$ is a weighting operator on the model space. This matrix can also be seen as the inverse of the square-root of the covariance matrix of the model, and contains prior

uncertainty information of the prior model parameters. In our implementation, since we want to separate the influence of the diagonal and off-diagonal terms of the covariance matrix, we choose a diagonal $\mathbf{W}_m$ matrix, $diag(\mathbf{W}_m^T\mathbf{W}_m) = 1/\sigma^2(\mathbf{m})$. The prior weighting model $\sigma^2(\mathbf{m})$ contains both the prior model uncertainty (variance) and the potential weighting function, and will be discussed in the application section. The covariances (off-diagonal terms) are implicitly taken into account through the Tikhonov term.

Does the $\mathbf{W}_m$ operator play a critical role in driving the inversion procedure towards a given minimum? This is a question we want to investigate. Note that the misfit function, mixing both data and model quantities, is dimensionless due to the introduction of the matrices $\mathbf{W}_d$ and $\mathbf{W}_m$, and through the hyper-parameter λ_1 dimension. In order to have three dimensionless terms in the sum, the hyper-parameter λ_1 has a dimension $[dim(h^2)/dim^2(\mathbf{m})]$, due to the dimensionality of the $\mathbf{D}$ operator, where the grid size h is for a 2D square regular cartesian grid. For a model described by velocity, the dimension of λ_1 is second squared $(m^2/(m/s)^2 = s^2)$.

Minimizing the misfit function classically leads to the normal equation system which can be written as

$$\mathcal{H}_m \Delta\mathbf{m} = -\mathcal{G}_m, \tag{1.33}$$

where the gradient and the Hessian of the misfit function are denoted $\mathcal{G}_m$ and $\mathcal{H}_m$ respectively. The gradient expression can be written with three components as

$$\mathcal{G}_m = -\sum_{ns}\left(\frac{\partial\mathbf{d}(\mathbf{m})}{\partial\mathbf{m}}\right)^T \mathbf{W}_d^T\mathbf{W}_d\left(\mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})\right) + \lambda_1\mathbf{D}\mathbf{m} + \lambda_2\mathbf{W}_m^T\mathbf{W}_m\left(\mathbf{m} - \mathbf{m}_p\right). \tag{1.34}$$

The sensitivity matrix $\mathbf{J} = \partial\mathbf{d}(\mathbf{m})/\partial\mathbf{m}$ is composed by the Fréchet derivatives of the synthetic data with respect to the model parameters. The data-term gradient is efficiently computed with an adjoint formulation (Plessix, 2006) without an explicit computation of the matrix $\mathbf{J}$. The two terms related to the model penalties are generally straightforward to compute and are simply added to the data-term gradient contribution, leading to negligible computer memory and CPU-time increase.

The Hessian matrix is based on the second derivative of the misfit function and is not computed in our implementation. Instead, we minimize our problem with a bounded quasi-Newton method using the L-BFGS-B routine (Byrd et al., 1995). This routine allows to take into account an approximate non-diagonal inverse Hessian from previous gradient and model vectors, and performs a line-search satisfying Wolfe's conditions. This bounded limited-memory quasi-Newton method is an efficient alternative to preconditioned steepest-descent or conjugate-gradient methods based only on gradients and/or approximate diagonal Hessian approaches. This cheap and efficient estimation of the influence of the inverse Hessian in the optimization improves focusing, partially corrects the descent direction from effects due to limited aperture illumination and frequency bandwidth and respects dimensionalities of the different parameter values (Brossier et al., 2009).

A major point for real data applications is the source-wavelet estimation. Our FWI is implemented in the time-domain for both the forward and the inverse problem. The source-wavelet estimation is however straightforwardly implemented in the frequency-domain by a linear inverse problem resolution. The computed and observed time-domain data are Fourier transformed to apply the Pratt (1999) (his equation 17) source estimation equation for each frequency. The Fourier coefficients of the wavelet are then transformed back to the time-domain and appropriately processed (anti-causal mute and/or band-pass filtering if required)

before performing FWI. This estimation is performed once before the optimization. In the following tests applied to synthetics, we use the exact source wavelet for fair comparisons, such that the results are not biased by potential errors from this estimation.

1.2.4 Application to Marmousi model

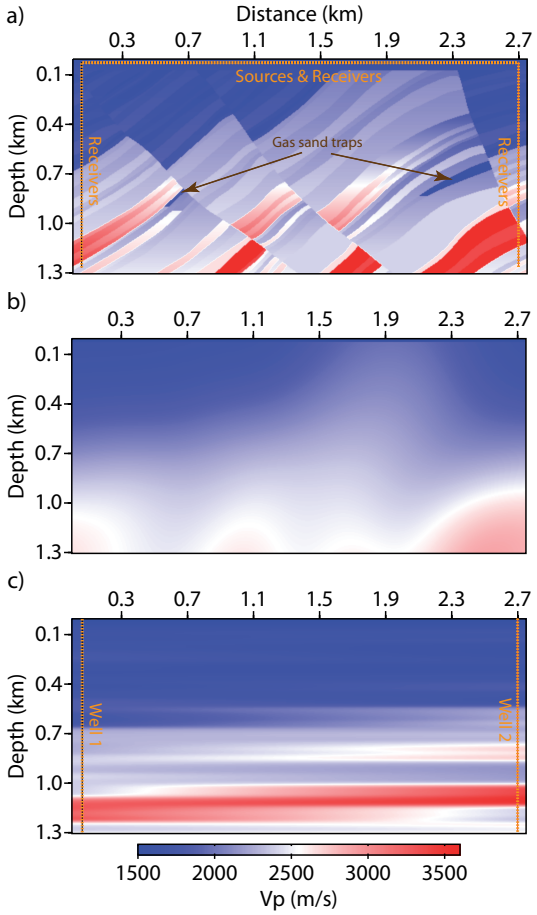


Figure 1.3: (a) The true V_p velocity model which is a small part of the Marmousi model and the acquisition geometry; (b) initial model for inversion which is a smooth model of the true model; (c) the prior model created by linear distance weighted interpolation in the x direction between the exact values inside two exploration wells and then gently smoothed.

In this section, we study the effect of prior information in FWI. In particular, we show how prior information allows to mitigate the lack of seismic illumination. A selected target zone of the Marmousi II P-wave velocity distribution (Martin et al., 2006) and a homogeneous density model are considered. The target exhibits two gas sand traps (Figure 1.3a). We consider a shallow-water configuration with a water depth of only 25 m. Our acquisition geometry contains 54 isotropic pressure-sources, located along a horizontal line at 15 m depth, every 50 m. The layout is the same for all shots, one fixed horizontal receiver line at 15 m depth and two fixed vertical lines of receivers inside two exploration wells at $x = 50$ m and $x = 2700$ m with a 10 m interval between sensors. The deepest receivers inside the wells are at $z = 1265$ m.

The grid is regular, with the grid size equal to 5 m, and it is consistent for both modeling and inversion. We do not consider any sources within the wells as this design is unusual and quite expensive. However we consider sensors inside the wells, which could be installed for exploration or monitoring purposes, and allowing to dramatically increase the illumination for velocity reconstruction. Note that our final test will be performed without these well sensors, to mimic a pure surface acquisition. A Ricker wavelet source with a central frequency of 10 Hz is used for all shots. The time seismograms are generated using finite-difference modeling in the time-domain with a fourth-order stencil in space and a second-order integration in time. Perfectly-Matching-Layer (PML) absorbing boundary conditions (Berenger, 1994) are used for non-reflecting boundaries. The first tests are performed using a PML on top, in order to mimic multiple-free data. The last test will consider a free-surface condition, modeling surface-multiples. The recorded pressure data are used as observed data, both at the surface and in wells. Figure 1.4a shows an example of a seismogram generated by a shot located at the center of the source line.

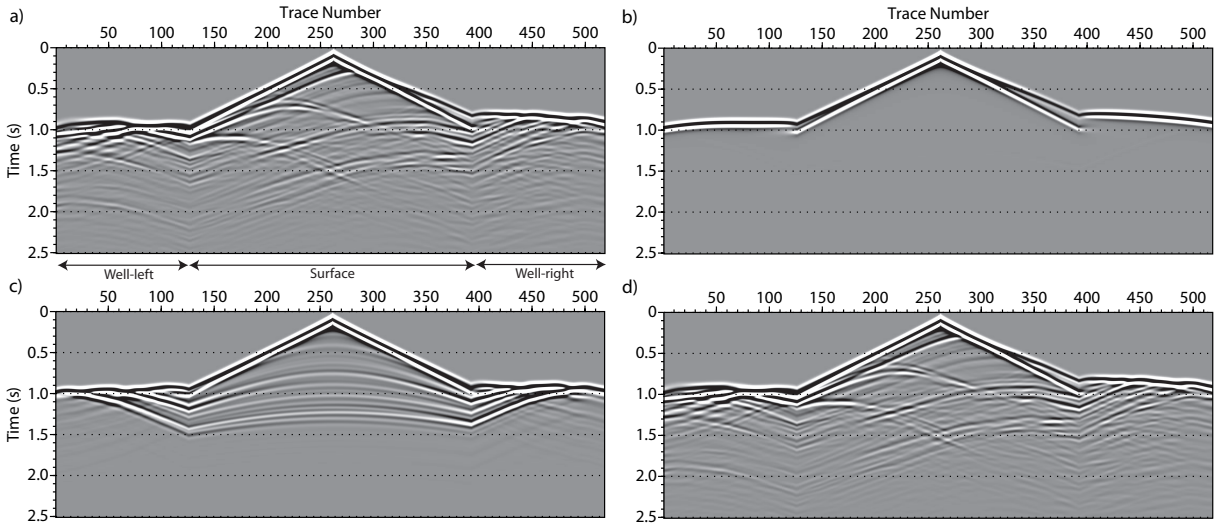


Figure 1.4: Seismograms of pressure data for the source located almost at the center of the Marmousi model $x = 1.4\ km$: (a) recorded inside the true model, (b) calculated inside the smooth initial model, (c) calculated inside the interpolated velocity model, and (d) computed inside the final model obtained through our dynamic approach.

In our study, the data weighting matrix $\mathbf{W}_d$ is chosen as identity $\mathbf{W}_d = \mathbf{I}\ [dim(data)]^{-1}$, where $dim(data)$ means the unit of pressure data. In order to have a dimensionless objective function, $\mathbf{W}_d$ should have a dimension which is the inverse of the data dimension. Note that for all further applications, the Tikhonov regularization parameter is kept fixed to a small value, imposing only a weak smoothing constraint, since we mainly focus on analyzing of the effects of the prior penalty term.

A smooth velocity model (Figure 1.3b), which mimics a time-tomography velocity model based on both first arrivals and reflected events, and referred to henceforth as *"smoothed velocity model"*, is used as the initial model for FWI. A time-domain FWI approach is used, involving all the frequencies of the spectrum (maximum 30 Hz in this case). No additional hierarchical approach such as the frequency-continuation approach of Bunks et al. (1995) is used in these

examples. This means that the weighting of each frequency is directly link to its amplitude in the spectrum. A first investigation (Figure 1.6a) is performed with noise-free data and a standard regularized FWI method, without considering a prior model (equivalent to $\lambda_2 = 0$). The result shows that the optimization is trapped in a local minimum. This issue can be related to cycle-skipping ambiguities due to the starting model inaccuracy, especially in the deepest part below 700 m and on the left part of the model until the second fault. Due to these inaccuracies, the target zones composed of the two reservoir areas are not well recovered with this configuration.

1.2.4.1 FWI with prior model and impact of prior weighting matrix ($\mathbf{W}_m$)

In our framework where well information does exist, the FWI method should use this non-seismic information as prior information for the inversion. We first need to build the prior velocity model $\mathbf{m}_p$ and the model weighting matrix $\mathbf{W}_m$ that contains the prior model uncertainty. In our study, we consider that the sonic-log measurements acquired in the two exploration wells provide an accurate estimation of the local vertical velocity. A prior model could have been created from interpolation of the well velocity, following picked horizons in the migrated section. Instead, we build a crude prior velocity model based on a linear interpolation between the two well locations without any migration and picking approach. This interpolated model (Figure 1.3c) from only the well data, henceforth called *"interpolated velocity model"*, even though being far from the true 2D structure of the Marmousi model, will be used as a prior velocity model for regularized inversion. As shown in the following test, this crude prior model allows to significantly help the inversion to converge, and when applied to real data, the more accurate the prior is, for example if it is derived using standard quantitative interpretation techniques, the better the FWI results will be.

The prior model has to be associated to the weighting matrix $\mathbf{W}_m$, in order to weight the penalty associated to the model residual $(\mathbf{m} - \mathbf{m}_p)$. As already mentioned, we use a diagonal weighting matrix containing both the uncertainty and some weighting. From how the prior model is built, we know that, quantitatively, the interpolated velocity values should be accurate close to the well positions, but they can be erroneous far from the wells, since the structure is highly heterogeneous. Therefore, we decided to build a weighting shape whose uncertainty values follow a Gaussian function with weak values near the wells and increasing values in the center of the area (Figure 1.5a). This is the prior weighting model A.

A key point in all additive regularized optimization schemes is the selection of the weighting hyper-parameters. As already mentioned, the λ_1 value chosen is small enough to ensure a slight smoothing of the results. In practice, to select the λ_2 hyper-parameter, we compute the misfit function for the starting model for $\lambda_2 = 1$. Based on the ratio γ between the prior-model misfit $\lambda_2 \mathcal{C}_{2_m}(\mathbf{m})$ and the data-term misfit $\mathcal{C}_d(\mathbf{m})$, we adjust the λ_2 value such that $10^{-3} < \gamma < 10^{-2}$. Therefore, by selecting this reasonable ratio of prior-model and data misfit terms, the FWI is prevented to minimize the model norm heavily at early iterations. In fact, an even stronger weight is applied to the data term in the global objective function. In this test, we choose to have the ratio $\gamma = 10^{-2}$.

Figure 1.6c shows the reconstructed velocity model after FWI, starting from the "smoothed velocity model", and using the "interpolated velocity model" as a prior model. We can see that the shallow left part of the reconstructed model has been strongly improved compared to Figure 1.6a. However, the deeper part of the result remains strongly dominated by the

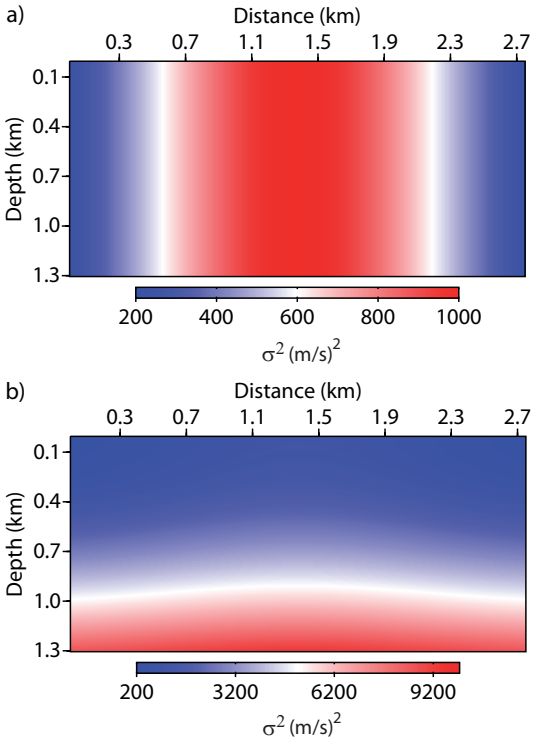


Figure 1.5: Two types of prior weighting model used for regularized inversion: (a) model A, the Gaussian function varying only in the x direction between two wells with maximum value at the center of model, and (b) model B, the same variation in x complemented by a quadratic evolution in the z direction (the Gaussian lateral variation could be seen now in the undulation of the white interface).

footprint of the interpolated velocity model used as a prior model. This footprint can be interpreted as an inappropriate relative weight between the prior penalty term and the data misfit term, for waves that illuminate this deeper part. The consistency of the two terms at shallow depth, leads to an improved reconstruction. In order to visually see the relative amplitude of the different terms of the gradient, the absolute value of the data-term gradient (Figure 1.7a), the prior-model term gradient (Figure 1.7b), and their ratio (prior-model/data) (Figure 1.7c) are computed at the first iteration. With increasing depth, the amplitude of the data-term gradient decreases, because the associated wave amplitudes, mainly in a reflection regime, decrease due to geometrical spreading, intrinsic attenuation and energy partitioning at interfaces. The ratio between the prior-model and the data gradients therefore shows that the deep part of the gradient is driven by the prior-model at the expense of the data term, because of the homogeneous weighting term with depth in the $\mathbf{W}_m$ matrix. To overcome this unfavorable balance between the data and the prior terms in the optimization, a weighting is required and can be implemented in two different ways. We can either modify the $\mathbf{W}_m$ matrix to decrease the weight in depth or change the data-term weighting $\mathbf{W}_d$ matrix such that the late arrivals have more weight in the data misfit and the data-gradient terms. This second weighting can be linked to the metric choice of the misfit function norm for the data term as defined in Jin et al. (1992).

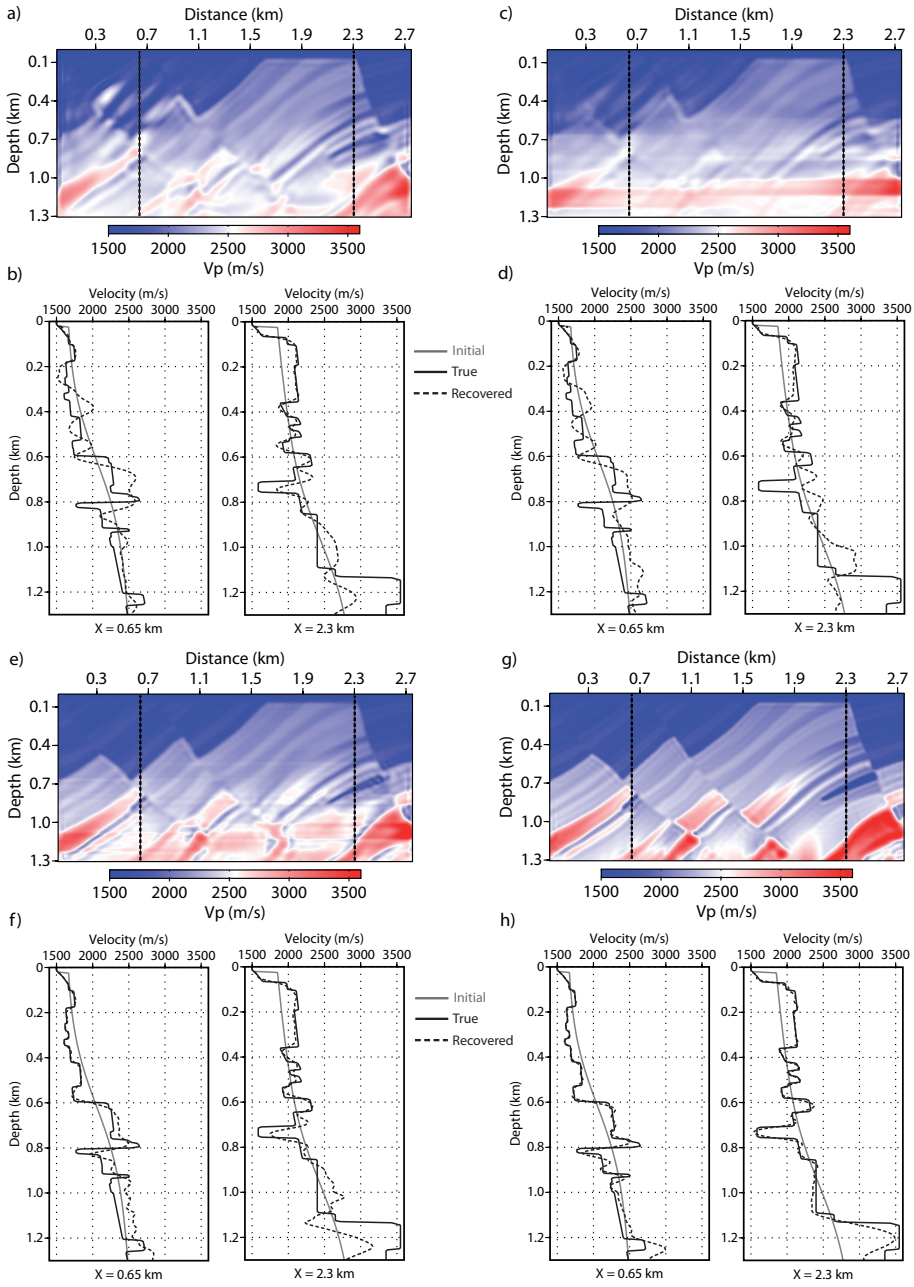


Figure 1.6: The recovered V_p models by FWI and two QC vertical logs passing through the two target areas at $x = 0.65\ km$ and $x = 2.3\ km$, (a) reconstructed model starting from the smooth initial model and *without a prior model*, small λ_1 and $\lambda_2 = 0$; (b) two vertical logs corresponding to the model (a); (c) reconstructed model starting from the smooth initial model and with the prior model, small λ_1, fixed λ_2, the *prior weighting model A* and the ratio between prior-model and data misfit terms $\gamma = 1 \times 10^{-2}$; (d) two vertical logs corresponding to the model (c); (e) reconstructed model starting from the smooth initial model and with the prior model, small λ_1, fixed λ_2 same as case (c), the *prior weighting model B*. Note the γ ratio is now decreased to 3×10^{-3}; (f) two vertical logs corresponding to the model (e); (g) reconstructed model starting from the smooth initial model and with the prior model, small λ_1, initial value of $\gamma = 3 \times 10^{-3}$, the prior weighting model B and using the *dynamic prior approach*. The prior model is removed from the inversion at the end of the procedure; (h) two vertical logs corresponding to the model (g).

In our study, we choose to involve the depth weighting in the model space and we use a rough but efficient approximation of the geometrical spreading to change the $\mathbf{W}_m$ matrix in depth: we propose to make the operator $\mathbf{W}_m^T\mathbf{W}_m$ decrease by a simple $1/z^2$ with respect to the depth z, in order to compensate for the propagating decay of the wave amplitude. This kind of depth weighting has been used in the Controlled Source Electromagnetic method (Plessix and Mulder, 2008) and gravity inversion applications (Li and Oldenburg, 1998). Plessix and Mulder (2008) have proposed the depth weighting matrix to compensate the exponential decay of the amplitude of electromagnetic waves and also geometrical spreading. This depth weighting was used as a preconditioning of the model parameter (Plessix and Mulder, 2008) and to counteract the geometric decay of the kernels in inversion (Li and Oldenburg, 1998). In our application, we use the same general principle, but our main goal is to make an appropriate balance between the prior-model norm and data misfit in depth. We combine this weighting and the uncertainty associated to distance away from the wells to build a new weighting matrix (Figure 1.5b), referred to as prior weighting model B.

FWI is now applied using the "smoothed velocity model" as the initial model and the "interpolated velocity model" as the prior model together with the prior weighting model B. The same hyper-parameter λ_2 is used, but note that the γ ratio value between prior-model and data misfit is decreased to around 3×10^{-3}. This weighting model B allows to successfully balance the gradient energy with the depth as shown in Figure 1.7d and 1.7e. The reconstructed velocity is shown in Figure 1.6e and exhibits a significant improvement compared to Figure 1.6a. This result first shows the importance of the prior weighting, which should contain appropriate uncertainty information, but should also ensure an appropriate balance between the prior misfit term and the data misfit term in the optimization. In this case, the prior term significantly helps the inversion to converge to the global minimum of the optimization problem, mitigating the cycle-skipping issues that the data misfit term cannot handle alone. In fact, adding the prior model penalty allows to change successive descent directions and helps the inversion to converge to the correct global minimum valley of the misfit function. This test shows that prior information allows to constrain inversion and, therefore, mitigates the non-uniqueness issue of ill-posed inverse problems.

1.2.4.2 Roles of initial versus prior models

In many geophysical inversions, it has been proven successful to choose the initial model equal to the prior model (see Oldenburg (1994); Routh and Oldenburg (1999) for electrical/electromagnetic inversion and Routh and Anno (2008); Miller et al. (2008) for time-lapse inversion), when it is chosen sufficiently accurate. In this part, we address the relative role of the prior and initial models in the inversion procedure, when only partial information is contained in the available models. A first natural idea could be the use of the prior velocity model (Figure 1.3c) as the initial model of FWI. Since this model helps the FWI when it is used as prior model, it could be a good candidate for the initial model of the inversion. Fundamental differences exist when using a particular velocity model as a prior model which has no direct impacts on the modeling of synthetic data or when using it as an initial model with direct consequences on the synthetic data. Figure 1.8 illustrates the inversion result derived using the classical regularized FWI (same tuning as Figure 1.6a) and the "interpolated velocity model" as initial model. We can clearly see that the inversion converges towards a local minimum, far from being a satisfactory result. Moreover, in this case the optimization process stops after

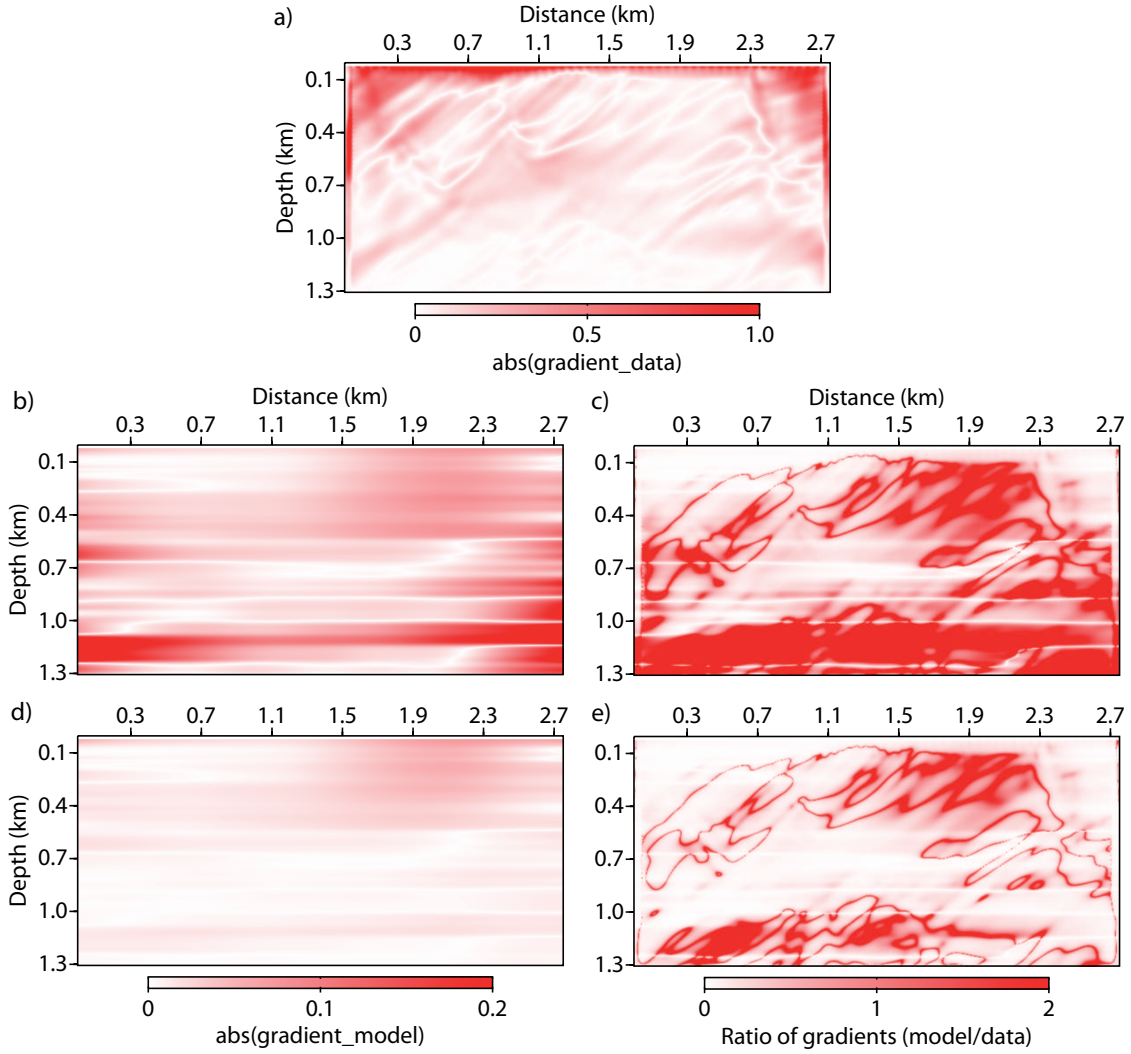


Figure 1.7: At first iteration of optimization, the absolute value of (a) the data-term gradient, (b) the prior-model term gradient , (c) the ratio between prior-model and data gradients, in case of using the prior weighting model A; (d) and (e) same as (b) and (c) respectively, but in case of using the prior weighting model B.

only a few iterations. The shallow part on the right-hand side of the model seems satisfactory but the left-hand side and the deeper parts seem to be badly handled by this initial model, built from interpolation in this strongly laterally-varying structure. One interpretation of this failure is related to the major difference in the meaning of the initial and of the prior model: the initial model must be localized in the attraction valley of the global minimum of the misfit function, often related in seismic as being kinematically accurate and not generating erroneous arrivals in the synthetic data computed using the wave equation (see Figure 1.4a and 1.4c). Indeed, it is much more difficult for the inversion workflow to suppress or shift a structure than to create a new one. On the contrary, the prior model is never explicitly used as an input for solving the wave equation and is only used to drive the optimization step in order to

minimize the total objective function. It can therefore contain any structure, complementary to the information contained in the initial model, that can drive inversion towards expected zones of the model space. In our case, the prior model allows FWI to be driven and partially fills in the lack of low wavenumbers that cannot be extracted from only the data. Both the "smoothed velocity model" and the "interpolated velocity model" contain partial information on the velocity model, that, when used alone, is not sufficient to converge towards the global minimum. Only an appropriate combination of both pieces of information, through the initial and prior models, allows to exploit the partial information included in both, and allows to significantly improve the results. Note however that for regions of poor seismic illumination, as the optimization is driven by the prior model, this model requires to be as accurate as possible (kinematically correct) to ensure good results. In this case, the prior model puts its imprint on the current model and then is partially used for solving the wave equation.

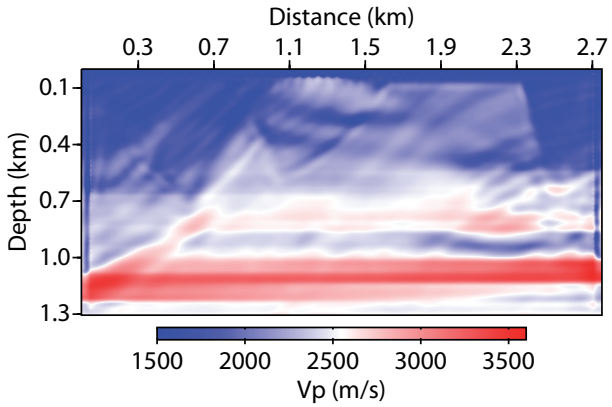


Figure 1.8: The recovered V_p model by standard FWI starting from an initial model equal to the interpolated velocity model, small λ_1 and $\lambda_2 = 0$.

1.2.4.3 Dynamic prior regularization parameter

In complex environments, the prior model derived from extra information on the target zone may be far away from the exact model we never reach. Even if the prior model can significantly improve results by driving the inversion in an appropriate direction, the final model can keep a significant footprint of the prior model structure and may prevent a significant expression of the data itself. As shown in Figure 1.6e, the result exhibits ghost interfaces coming from the interpolated prior model. These footprints of the prior setting do not honor the data itself. However, keeping a fixed hyper-parameter on the prior term of the misfit function prevents the results from being improved since the prior model is intrinsically wrong in such a case.

Thus, one can investigate a dynamic weighting of the prior information, in order to decrease the weight of the prior term (λ_2) during iterations of the optimization. We suggest a simple dynamic approach, considering a starting λ_2 value that is gradually decreasing until it reaches zero. This method allows to drive FWI towards the global minimum valley of the objective function at the beginning, due to the prior-model influence, and to finally leave only the data term to drive the final iterations of the optimization, by gradually decreasing the prior weight. The Tikhonov regularization term is kept constant as we only discuss here the reciprocal influences of the data misfit term and of the prior term. Our heuristic approach is

based on the decrease of the objective function with iterations. When the slope of the objective function curve becomes too small, and smaller than a specific threshold, the current λ_2 value is divided by a factor of two to reinforce the weight of data misfit term. Our criterion is based on the first derivative of the misfit function with respect to iterations, computed with a simple finite-difference stencil. The derivative value at each iteration is normalized by the first derivative value. During the optimization procedure, the corresponding derivative value is compared to the fixed threshold at each iteration. Every time that the derivative is smaller than the threshold, meaning that the misfit function curve with iterations is becoming too flat, the hyper-parameter λ_2 is changed. The key issue of this strategy is the value of the threshold at which the hyper-parameter λ_2 term must be decreased. We find that a few trials can narrow down this value quite rapidly from variation of the misfit function. This threshold value should be of the order of $10^{-4} \sim 10^{-3}$. Note that in our implementation, the L-BFGS-B optimization is restarted each time the hyper-parameter λ_2 is changed.

The dynamic method has a similar property to the multiplicative regularization and cooling regularization approaches (van den Berg et al., 1999; van den Berg et al., 2003). In Total Variation (TV) as the multiplicative constraint, the data objective function itself is defined as the weight of TV. Therefore, the regularization term has a large weighting parameter in the beginning of the optimization process, and gradually decreases as the objective function is minimized and the data fitted.

Figure 1.6g shows the recovered model obtained by this dynamic method, using the "smoothed velocity model" as the starting model, the "interpolated velocity model" as a prior model and the optimal weighting matrix B. As with the fixed λ_2 strategy, we can see that the reconstructed model is dramatically improved when compared to the one obtained by standard FWI. Moreover, the dynamic approach allows to mitigate the footprint of the prior model, since during the optimization the prior penalty weight decreases with respect to the data misfit term. Thus, the effect of the prior model is being reduced slowly and the misfit data term helps inversion to converge to a quasi-perfect final model. As a quality control, vertical profiles taken through the two gas sand traps (for the initial, true and recovered models) illustrate that the target velocity is recovered accurately (Figure 1.6h), compared to the result from standard FWI using the same initial model (Figure 1.6b). In addition, the computed seismogram inside the final model shows that the full seismic arrivals have been exploited during the optimization (see Figure 1.4d).

The objective function curves for the data term, the model norms and the λ_2 curve as a function of iteration are shown in Figure 1.9. In this case, the data are without noise and a very small stopping criterion is selected to fit the data as much as possible, leading to a large number of iterations. This stopping criterion, based on the flatness of the misfit function for two successive iterations, is the same for all the inversion tests, so that the results are comparable. The data objective function always decreases and by reducing λ_2, we try to prevent giving a high weight to the prior penalty term. Therefore, by reducing the model objective function value, an appropriate contribution of the prior model is kept during optimization. Note that there is no change on the small λ_1 value and that the Tikhonov term always exists, leading to a non-zero model objective.

The convergence of standard FWI without any prior model and FWI with dynamic prior weighting shows dramatic differences of the evolution of the total objective function (Figure 1.10). The standard FWI gets trapped in a local minimum and stops the optimization after 87 iterations. By including the prior model to the optimization, the path of descent is changed

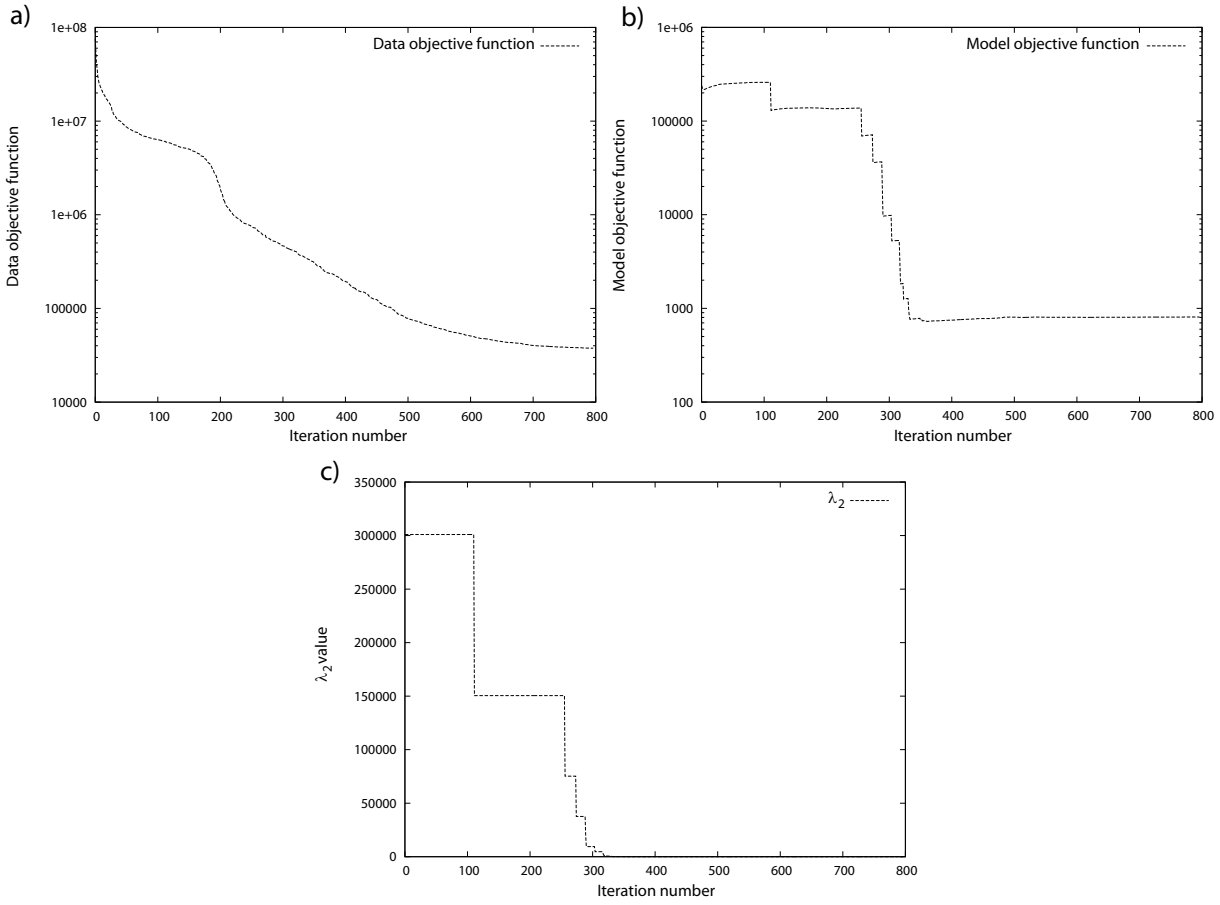


Figure 1.9: Evolution of (a) the data objective function, (b) the model objective function, and (c) λ_2 value with iterations in case of using the dynamic approach. Note that (a) and (b) curves are shown in logarithmic scale.

and the optimization procedure is not trapped by local minimum attraction basins. The beginning of the optimization appears quite equivalent for both approaches until iteration 40, even showing better convergence speed for the standard FWI. Please note however that even if the convergence rate is the same, the results may be different because the null-space is different between an inversion with and without the prior model penalty. After this step, due to different descent paths, standard FWI slows down convergence speed and rapidly stops. For FWI with dynamic prior weight (blue curve), we can observe a large decrease in the objective function between iterations 170 and 230. Looking at the updated model history at these iterations shows a significant improvement, associated to the prior model penalty use, in the shallow left part target, leading to a large decrease in the data misfit. In standard FWI, the data and Tikhonov terms of the misfit gradient alone are not able to solve this problem in the shallow left part of model.

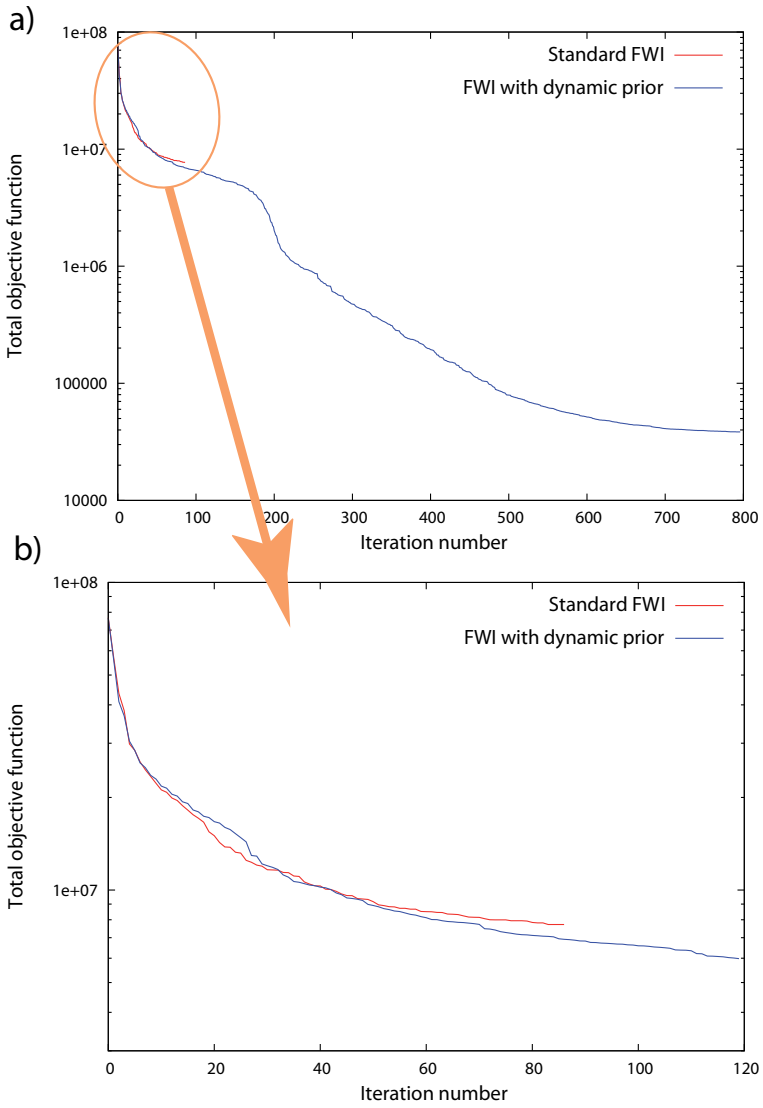


Figure 1.10: Comparison of the total objective function curves in case of using the standard FWI and the dynamic prior weighting FWI (a) for all iterations, (b) shown at early iterations.

1.2.4.4 Noisy data

In presence of noise, the ill-posedness of the inverse problem is increased. Therefore, we need to study the effect of noise on our proposed regularized FWI including the prior model penalty. We keep the same acquisition configuration, while an artificial Gaussian noise in the range of $0-30\ Hz$, the bandwidth of the source, is added to the true noise-free data. The signal-to-noise ratio is around 7 dB. Figure 1.11 illustrates an example of shot gathers used for FWI (we have used the suaddnoise procedure of Seismic Unix (Cohen and Stockwell, 2008)).

Three inversion tests starting from the "smoothed velocity model" (Figure 1.3b) are performed. The first one uses the standard FWI without the prior model (Figure 1.12a). The second test uses the "interpolated velocity model" as prior model with a fixed λ_2 value (Figure 1.12c), and the third one uses the dynamic prior weighting (Figure 1.12e). All the parameters

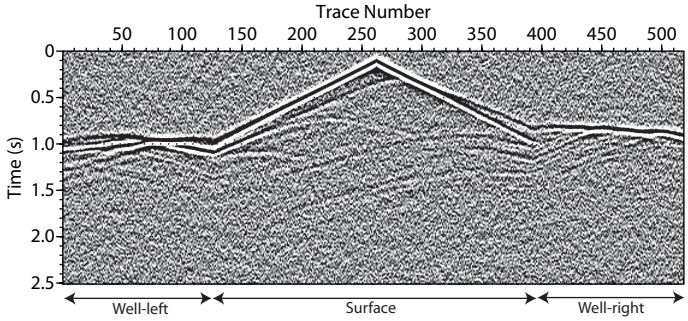


Figure 1.11: Noisy seismograms of pressure data for the source located almost at the center of the Marmousi model $x = 1.4\ km$; random Gaussian noise added to the synthetic seismograms in bandwidth of $0 - 30\ Hz$ and $SNR = 7\ dB$.

are chosen identical to those of the noise-free data set case, except the hyper-parameter λ_2, which is now increased in order to account for the noise energy in data. The λ_2 value is chosen such that the ratio between the prior-model penalty and the data misfit remains the same in the global misfit function, and equal to the noise-free case ($\gamma = 3 \times 10^{-3}$). In presence of noise, the data misfit function has a larger value than the noise-free case, therefore a higher λ_2 value is required.

The results with noisy data remain consistent with the noise-free tests: the prior penalty term still drives the inversion towards a more realistic and accurate final model, even though more noisy. The dynamic approach also remains an appropriate strategy as can be shown from the model reconstruction and from the two displayed vertical profiles (Figure 1.12f). This test for data with the presence of noise confirms the robustness of the approach for non-perfect data.

1.2.4.5 Surface acquisition and multiple-contaminated data

In this section, we apply our scheme to a less favorable frame: a free surface condition is used, meaning that surface multiples are now present in the data. Moreover, we suppress the sensors located in the wells, leading to a pure surface acquisition. Note that the sensors in the wells previously allowed to dramatically increase the illumination in this selected target of Marmousi, where the ratio maximum offset/depth is about two, instead of three to four as in classical FWI applications to exploit diving waves. In such a configuration with a small offset compared to the depth, the FWI behaves generally like a non-linear migration technique, as the low part of the wavenumber domain can not be retrieved (Plessix and Mulder, 2004) and we may question how the prior information may fill in this part of wavenumber domain.

The observed data used in this test are shown in Figure 1.13. The surface-related multiples can clearly be seen compared to the previous data set.

Two inversion tests starting from the "smoothed velocity model" are performed. The first one uses the standard FWI without the prior model (Figure 1.14a). The second test uses the "interpolated velocity model" as the prior model, the optimal prior weighting model B, and the dynamic approach (Figure 1.14c). Here, we assume that the sonic logs are still available and that we can use them for building the prior model (like in previous tests).

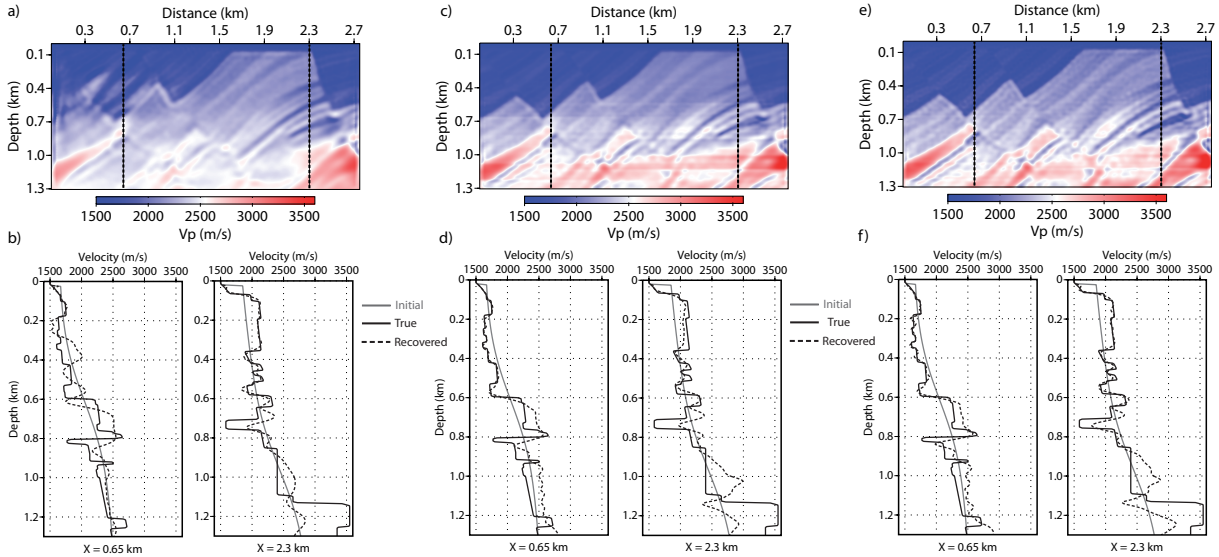


Figure 1.12: The recovered V_p models by FWI of the noisy data and two QC vertical logs passing through the two target areas at $x = 0.65\ km$ and $x = 2.3\ km$, (a) reconstructed model starting from the smooth initial model and without a prior model, small λ_1 and $\lambda_2 = 0$; (b) two vertical logs corresponding to the model (a); (c) reconstructed model starting from the smooth initial model and with the prior model, small λ_1, fixed adapted λ_2 to have the ratio $\gamma = 3 \times 10^{-3}$ at first iteration, and the prior weighting model B; (d) two vertical logs corresponding to the model (c); (e) reconstructed model starting from the smooth initial model and with the prior model, small λ_1, same initial λ_2 before it is decreased to zero (or same initial ratio $\gamma = 3 \times 10^{-3}$), the prior weighting model B and using the dynamic approach. The prior model is removed from the inversion at the end of the procedure; (f) two vertical logs corresponding to the model (e).

All the parameters are chosen identical to those of the previous tests. The starting λ_2 value is adapted, due to a difference in the data energy and trace number, to keep the same $\gamma = 3 \times 10^{-3}$ ratio between the prior-model and data misfit terms at the first iteration.

The result of standard FWI (Figure 1.14a) contains many anomalies and ringing artifacts in the shallow part. These effects could be associated to surface-related multiples coming from the free surface. Even if the main structures are recovered in the shallow part, the velocity model is strongly contaminated by artifacts at all depths leading to erroneous velocity values at the two reservoir depths (see the QC logs in Figure 1.14b). Moreover, due to the limited aperture coverage of the acquisition, the low part of the wavenumber spectrum is not recovered, and the structure seems depth stretched due to the initial model inaccuracy. Adding the prior model and the dynamic weighting allows to significantly improve the results (Figure 1.14c), canceling the shallow ringing effects associated to the surface-related multiples. In the shallow part of the target where illumination remains strong, the result is almost perfect. Deeper, because of the lack of illumination, some artifacts appear, but the prior model allows, at least partially, to fill in the low part of the wavenumber spectrum that can not be retrieved from short-spread reflection only. This leads to well-positioned structures, until at least 1 km depth (Figure 1.14d).

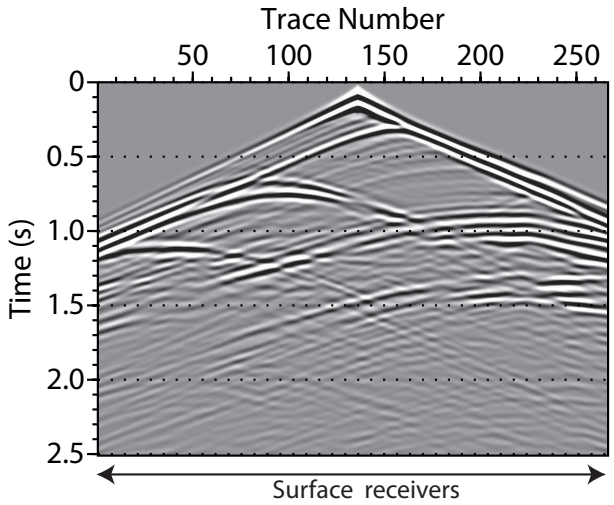


Figure 1.13: Seismograms of pressure data for the source located almost at the center of the Marmousi model $x = 1.4\ km$, recorded inside the true model with free surface condition and using the receivers only at the surface.

1.2.5 Discussion

So far, FWI has been considered essentially as a data-driven procedure with negligible contribution of prior model information and has therefore been investigated for seismic exploration purposes. As the knowledge of the target zone is increased, we may need to introduce more model-driven features in the optimization procedure, especially when we have poor illumination of the target zones.

The description of the misfit function with three terms should increase potential perspectives of the FWI as we may relax the illumination constraints of this approach at the expense of a better knowledge of expected features of the model we want to reconstruct.

The design of the hyper-parameters and, more specifically, the dynamic evolution during the inversion procedure, could be improved and robustness should be analyzed. One can say that this tuning is based on the misfit evolution for different damping laws. From the synthetic example we have studied, we have found that behaviors of the FWI for different acquisition configurations, namely the one with receivers at the surface and within two wells, and the one with receivers only at the surface with the surface-related multiples, are quite similar and, therefore, the tuning of hyper-parameters should not be highly sensitive to the application.

In addition, deep targets could benefit as well from the prior information once the overburden structure has been defined. In seismic exploration, sub-basalt and subsalt imaging is quite challenging and any extra piece of information could help to improve the illumination of the target. The introduction of prior information would allow to help recover poorly illuminated zones, thus broadening the application of the full waveform inversion.

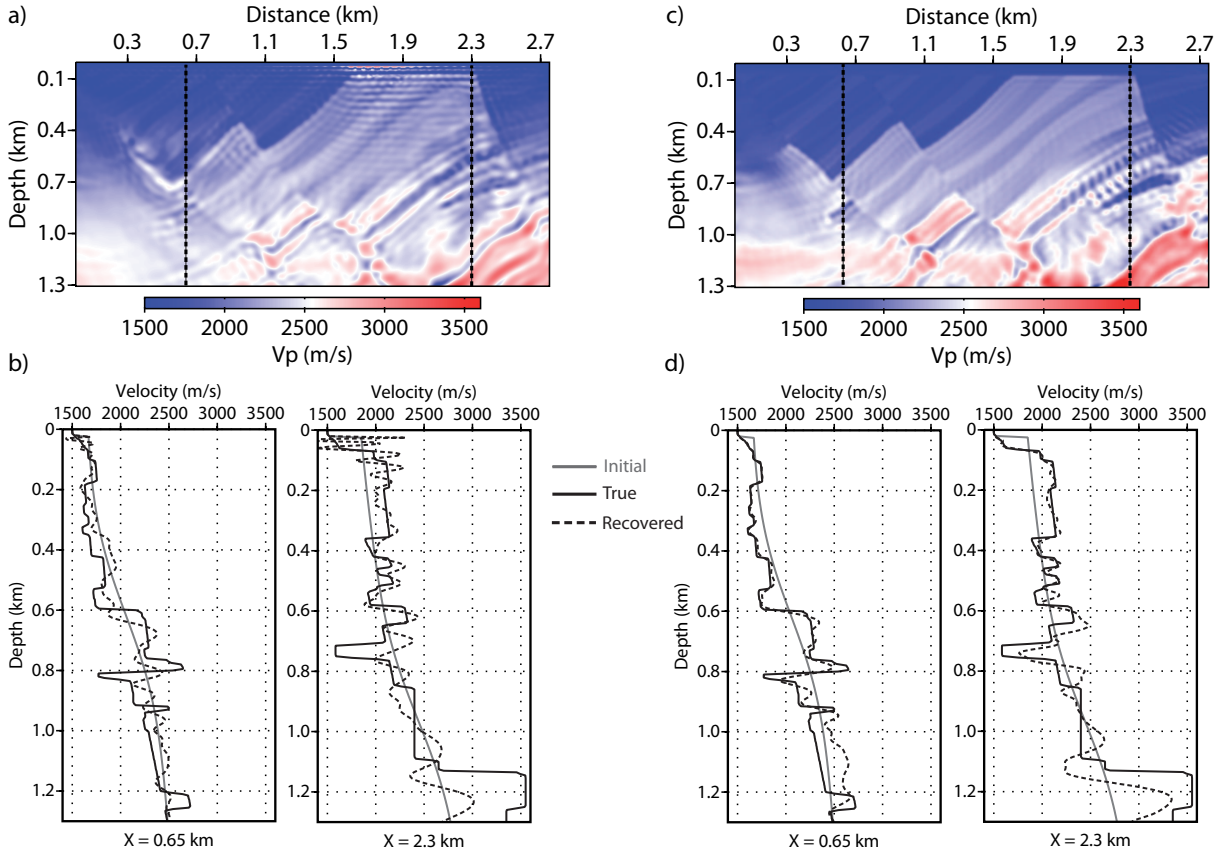


Figure 1.14: The recovered V_p models by FWI and two QC vertical logs passing through the two target areas at $x = 0.65\ km$ and $x = 2.3\ km$, in case of *free surface condition and using only the receivers at the surface*, (a) reconstructed model starting from the smooth initial model and without a prior model, small λ_1 and $\lambda_2 = 0$; (b) two vertical logs corresponding to the model (a); (c) reconstructed model starting from the smooth initial model and with interpolated velocity model (Figure 1.3c) as a prior model, small λ_1, initial λ_2 value chosen to have the ratio $\gamma = 3 \times 10^{-3}$ at first iteration, the prior weighting model B and using the dynamic approach; (d) two vertical logs corresponding to the model (c).

1.2.6 Conclusions

We have proposed a regularized FWI scheme that includes prior information as an optimization penalty term. Aside from the data misfit term, our misfit definition is composed of two penalty terms: the Tikhonov term to ensure smoothness and the prior model term to help the convergence towards expected models. Generally, this latter prior penalty term is not used in classical FWI implementation, but we show that adding this information reduces the non-uniqueness issue of the inverse problem that is a well-known difficulty of the full waveform inversion. This prior information can be deduced from non-seismic data, well-logging and geological constraints that are generally available for specific exploration applications and for monitoring during production. We show with that this prior information improves the well-posedness of the problem as compared to the standard FWI approaches, and allows to partially

mitigate potential kinematic inaccuracy of the starting model as well as illumination issues. The prior weighting operators and the prior model require an appropriate design: one has to properly balance the prior model term and the data term during the inversion. We have shown that well-balancing both pieces of information is crucial during the optimization process. The prior model definition, may include structures that can help drive the inversion towards the global minimum valley without being kinematically as accurate as the initial model for wave propagation. Moreover, these structures may prevent convergence in the final steps of the inversion and one can wish to decrease the importance of this prior information with respect to the data information. We have proposed a dynamic weighting of the prior term during the inversion in order to smoothly reduce the impact of the prior information, leaving the floor at the end only to the data itself (regardless of the smooth Tikhonov term).

During this investigation, we have shown the striking differences between the roles of the initial model and the prior model in this constrained FWI framework where generally only partial information is available: the initial model must respect the wave equation and the related kinematic features to be positioned in the global minimum valley of the misfit function; while the prior model does not have such obligations restrictively like the initial model. The construction of the initial model is quite delicate while the construction of the prior model could allow more freedom.

Future investigations will be focused on using prior model for time-lapse applications, in order to accurately obtain the physical parameter variations in a target zone. The design of the prior model, for more complex environments and real data applications, should also consider geostatistical approaches and/or standard quantitative interpretation techniques to honor the geological structures.

1.2.7 Acknowledgements

We would like to thank TOTAL Exploration & Production and SEISCOPE consortium for supporting this study. This work was performed by access to the high-performance computing facilities of CIMENT (Université de Grenoble) and to the HPC resources of GENCI-CINES under Grant 2011-046091. We acknowledge both of these facilities and the support of their staff. We would like to gratefully thank the Associate Editor, Aria Abubakar, three anonymous reviewers and Partha Routh, for their very constructive comments on the manuscript.

1.3 Estimation of regularization parameters

The main issue concerns the development is a reliable technique for choosing the regularization parameters. Several methods, such as *discrepancy principle, L-curve, Generalized Cross Validation (GCV)* and *NCP Analysis* have been proposed but no one has advised a general purpose algorithm which will always provide a proper estimation of these regularization parameters (Hansen, 1998; Castellanos et al., 2002; Hansen, 2010). In our paper, we suggested to approximately choose these parameters based on the ratio between model and data misfit functions at the beginning of optimization. In our regularization scheme, since the dynamic approach will be applied, probably very precise estimation of this regularization parameter at the beginning may probably not be necessary.

Here, we briefly introduce two different well-known parameter-choice techniques that could also be used in seismic data inversion:

Discrepancy principle: in this method the regularization parameter is chosen such that the data residual norm equals discrepancy in the data which is measured by the noise norm, $\|d_{cal}(\mathbf{m}_\lambda) - d_{obs}\|_2 = \|n\|_2$. This method is simple but the main disadvantage is that we often do not know the noise norm exactly, and therefore we must estimate the norm or the standard deviation of data noise. This might be difficult to obtain in practice and unfortunately the quality of the computed regularization parameter is very sensitive to the accuracy of the estimation of $\|n\|_2$ (Hansen, 2010).

L-curve: this technique is based on the balancing of two error components, a perturbation error coming from the inversion of the noise component in the data and a regularization error due to the introduction of the regularization filter (Hansen, 1992). The regularization parameter is defined from the corner of the L-curve plot. This curve is a plot of the regularized solution norm versus the residual norm, both in log scale $(log\|d_{cal}(\mathbf{m}_\lambda) - d_{obs}\|_2, log\|\mathbf{m}_\lambda\|_2)$.

It is important to note that all these parameter estimation methods could be expensive for FWI applications where we have an iterative optimization method and expensive computational tasks. The reader is referred to Hansen (1998) and Hansen (2010) for a detailed explanation of these techniques.

1.4 Building prior model using geostatistical techniques

For building the prior model from the available prior information, we can use more accurate techniques than just doing a simple interpolation between wells. One of the potential methods could be a geostatistical approach (Isaaks and Srivastava, 1989; Dubrule, 2003). In this technique, the spatial variation due to distance from wells can be accounted; therefore, this geostatistical interpolation could provide more reliable prior model. Indeed, in addtion to well data, the available geology data could also be integrated for building the prior model.

Kriging technique addresses the problem of interpolating a variable like velocity on the basis of a number of scattered data in the space (Matheron, 1970). Let us consider that we know the velocity computed at a well location from the sonic log. Then, we want to estimate the velocity value far away from the well location by using the value at the well. We can do that based on one variogram model. The variogram model contains the information related to variable variation with the distance. Obviously, we have an estimation error which could be provided with statistical models. Kriging is the generalization of this idea to the estimation of the variable value at one specific location x using not just one well but several wells away (Dubrule, 2003). In fact, this technique provides an estimation of the variable in the space based on a weighted average of the measured values in several scattered data points. These weights are directly linked to the spatial distance between the estimated point and the measured points. Kriging will consist in finding these weights such that minimizing the estimated variance.

This kriging technique could be used for constructing the prior model using distributed well data in 2D or 3D spaces. In addition, the geostatistic interpolation could be honored from the available geological horizons. It means the interpolation can be perfomed by following the

velocity interfaces. Indeed, for time-lapse applications, we have usually several sources of prior information on the studied field.

1.5 Other possibilities of dynamic methods

The dynamic weighting of prior model penalty term we propose, could also be implemented by other methods. The similar idea used in multiplicative regularization (van den Berg et al., 2003), where the data objective function is defined as the weight of total variation, could be used for the regularization parameter λ_2 of prior model penalty term. The ratio between prior model objective term and data part of objective function is important and this kind of information can be used to control the dynamic approach.

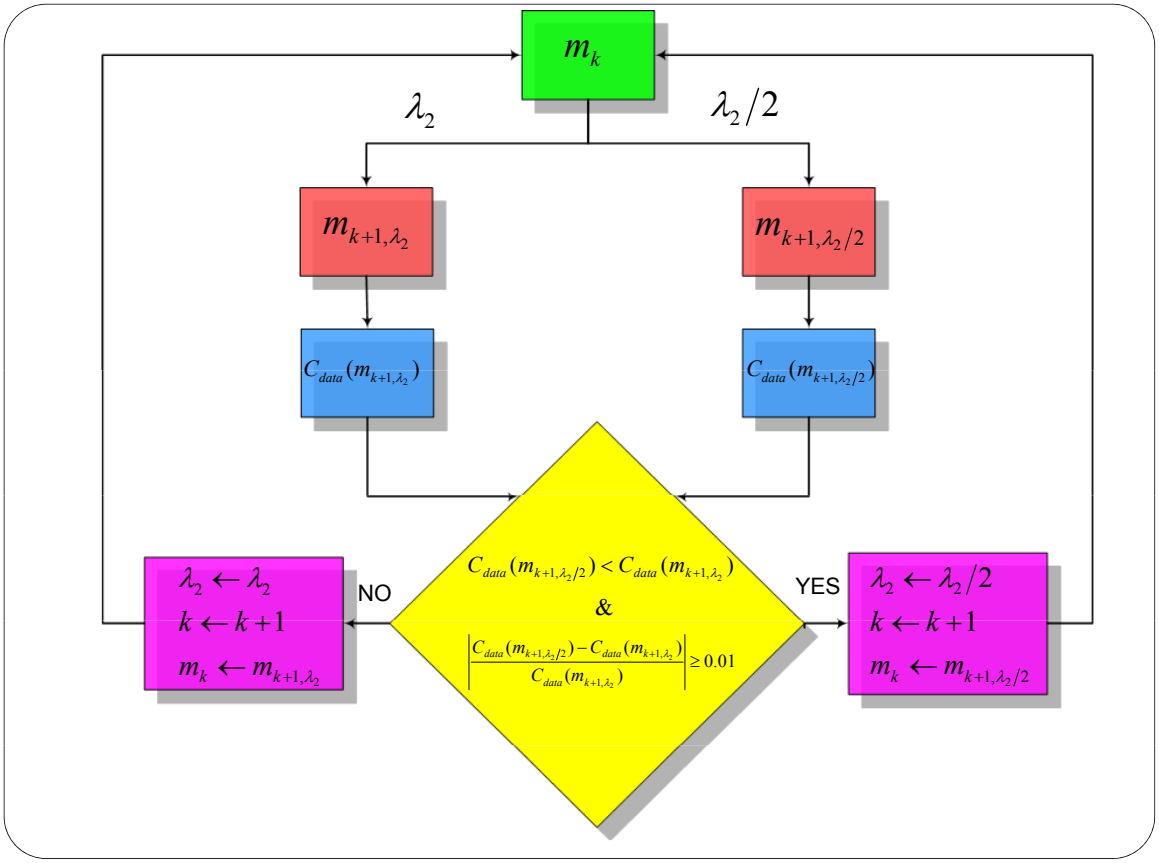


Figure 1.15: Lambda half-lambda algorithm based on choosing damping parameter in Levenberg–Marquardt algorithm. This algorithm could be an alternative for the dynamic weighting approach.

Another possibility, called here "lambda half-lambda" method, is derived from the idea of Levenberg–Marquardt algorithm (LMA) (Levenberg, 1944; Marquardt, 1963; Moré, 1978). The LMA is originally an optimization method modifying the Hessian ($\mathbf{B}_{lm} = \mathbf{B} + \epsilon\mathbf{I}$). The idea of choosing the damping parameter (ϵ) which was suggested by Marquardt (1963) can be

interesting for us as a dynamic approach for changing the regularization parameter λ_2. Our proposed "lambda half-lambda" algorithm is shown in Figure 1.15.

In this suggested algorithm, at each iteration, two optimizations must be done. Once with λ_2 value as a weight of prior model penalty term and one time with half of this current λ_2. After both optimizations, the data terms of the objective function in two cases are compared. If use of the half-lambda results in a more reduction in objective function then this is taken as the new value of λ_2 and the new optimum model is taken as that obtained with this half-lambda and the process continues; otherwise λ_2 is left unchanged and the new optimum is taken as the value obtained with λ_2 as regularization parameter. This algorithm can deliver a robust result (Figure 1.16) similar as derivative based method used for the dynamic approach (compare Figure 1.16c with Figure 1.6g). However, performing two optimizations at each step would be the main drawback of this algorithm.

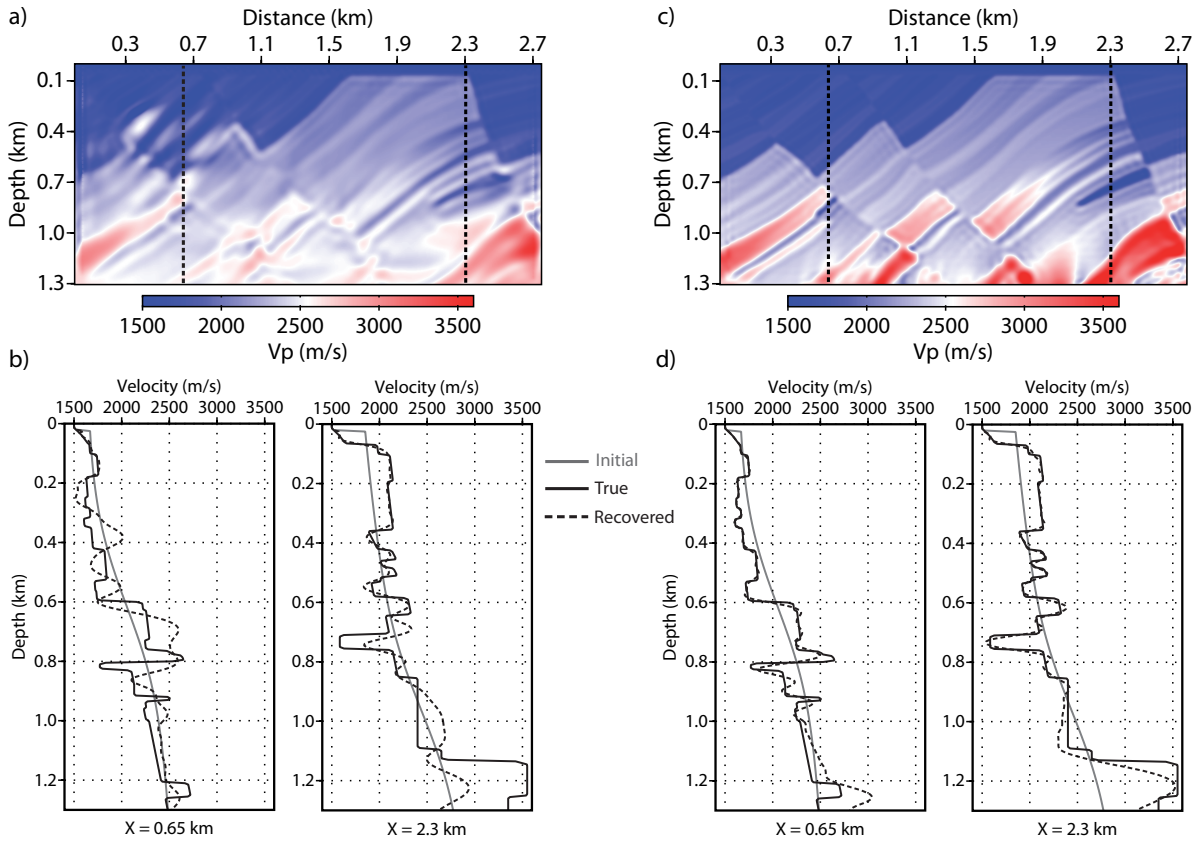


Figure 1.16: The recovered V_p models by FWI and two QC vertical logs passing through the two target areas at $x = 0.65\ km$ and $x = 2.3\ km$, (a) the reconstructed model by classical FWI without prior model, (b) two vertical logs correspond to model (a); (c) the reconstructed model by regularized FWI with dynamic prior weighting performed by "lambda half-lambda" method, (d) two vertical logs correspond to model (c).

1.6 Conclusions

We have proposed a regularized FWI scheme that includes prior information as an optimization penalty term. Aside from the data misfit term, our misfit definition is composed of two penalty terms: the Tikhonov term to ensure smoothness and the prior model term to help the convergence towards expected models. Using a synthetic benchmark, we show how the prior information improves the well-posedness of the inverse problem as compared to the standard FWI approaches, and allows to fill in partially the lack of low-wavenumber during inversion. The prior model definition may include structures that can help drive the inversion towards the global minimum valley. This prior model is not involved in the synthetic data computation and, therefore, is not required to be kinematically as accurate as the initial model for wave propagation.

However, structures of the prior model may prevent convergence in the final steps of the inversion and one can wish to decrease the importance of this prior information with respect to the data information. We have proposed the idea of dynamic weighting of the prior term during the inversion in order to smoothly reduce the impact of the prior information. This idea could be promising and the way of performing that in a proper way can be an open research study.

In time-lapse applications, this kind of prior information should be included for the inversion. First, in order to get more accurate and more precise baseline model, and second to use these prior data for inverting the small time-lapse variations. In the next chapter, we apply this regularized FWI algorithm on synthetic time-lapse examples and show how the time-lapse variations could be obtained in a robust way.

Chapter 2

Time-lapse Full Waveform Inversion

Contents

After the accurate baseline reconstruction through FWI using limited aperture and with the introduction of prior information, we may consider different strategies for the time-lapse inversion. In this chapter, we discuss different time-lapse strategies and we compare the robustness of these strategies on Marmousi synthetic datasets. We make different discussions about null-space issues and differences between baseline and monitor inversions. In addition, we propose a target-oriented time-lapse inversion based on the prior model and the prior weighting.

In a second part, we apply our regularized FWI method to another synthetic time-lapse datasets with steam injection, namely Dai model (Dai et al., 1995). In this second synthetic case, we investigate the sensitivity of the recoverd baseline model with respect to different initial and prior models. Finally, the sensitivity of time-lapse models obtained by differential FWI strategy with respect to the baseline models is studied and discussed.

2.1 Time-lapse seismic imaging using regularized FWI with prior model: which strategy?

Time-lapse seismic imaging using regularized full waveform inversion with prior model: which strategy?
Amir Asnaashari, Romain Brossier, Stéphane Garambois, François Audebert, Pierre Thore, and Jean Virieux
2013, *Geophysical Prospecting*, accepted.

2.1.1 Abstract

Full waveform inversion (FWI) is an appealing technique for time-lapse imaging, especially when prior model information is included into the inversion workflow. Once the baseline reconstruction is achieved, several strategies can be used to assess the physical parameter changes, such as parallel difference (two separate inversions of baseline and monitor datasets), sequential difference (inversion of the monitor dataset starting from the recovered baseline model), and double-difference (inversion of the difference data starting from the recovered baseline model) strategies. Using the synthetic Marmousi datasets, we investigate which strategy should be adopted to get more robust and more accurate time-lapse velocity changes in noise-free and noisy environments. In addition, we propose a target-oriented time-lapse imaging using regularized FWI including prior model and model weighting, if the prior information exists on the location of expected variations. This scheme applies strong prior model constraints outside of the expected areas of time-lapse changes, and relatively less prior constraints in the time-lapse target zones. In application of this process to the Marmousi model dataset, the local resolution analysis performed with spike tests shows that the target-oriented inversion prevents the occurrence of artifacts outside the target areas, which could contaminate and compromise the reconstruction of the effective time-lapse changes, especially when using the sequential difference strategy. In a strongly noisy case, the target-oriented prior model weighting ensures the same behavior for both time-lapse strategies, the double-difference and the sequential difference strategies, and leads to a more robust reconstruction of the weak time-lapse changes. Therefore,

we show that, in some configurations where the double-difference strategy is not applicable (for example, non-perfectly matched acquisition geometries), the target-oriented sequential difference strategy can provide the same robust result as the double-difference strategy.

Keywords

Monitoring, Full Waveform, Inversion, Time-lapse, double-difference, target-oriented

2.1.2 Introduction

Over several years, monitoring and time-lapse analysis has become an important tool for optimizing the oil and gas production and for evaluating CO_2 sequestration efficiency. Time-lapse data provide high-resolution images which enable us to track dynamic changes in physical properties, especially fluid parameter variations in target areas. Several successful time-lapse studies have been undertaken for near-surface geophysical problems with non-seismic data, GPR and electrical data (Ramirez et al., 1993; Day-Lewis et al., 2002, 2003; Singha and Gorelick, 2005; Oldenborger et al., 2007; Miller et al., 2008), for seismic monitoring (Lumley, 2001; Rickett and Lumley, 2001; Calvert, 2005; Hall, 2006; Lumley et al., 2008; Ayeni and Biondi, 2010), and for joint inversion of electromagnetic and seismic data of reservoir monitoring (Liang et al., 2012). Most of these studies are based on tomography problems where inversion techniques are needed.

Over the past few decades, full waveform inversion (FWI) has become a promising technique for velocity model building that reconstructs high-resolution velocity models of the subsurface through the extraction of the full information content of seismic data (Tarantola, 1984b; Pratt, 1999; Virieux and Operto, 2009). Since the FWI approach delivers high-resolution quantitative images of macro-scale physical parameters, it should appear as a quite attractive tool for monitoring purposes, even if it is not yet widely applied (Gosselet and Singh, 2008; Abubakar et al., 2009; Plessix et al., 2010; Thore et al., 2010; Romdhane et al., 2012; QueiBer and Singh, 2013). Gradient-based inversions used in FWI cannot provide yet uncertainties on the inverted values, in spite of few attempts (Fichtner and Trampert, 2011), compared to stochastic approaches. The computational cost of the a posteriori covariance matrix limits dramatically such kind of analysis for FWI applications. Nevertheless FWI approach could be one of promising techniques for time-lapse imaging, taking into account both the phase and amplitude of data and and it may be considered as less user dependence. Moreover, the final product of FWI relies on quantitative image of physical properties that should help interpretation of 4D changes.

Time-lapse variations are defined as a high-frequency details in the model at the reservoir scale, while FWI is known as a technique which is limited to rather low frequencies for exploration projects due to the computational cost. However, even with these low frequencies it can provide high resolution images. Today, by increasing the computer facilities and more developments in FWI algorithms, there is less problem to move onto inversion of intermediate frequencies. Nice applications have been shown at intermediate and high frequencies in both 2D and 3D problems (Shipp and Singh, 2002; Bansal et al., 2013; Lu et al., 2013). Moreover, for time-lapse applications, it is not necessary to invert an entire field model which would be very large (this is the case in general exploration project); the inverted model could be reduced to the interested zones. Therefore, by decreasing the size of model (unknown parameters), it

helps to move to a higher frequency inversion. Please note however that the problem of high frequency inversion is not only a computational cost issue: the amplitude and cycle-skipping issues become more and more important when moving to high frequencies.

Regularization or preconditioning may reduce the non-uniqueness of the ill-posed inverse problem. Several studies have been done on regularization schemes of FWI implementations (Abubakar et al., 2009; Herrmann et al., 2009; Loris et al., 2010; Guitton, 2011; Ma et al., 2012). All these regularization techniques do not directly use prior model information as a separate term in the data-oriented misfit function: they prefer to concentrate on the definition/restriction of the model space. One may want to use prior model information in the FWI scheme as is done in other velocity building techniques (Le Stunff and Grenier, 1998; Tarantola, 2005). Recently, some authors suggested to introduce prior models in the multiplicative (Hu et al., 2009) and additive (Wang et al., 2012; Asnaashari et al., 2013a) regularization terms of the FWI workflow. Taking into account prior information could be very valuable for monitoring purposes, where many types of precise prior information have been collected for the target zone, in a way similar to the model-based regularization used on time-lapse electrical resistivity tomography (Oldenborger et al., 2007; Miller et al., 2008). Adding prior information can also significantly improve the accuracy of the reconstruction of the baseline model (Asnaashari et al., 2013a).

The time-lapse reconstruction procedure can be divided into two steps: (1) the baseline and (2) the monitor model reconstructions. In order to obtain a robust high-resolution time-lapse model, it is necessary to reconstruct both baseline and monitor models in a robust and accurate way. For the second step of this approach, several workflows can be designed for the monitor reconstruction. One of them, called the *parallel difference strategy*, independently inverts the two datasets (baseline and monitor) starting from the same initial model. The time-lapse changes will be assessed by subtracting the final derived monitor model from that of the baseline (Plessix et al., 2010). The second approach, called the *sequential difference strategy*, uses the final baseline model as the starting model for inverting the monitor dataset (Asnaashari et al., 2012). An alternative strategy, called *double-difference or differential strategy*, consists in inverting only the difference dataset to recover a differential image. It is related to a double-difference tomography method widely used in geodesy and in seismology in order to improve earthquake source locations or to image receiver areas (Waldhauser and Ellsworth, 2000; Zhang and Thurber, 2003; Monteiller et al., 2005; Got et al., 2008), and used more recently for locating microseismic events (Zhou et al., 2010). The double-difference method in the form discussed here has been proposed for time-lapse waveform inversion of acoustic cross-well data in frequency domain (Watanabe et al., 2004) and inversion of elastic data in time domain (Denli and Huang, 2009).

The sequential difference strategy does not depend too heavily on the repeatability of the acquisition geometry between baseline and monitor surveys, which is a great advantage. On the contrary, the double-difference method demands a perfect match of the receiver and source positions between the two surveys. On the other hand, the sequential difference strategy cannot naturally focus on the target areas and on the time-lapse changes during inversion (as shown in this paper). Could we tune the sequential difference inversion for a better focusing on target zones using prior model information? In addition, could we reduce the effects of image noise outside of the expected target zones if prior data are available? These are some of main questions that we want to address in this study.

In the first part, we shall present the theoretical framework of our study. Then, using

synthetic data modeled for the Marmousi model, we shall investigate the different behaviors and robustness of time-lapse strategies in noise-free and noisy cases. The fundamental differences between the sequential difference and the double-difference strategies, in conventional (without prior model) and target-oriented modes, will be illustrated by a local resolution analysis. We underline the sensitivity analysis of obtained time-lapse images with respect to the inaccuracy of the recovered baseline model in a noisy environment. We shall illustrate how the target-oriented inversion can significantly improve the results of all time-lapse inversions and how the sequential difference strategy in target-oriented mode can provide the same robust result as the double-difference strategy if prior data are available in crucial areas.

2.1.3 Time-lapse strategies with regularized full waveform inversion including prior model

Full waveform inversion is an iterative approach to recover model parameters based on the local optimization of residuals between observed and computed wavefields at receiver positions for different seismic source locations. Any prior information which can be incorporated into the estimation of the misfit function reduces the ill-posedness issue of inverse problems and ensures more robust results. For that, we apply the time-domain regularized FWI algorithm with prior model penalty introduced in Asnaashari et al. (2013a), where two penalty terms based on model parameters are introduced into the misfit function. The misfit function $\mathcal{C}(\mathbf{m})$ can be expressed using the ℓ_2 norm as

$$\begin{aligned}\mathcal{C}(\mathbf{m}) &= \frac{1}{2}\sum_{ns}\Big\{\left(\mathbf{d}_{obs}-\mathbf{d}(\mathbf{m})\right)^T\mathbf{W}_d^T\mathbf{W}_d\left(\mathbf{d}_{obs}-\mathbf{d}(\mathbf{m})\right)\Big\}+\frac{1}{2}\lambda_1\{\mathbf{m}^T\mathbf{D}\mathbf{m}\}\\ &+ \frac{1}{2}\lambda_2\Big\{\left(\mathbf{m}-\mathbf{m}_p\right)^T\mathbf{W}_m^T\mathbf{W}_m\left(\mathbf{m}-\mathbf{m}_p\right)\Big\}\\ &= \mathcal{C}_d(\mathbf{m})+\lambda_1\mathcal{C}_{1_m}(\mathbf{m})+\lambda_2\mathcal{C}_{2_m}(\mathbf{m}),\end{aligned} \tag{2.1}$$

where the observed and calculated data at receiver positions are denoted by vectors $\mathbf{d}_{obs}$ and $\mathbf{d}(\mathbf{m})$, respectively. The matrix $\mathbf{W}_d$ is a weighting operator on the data misfit. This first term of the misfit function is obtained by adding the contribution over ns sources. The second term of the objective function corresponds to the standard Tikhonov regularization (Tikhonov and Arsenin, 1977) where the modulus of the gradient (roughness) of the model is minimized, in order to search for smooth models. Practically, this term is formulated as the application of Laplacian operator $\mathbf{D}$. A five-point finite-difference stencil is used for the discretization of this operator. The third term of the objective function is related to the prior model $\mathbf{m}_p$ designed from available prior information, such as sonic logs and geological information. This term estimates residuals between the current model at a given iteration and the prior model considered at that same iteration. The matrix $\mathbf{W}_m$ is a weighting operator in the model space. This matrix contains the prior uncertainty information and the potential weighting of this information. This weighting operator plays a crucial role when reconstructing the model for target-oriented imaging. In our misfit function, the prior-model variance information is included in the diagonal $\mathbf{W}_m$ matrix and the covariances are implicitly considered in the Tikhonov term. The two regularization hyper-parameters λ_1 and λ_2 allow each penalty term to be weighted with respect to each other and to the data term. Designing the prior model and estimating the weighting matrix is an important step before the optimization. Moreover, hyper-parameters should also be estimated. These issues come with the new definition of the misfit function.

The gradient of misfit function now has three terms, one data term and two model terms, expressed as

$$\mathcal{G}_m = -\sum_{ns} \mathbf{J}^T \mathbf{W}_d^T \mathbf{W}_d \left(\mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})\right) + \lambda_1 \mathbf{D}\mathbf{m} + \lambda_2 \mathbf{W}_m^T \mathbf{W}_m \left(\mathbf{m} - \mathbf{m}_p\right), \tag{2.2}$$

where the Jacobian matrix $\mathbf{J} = \partial \mathbf{d}(\mathbf{m})/\partial \mathbf{m}$ is composed by the Fréchet derivatives of the synthetic data with respect to the model parameters. The data term of the gradient is efficiently computed with an adjoint formulation (Plessix, 2006) without an explicit computation of $\mathbf{J}$. Then, two model penalty terms are readily added to the data term without extra computer efforts as they can be easily differentiated. The quasi-Newton procedure involving the L-BFGS-B scheme (Byrd et al., 1995) is used for the optimization and updates the model parameter vector $\mathbf{m}$. An approximate non-diagonal inverse Hessian (second derivative of the misfit function) from previous gradients and model vectors is considered in this scheme. This bounded limited-memory quasi-Newton method is an efficient alternative to preconditioned steepest-descent methods, because of improved focusing and partially correcting the descent direction affected by limited aperture of illumination (Brossier et al., 2009). The bounded values can be approximately estimated from the prior information. A detailed review and applications of this regularized FWI scheme including prior penalty can be found in Asnaashari et al. (2013a).

In the following, we briefly introduce three different strategies for time-lapse imaging with this new regularized FWI approach, considering initial and prior models.

2.1.3.1 Parallel difference strategy

The parallel difference strategy considers two independent inversions, where the baseline and monitor datasets are processed separately using the same starting and prior models. After inversion, the time-lapse variation is simply the subtraction between the recovered (rec.) monitor model and the reconstructed (rec.) baseline model (Plessix et al., 2010), as explained by the diagram in Figure 2.1. The main advantage of this strategy is its applicability to acquisition geometries that do not necessarily match between the two surveys. As the two inversions are performed independently, the main drawback is the potential interpretation of non-repeated inversion artifacts as a false time-lapse response.

2.1.3.2 Sequential difference strategy

Because one expects the time-lapse response to be localized and of small amplitude in most of the monitoring applications, the baseline model is a natural good candidate for the starting model of the inversion of the monitor dataset. Starting from the baseline model should reduce the number of necessary iterations for the inversion of the monitor dataset. Therefore, once the baseline reconstruction is completed, this baseline model is subsequently used as the initial model during the inversion of the monitor dataset. At the end, the time-lapse model is obtained by a subtraction between the two recovered models. The schematic of sequential difference strategy is shown in Figure 2.2. This approach is similar to the previous strategy, but with a different initial model. As the baseline model is close to the monitor model due to small and localized time-lapse variations, the recovered baseline could also be a good candidate for the

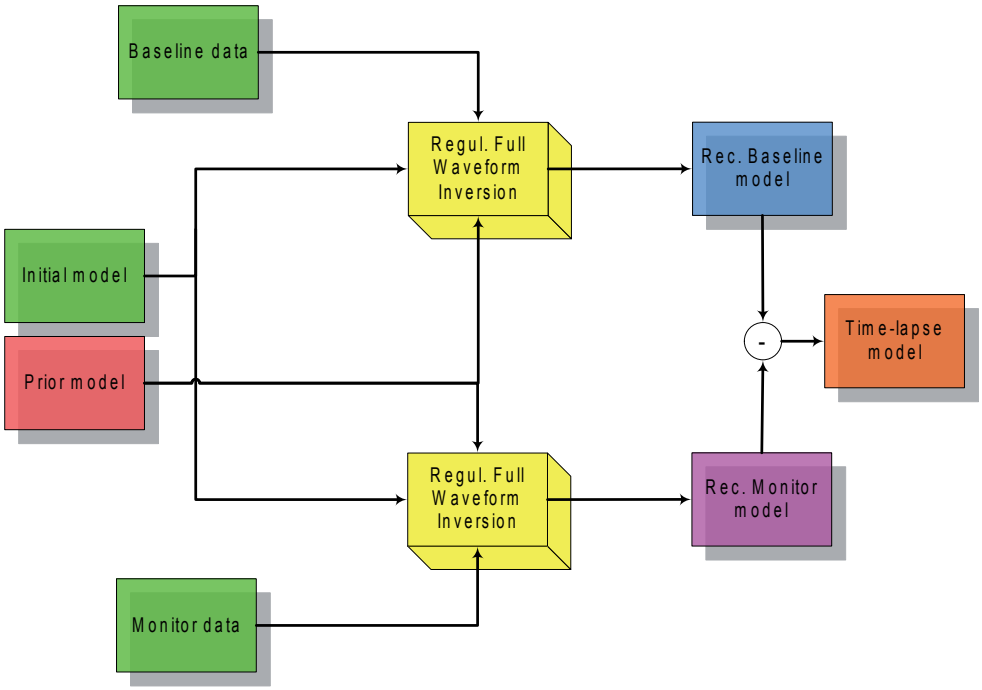


Figure 2.1: Diagram of parallel difference full waveform inversion.

prior model. The role of the prior model is to regularize the inversion and to find the monitor model not far away from the initial one (baseline model). If more prior data on the changes are available from boreholes and non-seismic data, they would be integrated to design a new prior model. However, the prior model alone is not sufficient, and it is important to consider an appropriate model weighting matrix $\mathbf{W_m}$. The prior information related to the location of time-lapse changes should be included in the $\mathbf{W_m}$ matrix. We discuss this issue in more detail in the target-oriented FWI section.

2.1.3.3 Double-difference strategy

In the double-difference strategy, instead of minimizing the difference between the full observed and calculated data, we attempt to minimize the difference of the difference data between two sets of data (Watanabe et al., 2004; Denli and Huang, 2009), yielding the expression

$$\Delta\mathbf{d} = (\mathbf{d}_{obs_m} - \mathbf{d}_{obs_b}) - (\mathbf{d}_{calc_m} - \mathbf{d}_{calc_b}) = \mathbf{d}_{obs_{time-lapse}} - \mathbf{d}_{calc_{time-lapse}}, \tag{2.3}$$

where the monitor and baseline observed data are denoted respectively by $\mathbf{d}_{obs_m}$ and $\mathbf{d}_{obs_b}$, while the computed data for these two experiments are denoted by $\mathbf{d}_{calc_m}$ and $\mathbf{d}_{calc_b}$. Based on equation 2.3, the double-difference inversion focuses on the difference data, *i.e.* changes in data due to the time-lapse variation. In order to use the standard regularized FWI algorithm at our disposal, we can rewrite equation 2.3 in another way,

$$\Delta\mathbf{d} = (\mathbf{d}_{obs_m} - \mathbf{d}_{obs_b} + \mathbf{d}_{calc_b}) - \mathbf{d}_{calc_m} = \mathbf{d}_{composite} - \mathbf{d}_{calc_m}. \tag{2.4}$$

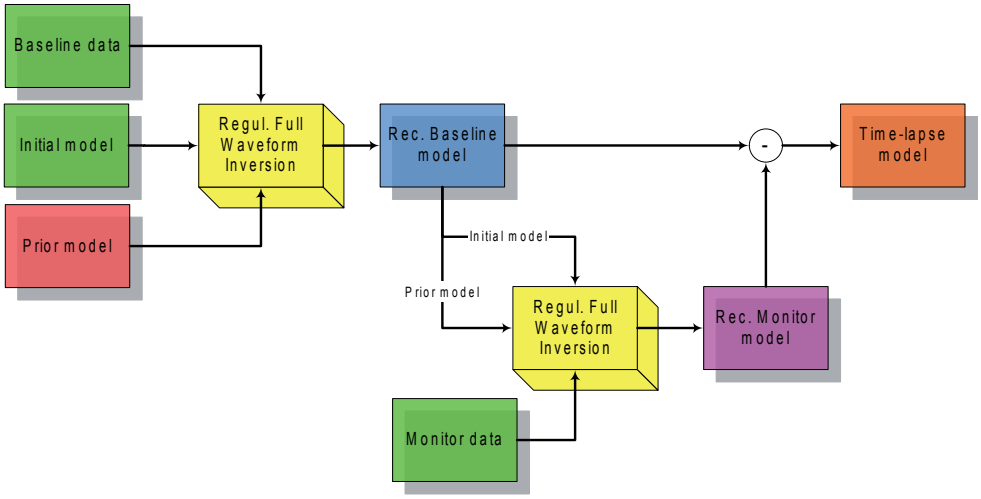


Figure 2.2: Diagram of sequential difference full waveform inversion.

Therefore, for the double-difference analysis, we first need the elaboration of a composite dataset defined as

$$\mathbf{d}_{composite} = \mathbf{d}_{obs_m} - \mathbf{d}_{obs_b} + \mathbf{d}_{calc_{rec-b}}, \tag{2.5}$$

which is composed of (a) the time-lapse difference observed data ($\mathbf{d}_{obs_m} - \mathbf{d}_{obs_b}$) which should only contain the time-lapse changes between the two datasets and (b) the computed data $\mathbf{d}_{calc_{rec-b}}$ estimated using a forward modeling in the recovered baseline model (rec-b). This composite dataset $\mathbf{d}_{composite}$ can be used as a new observed dataset $\mathbf{d}_{obs}$ in equation 2.1, which allows minimizing the double-difference residual 2.3 with a standard regularized FWI algorithm. The reconstructed baseline model is the natural choice for the initial model for this inversion. Finally, the time-lapse model changes $\delta\mathbf{m}_{time-lapse} = \mathbf{m}_{composite} - \mathbf{m}_{rec-b}$ can be obtained. A schematic of the proposed algorithm is shown in Figure 2.3.

Please note that the difference observed data are computed by a raw subtraction, without any time warping. The difference data may contain artificial 4D effects along time axis, due to slightly time-shift related to 4D changes. This issue does not make a problem, since FWI performs inversion in depth domain and not in time domain. These artificial 4D effects are associated to the time-lapse anomaly at specific depth, therefore can be focalized at that depth. Hence, because of inversion in depth domain, there is no need to align baseline and monitor traces in time.

The main advantage of the double-difference strategy is that the unexplained events (non-fitted during inversion) of the baseline data at the baseline reconstruction step have less impact on the time-lapse perturbation reconstruction, as compared to the sequential difference strategy. The composite dataset is free from any unexplained baseline events. In contrast, the sequential difference uses the full real monitor data that may contain unexplained baseline/monitor events. As long as these unexplained events stay common for both baseline and monitor data, they will not produce time-lapse artifacts. This feature, as well as the importance of an accurate recovered baseline model, will be discussed in detail in the synthetic example section.

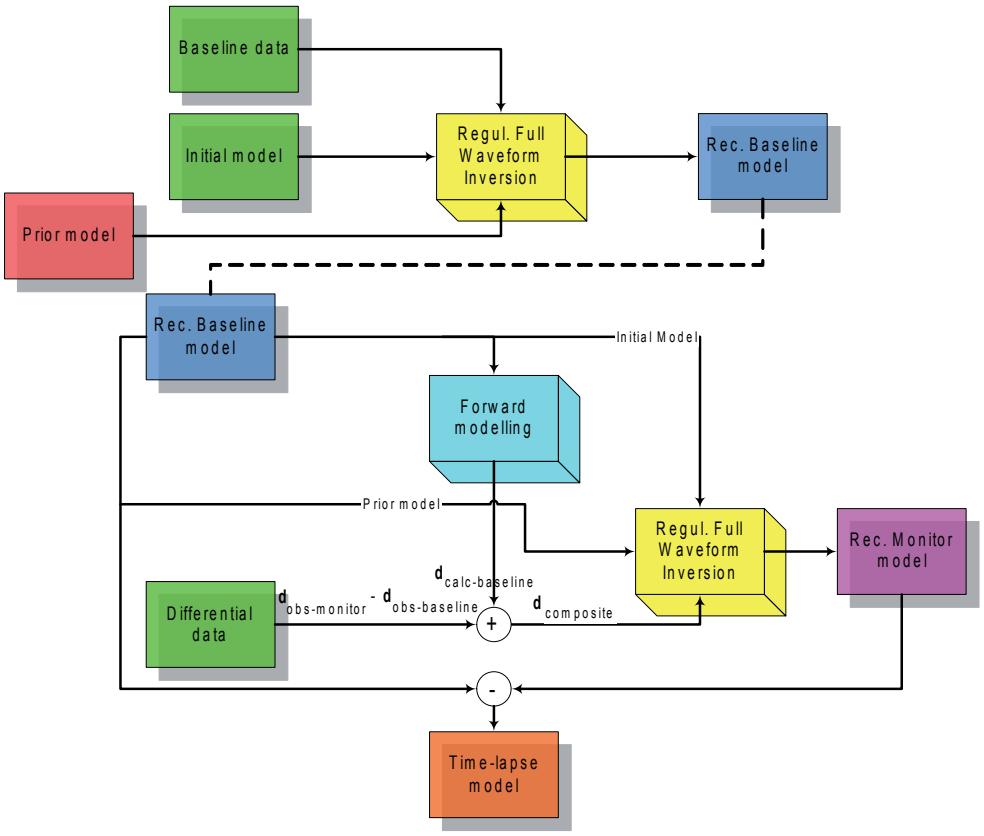


Figure 2.3: Diagram of double-difference (differential) full waveform inversion.

For double-difference, it is essential to have perfectly matched surveys (e.g. permanent sensors or ocean-bottom cable (OBC) configurations) to perform subtraction between datasets. In case of non-matching surveys, we need to tackle extra pre-processing steps and make data interpolation to match receiver and source positions between two surveys. In parallel difference and sequential difference strategies, because of performing two separate inversions, exact-matching issue is less required.

2.1.4 Target-oriented inversion

The sequential difference strategy cannot naturally focus only on target areas and on the time-lapse data changes during inversion. Is there any way to make the sequential difference inversion focus on the target zones? In addition, could we reduce effects of image noise outside of the expected target zones?

As the baseline model is close to the monitor model, since the time-lapse variations are small and localized, the recovered baseline model could also be a good candidate for the prior model in both the sequential difference and the double-difference strategies. If the variations are localized only at the reservoir and assuming no overburden changes, we can constrain the inversion outside of the area of expected changes through the prior-model misfit term of the objective function (the location of these expected changes may come from external

information or be known after previous attempts of different inversions). In other words, we may give strong weights to the prior model (here, the recovered baseline model) for the parts of the model where no time-lapse variations are expected to occur, and relatively less prior weights in the target areas. By doing so, the inversion tends to update the monitor model only inside the target areas where the weaker prior constraints leave more relative weight in the cost function to the data-misfit term with respect to the model term. Of course, hyper-parameters will balance the global contributions between data and model terms. Thus, this target-oriented inversion can be easily incorporated into the double-difference and sequential difference strategies, using the target-oriented model weighting matrix $\mathbf{W}_m$. This is another advantage of these strategies compared to the less interesting parallel difference strategy. The target-oriented procedure could be used as a final focusing strategy. The non-seismic prior data such as well data and geological models, and a migrated section of difference dataset, together, can help to approximately identify the areas of expected changes. The diagonal model weighting matrix is computed based on $diag(\mathbf{W}_m^T \mathbf{W}_m) = 1/\sigma^2(\mathbf{m})$, where $\sigma^2(\mathbf{m})$ is related to the prior uncertainty. In the expected target area, the prior uncertainty associated to the prior model (recovered baseline model) is relatively higher than outside of the target area where no changes are expected. This target-oriented inversion mode can be integrated with two of the time-lapse strategies, the sequential difference and the double-difference ones. This target-oriented option can be performed, if the prior information exists on the location of expected time-lapse variations.

2.1.5 Application to the Marmousi model

In this section, we study the robustness and behavior of the different time-lapse imaging strategies presented before, in the case of both noise-free and noisy synthetic datasets modeled on a version of the Marmousi2 model (Martin et al., 2006). A selected target zone of the Marmousi2 P-wave velocity model and a homogeneous density model are considered as the true baseline models (Figure 2.4a). The true monitor velocity model is created from the baseline model through a relative (40 m/s) variation of velocity inside two gas reservoirs (Figure 2.4b). The velocity changes are around 2% of the baseline velocity variation inside the reservoirs, which is realistic. We assume that the two reservoirs are fully saturated by gas at the beginning, and after production the gas is partially replaced by water. This means, in a second phase of the exploitation, that the reservoirs are partially saturated by gas and water. Due to this issue, there is an increasing change in V_p velocity. A surface acquisition geometry with a free-surface condition is used to generate the synthetic data, with seventy-seven isotropic explosive sources, located along a horizontal line at a depth of 16 m, equally spaced by a distance of 50 m. A horizontal receiver line at a depth of 15 m with a sensor interval of 10 m completes the acquisition. All receivers record all shots. A Ricker wavelet source with a central frequency of 10 Hz is used, for baseline and monitor surveys. The time seismograms are generated using 2D acoustic finite-difference modeling in the time domain with a fourth-order stencil in space and a second-order integration in time. Perfectly-matching-layer (PML) absorbing boundary conditions (Berenger, 1994) are used for non-reflecting boundaries except at the top where a free-surface condition is implemented. The observed data for the inversion are the recorded pressure data.

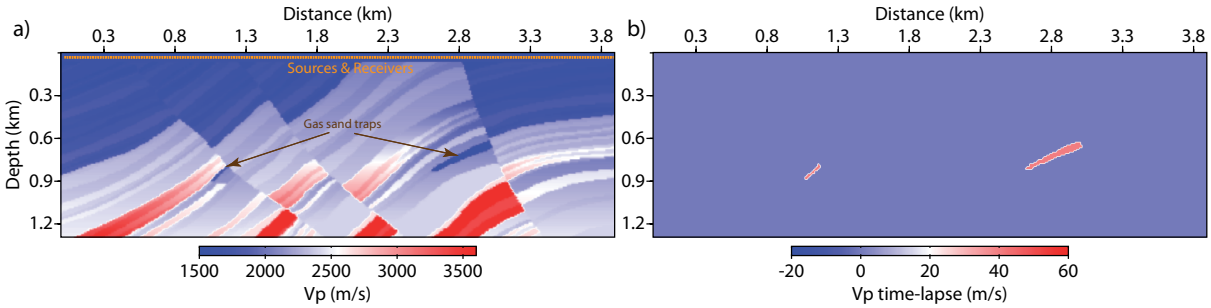


Figure 2.4: (a) The true V_p baseline model selected from Marmousi model and the acquisition geometry; (b) the true time-lapse model.

2.1.5.1 Noise-free data

Figure 2.5a shows an example of baseline seismograms generated by a shot located at the center of the source line. Clearly, surface multiples are present in the data and increase the complexity of the data for inversion. Figure 2.5b shows the noise-free difference data (time-lapse data) obtained by subtraction between the monitor and the baseline data. Note that the amplitude of the difference seismograms is amplified by a factor of 10 so as to be at the same scale as the baseline seismograms, for visualization purposes. The time-lapse signal is weak due to very small velocity variation of the two small target reservoirs. This low energy time-lapse signal may induce difficulties for the recovering of time-lapse changes. The baseline reconstruction should be done as the first step.

2.1.5.2 Baseline reconstruction

The main challenge for monitoring remains the necessity to derive a robust high-resolution baseline model which will be key for the reconstruction of time-lapse parameter variations (Asnaashari et al., 2011). Therefore, an accurate and robust reconstruction of the baseline model is crucial. Recently, Asnaashari et al. (2013a) have proposed a dynamic prior weighting FWI designed to impose a strong contribution of the prior model term in the objective function at the beginning of optimization and then gradually reduce its influence and increase the impact of the data term. Combining the prior information and a dynamic prior weighting allows to drive the inversion to the valley of the quasi-global minimum, to mitigate the cycle-skipping issues, and to progressively give more importance to the data-misfit term at the late iterations of optimization.

The same dynamic approach is followed for the inversion of the baseline data using the regularized 2D acoustic full waveform inversion. Our FWI is implemented in the time domain for both the forward and the inverse problem, involving all the frequencies of the spectrum (between 1 Hz to maximum 25 Hz in this case). No additional hierarchical approach such as the frequency-continuation approach of Bunks et al. (1995) is used. This means that the weighting of each frequency is directly linked to its amplitude in the spectrum. The full offset range is included for the inversion. The forward modeling is performed by 2D acoustic finite-difference method with a fourth-order stencil in space and a second-order integration in time.

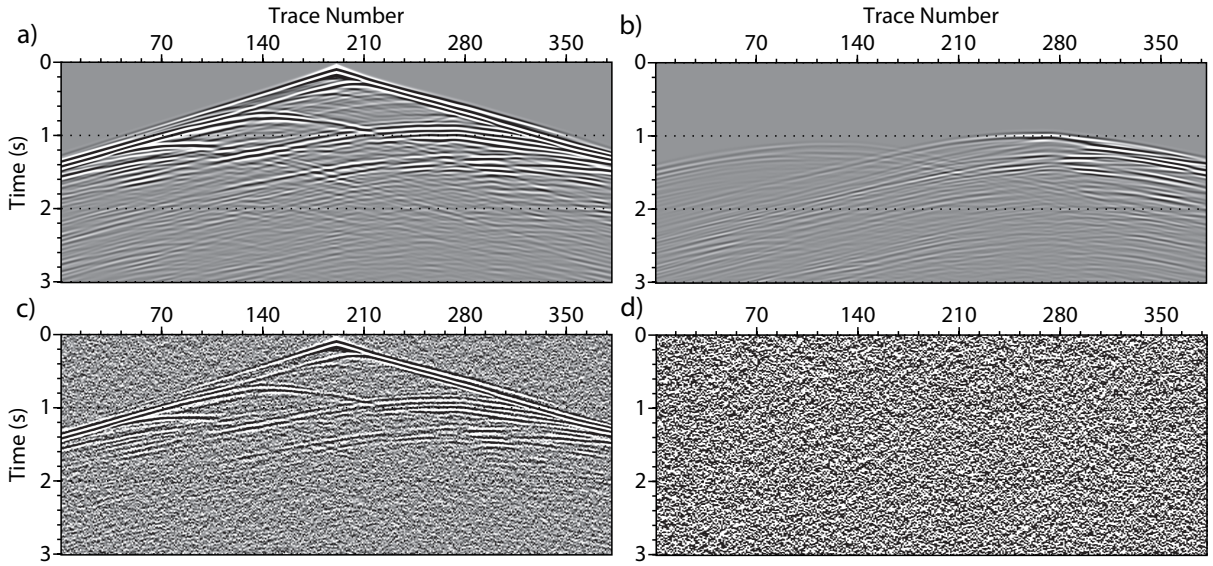


Figure 2.5: (a) The noise-free observed baseline seismograms for the source located at the center of Marmousi model; (b) the noise-free difference data between monitor and baseline datasets; (c) the noisy baseline seismograms obtained with an artificial Gaussian band-limited noise in the bandwidth of source wavelet, which has been added to the noise-free data with S/N=6 dB; (d) the noisy difference data, the time-lapse seismic events are blinded by high level of random noise. Please note that the amplitude of (b) and (d) seismograms are amplified by factor 10 to be plotted at the same scale of (a) and (c), for visualization purpose.

For this model, performing of forward modeling for all the shots takes around 15 seconds of cpu elapsed-time, using a shot-based parallelization FWI algorithm (on 77 mpi-processors for this test). Therefore, performing FWI on the baseline data takes roughly 2-3 hours of cpu elapsed-time (note that for the sequential difference or double-difference strategies due to starting from the recovered baseline model and less iteration numbers, this computational time will reduce to less than 20 minutes). Please note that our code is a prototype code and is not optimized for production.

A strongly smoothed initial model (Figure 2.6a) obtained by smoothing the true baseline model, which mimics a time-tomography velocity model based on both first arrivals and reflected events, is used as the initial model. In our study, we consider that the sonic-log measurements acquired in the two exploration wells, located at two sides of the model, provide an accurate estimation of the local vertical velocity. We build a crude prior velocity model (Figure 2.6b) based on a linear interpolation between the two well locations. However, more accurate techniques such as geostatistical approach or structural interpolation based on migration could be used. From the local velocity measurements, the interpolated prior model is accurate near the well locations and degradates with distance from well, as the structure is highly heterogeneous. Therefore, we decided to build a weighting shape whose uncertainty values follow a Gaussian function in x direction with weak values near the wells and increasing values in the center of the area. Then, this uncertainty model has been combined with a depth weighting

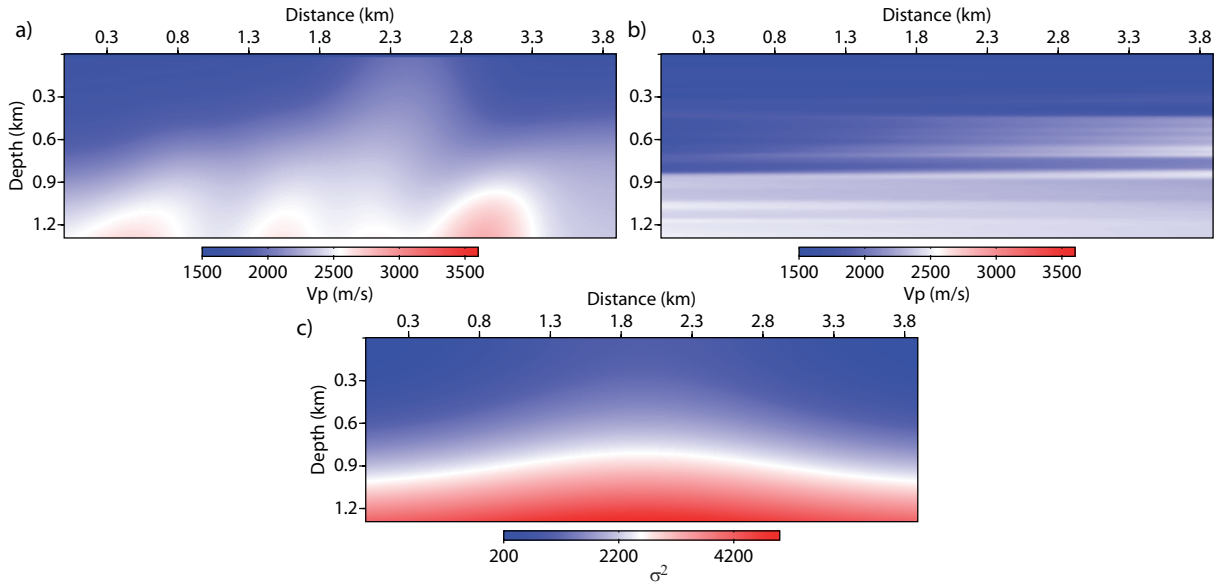


Figure 2.6: (a) The strongly smoothed initial model used for the baseline and the parallel difference inversions. (b) The prior model used for the baseline and the parallel difference inversions. This prior model is built by a linear interpolation between the two exploration wells at two sides of model and is slightly smoothed. (c) The prior model weighting σ used for the baseline and the parallel difference inversions. This model weighting contains the Gaussian function varying only in the x direction between two wells with maximum value at the center of model, then this variation is complemented by a quadratic evolution in the z direction.

(z^2), in order to have an appropriate balance between the data-misfit and prior-model terms by going into depth (it can also be seen as compensating for the propagating decay of the wave amplitude). This $\sigma^2(\mathbf{m})$ model (Figure 2.6c) is used to include the prior model into the misfit function with a spatial varying weight (recall $diag(\mathbf{W}_m^T\mathbf{W}_m) = 1/\sigma^2(\mathbf{m})$).

For selection of regularization parameters, the λ_1 value is chosen as a small value to ensure a slight smoothing of the results. To select the λ_2 hyper-parameter, the misfit function is computed for the starting model for $\lambda_2 = 1$. Based on the ratio γ between the prior-model misfit $\lambda_2\mathcal{C}_{2_m}(\mathbf{m})$ and the data-term misfit $\mathcal{C}_d(\mathbf{m})$, we adjust the λ_2 value such that $10^{-3} < \gamma < 10^{-2}$. Therefore, by selecting this reasonable ratio of prior-model and data-misfit terms, the FWI is prevented to minimize the model norm heavily at early iterations. In this test, we choose to have the ratio $\gamma = 10^{-3}$. As mentioned before, the dynamic weighting approach is performed to gradually reduce the λ_2 value during the optimization. The reader is referred to Asnaashari et al. (2013a) for more details. The stopping criterion for L-BFGS-B is based on the flatness of the misfit function for two successive iterations, ensuring a convergence. The criterion is the same for all the following inversion tests performed in this study, so that the results are comparable on the data side.

The obtained baseline model is shown in Figure 2.7a. As a quality control (QC), the normalized V_p error computed at each grid point is plotted (Figure 2.7b). The theoretical error associated to the recovered model is computed by the expression $V_p^{err} = (V_p^{true} - V_p^{rec})/V_p^{true}$

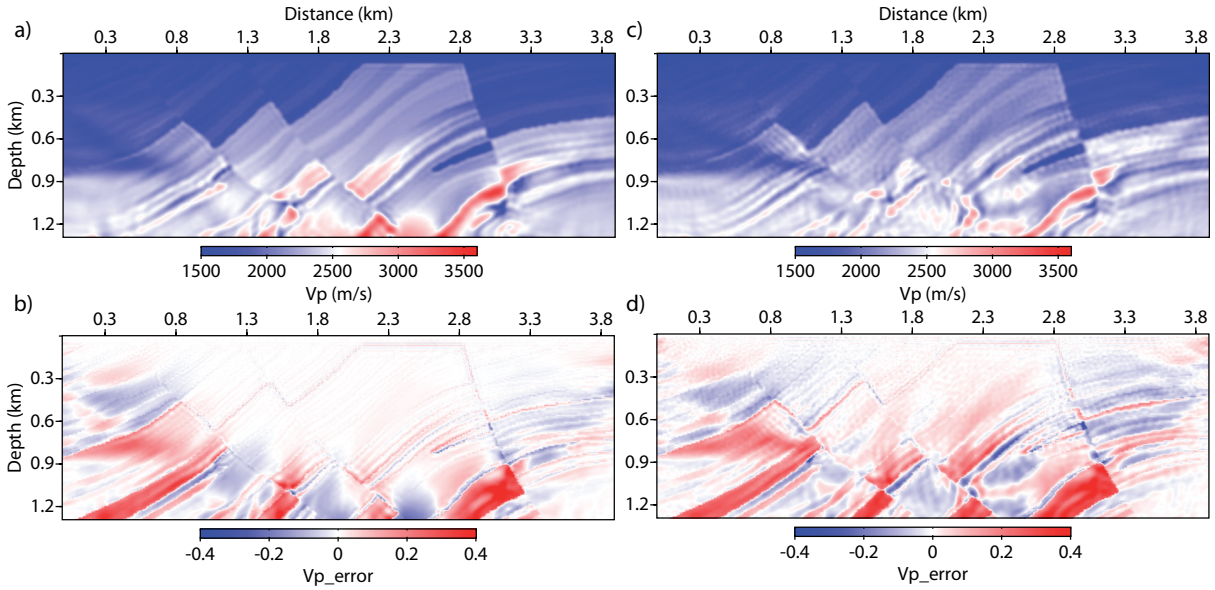


Figure 2.7: (a) The recovered baseline model by regularized FWI using the interpolated prior velocity model with inversion of noise-free data; (b) the normalized V_p error at each grid point, $V_p^{err} = (V_p^{true} - V_p^{rec})/V_p^{true}$, used as a quality control for the recovered model (a); (c) similar as (a) in case of strongly noisy data with $S/N = 4.5\ dB$; (d) the normalized V_p error for the recovered model (c).

as we know the true model.

At larger depth and close to the boundaries of the model, the error is slightly higher due to the lack of illumination. However, inside the reservoirs and in the overburden, the reconstruction of the velocity model converges to a quasi-perfect model thanks to the influence of the prior model. We insist that such a precise and accurate result cannot be obtained without considering the prior model and the dynamic approach. Adding the prior model into inversion scheme leads to an accurate baseline model which is crucial for time-lapse inversion.

2.1.5.3 Time-lapse imaging

In this part, we perform a robustness study of the three proposed time-lapse imaging strategies (parallel difference, sequential difference, double-difference), particularly in the presence of ambient noise. In the first investigation, we focus on the noise-free case using the conventional modes (*i.e.* without target-oriented option). For the parallel difference strategy, we invert the monitor dataset with the same configuration and inversion parameters as for the baseline reconstruction, starting from the same smooth initial model, using the same interpolated prior velocity model, and the same parameters of the dynamic approach. Figures 2.8a and 2.8b show the time-lapse model obtained by the parallel difference strategy, with two vertical QC logs passing through the two gas reservoirs. The result shows that inverting both data independently from the same initial model does not make it possible to focus only on time-lapse changes. The two reservoir variations are approximately recovered, but many high amplitude anomalies

associated to inversion artifacts are present in the final time-lapse model. This problem is mainly linked to the ill-posedness and the non-uniqueness issues of the inverse problem. Since the complexity of data is not the same for the two datasets, the two independent inversions do not converge to similar models and do not have the same path of optimization. Thus, different artifacts are created for the two inversions. In terms of computational costs, the parallel difference strategy is not interesting, because both inversions start from the smooth initial model.

Once the recovered baseline model is obtained, we can perform the sequential difference and the double-difference strategies starting from this baseline model. Since the expected monitor model must be close to the baseline model, due to the small and localized time-lapse variations, the recovered baseline could also be a good candidate for the prior model in both strategies. This is the case in this study. In the conventional mode (without target-oriented option), a homogeneous weighting is applied to the model, with a low influence of the prior model penalty term in the objective misfit function. The sequential difference is able to detect the time-lapse variation, but it cannot focus solely on the target areas (Figures 2.8c and 2.8d). The result is less contaminated by inversion artifacts as compared to the parallel difference strategy. Figure 2.8e shows the time-lapse result produced in the noise-free case with the conventional double-difference strategy. It appears that both target zones are very well recovered and that most of the difference data energy is focalized inside the correct perturbation zones, although some low-value anomalies exist below 1 km depth at the center of the model. This may be due to small inaccuracies in the reconstructed baseline model, that are not completely recovered in this part mostly due to the lack of illumination. However, such time-lapse incorrect anomalies may be considered as acceptable. For the small reservoir, since its thickness is smaller than the spatial wavenumber, there is a small shift in the position of the top reservoir (please note that the maximum frequency included in the time-domain inversion is around 25 Hz).

For the target-oriented inversion (in the following), we need to design the model weighting matrix $\mathbf{W}_m$ based on the location of expected changes. Based on non-seismic prior information, well-known the depth of reservoirs, and the migrated section of difference dataset, it is possible to roughly detect the area of time-lapse changes. After detecting the approximate location of variations, the prior uncertainty model $\sigma^2(\mathbf{m})$ (Figure 2.9) is built by a summation between a small constant uncertainty value everywhere in the model and two Gaussian functions varying in x and z direction which are centered roughly at the center of variations. These two Gaussian functions, which contain high relative values with respect to the small values of other places, present the areas where we expect to be updated during the target-oriented inversion. This is a roughly design of the target-oriented prior uncertainty model, however more precise techniques could be used in real case applications. We use this target-oriented weighting model for all following target-oriented inversions.

As the parallel difference may not be very interesting anymore, in the following we will concentrate on the two other strategies.

2.1.5.4 Local resolution analysis

In this section, we want to analyze the different behaviors of the two time-lapse strategies, in conventional and target-oriented modes. For that, we perform a local resolution analysis. The Hessian function is related to the point-spread functions (PSFs) or the spike tests; this relation is important information brought by the Hessian in a local resolution analysis. Spike

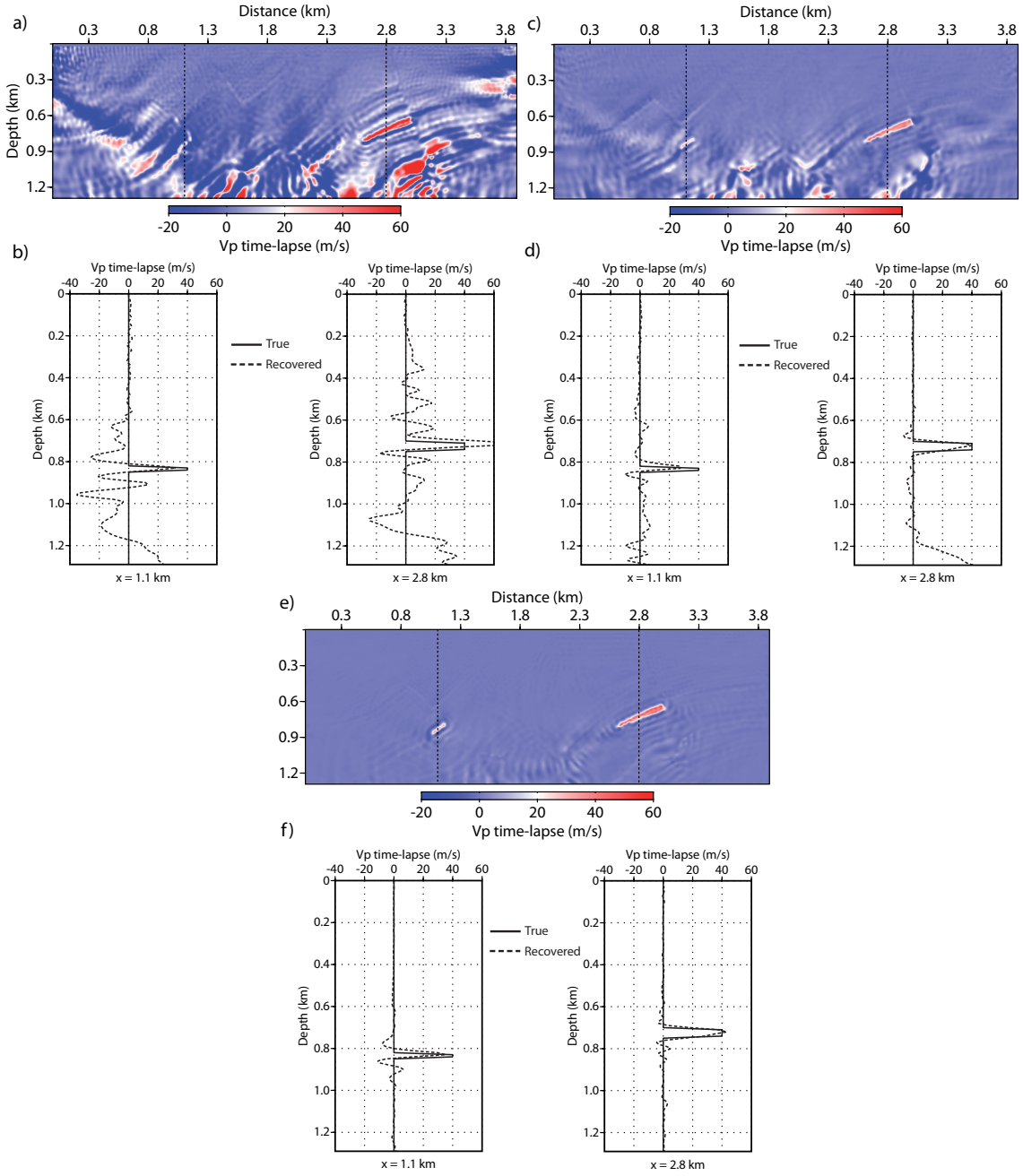


Figure 2.8: The recovered time-lapse V_p models by regularized FWI of noise-free data and two QC vertical logs passing through the two target areas at $x = 1.1\ km$ and $x = 2.8\ km$: (a) using the parallel difference strategy; (b) two vertical logs corresponding to the model (a); (c) using the sequential difference strategy; (d) two vertical logs corresponding to the model (c); (e) using the double-difference approach; (f) two vertical logs corresponding to the model (e).

tests are commonly used as a diagnostic tool for resolution analysis in linearized tomographic problems (Menke, 1984; Fang et al., 2010; Fichtner and Trampert, 2011). They constitute a

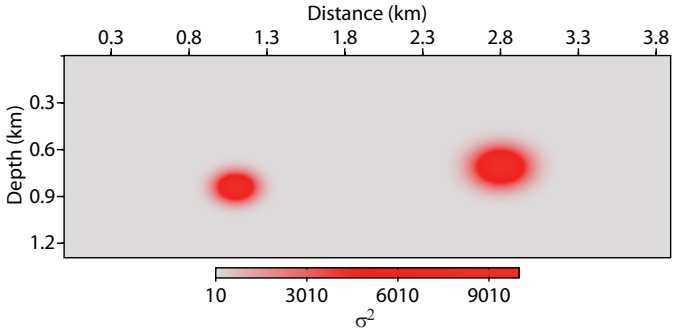


Figure 2.9: The target-oriented σ model is built based on the location of time-lapse changes coming from the migrated section, the available prior well data and geological model.

numerical way for estimating the effects of the off-diagonal terms of the Hessian function on the recovering model parameters, for one row and one column of the Hessian matrix related to the point model perturbation.

We shall consider a true monitor model which contains only two point time-lapse perturbations that are point-localized inside the two reservoirs. In a first investigation, without any target-oriented prior weighting, the time-lapse images (two spike functions) are recovered by the inversion from three different configurations: 1) The time-lapse inversion (eighter sequential difference or double-difference) starting from the true baseline model used as the starting monitor model, shown in Figure 2.10a; 2) the sequential difference strategy starting from the recovered baseline model (Figure 2.7a) used as the starting monitor model, shown in Figure 2.10b; and 3) the double-difference strategy starting from the recovered baseline model (Figure 2.7a) used as the starting monitor model, shown in Figure 2.10c. The results are shown here after 10 iterations. The spikes recovered by the double-difference strategy (from the recovered baseline model) appear to be very similar to the ones obtained with the true baseline model, showing the robustness of the strategy with respect to the starting model. In addition, the total energy of the time-lapse perturbations is better located and better focused at their positions with the double-difference method, as compared to the sequential difference strategy.

When using the true baseline model as the initial model for time-lapse inversion (Figure 2.10a), the two strategies bring quite identical results, because the calculated baseline data $\mathbf{d}_{calc_b}$ are equivalent to the observed baseline data $\mathbf{d}_{obs_b}$, so the composite dataset and the observed monitor data are identical (for noise-free data). Let us clarify the differences between the composite and the monitor dataset, when the recovered base model is used. The monitor dataset can be written as the time-lapse perturbation data plus the baseline data (simply by mathematical adding and subtracting the baseline data to and from the monitor data),

$$\mathbf{d}_{obs_m} = \mathbf{d}_{obs_m} - \mathbf{d}_{obs_b} + \mathbf{d}_{obs_b}. \tag{2.6}$$

The expression is similar to the one for the composite data (2.5), but the calculated recovered baseline data has been replaced with the observed baseline data. The misfit data computed at the first iteration (when $\mathbf{d}_{calc_m} = \mathbf{d}_{calc_{rec-b}}$) of optimization for the two strategies are expressed as

$$\Delta\mathbf{d}_{double-difference} = \mathbf{d}_{obs_m} - \mathbf{d}_{obs_b}, \tag{2.7}$$

$$\Delta \mathbf{d}_{sequential} = (\mathbf{d}_{obs_m} - \mathbf{d}_{obs_b}) + (\mathbf{d}_{obs_b} - \mathbf{d}_{calc_{rec-b}}). \tag{2.8}$$

Equation 2.7, which follows from equation 2.4, and equation 2.8, derived from equation 2.6 and the data misfit equation $(\mathbf{d}_{obs_m} - \mathbf{d}_{calc_m})$, reveal the main differences between the sequential difference and the double-difference strategies. The sequential difference strategy, in addition to the time-lapse data (first part of the expression, common with the expression for the double-difference approach), tries to recover the misfit related to the baseline events that have not been fully reconstructed before (second part of the expression). This is the drawback of the sequential difference strategy, that leaves a potential risk of creating spurious time-lapse variations. However, the double-difference strategy focuses on inverting difference data and requires as a starting model a baseline model which accurately enough describes the background baseline kinematics and structures, otherwise the time-lapse energy cannot be localized and focalized properly at the correct positions.

We should mention that the problem of sequential difference is not related to a convergence of inversion in a wrong direction. This issue rather appears as a mixing of two different effects: the updating of baseline structure and the recovering of time-lapse changes. If it would be possible to qualify the time-lapse variations with another smart way rather than a simple subtraction between the two obtained models (baseline and monitor), we could better distinguish 4D changes from other effect. In the following, as we rely on model-difference to build the 4D variations, the baseline improvement during the monitor inversion will be considered as 4D artifact.

In a second investigation, the same tests are performed, with the baseline model as a prior model (*i.e.* the true baseline model used as the prior model in the test of Figure 2.10d and the recovered baseline model used as the prior model in the tests of Figures 2.10e and 2.10f) and applying the target-oriented prior model weighting (Figure 2.9). The results are shown in Figure 2.10 (bottom panel). The target-oriented inversion prevents the apparition of artifacts outside of the expected target areas, especially for the sequential difference strategy, more prone to artifacts. For both time-lapse strategies, target-orienting the inversion can significantly improve the results compared to the conventional inversion (top and bottom panels of Figure 2.10). By applying a stronger model weighting outside the target zones, the second term of the sequential-difference data misfit (2.8) (*i.e.* $\mathbf{d}_{obs_b} - \mathbf{d}_{calc_{rec-b}}$) plays a minor role in driving the inversion procedure. The model constraints outside the target do not allow these non-fully fitted baseline data to update the monitor model. For target-oriented inversion, the dynamic prior weighting cannot be applied and the prior regularization parameter λ_2 has to be kept constant until the end of optimization.

2.1.5.5 Noisy data S/N=6 dB

In this section, we study the robustness of the two aforementioned time-lapse strategies in presence of random noise. In particular, we present a sensitivity analysis of the time-lapse models obtained by the sequential difference strategy and the double-difference strategy with respect to the inaccuracy of the recovered baseline model. How accurate should the baseline model be? This is the question that is addressed in this section.

An artificial Gaussian noise in the range of $1-25\ Hz$ (the bandwidth of the source wavelet) has been added to the true noise-free data (we have used the 'suaddnoise' procedure of Seismic Unix (Cohen and Stockwell, 2008)). The signal-to-noise ratio is around 6 dB. Figures 2.5c

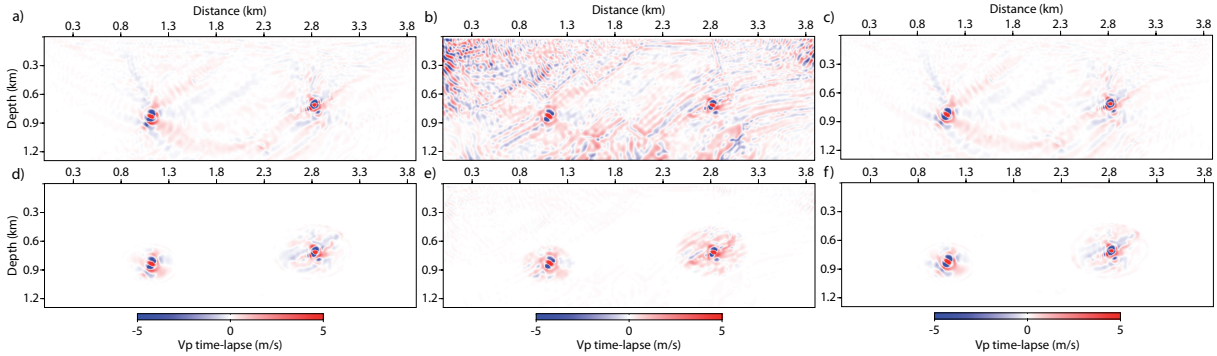


Figure 2.10: Local resolution analysis with spike tests, top panel without target-oriented inversion: (a) using time-lapse inversion (sequential difference or double-difference) starting from the true baseline model (Figure 2.4a); (b) using the sequential difference strategy starting from the recovered baseline model (Figure 2.7a); (c) using the double-difference strategy starting from the recovered baseline model (Figure 2.7a); (d), (e), and (f) similar as (a), (b), and (c) in case of using the target-oriented model weighting (Figure 2.9).

and 2.5d show the noisy baseline and the difference seismograms, respectively. The difference seismograms are amplified by a factor of 10 to be plotted at the same scale as the baseline seismograms. Clearly, the low-energy time-lapse signal is below the level of random noise. First of all, the noisy baseline data are inverted by the dynamic regularized FWI including the prior model built by well logs (Figure 2.6b). The recovered baseline model and its corresponding error map are illustrated in Figures 2.11a and 2.11b. As expected, the result with noisy data is less accurate compared to the result for the noise-free case, since the inverse problem becomes more ill-posed. However, using the prior model helps to fill in the lack of low wavenumbers and ensures that the inversion converges to a more robust model. To perform the sensitivity analysis, in addition to this final recovered model, another baseline model with less accuracy is used. One of the intermediate updated models, obtained during optimization but before reaching the convergence, is chosen (Figure 2.11c).

2.1.5.6 Sensitivity to the inaccuracy of baseline model

We perform four different inversions using the conventional sequential difference and the conventional double-difference strategies starting from the intermediate recovered baseline model (Figure 2.11c) and from the final obtained baseline model (Figure 2.11a). The results are shown in Figure 2.12. The double-difference strategy delivers almost the same results in both cases and the time-lapse perturbations are reconstructed quite well. Nevertheless, the result obtained with a more accurate baseline model is more precise, especially for the smaller reservoir. The results of the sequential difference are completely different when different baseline models are used as the initial monitor model. When the accurate baseline model is chosen, this method converges to similar result as the double-difference one. However, there are some small anomalies below 1 km depth and the smaller reservoir is less recovered. When the inaccurate baseline model is used as the initial model, the time-lapse model obtained by the sequential

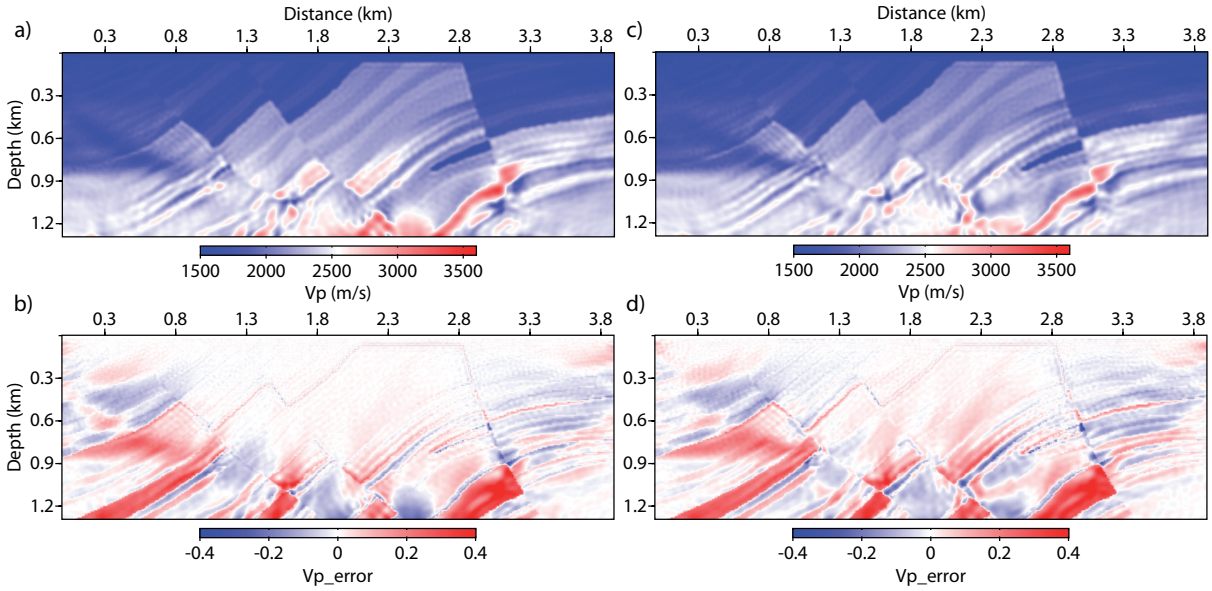


Figure 2.11: The recovered baseline model by regularized FWI using the interpolated prior velocity model with inversion of noisy data $S/N = 6\ dB$: (a) the final obtained model; (b) the normalized V_p error at each grid point, $V_p^{err} = (V_p^{true} - V_p^{rec})/V_p^{true}$, used as a quality control for the recovered model (a); (c) the intermediate updated model during optimization with less accuracy compared to (a); (d) the normalized V_p error for the recovered model (c).

difference strategy is affected by several artifacts, inside and outside the target zones. These artifacts are caused by the non-fully explained events of the baseline data ($\mathbf{d}_{obs_b} - \mathbf{d}_{calc_{rec-b}}$), and FWI keeps updating the model in the time-lapse inversion due to these residuals. The sequential difference strategy attempts to recover the parts of the model which have not been reconstructed during the baseline reconstruction step. Therefore, this strategy can only be attractive when one has confidence in the baseline model reconstruction, which would be difficult case on real application. On the other hand, the double-difference strategy is less sensitive to the inaccuracy of the recovered base model.

2.1.5.7 Strongly noisy data S/N=4.5 dB

What will be the efficiency of our different strategies for strongly noisy data? Does the added noise change the behavior of the two time-lapse strategies in the conventional mode and the target-oriented mode? Same as before, an artificial Gaussian noise in the range of $1 - 25\ Hz$ has been added to the true noise-free data with signal-to-noise ratio around $4.5\,dB$. The noisy baseline data are inverted by the dynamic regularized FWI including the prior model. The recovered baseline model and its corresponding error map are illustrated in Figures 2.7c and 2.7d. It appears that the baseline model is not recovered accurately, even at shallow depths, due to the high level of noise.

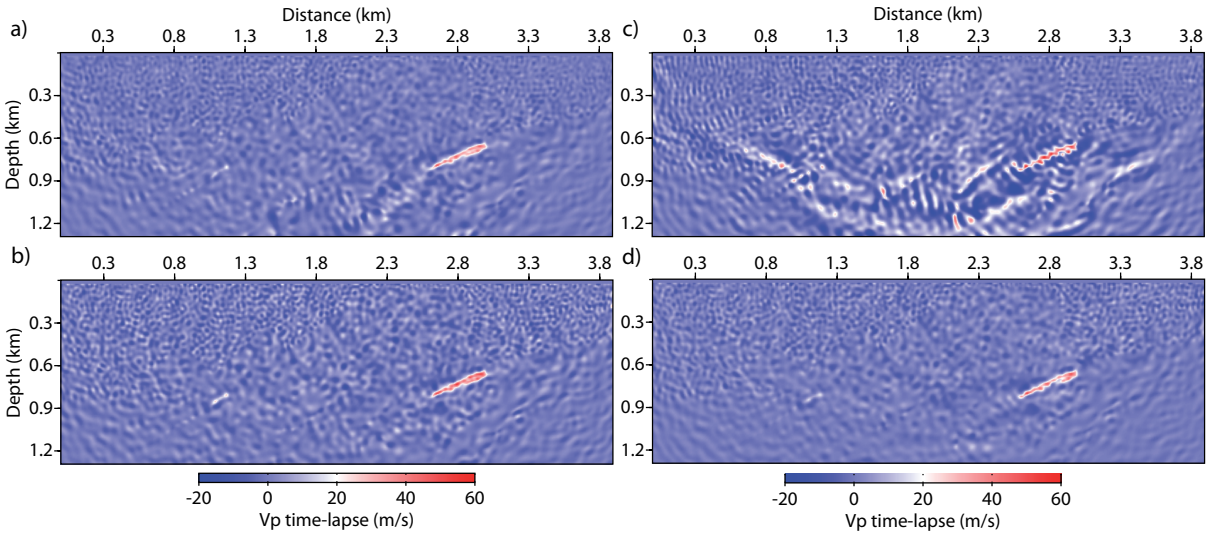


Figure 2.12: Sensitivity of time-lapse models with respect to the inaccuracy of baseline model, in case of noisy data $S/N = 6\ dB$, time-lapse models obtained by: (a) the sequential difference strategy starting from the final recovered baseline model (Figure 2.11a) as the starting model; (b) the double-difference strategy starting from the final recovered baseline model; (c) the sequential difference strategy starting from the intermediate recovered baseline model (Figure 2.11c) as the starting model; (d) the double-difference strategy starting from the intermediate recovered baseline model.

2.1.5.8 Conventional time-lapse inversions

In a first investigation, the conventional sequential difference and double-difference strategies are tested in this strongly noisy environment (*i.e.* without target-oriented model weighting). Figure 2.13 shows the results of these inversions. In the conventional mode, the result of the double-difference strategy is more robust than the one from the sequential strategy: however, in both cases, most of the recovered model is dominated by the uncorrelated noise. The result of the sequential difference strategy is contaminated by the artifacts so that it is difficult to interpret even the location of the real time-lapse variations. To improve this result, we propose to apply the sequential difference strategy in target-oriented mode with an appropriate model weighting matrix.

2.1.5.9 Target-oriented time-lapse inversions

In a second investigation, the two time-lapse strategies are applied in target-oriented inversion mode. The results are shown in Figure 2.13. The target-oriented inversions can prevent the apparition of perturbations outside the expected target zones: most of the image noise artifacts and artifacts due to continuing baseline updates (in sequential difference case) are removed, thanks to the target-oriented weighting matrix $\mathbf{W}_m$. Therefore, the sequential difference strategy in target-oriented mode behaves similar to the double-difference strategy and can deliver

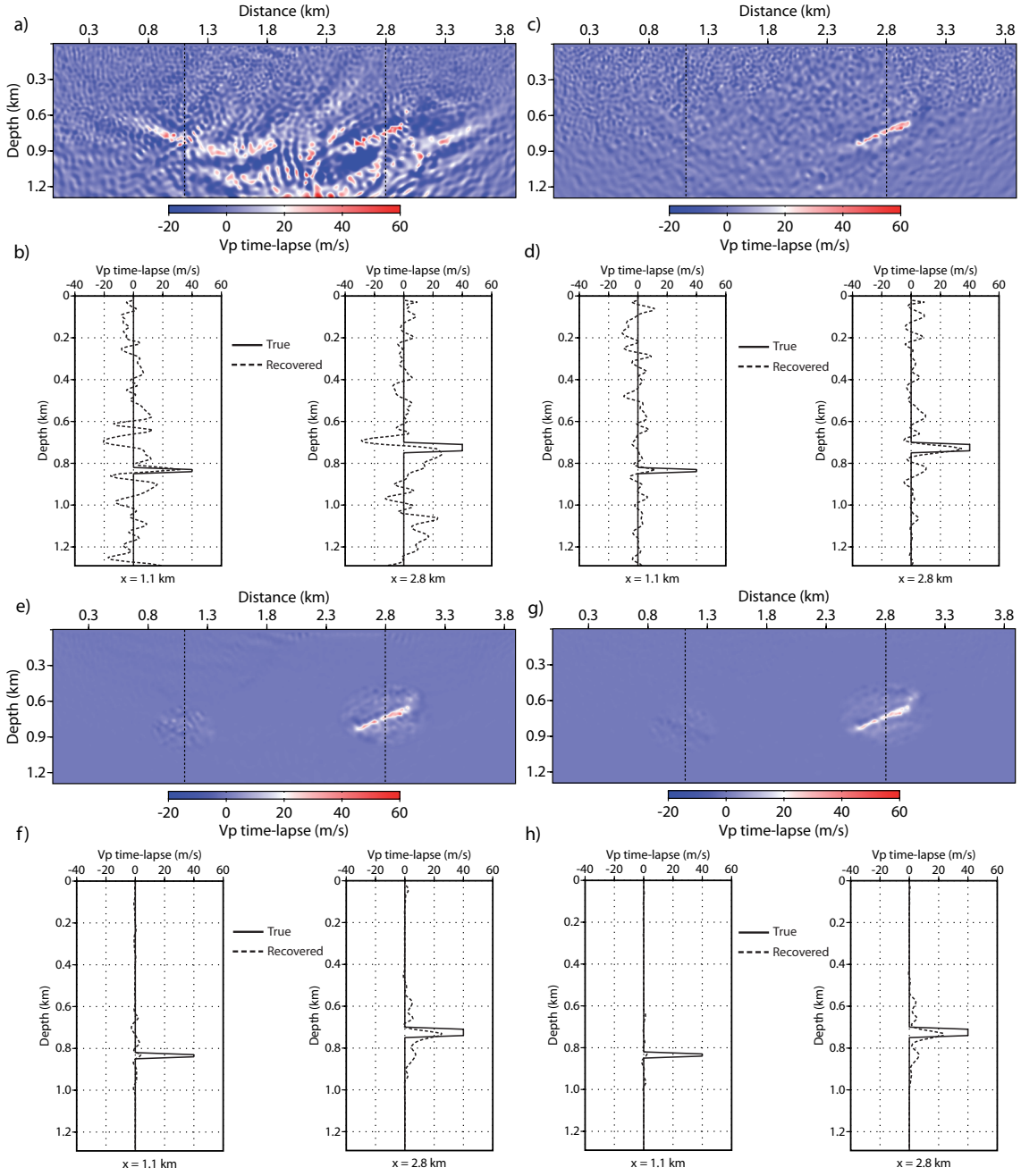


Figure 2.13: The recovered time-lapse V_p models by regularized FWI of the strongly noisy data ($S/N = 4.5\ dB$) and two QC vertical logs passing through the two target areas at $x = 1.1\ km$ and $x = 2.8\ km$ using: (a) the conventional sequential difference strategy; (b) two vertical logs corresponding to the model (a); (c) the conventional double-difference strategy; (d) two vertical logs corresponding to the model (c); (e) the target-oriented sequential difference strategy; (f) two vertical logs corresponding to the model (e); (g) the target-oriented double-difference strategy; (h) two vertical logs corresponding to the model (g).

the same robust results as the double-difference method, if some information exists on the location of the expected perturbed zones.

In this case, the recovered monitor model is almost frozen and is not allowed to be updated outside of the target zones, due to the prior model and model weighting matrix. Therefore, the unexplained baseline events cannot contaminate the reconstruction of the time-lapse perturbations. In this case, the velocity variations in the larger reservoir are better recovered and better positioned with the sequential difference strategy in target-oriented mode, compared to the conventional mode (see Figures 2.13b and 2.13f). Please note that another numerical tests have shown that less accurate knowledge of the target positions in the prior weighting (wider expected target area) still leads to robust results, compared to the conventional approach. Clearly, the smaller reservoir change is missed in all inversion sequences because of the high level of noise, the small size of the perturbation area, and the inaccuracy of the baseline model.

In some situations where the double-difference method is not applicable, for example non-perfectly matched acquisition geometries, we must consider the sequential difference strategy for time-lpase inversion. By performing this strategy in target-oriented mode, it is possible to obtain a time-lapse result as robust as that would be delivered by the double-difference strategy. In this case, we can focus only on the time-lapse areas and reduce the drawbacks of the conventional sequential strategy.

2.1.6 Discussion

Nowadays, techniques such as full waveform inversion and amplitude versus offset (AVO) increase our capabilities for better characterizing reservoir changes spatially and quantitatively. We have shown in this paper that the double-difference and the sequential difference strategies could be promising techniques for time-lapse imaging, however, with some advantages and drawbacks.

When the quality of data is acceptable in terms of signal-to-noise ratio, the double-difference strategy can be more interesting than the sequential one, since it can focus on the target areas. In the case of very strongly noisy data, the double-difference strategy is more sensitive to the non-repeatable noise at the subtraction step. The subtraction between two datasets increases the standard deviation of noise, and the amplitude of noise overwhelms the low-energy true time-lapse signal. Therefore, the time-lapse energy cannot be detected by the inversion scheme and the inversion is strongly driven by the noise. This is what happens for the smaller reservoir in a strongly noisy environment in our synthetic examples. Making the crude subtraction between two datasets may not be a good idea in a noisy environment and we need other ways to preserve small time-lapse responses in the middle of a high level of noise: pattern recognition techniques, etc., for instance. In this case, the target-oriented sequential strategy could be an alternative. In the sequential strategy, there is no subtraction between two datasets, therefore the standard deviation of random noise will not change and it makes this strategy less sensitive to the random noise. On the other side, by including a target-oriented prior model in this strategy, we can make the sequential strategy to focus on the time-lapse variation areas. Therefore, the target-oriented sequential strategy can behave similarly to the double-difference strategy, but without its sensitivity to noise. In addition, we benefit from the applicability of the sequential strategy to the non-perfectly matched acquisition geometries. In all cases, adding the prior information into the time-lapse inversion has been found to be really helpful.

For double-difference, it is essential to have perfectly matched surveys (e.g. permanent sensors or OBC configurations) to perform subtraction between datasets, if not, it is possible to perform extra pre-processing steps and make data interpolation to match two surveys.

2.1.7 Conclusions

We have studied the robustness of three strategies for time-lapse imaging with regularized FWI in noise-free and noisy environments. Using the Marmousi synthetic case, it has been shown that the parallel difference strategy is highly sensitive to the differential artifacts in the images coming from two independent inversions. The sequential difference strategy can be more attractive when the baseline model is accurately recovered, because this method attempts to recover the parts of the model which have not been fully reconstructed before. Therefore, this method is more sensitive to the inaccuracy of the recovered baseline model. In case of a low level of noise, the double-difference strategy is more robust than the sequential one, since the double-difference inversion only focuses on the difference (time-lapse) data. As shown with local resolution analysis, the difference data are better localized and better focalized at their correct positions using the double-difference strategy. In addition, we have illustrated that the double-difference strategy is less sensitive to the inaccuracy of the reconstructed baseline model used as initial model.

We also propose a target-oriented time-lapse imaging with regularized FWI including prior model: in the misfit function, we apply strong prior model constraints outside the areas of expected 4D changes and we apply relatively smaller prior constraints in the target zones. The target-oriented inversion steers the recovery of the monitor model towards the target areas, where the smaller prior model constraints give more relative weight, in the misfit function, to the data-misfit term. This target-oriented option could be performed as a final focused strategy in real application, if the prior information exists on the location of expected time-lapse variations. In the application to the Marmousi dataset, the local resolution analysis with spike tests shows that the target-oriented inversion prevents the apparition of artifacts outside the target areas. In the strongly noisy case, the target-oriented weighting matrix $\mathbf{W}_m$ leads to a similar behavior for the sequential and the double-difference time-lapse strategies and delivers more robust images. In particular, for the sequential difference approach, the unexplained baseline events cannot contaminate reconstruction of the time-lapse perturbations and the time-lapse variations are better recovered and better positioned with the target-oriented inversion mode than with the conventional one. The next step will be the analysis of the reconstruction capacities of these strategies on real field datasets.

2.1.8 Acknowledgements

We would like to thank TOTAL Exploration & Production and SEISCOPE consortium for supporting this study. This work was performed by access to the high-performance computing facilities of TOTAL (Rostand), CIMENT (Université de Grenoble, France), and GENCI-CINES under Grant 2012-046091. We would like to gratefully thank the Associate Editor, Milana Ayzenberg, Yunsong Huang and an anonymous reviewer, for their very constructive comments on the manuscript.

2.2 Supplementary tests of target-oriented inversion

In real applications, it may be difficult to identify the area where time-lapse variations are expected to occur, as illustrated in the paper where two small and well-separated targets were considered. Consequently, it would be better to have a larger (wider) target area which is extended in horizontal axis along the reservoir. We assume here that the depth of reservoir is known, but that we have no prior information about potential lateral extension. Hence, we can design the target-oriented prior model weighting in the manner illustrated by Figure 2.14 for the Marmousi application.

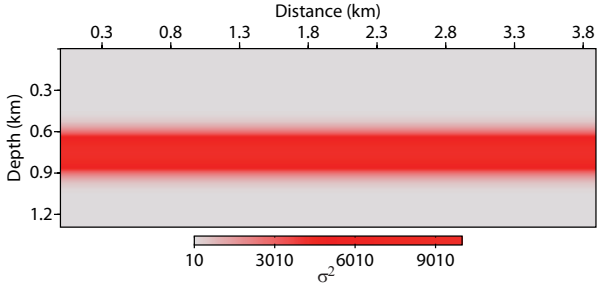


Figure 2.14: Another kind of target-oriented $\mathbf{W_m}$ which includes prior information where only the depth of reservoirs is known.

This target-oriented weighting model is applied to the strongly noisy data ($S/N = 4.5\ dB$) using the sequential difference and the differential strategies. Clearly, the obtained time-lapse results appear improved compared to the conventional inversion (without any target-oriented weighting), even when taking into account the larger target area (compare Figures 2.15a and 2.15b with Figures 2.13a and 2.13c). The time-lapse variation corresponding to the large reservoir is properly recovered in the target-oriented sequential difference strategy. For both target-oriented inversions, we obtain almost similar results. However, the uncorrelated noise is also reconstructed in the rectangular target area for both strategies, sequential and differential. Even with this wider rectangular shape for target weighting, the sequential difference strategy is forced to invert only the misfit related to the time-lapse signals. In fact, by including the target-oriented prior weighting into the inversion scheme, the objective function shape will change and monitor inversion is forced to stay inside the local minimum around the baseline model point. In other words, reducing the size of model space for the time-lapse inversion, implies that the new objective function has less local minima.

2.3 Application to the Dai model

In this section, we apply our time-lapse inversion to another synthetic and realistic dataset. Dai et al. (1995) have investigated the influence of steam injection onto a reservoir. We design a similar model, named here as the Dai model, with slight changes such as curved and tilted layers in order to consider more realistic media than only a model with flat layers, as the original Dai model investigated. The true P-wave velocity models for our acoustic time-lapse analysis are shown in Figure 2.16. Baseline model consists of eight layers; all of them being saturated with water except the sixth one which is saturated with oil. After steam injection into this

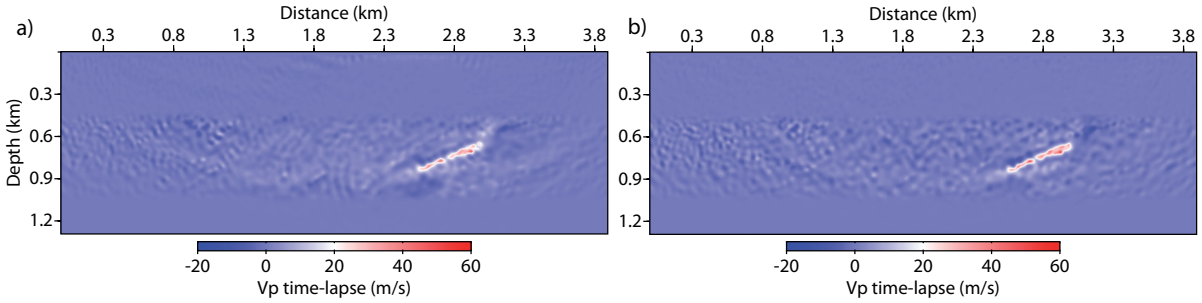


Figure 2.15: Time-lapse results obtained with new target-oriented prior model weighting: (a) the sequential difference strategy, (b) the differential strategy. Even with taking into account less accurate target-oriented prior information, the target-oriented inversion provides an improvement of time-lapse variation as compared to the conventional inversion result.

layer, the heat leads to two concentric areas, one saturated with steam and the outer one with heated oil. The V_p velocity values were computed considering a poroelastic approach taking into acount the rock physic and fluid parameters (Dupuy, 2011) and then the seismograms are generated by an acoustic modeling.

2.3.1 Sensitivity of recovered baseline model with respect to initial and prior models

A sensitivity study on this synthetic model has been done. In this synthetic test, the accuracy of baseline models obtained by regularized FWI with respect to different initial and prior models is studied.

The acquisition configuration considers 27 sources, located along a horizontal line with interval of 25 m, below the surface at a depht of 75 m. We consider one line of receivers at the same depth as sources and two additional lines of receivers inside two vertical boreholes at the horizontal positions $x = 75\ m$ and $x = 725\ m$ with a spacing of 12.5 m between sensors. We do not consider sources in wells, but receivers could be installed permanently in wells for time-lapse imaging and will increase dramatically our illumination for the baseline reconstruction. An explosive Ricker source with a central frequency of 20 Hz is used for all shots. The time seismograms are generated using an acoustic finite-difference modeling in time domain with a fourth-order stencil in space and a second-order integration in time (Levander, 1988). Perfectly-Matching-Layer (PML) absorbing boundaries condition are set all around the model, even at the zero-level surface. Figure 2.17 shows seismograms of the shot located at the center of the source line before steam injection (baseline model).

Surface receivers record direct and reflected waves but no diving waves because of the short offsets used, while the receivers inside wells record the reflected and the transmitted waves. Therefore, considering the receivers inside wells could mimic an enlargement of the horizontal offset range, and could increase our spatial wavenumber sampling.

In this section, we only focus on the reconstruction of the baseline model. We analyze the effect of initial models and the impact of prior information on a constrained inversion. In particular, we study how the accuracy of the reconstructed models could be improved thanks

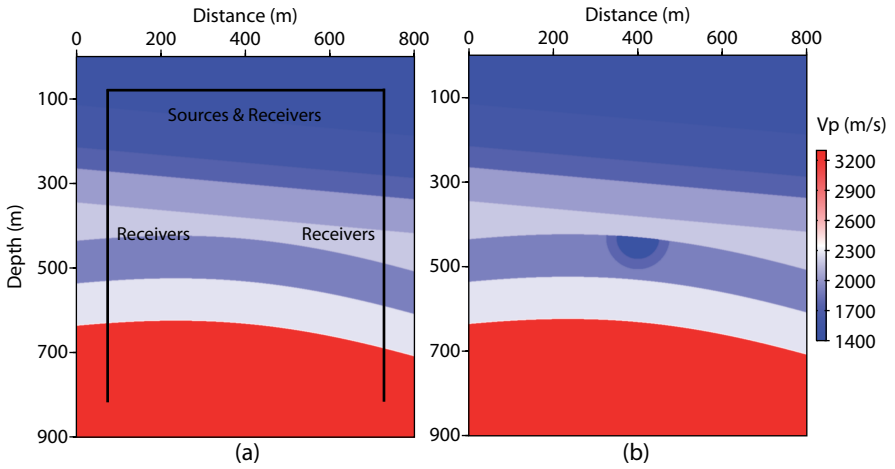


Figure 2.16: The true synthetic models and acquisition geometry: (a) baseline model, before steam injection, (b) monitor model, after injection.

to prior information, which mitigates the lack of illumination. We perform sensitivity tests using three different quality of initial models. For each initial model, three tests are performed without prior and with two different types of prior model, leading to nine configurations.

The three initial models are built from a smoothed version of the true baseline model, using different smoothness factors. A Gaussian smoothing function is used. The less smoothed one with a smoothness length of 150 m (in both directions) is referenced as m_{0a} (Figure 2.18a). A highly smoothed version (smoothness length of 300 m)) is referenced as m_{0c} (Figure 2.18c). This model does not ensure a correct kinematic consistency of all arrivals. A model of intermediate quality m_{0b} is also considered with a smoothness length of 250 m (Figure 2.18b). In this time-lapse application, we consider to have two acquisition wells and associated sonic-log that give an accurate estimation of local velocity. Based on this sonic-log prior information, we build two different prior models for inversion: a first prior model m_{p1}, called smooth prior model, combines local well measurements near the well positions and the smooth background of the starting model (Figure 2.19a). A second prior model m_{p2}, called 1D prior model, is built by horizontal linear interpolation between the measured velocity values inside wells (Figure 2.19b).

The formulation suggested by Asnaashari et al. (2013a) is used for inversion (equation 1.29),

$$\mathcal{C}(\mathbf{m}) = \mathcal{C}_d(\mathbf{m}) + \lambda_1 \mathcal{C}_{1_m}(\mathbf{m}) + \lambda_2 \mathcal{C}_{2_m}(\mathbf{m}). \tag{2.9}$$

This equation has been explained in details in Chapter 1.

The model weighting $\mathbf{W}_m$, which allows to weight the penalty associated to the model residual $(\mathbf{m} - \mathbf{m}_p)$, is selected as a diagonal matrix. For this example, the uncertainties are weak near wells where sonic-log measurement have been done, while increasing as one moves away from wells. The shape of this uncertainty is built with a Gaussian function varying in x direction, with maximum uncertainty in the center and minimum uncertainty near wells. Then, this uncertainty model is combined with a depth weighting (Figure 2.19c). This final σ model is used to compute the model weighting $(diag(\mathbf{W}_m^T \mathbf{W}_m) = 1/\sigma^2(\mathbf{m}))$.

The regularization parameter λ_2 has been chosen in such a way that having a ratio between

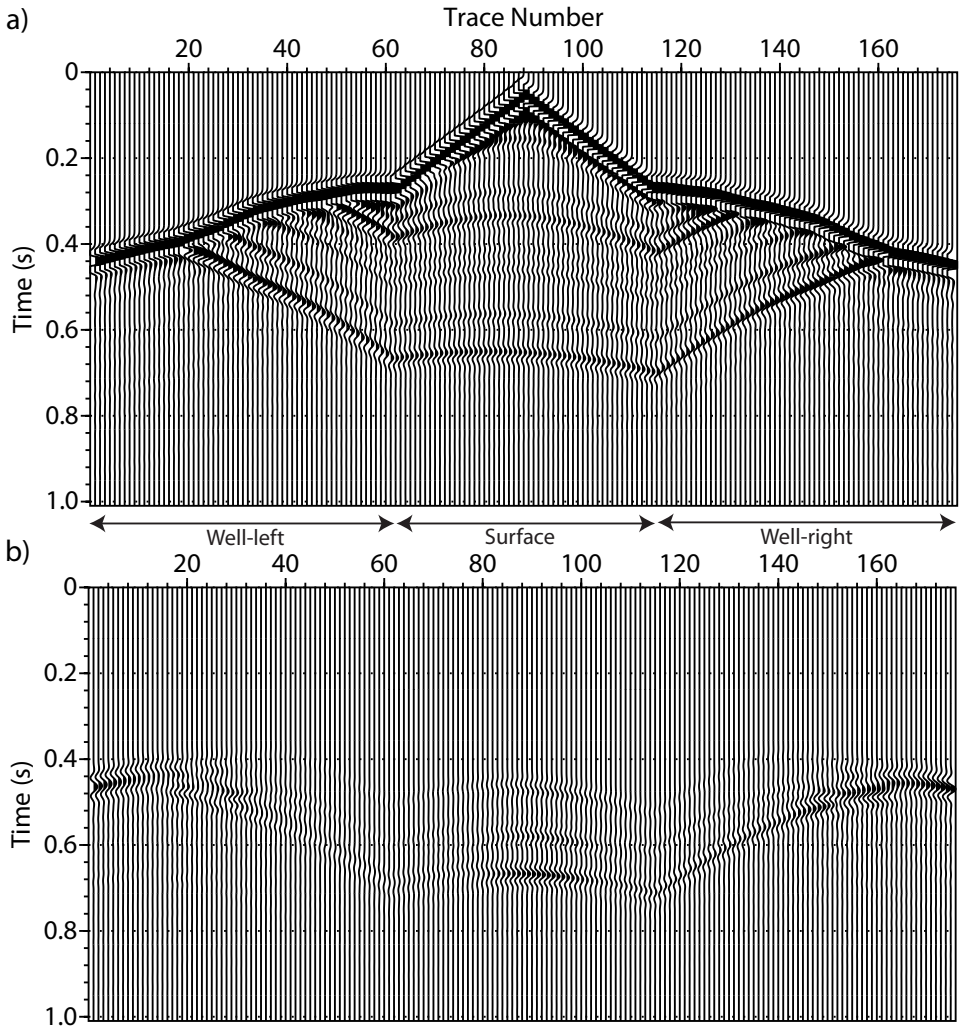


Figure 2.17: The seismogram of pressure data for the source located at the center of model $x = 400\ m$; the first 61 traces related to receivers inside well located at the left part of model, trace numbers 62-114 corresponding to surface receiver line, and 115-175 correspond to receivers inside right well: (a) baseline dataset, (b) differential dataset due to the injected steam.

prior-model and data terms of objective function around 1×10^{-2} at first iteration, in case of the m_{oc} starting model. Then λ_2 are kept fixed for others tests. For all cases, Tikhonov regularization parameter λ_1 is chosen as a small at value before numerical instabilities.

The baseline dataset is first inverted through a regularized FWI without any prior penalty ($\lambda_2 = 0$) using the three initial models. The recovered models for each starting model and one associated vertical profile at position $x = 400\ m$ are shown in Figure 2.20. As expected, the recovered models from starting models m_{0a} and m_{0b} allows to converge toward acceptable and similar solutions, as the kinematic of arrivals is preserved. For model m_{0c}, inversion converges toward a local minimum in the right part where illumination is weaker, and due to kinematic issue, that can lead to cycle-skipping issues. This test illustrates that poorly illuminated targets are much more sensitive to an inaccurate initial model. Indeed, the classical FWI converges to a secondary minimum instead of global minimum and this problem comes from the non-

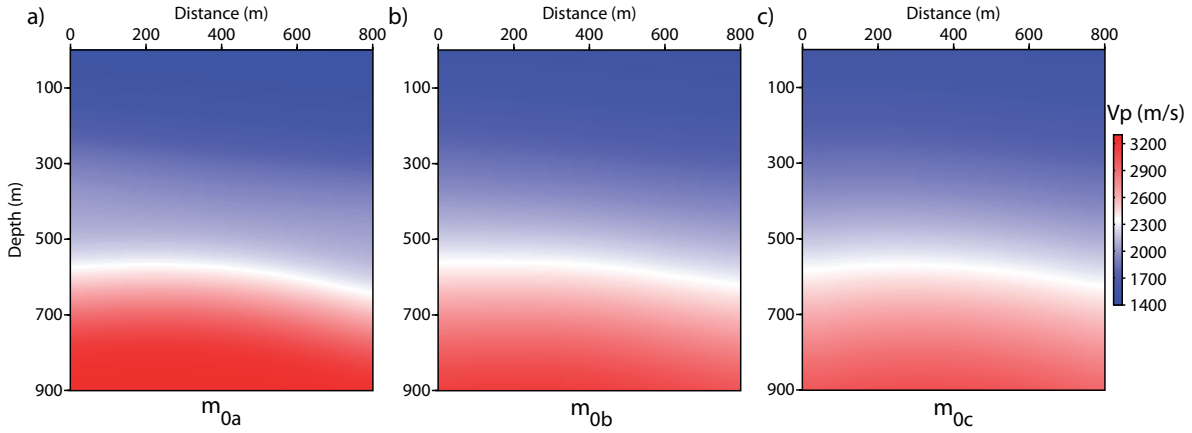


Figure 2.18: The smoothed versions of true baseline model by different smoothness factors are used as three initial models for FWI, (a) less smoothed (with 150 m length of Gaussian smoothing), the accurate initial model m_{0a}, (b) medium smoothed (with 250 m smoothness length), intermediate model m_{0b}, and (c) highly smoothed (with 300 m smoothness length), the approximate model m_{0c}.

uniqueness of ill-posed inverse problem.

The question can be raised up whether there is any way to improve the obtained model from the inaccurate initial model. Could the prior information solve these mentioned problems? Therefore, the baseline dataset is inverted by regularized inversion using the smooth prior model m_{p1} for all three initial model cases. Figure 2.21e shows that adding prior model to the inversion leads to a significant improvement on the obtained model when using a kinematically inaccurate initial model. Compared to the case of without prior model, the recovered baseline model is closer to global minimum and there is no issue of cycle-skipping and partial illumination. In fact by adding prior model misfit to the objective function, this model penalty term changes the direction of gradient of objective function as compared to the case when only the gradient is driven by data misfit and Tikhonov terms. This change in gradient helps the inversion to converge to the correct valley of descent. Therefore, regularization and prior model, when adequately introduced, can solve the non-uniniqueness of ill-posed problem. However, we should mention that the recovered baseline models for both other cases, the accurate initial m_{0a} (Figure 2.21a) and the intermediate initial model m_{0b} (Figure 2.21c), are degradated compared to the cases of without prior model, because of very high value of λ_2 which was the optimal value found for the third initial model case, because of using less appropriate prior model, and also not applying the dynamic prior weighting. The recovered logs in all three cases show that the interfaces especially the deeper curvature interfaces are reconstructed very well but going into depth, below the $z = 400$ m, problem of lack of low wavenumber exists. The λ_2 should be chosen for each initial model case independently, based on the ratio between prior-model and data-misfit terms. Here, we kept it as a constant for all cases to have comparable results.

Is there any way to compensate the problem of lack of low wavenumber by changing the prior model? Hence, the previous regularized inversions are performed again using the 1D prior model m_{p2} which is a linear interpolation between velocities of two wells. Figure 2.22 shows the results for the case of 1D prior model. The first observation is that even with including

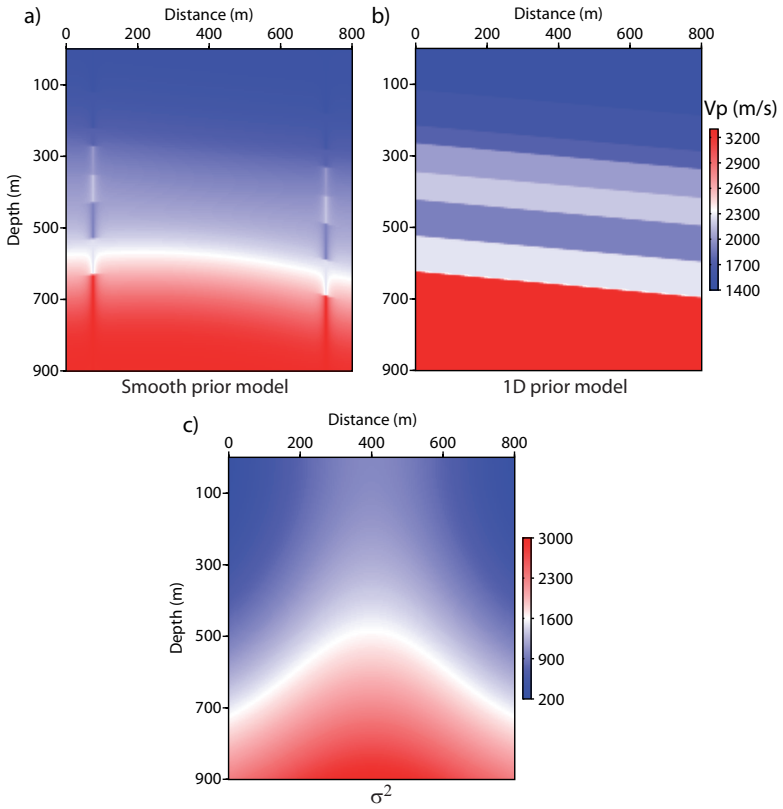


Figure 2.19: Based on prior information and well measurements, it is possible to design prior models in two different ways: (a) smooth prior model, and (b) 1D prior model. (c) The uncertainty attached to prior model is chosen as Gaussian function varying in x direction which is combined with a depth weighting in z direction, (σ^2) model.

deeper flat interfaces inside the prior model, the inversions are able to reconstruct the curvature interfaces based on the data-misfit term of objective function. In addition, compared to the case of inversion with smooth prior model, this new prior model can compensate the lack of low wavenumber in the deeper part of model. Even though, at the center and inside the seventh layer of recovered models, we still have problem of higher velocity anomaly, which is due to illumination issue and not using the dynamic weighting approach. Later, we show how the dynamic prior weighting approach can solve this problem.

To sum up, adding prior model helps FWI to go to the global minimum and makes inversion less sensitive to the initial models. In addition, lack of low wavenumber could be partially compensated by choosing appropriate prior model. In the next section, the sensitivity of time-lapse inversion with respect to these baseline models will be discussed.

2.3.2 Time-lapse reconstruction by Differential FWI

Once the baseline models have been correctly obtained, the time-lapse inversion can be performed. Here, we focus on the differential (double-difference) FWI strategy, which has been found to be more reliable and more robust. The nine baseline models obtained from different

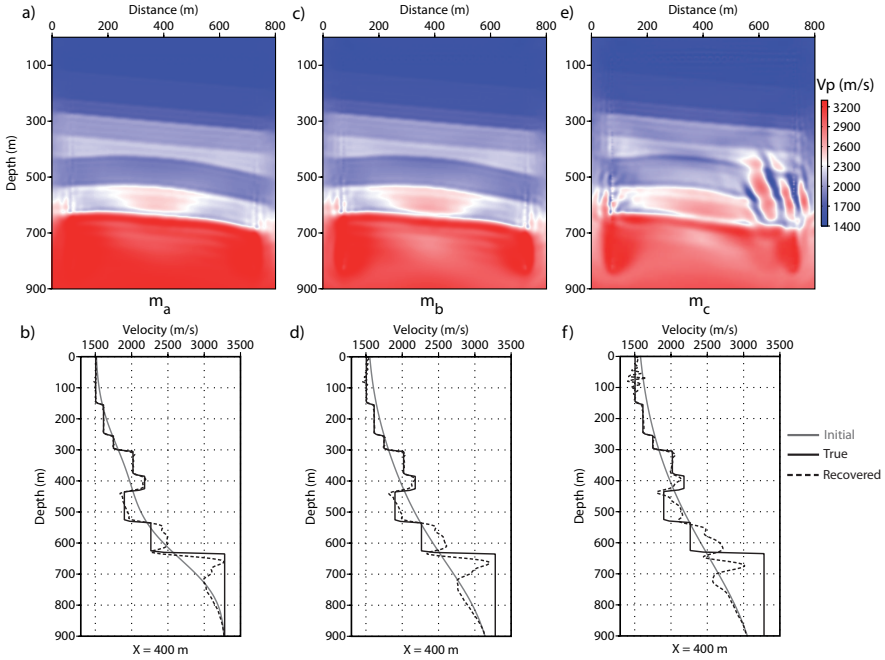


Figure 2.20: The recovered baseline V_p models obtained by regularized FWI *without prior model*, i.e. small λ_1 and $\lambda_2 = 0$ in all cases and the vertical logs associated to the center of each model ($x = 400\ m$); (a) and (b) in case of the accurate initial model m_{0a}, (c) and (d) the intermediate initial model m_{0b}, and (e) and (f) the approximate initial model m_{0c}.

inversion configurations in previous section are used to perform nine differential inversions. In this investigation, no prior model is applied in time-lapse inversion scheme for all nine time-lapse configurations, in order to see the impact of different reference models (starting models for the differential strategy) on the time-lapse inversion. Figure 2.23 illustrates the obtained time-lapse variation models from the baseline models (Figure 2.20) which are recovered without any prior model. Obviously, the accurate initial model used for the baseline inversion leads to more accurate baseline model and consequently this one results to get a robust time-lapse variation (Figure 2.23a). In case of inaccurate baseline model, the obtained time-lapse model contains several artifacts and anomalies (Figure 2.23c).

It has been shown that by using the prior model for the baseline inversion, especially in case of the inaccurate initial model, we honor a significant improvement in recovering the baseline model. Therefore, it is certainly expected that the time-lapse model would be improved (compare Figure 2.24c and 2.23c). The time-lapse result of second initial model with smooth prior model (Figure 2.24b) is degradated and less accurate than the same case obtained without any prior model (Figure 2.23b). First, this observation shows the sensitivity and robustness of time-lapse variation recovering with respect to the reference model used for time-lapse inversion. Second, the recovered baseline model, in case of second initial model (m_{0b}) and using the smooth prior model, is less accurate as we have not performed the dynamic prior weighting. Later, we will show how the dynamic approach leads to an almost perfect baseline model.

In the last configurations, the differential inversions (Figure 2.25) are performed from the baseline models obtained by including the 1D prior model (Figure 2.22). In all three cases, we obtain almost similar results. These results are more accurate at reservoir regions compared

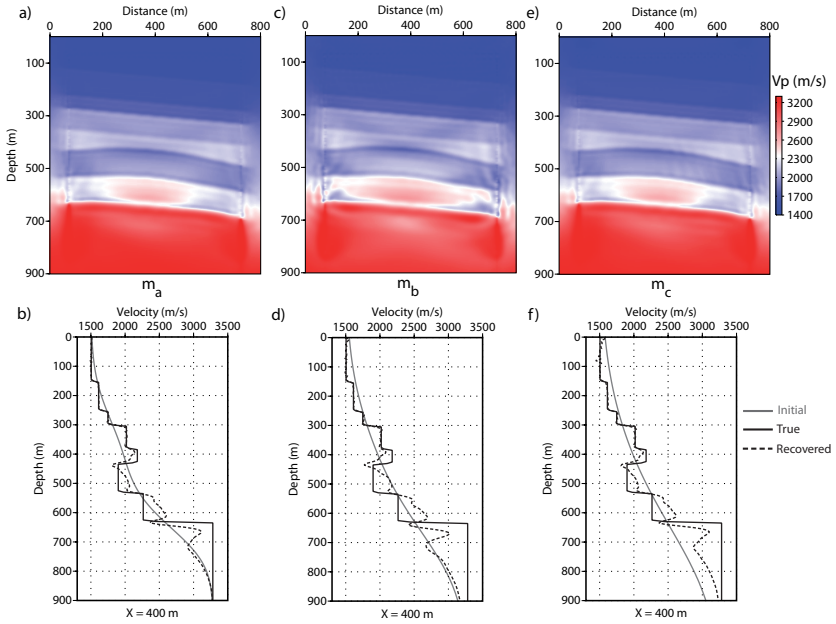


Figure 2.21: The recovered baseline V_p models derived by regularized FWI *with smooth prior model*, small λ_1 and fixed λ_2 in all cases and the vertical logs associated to the center of each model ($x = 400\ m$) in case of (a) and (b) the accurate initial model m_{0a}, (c) and (d) the intermediate initial model m_{0b}, and (e) and (f) the approximate initial model m_{0c}.

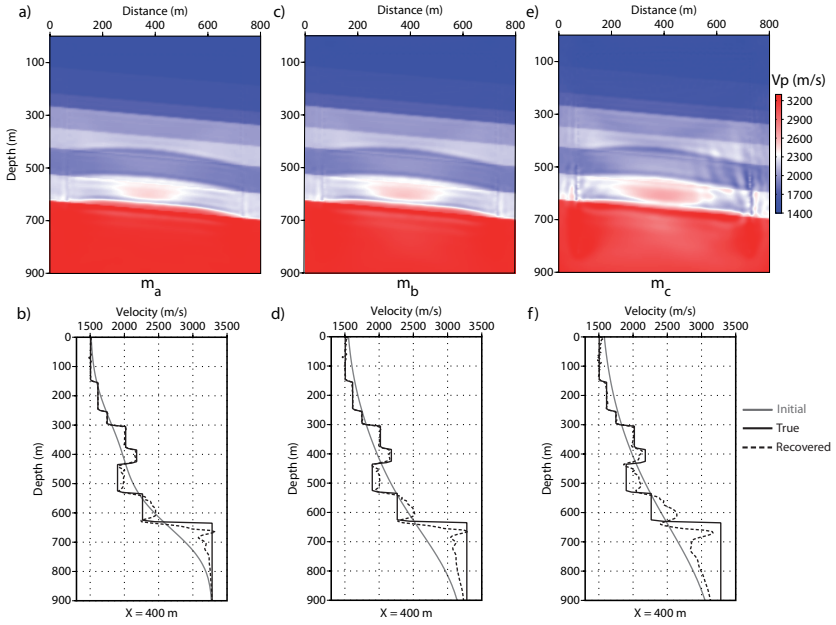


Figure 2.22: The recovered baseline V_p models derived by regularized FWI *with 1D prior model*, small λ_1 and fixed λ_2 in all cases and the vertical logs associated to the center of each model ($x = 400\ m$) in case of (a) and (b) the accurate initial model m_{0a}, (c) and (d) the intermediate initial model m_{0b}, and (e) and (f) the approximate initial model m_{0c}.

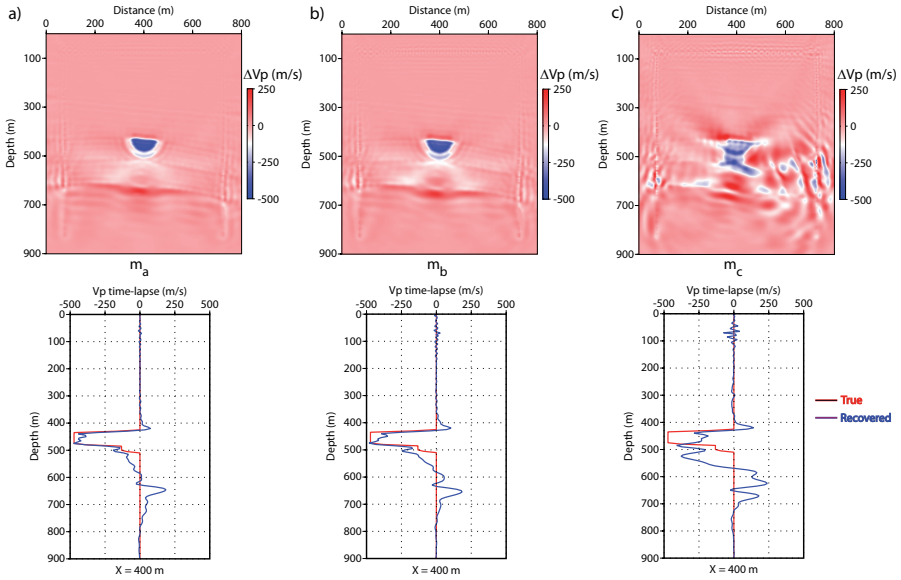


Figure 2.23: Time-lapse variation models and the QC vertical logs at $x = 400\ m$ obtained by the differential FWI with small Tikhonov regularization starting from the recovered baseline models obtained without using any prior model: (a) from the baseline model shown in Figure 2.20a, (b) from the baseline model shown in Figure 2.20c, (c) from the baseline model shown in Figure 2.20e.

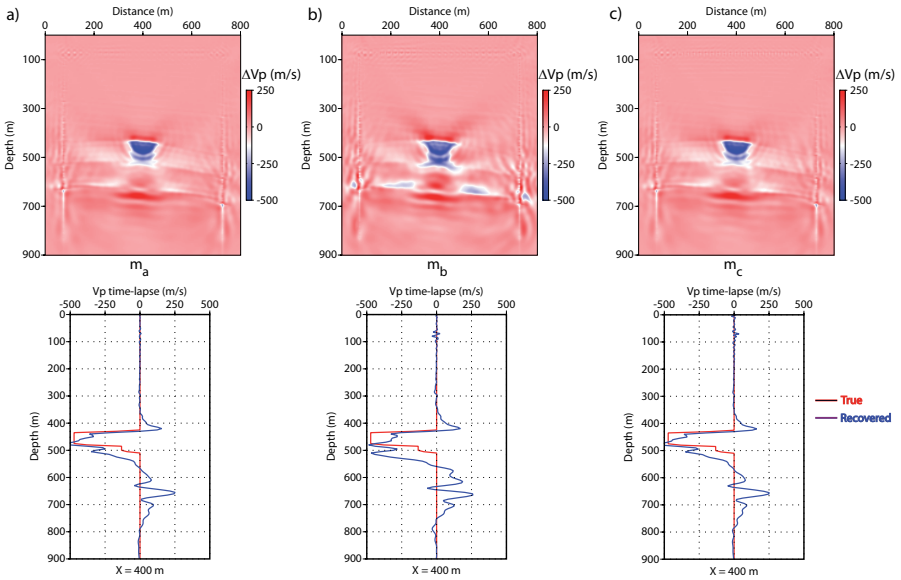


Figure 2.24: Time-lapse variation models and the QC vertical logs at $x = 400\ m$ obtained by the differential FWI with small Tikhonov regularization starting from the recovered baseline models obtained with the smooth prior model: (a) from the baseline model shown in Figure 2.21a, (b) from the baseline model shown in Figure 2.21c, (c) from the baseline model shown in Figure 2.21e.

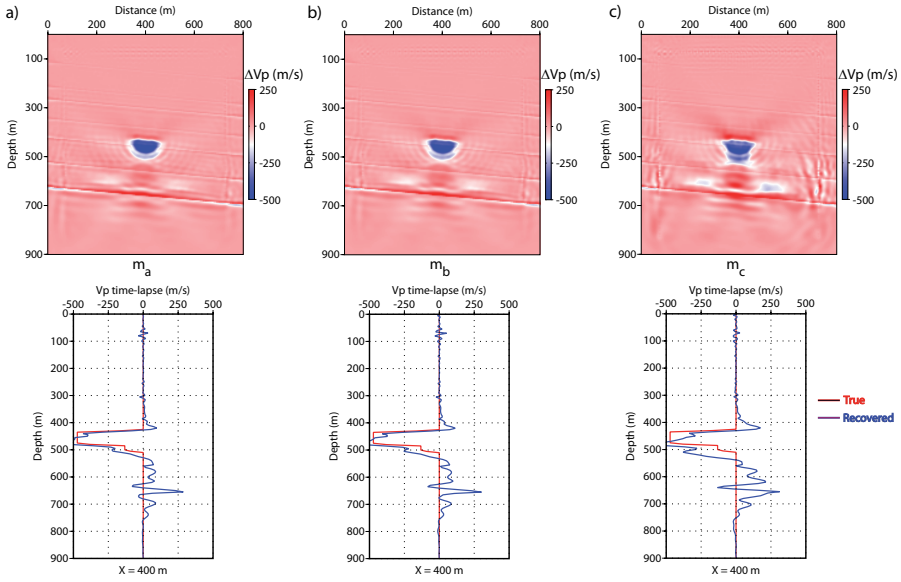


Figure 2.25: Time-lapse variation models and the QC vertical logs at $x = 400\ m$ obtained by the differential FWI with small Tikhonov regularization starting from the recovered baseline models obtained with the 1D prior model: (a) from the baseline model shown in Figure 2.22a, (b) from the baseline model shown in Figure 2.22c, (c) from the baseline model shown in Figure 2.22e.

to the results obtained from the smooth prior model (Figure 2.24). However, we can see the footprints of the 1D prior model used for the baseline inversion. Below the real time-lapse variation and near to the last high contrast interface (depth 650 m), there is a high velocity variation which is incorrect. This anomaly coming from the inaccuracy of baseline model could be removed by performing the target-oriented inversion using the prior model and prior weighting.

It has been shown that the time-lapse inversion is sensitive to the baseline models used for the inversion. When more accurate baseline model is used, more robust time-lapse variation can be recovered. To get more accurate baseline model, it is necessary to include the appropriate prior information into the inversion scheme.

Here, in order to remove the footprints of the prior model on the baseline inversion and also better improve the baseline result, we apply the dynamic prior weighting approach to the case where the approximate initial model m_{0c} and the 1D prior model are used (Figure 2.26a) for the baseline inversion. This approach leads to improve the baseline result and inversion converges to the quasi-perfect model. In case of fixed prior weighting, there is a large high velocity zone approximately at $600m$ depth (Figure 2.22e), which is incorrect and cannot be solved by optimization due to a fixed prior weighting. By performing the dynamic approach, after solving the cycle-skipping problem, the effect of 1D prior model gradually reduces. This leads to decrease the data residual as much as possible and delivers more precise model. It is possible to compare the baseline result of the classical FWI and the dynamic prior weighting FWI starting both from the same inaccurate initial model (Figure 2.20e and 2.26a). Then, if we perform time-lapse inversion (differential FWI) starting from this accurate baseline model, the most accurate time-lapse variation can be delivered (Figure 2.26b).

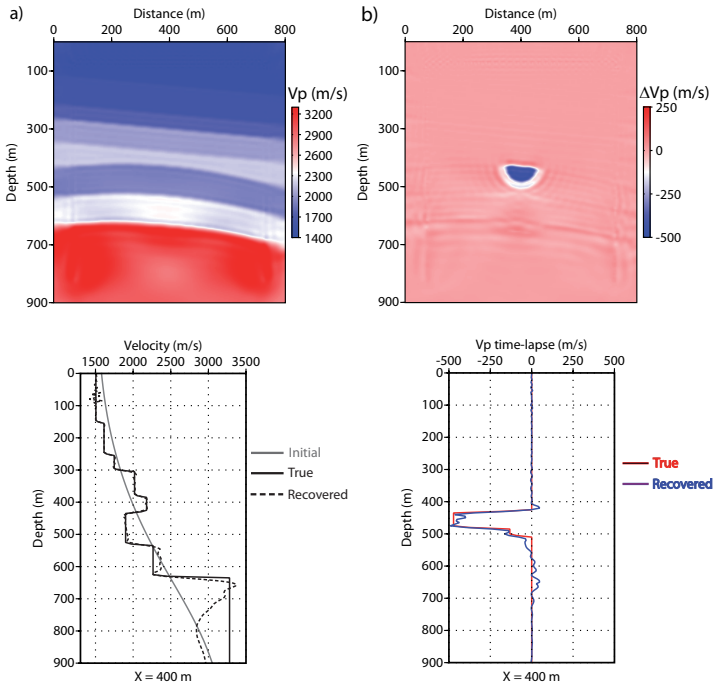


Figure 2.26: (a) The baseline model obtained by the dynamic prior weighting and the QC vertical logs at $x = 400\ m$. The inversion starts from the approximate initial model m_{0c} and with the 1D prior model. (b) The time-lapse model from the recoverd baseline model (a) and its corresponding vertical log.

Now, we would like to perform the target-oriented time-lapse inversion (differential FWI) in order to focus only on the target region. The recovered baseline model is used as the starting and the prior model for the inversion. We design a prior weighting model $\mathbf{W}_m$ to introduce strong weight of the prior model outside of the interested target and less weight inside it. As outside of the interested target, there is no variation. Figure 2.27 shows the σ^2 model (recall $diag(\mathbf{W}_m^T\mathbf{W}_m) = 1/\sigma^2(\mathbf{m})$). The target-oriented method is applied to invert the differential dataset starting from one less accurate recovered baseline model and the quasi-perfect model obtained by the dynamic approach (Figure 2.26a). The less accurate baseline model is chosen from that obtained by starting from the approximate initial model m_{0c} and using the fixed weighting and the smooth prior model (Figure 2.21e). The results of target-oriented inversion are shown in Figure 2.28. The impressive results illustrate how the target-oriented inversion can focus on the recovering of variations inside the target and how it can remove the artifacts outside of the interested zone (Figures 2.28a and 2.28b to be compared with Figures 2.24c and 2.26b, respectively). Target-oriented inversion provides almost similar time-lapse results which are now less sensitive to the reference baseline model. This leads to have an easier and consistent quantitative interpretation.

2.4 Discussion

In this chapter, we mainly focused on the quality of recovered baseline model, for the time-lapse strategies. However, the quality of baseline data-fitting would be important as well. How well

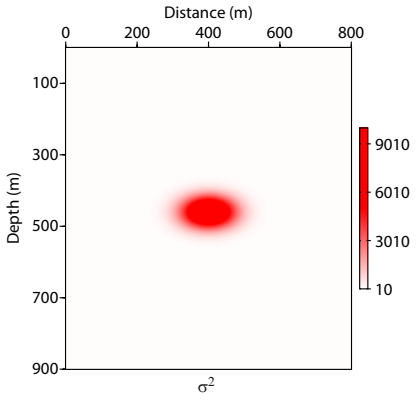


Figure 2.27: The target-oriented σ model built based on the available prior information on the location of injected steam area.

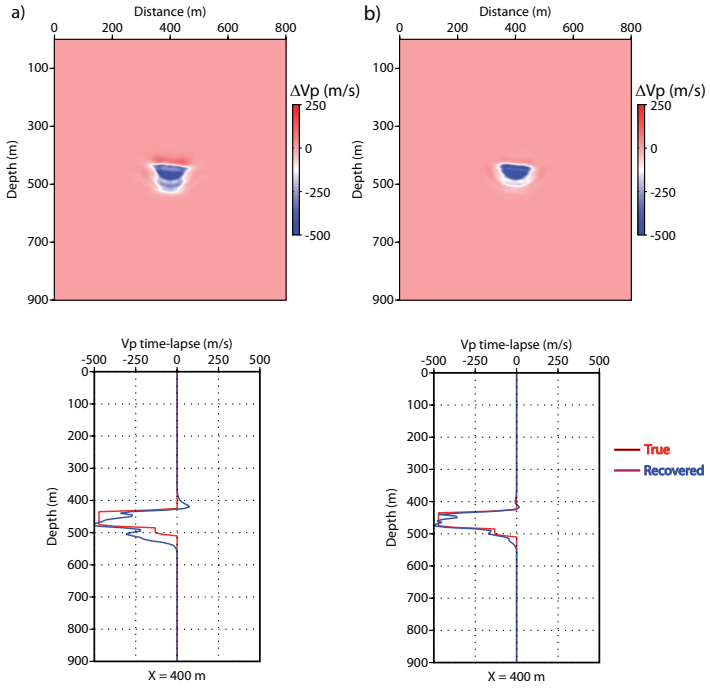


Figure 2.28: Results of the target-oriented time-lapse inversion starting from two different baseline models: (a) starting from the baseline, obtained by using fixed weighting and the smooth prior model (Figure 2.21e); (b) starting from the baseline, obtained by dynamic approach (Figure 2.26a.)

the baseline data are explained? In synthetic case, when the true model is known, the good quality of model can ensure the good quality of fitting data. In reality, the good quality of the data fitting cannot totally ensure the accuracy of obtained model (due to non-uniqueness issue). However, it can tell that the obtained model is located in the interested valley of misfit function. Suppose that the baseline data are fitted very well, but with a wrong baseline model (although this assumption would be rare if the data contain several arrivals and complexity).

Is it possible to recover the time-lapse variation on this wrong baseline model? It appears that the differential strategy, since it inverts the difference data, could find the time-lapse variation, but probably not at appropriate depth. The difference data should be propagated inside the baseline model (which is non-kinematic and wrong here), leading to misposition time-lapse perturbation. The differential strategy cannot change the background velocity (i.e. the kinematic of velocity model used for starting differential inversion) and it cannot update the parts of model which are invariant between baseline and monitor model.

The sequential strategy, since it uses directly the monitor dataset which contains time-lapse data and repeatable baseline data, could possibly change the repeatable parts of model between baseline and monitor model. By using monitor dataset, because of time-lapse data, the monitor data space will change with respect to the baseline data space. Therefore, there would be a possibility to move into another local minimum and repeatable baseline data are better explained by other model parameter values. Hence, the repeatable parts of velocity (between two models) can be updated as well in good or wrong directions. In this case, by simple subtracting between two obtained models, baseline and monitor, all these variations appearing could be seen as incorrect time-lapse variations, which would be difficult to distinguish them from real 4D changes.

The residuals left in the baseline reconstruction step ($\mathbf{d}_{obs_b} - \mathbf{d}_{calc_{rec-b}}$) exist in the misfit function of the sequential difference strategy. If these baseline residuals are still left unchanged during monitor inversion, there must be no problem for identifying the time-lapse variation, because the model parameters related to these residuals stay invariant for both recovered models (baseline and monitor). However, if these residuals are recovered during monitor inversion, the final time-lapse model will be affected by spurious variations. These spurious time-lapse variations come from the fact that we make a simple subtraction (point by point) between two depth images (baseline and monitor) to get the time-lapse model. A question can be raised up, is there any other way to extract the real time-lapse variation from two images? For example, by computing the coherency between two images, one can extract the real time-lapse variation. Similar as the data space in FWI, where several misfit functions between observed and calculated data have been introduced such as difference-based or correlation-based misfit (van Leeuwen and Mulder, 2008, 2010) to better extract the information, we may need a better difference method to extract time-lapse vaiation from two obtained images. This would be considered as a perspective for depth domain time-lapse imaging.

In target-oriented mode, by applying the target-oriented prior model weighting, the misfit function is changed. The standard misfit function contains several local minima where it is more possible to get trapped in. By constraining inversion to the target areas, the misfit function is reshaped and it seems to become more convex. Therefore, the paths of mimimizing both misfit functions, in conventional and target-oriented modes, are not the same at all which leads to get different time-lapse results even inside the target areas. In other words, we can mention that by adding the prior model information to the inversion, the model parameter space is reduced. Hence, the time-lapse inversion becomes more well-posed.

2.5 Conclusions

In this chapter, we have studied the robustness of three strategies for time-lapse imaging with regularized FWI. Using the Marmousi synthetic case, it has been shown that the parallel

difference strategy is highly sensitive to the differential independent artifacts in the images. The sequential difference strategy can be more attractive when the baseline model is accurately recovered, since this method attempts to recover the parts of the model which have not been fully reconstructed before. The differential strategy is more robust than the sequential one, since the differential inversion only focuses on the differential (time-lapse) data.

In synthetic steam injection model, we have shown the major impact of the appropriate prior model on the inversion, in order to reduce the sensitivity of baseline inversion to the inaccurate smooth initial model. Consequently, this reconstructed robust baseline model leads to recover the time-lapse variation in more accurate and more robust way.

We also propose a target-oriented time-lapse imaging with regularized FWI including prior model and prior weighting model. In the application to the Marmousi and the Dai datasets, it has been shown that the target-oriented inversion prevents the apparition of artifacts outside of the target areas and leads to retrieve a more precise time-lapse model for quantitative and qualitative interpretations.

When the quality of data is acceptable in terms of signal-to-noise ratio, the differential strategy can be more interesting than the sequential one, since it can focus on the target areas. In the case of very strongly noisy data, the double-difference strategy is more sensitive to the non-repeatable noise at the subtraction step. The high noise level can mask the weak time-lapse signals. In this case, probably using the target-oriented sequential difference could be an alternative strategy.

In the next chapter, we study the time-lapse FWI in frequency domain. Which sampling frequency can be efficient to recover baseline accurately, and consequently efficient for reconstructing small time-lapse perturbations? What would be the sensitivity of different time-lapse strategies to the different sampling frequencies for inversion? We investigate whether limited number of inverted frequencies is sufficient for the time-lapse inversion or not. We are interested to perform time-lapse inversion in frequency domain, especially in terms of sparce sampling and data compression which could be crucial to reduce computational cost for 3D inversion as a perspective.

Chapter 3

Time-lapse FWI and Frequency Sampling Strategy

Contents

This chapter contains one article which is in preparation and will be submitted to *Geophysical Prospecting*. In earlier chapters, the baseline reconstruction and several time-lapse strategies have been discussed. The inversion was performed in the time domain. In this paper, we study the time-lapse inversion in frequency domain. Moving to frequency domain, it can keep the

data complexity low and can also provide sparce sampling of data for FWI. In other words, data in frequency domain are compressed as compared to time domain data. Therefore, frequency domain inversion could be crucial for 3D inversion. However, number of selected frequencies for inversion is important, and there is a trade-off between frequency-domain inversion and time-domain inversion, based on their computational expenses. In time-lapse applications, the accuracy of recovered time-lapse changes by these two different domain inversions would also be important.

Several studies will be investigated in order to find which domain, time or frequency, would be more efficient for a time-lapse application. We propose to compare the results of an inversion approach in frequency domain inverting all the frequencies simultaneously (similar to the time domain inversion) and several decimations on the selected frequencies on the Marmousi synthetic data. Which sampling frequency can be suitable to recover baseline accurately, and consequently efficient for reconstructing small time-lapse perturbations? What are the differences in resolution between time-lapse models derived from different sampling frequency strategies? In addition, the sensitivity of time-lapse strategies to the multi-group multi-frequency inversion will be studied in case of noise-free and noisy data.

Time-lapse full waveform inversion in frequency domain: which frequency decimation?
Amir Asnaashari, Romain Brossier, Stéphane Garambois, François Audebert, and Jean Virieux
2013, *Geophysical Prospecting*, in preparation.

3.1 Introduction

Monitoring is an important issue for oil and gas production and CO_2 sequestrations to track the time-lapse variations in target areas. Full waveform inversion (FWI) is an attractive technique for velocity model building that reconstructs high resolution velocity models of the subsurface through the extraction of the full information content of seismic data (Tarantola, 1984b; Virieux and Operto, 2009). Since the FWI approach delivers high resolution quantitative images of macro-scale physical parameters, it could be a good candidate for monitoring applications to recover the parameter variation through a time evolution.

The seismic waveform inversion has been used for the reconstruction of the model parameters during last decade. This inversion formulation can be done either in time or frequency domain. In 1980s, the time-domain formulation has been initiated by Lailly (1983b); Tarantola (1984a) and Mora (1987). Later, Pratt and Worthington (1990); Pratt (1990b) and Pratt (1999) have proposed the same idea of inversion problems in the frequency domain. Always, there is a question about which domain is more robust and efficient for inversion. Time-domain and frequency-domain methods are analytically equivalent when all the frequency components are inverted simultaneously (Pratt et al., 1998).

The advantages of time-domain modeling and inversion are its less demanding for computer memories for the forward modeling. The interesting point of this method is its well-posedness in terms of frequency component, since the gradient is calculated in time domain, which means

that all frequency contents are used simultaneously. However, by taking into account all frequency contents, the complexity of data incearses. Freudenreich and Singh (2000) have shown that in a simple example, the frequency-domain approach works fine for wide-angle data but fails in the case of limited offset range, while the time-domain approach appears more robust.

Most of the advantages of frequency domain are related to the forward modeling and also sparce sampling for FWI in 3D. One of the main reasons to move to the frequency domain to solve the Hemholtz equation is that, the matrix of discrete operator of wave propagation can be factorized only once for all the sources. In case of a large number of sources, the time-domain strategy is less interesting because of the increased computation time. Therefore if we have a significant number of sources, the frequency domain is the method to choose (Marfurt, 1984), since the LU decomposition is performed only once per frequency (Amestoy et al., 2003; Press et al., 2007; MUMPS-team, 2009). However, in frequency domain, it is important to correctly select the number of frequencies to be inverted. It has been shown that the large aperture seismic surveys could be inverted using only a limited number of frequencies (Pratt, 1990a; Sirgue and Pratt, 2004), thus asking for a lesser number of forward modeling solutions. However, the LU decomposition should be performed for each frequency and in the case of multifrequency inversion, including a large number of frequencies, the computational cost is expensive. In case of choosing a smaller number of frequencies, the inverse problem remains ill-posed and there is a risk to be trapped in local minima. Consequently, the main question to address, with such an approach, regards the optimal number of frequencies for an efficient FWI. Perhaps, a limited number of frequencies could be sufficient for the reconstruction of model parameters for an exploration project, but is it the case for monitoring applications where small variations and changes should be detected?

The process of time-lapse FWI could be performed in two steps: first the baseline reconstruction and second the monitor reconstruction. In the second step of monitoring, several methods can be used for the monitor reconstruction. The *differential method* (double difference) is one of the proposed approaches for such applications (Watanabe et al., 2004; Denli and Huang, 2009; Asnaashari et al., 2012). This procedure is focused on inverting the differential dataset from the reconstructed baseline model obtained at the first step. A second approach, called the *parallel difference method*, independently inverts the two datasets (baseline and monitor) starting from the same initial model. The time-lapse changes will be assessed by subtracting the final derived monitor model from that of the baseline (Plessix et al., 2010). A third approach, called the *sequential difference method*, uses the final baseline model as a starting model for inverting the monitor dataset (Asnaashari et al., 2012).

Which sampling frequency can be optimal for recovering the baseline accurately, and consequently efficient for reconstructing small time-lapse perturbations? What would be the sensitivity of different time-lapse strategies to the different sampling frequencies for inversion? What are the differences in resolution between models derived from different sampling frequency strategies?

In this paper, we propose to compare on the Marmousi synthetic dataset the results of an inversion approach in frequency domain inverting all the frequencies simultaneously, with the results of inversion after several decimations on the selected frequencies, for baseline reconstruction and time-lapse recovering. In addition, the sensitivity of time-lapse strategies to the multi-group multi-frequency inversion will be studied in case of noise-free and noisy data.

3.2 Formulation in frequency domain

Full waveform inversion is an iterative approach aiming to recover model parameters based on the local optimization of residuals between the observed and the calculated wavefields at receiver positions. The forward and inverse problem can be written either in time or in frequency domain.

In this section, we briefly introduce the waveform inversion in the frequency domain using multifrequency approaches (Pratt and Worthington, 1990; Liao and McMechan, 1996). Before doing inversion, we need to compute precisely the calculated data which is produced by the forward modeling engine.

3.2.1 Forward modeling

Frequency-domain forward modeling is of special interest for multi-source experiments, because of its computational efficiency if there are many sources (Pratt and Worthington, 1990; Stekl and Pratt, 1998). Several techniques of discretization of the Wave Equation have been studied, such as finite-differences (FD) or finite-elements (FE). Among them, one of the possible methods is the finite element Discontinuous Galerkin (DG) method (Käser and Dumbser, 2006; Dumbser and Käser, 2006; Brossier et al., 2008). This DG method allows the use of triangular/tetrahedral meshes, which is suitable for the handling of strong physical contrasts in the medium, including liquid/solid contact. However this method demands higher computational costs compared to the finite-difference and is more difficult to be implemented. Therefore in this study, we use the finite-difference method to solve the acoustic wave equation.

A classical five-point stencil for 2D acoustic wave equation requires at least 10 gridpoints per shortest wavelength (Pratt and Worthington, 1990). Jo et al. (1996) have proposed a mixed-grid stencil method where the required number of gridpoints per shortest wavelength reduces to four points for accurate modeling. This approach consists of a linear combination of two discretizations of the second derivative operator, the discretizations on the classical Cartesian coordinate system and the 45° rotated system. This mixed-grid stencil is one of the strategies proposed to increase the accuracy of the numerical operators. In fact, we can increase the accuracy by incorporating gridpoints in the FD stencil along several directions (Jo et al., 1996; Stekl and Pratt, 1998; Shin and Sohn, 1998). An alternative strategy to increase the accuracy of FD operators is to go to higher order operators (Hustedt et al., 2004). The fourth-order staggered-grid method with 13-point stencil has been suggested by Hustedt et al. (2004).

It has been shown that the accuracy of the fourth-order staggered-grid stencil is slightly better than the mixed-grid stencil when averaging of the mass acceleration term is applied to the staggered-grid stencil. However, the mixed-grid stencil using 9-point stencil appears to be more efficient than the 13-point staggered-grid stencil in terms of CPU and memory performance for matrix factorization (Hustedt et al., 2004). In our study, we use the mixed-grid stencil to solve the acoustic wave equation in frequency domain. In the following, I explain briefly this method.

3.2.1.1 Finite-difference frequency domain (FDFD)

The acoustic wave equation is first expressed as a first-order hyperbolic system (Virieux, 1984) including the pressure and the particle velocity wavefields. The system of hyperbolic equation

in frequency domain is given by:

$$\begin{aligned}
\frac{-i\omega\xi_x(x)}{K(x,z)}P_x(x,z,\omega) &= \frac{\partial Q(x,z,\omega)}{\partial x} + s(x,z,\omega) \\
\frac{-i\omega\xi_z(z)}{K(x,z)}P_z(x,z,\omega) &= \frac{\partial R(x,z,\omega)}{\partial z} \\
-i\omega Q(x,z,\omega) &= \frac{b(x,z)}{\xi_x(x)}\frac{\partial P(x,z,\omega)}{\partial x} \\
-i\omega R(x,z,\omega) &= \frac{b(x,z)}{\xi_z(z)}\frac{\partial P(x,z,\omega)}{\partial z},
\end{aligned} \tag{3.1}$$

where $Q(x,z,\omega)$ and $R(x,z,\omega)$ are particle velocity wavefields on Cartesian grid. The unphysical pressure components $P_x(x,z,\omega)$ and $P_z(x,z,\omega)$ are used to separate the horizontal and vertical derivatives and also to take into account for the PML absorbing condition (Operto et al., 2002). The real pressure wavefield can be obtained by $P(x,z,\omega) = P_x(x,z,\omega) + P_z(x,z,\omega)$. $b(x,z)$, $K(x,z)$ and $s(x,z,\omega)$ represent the inverse of density (buoyancy), bulk modulus and the pressure source term, respectively. The coefficients $\xi_x(x) = 1 + i\gamma_x(x)/\omega$ and $\xi_z(z) = 1 + i\gamma_z(z)/\omega$ contains the PML damping values (γ) which are equal to zero outside the PML layers.

The system 3.1 is discretized using second-order centered finite difference. After discretization and elimination of particle velocities from the coupled first-order equations, we can end up with a parsimonious second-order staggered-grid formulation (Luo and Schuster, 1990; Hustedt et al., 2004). The mixed-grid method combines two second-order staggered-grid stencils on the classical Cartesian coordinate system and the 45° rotated system (Jo et al., 1996). The two second-order derivations provide a tool to discretize the frequency-domain second-order acoustic wave equation. At the end, five C coefficients corresponding to the staggered-grid five-point stencil on a classical grid are exteracted (Figure 3.1). As another alternative, the spatial derivatives of the system 3.1 can be discretized along the rotated axes (x', z'). It leads to have R coefficients (see Figure 3.1). The mixed-grid stencil method combines two mentioned staggered-grid stencils using the averaging coefficient to take into account the coupling between the grid points on the classical Cartesian and the rotated grids.

3.2.1.2 Direct solver for wave equation

At the end, the discrete system of wave equation can be expressed in a matricial form:

$$\mathbf{A}(\mathbf{m}(\mathbf{x},\mathbf{z}),\omega)\mathbf{u}(\mathbf{x},\mathbf{z},\omega) = \mathbf{s}(\mathbf{x},\mathbf{z},\omega), \tag{3.2}$$

where $\mathbf{A}(\mathbf{m}(\mathbf{x},\mathbf{z}),\omega)$ is the impedance matrix and $\mathbf{m}(\mathbf{x},\mathbf{z})$ denotes the model parameters. ω is the angular frequency. The solution of this equation is $\mathbf{u}$, the pressure wavefield vector, and $\mathbf{s}$ is the vector of source term. For the mixed-grid stencil, the impedance matrix $\mathbf{A}$ contains $(n_x \times n_z)^2$ elements but only $(9 \times n_x \times n_z)$ are non-zero. The system of linear equation in matricial form is solved by decomposition of the matrix $\mathbf{A}$ into two upper-diagonal ($\mathbf{U}$) and lower-diagonal ($\mathbf{L}$) matrices (Press et al., 2007). The method is known as LU decomposition (Amestoy et al., 2003; MUMPS-team, 2009). The equation $\mathbf{Au=s}$ is transformed into $\mathbf{LUu=s}$, which is solved in two steps. Firstly $\mathbf{Ly=s}$ and then $\mathbf{Uu=y}$. The matrix $\mathbf{A}$ is only dependent

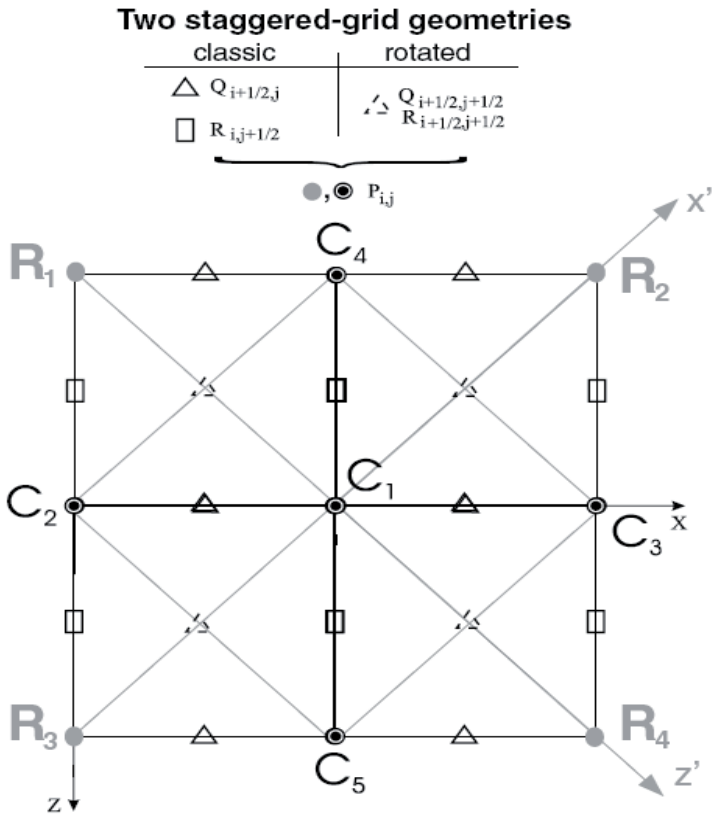


Figure 3.1: Schematic view of two staggered-grid stencils, classical and rotated grids (after Hustedt et al. (2004)).

on the acoustic model parameters and on frequency, and is source-independent. Therefore the factorization is performed only once (for each frequency) with a multi-frontal method and the wavefields generated for multiple sources are computed effectively by forward and backward substitutions (Duff and Reid, 1983; Amestoy et al., 2000). This is the main advantage of the factorization of the impedance matrix for a large number of sources (Marfurt, 1984).

3.2.2 Inverse problem

The inverse problem relies on an iterative local optimization problem that is generally introduced as a linearized least-squares problem. The optimization attempts to minimize the residuals between the observed and the modeled wavefields at receivers (Tarantola, 1987). We can define the residual vector $\Delta\mathbf{d}$, considering one seismic source and one frequency,

$$\Delta\mathbf{d}(\omega) = \mathbf{d}_{obs}(\omega) - \mathbf{d}^k_{calc}(\omega), \tag{3.3}$$

where the data misfit error $\Delta\mathbf{d}(\omega)$ denotes the difference between the observed data $\mathbf{d}_{obs}$ and the modeled data $\mathbf{d}_{calc}$ computed in the model $\mathbf{m}^k$ at the iteration k of the inversion. The synthetic data $\mathbf{d}_{calc}$ is obtained by applying a sampling operator $\mathbf{S}$ to the incident wavefield $\mathbf{u}$ resulting from the forward problem resolution $\mathbf{A}\mathbf{u} = \mathbf{s}$. The data vector contains complex pressure wavefield in frequency domain recorded at receiver positions.

The inversion is based on the ℓ_2 norm of the data residuals in a multifrequency approach. The objective function can be written as

$$\mathcal{C}(\mathbf{m}^k) = \frac{1}{2}\sum_{n\omega}\sum_{ns}\Delta\mathbf{d}^{\dagger}(\omega)\mathbf{W}_d\Delta\mathbf{d}(\omega), \tag{3.4}$$

where superscript † indicates the adjoint operator (transposed conjugate). $\mathbf{W}_d$ is the weighting matrix in data space and contains prior information on the data (Tarantola, 2005). The summations in equation 3.4 are performed over the ns sources and over $n\omega$ simultaneously inverted frequencies. In this case, the unknown model parameters $\mathbf{m}$ describe the P-wave velocity model. In this approach, there is a summation over all the selected frequencies for inversion, in other words all frequencies are inverted at the same time (simultaneously).

The model penalty terms, the Tikhonov regularization and the prior model terms, can be added to this data objective function in the same way as presented in Asnaashari et al. (2013a). Here, only the effect of frequency decimation on the inversion will be studied. Since this frequency decimation effect might be affected by the model regularization terms, these additional model penalties are not applied in this study.

Minimizing the misfit function classically leads to the normal equation system which can be written as,

$$\mathcal{H}_m^{k-1}\Delta\mathbf{m}^k = -\mathcal{G}_m^{k-1}, \tag{3.5}$$

where the gradient and the Hessian of the misfit function at iteration $k-1$ are denoted $\mathcal{G}_m^{k-1}$ and $\mathcal{H}_m^{k-1}$ respectively. The gradient of the objective function with respect to the model parameters can be written as

$$\begin{aligned}
\mathcal{G}_{\mathbf{m}} &= \frac{\partial\mathcal{C}(\mathbf{m})}{\partial\mathbf{m}} \\
&= -\frac{1}{2}\sum_{n\omega}\sum_{ns}\left[\left(\frac{\partial\mathbf{d}_{calc}(\mathbf{m})}{\partial\mathbf{m}}\right)^{\dagger}\mathbf{W}_d(\mathbf{d}_{obs}-\mathbf{d}_{calc}(\mathbf{m})) + (\mathbf{d}_{obs}-\mathbf{d}_{calc}(\mathbf{m}))^{\dagger}\mathbf{W}_d\left(\frac{\partial\mathbf{d}_{cal}(\mathbf{m})}{\partial\mathbf{m}}\right)\right] \\
&= -\sum_{n\omega}\sum_{ns}\Re\left[\left(\frac{\partial\mathbf{d}_{calc}(\mathbf{m})}{\partial\mathbf{m}}\right)^{\dagger}\mathbf{W}_d(\mathbf{d}_{obs}-\mathbf{d}_{cal}(\mathbf{m}))\right] \\
&= -\sum_{n\omega}\sum_{ns}\Re\left[\mathbf{J}^{\dagger}\mathbf{W}_d\Delta\mathbf{d}\right].
\end{aligned} \tag{3.6}$$

$\Re$ denotes the real part of a complex number. The sensitivity matrix $\mathbf{J} = \partial\mathbf{d}(\mathbf{m})/\partial\mathbf{m}$ is composed by the Fréchet derivatives of the synthetic data with respect to the model parameters. In our implementation, we never compute explicitly the matrix $\mathbf{J}$. The gradient with respect to the model parameters $\mathbf{m} = \{m_i\}_{i=1,N}$, where N denotes the number of unknowns, can be efficiently derived from the adjoint-state formulation using the back-propagation technique (Plessix, 2006). This formulation gives for the i^{th} component of the gradient $\mathcal{G}$:

$$\mathcal{G}_{m_i} = -\sum_{n\omega}\sum_{ns}\Re\Big\{\mathbf{u}^t\Big(\frac{\partial\mathbf{A}}{\partial m_i}\Big)^t\lambda^*\Big\}, \tag{3.7}$$

where ^{t} and * denote the transpose and conjugate operators, respectively. The gradient can be computed as a product between the incident wavefield $\mathbf{u}$ from the source, and the adjoint back-propagated wavefield λ^* which is computed as $\mathbf{A}^t\lambda^* = \mathbf{S}^t\mathbf{W}_d\Delta\mathbf{d}^*$, using residuals at

Algorithm 3.1 Simultaneous all inverted frequencies algorithm for FWI

while (NOT *convergence* AND $k < iter_{max}$) **do**
  for $\omega_i = 1$ to $n\omega$ **do**
    Compute incident wavefields $\mathbf{u}$ from sources
    Compute residual vectors $\Delta\mathbf{d}$ and objective function $\mathcal{C}^{k-1}$
    Compute adjoint back-propagated wavefields $(\mathbf{A}^{-1})^t\mathbf{S}^t\mathbf{W}_d\Delta\mathbf{d}^*$ from receivers
    Build gradient vector $\mathcal{G}^{k-1}$ and make summation over all frequencies
  end for
  Compute perturbation vector $\Delta\mathbf{m}^k$ and optimal step length $\alpha^k$with L-BFGS-B optimization
  Update model $\mathbf{m}^k = \mathbf{m}^{k-1} + \alpha^k\Delta\mathbf{m}^k$
end while

receiver positions as a composite source. Therefore, for computing the gradient, only two forward problems per shot are required. The radiation pattern of the scattering of the model parameter m_i is represented by the sparse matrix $\partial\mathbf{A}/\partial m_i$.

The gradient of the objective function is then used in an optimization algorithm to update the model vector with the perturbation vector $\Delta\mathbf{m}^k$ through the expression

$$\mathbf{m}^k = \mathbf{m}^{k-1} + \alpha^k\Delta\mathbf{m}^k, \tag{3.8}$$

where the step length at iteration k is denoted by α^k.

The Hessian matrix is based on the second derivative of the misfit function and is not computed in our implementation. Instead, we minimize our problem with a bounded quasi-Newton method using the L-BFGS-B routine (Byrd et al., 1995). This routine allows to take into account an approximate non-diagonal inverse Hessian from previous gradient and model vectors, and performs a line-search satisfying Wolfe's conditions. This bounded limited-memory quasi-Newton method is an efficient alternative to preconditioned steepest-descent or conjugate-gradient methods based only on gradients and/or approximate diagonal Hessian approaches (Brossier et al., 2009). The summary of simultaneous frequency inversion is shown in Algorithm 3.1.

Time-domain approach behaves similar as the frequency-domain approach when all frequencies are taken for the inversion.

3.2.3 Time-lapse FWI strategies

Once the baseline model is reconstructed, several strategies could be used for time-lapse procedure. In this section, we briefly introduce three different workflows for time-lapse imaging with FWI (details can be found in Asnaashari et al. (2013b)).

3.2.3.1 Parallel difference FWI

The parallel difference method considers independent inversion of the baseline and monitor datasets, using a similar starting model. After inversion, the time-lapse response can be obtained by making a subtraction between the recovered monitor and the reconstructed baseline

models. The main advantage of this approach is its applicability to acquisition geometries that do not match between the two experiments. As the two inversions are performed independently, a drawback is the potential interpretation of inversion artifacts as a real time-lapse response.

3.2.3.2 Sequential difference FWI

As the time-lapse response in data is often weak with regard to the full data complexity, the sequential difference method uses the recovered baseline model as a starting model for the monitor data inversion. This means that the baseline is reconstructed first, and subsequently the baseline model is used to invert the monitor dataset. As the time-lapse response is weak, the baseline model should be a good candidate for a starting model and should prevent a large number of iterations in the second step. The main drawback is that this strategy cannot focus only on the differential signals caused by the time-lapse variation.

3.2.3.3 Differential FWI

In the differential method, we attempt to minimize the residual of the differential data between baseline and monitor datasets (Watanabe et al., 2004; Denli and Huang, 2009; Asnaashari et al., 2013b), giving the expression

$$\Delta \mathbf{d} = (\mathbf{d}_{obs_m} - \mathbf{d}_{obs_b}) - (\mathbf{d}_{calc_m} - \mathbf{d}_{calc_b}), \tag{3.9}$$

where $\mathbf{d}_{obs_m}$ and $\mathbf{d}_{obs_b}$ are the monitor and baseline observed data at each frequency respectively, and $\mathbf{d}_{calc_m}$ and $\mathbf{d}_{calc_b}$ are the computed data for these experiments.

For the differential analysis, we first need the construction of a composite dataset defined as

$$\mathbf{d}_{composite} = \mathbf{d}_{obs_m} - \mathbf{d}_{obs_b} + \mathbf{d}_{calc_{rec-b}}, \tag{3.10}$$

which is composed of (a) the time-lapse differential observed data ($\mathbf{d}_{obs_m} - \mathbf{d}_{obs_b}$) which should only represent the time-lapse changes of the two datasets and (b) the calculated data $\mathbf{d}_{calc_{rec-b}}$ computed using forward modeling in the reconstructed baseline model. This composite dataset $\mathbf{d}_{composite}$ can be used as a new observed dataset $\mathbf{d}_{obs}$ in equation 3.3, which is now used to minimize the differential residual 3.9 with a standard FWI algorithm. Finally, the time-lapse model changes $\delta \mathbf{m}_{time-lapse} = \mathbf{m}_{composite} - \mathbf{m}_{rec-b}$ can be computed.

3.3 Sensitivity of reconstructed baseline with respect to frequency decimation

In this section, we perform the sensitivity analysis of the recovered baseline model with respect to the sampling frequency used for inversion. This study is done on Marmousi acoustic synthetic data, noise-free and noisy data. The inversion is performed by taking all the selected frequencies at the same time. In other words, the misfit function and its gradient are computed with summation on all the selected frequencies at each iteration.

The true baseline model is identical to that chosen in Asnaashari et al. (2013b) (Figure 2.4a). The monitor velocity model is created from the baseline model through a relative (40 m/s)

variation of velocity inside two gas reservoirs (Figure 2.4b). A surface acquisition geometry with a free-surface condition is used to generate the synthetic data, with seventy-seven isotropic explosive sources, located along a horizontal line at a depth of 16 m, equally spaced by a distance of 50 m. A horizontal receiver line at a depth of 15 m with a sensor interval of 10 m is used for recording the pressure data for each shot. A Ricker wavelet source with a central frequency of 10 Hz is used, for baseline and monitor surveys. The time seismograms are generated using acoustic finite-difference modeling in the frequency domain with a mixed staggered-grid 9-point stencils (Jo et al., 1996; Hustedt et al., 2004). To generate the time seismograms, the forward modeling in frequency domain is performed between 1 to 30 Hz (frequency bandwidth of source wavelet) with a sampling rate of $df = 0.2\ Hz$, and then the Inverse Fourier Transform is performed to come back to time domain. This sampling frequency is the Nyquist frequency, in other words the maximum recording time is 5 s. As the inversion will be done in frequency domain, the complex frequency data for each simulated frequency were saved at all receiver positions, to be used directly for inversion. The time domain seismograms are only needed to generate synthetic noisy data later. Perfectly-matched layer (PML) absorbing boundary conditions (Berenger, 1994) are used for non-reflecting boundaries except at the top where a free-surface condition is implemented.

3.3.1 Noise-free data

In the first investigation, a noise-free dataset is investigated. Six inversion configurations are done with different sampling frequencies. The inversion starts from 2 Hz until 30 Hz for all cases, since below 2 Hz the energy of source signal is very low. In the first case, all frequencies between 2 and 30 Hz with $df = 0.2\ Hz$ are selected to be inverted simultaneously, this strategy being denoted $B0$. This case could be seen as a time-domain inversion using all the frequency components of the signals. In the second case, a decimation on the selected frequencies is performed and only one frequency out of five is selected, which means $df = 1\ Hz$ (strategy denoted $B1$). For the third case, one out of ten frequencies is selected and so on. Table 3.1 shows the sampling frequency and the number of frequencies used for inversion of each configuration.

	B0	B1	B2	B3	B4	B5
sampling frequency (df) Hz	0.2	1.0	2.0	3.0	4.0	5.0
Number of frequencies	138	28	14	10	7	6

Table 3.1: Sampling frequency and number of frequencies are selected for inversion at each configuration.

For all the computations, a smoothed version of the true baseline model is used as the initial model (Figure 3.2). The smoothness factor used for smoothing the true model is smaller than that used for building the initial model in Asnaashari et al. (2013b), therefore this initial model is kinematically accurate enough for inverting all the frequencies at the same time. Therefore, there is no need to use any regularizations and/or the prior models. This strategy has been chosen, in order to study only the effect of frequency decimation of the data on the inversion without studying the impacts from the model part of objective function. Otherwise, making clear conclusion would be difficult and be affected by several factors. The stopping criterion

of L-BFGS-B optimization is selected to keep constant for all configurations. The stopping criterion is based on the flatness of the objective function for two consecutive iterations.

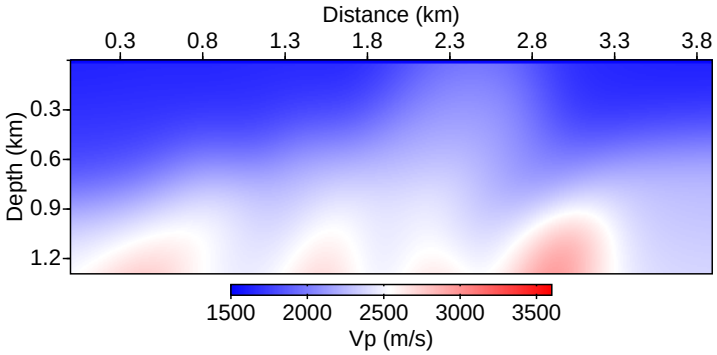


Figure 3.2: The V_p initial model used for inversion. This model is a smoothed version of true baseline model.

Figure 3.3 shows the results of baseline reconstruction for $B0$ to $B5$ configurations. When all frequencies are selected to be inverted ($df = 0.2\ Hz$), almost the perfect model can be obtained with high resolution. Near the border and the bottom of model, the model presents less accuracy due to lack of illumination in those parts. Increasing the sampling frequency results in the degradation and a loss of resolution and accuracy of the derived model. In addition, we can see many inversion artifacts. Due to large sampling frequency (Figure 3.3e and 3.3f), most of the wavenumbers, especially the low and intermediate wavenumbers between two selected frequencies are not recovered, therefore the optimization procedure goes to local minina and several ringing artifacts (high frequency artifacts) appear on the model. By increasing the sampling frequency for inversion and reducing the number of inverted frequencies, the ill-posedness issue of inverse problem increases, and therefore there is higher risk of getting trapped in local minima. The aspect concerning to a hierarchy in the data, introducing data gradually from low to high frequency into the inversion scheme, will be discussed in section 3.5.

Table 3.2 shows the average computational time (Elapsed-time) per core for performing one iteration in several discussed configurations. For all of the configurations, we perfom the inversion with parallel calculation on 8 cores. It can be seen that by increasing the number of inverted frequencies, the computational time will increase almost linearly.

	B0	B1	B2	B3	B4	B5
Number of frequencies	138	28	14	10	7	6
Average Elapsed time per core and per iteration (s)	332.5	68.7	32.3	24.1	15.5	14.1

Table 3.2: Average Elapsed time per core for performing one iteration for each configuration.

To have a quantitative interpretation of the results, the root mean square (RMS) of velocity error is computed based on equation 3.11. This RMS value (η) is used as a model quality factor.

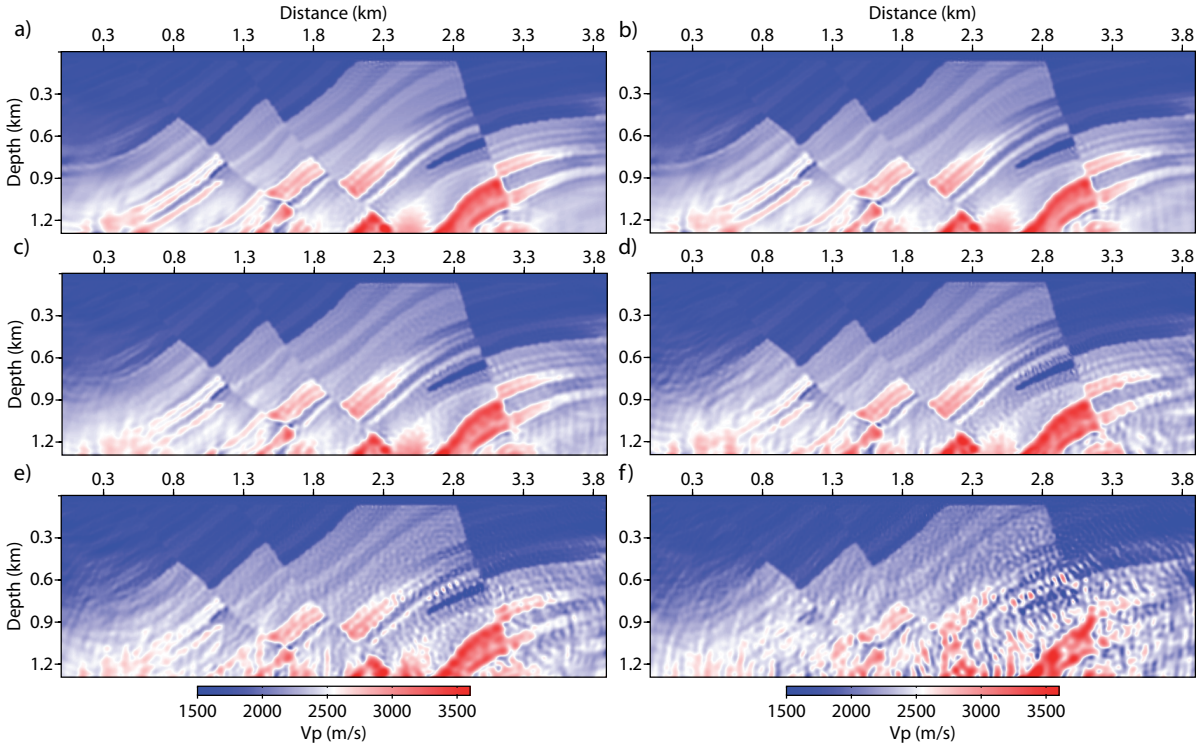


Figure 3.3: Recovered baseline models obtained by FWI of noise-free data in frequency domain and by inverting all selected frequencies at the same time: (a) $B0$ ($df = 0.2\ Hz$), (b) $B1$ ($df = 1\ Hz$), (c) $B2$ ($df = 2\ Hz$), (d) $B3$ ($df = 3\ Hz$), (e) $B4$ ($df = 4\ Hz$), (f) $B5$ ($df = 5\ Hz$).

$$\eta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(m_{i_{rec-b}} - m_{i_{true-b}}\right)^2}, \tag{3.11}$$

where N, m_{rec-b} and m_{true-b} represent the number of unknown model parameter, the recovered and the true baseline model values respectively. Here, the model parameters only contain the P-wave velocity parameters.

The RMS of V_p error versus sampling frequency is plotted in Figure 3.4 for each obtained baseline models . By increasing the sampling frequency and reducing the number of inverted frequencies, the RMS of error increases which means the model quality decreases (as we can observe from Figure 3.3). From Figures 3.3 and 3.4, one can decide that the first three results can be acceptable and they are close to the true model. It means, for the noise-free case, even with quasi-large sampling frequency, we can get acceptable results, *i.e.* we can recognize appropriately the different geological layers. This conclusion is suitable for any exploration project, but may be not valid for time-lapse application. To answer this question, a sensitivity analysis of time-lapse FWI with respect to different obtained baseline models and with respect to frequency decimation will be performed in section 3.4.

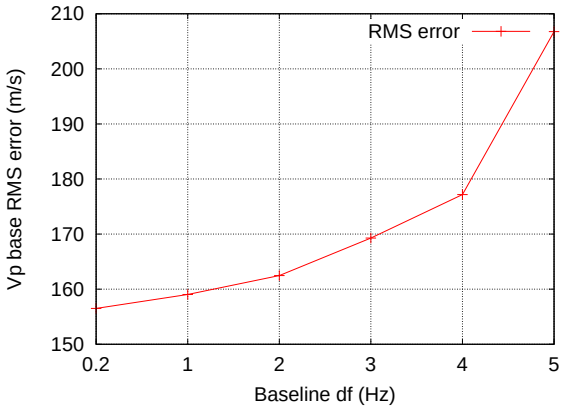


Figure 3.4: RMS of V_p baseline error value for each sampling frequency used for inversion of noise-free data.

3.3.2 Noisy data

Is inversion still robust to the coarse sampling frequency in case of noisy data? In this section, the sensitivity of the recovered baseline model with respect to the sampling frequency is performed in presence of random noise.

An artificial Gaussian noise in the range of $2-30\ Hz$ (the bandwidth of the source wavelet) has been added to the true noise-free data in time domain (using the 'suaddnoise' procedure of Seismic Unix (Cohen and Stockwell, 2008)). The signal-to-noise ratio is around 8 dB. Then the frequency components of all the traces are extracted by Fourier transform, and these complex datasets in frequency domain are used as observed data for inversion.

Different inversion configurations with different sampling frequencies have been investigated, similar to the noise-free case. We stop the investigation at $df = 3\ Hz$, because beyond this point the baseline model is very poorly recovered with a huge number of artifacts. The results of baseline inversion are shown in Figure 3.5. The degradation of resolution and accuracy when increasing the sampling frequency is fairly faster in this noisy environment than in the noise-free case (compare Figures 3.5 and 3.3). In noisy environment, the ill-posedness of the inverse problem increases. The higher values of RMS of V_p error in noisy case confirm this statement (Figures 3.6 and 3.4). In this case, by reducing the number of frequencies to be inverted simultaneously (or increasing sampling frequency), the gradient of objective function is more noisy. In fact, by making summation over the gradients computed for each frequency component, it is possible to reduce the impacts of random noise on inversion procedure (due to increasing signal-to-noise ratio). Therefore, if the number of frequencies increases (using smaller sampling frequency), there are lesser random noise artifacts on the recovered model. We can clearly see this issue in Figures 3.5a to 3.5d. On the other hand, using higher number of frequencies, the computational time and memory resources will rise. Therefore, finding the trade-off between acceptable accuracy and computational expenses is necessary.

In the next section, we will see whether such accuracy of recovered baseline models is sufficient to find small time-lapse variation or not.

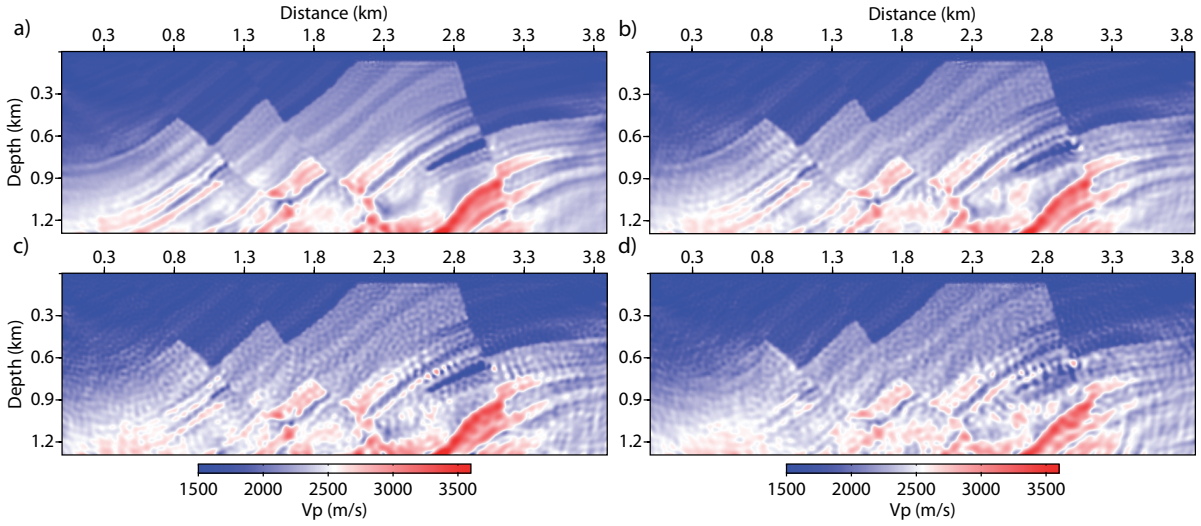


Figure 3.5: Recovered baseline models obtained by FWI of noisy data in frequency domain and by inverting all selected frequencies at the same time: (a) $BN0$ ($df = 0.2\ Hz$), (b) $BN1$ ($df = 1\ Hz$), (c) $BN2$ ($df = 2\ Hz$), (d) $BN3$ ($df = 3\ Hz$).

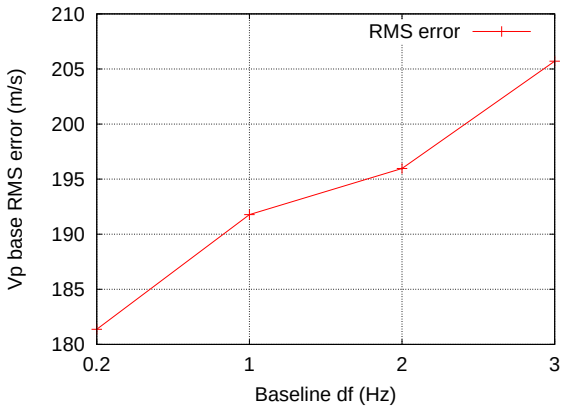


Figure 3.6: RMS of V_p baseline error value for each sampling frequency used for inversion of noisy data.

3.4 Sensitivity of time-lapse models with respect to frequency decimation

Which baseline sampling frequency can deliver more robust and precise enough baseline model for time-lapse inversion? As proposed in this section, a sensitivity analysis of obtained time-lapse model with respect to: 1) baseline frequency decimation used to get the baseline model, and 2) time-lapse frequency decimation on the monitor or composite datasets is performed. It has been shown that even $df = 2\ Hz$ can give an acceptable baseline model (at least in noise-free case), but is this precision of baseline model sufficient to recover small time-lapse variations? The second question is related to which sampling frequency is required in second

step of time-lapse imaging (monitor reconstruction) to get correctly the desired variations. Do we need the same sampling frequency for time-lapse imaging as the one used for baseline reconstruction? Does the selection of frequencies depend on which time-lapse strategy is used? These types of questions will be answered in this section, for noise-free and noisy environments.

3.4.1 Noise-free data

In (Asnaashari et al., 2013b), it has been shown that the parallel difference strategy is very sensitive to the inversion artifacts coming from two independent inversions, which makes it less interesting for time-lapse imaging. However, we should mention for the applications where there is a large time shifts between arrival events in seismograms of baseline and monitor datasets, it is necessary to use the parallel difference strategy (starting both inversion from a same smooth initial model). The kinematics between the two datasets are different, which means there have different background velocity models.

For sensitivity analyses of synthetic Marmousi dataset, we mainly focus on the sequential difference and differential strategies. Here, only for noise-free dataset, some of parallel difference results are shown (Figure 3.7). For each case shown in Figure 3.7, frequency components of both baseline and monitor datasets are extracted with same sampling frequency and the two inversions start from the same smooth initial model (Figure 3.2). It appears that when all the frequencies are inverted ($df = 0.2\ Hz$), the obtained time-lapse model is robust and the location of the two reservoirs is well resolved, although some artifacts are visible due to the two independent inversions. By decreasing the number of frequencies, the time-lapse model starts to degrade, especially in the zones presenting a lesser coverage of illumination. In this case, the convergence path of optimization procedure for the two datasets is totally different and this leads to have different recovered velocity values (for baseline and monitor models). Therefore, it is really difficult to even detect the location of true time-lapse variations.

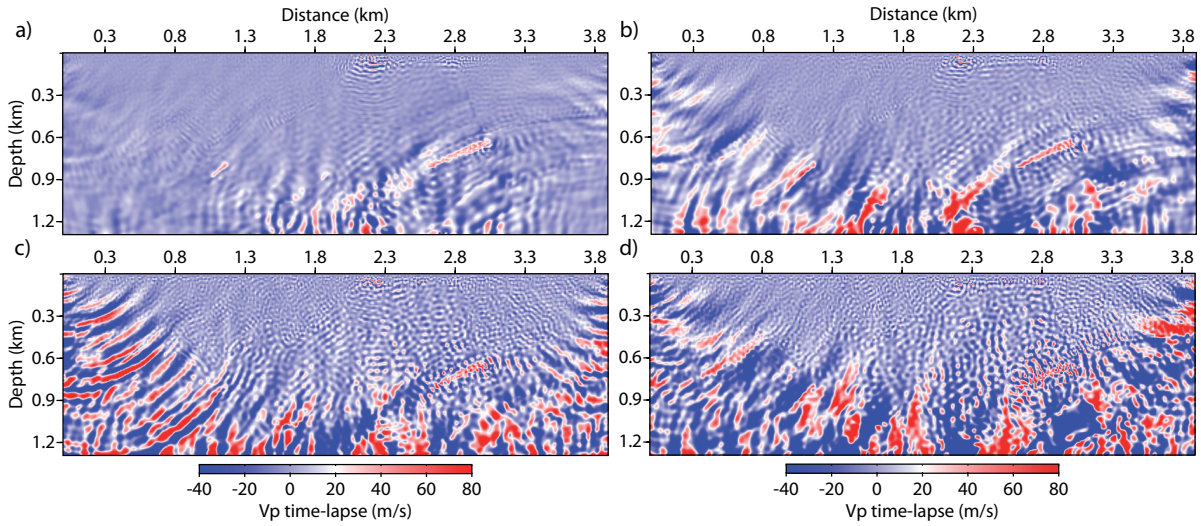


Figure 3.7: Time-lapse variation models obtained from the parallel difference strategy applied to noise-free dataset, using similar sampling frequency for both baseline and monitor reconstruction: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$, (d) $df = 3\ Hz$.

For the sequential difference and differential strategies, the recovered baseline model is used as an initial model to invert respectively monitor and composite datasets. With six different sampling frequencies, six different baseline models have been obtained. The models with RMS of V_p error less than 200 m/s are selected for time-lapse inversion (first five models in Figure 3.3). For each baseline model, six tests are done using six different sampling frequencies (same as Table 3.1), leading to thirty configurations for each time-lapse strategy, sequential difference and differential.

Only the most interesting results are displayed here. Figures 3.8 and 3.9 show the results of time-lapse variation obtained by the sequential difference and differential strategies, respectively, for the baseline model $B0$ (Figure 3.3a) obtained by inverting all the frequencies ($df = 0.2\ Hz$). The sequential difference strategy is more sensitive to the time-lapse sampling frequency than the differential strategy. In the sequential difference strategy, increasing the sampling frequency results in decreasing the robustness of the obtained time-lapse model. A coarse time-lapse sampling frequency cannot deliver a robust time-lapse variation model (even starting from the most accurate baseline model $B0$). However, in the differential method, the rate of degradation of the time-lapse model is slower and even with a large time-lapse sampling frequency of $df = 5\ Hz$ (Figure 3.9f), acceptable time-lapse results can be recovered although the velocity of the smaller reservoir is under-estimated. The reason is that the differential strategy inverts directly the difference data related to the time-lapse variations. In clean (noise-free) dataset even with a coarse sampling frequency, we are able to roughly sample the difference data while in the sequential strategy, the sampled difference data are contaminated by data residuals coming from non-fully fitted baseline events.

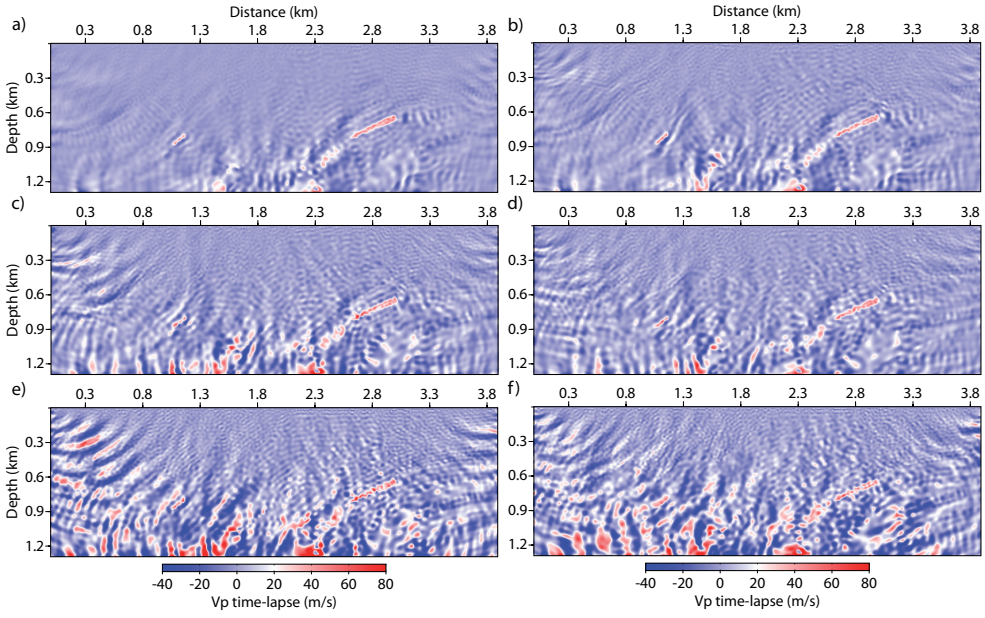


Figure 3.8: Time-lapse variation models obtained from the sequential difference strategy applied to noise-free dataset, starting from the recovered baseline model **B0** (**df = 0.2 Hz**) and frequency decimation on monitor dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$, (d) $df = 3\ Hz$, (e) $df = 4\ Hz$, (f) $df = 5\ Hz$.

Figures 3.10 and 3.11 show the results of time-lapse variation obtained by the sequential

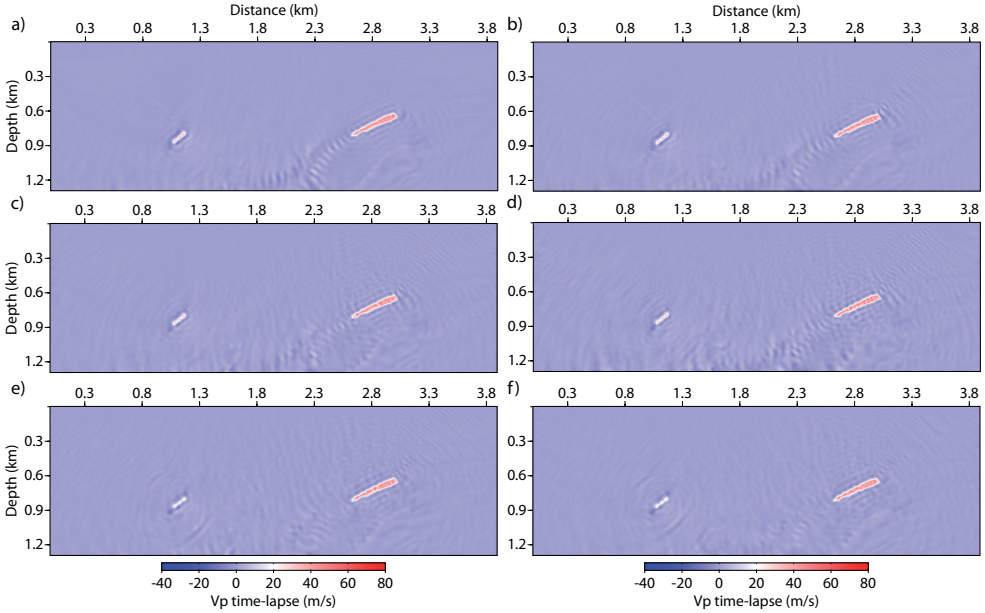


Figure 3.9: Time-lapse variation models obtained from the differential strategy applied to noise-free dataset, starting from the recovered baseline model **B0** (**df = 0.2 Hz**) and frequency decimation on composite dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$, (d) $df = 3\ Hz$, (e) $df = 4\ Hz$, (f) $df = 5\ Hz$.

difference and differential strategies, respectively, for the baseline model $B1$ ($df = 1\ Hz$) (Figure 3.3b). This baseline model seems accurate enough visually and in terms of velocity error RMS, but for the sequential difference strategy in time-lapse application it is actually not satisfactory. Indeed, the sequential difference is very sensitive to the inaccuracy of baseline model (Asnaashari et al., 2013b). When the time-lapse sampling frequency does not match with and is smaller than the baseline sampling frequency, it seems the sequential difference method attempts to recover the frequency components of baseline model which have not been recovered before. The monitor dataset contains the time-lapse information and the repeatable information between baseline and monitor. Therefore, when the monitor dataset is used for inversion with denser frequency samplings, these new frequency components are able to update the model outside of the reservoir variation. Thus, there is a risk to interpret them as a potential time-lapse variation.

When the time-lapse sampling frequency is larger than the baseline sampling frequency, due to the missing frequencies between the baseline and monitor datasets, the monitor dataset used for time-lapse inversion contains less repeatable information than the baseline inversion. In other words, the redundant data, which can constrain the sequential difference inversion around the baseline points, are removed from the observed data. Therefore, due to the lack of these repeatable data constraints, the optimization procedure in time-lapse step goes to the incorrect direction of changing the model outside of time-lapse target. As we can see from Figure 3.10, for sequential difference strategy, only minimal artifacts are visible and a more robust result is obtained when the same sampling frequency is used for both baseline and time-lapse reconstructions $df_{time-lapse} = df_{baseline}$. The differential strategy is less sensitive to the time-lapse sampling frequency, because only the differential dataset is inverted during

optimization. However, using a larger number of frequencies of differential data leads to a better focusing of the energy of the time-lapse signal on the area of changes, especially for the smaller reservoir.

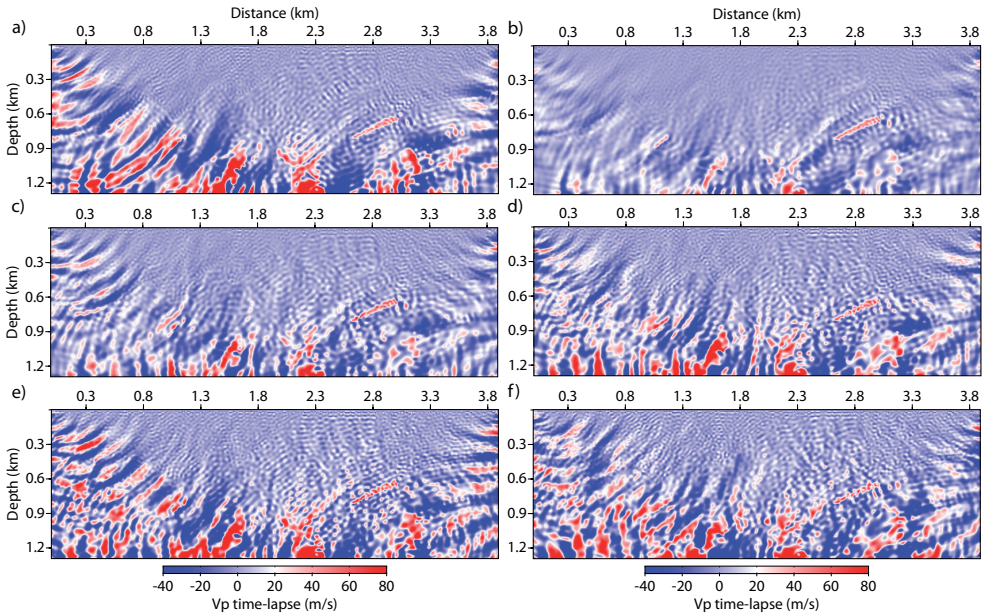


Figure 3.10: Time-lapse variation models obtained by the sequential difference strategy applied to noise-free dataset, starting from the recovered baseline model **B1** (**df = 1 Hz**) and frequency decimation on monitor dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$, (d) $df = 3\ Hz$, (e) $df = 4\ Hz$, (f) $df = 5\ Hz$.

The time-lapse models obtained from the baseline model $B3$ ($df = 3\ Hz$) (Figure 3.3d) are shown in Figures 3.12 and 3.13. All the already discussed problems can be seen in this case as well. In the sequential difference strategy, even with the same sampling frequency for time-lapse and baseline reconstruction, the true time-lapse variation is not recovered. The reconstruction of time-lapse inside the reservoirs is affected by many artifacts due to the inaccuracy of the baseline model. However, the differential strategy is still robust despite the inaccuracy of the baseline model. We can clearly observe in Figure 3.13 that the time-lapse energy spreads everywhere because of the inaccuracy of the baseline model. The time-lapse model contains ringing effects and discontinuity on time-lapse variation. This effect is reduced when the same sampling frequency is used for both time-lapse and baseline step (Figure 3.13d). However, this issue is less evident for the differential strategy than for the sequential one, therefore this requirement is less important for the differential to be satisfied.

In order to have a better global view and a quantitative interpretation of the time-lapse results, the RMS of time-lapse V_p error is computed,

$$\epsilon = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\delta m_{i_{rec-timelapse}} - \delta m_{i_{true-timelapse}}\right)^2}. \tag{3.12}$$

In this equation, N, $\delta m_{rec-timelapse}$ and $\delta m_{true-timelapse}$ represent the number of unknown model parameters, the recovered and the true time-lapse variation model values respectively.

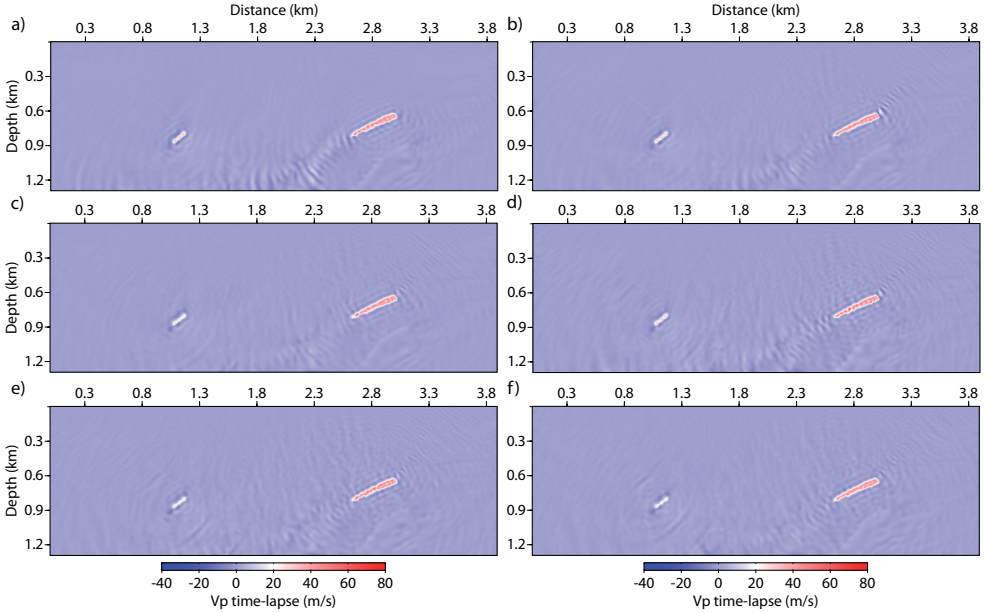


Figure 3.11: Time-lapse variation models obtained by the differential strategy applied to noise-free dataset, starting from the recovered baseline model **B1** (**df = 1 Hz**) and frequency decimation on composite dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$, (d) $df = 3\ Hz$, (e) $df = 4\ Hz$, (f) $df = 5\ Hz$.

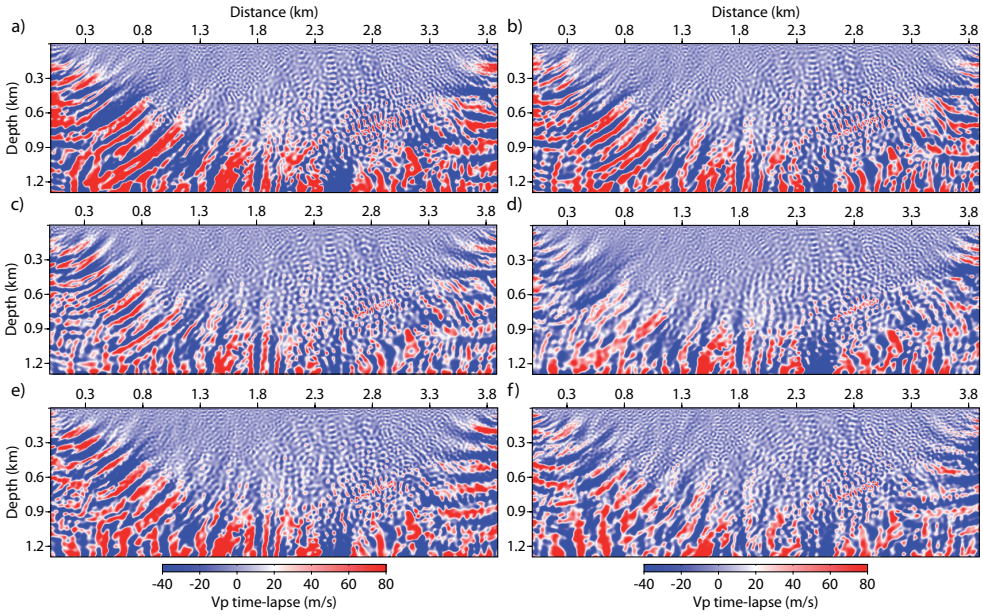


Figure 3.12: Time-lapse variation models obtained by the sequential difference strategy applied to noise-free dataset, starting from the recovered baseline model **B3** (**df = 3 Hz**) and frequency decimation on monitor dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$, (d) $df = 3\ Hz$, (e) $df = 4\ Hz$, (f) $df = 5\ Hz$.

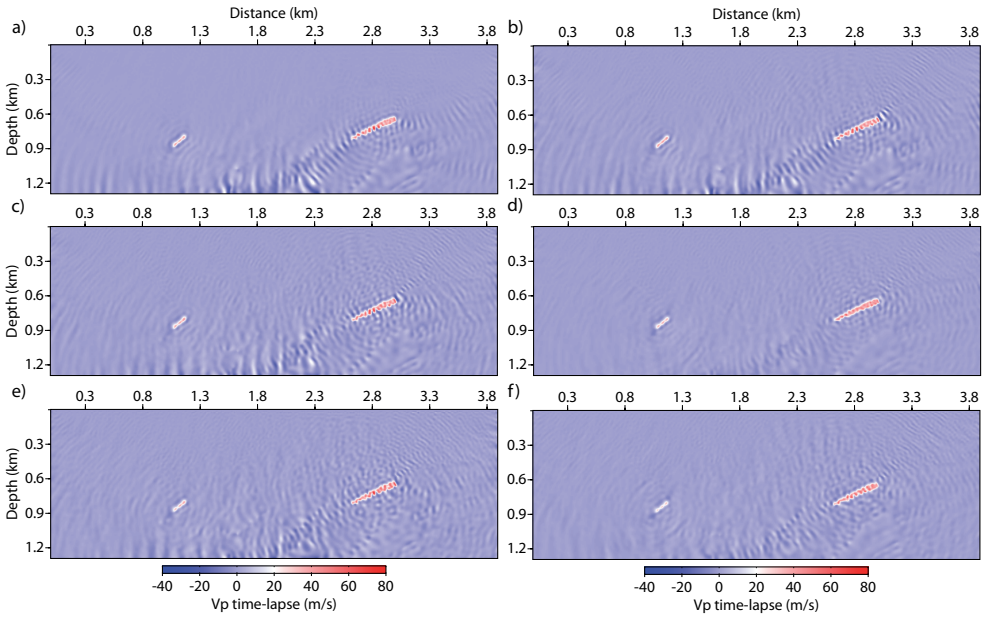


Figure 3.13: Time-lapse variation models obtained by the differential strategy applied to noise-free dataset, starting from the recovered baseline model **B3** (**df = 3 Hz**) and frequency decimation on composite dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$, (d) $df = 3\ Hz$, (e) $df = 4\ Hz$, (f) $df = 5\ Hz$.

The RMS of time-lapse velocity errors for the sequential and differential strategy are plotted in 2D maps (Figure 3.14). The horizontal and vertical axes of this map represent the time-lapse and baseline sampling frequencies, respectively.

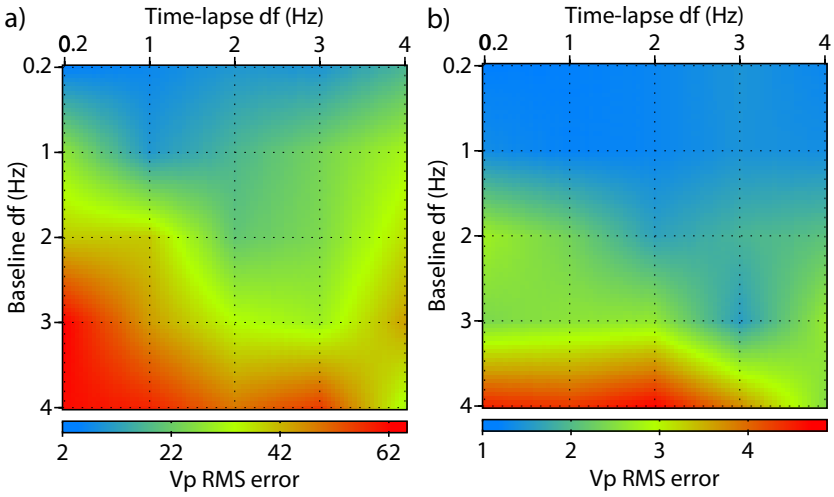


Figure 3.14: RMS map of time-lapse V_p error versus time-lapse and baseline sampling frequency used for inversion of noise-free data: (a) sequential difference strategy, (b) differential strategy.

Generally, the differential strategy is less sensitive to the frequency decimation in baseline and time-lapse step compared to the sequential one, as shown by very low RMS errors (at least

for noise-free case). The minimum of RMS error for each baseline model happens when a same sampling frequency is used for time-lapse step (it means that on the main diagonal line of map, $df_{time-lapse} = df_{baseline}$). This issue is more important for the sequential difference because of larger differences between RMS values on this line and other points of the map. Another piece of information that we can extract from comparing both maps is that the sequential method is more sensitive to the inaccuracy of baseline model (or to the baseline frequency decimation) than the differential strategy. The RMS error increases faster in vertical direction in case of sequential strategy.

3.4.2 Noisy data

Similar sensitivity tests have been conducted for noisy data. Here, only three baseline models have been chosen (*i.e.* $BN0$, $BN1$, $BN2$ in Figure 3.5), because for $df > 2\ Hz$ the recovered baseline model is poorly recovered and has a high RMS error value. Also, only the first three time-lapse sampling frequency values are selected for time-lapse inversion, leading to nine time-lapse inversion configurations for each strategy (sequential difference and differential).

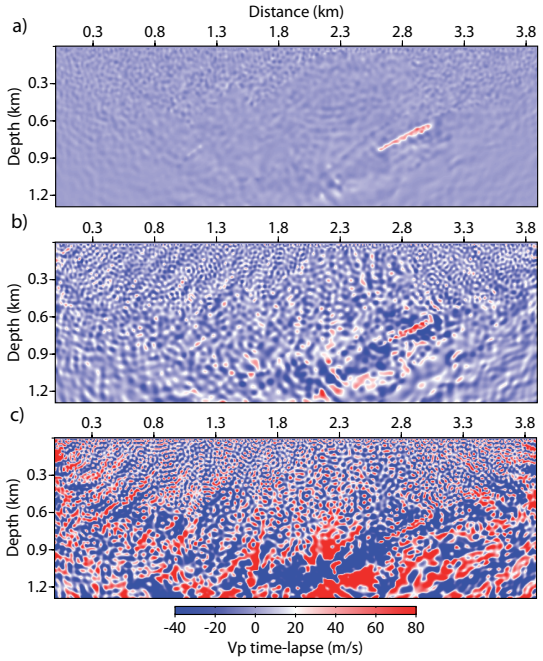


Figure 3.15: Time-lapse variation models obtained by the sequential difference strategy applied to noisy dataset, starting from the recovered baseline model **BN0** (**df = 0.2 Hz**) and frequency decimation on monitor dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$.

Figures 3.15 and 3.16 show the results of time-lapse variations obtained by the sequential difference and differential strategies, respectively, for the baseline model $BN0$ (Figure 3.5a). In presence of random noise, the complexity of data increases, therefore we need to take into account more frequencies to reduce the ill-posedness issue. When all the time-lapse frequencies are used (Figure 3.15a and 3.16a), both strategies converge to almost similar results, although, the sequential strategy tries to recover a little below the bigger reservoir. Also, due to high level

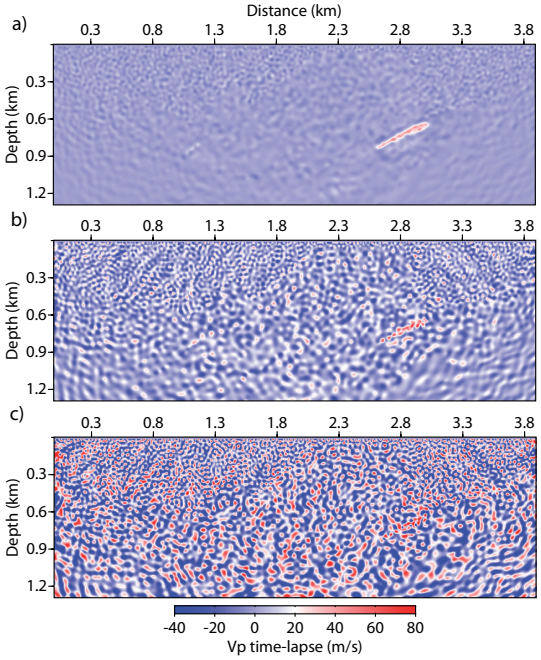


Figure 3.16: Time-lapse variation models obtained by the differential strategy applied to noisy dataset, starting from the recovered baseline model **BN0** (**df = 0.2 Hz**) and frequency decimation on composite dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$.

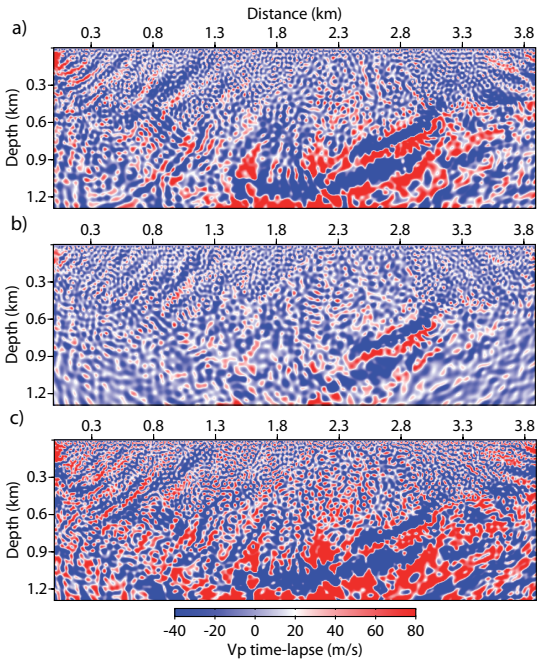


Figure 3.17: Time-lapse variation models obtained by the sequential difference strategy applied to noisy dataset, starting from the recovered baseline model **BN1** (**df = 1 Hz**) and frequency decimation on monitor dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$.

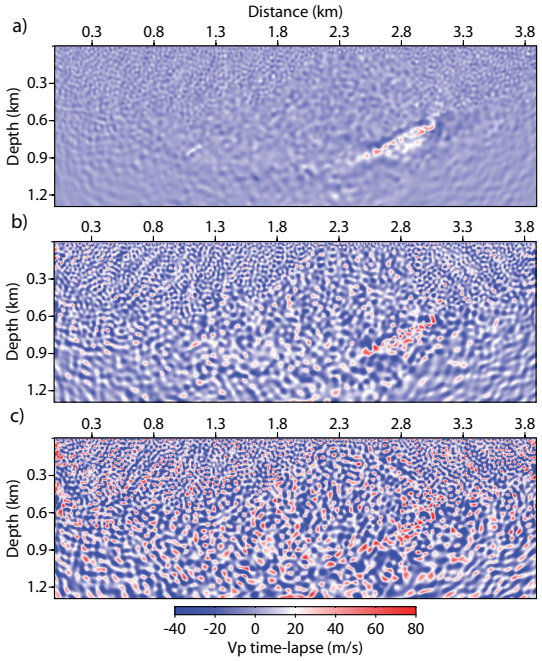


Figure 3.18: Time-lapse variation models obtained by the differential strategy applied to noisy dataset, starting from the recovered baseline model **BN1** (**df = 1 Hz**) and frequency decimation on composite dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$.

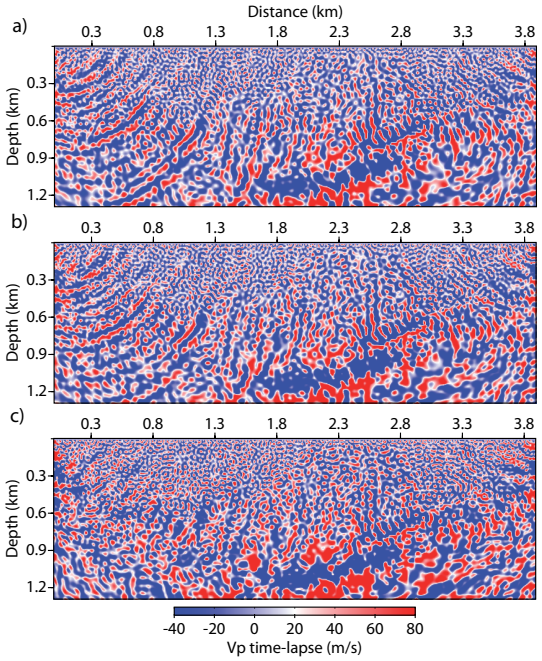


Figure 3.19: Time-lapse variation models obtained by sequential difference strategy applied to noisy dataset, starting from the recovered baseline model **BN2** (**df = 2 Hz**) and frequency decimation on monitor dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$.

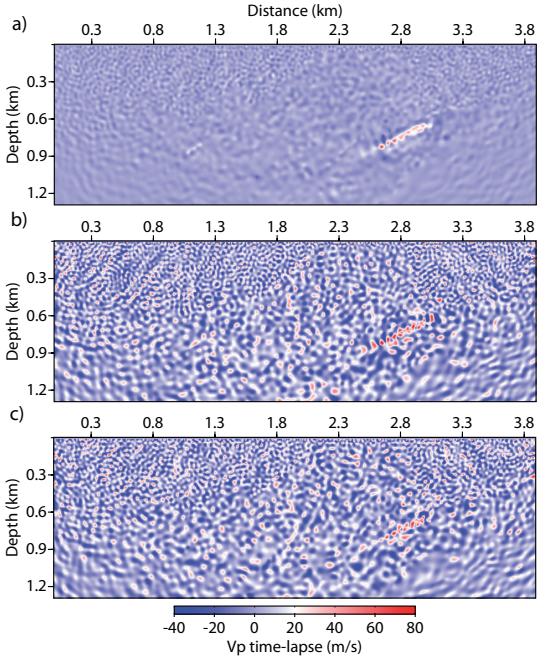


Figure 3.20: Time-lapse variation models obtained by differential strategy applied to noisy dataset, starting from the recovered baseline model **BN2** (**df = 2 Hz**) and frequency decimation on composite dataset: (a) $df = 0.2\ Hz$, (b) $df = 1\ Hz$, (c) $df = 2\ Hz$.

of noise, the smaller reservoir is not well detected. By reducing the number of inverted time-lapse frequencies, both strategies are affected by the presence of noise. In case of $df = 1\ Hz$, the location of the larger reservoir can still be detected but for $df = 2\ Hz$ it is very difficult to identify the location of the true time-lapse variation. It appears that the differential strategy is driven by random noise. By reducing the number of simultaneous inverted frequencies, the signal-to-noise ratio decreases and this affects the computation of the gradient of objective function; therefore the low energy time-lapse signal has a lesser impact on the model updating and the optimization is controlled by random noise.

Comparison of the time-lapse results for other baseline models in noisy case shows the consistency of our conclusions between noisy and noise-free cases, where the noisy case is more sensitive to the time-lapse frequency decimation, due to the complexity of data. Another interesting point is that the artifacts in the sequential strategy have a structured shape while in the differential case they are similar to high frequency artifacts and noisy images. This means the sequential difference strategy goes to another local minimum and attempts to recover the baseline structure (either in good recovering direction or wrong direction), but the differential strategy is driven by high level of the noise in differential dataset.

In case of baseline $BN2$ ($df = 2\ Hz$) (Figures 3.19 and 3.20), since the baseline model is not accurate and precise enough and also because of noise, the sequential difference cannot converge to robust results. However, the differential strategy at least for $df_{timelapse} = 0.2\ Hz$ can provide the correct time-lapse variation of the larger reservoir.

Time-lapse velocity RMS errors for the sequential and differential strategies are plotted in 2D maps (Figure 3.21). Generally, for the same inversion configuration, the RMS error

is higher in presence of noise than in the noise-free case. Similar to the noise-free case, for the sequential difference strategy, the minimum of RMS error for each baseline model occurs when the same sampling frequency is also used for time-lapse step ($df_{time-lapse} = df_{baseline}$). In noisy environment, we can see that the differential strategy is more sensitive to the time-lapse frequency decimation than to the baseline sampling frequency. There is a faster and larger variation of RMS error in horizontal direction than in the vertical one, because reducing an already limited number of frequencies leads to a decrease of the signal-to-noise ratio of the differential dataset.

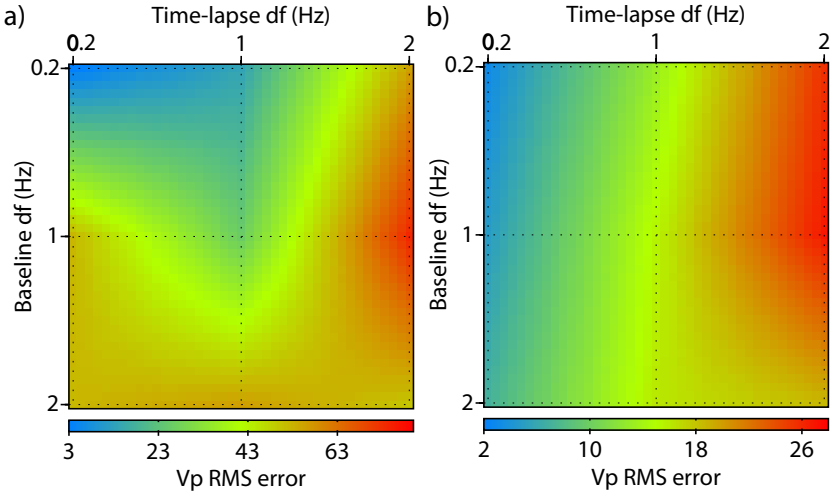


Figure 3.21: RMS map of time-lapse V_p error versus time-lapse and baseline sampling frequency used for inversion of noisy data: (a) sequential difference strategy, (b) differential strategy.

After these sensitivity analyses, we can conclude that to get accurate time-lapse variations, it is necessary to increase the number of inverted frequencies in order to obtain a more precise baseline model. This issue is more crucial for the sequential difference strategy. On the other hand, in order to have less time-lapse artifacts in the inversion process, a same sampling frequency for the baseline and time-lapse inversions should be used (related to the data repeatability issue), neither greater nor smaller. The differential strategy is less sensitive to the baseline frequency decimation. The rate of model degradation is slower compared to the sequential difference cases. For noisy data, we need a smaller baseline sampling frequency to get an accurate baseline model and also a smaller time-lapse sampling frequency for both time-lapse strategies. Therefore, it is crucial to increase the number of inverted frequencies for time-lapse applications as compared to the general exploration project. On the other hand, in simultaneous frequency inversion due to high computational cost for LU decomposition of impedance matrix and solving the wave-equation for each frequency component, it might be better to move to time-domain modeling and inversion for time-lapse applications, so as to solve the wave-equation once for all the frequencies. However, an efficient time-domain inversion algorithm is required.

3.5 Multi-group multi-frequency strategy

In frequency domain, several hierarchical multi-scale strategies that proceed from low frequencies to higher frequencies have been proposed to reduce the non-linearity of the inverse problem and the cycle-skipping problem (Pratt and Worthington, 1990; Bunks et al., 1995; Pratt, 1999; Sirgue and Pratt, 2004; Brossier et al., 2009). In the previous section, we used the approach of inverting all the selected frequencies simultaneously. There is however a possibility to combine the two approaches, low to high frequencies and simultaneous frequency inversion, by a multi-group approach where each of this group contains a small number of frequencies (Brossier et al., 2009; Brossier, 2011). The inversion within each group of frequencies is done by inverting all the frequencies of the group at the same time, then after a limited number of optimization iterations, inversion moves to second group of frequency and so on. In order to have a more robust result, we can have redundant information between groups. Therefore, we can repeat the last frequency of each group on the first frequency index of the next group (making overlap between groups). This approach, because of solving first for low wavenumbers and gradually moving to higher frequencies, could reduce the cycle-skipping problem. The attraction basin of the objective function should start from a wide valley and gradually move to a narrow one. In this section, some tests are performed to evaluate whether the time-lapse strategies are compatible with these consecutive multi-group inversions. Which frequency strategy, simultaneous inversion of all selected frequencies or gradually sweeping frequencies, can deliver more robust time-lapse images?

3.5.1 Baseline model reconstruction

It has been shown before that the best case of simultaneous frequency inversion is the case when all the frequencies (138 frequency components) are used for time-lapse application. However, the sampling frequency $df = 1\ Hz$ (28 frequency components) delivered acceptable time-lapse results for both time-lapse strategies, at least in noise-free case, with huge reduction in computational time (Table 3.2). Instead of doing forward modeling for 138 frequencies at each iteration, it is only performed for 28 frequencies, therefore we have a large gain in computational time. Thus, $df = 1\ Hz$ is chosen to select the frequencies of multi-group multi-frequency strategy. We choose to have three frequency components at each group with last frequency repeating for the next group. Table 3.3 shows the selected frequencies for each group.

	G1	G2	G3	...	G12	G13
Frequency (Hz)	2, 3, 4	4, 5, 6	6, 7, 8	... ...	24, 25, 26	26, 27, 28

Table 3.3: Table of multi-group multi-frequency, each group contains three frequencies and there is an overlap of one frequency index between two consecutive groups.

This multi-group inversion (Algorithm 3.2) has been applied to baseline noise-free and noisy datasets. For the baseline reconstruction, the maximum number of optimization iterations for each inversion group is chosen to be 20 iterations, although there is a possibility that the L-BFGS-B procedure can stop before if it satisfies the convergence criterion. The final model of each group is used as an initial model for the next group. Figure 3.22 shows the final results for

Algorithm 3.2 Multi-group multi-frequency FWI algorithm

for $G_i = 1$ to ng **do**
  while (NOT *convergence* AND $k < iter_{max}$) **do**
    for $\omega_i = 1$ to $n\omega$ **do**
      Compute incident wavefields $\mathbf{u}$ from sources
      Compute residual vectors $\Delta\mathbf{d}$ and objective function $\mathcal{C}^{k-1}$
      Compute adjoint back-propagated wavefields $\mathbf{A}^{-1}\mathbf{S}^t\mathbf{W}_d\Delta\mathbf{d}^*$ from receivers
      Build gradient vector $\mathcal{G}^{k-1}$ and make summation over frequencies
    end for
    Compute perturbation vector $\Delta\mathbf{m}^k$ and optimal step length $\alpha^k$with L-BFGS-B optimization
    Update model $\mathbf{m}^k = \mathbf{m}^{k-1} + \alpha^k\Delta\mathbf{m}^k$
  end while
end for

this multi-group inversion in noise-free and noisy cases. Comparison between these results and the results of simultaneous inversion (Figure 3.3b and 3.5b) illustrates that for noise-free case the two inversion strategies deliver almost the similar baseline model with small differences (see the QC vertical logs in Figure 3.23). For noisy case, the result of multi-group inversion is less accurate and contains more random artifacts due to random noise. This can be explained by the stack of a smaller number of frequencies at each iteration. In simultaneous inversion, a summation of gradient of objective function over all the selected frequencies at each iteration is performed, while in multi-group inversion there is a summation only over three frequencies (for each group). Therefore the gradient and consequently the Hessian contain more imprints of noise.

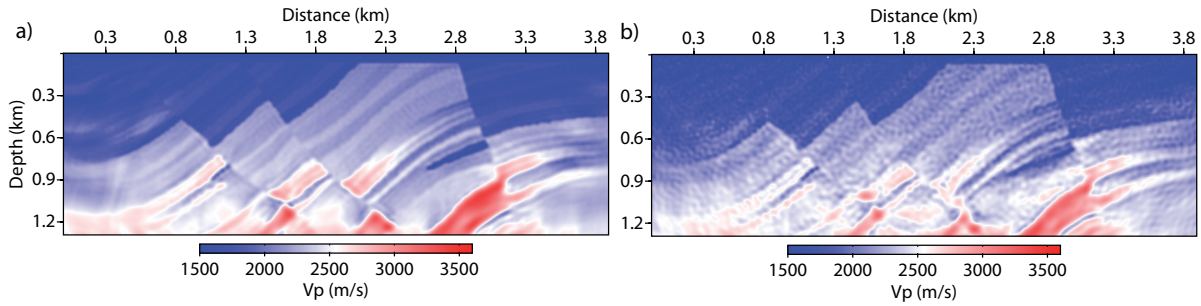


Figure 3.22: Baseline models obtained by multi-group multi-frequency inversion of: (a) noise-free data, (b) noisy data.

3.5.2 Time-lapse model reconstruction

The recovered baseline models are used to recover time-lapse variations through the sequential difference and the differential strategies. At the time-lapse stage, five iterations are chosen for the maximum number of iterations. As there is no large variation between baseline and monitor models, we need a smaller number of iterations for the time-lapse reconstruction. The

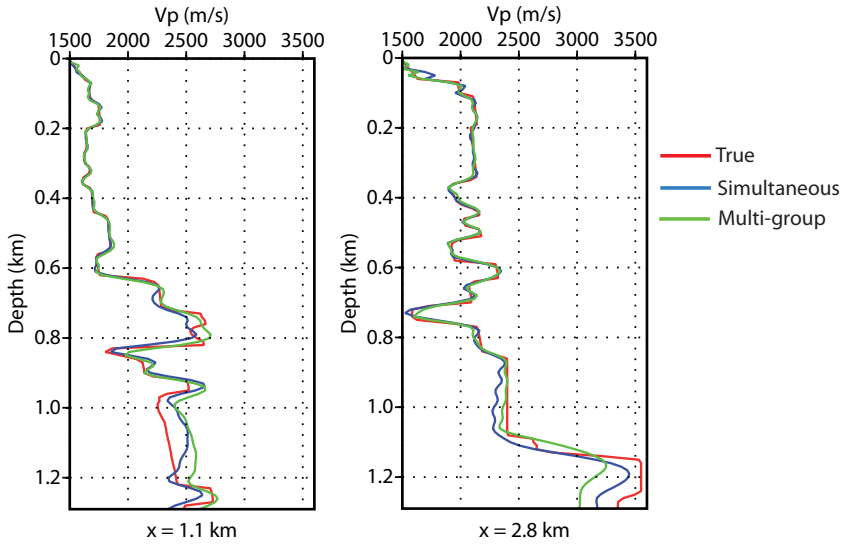


Figure 3.23: Two QC vertical logs passing through the two target areas at $x = 1.1\ km$ and $x = 2.8\ km$, comparing the true baseline model, the recovered model using simultaneous frequency inversion (Figure 3.3b) and the recovered model using multi-group inversion (Figure 3.22a) in noise-free case.

results of the noise-free case are shown in Figure 3.24. It appears these results are less accurate and less robust than the time-lapse results obtained for simultaneous inversion (Figures 3.10b and 3.11b and compare the QC vertical logs in Figure 3.25). We can see that the differential result is less accurate and contains large wavelength artifacts especially at the border. The problem shown here is that the high resolution baseline model is used as the starting model for time-lapse imaging and at the first step (group 1), the inversion tries to reconstruct the low wavenumbers on the high resolution baseline model. Probably, the reconstruction of long wavelengths on the model which already contains short wavelengths makes problems. In the differential strategy, these problems and artifacts are less important than in the sequential difference, as the differential inversion concentrates on the low energy difference datasets.

In noisy environment, in addition to the previous problems we should add other problems such as the increasing ill-posedeness issue caused by the presence of noise and also the fact that we are using a smaller number of frequencies at each iteration (thus reducing signal-to-noise ratio). Therefore, the obtained time-lapse models are not robust and trustable at all (Figure 3.26).

To conclude this part, it seems the multi-group multi-frequency inversion can provide the acceptable precision of baseline model (or generally for exploration project), but this approach can be less interesting for time-lapse applications. For time-lapse applications, starting from a high resolution baseline model and using multi-group inversion from low to high frequency, it appears the reconstruction of low frequency components on the initial high frequency model causes problems. This leads to the introduction of many artifacts on the time-lapse model, even in the noise-free case. Therefore, inversion of all the selected frequencies at the same time appears more compatible with both the sequential difference and the differential strategies.

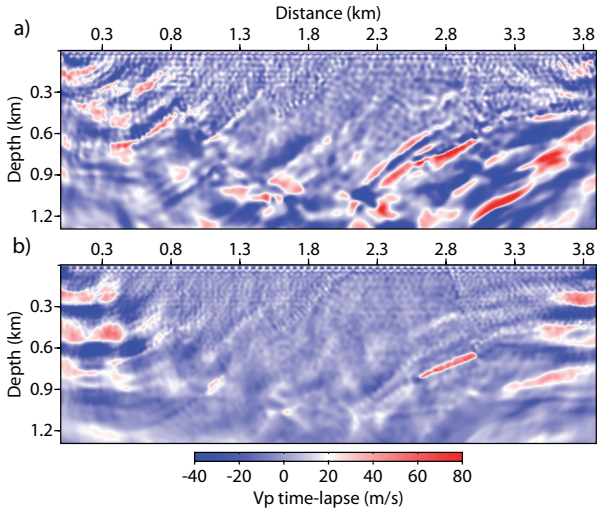


Figure 3.24: Time-lapse variation models obtained by multi-group multi-frequency inversion of noise-free dataset through: (a) the sequential difference, (b) the differential strategy.

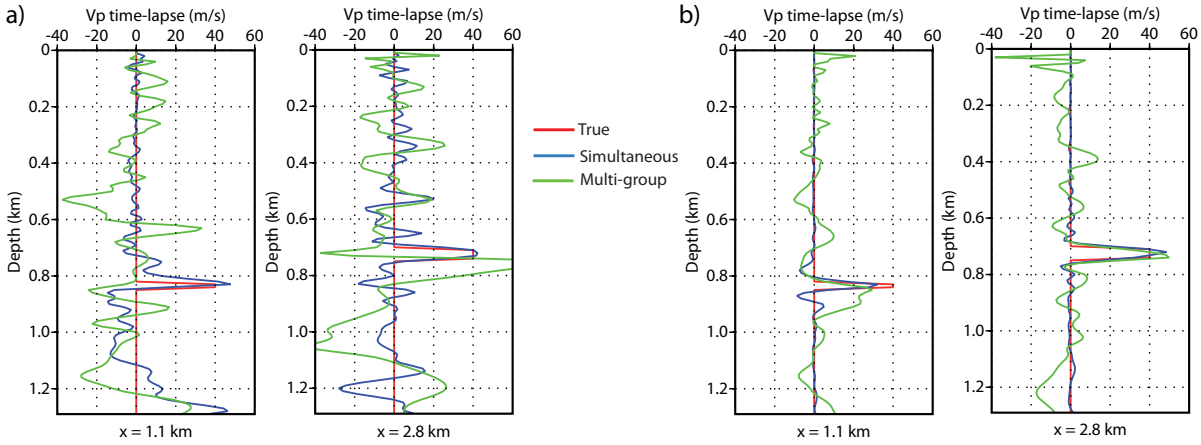


Figure 3.25: Two QC vertical logs passing through the two target areas at $x = 1.1\ km$ and $x = 2.8\ km$, comparing the true time-lapse model, the recovered time-lapse model using simultaneous frequency inversion and the recovered model using multi-group time-lapse inversion in noise-free case with: (a) the sequential difference strategy, (b) the differential strategy.

3.6 Conclusions

For simultaneous frequency inversion, we can conclude that getting an accurate time-lapse variations requires to increase the number of inverted frequencies in order to obtain a more precise baseline model. This fact is more important for the sequential difference strategy. On the other hand, in order to have less time-lapse artifacts in the inversion process, a same sampling frequency for the baseline and time-lapse inversions should be used (related to the data repeatability issue). The differential strategy is less sensitive to the baseline frequency decimation. The speed of model degradation is slower compared to the sequential difference cases.

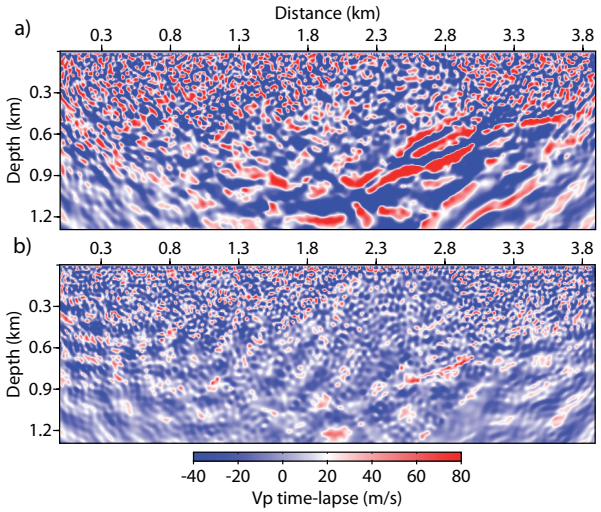


Figure 3.26: Time-lapse variation models obtained by multi-group multi-frequency inversion of noisy dataset through: (a) the sequential difference, (b) the differential strategy.

For noisy data, we need a smaller baseline sampling frequency to get accurate baseline model and also smaller time-lapse sampling frequency for both time-lapse strategies. Therefore, it is crucial to increase the number of inverted frequencies for time-lapse applications as compared to the general exploration project. Due to the large number of inverted frequencies, it might be better to move into time-domain modeling and inversion for time-lapse applications to solve the wave-equation once for all the frequencies.

Multi-group multi-frequency inversion is another interesting strategy for FWI in frequency domain, but this approach appears less interesting for time-lapse applications. Starting from a high resolution baseline model and using multi-group inversion from low to high frequencies of the monitor or composite datasets, the multi-scale strategy does not work properly, since the inversion attempts to reconstruct low wavenumbers on a model that already contains high wavenumbers. This leads to the generation of many artifacts on the time-lapse model, even in the noise-free case. In this case, it is possible that considering a prior information could mitigate the problem. We can conclude that the inversion of all the selected frequencies at the same time is more compatible with both the sequential difference and the differential strategies. Probably, the inversion of low to high frequency hierarchy could be more compatible with the parallel difference strategy, where two independent inversions (baseline and monitor) start from a same smooth initial model. In this case, the wavenumbers are gradually recovered from low to high contents, independently, for two datasets.

These sensitivity analyses provided in this study suggest that a more interesting approach for time-lapse imaging could be the full waveform inversion conducted in the time domain where all the frequency components are inverted together at the same time. An efficient time-domain inversion algorithm is required. If we have reasonable number of sources, doing inversion in time domain is really less expensive and less time-consuming compared to doing inversion in frequency domain for all the frequencies. Probably, including proper prior information into the time-lapse inversion may reduce demanding of large number of frequencies to be considered. If it is possible, in this case, we can keep the frequency-domain inversion and reach the aim of data compressing. This would be considered as a future research.

Chapter 4

Application to Real Field Data

Contents

In this chapter, the different time-lapse inversion strategies presented before are applied to real field data where steam injection has been monitored.

4.1 Introduction

The field is located in North Canada. A geology overview of this field is shown in Figure 4.1. The lower Cretaceous McMurry Fm layer is the principal oil sand reservoir. These Fluvial and estuarine channelized deposits are located on top of Devonian unconformity. An unconformity is a buried erosional or non-depositional surface separating two rock masses or strata of different ages, indicating that sediment deposition was not continuous. The high velocity of the Devonian layer implies that it will be considered as a basement. The main reservoir contains unconsolidated sands at depth of 250 to 500 m with a thickness of 30 to 70 m. Due to the high viscosity of oil sands, a thermal production is required. It consists in injecting high temperature steam into the reservoir at different places, to change the property of oil, especially its viscosity in order to make it less viscous and facilitate the production.

In literature, rock physics studies show that thermal enhanced oil recovery (EOR) processes, especially steam injection, should be visible in repeated surface seismic surveys (Greaves and Fulp, 1987; Lumley, 1995a; Jenkins et al., 1997; Lumley, 2001). Indeed, the injected steam in a reservoir can significantly decrease the rock seismic impedance, due to substitution of fluid (presence of free gas) and also to associated temperature change.

Several 3D monitoring surveys have been acquired on this field at different times, in order to study the variations of reservoir and analyze the geometry of injected steam area. In our study, we deal with two monitor datasets which were acquired after the initial steam injection. I should mention that the M1 dataset already contains the effect of injected steam, because this dataset does not correspond to the ideal baseline case acquired before steam injects in. Between these two seismic acquisition times, more steam has been reinjected into the reservoir. Therefore, the goal of this time-lapse study is to identify and detect potential velocity anomalies related to this new injection.

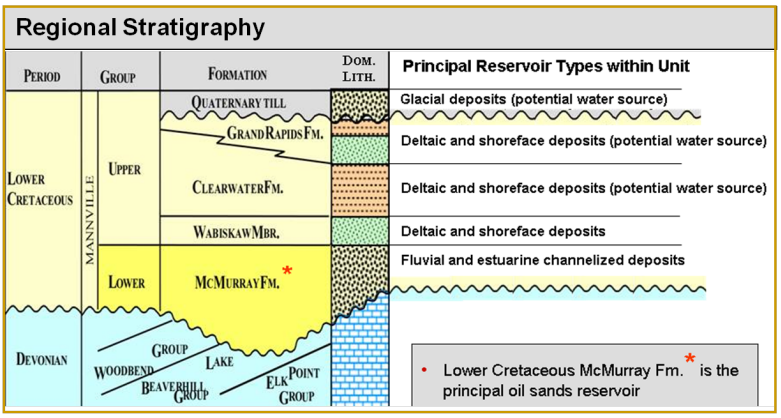


Figure 4.1: Geological overview of the field. Lower Cretaceous McMurry Fm is the principal oil sands reservoir.

4.2 Seismic datasets

For this study, two 2D repeated datasets (M1 and M2) are provided. The original datasets were acquired as a 3D land acquisition, whose location of receivers is shown in Figure 4.2. The 2D acquisition line studied here, which intersects steam chambers, has been extracted from the 3D dataset by interpolation. This step was performed by CGG company using a 5D interpolation (inline, crossline, offset, azimuth and time) strategy.

The obtained pseudo 2D line consists of a series of virtual shot and receiver points spaced by 10 m all along the line (Figure 4.3). There are 140 shot points and 140 receivers. As all the receivers are active for each shot, the total number of interpolated traces is $140 \times 140 = 19600$. The minimum and maximum offsets correspond to 5 m and 1395 m, respectively.

Several processing steps were performed to prepare the datasets for the 5D interpolation (by CGG):

- M2 seismic moved to the M1 floating datum by applying M1 regional correction. Indeed,

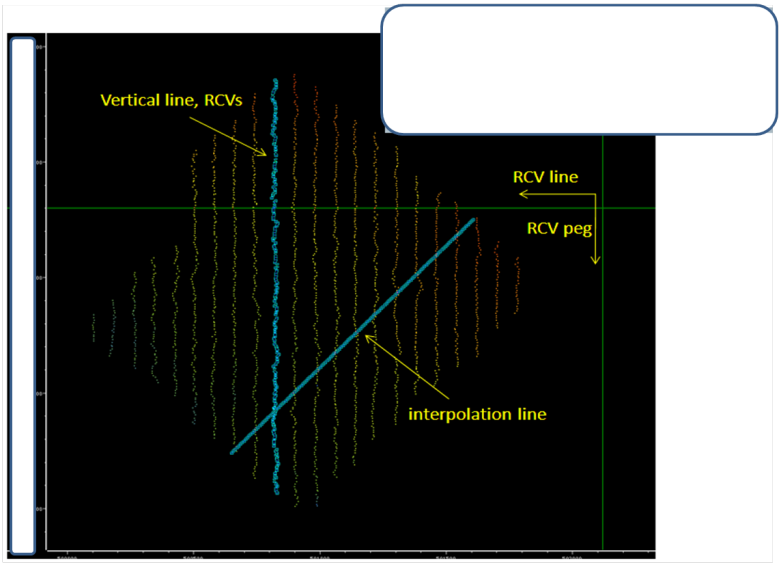


Figure 4.2: Receiver point map for M1 and M2 and location of the interpolation line.

since the geology was not changed between the two acquisitions, the same (M1) velocities for M1 and M2 are applied (statics and regional corrections), as it is the use for 4D data.

- An anti-aliasing filter, Butterworth 200 Hz, is applied.
- Low-cut filter, Butterworth 13 Hz, is used to attenuate the low frequency noise.
- NMO correction: for M2, which is at the floating datum, M1 velocities are applied, in order to process the two monitors in a comparable way. The velocity used is a 1D velocity without any lateral variation.
- Mute for 5D interpolation: for 5D interpolation, it is recommended not to integrate stretching associated with long offsets. Gathers are hence slightly muted before being used for interpolation. The mute law depends only on offset.
- Data selection for interpolation: only the traces whose distance from the interpolation line is less than 200 m are kept for the interpolation.
- 5D interpolation: this interpolation performs multi-dimensional Fourier data mapping. Dimensions are inline, crossline, azimuth, offset and time.

I should mention that these two processed datsets are the only datasets that were provided and I have to deal with them for FWI. It seems the datasets are processed towards the classical reflection processing. I have no access to the raw datasets.

At the end, M2 gathers have been moved back to their proper initial floating datum. The interpolation line prestack data are at the floating datum. CGG also provide us the interpolation line stacks (Figure 4.4). Stacks are at the final datum. For M2, the M1 velocities have been applied at the M1 floating datum, before moving back to the final datum. Because of that there is a time-shift (about 0.1 s) between stack and prestack data (compare Figure 4.4 and Figure 4.5).

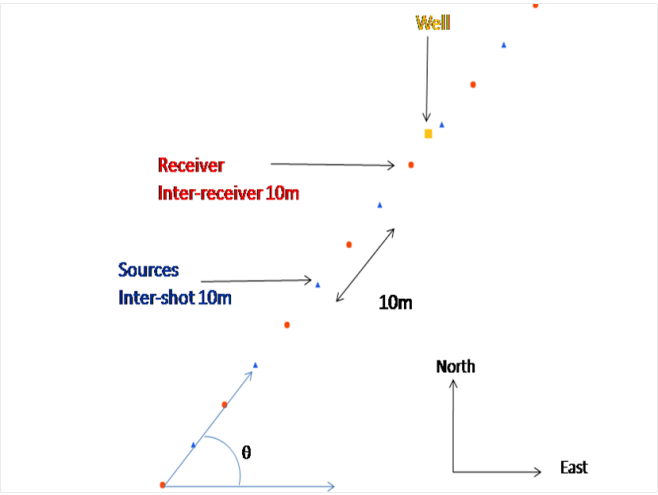


Figure 4.3: 2D interpolation schematic line.

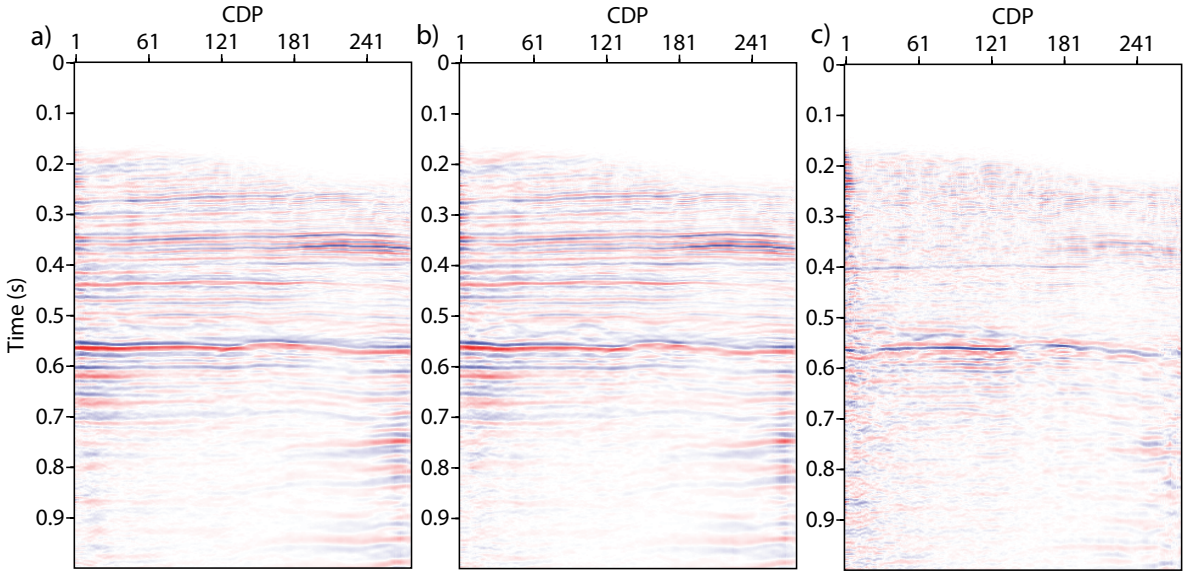


Figure 4.4: Interpolated 2D line stack (provided by CGG) for (a) M1, (b) M2 and (c) difference between the two. Please note that the difference section is multiplied by a factor of 2 to be plotted at the same scale as others, for visualization purpose.

After making a raw subtraction between M1 and M2 stack images, the main 4D signals can be seen around the main reflection event (Figure 4.4c). This is the interested area which is looked for. In addition, at time 0.4 s , another anomaly can be observed in the overburden zone which is probably affected by the production activities of the field.

The interpolated prestack data exhibit a frequency spectrum content between 13 Hz to 170 Hz. For performing 2D acoustic FWI, the data are filtered below 80 Hz. The filtered seismic data are shown in Figure 4.5 for the shot located at the center of the 2D line. We can clearly see the time-lapse seismic differences due to injection of steam (Figure 4.5c). It seems

that the main reflection event corresponds to the reflection from the Devonian layer (base of reservoir). After new injection, due to time-lapse variation, there are some changes in phase and amplitude of seismic response around this region. During processing step of early raw dataset, the surface waves were attenuated, but we can see that the difference data contain the imprint of surface waves. The surface waves may not be identical between two acquistions, because of seasonal changes and temperature variation of weathering zone.

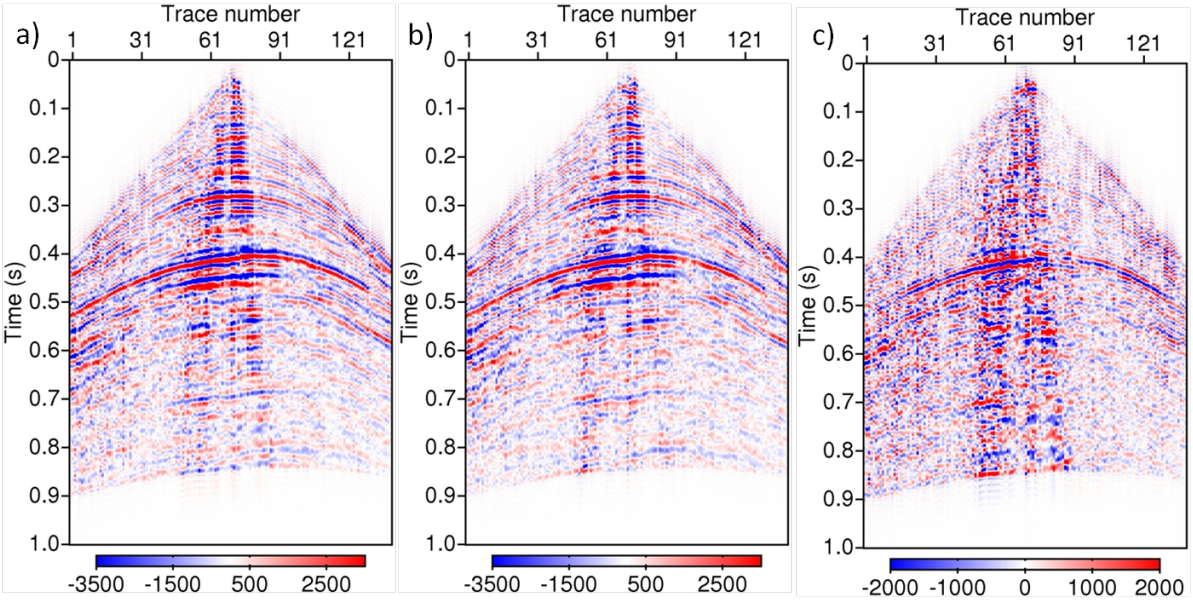


Figure 4.5: Seismic data for one shot located at the center of 2D line, (a) M1, (b) M2, and (c) the difference dataset.

4.3 Main problems

For these datasets, there are three main difficulties:

- Absence of direct waves and diving waves: as it can be seen in Figure 4.5, the direct waves and diving waves were removed from the dataset (probably at muting step before 5D interpolation).
- Problem of low frequency content: as mentioned, because of low frequency noise attenuation, a low-cut filter was used. The minimum frequency content in the datasets is around 13 Hz.
- Problem of source estimation: since there are no direct waves or first arrivals in the dataset, we have a problem to estimate the source wavelet. I must estimate source wavelet on a reflected event. This issue will be explained later.

It appears that the performed processing step on these datasets is more oriented towards the processing of classical reflection analysis. Therefore, the inversion of these datasets by FWI

is not ideal, but these were the only datasets available. In this configuration, it is expected that the FWI method behaves like a quantitative non-linear migration technique, because the low content of the wavenumber cannot be retrieved (Plessix and Mulder, 2004).

4.4 Inversion of M1 dataset

First of all, the M1 dataset must be inverted to obtain the recovered M1 model. In this study, regularized 2D acoustic FWI in time domain is used. During inversion, only P-wave velocity is inverted. A homogeneous density model with 2000 g/cm^3 is considered for the forward modeling. Before starting the inversion, we need to have an initial velocity model and to estimate a source wavelet.

4.4.1 Initial model

A velocity model (Figure 4.6a) has been provided by Total. This 2D velocity model was extracted from the interpolated velocity cube which has been built from interpolation between several well logs located in this field. This velocity model is built at the base time, that is, before any injection and it does not contain any effects of injected steam. Note that the origin $(x = 0, z = 0)$ corresponds to the position of the first shot. Is this velocity model accurate enough?

The shallow part of this interpolated model has a low velocity value. If the forward modeling is performed with this model, the direct waves contaminate the reflection events. Because of absence of the direct waves in the real dataset, we should modify the low velocity zone at the top in such a way that the direct arrivals do not appear in the calculated data. The velocity of low velocity zone at the top of model has been replaced by a value of 1850 m/s. This value comes from a velocity analysis (semblance method) of reflected events. Then, during inversion, the calculated data are muted in order to remove the direct waves, so that they are consistent with the observed data.

However, this modified velocity model, coming from the interpolation of well logs, is not kinematically accurate for all the arrivals (especially for the main reflection event, please see Figure 4.7). Figure 4.7 shows the observed data and the calculated data inside the modified model (Figure 4.6b). Clearly, it can be seen that most of the reflection events observed in the calculated and observed data are time-shifted, except the first reflection event. Therefore, this interpolated velocity model cannot be directly used as an initial model in inversion optimization, due to a large cycle-skipped issue. Probably, a smoothed version of this modified model (Figure 4.11a) could be a good candidate for the initial model.

4.4.2 Source-wavelet estimation

Estimation of source wavelet is a crucial step before inversion. In our case, where the real provided data do not contain direct arrivals and diving waves, it would be difficult to estimate the wavelet. The estimation must be done on the reflection events, a process which is limited and increases uncertainties. If the smoothed initial model is used for the source-wavelet estimation, it is not possible to have proper reflection events in calculated data which mimic the observed reflection events. Therefore, we have to use the modified interpolated velocity model without

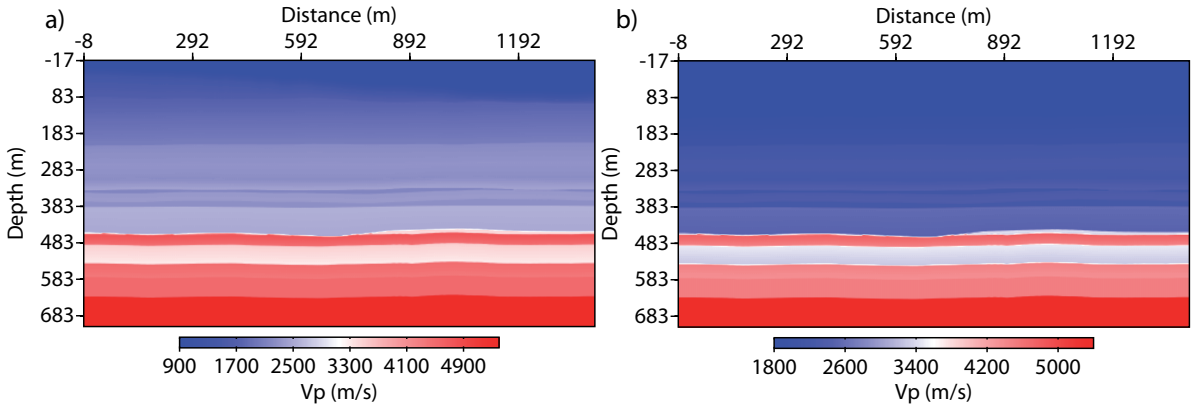


Figure 4.6: (a) 2D velocity model extracted from the interpolated velocity cube which was built by interpolation between well logs. (b) Modified velocity model at near surface. Low velocity at top of model has been replaced by a value of 1850 m/s. This value comes from a velocity analysis (semblance method) and is also provided by CGG.

any smoothing. The wavelet estimation on reflection seismic events is done considering the following points:

- In order to be less sensitive to the effects of reflectivity variations and converted waves on the amplitude and phase of source wavelet, the estimation is performed only on the first reflection event (shown by an arrow in Figure 4.7) and for the short offset traces (offset in this range ± 200 m). If the estimation is performed on the deeper reflection event, due to energy partition at previous reflectors, we have a high uncertainty on the estimated source. In the observed data, for far offset, partial energy of P-wave is converted to S-wave, a mechanism not taken into account in our acoustic modeling. Hence, for source estimation, only the short offsets are taken into account.
- It is assumed that the velocity contrast related to the first reflector, included into the interpolated model, mimics the real contrast, otherwise, the effect of differences in reflectivity coefficients (observed and calculated data) will be included in amplitude of source wavelet.
- I should also assume that the wave propagation does not create too much dispersion phenomena.

Our acoustic FWI is implemented in the time domain for both the forward and the adjoint problems. The source-wavelet estimation is however straightforwardly implemented in the frequency domain by a linear inverse problem resolution. The computed and observed time-domain data are Fourier transformed to apply the Pratt (1999) (his equation 17) source estimation equation for each frequency. The Fourier coefficients of the wavelet are then transformed back to the time domain and appropriately processed (anti-causal mute and band-pass filtering) before performing FWI. This estimation is performed once before the optimization.

In this case, the observed data and the calculated data are windowed to select the first reflection (using hyperbola formulation and a Hanning window). The estimation is performed

on short offset traces for all shots and then the average source wavelet is computed from them (Figure 4.8). The wavelet behavior is similar for all the shots (as we are informed), therefore we use a global averaging.

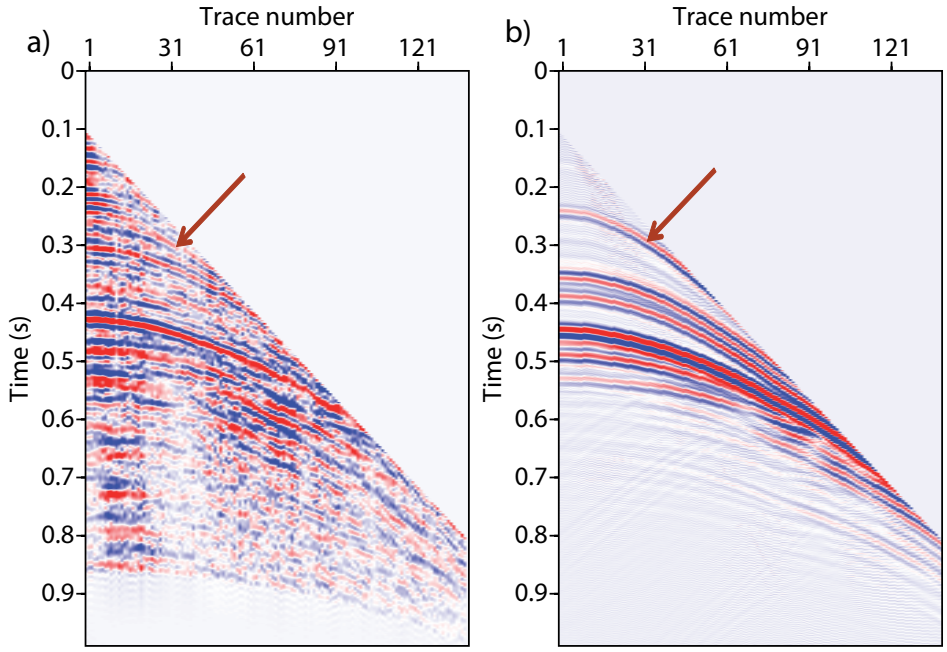


Figure 4.7: (a) The observed shot gather for the first shot, (b) the calculated data computed inside the modified interpolated velocity model using an estimated source wavelet. Please note that the direct waves are muted in both datasets. The arrow shows the first reflection event which is used to estimate the source wavelet.

As mentioned, the non-smoothed model cannot be used as the initial model for the inversion because of inaccurate kinematic issue. To overcome this problem, this model is smoothed by an intermediate smoothness factor (Figure 4.11a) and is considered as the initial model for M1 reconstruction. In order to decrease the uncertainty related to the estimated amplitude of wavelet inside this smooth model, we perform a simple amplitude scaling based on the misfit function at first iteration. The sensitivity analysis on the amplitude of wavelet is studied. The misfit functions (ℓ_2 norm of misfit between observed and calculated data) are computed at first iteration using different multiplicative factors (α) which are applied to the estimated wavelet. Figure 4.9 shows the plot of the misfit function curve with respect to the different α factors. The minimum of objective function is provided by $\alpha = 8$. Therefore, the estimated source wavelet which will be considered for FWI, is multiplied by this factor to have a proper amplitude.

4.4.3 Inversion without prior model

The provided M1 dataset (before it has been filtered below 80 Hz) is migrated (Kirchhoff depth migration) using the smooth initial model (migration procedure of Seismic Unix package is used (Cohen and Stockwell, 2008)). Migration technique puts the reflectors at their positions

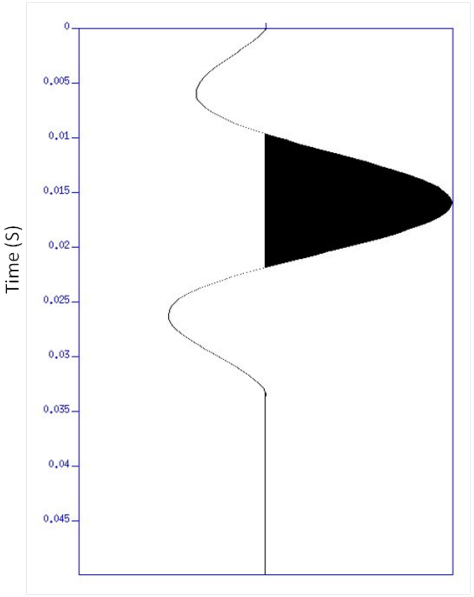


Figure 4.8: Average estimated source wavelet computed by windowing on the first reflection event and using near offset traces. The estimation is performed on all shots. The estimated wavelet is appropriately processed by an anti-causal mute and band-pass filtering.

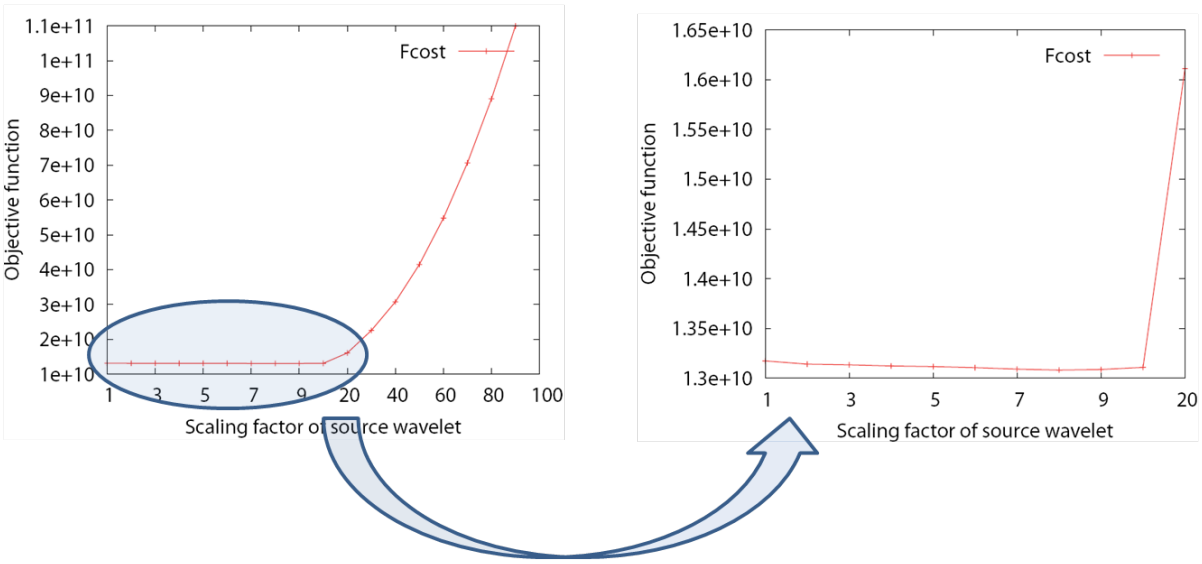


Figure 4.9: The objective function computed at first iteration using different amplitude multiplicative factors (α) of estimated source. From this plot, the minimum objective function is provided by $\alpha = 8$.

(Figure 4.10). As the dataset does not contain the diving waves and low frequency contents, we expect that FWI provides a similar result as the migrated section but no more details. However, the advantage of FWI is that it delivers a quantitative image as compared to the migration process, which may be of great interest to monitor subtle velocity changes. In this

configuration, FWI will behave like as a quantitative iterative non-linear migration technique (Plessix and Mulder, 2004).

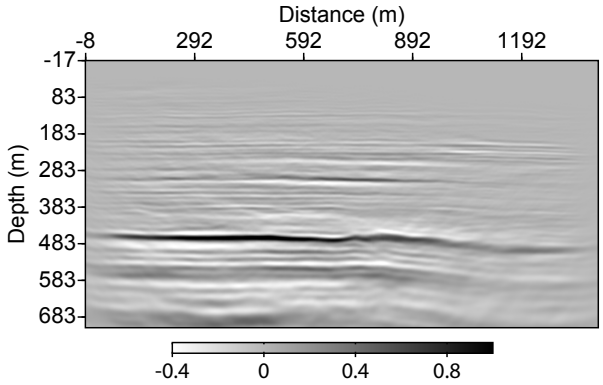


Figure 4.10: Migrated section of M1 dataset using the smoothed initial velocity model (Figure 4.11a).

In the first step, the M1 dataset is inverted by FWI without including any prior model information into the inversion scheme. For that, I consider a small value for Tikhonov regularization to ensure a small smoothness on the result. There is no large variation for the topography, we have almost 70 m differences in elevation between position of first shot to the last one. However, Due to non-flatness of the surface, the free-surface conditions cannot be considered in our finite-difference implementation (our code can consider the free-surface conditions only for a flat topography). However, in this land dataset, the surface waves and the multiples were almost removed during processing step.

The 2D acoustic FWI in time domain is used to invert the M1 dataset starting from the smoothed initial model. The inversion parameters, such as grid distance ($h = 3\ m$) and time-sample interval ($dt = 0.33\ ms$), are chosen in such that the inversion can be done until 80 Hz. Here, no multi-scale approach on frequency content such as the frequency-continuation approach of Bunks et al. (1995) is used. L-BFGS-B routine is used for the optimization to get model parameter perturbations (Byrd et al., 1995). An Identical stopping criterion based on flatness of misfit function is used to stop the optimization for all the following inversions.

For M1, the inversion stopped after 35 iterations which leads to have 18% reduction of data-misfit function at the end with respect to the initial data-misfit function. The recovered M1 velocity model is shown in Figure 4.11b. Clearly, it can be seen that the inversion procedure reconstructs the reflectors inside the initial model. The obtained result of FWI is consistent with the stack migrated section, but slight depth shifts between them are observed (compare Figures 4.11b and 4.10). As expected, the FWI provides similar result as the migration technique, but a quantitative image. Clearly, it appears that the background velocity used in the initial model does not essentially change during inversion, due to lack of low frequency contents and transmission waves. The low velocity zone, at depth of around 430 m, is related to the main reflection event in dataset. We cannot clearly interpret this low velocity zone, but it probably corresponds to the production activities of the reservoir. Moreover, a slide-lobe effect around this main reflector can also intensify velocity variation in this zone, due to the band-limited frequency and limited aperture inversion. The base of this low velocity zone (the red refletor) might be the base of the reservoir, corresponding to the lithological basement (Devonian layer).

Probably, the W shape which can be seen in the obtained model, could be related to the sensitivity kernel for a reflector line. Indeed, for this dataset, there are less transmission events and the inversion is dominated by the reflection data. Therefore, there is less contribution of a transmission sensitivity kernel to recover the low content wavenumbers.

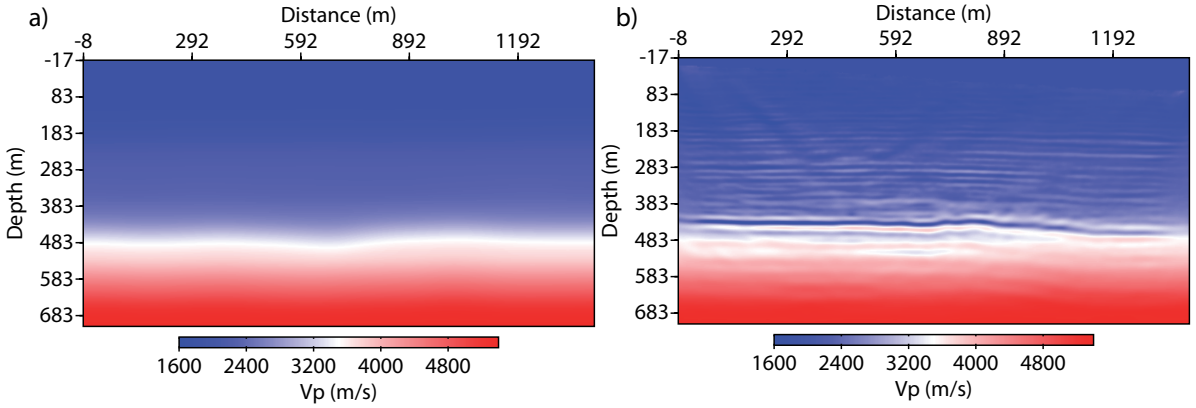


Figure 4.11: (a) The smoothed initial model which comes from smoothing the modified interpolated velocity model, (b) the recovered M1 velocity model obtained by FWI without using any prior model.

As a quality control (QC), the observed data and the final calculated data are plotted in Figure 4.12 to compare them. In Figure 4.12a, the observed and final calculated data are normalized by maximum value of the observed data for each trace. This plot can be used as an indicator to compare how the relative amplitude between the recorded and calculated data is fitted. In Figure 4.12b, both datasets are normalized independently by their own maximum value. This could be used as another indicator to compare how the phases of signals are fitted after inversion. Please note that the inversion is performed without any normalization of data. For the trace number less than 50, the predicted phases have a phase shift with respect to the recorded data, especially for the main reflection. In addition, the amplitude of data is poorly explained, certainly due to not considering a proper density model. When we observe a scan of all the traces (a video of these plots for all shots), for some shots the phases are not accurately predicted with this final model.

These observations leads to ask for the search of a better reconstruction of M1 model, which is a crucial step for monitor reconstruction, as shown in Chapter 1. As suggested in Chapter 1, the prior model can be accounted for the inversion to better and more accurately recover the model parameters.

4.4.4 Inversion with prior model

The modified interpolated velocity model can be a good candidate to be used as the prior model (Figure 4.13a), because it has been built from well logs located in this field. From the way of designing the 3D cube of interpolated velocity model, there is a vertical log located at this 2D line at distance of $x = 1040\ m$ and another well near to this 2D line around $x = 250\ m$. If we assume there would not be any other wells to contribute for building this 2D velocity section,

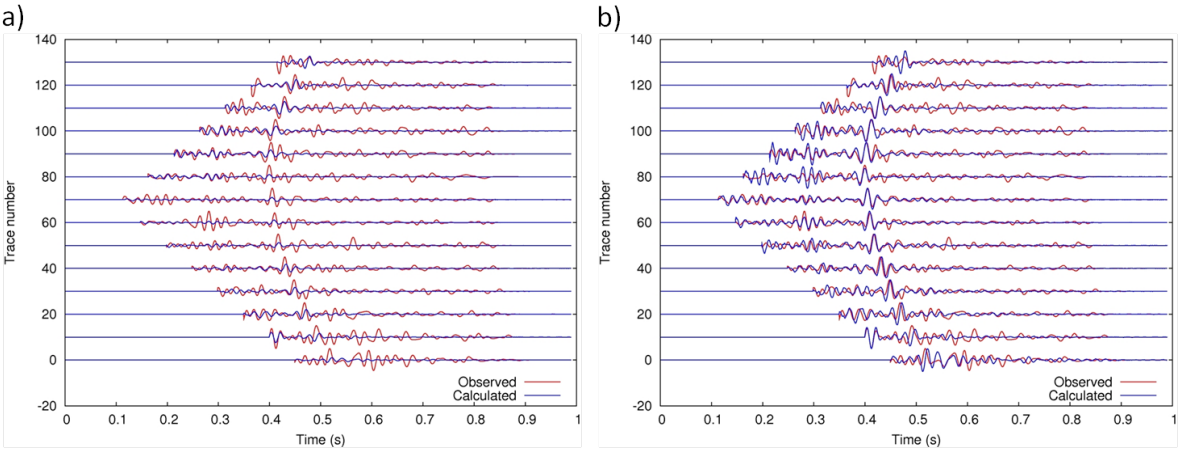


Figure 4.12: Comparison of the final calculated and observed data for the centered shot for inversion of M1 dataset without using any prior model: (a) both data are normalized by maximum value of observed data (trace by trace), indicator to compare the relative amplitude; (b) each data are normalized by their own maximum value (trace by trace), independently. This is an indicator to compare the phase of signals. Please note the normalization is done only for visualization purpose.

there is a lesser uncertainty near to these wells and higher uncertainty between these two well locations. Therefore, the prior uncertainty model σ^2 is designed in the same way as suggested in Asnaashari et al. (2013a). The second type of uncertainty is related to the surface acquisition and geometrical spreading of data. After combining these two kinds of uncertainties, we can end up with the prior uncertainty model shown in Figure 4.13b.

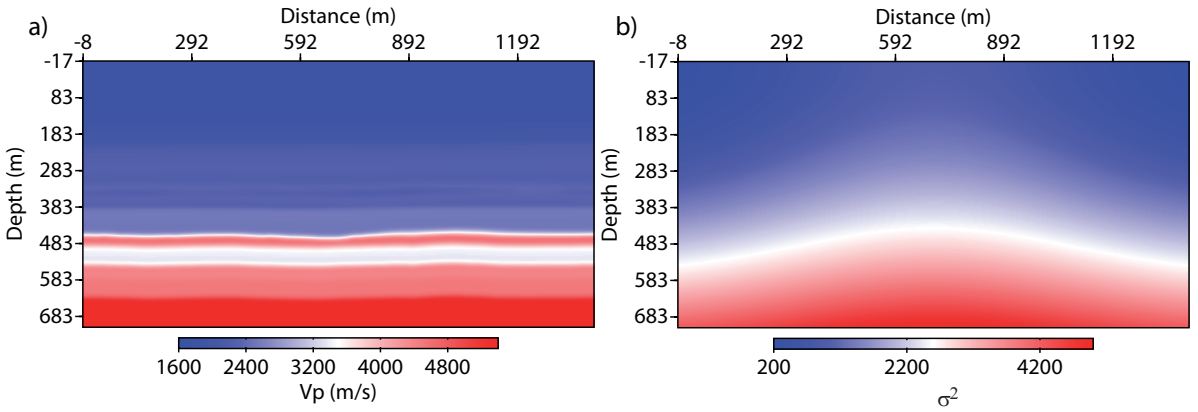


Figure 4.13: (a) The modified interpolated velocity model, slightly smoothed, is used as the prior model. (b) Prior uncertainty model (σ^2), the Gaussian function varying in the x direction with maximum value at the center of model, which is complemented by a quadratic evolution in the z direction.

The regularized FWI including a prior model suggested by Asnaashari et al. (2013a) has been used to invert the M1 dataset. The prior regularization parameter λ_2 is chosen such that to have a ratio between the prior-model misfit term and the data-misfit term around $\gamma = 5 \times 10^{-3}$, at first iteration. Therefore, by selecting this reasonable ratio between prior-model and data-misfit terms, the FWI is prevented to minimize the model norm heavily at early iterations. Then, during the optimization the dynamic weighting method is used to gradually reduce the effect of the prior model.

By including the prior model and using the dynamic weighting approach, the inversion goes through more iterations (89 iterations) than the classical FWI in previous inversion. In addition, this new inversion leads to have 22% reduction of data-misfit function at the end with respect to the initial data-misfit function. Comparing to the classical FWI (without prior model), the data-misfit term is more reduced. The recovered M1 velocity model is shown in Figure 4.14. Comparing Figures 4.14 and 4.11b illustrates that the velocity below the main reflector is better recovered thanks to the prior model. The effect of the prior model can be seen in the final obtained model (below 450 m depth). A QC vertical log at $x = 1040$ m confirm this issue (Figure 4.15). The lower velocity zone below the Devonian reflector (which can be seen as a white layer in the prior model) affects the recovering of M1 model.

In addition, comparing the observed data and final calculated data shows that by adding the prior model into inversion, the phases are better estimated and the amplitudes are better fitted (compare Figures 4.16 and 4.12). For this specific shot at the center of acquisition line, we can clearly see that for the trace number less than 50, especially, the phases of signals are better predicted by using the prior model information.

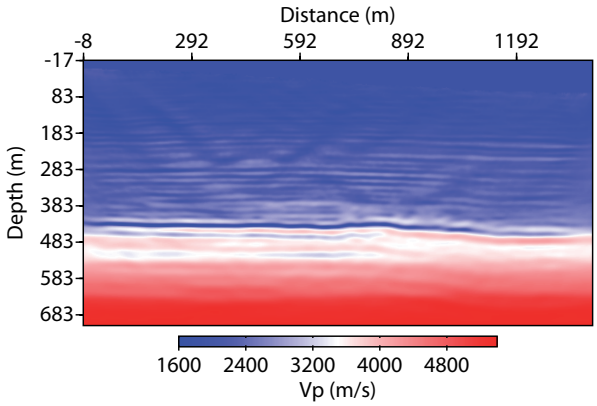


Figure 4.14: The recovered M1 velocity model obtained by regularized FWI using the prior model (Figure 4.13) and a dynamic weighting approach to gradually decrease the importance of the prior model during inversion.

For these datasets where there is a lack of low frequency content and of transmission waves, the prior model appears crucial to partially retrieve the low spatial wavenumbers. If we have a more precise prior model (after injection), probably we can improve the result of M1. The prior model, which is already used, does not contain any effect of injected steam because it is built from well logs acquired before steam injection. Due to this issue, we cannot put stronger weight into this prior model for the optimization.

I should mention that, since the inversion updates only the velocity model without consid-

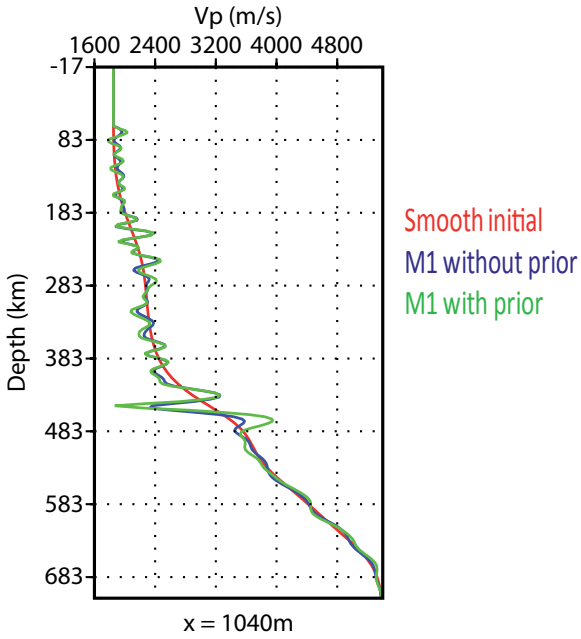


Figure 4.15: QC vertical logs at $x = 1040\ m$ coming from the models: the smooth initial model, the recovered M1 model without prior model, and the recovered M1 model using the prior model and dynamic weighting approach. The effect of the prior model can be easily seen on the recovering of velocity model, especially for the main reflector and below it.

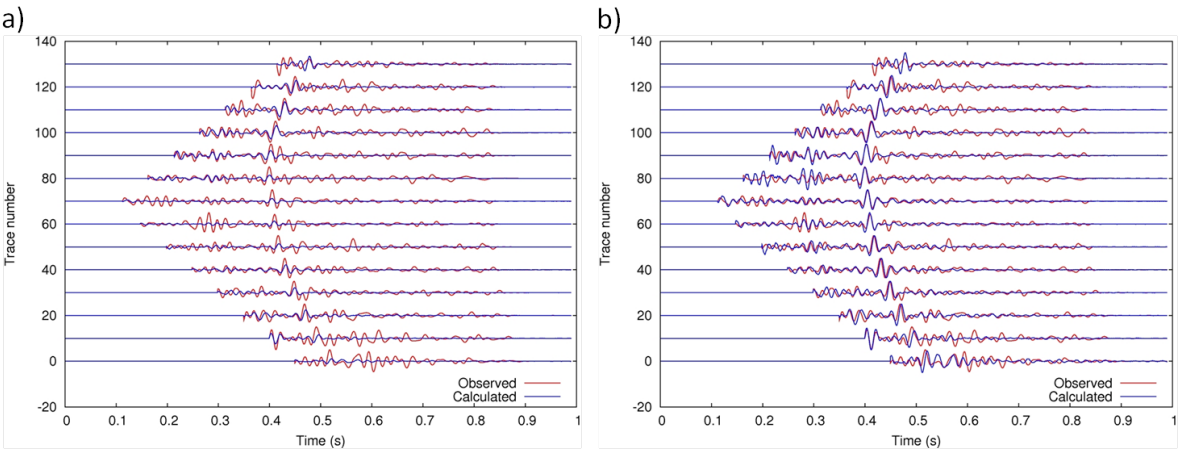


Figure 4.16: Comparison of the final calculated and observed data for the centered shot for inversion of M1 dataset using the prior model: (a) both data are normalized by maximum value of observed data (trace by trace), indicator to compare the relative amplitude; (b) each data are normalized by their own maximum value (trace by trace), independently. This is an indicator to compare the phase of signals. Please note the normalization is done only for visualization purpose.

ering updating the density model, the effect of impedance is entirely carried by the velocity. In other words, the variation of density is transferred explicitly to the update of velocity; therefore the obtained velocity model contains the effect of impedance.

As another quality control, Common Image Gathers (CIG) are computed by Kirchhoff migration of the original M1 dataset (before filtering for FWI purpose). The CIGs are computed using the smooth initial model and the final FWI model derived with prior model. This final FWI model has been slightly smoothed before using for computing the arrival time related to the computed rays. An example of some CIGs are shown in Figure 4.17. Comparison of CIGs in two models (Figures 4.17a and 4.17b) shows that the reflector events become more flat after FWI inversion. I should mention, however, that in some positions and some other CIGs, the smooth initial velocity model can provide better and flatter CIGs as compared to high resolution FWI result. Generally, the Kirchhoff migration technique better works using a smooth model. Since the high resolution FWI result already contains the velocity contrasts (the reflectors), Kirchhoff migration does not work properly to compute the associated ray arrival times, especially near the interested zone in the FWI model.

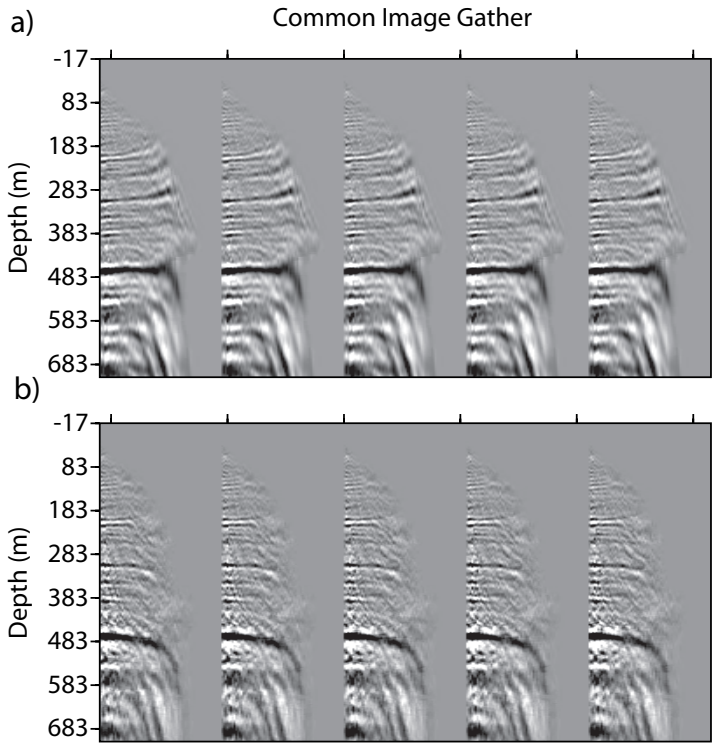


Figure 4.17: An example of some Common Image Gathers (CIG) of M1 dataset which is migrated using: (a) the smooth initial model, (b) the final FWI model with prior model. It seems that the CIG gathers become more flat after FWI, however for the main reflector it is not the case.

4.5 Time-lapse inversion

Once the M1 velocity model is obtained, we focus on the recovery of the time-lapse variations. Here, the three different time-lapse strategies which are discussed in Asnaashari et al. (2013b) have been used. The sequential difference and the differential (double-difference) strategies are first applied in the conventional mode (without target-oriented option), and then using the target-oriented prior weighting. The recovered M1 model, obtained by regularized inversion including the prior model (Figure 4.14), is used as the initial model for the sequential difference and the differential inversion.

As a first step, the difference dataset between two surveys data is migrated using the slightly smoothed M1 model (Figure 4.18). By doing so, we can have some ideas about the location of expected time-lapse variations. In the right part, around the reservoir depth, three elliptical zones (black areas) can be identified, which are probably related to the steam chambers after reinjection. In the left part, it appears that we expect to have some time-lapse variations. In the overburden zone, we can see a continuous reflector which is coherent with our observation from the difference stacked data (compare Figures 4.18 and 4.4c).

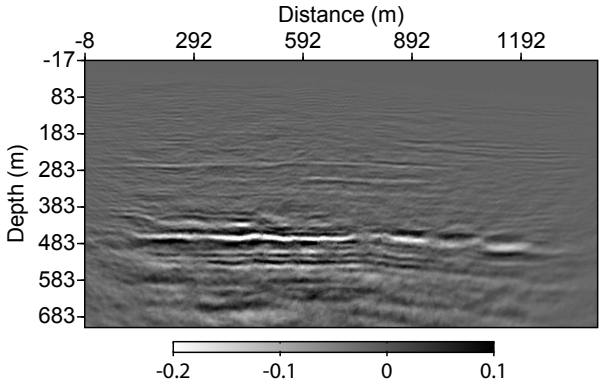


Figure 4.18: Migrated section of difference dataset using the recovered M1 velocity model. This velocity model is slightly smoothed before performing migration. This section can show the location of the expected 4D changes.

4.5.1 Conventional inversion

Parallel difference: for the parallel difference strategy, we invert the M2 dataset with the same configuration and inversion parameters as for the M1 reconstruction, starting from the same smooth initial model, using the same prior model, and the same parameters of the dynamic approach. Figure 4.19a shows the time-lapse model obtained by the parallel difference strategy after 85 iterations for the second inversion. At reservoir depth, three areas of negative velocity variations are recovered (shown by three arrows in Figure 4.19a). These three zones are affected by injection of steam. The P-wave velocity appears to decrease, certainly due to reinjection of steam in the period between the two acquisitions. The presence of steam and increase of temperature lead to have a velocity decrease and an expansion of steam chambers. These patches of injected steam are validated by the configuration of injected equipments and wells. The result of parallel difference for these steam chambers are consistent with the migrated section of difference data.

Underneath these steam zones, there are two continuous lines with positive and negative velocity variations. These are the 4D artifacts related to the parallel difference strategy, in some places of these continuous lines. Even in the migrated section, we cannot observe this strong continuity everywhere along the reservoir. These artifacts are coming from the reconstruction of main reflector, through two separate inversions. The inversion of M1 and M2 datasets, starting from the same smooth initial model, cannot probably converge to reconstruct the main reflector at the same depth. Because of that, we have strong continuous variations, due to a simple subtraction between two final images. In addition, this problem could also be caused by side-lobe issue because of band-limited inversions. In distance $x = 300 - 600\ m$ and at depth of around the main reflector, a velocity variation is observed which can be seen in the migrated section as well. It seems this kind of variation exists in the dataset and the parallel difference FWI strategy tries to recover these variations, but we cannot yet interpret them. The reservoir specialists of this project do not expect any positive velocity variations in this area.

Sequential difference: the recovered M1 model, obtained by including the prior model penalty, is used as an initial monitor and a prior monitor model to invert the real M2 dataset. In a first investigation, the sequential difference strategy is applied in conventional mode, *i.e.* without any target-oriented prior weighting. A homogeneous prior weighting model is used with a low weight of the prior model into the inversion. Figure 4.19b shows the time-lapse result after 19 iterations. It is difficult to distinguish the time-lapse variations area, without knowing the result of the parallel difference strategy. The obtained time-lapse variation model is dominated by a strong negative velocity variation line (blue line). This sequential difference strategy attempts to continue updating the M1 model during the M2 inversion. The second inversion tries to fit the data residuals left from previous inversion, especially for the main reflector event. However, the real time-lapse variations (patches of injected steam) are recovered by this strategy, but the 4D artifact is dominating the time-lapse result. I should mention that this issue (keep updating the common events between two surveys) related to the sequential difference strategy is not due to convergence in a wrong direction. Since the qualification of time-lapse variations is based on a simple subtraction between two final models, this continued updating of the model is considered as a 4D artifact.

Differential: in this strategy, the composite dataset is inverted which is focusing on inversion of the difference dataset. Similar to the sequential difference configuration, the obtained M1 model is used as the initial and prior monitor model for this inversion. In a first step, the differential strategy is applied in conventional mode. The obtained time-lapse result after 61 iterations is shown in Figure 4.19c. In this case, the time-lapse dataset is better focused and better focalized at their position by the double-difference inversion. Comparing to the result obtained by the parallel difference, we can observe that the three steam anomalies (shown with three arrows) are better identified and better focused by the differential strategy (see Figures 4.19a and 4.19c). In addition, the 4D artifacts, continuous lines underneath of 4D responses in parallel difference result, do not exist in the differential result. In the differential case, below the three steam chambers, velocity increases. Perhaps these velocity increases might be due to migration of steam and/or probably fluid substitution between two acquisition times, it would be difficult to interpret these increasing velocity anomalies.

It seems that this result is more coherent with the migrated section of the difference data, particularly at the right part of model where the steam injection zones are identified. However, the strong continous 4D anomaly, which is present in the migrated section (distance $x =$

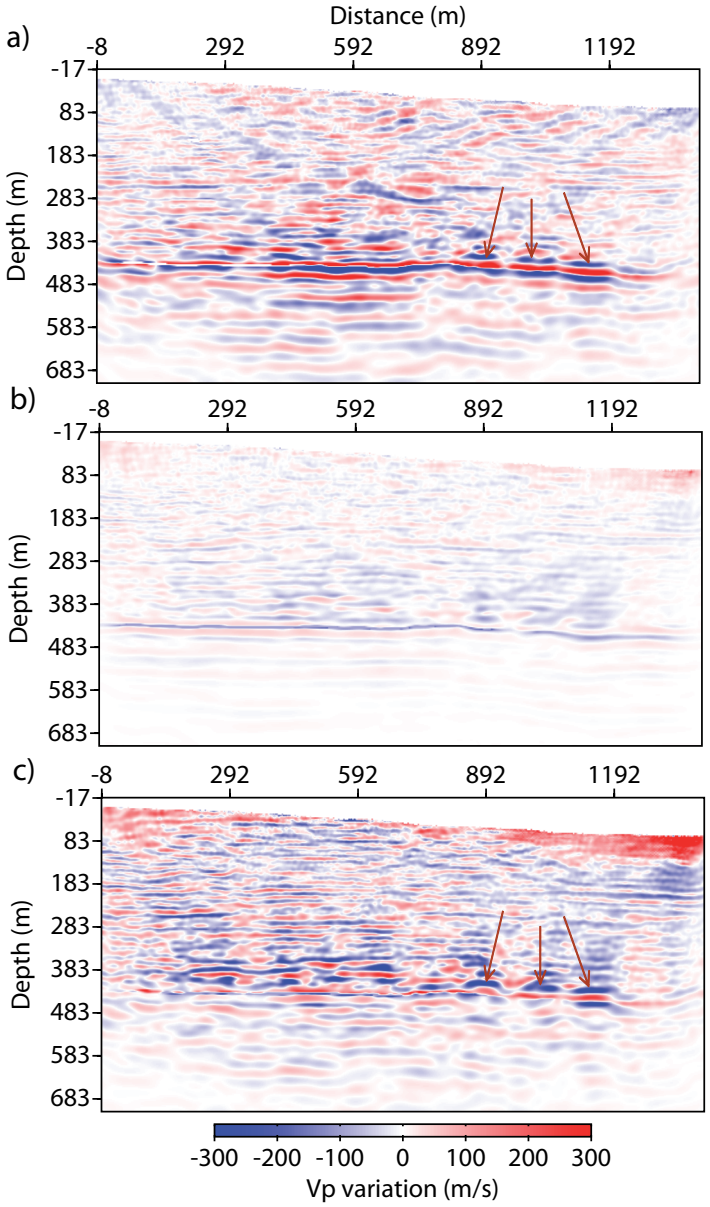


Figure 4.19: Time-lapse variation velocity models obtained by different strategies: (a) the parallel difference, (b) the conventional sequential difference without using the target-oriented option, (c) the conventional differential (double-difference) strategy. The arrows show the location of negative velocity variations caused by reinjection of steam.

$300-600\ m$), is not recovered by the differential strategy. I should mention that this continous 4D anomaly is not expected by the specialists. It is possible to have some steam injection activities above this area (as it can be underlined by some negative velocity anomalies in the differential result), but a strong continous line is not expected to have the positive velocity variation (as we observe in the migrated section for this area).

Comparison of the obtained time-lapse results (obtained by the parallel difference and the

conventional differential strategy) and the migrated section of difference dataset shows that probably time-lapse variations, below the reservoir zone presence in the migrated image (below 550 m depth), are migration artifacts due to velocity variations occuring in the reservoir areas (push down-pull up effect). Due to changes in the reservoir area because of steam injection and oil production, the image of underneath areas are affected. This kind of artifacts are not observed in the FWI time-lapse results, because the M2 velocity model is iteratively improved during inversion and the time-lapse anomalies are recovered. Therefore, the time-lapse signals are focalized at their correct depth positions.

Above each velocity decreasing zone (due to steam reinjection), in both sequential and differential results, a vertical column of negative velocity variations can be observed. There is a possibility to interpret them as a leakage of steam above the injected points. After discussion with reservoir specialists of this project, there is no evidence of the leakage of steam in this area. Therefore, these anomalies could be interpreted as inversion artifacts. In the overburden zone, there are several low wavenumber artifacts. These artifacts are probably produced by two causes: 1) the difference dataset contains imprint of surface waves which are not identical between two surveys after processing, so they are not completely cancelled out by data subtraction (clearly, it can be seen in Figure 4.5c); 2) Since the data were acquired at different seasons, we have differences on seasonal noise. This issue can also produce some 4D artifacts.

4.5.2 Target-oriented inversion

Is there any way to improve the real time-lapse responses and reduce the artifacts? In a second investigation, the two time-lapse strategies (sequential and differential) are applied using a target-oriented inversion mode. The recovered M1 model is used as the prior model. For the target-oriented inversion, we need to design the model weighting matrix $\mathbf{W}_m$ based on the location of expected changes. Based on the previous inversion results, the migrated section of difference dataset, and information about the depth of reservoir, it is possible to build the target-oriented model weighting. Two Gaussian functions varying in z direction are laterally used all along the reservoir and in the overburden zone (where the 4D anomaly can be observed in the migrated section). These Gaussian functions are summed up with a constant small value at other places. The target-oriented $\sigma^2(\mathbf{m})$ model (Figure 4.20) applies strong weights of the prior model (M1 model) everywhere outside these target zones and relatively lower constraints inside these expected variation areas (recall $diag(\mathbf{W}_m^T\mathbf{W}_m) = 1/\sigma^2(\mathbf{m})$). The first target zone in the overburden area is considered to recover the small 4D variation in this area which is observed in the migrated section (Figure 4.18) and the difference stacked data (Figure 4.4c).

The results of time-lapse model obtained using this method, for the sequential and differential inversions, are shown in Figure 4.21. The artifacts in the overburden zone are reduced for both strategies. In sequential difference case, the target-oriented inversion better focuses on the area of expected variations and provides more focused results. At least compared to the conventional case, the areas affected by injected steam can be identified. In target-oriented differential, the 4D anomalies are better recovered; it is possible to detect another potential steam chamber at the left part of model. The expected 4D variation in the overburden zone seems to be also recovered (depth around 280 m) and it is easier to identify this variation area by removing other artifacts in the overburden zone. By using a target-oriented prior weighting, the model space is reduced for inversion and hence, the time-lapse inversion becomes better-posed. Although, the value of velocity variation related to the artifacts (discussed before as

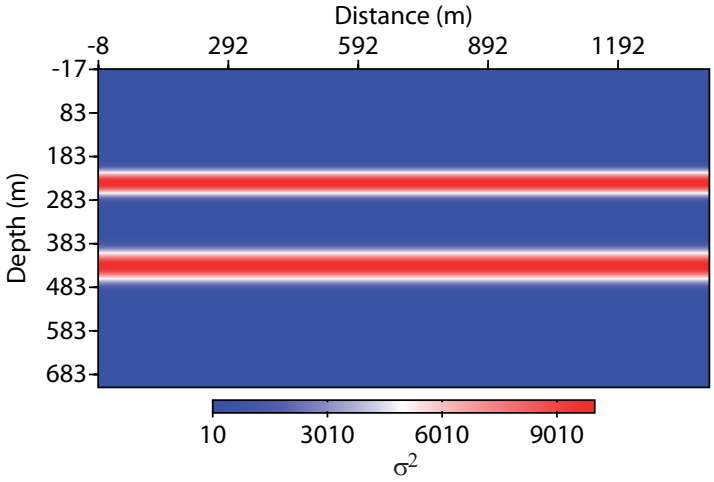


Figure 4.20: The target-oriented σ^2 model which is built by two Gaussian functions varying in z direction all along the reservoir and in the overburden zone (where the 4D anomaly can be observed in the migrated section). These Gaussian functions are summed up with a constant small value at other places.

a potential steam leakage) decreases more. They are more emphasized in this target-oriented differential inversion, which are probably incorrect (discussion with reservoir specialists).

Figure 4.22 shows QC vertical logs at $x = 1040\ m$ coming from the time-lapse models obtained by: the parallel difference, the target-oriented sequential difference, and the target-oriented differential strategy. The marker shows the negative anomaly due to the reinjection of steam. Using target-oriented inversion, the 4D artifacts in the overburden, particularly near surface zone, are removed. The value of negative velocity variations, obtained by the parallel difference and the differential, are almost coherent. Below the marker (steam area), the differential strategy provides a better result as compared to the strong fake negative anomaly in the parallel difference result (at depth around 460 m in the log).

As another QC to verify how the time-lapse data are fitted by different strategies, the final calculated and the observed difference data are plotted together. The final calculated difference data are computed by a raw subtraction between final calculated M2 (obtained from three strategies) and final calculated M1 datasets. Figures 4.23, 4.24 and 4.25, respectively, show this comparison between the observed and calculated difference data, for the parallel difference, the target-oriented sequential difference and the target-oriented differential strategies. First, in terms of prediction of phases, it seems that both target-oriented strategies better predict the phase of the difference data as compared to the parallel difference, especially, for the wavelets near the main reflection and for the traces with number less than 60. In addition, the target-oriented strategies do not allow inversion to fit the noise, which can be seen at earlier time (please compare trace number 70 in Figures 4.23b, 4.24b and 4.25b). This type of noise is related to the surface waves and seasonal temperature variation. Generally, the target-oriented inversion appears to be more robust than the parallel difference strategy. However, it is possible to have some traces which are out of phase for the target-oriented differential strategy, for example trace number 80. It seems the receiver of trace number 0 did not work properly during acquisition, since this observed trace is not coherent with other traces.

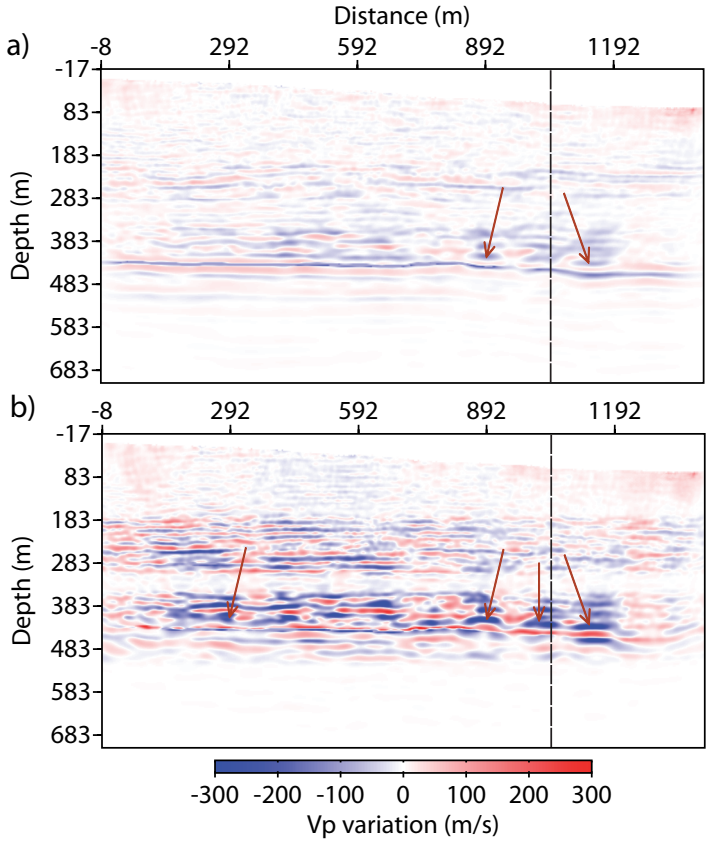


Figure 4.21: Time-lapse variation velocity models obtained by the target-oriented inversion: (a) the sequential difference, (b) the differential (double-difference) strategy. The arrows show the location of negative velocity variations caused by reinjection of steam. The vertical dashed line shows the location of vertical well at $x = 1040\ m$ to illustrate QC vertical logs.

A comparison of the relative amplitude between the calculated and observed difference data shows that the target-oriented differential inversion can better fit the amplitude (see Figures 4.23a, 4.24a and 4.25a). The reason is that the differential strategy directly inverts the difference dataset, without any contribution from residuals left from previous inversion (recovering M1).

As the last data comparison, the final calculated difference data in case of the conventional and the target-oriented differential strategies are plotted in Figure 4.26. It shows how the target-oriented inversion behaves differently from the conventional mode in fitting the dataset. For the earlier arrivals, the calculated data appears different which can slightly affect the simulation of late arrivals as well. If we assume that there is no change in overburden for this field (except in the first considered target zone), the difference data existing at earlier time will be considered as noise. By focusing inversion on the target zone, there is a possibility to avoid fitting this 4D noise.

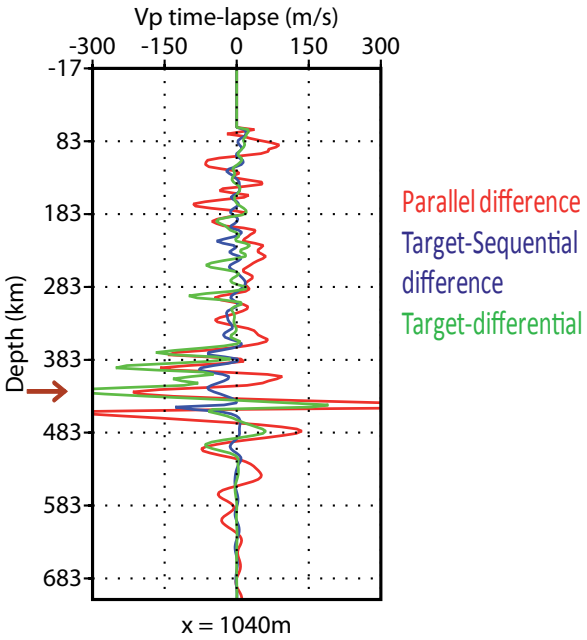


Figure 4.22: QC vertical logs at $x = 1040\ m$ coming from the time-lapse models obtained by: the parallel difference, the target-oriented sequential difference, and the target-oriented differential strategy. The marker shows the negative anomaly due to the reinjection of steam.

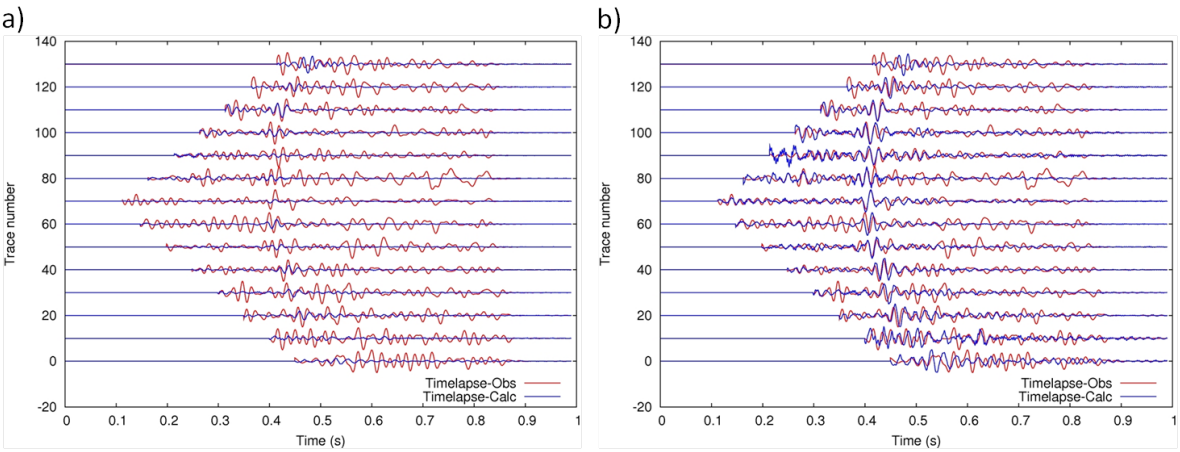


Figure 4.23: Comparison of the final calculated and observed difference data for the centered shot for the **parallel difference inversion**: (a) both data are normalized by maximum value of observed difference data (trace by trace), indicator to compare the relative amplitude; (b) each data are normalized by their own maximum value (trace by trace), independently. This is an indicator to compare the phase of signals. Please note the normalization is done only for visualization purpose.

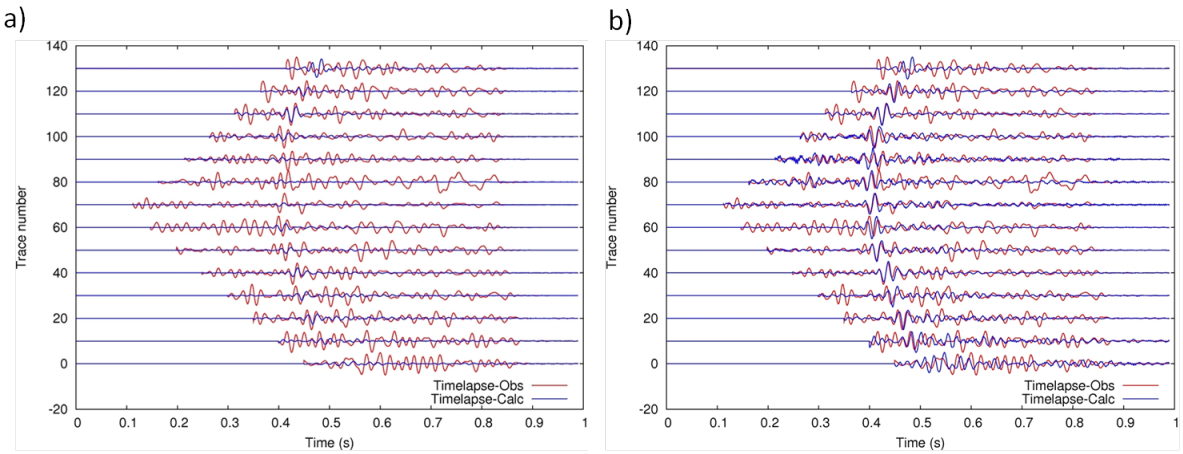


Figure 4.24: Comparison of the final calculated and observed difference data for the centered shot for the **target-oriented sequential difference inversion**: (a) both data are normalized by maximum value of observed difference data (trace by trace), indicator to compare the relative amplitude; (b) each data are normalized by their own maximum value (trace by trace), independently. This is an indicator to compare the phase of signals. Please note the normalization is done only for visualization purpose.

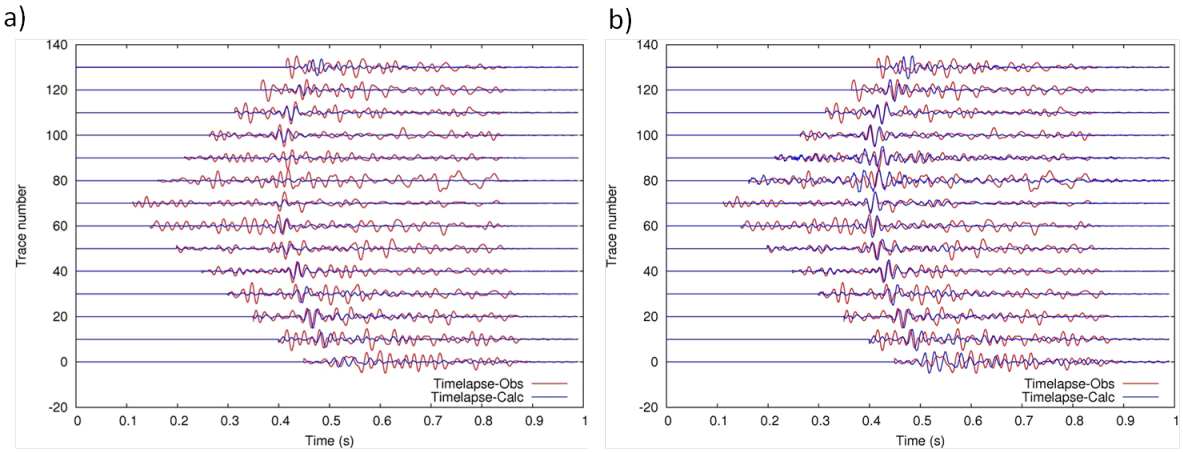


Figure 4.25: Comparison of the final calculated and observed difference data for the centered shot for the **target-oriented differential inversion**: (a) both data are normalized by maximum value of observed difference data (trace by trace), indicator to compare the relative amplitude; (b) each data are normalized by their own maximum value (trace by trace), independently. This is an indicator to compare the phase of signals. Please note the normalization is done only for visualization purpose.

4.6 Conclusions

In this real field steam injection datasets, we have applied several time-lapse strategies which were proposed in previous chapters. The provided filtered datasets have not been processed to

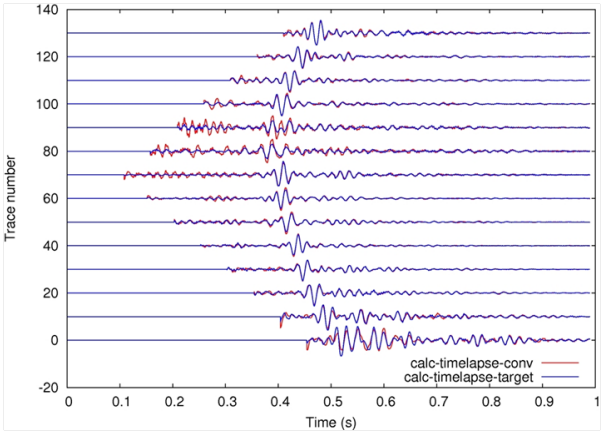


Figure 4.26: Comparison of the final calculated difference data between the conventional and the target-oriented differential strategy: both data are normalized by maximum value of calculated difference data coming from the conventional differential strategy (trace by trace).

fit FWI requirements. During early processing steps, the low frequency content (until 13 Hz) has been filtered out. Moreover, the real dataset does not contain the direct and transmission waves (due to muting and to an intermediate offset acquisition), as the data were initially acquired with a reflection survey design.

For the inversion, due to the muting of the direct waves, we must estimate the source wavelet on reflection events. This step is performed under some assumptions and estimation on first reflection event for short offsets, which probably leads to an estimated source-wavelet that contains large uncertainty. We assumed that the velocity contrast related to the first reflector, included into the interpolated model, mimics the real contrast and there is no dispersion effect.

Due to the lack of low frequency content in the data and lack of diving and transmission waves, the FWI method can be considered as a quantitative non-linear migration technique, as the lower part of the wavenumbers cannot be retrieved from the data part. In order to partially balance this issue, it has been shown that it is necessary to add prior model information into the inversion. By including the interpolated velocity model deduced from well information as the prior model and using the dynamic weighting approach, the regularized FWI improves the reconstruction of M1 model and the data-misfit function is further decreased compared to the case when no prior model was used. It is clearly shown here that a precise prior model is required to increase accuracy of inverted velocity models. In a second step, we focused on the monitoring approach, and applied our different strategies to two datasets between steam injection was done.

The parallel difference strategy reconstructs the time-lapse variation zones caused by reinjection of steam. However, the 4D anomalies are less focused and the time-lapse model contains several artifacts coming from reconstruction of the main reflector, through two separate inversions. We obtain strong continuous variations, when looking the simple subtraction between two final images. In addition, this problem could also be caused by side-lobe issue because of band-limited inversions.

The recovering of residuals left from M1 inversion is mainly an issue when the conventional sequential difference inversion was performed. Therefore, the time-lapse model obtained by

this inversion strategy is mainly affected by this issue and it is difficult to identify the real 4D variation zones due to steam injection. The second inversion (in sequential difference) attempts to better explain the amplitude of the main reflection event which is not fully explained during inversion of M1 data.

The conventional differential strategy provides a better time-lapse result around the reservoir depth, with less artifacts below the three steam chambers as compared to the parallel difference. The result obtained by the differential strategy suffers from several artifacts in the overburden. One of the reasons could be the inaccuracy of the recovered M1 model which does not allow the difference data energy to be properly focalized at the correct positions. Another reason could be related to the presence of the imprint of surface waves in the difference dataset and related to differences in the noise content between the two datasets.

In the last investigation, the target-oriented inversion using a target prior weighting all along the reservoir is applied. The target inversions, for both sequential and differential strategies, deliver more robust time-lapse results by focusing only on the target zones. Especially, for the target-oriented sequential difference case, at least some of the steam areas can be identified, as compared to the conventional case. Including the target-oriented prior model does not allow inversion to fit the 4D noise outside the target areas. Comparing the final calculated and the observed difference dataset shows that the phase and amplitude of signals are better explained and fitted by the target-oriented differential strategy.

We conclude that there is no unique strategy which can work properly for all the cases. The use of specific time-lapse strategy depends on the application case, on the quality of the data, on how matched the two acquisition surveys are, and on how accurately the baseline model is recovered. For example, in this case study, the parallel difference strategy could provide better results than the conventional sequential difference strategy. By neglecting the continuous line (4D artifacts corresponding to the main reflector), the parallel difference could recover the patches of injected steam which were not detected by the sequential difference. While in synthetic cases shown in earlier chapters, usually the conventional sequential difference delivers more robust result than the parallel difference.

In this study, we did not take into account the density model since it was not provided. In this configuration, since the dataset contains essentially reflection data, it is important to consider the density model, even at least for computing more accurately the calculated data. In addition, recovering a density model could be crucial for this dataset, because the presence of steam induces density change as well. In this case, the amplitude of reflected data could be better explained during inversion.

Conclusions and Perspectives

General conclusions

In this manuscript, it has been shown that full waveform inversion is a promising technique to be used for time-lapse applications, in order to recover small time-lapse changes quantitatively. I first developed a methodology designed to include prior model information into the classical FWI framwork, and showed how crucial it is for time-lapse inversion (Chapters 1 and 2). Three time-lapse strategies (parallel difference, sequential difference and double-difference (differential)) are proposed for time-lapse inversion. On Marmousi synthetic dataset, I showed that the double-difference strategy can deliver more robust and more accurate time-lapse velocity variations (Chapter 2). I also proposed a target-oriented time-lapse inversion based on the specific design of the prior weighting model, in order to make inversion focus only on the expected area of changes. This target-oriented option can be used with the sequential difference and the double-difference strategies, where the recovered baseline model is used as the prior model (Chapter 2).

A sensitivity study time-lapse variations reconstruction according to different frequency decimations (with frequency-domain FWI) showed that, as anticipated, increasing the number of inverted frequencies increases the accuracy of time-lapse variations. In this case, probably it may be more efficient (in terms of computational cost) to use time-domain FWI for time-lapse applications (Chapter 3). Finally, the suggested baseline recosntruction method including the prior model and the proposed time-lapse strategies were applied to a real field time-lapse dataset (steam injection). The negative velocity variations due to the reinjection of steam are quantitatively recovered (Chapter 4). The differential strategy provides a better time-lapse variation result with less 4D artifacts and compared to the classical migrated images, the FWI delivers a quantitative image of changes.

Including prior model

It has been shown that including prior model information is crucial for time-lapse applications, where there are several kinds of prior information. I have shown the major impact of the appropriate prior model on the inversion, mainly in reducing the sensitivity of baseline inversion to the inaccurate smooth initial model. Consequently, this obtained robust baseline model leads to recover the time-lapse variation in more accurate and more robust ways.

The prior information can be deduced from non-seismic data, well-logging and geological constraints that are generally available for specific exploration applications and for monitoring during production. It has been shown that this prior information improves the well-posedness of the problem as compared to the standard FWI approaches, and allows to partially mitigate potential kinematic inaccuracy of the starting model. Therefore, the inversion becomes less

sensitive to the starting model in terms of cycle-skipping issues. The introduction of prior information would allow to help recover poorly illuminated zones, thus broadening the application of the full waveform inversion.

The prior model definition, may include structures that can help drive the inversion towards the global minimum valley without being kinematically as accurate as the initial model for wave propagation, since the prior model is not used directly to solve the wave equation. However, in poorly illuminated areas where there is less coverage of seismic data, the inversion is driven mainly by the prior model term. In these areas, the updated model will be affected by the structures of the prior model, therefore the design of prior model in these areas must be more precise in terms of kinematic. Moreover, the structures included in prior model may prevent convergence in the final steps of the inversion (for the areas that we have good data coverage) and one may wish to decrease the importance of this prior information with respect to the data information. I have proposed a dynamic weighting of the prior term during the inversion in order to smoothly reduce the impact of the prior information. Since the data-misfit term decreases during the optimization, the weight of prior-model misfit should be decreased as well. The idea of dynamic prior weighting is interesting, but performing it in smart and more efficient ways could be an open question for future works.

Time-lapse inversion strategies and applications

I have studied the robustness of three strategies for time-lapse imaging with regularized FWI in noise-free and noisy synthetic cases. Using the Marmousi model, it has been shown that the parallel difference strategy is highly sensitive to the differential artifacts in the images coming from two independent inversions. Since this strategy considers two inversions starting from a same smooth initial model, it requires a large number of iterations for both baseline and monitor inversions, therfore it would be less interesting in terms of computational costs. The sequential difference strategy is more sensitive to the inaccuracy of the recovered baseline model (which is used as the initial monitor model), because this method attempts to recover the parts of the model which have not been fully reconstructed before. The continuous updating of baseline structures can contaminate the final time-lapse variation model and there is a risk to interpret them as potential 4D variations. This strategy can be more attractive when the baseline model is accurately recovered. In real and practical applications, it would be difficult to ensure the high accuracy of the recovered baseline model, however adding the prior model can improve the accuracy of the obtained model. I should mention that this problem of sequential difference is not related to a convergence of inversion in a wrong direction. This issue rather appears as a mixing of two different effects: the updating of baseline structure and the recovering of time-lapse changes. If it would be possible to qualify the time-lapse variations with another smart way (a kind of intercorrelation technique) rather than a simple subtraction between the two obtained models (baseline and monitor), we could better distinguish 4D changes from other effects.

It has been found that the double-difference strategy is more robust than the sequential one, since the double-difference inversion only focuses on the difference (time-lapse) data, at least in the case of a low level of noise. As shown with local resolution analysis, the difference data are better localized and better focalized at their correct positions using the double-difference strategy as compared to the sequentail difference strategy. In addition, I have illustrated that the double-difference strategy is less sensitive to the inaccuracy of the reconstructed baseline model used as the initial monitor model. In high level of noise environments, it seems that

the double-difference strategy could be more sensitive to the uncorrelated noise which is not cancelled out at the subtraction step. By making a simple data substraction, the standard deviation of random noise will increase and this uncorrelated noise may mask the low energy real time-lapse signals. In this case, we need another way to extract the real time-lapse signals from the high level 4D noise, before perfoming the time-lapse inversion. For double-difference, it is essential to benefit from perfectly matched surveys (e.g. permanent sensors or OBC configurations) to perform subtraction between datasets. In other cases, it is possible to perform extra pre-processing steps and make data interpolation to match two surveys. However, the interpolated data may suffer a bit from interpolation errors which can possibly influence the recovered time-lapse variations.

I also proposed a target-oriented time-lapse imaging with regularized FWI including prior model. For that, in the misfit function, I apply strong prior model constraints outside the areas of expected 4D changes and relatively smaller prior constraints in the target zones. The target-oriented inversion steers the recovery of the monitor model towards the target areas, where the smaller prior model constraints give more relative weight to the data-misfit term. This target-oriented option could be performed as a final focused strategy in real applications, if prior information exists on the location of expected time-lapse variations and when there is no change in overburden. Application to the Marmousi noisy dataset shows that the target-oriented inversion prevents the apparition of image-noise and artifacts outside the target areas. The target-oriented weighting matrix $\mathbf{W}_m$ leads to a similar behavior for the sequential and the double-difference time-lapse strategies and delivers more robust images for quantitative and qualitative interpretations. In particular, for the sequential difference approach, the unexplained baseline events cannot contaminate reconstruction of the time-lapse perturbations and the time-lapse variations are better recovered and better positioned with the target-oriented inversion mode than with the conventional one.

In target-oriented mode, by applying the target-oriented prior model weighting, the misfit function has changed. Before, the misfit function contained several local minima where it was more possible to get trapped in. By constraining inversion to the target areas, the misfit function is reshaped and it seems to become more convex. Therefore, the paths of mimimizing both misfit functions, in conventional and target-oriented modes, are not the same at all which leads to different time-lapse results even inside the target areas. In other words, I can mention that by adding the prior model information to the inversion, the model parameter space is reduced. Hence, the time-lapse inversion becomes more well-posed.

In our particular real case, due to the lack of low frequency content in the data and lack of diving and transmission waves, the FWI method acts as a quantitative non-linear migration technique, as the lower part of the wavenumbers cannot be retrieved from the data part. In order to partially balance this issue, it has been shown that it is necessary to add prior model information into the inversion.

In our real case, the result of the sequential difference strategy is not satisfactory in terms of obtained value of velocity variations. The recovering of residuals left from the previous inversion was mainly an issue when the conventional sequential difference inversion was performed. The parallel difference strategy reconstructs the time-lapse variation zones caused by reinjection of steam. However, the 4D anomalies are less focused and the time-lapse model contains several artifacts coming from reconstruction of the main reflector, through two separate inversions. The conventional differential strategy provides a better time-lapse result around the reservoir depth, with less artifacts compared to the parallel difference. Including the target-oriented

prior weighting into time-lapse inversion scheme has improved the results of the time-lapse inversions, particularly for the sequential difference strategy. In this case, the zones affected by steam injection are better identified. In addition, the target-oriented differential strategy provides a result with less 4D noise artifacts in the near-surface zone and delivers a more robust time-lapse result. The obtained time-lapse velocity variation model is consistent with the classical migrated image of the difference dataset, however the FWI method delivers a quantitative image of parameter changes.

I conclude that there is no unique strategy which can work properly for all the real cases. Using a specific time-lapse strategy depends on the application case. The selection of a proper strategy depends on the quality of data, on how well-matched the two acquisition surveys are, on how accurate the baseline model is recovered and/or on how well baseline data is fitted. In real applications, for using the sequential difference strategy, it is important to know how accurate the baseline data are fitted and explained during baseline reconstruction. If the unexplained residuals left are large (especially for the seismic events related to the target zones), there would be a risk that the low signals of time-lapse varitions will not be taken into account for the inversion (masked by the large unexplained residuals). If this issue occurs for the events corresponding to the target areas, the target-oriented approach can slightly reduce this issue, but cannot solve it completely.

Frequency sampling strategies

A frequency decimation strategy has been investigated in order to compress the inverted data and keep a low complexity of them. For simultaneous frequency inversion, I can conclude that to get accurate time-lapse variations, it is necessary to increase the number of inverted frequencies in order to obtain a more precise baseline model. On the other hand, in order to have less 4D artifacts in the inversion process, the same sampling frequency for the baseline and time-lapse inversions should be considered (this is related to the time-lapse data repeatability issue). We need a smaller baseline sampling frequency to get accurate baseline model and also a smaller time-lapse sampling frequency for time-lapse inversions. Therefore, it is crucial to increase the number of inverted frequencies for time-lapse applications compared to a general exploration project. Due to the large number of inverted frequencies, it might be better to move into time-domain modeling and inversion for time-lapse applications to solve the wave-equation once for all the frequencies and where all the frequency components are inverted at the same time. An efficient time-domain inversion algorithm is required. If we have a reasonable number of sources, doing inversion in time domain is really less expensive and less time-consuming as compared to performing inversion in frequency domain for a large number of frequencies.

However, the observation from this frequency investigation shows that the differential strategy might be performed in frequency domain with a lesser number of frequencies as compared to the sequential difference strategy, since the differential strategy is less sensitive to the inaccuracy of the recovered baseline model. Therefore, in this case, probably we can reach the goal of data compressing, which would be crucial for 3D inversion. On the other hand, probably we should make more efforts to reduce the size of data space for time-lapse applications by other techniques like as source encoding and selection of sources based on their sensitivity to the time-lapse variations, rather than data compressing by a frequency selection.

Multi-group multi-frequency inversion is another interesting strategy for FWI in frequency domain, but this approach appears less interesting for time-lapse applications. Starting from a high resolution baseline model and using multi-group inversion from low to high frequency of

monitor or composite datasets does not work properly, due to the construction of low wavenumbers on the model which already contains high wavenumbers. This leads to generate many artifacts on time-lapse model, even in noise-free case. In this case, the parallel difference strategy may be more compatible with this low to high frequency hierarchy than the two other time-lapse strategies, since both inversions (baseline and monitor) gradually reconstruct the model from low to high wavenumbers. On the other hand, inversion of all the selected frequencies at the same time is more compatible with the sequential difference and the differential strategies.

Perspectives

As a general perspective, it is suggested to extend these 2D acoustic time-lapse inversion strategies to an elastic time-lapse inversion and/or to 3D inversion including a prior model which is designed through more precise techniques. By reconstructing more macro-scale elastic parameters (multi-parameter inversion), we can better constrain the estimation of micro-scale parameters and fluid properties (and their variations in time-lapse applications).

In order to move to intermediate and high frequency inversion, which is crucial for time-lapse applications, it is necessary to reduce the computational costs and required computing memories. To do so, it is suggested to use an efficient time-domain inversion, considering a source-selection technique and reducing the size of the inverse problem to the size of areas of interest.

Prior model

The design of the prior model, for more complex environments and real data applications, should also consider geostatistical approaches and/or standard quantitative interpretation techniques to honor the geological structures (Isaaks and Srivastava, 1989; Dubrule, 2003). Indeed, in addtion to well data, the available geology data could also be integrated for building the model. Moreover, using the prior model penalty term can be extended in 3D FWI.

As a perspective, if we can access the diagonal term of the Hessian, it can be possible to compute the a posteriori covariance matrix. Recently, Fichtner and Trampert (2011) and Demanet et al. (2012) suggested how to approximately compute the Hessian in FWI. In this case, the a posteriori covariance matrix obtained from the baseline reconstruction step can be used as the prior information and preconditioning for the monitor inversion. In addition, the diagonal of the Hessian can be used to balance the prior-model misfit and the data-misfit term for the optimization.

In target-oriented method, it can be suggested to use a dynamic method for changing the target prior weighting model during inversion, starting the inversion with wider and larger target areas, and gradually reducing the size of interested areas.

Time-lapse inversion

In order to better fit the amplitude of time-lapse datasets, it is crucial to take into account a proper density model for forward modeling. Moreover, the inversion of a density model is suggested for future work. In fluid substitution, particularly in the case of gas and water, there would also be significant changes on the density model. Probably, it would be better to invert the impedance parameter.

In this thesis, as a first step for time-lapse FWI, the 2D acoustic inversion was investigated. This time-lapse inversion with introducing of the prior model is suggested to be extended to the elastic approximation and/or to 3D inversion. In case of 3D inversion, the direct solver encounters the computational memory limitation in frequency domain, therefore we should move to the time domain modeling, and especially for time-lapse applications, inversion should also be performed in time domain.

In the case of 3D full waveform inversion, it is necessary to use the data compressing techniques and to reduce the computational cost in order to be able to invert higher frequencies to detect small time-lapse variations. I suggest to reduce the number of sources by selection of sources which are sensitive to the time-lapse variations using elastic wave sensitivity and visibility analysis (Denli and Huang, 2008; Shabelansky et al., 2011). Another computational saving method could be to use a grid injection method to simulate the monitor seismogram based on the computation of the perturbed seismogram only in the variation zones under some assumptions (Robertsson and Chapman, 2000; Robertsson et al., 2000; Borisov and Singh, 2013).

In terms of optimization methods, I suggest to use the truncated Newton routine. Métivier et al. (2013) have shown that the truncated Newton method is less sensitive to the starting inversion at higer frequencies compared to the Gauss-Newton or quasi-Newton methods. Therefore, this optimization technique is less sensitive to the accuracy of initial model. Hence, the truncated Newton method may reduce the sensitivity of time-lapse inversion with respect to inaccuracy of baseline model. In addition, it has been shown that this technique is more robust to noise (Métivier et al., 2013).

In noisy case, to better subtract two datasets for double-difference inversion, it could be possible to use pattern recognition method or co-kriging method to identify the coherent signals between baseline and monitor datasets. By doing so, it is possible to remove the uncoherent noise between two datasets before performing subtraction.

To better qualify the time-lapse changes from two obtained models (baseline and monitor), instead of simple subtraction, we need a smart way to compute the coherency between two images and to extract the real time-lapse variations from other effects such as continuous updating the baseline model (especially for the sequential difference strategy). Similar to the data space in FWI, where several misfit functions between observed and calculated data have been introduced such as difference-based or correlation-based misfit (van Leeuwen and Mulder, 2008, 2010; Masoni et al., 2013) to better extract the information, we may need a better difference method to extract time-lapse model from two obtained images.

For our real application (steam injection), due to specific configuration of dataset containing mainly the reflection events, I suggest to use the waveform inversion of reflected waves in the data domain which is recently proposed by Brossier et al. (2013).

Since for the time-lapse applications, other kinds of data such as electromagnetic and production data are available, joint inversion of seismic data with these datasets can be suggested which could provide a more robust time-lapse model (Liang et al., 2012).

Downscaling

Once the time-lapse variations of macroscale parameters ($\Delta V_p, \Delta V_s, \Delta\rho, \Delta Q_p, ...$) are recovered by time-lapse FWI strategies, it is possible to obtain fluid and poroelastic parameter variations, such as porosity and saturation, by downscaling approaches (Gassmann, 1951; Berryman

et al., 2002; Avseth et al., 2005; Rubino and Velis, 2011; Dupuy, 2011; De Barros et al., 2012). In Appendix A.2, we present a paper related to the downscaling approaches, which is still a work in progress.

Bibliography

Abubakar, A., Hu, W., Habashy, T. M., and van den Berg, P. M. (2009). Application of the finite-difference contrast-source inversion algorithm to seismic full-waveform data. *Geophysics*, 74(6):WCC47–WCC58.

Amestoy, P. R., Duff, I. S., and L'Excellent, J. Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering*, 184:501–520.

Amestoy, P. R., Duff, I. S., L'Excellent, J. Y., and Koster, J. (2003). *Multifrontal massively parallel solver (MUMPS version 4.3): Users'guide*. http://enseeiht.fr/apo/MUMPS/.

Anderson, R. N., Boulanger, A., He, W., Xu, L., Flemings, P. B., Burkhart, T. D., and Hoover, A. R. (1997). 4-D time-lapse seismic monitoring in the south Timbalier 295 field, Gulf of Mexico. In *SEG Technical Program Expanded Abstracts 1997*, pages 868–871.

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2011). Sensitivity analysis of time-lapse images obtained by differential waveform inversion with respect to reference model. *SEG Technical Program Expanded Abstracts*, 30(1):2482–2486.

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2012). Time-lapse imaging using regularized FWI: a robustness study. *SEG Technical Program Expanded Abstracts*, 31(1):1–5.

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2013a). Regularized seismic full waveform inversion with prior model information. *Geophysics*, 78(2):R25–R36.

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2013b). Time-lapse seismic imaging using regularized full waveform inversion with prior model: which strategy? *Geophysical Prospecting*, accepted.

Avseth, P., Mukerji, T., and Mavko, G. (2005). *Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk*. Cambridge University Press, Cambridge, UK.

Ayeni, G. and Biondi, B. (2010). Target-oriented joint least-squares migration/inversion of time-lapse seismic data sets. *Geophysics*, 75(3):R61–R73.

Bansal, R., Routh, P., Krebs, J., Lee, S., Baumstein, A., Anderson, J., Downey, N., Lazaratos, S., Lu, R., and Saldarriaga, S. (2013). Full wavefield inversion of ocean bottom node data. In *EAGE Technical Program Expanded Abstracts 2013*, page We1104.

Beasley, C. J., Chambers, R. E., Workman, R. L., Craft, K. L., and Meister, L. J. (1997). Repeatability of 3-D ocean-bottom cable seismic surveys. *The Leading Edge*, 16:1281–1285.

Ben Hadj Ali, H. (2009). *Three dimensional visco-acoustic frequency domain full waveform inversion.* PhD thesis, Université de Nice-Sophia-Antipolis.

Berenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.

Berryman, J., Berge, P., and Bonner, B. (2002). Estimating rock porosity and fluid saturation using only seismic velocities. *Geophysics*, 67:391–404.

Bertrand, A. and MacBeth, C. (2003). Sea water velocity variations and real time reservoir monitoring. *The Leading Edge*, 22:351–355.

Beydoun, W. B. and Mendes, M. (1989). Elastic ray-Born ℓ^2 - migration/inversion. *Geophysical Journal*, 97:151–160.

Beydoun, W. B. and Tarantola, A. (1988). First Born and Rytov approximation: Modeling and inversion conditions in a canonical example. *Journal of the Acoustical Society of America*, 83:1045–1055.

Borisov, D. and Singh, S. (2013). Time-lapse elastic 3D full waveform inversion using grid injection method. In *EAGE Technical Program Expanded Abstracts 2013*, page Th P10 15.

Born, M. and Wolf, E. (1980). *Principles of optics.* Pergamon Press, 6^{th} edition edition.

Born, M. and Wolf, E. (1993). *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light; sixth edition with corrections.* Pergamon Press.

Brossier, R. (2009). *Imagerie sismique à deux dimensions des milieux visco-élastiques par inversion des formes d'onde: développements méthodologiques et applications.* PhD thesis, Université de Nice-Sophia-Antipolis.

Brossier, R. (2011). Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance. *Computers & Geosciences*, 37(4):444 – 455.

Brossier, R., Operto, S., and Virieux, J. (2009). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):WCC105–WCC118.

Brossier, R., Operto, S., and Virieux, J. (2013). Toward data-domain waveform inversion of reflected waves. In *EAGE Technical Program Expanded Abstracts 2013*, page Workshop F01.

Brossier, R., Virieux, J., and Operto, S. (2008). Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling. *Geophysical Journal International*, 175(2):541–559.

Bunks, C., Salek, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.

Byrd, R., Lu, P., and Nocedal, J. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 16:1190–1208.

Calvert, R. (2005). *Insights and methods for 4D reservoir monitoring and characterization.* SEG Distinguished Instructor Series No. 8.

Castagna, J. P., Batzle, M. L., and Kan, T. K. (1993). Rock physics - the link between rock properties and avo response, in offset-dependent reflectivity-theory and practice of avo analysis, edited by Castagna J.P. and Backus M.M.

Castellanos, J. L., Gomes, S., and Guerra, V. (2002). The triangle method for finding the corner of the L-curve. *Applied Numerical Mathematics*, 43:359–373.

Červený, V. (2001). *Seismic Ray Theory.* Cambridge University Press, Cambridge.

Chauris, H. and Noble, M. (2001). 2D velocity macro model estimation from seismic reflection data by local Differential Semblance Optimization: applications on synthetic and real data. *Geophysical Journal International*, 144:14–26.

Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002a). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I: Theoretical aspects. *Geophysics*, 67(4):1202–1212.

Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002b). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part II: Applications on synthetic and real data. *Geophysics*, 67(4):1213–1224.

Chauris, H., Noble, M., and Taillander, C. (2008). Which initial model do we need for full waveform inversion? In *Expanded Abstracts, 70th Annual EAGE Conference & Exhibition, WO11: Full waveform inversion - Current status and perspectives, Roma.* EAGE.

Chavent, G. (1974). Identification of parameter distributed systems. In Goodson, R. and Polis, M., editors, *Identification of function parameters in partial differential equations*, pages 31–48. American Society of Mechanical Engineers, New York.

Chavent, G. (2009). *Nonlinear least squares for inverse problems.* Springer Dordrecht Heidelberg London New York.

Chavent, G. and Jacewitz, C. A. (1995). Determination of background velocities by multiple migration fitting. *Geophysics*, 60(2):476–490.

Claerbout, J. (1971). Towards a unified theory of reflector mapping. *Geophysics*, 36:467–481.

Claerbout, J. F. (1976). *Fundamentals of Geophysical Data Processing.* McGraw-Hill Book Co.

Coates, R. T. and Chapman, C. H. (1990). Ray perturbation theory and the Born approximation. *Geophysical Journal International*, 100:379–392.

Cohen, J. K. and Stockwell, J. J. W. (2008). CWP/SU: Seismic Unix release No. 41: an open source software package for seismic research and processing. *Center for Wave Phenomena, Colorado School of Mines.*

Dai, N., Vafidis, A., and Kanasewich, E. (1995). Wave propagation in heteregeneous porous media: A velocity-stress, finite-difference method. *Geophysics*, 60(2):327–340.

Day-Lewis, F. D., Harris, J. M., and Gorelick, S. M. (2002). Time-lapse inversion of crosswell radar data. *Geophysics*, 67(6):1740–1752.

Day-Lewis, F. D., Jr., J. W. J. L., Harris, J. M., and Gorelick, S. M. (2003). Time-lapse imaging of saline-tracer transport in fractured rock using difference-attenuation radar tomography. *Water Resources Research*, 39(10):1290.

De Barros, L., Dupuy, B., O'Brien, G., Virieux, J., and Garambois, S. (2012). *Using a poroelastic theory to reconstruct subsurface properties: numerical investigation.* Seismic Waves, Research and Analysis, Intech, Masaki Kanao (Ed.).

Demanet, L., Letourneau, P. D., Boumal, N., Calandra, H., Chiu, J., and Snelson, S. (2012). Matrix probing: a randomized preconditioner for the wave-equation hessian. *Applied Computational Harmonic Analysis*, 32:155–168.

Denli, H. and Huang, L. (2008). Elastic-wave sensitivity analysis for seismic monitoring. In *SEG Technical Program Expanded Abstracts 2008*, pages 30–34.

Denli, H. and Huang, L. (2009). Double-difference elastic waveform tomography in the time domain. *SEG Technical Program Expanded Abstracts*, 28(1):2302–2306.

Dubrule, O. (2003). *Geostatistics for seismic data integration in earth models.* SEG Distinguished Instructor Series No. 6.

Duff, I. S. and Reid, J. K. (1983). The multifrontal solution of indefinite sparse symmetric linear systems. *ACM Transactions on Mathematical Software*, 9:302–325.

Dumbser, M. and Käser, M. (2006). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes II: The Three-Dimensional Isotropic Case. *Geophysical Journal International*, 167(1):319–336.

Dupuy, B. (2011). *Propagation des ondes sismiques dans les milieux multiphasiques hétérogènes : modélisation numérique, sensibilité et inversion des paramètres poroélastiques.* PhD thesis, Université Joseph Fourier Grenoble.

Fang, Y., Cheney, M., and Roecker, R. (2010). Imaging from sparse measurements. *Geophysical Journal International*, 180(3):1289–1302.

Farquharson, C. G. and Oldenburg, D. W. (1998). Non-linear inversion using general measures of data misfit and model structure. *Geophysical Journal International*, 134:213–227.

Fichtner, A., Kennett, B. L. N., Igel, H., and Bunge, H. P. (2010). Full waveform tomography for radially anisotropic structure: New insights into present and past states of the Australasian upper mantle. *Earth and Planetary Science Lettters*, 290(3-4):270–280.

Fichtner, A. and Trampert, J. (2011). Resolution analysis in full waveform inversion. *Geophysical Journal International*, 187:1604–1624.

Forgues, E. and Lambaré, G. (1997). Parameterization study for acoustic and elastic ray+born inversion. *Journal of Seismic Exploration*, 6:253–278.

Freudenreich, Y. and Singh, S. (2000). Full waveform inversion for seismic data - frequency versus time domain. In *EAGE Technical Program Expanded Abstracts 2000*, page C54.

Gassmann, F. (1951). Über die elastizität poröser medien. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 96:1–23.

Gauthier, O., Virieux, J., and Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms: numerical results. *Geophysics*, 51(7):1387–1403.

Gelis, C., Virieux, J., Grandjean, G., Leparoux, D., and Operto, S. (2004). Full waveform elastic inversion in a space frequency domain formulation: a powerful geotechnical tool for superficial reconstruction. In *EOS Trans. AGU*, volume 85. American Geophysical Union.

Gholami, Y., Brossier, R., Operto, S., Prieux, V., Ribodetti, A., and Virieux, J. (2013a). Which parametrization is suitable for acoustic VTI full waveform inversion? - Part 2: application to Valhall. *Geophysics*, 78(2):R107–R124.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A., and Virieux, J. (2013b). Which parametrization is suitable for acoustic VTI full waveform inversion? - Part 1: sensitivity and trade-off analysis. *Geophysics*, 78(2):R81–R105.

Gosselet, A. and Singh, S. (2008). 2D full wave form inversion in time-lapse mode: CO2 quantification at Sleipner. In *EAGE, 70th Conference and Exhibition, Expanded abstracts*, page W075.

Got, J. L., Monteiller, V., Virieux, J., and Operto, S. (2008). Potential and limits of double-difference tomographic methods. *Geophysical Prospecting*, 56(4):477–491.

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

Greenhalgh, S., Zhou, B., and Green, A. (2006). Solutions, algorithms and inter-relations for local minimization search geophysical inversion. *Journal of Geophysics and Engineering*, 3(2):101–113.

Guitton, A. (2011). A blocky regularization scheme for full waveform inversion. *SEG Technical Program Expanded Abstracts*, 30(1):2418–2422.

Guitton, A., Ayeni, G., and Gonzales, G. (2010). A preconditioning scheme for full waveform inversion. *SEG Technical Program Expanded Abstracts*, 29(1):1008–1012.

Hale, D. (2013). Dynamic warping of seismic images. *Geophysics*, 78(2):S105–S115.

Hall, S., MacBeth, C., Barkved, O. I., and Wild, P. (2005). Cross-matching with interpreted warping of 3D streamer and 3D ocean-bottom-cable data at Valhall for time-lapse assessment. *Geophysical Prospecting*, 53(2):283–297.

Hall, S., Michelena, R., Mavko, G., Wild, P., and MacBeth, C. (2002). Time-lapse seismic monitoring of compaction and subsidence at Valhall through cross-matching and interpreted warping of 3D streamer and OBC data. In *SEG Technical Program Expanded Abstracts 2002*, pages 1696–1699.

Hall, S. A. (2006). A methodology for 7D warping and deformation monitoring using time-lapse seismic data. *Geophysics*, 71(4):O21–O31.

Han, D.-H., Nur, A., and Morgan, D. (1986). Effects of porosity and clay content on wave velocities in sandstones. *Geophysics*, 51(11):2093–2107.

Hansen, C. (1998). *Rank-defficient and discrete ill-posed problems - Numerical aspects of linear inversion*. Society for Industrial and Applied Mathematics - Mathematical modeling and computation.

Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34(4):561–580.

Hansen, P. C. (2010). *Discrete Inverse Problems: Insight and Algorithms*. SIAM.

Harris, J. M., Langan, R. T., Fasnacht, T., Melton, D., Smith, B., and Sinton, J. andTan, H. (1996). Experimental verification of seismic monitoring of CO2 injection in carbonate reservoirs. In *SEG Technical Program Expanded Abstracts 1996*, pages 1870–1872.

Harris, P. E. and Henry, B. (1998). Time lapse processing: A North Sea case study. In *SEG Technical Program Expanded Abstracts 1998*, pages 1–4.

He, W., Guerin, G., Anderson, R. N., and Mello, U. T. (1998). Time-dependent reservoir characterization of the LF sand in the south Eugene Island 330 field, Gulf of Mexico. *The Leading Edge*, 17:1434–1438.

Herrmann, F. J., Erlangga, Y. A., and Lin, T. T. Y. (2009). Compressive simultaneous full-waveform simulation. *Geophysics*, 74(4):A35–A40.

Hu, W., Abubakar, A., and Habashy, T. M. (2009). Simultaneous multifrequency inversion of full-waveform seismic data. *Geophysics*, 74(2):R1–R14.

Hudson, J. A. and Heritage, J. R. (1981). *Use of the Born approximation in seismic scattering problems*. Cambridge University Press.

Hustedt, B., Operto, S., and Virieux, J. (2004). Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling. *Geophysical Journal International*, 157:1269–1296.

Isaaks, E. H. and Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press.

Jenkins, S. D., Waite, M. W., and Bee, M. F. (1997). Time-lapse monitoring of the Duri steamflood: A pilot and case study. *The Leading Edge*, 16:1267–1273.

Jin, S., Madariaga, R., Virieux, J., and Lambaré, G. (1992). Two-dimensional asymptotic iterative elastic inversion. *Geophysical Journal International*, 108:575–588.

Jo, C. H., Shin, C., and Suh, J. H. (1996). An optimal 9-point, finite-difference, frequency-space 2D scalar extrapolator. *Geophysics*, 61:529–537.

Johnstad, S. E., Seymour, R. H., and Smith, P. J. (1995). Seismic reservoir monitoring over the Oseberg field during the period 1989-1992. *First Break*, 13(5):169–183.

Käser, M. and Dumbser, M. (2006). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes I: The Two-Dimensional Isotropic Case with External Source Terms. *Geophysical Journal International*, 166:855–877.

Lailly, P. (1983a). The seismic inverse problem as a sequence of before stack migrations. In Bednar, R. and Weglein, editors, *Conference on Inverse Scattering, Theory and application, Society for Industrial and Applied Mathematics, Philadelphia*, pages 206–220.

Lailly, P. (1983b). The seismic problem as a sequence of before-stack migrations. In Bednar, J., editor, *Conference on Inverse Scattering: Theory and Applications*. SIAM, Philadelphia.

Landrø, M. (2001). Discrimination between pressure and fluid saturation changes from time-lapse seismic data. *Geophysics*, 66(3):836–844.

Le Stunff, Y. and Grenier, D. (1998). Taking into account a priori information in 3D tomography. In *SEG Technical Program Expanded Abstracts*, volume 17, pages 1875–1878. SEG.

Levander, A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least-squares. *Quarterly of Applied Mathematics*, 2:162–168.

Li, Y. and Oldenburg, D. W. (1998). 3-D inversion of gravity data. *Geophysics*, 63(1):109–119.

Liang, L., Abubakar, A., and Habashy, T. M. (2012). Joint inversion of time-lapse crosswell electromagnetic, seismic, and production data for reservoir monitoring and characterization. In *SEG Technical Program Expanded Abstracts*, volume 31(1), pages 1–7.

Liao, Q. and McMechan, G. A. (1996). Multifrequency viscoacoustic modeling and inversion. *Geophysics*, 61:1371–1378.

Loris, I., Douma, H., Nolet, G., Daubechies, I., and Regone, C. (2010). Nonlinear regularization techniques for seismic tomography. *Journal of Computational Physics*, 229:890–905.

Lu, R., Lazaratos, S., Wang, K., Cha, Y., Chikichev, I., and Prosser, R. (2013). High-resolution elastic FWI for reservoir characterization. In *EAGE Technical Program Expanded Abstracts 2013*, page Th1002.

Lumley, D., Adams, D., Wright, R., Markus, D., and Cole, S. (2008). Seismic monitoring of CO2 geo-sequestration: realistic capabilities and limitations. In *SEG Technical Program Expanded Abstracts*, volume 27(1), pages 2841–2845.

Lumley, D. E. (1995a). 4-D seismic monitoring of an active steamflood. In *SEG Technical Program Expanded Abstracts 1995*, pages 203–206.

Lumley, D. E. (1995b). *Seismic time-lapse monitoring of subsurface fluid flow*. PhD thesis, Stanford University.

Lumley, D. E. (2001). Time-lapse seismic reservoir monitoring. *Geophysics*, 66(1):50–53.

Luo, Y. and Schuster, G. T. (1990). Parsimonious staggered grid finite-differencing of the wave equation. *Geophysical Research Letters*, 17(2):155–158.

Ma, Y., Hale, D., Gong, B., and Meng, Z. J. (2012). Image-guided sparse-model full waveform inversion. *Geophysics*, 77(4):R189–R198.

MacKay, S., Fried, J., and Carvill, C. (2003). The impact of water-velocity variations on deepwater seismic data. *The Leading Edge*, 22:344–350.

Malinowski, M. and Operto, S. (2008). Quantitative imaging of the Permo-Mesozoic complex and its basement by frequency domain waveform tomography of wide-aperture seismic data from the Polish basin. *Geophysical Prospecting*, 56:805–825.

Malinowski, M., Operto, S., and Ribodetti, A. (2011). High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion. *Geophysical Journal International*, 186(3):1179–1204.

Marfurt, K. (1984). Accuracy of finite-difference anf finite-element modeling of the scalar and elastic wave equations. *Geophysics*, 49:533–549.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441.

Martin, G. S., Wiley, R., and Marfurt, K. J. (2006). Marmousi2: An elastic upgrade for marmousi. *The Leading Edge*, 25(2):156–166.

Masoni, I., Brossier, R., Virieux, J., and Boelle, J. (2013). Alternative misfit functions for FWI applied to surface waves. In *EAGE Technical Program Expanded Abstracts 2013*, page Th P10 13. EAGE.

Matheron, G. (1970). The theory of regionalized variables and its applications. Technical Report Fasc. 5, Centre de Géostatistique de l'Ecole des Mines de Paris.

McMechan, G. (1989). A review of seismic acoustic imaging by reverse time migration. *International Journal of Imaging Systems and Technology*, 1:18–21.

Mead, J. L. and Renaut, R. A. (2009). A Newton root-finding algorithm for estimating the regularization parameter for solving ill-conditioned least squares problems. *Inverse Problems*, 25(2):025002.

Menke, W. (1984). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, Inc., Orlando, USA.

Métivier, L., Brossier, R., Virieux, J., and Operto, S. (2013). Full waveform inversion and the truncated newton method. *SIAM Journal On Scientific Computing*, 35(2):B401–B437.

Meunier, J. and Herculin, S. (2003). Evaluation and handling of positioning differences in 4D seismic. In *SEG Technical Program Expanded Abstracts 2003*, pages 75–78. SEG.

Miller, C. R., Routh, P. S., Brosten, T. R., and McNamara, J. P. (2008). Application of time-lapse ERT imaging to watershed characterization. *Geophysics*, 73(3):G7–G17.

Moldoveanu, N., van Baaren, P., Addessi, D., Stubbington, L., and Combee, L. (1996). Repeatability of the seismic experiments for 4-D seismic in transition zone surveys. In *SEG Technical Program Expanded Abstracts 1996*, pages 5–8. SEG.

Monteiller, V., Got, J.-L., Virieux, J., and Okubo, P. (2005). An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano. *Journal of Geophysical Research*, 110(B12306):doi:10.1029/2004JB00346.

Mora, P. R. (1987). Nonlinear two-dimensional elastic inversion of multi-offset seismic data. *Geophysics*, 52:1211–1228.

Mora, P. R. (1988). Elastic wavefield inversion of reflection and transmission data. *Geophysics*, 53:750–759.

Mora, P. R. (1989). Inversion = migration + tomography. *Geophysics*, 54(12):1575–1586.

Moré, J. (1978). The levenberg-marquardt algorithm: Implementation and theory. In Watson, G., editor, *Numerical Analysis*, volume 630 of *Lecture Notes in Mathematics*, pages 105–116. Springer Berlin / Heidelberg. 10.1007/BFb0067700.

MUMPS-team (2009). *MUMPS - MUltifrontal Massively Parallel Solver users' guide - version 4.9.2 (November 5, 2009)*. ENSEEIHT-ENS Lyon, http://www.enseeiht.fr/apo/MUMPS/ or http://graal.ens-lyon.fr/MUMPS.

Nocedal, J. (1980). Updating Quasi-Newton Matrices With Limited Storage. *Mathematics of Computation*, 35(151):773–782.

Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization*. New York, US : Springer.

Nur, A., Tosaya, C., and Thanh, D. V. (1984). Seismic monitoring of thermal enhanced oil recovery processes. In *SEG Technical Program Expanded Abstracts 1984*, page RS6.

Oldenborger, G. A., Knoll, M., Routh, P., and LaBrecque, D. (2007). Time-lapse ERT monitoring of an injection/withdrawal experiment in a shallow unconfined aquifer. *Geophysics*, 72(4):F177–F187.

Oldenburg, D. W. (1994). Practical strategies for the solution of large-scale electromagnetic inverse problems. *Radio Science*, 29(4):1081–1099.

Operto, S., Brossier, R., and Virieux, J. (2007). Documentation of FWM2DPSV programs: 2D P-SV finite-difference time-domain modelling of elastic wave propagation. Technical Report TR 07, SEISCOPE Project.

Operto, S. and Virieux, J. (2006). SEISCOPE consortium: seismic imaging of complex structures from multicomponent global offset data by full waveform inversion. Technical report, UMR Géosciences Azur - CNRS. Consortium SEISCOPE.

Operto, S., Virieux, J., Dessa, J. X., and Pascal, G. (2006). Crustal imaging from multifold ocean bottom seismometers data by frequency-domain full-waveform tomography: application to the eastern Nankai trough. *Journal of Geophysical Research*, 111(B09306):doi:10.1029/2005JB003835.

Operto, S., Virieux, J., Hustedt, B., and Malfanti, F. (2002). Adaptative wavelet-based finite-difference modelling of SH-wave propagation. *Geophysical Journal International*, 148:1–28.

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(2):495–503.

Plessix, R. E. (2009). Three-dimensional frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, 74(6):WCC53–WCC61.

Plessix, R. E. (2012). Waveform inversion overview: Where are we? and what are the challenges? In *Expanded Abstracts*. EAGE.

Plessix, R. E., Michelet, S., Rynja, H., Kuehl, H., Perkins, C., de Maag, J. W., and Hatchell, P. (2010). Some 3d applications of full waveform inversion. In *Expanded Abstracts*. EAGE.

Plessix, R. E. and Mulder, W. A. (2004). Frequency domain finite difference amplitude preserving migration. *Geophysical Journal International*, 157:975–987.

Plessix, R. E. and Mulder, W. A. (2008). Resistivity imaging with controlled-source electromagnetic data: depth and data weighting. *Inverse Problems*, 24:034012.

Plessix, R. E. and Perkins, C. (2010). Full waveform inversion of a deep water ocean bottom seismometer dataset. *First Break*, 28:71–78.

Porter-Hirsche, J. L. and Hirsche, K. W. (1998). Repeatability study of land data acquisition and processing for time-lapse seismic. In *SEG Technical Program Expanded Abstracts 1998*, pages 9–11. SEG.

Pratt, R. G. (1990a). Frequency-domain elastic modeling by finite differences: a tool for crosshole seismic imaging. *Geophysics*, 55(5):626–632.

Pratt, R. G. (1990b). Inverse theory applied to multi-source cross-hole tomography. part II : elastic wave-equation method. *Geophysical Prospecting*, 38:311–330.

Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part I : theory and verification in a physic scale model. *Geophysics*, 64:888–901.

Pratt, R. G., Shin, C., and Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362.

Pratt, R. G. and Shipp, R. M. (1999). Seismic waveform inversion in the frequency domain, part II: Fault delineation in sediments using crosshole data. *Geophysics*, 64:902–914.

Pratt, R. G., Song, Z. M., Williamson, P. R., and Warner, M. (1996). Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, 124:323–340.

Pratt, R. G. and Symes, W. (2002). Semblance and differential semblance optimisation for waveform tomography: a frequency domain implementation. In *Journal of Conference Abstracts*, volume 7(2), pages 183–184. Cambridge publications.

Pratt, R. G. and Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part I: acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 3 edition.

Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O., and Kommedal, J. (2011). On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophysical Journal International*, 187:1495–1515.

Prieux, V., Brossier, R., Operto, S., and Virieux, J. (2012). Two-dimensional anisotropic visco-elastic full waveform inversion of wide-aperture 4C OBC data from the Valhall field. In *Expanded Abstracts*. EAGE.

Prieux, V., Brossier, R., Operto, S., and Virieux, J. (2013a). Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 1: imaging compressional wavespeed, density and attenuation. *Geophysical Journal International*, doi: 10.1093/gji/ggt177.

Prieux, V., Brossier, R., Operto, S., and Virieux, J. (2013b). Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 2: imaging compressional and shear-wave velocities. *Geophysical Journal International*, doi: 10.1093/gji/ggt178.

Prieux, V., Lambaré, G., Operto, S., and Virieux, J. (2013c). Building starting model for full waveform inversion from wide-aperture data by stereotomography. *Geophysical Prospecting*, 61(Issue supplement: 60 year anniversary issue):109–137.

Pullin, N., Matthews, L., and Hirsche, K. (1987). Techniques applied to obtain very high resolution three-dimensional seismic imaging at an Athabasca tar sands thermal pilot. *The Leading Edge*, 6:10–15.

QueiBer, M. and Singh, S. C. (2013). Full waveform inversion in the time lapse mode applied to CO2 storage at Sleipner. *Geophysical Prospecting*, 61(3):537–555.

Raiga-Clemenceau, J., Martin, J., and Nicoletis, S. (1988). The concept of acoustic formation factor for more accurate porosity determination from sonic transit time data. *Log Analyst*, 219:54–60.

Ramirez, A., Daily, W., LaBrecque, D., Owen, E., and Chesnut, D. (1993). Monitoring an underground steam injection process using electrical resistance tomography. *Water Resources Research*, 29(1):73–87.

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A., and dell'Aversana, P. (2004). Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions: application to a thrust belt. *Geophysical Journal International*, 159:1032–1056.

Raymer, L., Hunt, E., and Gardner, J. (1980). An improved sonic transit time-to-porosity transform. In *Abstracts Soc Professional Well Log Analysis (SPWLA), 21st Ann. Logg. Symp., Paper P July*.

Rennie, J., Alexandre, R., and Ronen, S. (1997). Sensitivity of repeat 3-D seismic surveys to geometry variations-A controlled experiment. In *SEG Technical Program Expanded Abstracts 1997*, pages 91–95. SEG.

Ribodetti, A. and Virieux, J. (1996). Asymptotic theory for imaging the attenuation factors Q_p and Q_s. In *Inverse Problems of Wave Propagation and Diffraction, Proceedings, Aix-les-Bains, France 1996*, pages 334–353. Springer-Verlag.

Richtmyer, R. D. and Morton, K. W. (1967). *Difference methods for initial value problems.* Wiley-Interscience, Kreiger, New York; réimpression (1994).

Rickett, J. and Lumley, D. E. (1998). A cross-equalization processing flow for off-the-shelf 4-D seismic data. In *SEG Technical Program Expanded Abstracts 1998*, pages 16–19.

Rickett, J. E. and Lumley, D. E. (2001). Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the Gulf of Mexico. *Geophysics*, 66(4):1015–1025.

Robertsson, J. and Chapman, C. (2000). An efficient method for calculating finite-difference seismograms after model alterations. *Geophysics*, 65(3):907–918.

Robertsson, J., Ryan-Grigor, S., Sayers, C. M., and Chapman, C. (2000). A finite-difference injection approach to modeling seismic fluid flow monitoring. *Geophysics*, 65(3):896–906.

Romdhane, A., Ravaut, C., and Querendez, E. (2012). CO2 monitoring at the Sleipner field with full waveform inversion. In *Geophysical Research Abstracts, EGU 2012*, volume 14. EGU.

Ronen, S., van Waard, R., and Keggin, J. (1999). Repeatability of sea bed multi-component data. In *SEG Technical Program Expanded Abstracts 1999*, pages 1330–1333. SEG.

Routh, P. S. and Anno, P. D. (2008). Time-lapse noise characterization by inversion. *SEG Technical Program Expanded Abstracts*, 27(1):3143–3147.

Routh, P. S. and Oldenburg, D. W. (1999). Inversion of controlled source audio-frequency magnetotellurics data for a horizontally layered earth. *Geophysics*, 64(6):1689–1697.

Roux, P., Iturbe, I., Nicolas, B., Virieux, J., and Mars, J. I. (2011). Travel-time tomography in shallow water: Experimental demonstration at an ultrasonic scale. *The Journal of the Acoustical Society of America*, 130(3):1–10.

Rubino, J. and Velis, D. (2011). Seismic characterization of thin beds containing patchy carbon dioxide-brine distributions: a study based on numerical simulations. *Geophysics*, 76(3):R57–R67.

Rudin, L., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268.

Sava, P. and Biondi, B. (2004). Wave-equation migration velocity analysis. ii. subsalt imaging examples. *Geophysical Prospecting*, 52(6):607–623.

Sears, T., Singh, S., and Barton, P. (2008). Elastic full waveform inversion of multi-component OBC seismic data. *Geophysical Prospecting*, 56(6):843–862.

Sears, T. J., Barton, P. J., and Singh, S. C. (2010). Elastic full waveform inversion of multicomponent ocean-bottom cable seismic data: Application to alba field, u. k. north sea. *Geophysics*, 75(6):R109–R119.

Shabelansky, A. H., Malcolm, A., and Fehler, M. (2011). Visibility analysis using reverse time wave sensitivity for time-lapse target-oriented imaging. In *SEG Technical Program Expanded Abstracts 2011*, pages 4155–4159.

Shen, P. and Symes, W. W. (2008). Automatic velocity analysis via shot profile migration. *Geophysics*, 73(5):VE49–VE59.

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

Shin, C. and Sohn, H. (1998). A frequency-space 2-D scalar wave extrapolator using extended 25-point finite-difference operator. *Geophysics*, 63:289–296.

Shipp, R. M. and Singh, S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, 151:325–344.

Singha, K. and Gorelick, S. M. (2005). Saline tracer visualized with three-dimensional electrical resistivity tomography: Field-scale spatial moment analysis. *Water Resources Research*, 41(5):W05023.

Sirgue, L., Barkved, O. I., Dellinger, J., Etgen, J., Albertin, U., and Kommedal, J. H. (2010). Full waveform inversion: the next leap forward in imaging at Valhall. *First Break*, 28:65–70.

Sirgue, L., Etgen, T. J., Albertin, U., and Brandsberg-Dahl, S. (2007). System and method for 3D frequency-domain waveform inversion based on 3D time-domain forward modeling. *US Patent Application Publication*, US2007/0282535 A1.

Sirgue, L. and Pratt, R. G. (2004). Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. *Geophysics*, 69(1):231–248.

Sonneland, L., Veire, H. H., Raymond, B., Signer, C., Pedersen, L., Ryan, S., and Sayers, C. (1997). Seismic reservoir monitoring on gullfaks. *The Leading Edge*, 16:1247–1252.

Stekl, I. and Pratt, R. G. (1998). Accurate viscoelastic modeling by frequency-domain finite difference using rotated operators. *Geophysics*, 63:1779–1794.

Symes, W. W. (2008). Migration velocity analysis and waveform inversion. *Geophysical Prospecting*, 56:765–790.

Symes, W. W. and Carazzone, J. J. (1991). Velocity inversion by differential semblance optimization. *Geophysics*, 56:654–663.

Taillandier, C., Noble, M., Chauris, H., and Calandra, H. (2009). First-arrival travel time tomography based on the adjoint state method. *Geophysics*, 74(6):WCB1–WCB10.

Tape, C., Liu, Q., Maggi, A., and Tromp, J. (2010). Seismic tomography of the southern california crust based on spectral-element and adjoint methods. *Geophysical Journal International*, 180:433–462.

Tarantola, A. (1984a). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.

Tarantola, A. (1984b). Linearized inversion of seismic reflection data. *Geophysical Prospecting*, 32:998–1015.

Tarantola, A. (1986). A strategy for non linear inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.

Tarantola, A. (1987). *Inverse problem theory: methods for data fitting and model parameter estimation*. Elsevier, New York.

Tarantola, A. (2005). *Inverse Problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, Philadelphia.

Thore, P., Allouche, H., Lys, P., and Tarrass, I. (2010). Retrieving 4D signal in complex media using the full waveform inversion paradigm. In *72nd EAGE Conference & Exhibition*. EAGE.

Tikhonov, A. and Arsenin, V. (1977). *Solution of ill-posed problems*. Winston, Washington, DC.

Tromp, J., Tape, C., and Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160:195–216.

Tura, A. and Lumley, D. E. (1999). Estimating pressure and saturation changes from timelapse AVO data. In *SEG Technical Program Expanded Abstracts 1999*, pages 1655–1658.

van den Berg, P., Abubakar, A., and Fokkema, J. (2003). Multiplicative regularization for contrast profile inversion. *Radio Science*, 38:23.1–23.10.

van den Berg, P. M., van Broekhoven, A. L., and Abubakar, A. (1999). Extended contrast source inversion. *Inverse Problems*, 15:1325–1344.

van Leeuwen, T. and Mulder, W. (2008). Velocity analysis based on data correlation. *Geophysical Prospecting*, 56(6):791–803.

van Leeuwen, T. and Mulder, W. A. (2010). A correlation-based misfit criterion for wave-equation traveltime tomography. *Geophysical Journal International*, 182(3):1383–1394.

Virieux, J. (1984). SH wave propagation in heterogeneous media, velocity-stress finite difference method. *Geophysics*, 49:1259–1266.

Virieux, J. (1986). P-SV wave propagation in heterogeneous media, velocity-stress finite difference method. *Geophysics*, 51:889–901.

Virieux, J. and Lambaré, G. (2007). Theory and observations - body waves: ray methods and finite frequency effects. In Romanovitz, B. and Diewonski, A., editors, *Treatise of Geophysics, volume 1: Seismology and structure of the Earth*. Elsevier.

Virieux, J. and Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.

Waldhauser, F. and Ellsworth, W. (2000). A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America*, 90(6):1353–1368.

Wang, C., Yingst, D., Bloor, R., and Leveille, J. (2012). Vti waveform inversion with practical strategies: Application to 3d real data. In *SEG Technical Program Expanded Abstracts 2012*, pages 1–6. Society of Exploration Geophysicists.

Wang, H., Singh, S., and Calandra, H. (2013). Integrated inversion with combined wave-equation tomography and FWI. In *EAGE Technical Program Expanded Abstracts 2013*, page Workshop WA02.

Wang, Z. and Nur, A. (1990). Wave velocities in hydrocarbon-saturated rocks. *Geophysics*, 55:723–733.

Watanabe, T., Shimizu, S., Asakawa, E., and Matsuoka, T. (2004). Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring. *SEG Technical Program Expanded Abstracts*, 23(1):2323–2326.

Williamson, P. R., Cherrett, A. J., and Sexton, P. A. (2007). A new approach to warping for quantitative time-lapse characterisation. In *EAGE Technical Program Expanded Abstracts 2007*, page P064.

Wu, R. S. and Aki, K. (1985). Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics*, 50(4):582–595.

Wyllie, M., Gregory, J., and Gardner, L. (1956). Elastic wave velocities in heterogneous and porous media. *Geophysics*, 21:41–70.

Zabihi Naeini, E. (2012). Acquisition footprint removal from time-lapse datasets. In *EAGE Technical Program Expanded Abstracts 2012*, page P292. EAGE.

Zabihi Naeini, E., Hoeber, H., and Campbell, S. (2012). A new approach to reducing multiple leakage on time-lapse datasets. In *EAGE Technical Program Expanded Abstracts 2012*, page E010. EAGE.

Zabihi Naeini, E., Hoeber, H., King, B., Rumyantsev, D., and Buker, F. (2009). Simultaneous multi-vintage 4D binning. In *EAGE Technical Program Expanded Abstracts 2009*. EAGE.

Zabihi Naeini, E., Hoeber, H., Poole, G., Buker, F., and van Schaack, M. (2010). Simultaneous multi-vintage multi-parameter time-lapse matching. In *EAGE Technical Program Expanded Abstracts 2010*, page B038. EAGE.

Zhang, H. and Thurber, C. H. (2003). Double difference tomography: The method and its application to the Hayward fault, California. *Bulletin of the Seismological Society of America*, 93(5):1875–1889.

Zheng, Y., Barton, P., and Singh, S. (2013). Time-lapse waveform inversion for a compacting reservoir. In *EAGE Technical Program Expanded Abstracts 2013*, page Th P10 16.

Zhou, R., Huang, L., and Rutledge, J. (2010). Microseismic event location for monitoring CO2 injection using double-difference tomography. *The Leading Edge*, 29(2):208–213.

Appendix A

Supplementary Publications

Contents

This appendix contains two supplementary articles which were presented and prepared during this thesis.

The first one is an Extended Abstract which was presented in SEG annual meeting 2011 and it has been honored as "**SEG Award of Merit**", Best Student paper presented in SEG 2011.

The second one is a paper related to the downscaling approaches after getting the time-lapse variations of macro-scale parameters. This paper is still in progress. A suggested workflow can be used for the downscaling approach, however our applications to realistic synthetic cases are not complete yet and we have not shown here. The ideal workflow can be using a poro-elastic dataset and performing acoustic and/or elastic inversion to get macro-scale parameters, and then perform downscaling approach. In the case of poro-elastic dataset, the recovered V_p and/or V_s contain already an influence of porous parameters.

Downloaded 05/04/13 to 193.48.255.141. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Sensitivity analysis of time-lapse images obtained by differential waveform inversion with respect to reference model

Amir Asnaashari [†*], *Romain Brossier* [†], *Stéphane Garambois* [†], *François Audebert* [‡], *Pierre Thore* [‡], *and Jean Virieux*[†]
[†] *Institut des Sciences de la Terre (ISTerre), Université Joseph Fourier & CNRS*
[‡] *TOTAL E&P*

SUMMARY

For monitoring purposes, one of the promising techniques dedicated to assess physical properties changes in target regions is the differential waveform inversion, both in the acoustic and elastic cases. A central question of this technique regards the choice of the reference model. One solution could be the use of the reconstructed baseline image provided through the standard Full Waveform Inversion (FWI) procedure of initial data. However, how the accuracy of the baseline reconstructed image will affect the precision of further time-lapse images is of crucial importance. Here, we present a sensitivity analysis of time-lapse images obtained from differential inversion, with respect to various reference models. Density, P- and S-wave velocity changes could be converted into fluid property changes thanks to an empirical downscaling relationship. For accurate estimation of fluid parameter changes, the construction of highly resolved time-lapse images presenting acceptable errors is a key issue for the downscaling procedure. We illustrate on a specific synthetic example that the sensitivity analysis over the reference model variation provides linear convergence towards the time-lapse image obtained when using the exact baseline. An accurate baseline reconstruction is essential and could benefit from other data collected for monitoring purposes.

INTRODUCTION

FWI is a data fitting procedure aiming to develop high resolution quantitative images of the subsurface, through the extraction of the full information content of the seismic data (Tarantola, 1984). Beside the exploration application, the FWI method can be also used for monitoring applications, such as oil and gas reservoirs, steam injection, CO_2 sequestration, in order to obtain a quantitative image of changes of physical properties in target regions from successive seismic experiments.

The conventional difference method for time-lapse inversion needs to independently invert the two data sets (baseline and monitor sets) and to subtract the final derived monitor model from that of the baseline one in order to obtain a perturbation image of property changes (Plessix et al., 2010). This procedure might not be so robust because spurious features on both baseline and monitor images could potentially contaminate the differential model.

An alternative strategy consists in inverting only the differential data set to recover a differential image. The differential (double difference) method is widely used in geodesy and in seismology in order to improve earthquake source locations or to image receiver areas (Monteiller et al., 2005). This procedure has been proposed for time-lapse waveform inversion of acoustic data in frequency domain (Watanabe et al., 2004) and inversion of elastic data in time domain (Denli and Huang, 2009). The main advantage of the differential method compared to the difference approach is that the common noise between surveys can be rejected by data differentiation (Watanabe et al., 2004). The final time-lapse image is then more robust to noise contamination.

In this study, we investigate the influence of the reference model on the quality of the time-lapse V_p reconstruction by differential acoustic FWI. In particular, we address the required accuracy of the V_p reference model to delineate high resolution time-lapse image of steam injection in the acoustic version of the Dai et al. (1995) model.

STANDARD FULL WAVEFORM INVERSION

Standard FWI is an iterative optimization problem that is generally introduced as a linearized least-squares problem which attempts to minimize the residuals between the observed and the modeled wavefields (Tarantola, 1987). The inverse problem can be formulated in the frequency domain (Pratt and Worthington, 1990), and the associated objective function to be minimized is defined by

$$\mathscr{C} = \sum_{if=1}^{nf} \sum_{is=1}^{ns} \frac{1}{2} \Delta \mathbf{d}^{\dagger} \Delta \mathbf{d}, \tag{1}$$

where the data misfit error $\Delta \mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}_{calc}$ denotes the difference between the observed data $\mathbf{d}_{obs}$ and the modeled data $\mathbf{d}_{calc}$ computed in the model $\mathbf{m}^{(n)}$ at the iteration n of the inversion. Superscript † indicates the adjoint (transposed conjugate). The summations in equation (1) are performed over the ns sources and a group of nf simultaneously inverted frequencies. The synthetic data $\mathbf{d}_{calc}$ is obtained by applying a sampling operator $\mathbf{S}$ to the incident wavefield $\mathbf{u}$ resulting from the forward problem resolution $\mathbf{A}\mathbf{u} = \mathbf{s}$, where the impedance matrix $\mathbf{A}$ is the forward problem operator and $\mathbf{s}$ represents the source term.

The gradient $\mathscr{G}$ of the objective function with respect to the model parameters $\mathbf{m} = \{m_i\}_{i=1,N}$, where N denotes the number of unknowns, can be derived from the adjoint-state formulation using the back-propagation technique (Plessix, 2006). This formulation gives for the i^{th} component of the gradient $\mathscr{G}$:

$$\mathscr{G}_{m_i} = \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{u}^t \frac{\partial \mathbf{A}^t}{\partial m_i} \lambda^* \right\}, \tag{2}$$

where ^{t} and * denote the transpose and conjugate operators, respectively, and $\Re$ denotes the real part of a complex number. The gradient can be computed as a product between the incident wavefield $\mathbf{u}$ from the source, and the adjoint back-propagated

© 2011 SEG

Downloaded 05/04/13 to 193.48.255.141. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

wavefield λ^* which is computed as $\mathbf{A}\lambda^* = \mathbf{S}^t\Delta\mathbf{d}^*$, using residuals at receiver positions as a composite source. Therefore, for computing gradient, only two forward problems per shot are required. The radiation pattern of the scattering by the model parameter m_i is represented by the sparse matrix $\partial\mathbf{A}/\partial m_i$.

The gradient of the objective function is then used in an optimization algorithm to update the model vector with the perturbation vector $\delta\mathbf{m}^{(n)}$ through the expression

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \alpha^{(n)}\delta\mathbf{m}^{(n)}, \tag{3}$$

where the step length at iteration n is denoted by $\alpha^{(n)}$. In this study, a quasi-Newton L-BFGS (Nocedal, 1980) optimization scheme with line search is used in the FWI algorithm (Brossier, 2011).

DIFFERENTIAL WAVEFORM INVERSION

In differential method, instead of minimizing the difference between observed and computed data, we attempt to minimize the difference of the differential data between two sets of data (Watanabe et al., 2004; Denli and Huang, 2009), giving us the expression

$$\Delta\mathbf{d} = (\mathbf{d}_{obs_{monitor}} - \mathbf{d}_{obs_{baseline}}) - (\mathbf{d}_{calc_{monitor}} - \mathbf{d}_{calc_{baseline}}), \tag{4}$$

where $\mathbf{d}_{obs_{monitor}}$ and $\mathbf{d}_{obs_{baseline}}$ are the observed data from monitor and baseline surveys respectively, and $\mathbf{d}_{calc_{monitor}}$ and $\mathbf{d}_{calc_{baseline}}$ are the computed data for these experiments.

For the differential analysis, we first need the construction of a composite data set defined as

$$\mathbf{d}_{composite} = \mathbf{d}_{obs_{monitor}} - \mathbf{d}_{obs_{baseline}} + \mathbf{d}_{calc_{ref}}, \tag{5}$$

which is composed of (a) the time-lapse differential observed data $(\mathbf{d}_{obs_{monitor}} - \mathbf{d}_{obs_{baseline}})$ which should only represent the time-lapse changes of the two data sets and (b) the simulated data $\mathbf{d}_{calc_{ref}}$ computed using forward modeling in one arbitrary reference model. This reference model should explain as much as possible the $\mathbf{d}_{obs_{baseline}}$ data. This composite data set $\mathbf{d}_{composite}$ can be used as a new observed data set $\mathbf{d}_{obs}$ in equation (1), which is now equivalent to minimize the differential residual (4) with a standard FWI algorithm.

The differential approach requires that the acquisition surveys of the two experiments match or could be matched easily to perform the difference of the data sets. It means that this method is dedicated to permanent monitoring systems, where sensors are deployed once and where the repeatability of the source should be somehow fulfilled. The meaning of such source repeatability is a question we shall not address here.

The full procedure of differential approach can be summarized in the algorithm 1. Results of this algorithm depend on the reference model and one may question the relation between the accuracy in the time-lapse image and the selection of the reference model.

Algorithm 1 Preprocessing and inversion procedure with the differential algorithm

1: Creating the reference model $\mathbf{m}_{ref}$
2: $\mathbf{d}_{calc_{ref}}$ are computed in $\mathbf{m}_{ref}$
3: The composite data $\mathbf{d}_{composite}$ are computed (equation (5))
4: A standard FWI procedure is used to invert $\mathbf{d}_{composite}$ data from the $\mathbf{m}_{ref}$ model that gives $\mathbf{m}_{composite}$
5: The time-lapse model changes $\delta\mathbf{m}_{timelapse}$ can be computed from model $\delta\mathbf{m}_{timelapse} = \mathbf{m}_{composite} - \mathbf{m}_{ref}$

SENSITIVITY ANALYSIS

Extracting porosity and saturation for reservoir modeling is one task of the time-lapse processes. A downscaling procedure is essential to get fluid property changes from seismic parameters P- and S-wave time-lapse images. In order to get accurate changes in fluid properties, reconstruction of highly resolved and accurate differential images is crucial.

One of the conditions that has an important effect on final time-lapse image is the reference model that is used for differential inversion. As for full waveform inversion, the starting model is a crucial point to ensure convergence toward the global minimum. However, in differential FWI for reservoir monitoring, the inversion procedure focuses only on the scattered wavefield due to the time-lapse changes of the reservoir target that affects mostly the high frequency content of the data. Therefore, because of this lack of low frequency content in the time-lapse data, the selection of the starting model is crucial for the quality of the time-lapse image.

The reconstruction of the baseline model through the application of a standard FWI to the baseline data set will provide one possible starting model. Is this model enough for accurate differential time-lapse imaging? In order to address this point, the sensitivity of time-lapse image with respect to the reference model is studied. Several differential inversions are performed using the same differential data set $(\mathbf{d}_{obs_{monitor}} - \mathbf{d}_{obs_{baseline}})$, but using different reference models. We vary these starting models from the FWI baseline model to the exact baseline model using a linear interpolation at each point of the medium.

The quality control of final images needs an objective criterion. Let us define the medium perturbation as

$$\delta\mathbf{m}_{composite} = \mathbf{m}_{composite} - \mathbf{m}_{true-baseline}, \tag{6}$$

where $\mathbf{m}_{composite}$ is the model parameter obtained by differential inversion of composite data set. Then, the misfit value between $\delta\mathbf{m}_{composite}$ and true time-lapse model changes can be computed using the expression

$$\varepsilon = \sqrt{\sum_{i=1}^{N}\left(\delta m_{i_{composite}} - \delta m_{i_{true-timelapse}}\right)^2}, \tag{7}$$

where N denotes the number of model unknowns. In order to have a better comparison, we normalize this misfit value using this equation,

$$\gamma = \frac{\varepsilon - \varepsilon_{min}}{\varepsilon_{max} - \varepsilon_{min}} \times 100, \tag{8}$$

© 2011 SEG

Downloaded 05/04/13 to 193.48.255.141. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

where ε_{min} represents the minimum ε that can be achieved from differential inversion with the true baseline used as a reference model. The quantity ε_{max} indicates the maximum misfit value obtained from the FWI baseline model in our case. A second criterion is introduced for the analysis of the quality of the reference model through

$$\mu = \sqrt{\sum_{i=1}^{N} \left(m_{i_{ref}} - m_{i_{true-baseline}}\right)^2}, \quad (9)$$

where we normalize this model variation through the expression

$$\eta = \frac{\mu}{\mu_{max}} \times 100. \quad (10)$$

In equation (10), μ_{max} denotes the maximum misfit value obtained when considering the reconstructed baseline image.

The acceptable upper limit of the function η leads to time-lapse images with an acceptable accuracy for downscaling. For an estimation of this acceptable range, we have to consider the downscaling procedure for computing fluid properties through empirical relationship between macro-scale elastic parameters (like wave speed) and micro-scale properties related to fluid parameters. This range is highly dependent on the accuracy that we expect for computing fluid parameters, since the time-lapse model should be used to compute the changes in porosity, fluid saturation and pore pressure using this empirical relationship (Landrø, 2001).

Therefore, considering the required accuracy on fluid parameters like porosity, we can estimate the required accuracy on the macro-scale velocity time-lapse models. From the relation between time-lapse models accuracy / reference model accuracy, we can determine the required reference model for differential FWI procedure to ensure a robust work-flow from time-lapse composite data until microstructure characterization.

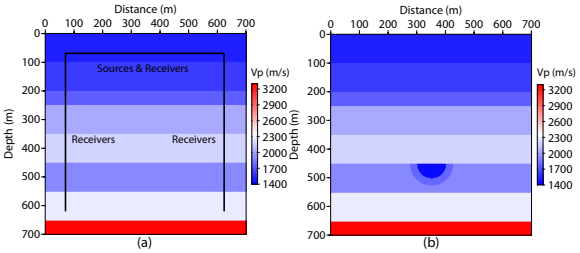


Figure 1: True V_p models: (a) the baseline, (b) the monitor.

SYNTHETIC APPLICATION

We assess the previous methodology on a synthetic example. We considered an acoustic synthetic version of the field example provided by Dai et al. (1995). The true P-wave velocity models for our acoustic time-lapse analysis are shown in Figure 1. Baseline model consists of eight layers; the sixth one is saturated with oil. After steam injection into this layer, the heat leads to two concentric areas at depth of 450 m, one saturated with steam and the outer one with heated oil.

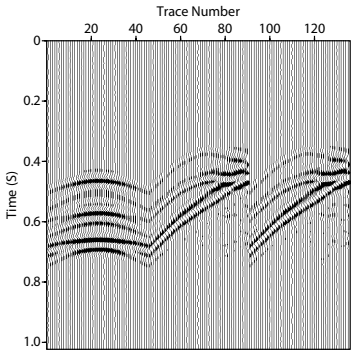


Figure 2: Differential seismogram of pressure data for the source located at the center $x = 350\ m$.

Our time-lapse configuration considers 23 sources, located along a horizontal line with interval of 25 m, below the surface at 75 m depth. We consider one line of receivers at the same depth as the sources and two vertical lines of receivers inside two boreholes at $x = 75\ m$ and $x = 625\ m$. An explosive Ricker source with a central frequency of 20 Hz is used for all shots. The time seismograms in baseline and monitor models are generated using a P_0 finite element Discontinuous Galerkin forward modeling (Brossier, 2011) and Perfectly-Matching-Layer absorbing boundaries condition all around the model. For all FWI tests, we use the same set-up and invert the data sets between $10-60$ Hz using sequential inverted frequencies with the inversion algorithm proposed by Brossier et al. (2009).

The baseline data set is first inverted through standard FWI to reconstruct the baseline image. Then, using linear interpolation between this reconstructed model and the true baseline model, three other models are created. Thus we ended up testing five reference models: the FWI baseline model (Figure 4.a), the true baseline model (Figure 4.e), and three intermediate models (Figure 4.b-d). These five reference models are used as an initial model for the differential inversion of differential data set (Figure 2) to study the sensitivity of time-lapse image. The normalized error γ (equation (8)) for time-lapse images is computed for each result, and these error values are plotted, in Figure 3, versus the accuracy of the reference model described by the parameter η.

The results and sensitivity curve show that the quality of the reference model directly drives the quality of the time-lapse image, what we expected. A linear relation also appears in Figure 3, meaning that the time-lapse imaging in this case can be related to a linear regime of the inversion. From downscaling constraints and Figure 3, we can determine the required accuracy of the reference model for micro-scale characterization.

In addition, it should be mentioned that reconstruction of the baseline model from standard FWI only could potentially not be sufficient for being used as a reference for differential inversion. However, for monitoring purposes, extra-information can be used like well data and geological constraints. These data should be used as a priori information in the framework of constrained FWI in order to build high resolution baseline models

© 2011 SEG

for time-lapse imaging.

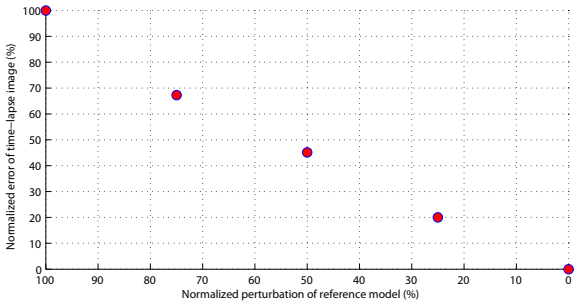


Figure 3: Sensitivity curve. Vertical axis indicates the γ value and horizontal one denotes the value of η. $\eta = 100\%$ and $\eta = 0\%$ represent the cases i) when the reconstructed baseline image through standard FWI, and ii) when the true baseline are respectively used as reference models.

CONCLUSIONS

We have studied the sensitivity of P-wave time-lapse imaging in terms of reference model accuracy for differential waveform inversion. Through a synthetic test, we have illustrated the relation between reference model accuracy and time-lapse image accuracy. This relation can be used to determine the required accuracy on the reference model in the framework of downscaling procedure for fluid characterization in reservoirs. This accuracy of reference model in differential approach is much more important than in standard full waveform inversion, since the inversion focuses on time-lapse response that affect mostly the high frequency content of the data. We propose that this level of accuracy of the reference model cannot be reached from standard FWI of recorded seismic baseline data only, but should benefit from available a priori information like well data and geological constraints. It will allow better building the accurate reference model from constrained FWI.

Perspectives of the work will focus on visco-elastic differential FWI and target oriented imaging on synthetic and real data.

ACKNOWLEDGMENTS

This study has been funded by TOTAL Exploration & Production. This work was performed using HPC resources from GENCI-CINES (Grant 2010-046091).

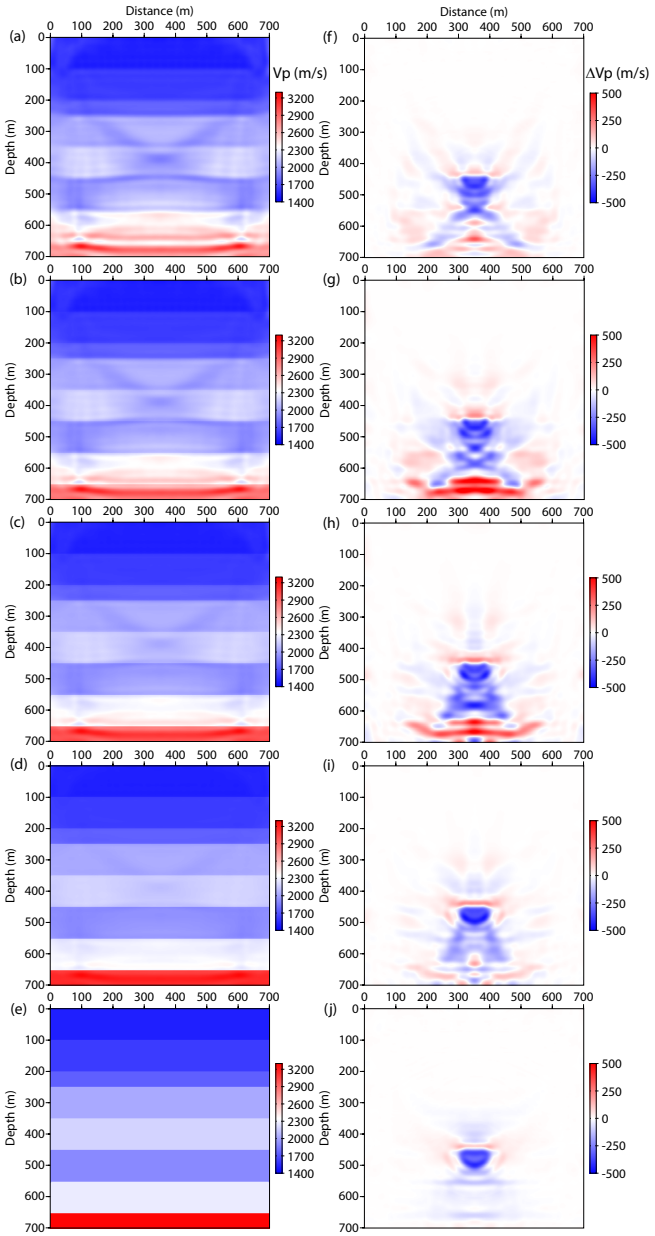


Figure 4: Reference and time-lapse V_p models: left panel shows the evolution of reference models from (a) reconstructed baseline image to (e) the true baseline model, and right panel illustrates each time-lapse image corresponding to the reference model at the left part.

Downloaded 05/04/13 to 193.48.255.141. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

© 2011 SEG

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2011 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Brossier, R., 2011, Two-dimensional frequency-domain viscoelastic full waveform inversion: Parallel algorithms, optimization and performance: Computers & Geosciences, **37**, no. 4, 444–455, doi:10.1016/j.cageo.2010.09.013.

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: Geophysics, **74**, no. 6, WCC105–WCC118, doi:10.1190/1.3215771.

Dai, N., A. Vafidis, and E. Kanasewich, 1995, Wave propagation in heteregeneous porous media: A velocity-stress, finite-difference method: Geophysics, **60**, 327–340, doi:10.1190/1.1443769.

Denli, H., and L. Huang, 2009, Double-difference elastic waveform tomography in the time domain: 79th Annual International Meeting, SEG, Expanded Abstracts, 2302–2306.

Landrø, M., 2001, Discrimination between pressure and fluid saturation changes from time-lapse seismic data: Geophysics, **66**, 836–844, doi:10.1190/1.1444973.

Monteiller, V., J.-L. Got, J. Virieux, and P. Okubo, 2005, An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano: Journal of Geophysical Research, **110**, B12, B12306, doi:10.1029/2004JB003466.

Nocedal, J., 1980, Updating quasi-Newton matrices with limited storage: Mathematics of Computation, **35**, no. 151, 773–782, doi:10.1090/S0025-5718-1980-0572855-7.

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**, no. 2, 495–503, doi:10.1111/j.1365-246X.2006.02978.x.

Plessix, R.-E., S. Michelet, H. Rynja, H. Kuehl, C. Perkins, J. W. de Maag, and P. Hatchell, 2010, Some 3D applications of full-waveform inversion: Presented at the 72nd Annual International Conference and Exhibition, EAGE.

Pratt, R. G., and M. H. Worthington, 1990, Inverse theory applied to multisource crosshole tomography, Part I: Acoustic wave-equation method: Geophysical Prospecting, **38**, no. 3, 287–310, doi:10.1111/j.1365-2478.1990.tb01846.x.

Tarantola, A., 1984, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, no. 6, 998–1015, doi:10.1111/j.1365-2478.1984.tb00751.x.

———, 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier.

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: 74th Annual International Meeting, SEG, Expanded Abstracts, 2323–2326.

Downloaded 05/04/13 to 193.48.255.141. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Inversion of poroelastic parameters using a downscaling method

Bastien Dupuy*, Amir Asnaashari†, Stéphane Garambois† and Jean Virieux†

**NTNU, Trondheim, Norway. E-mail: bastien.dupuy@ntnu.no*

†*ISTERRE, Université Joseph Fourier, CNRS, Grenoble, France.*

(June 20, 2013)

Running head: **Downscaling poroelastic properties**

ABSTRACT

We present a two-steps method to invert poroelastic parameters using a first step of seismic imaging that allows to obtain visco-elastic effective quantitative properties. From these visco-elastic parameters, we develop a semi-global optimization method to estimate microscale properties (porosity, mechanical moduli, fluid phase properties, saturation). The sensitivity studies show that we need to keep a well-determined system to obtain accurate estimation. Making appropriate assumptions on fluid and/or solid phases parameters, we can then invert pretty well the porosity and mechanical moduli, the fluid properties or the saturation from P- and S-wave velocities and quality factors. We suggest a workflow for downscaling the time-lapse parameter variation: after a first step of skeleton characterization on the baseline model using P- and S-wave velocity results obtained by FWI of poroelastic data, we can invert from the time-lapse P-wave and S-wave velocity, the fluid properties and saturation changes. Therefore, it is possible to downscale the effective properties obtained by classic seismic imaging methods to obtain high resolution images of microscale properties.

INTRODUCTION

The interpretation of the recorded seismograms has been constantly improved in the last decades. High resolution seismic reflection methods are useful for subsurface structure characterization but quantitative methods as arrival time tomography are essential to obtain quantitative acoustic, elastic or visco-elastic properties. Full-waveform inversion (FWI) methods are the last improvement of these seismic imaging methods and can provide high-resolution images of these macroscale effective elastic properties (basically, P- and S-waves velocities). The aim of this work is to propose to go until the microscale in the seismic imaging, by downscaling the effective medium properties found by FWI to obtain microscale poroelastic parameters. We use a two-step inverse approach: from macroscale parameters (velocities and quality factors) determined by a first inversion step, we recover microscale parameters in a second step by a downscaling method. We move from a macroscopic scale to a microscopic scale for the description of the medium.

Numerous models aim to link P- and S-wave velocities (V_P, V_S) to porosity (ϕ). Mainly, when porosity increases, P- or S-wave velocity decreases, pressure and shear moduli decreases when the pore volume increases. Empirical relations make a link between V_P and ϕ, however they are dependent on the rock type because they are based on experimental measurements. For example, Han et al. (1986) have established linear relations between V_P, V_S and the porosity and the clay content for clastic sediments during their full diagenetic evolution (from unconsolidated sands to sandstones). Other empirical relations exist, as Wyllie et al. (1956), Raymer et al. (1980) or Raiga-Clemenceau et al. (1988). However, in these relations, P-wave velocity is only linked to the porosity while, in real media, the drained medium (skeleton) rigidity plays an important role as well. This skeleton rigidity is

depending also on the mineral grains rigidity and the grains arrangement (geometry, compaction...). Moreover, we need to consider the rigidity of fluids inside pores. On the other hand, numerous empirical relations between V_P and V_S have been established for various rocks (as an example, summarized by Castagna et al. (1993) for limestones, sandstones, shales and dolomites).

Several authors address the approaches to invert poroelastic parameters from macroscale data deduced from seismic signals (velocities, attenuations, AVO...). Berryman et al. (2002) make a relationship between the propagation velocities V_P and V_S to the porosity and the water saturation for several rocks based upon experimental data. Tang and Cheng (1996) settle all other parameters and deduce the permeability from the Stoneley waves generated in boreholes. Bosch (1999) introduces a stochastic approach (density probability functions using Monte-Carlo methods) to invert jointly geophysical data (seismic, electromagnetic and gravimetric data). This kind of approach is also used by Chotiros (2002) on seismic data for saturated sands. He shows that the interpretation of data by visco-elastic models is impossible due to energy losses of waves reflections. Then, he uses a poroelastic modelization for his inversion to deduce the drained bulk modulus K_D and the drained shear modulus G_D from P- and S-wave velocities and attenuations (V_P, V_S, Q_P and Q_S). He also puts forwards some improvements introducing more complex models (composite materials or porosity variation with fluid pressure).

In the same manner, Bachrach (2006) inverts seismic data to deduce porosity and saturation images. Gunning and Glinsky (2007) try also to invert grains geometrical characteristics as well as porosity to deduce the medium permeability using variable offset AVO data. As well, van Dalen et al. (2010) show that it is possible to invert simultaneously the permeability and the porosity using informations contained in the frequency and angle

dependencies of reflection coefficients (PP and PS). Rubino and Velis (2009) describe a pre-stack spectral inversion method using thin layers AVA data to estimate effective properties of partially saturated media and Rubino and Velis (2011) link these effective properties to a saturation estimation.

In this paper, we are followed by Gassmann (1951) relations which allow to compute the variations of the bulk modulus of the drained medium (and then, the variations of V_P) during a fluid substitution in the porous medium. For that, we need to know the fluid properties (bulk modulus, density and viscosity) (there are summaries for classical fluids in Batzle and Wang (1992) or Mavko et al. (2009)). However, the Gassmann relations require the following assumptions:

- the moduli are computed at low frequency. To have frequency dependent moduli, we can consider complex media (double porosity, patchy saturation, see Pride et al. (2004)) to have a broad frequency band.

- the medium is assumed to be isotropic. Brown and Korringa (1975) have extended the relations to poroelastic anisotropic media, but we have to know properly the anisotropy parameters which is difficult even with well-constrained laboratory experiments.

- the skeleton is constituted by identical grains. To have various mineralogical compositions, we can compute the effective skeleton moduli by averages ((Reuss-Voigt or Hashin-Shtrickman means).

- the pores are saturated with one fluid phase. Domenico (1976) and Brie et al. (1995) for instance, use simple averages, partly empirical to compute the properties of the

effective fluid. However, the patchy saturation theory (White, 1975; Dutta and Odé, 1979) allows to take into account the spatial repartition of each fluid phase, which can have a strong influence on seismic waveforms dispersion and attenuation.

Gassmann relations are the basis for the downscaling of poroelastic parameters. In case of complex porous media (double porosity, partially saturated, visco-poroelastic...), these Gassmann equations can be generalized using frequency-dependent moduli (see Pride et al. (2004); Dupuy et al. (2012)). Inverted macroscale parameters are classically P- and S-wave velocities, extracted from seismograms by arrival times tomography. By developing FWI techniques, we can hope to obtain a better estimation for quality factors Q_P and Q_S (Malinowski et al., 2011). The inversion of other parameters sets could be investigated, as impedances (acoustic, elastic...), the Lame parameters or parameters coming from AVO (reflection and transmission coefficients, intercepts, gradients...). Moreover, the use of S-waves is crucial for several reasons:

- it allows to have more data which is important because the inversion system is highly under-determined,
- the effects of "non-fluid" parameters (porosity, compaction, clay content...) on V_P and V_S are similar, but only the fluid saturation has different effects on V_P and V_S (Berryman et al., 2002; Avseth et al., 2005).

After the description of the forward models, we explain the inverse theory and the semi-global optimization algorithm used to solve the inverse problem. Then, sensitivity studies describe first which parameters can be inverted from which data. Then, we apply this downscaling step to synthetic realistic cases. After the computation of seismograms

of poroelastic wave propagation in the reference medium (described by microscale porous parameters), an inversion step allows to recover macroscale parameters by FWI. Finally, the last step consists to recover microscale parameters by inversion of macroscale parameters. This last inversion step is based on global optimization algorithms.

FORWARD MODELING: UPSCALING APPROACHES

Porous media homogenization (forward model)

The description of a porous media requires an homogenization approach of both fluid and solid phases in order to deduce an equivalent medium (Burridge and Vargas, 1979). The porosity $\phi = V_V/V_T$, which is the ratio between void and total volumes, will define respective proportions of fluid and solid phases.

The fluid and associated flows through the solid matrix, are described by an bulk modulus K_f, a density ρ_f and a viscosity η. The viscosity can be formulated with the intrinsic permeability k_0 introduced in the Darcy law. Auriault et al. (1985) and Johnson et al. (1987) have generalized the Darcy law with a dynamic permeability $k(\omega)$ depending on the pulsation ω. This permeability, which is a complex number, has a significant frequency dependence as defined through the dispersive relation

$$k(\omega) = \frac{k_0}{\sqrt{1 - \frac{1}{2}\imath\frac{\omega}{\omega_c}} - \imath\frac{\omega}{\omega_c}}. \tag{1}$$

The cut-off pulsation ω_c allows separating the low frequency domain where viscous effects are dominant from the high frequency one where inertial effects prevail. Using Archie law,

the pulsation ω_c is defined as

$$\omega_c = \frac{\eta}{\rho_f k_0 \phi^{-m}}, \tag{2}$$

where the cementation exponent m is related to the electrical formation factor and to the pore tortuosities (Brown, 1980). Then, we can introduce the flow resistance density term $\tilde{\rho}(\omega)$ which describes the dynamic loss of energy due to the fluid flow with an explicit frequency dependence. The term is responsible for the intrinsic scattering of waves in the Biot poroelasticity theory (Biot, 1956) and it is expressed as

$$\tilde{\rho}(\omega) = \frac{\imath\eta}{\omega\ k(\omega)}. \tag{3}$$

To consider partially saturated media (several fluid phases), the simplest approach consists to build an effective fluid phase (computed by harmonic and arithmetic averages) to apply the Biot-Gassmann theory for a saturated medium. Domenico (1976) or Berryman et al. (2000) propose to compute equivalent parameters by average weighted on volume fraction of the fluid phase i. For densities, we use an arithmetic average (Voigt, 1889) as

$$\rho_f \;=\; \sum_i V_i\ \rho_{f_i}\ . \tag{4}$$

For the bulk fluid moduli, the arithmetic (Voigt, 1889) and harmonic (Reuss, 1929) averages are the low and high limits, as

$$\begin{aligned} K_f &= \sum_i V_i\ K_{f_i}\ , \\ \frac{1}{K_f} &= \sum_i \frac{V_i}{K_{f_i}}\ . \end{aligned} \tag{5}$$

Classically, the harmonic average is used (Hill, 1952), but for liquid/gas mixture, as an example, there is four orders in magnitude between K_f for water and for air ($10^5\ Pa$ for air and $10^9\ Pa$ for water). For a water saturation less than 99.9 %, harmonic averages give underestimated results. Brie et al. (1995) formulate the effective K_f as

$$K_f \quad = \quad (K_{f_l} - K_{f_g})\ V_l^e + K_{f_g}\ , \tag{6}$$

where the subscript l is for the liquid phase and g for the gas phase. The exponent e can be taken equal to 5 to have a good agreement with experimental results of Johnson (2001) (Carcione et al., 2006).

In the same way, for a fluid phase constituted by a gas phase (viscosity η_g and volume fraction V_g) and a liquid phase (viscosity η_l and volume fraction $1-V_g$), Teja and Rice (1981) give a formula (also used by Carcione et al. (2006)) to compute the effective viscosity η as

$$\eta = \eta_g \left(\frac{\eta_l}{\eta_g}\right)^{(1-V_g)} \ . \tag{7}$$

The solid skeleton is entirely described by the association of grains (defined by an bulk modulus K_s, a shear solid modulus G_s and a solid density ρ_s) in a solid frame. A lot of effective medium theories can be used to compute the effect mechanical moduli K_D and G_D (Berryman, 1995; Mavko et al., 2009). In consolidated materials (typically, rocks as sandstones, limestones, shales...), Pride (2005) has shown that the use of relations including

a general consolidation parameter cs matches well the real data:

$$\begin{aligned} K_D &= K_s \frac{1-\phi}{1+cs\ \phi}, \\ G_D &= G_s \frac{1-\phi}{1+\frac{3}{2}\ cs\ \phi} \end{aligned} \tag{8}$$

For unconsolidated materials (near surface sediments as sands or clays), the skeleton mechanical behavior is mainly controlled by the grains contacts. Various models derived from Walton (1987) allow to compute K_D and G_D in a sphere packing assembly with respect to the effective stress and a compliance parameters (Pride, 2005). In our downscaling process, we test two parametrizations: inversion of consolidation parameter cs or inversion of effective moduli K_D and G_D.

The density of the porous medium is the arithmetic mean of fluid and solid phases weighted by their own volumes via the porosity, such that

$$\rho = (1-\phi)\ \rho_s + \phi\ \rho_f. \tag{9}$$

The introduction of the undrained bulk modulus K_U, the Biot C modulus and the fluid storage coefficient M allows the explicit description of the homogenized porous medium through the Gassmann relations (Gassmann, 1951). Relationships between coefficients K_U, C and M and the modulus functions of K_D, K_s, K_f and ϕ are given by

$$\begin{aligned} K_U &= \frac{\phi\ K_D + (1-(1+\phi)\ K_D/K_s)\ K_f}{\phi\ (1+\Delta)}, \\ C &= \frac{(1-K_D/K_s)\ K_f}{\phi\ (1+\Delta)}, \\ M &= \frac{K_f}{\phi\ (1+\Delta)}, \end{aligned} \tag{10}$$

with the additional expression

$$\Delta = \frac{1-\phi}{\phi}\frac{K_f}{K_s}\left(1 - \frac{K_D}{(1-\phi)\ K_s}\right). \tag{11}$$

The shear modulus of the porous medium G is independent of the fluid characteristics and, therefore, equal to the shear modulus of the drained solid skeleton through the relation 8 where only the porosity ϕ and the consolidation parameter cs are involved.

Effective viscoelastic properties (velocities and quality factors)

The Biot poroelastodynamic theory (Biot, 1956) (formulated in frequency by Pride et al. (1992)) predicts three wave types: a compressional and a shear waves similar to those propagating inside an elastic body and a slow compressional wave, called Biot wave, which is strongly diffusive and attenuated at low frequencies. This Biot wave behaves as either a diffusive signal or a propagative wave depending on the frequency content of the source with respect to the cut-off pulsation or characteristic frequency, as defined in equation 2. The slowness of the shear wave is given by the following equation (Pride, 2005)

$$s_S^2 = \frac{\rho - \rho_f^2/\tilde{\rho}}{G}, \tag{12}$$

while slownesses of compressional waves, the P and Biot waves, are given by

$$s_P^2 = \gamma - \sqrt{\gamma^2 - \frac{4(\rho\tilde{\rho} - \rho_f^2)}{HM - C^2}} \quad , \quad s_{Biot}^2 = \gamma + \sqrt{\gamma^2 - \frac{4(\rho\tilde{\rho} - \rho_f^2)}{HM - C^2}} \tag{13}$$

where γ and H terms are

$$\gamma = \frac{\rho M + \tilde{\rho} H - 2\rho_f C}{HM - C^2} \ , \qquad H = K_U + \frac{4}{3} G \ . \tag{14}$$

Then, we can deduce the effective visco-elastic velocities and quality factors for P-, S- and Biot waves as:

$$\begin{aligned} V_{P,Biot,S}(\omega) &= \frac{1}{\mathcal{R}e(s_{P,Biot,S}(\omega))} \ . \\ Q_{P,Biot,S}(\omega) &= \frac{1}{2} \frac{\mathcal{R}e\ (s_{P,Biot,S}(\omega))}{\mathcal{I}m\ (s_{P,Biot,S}(\omega))} \ . \end{aligned} \tag{15}$$

The visco-elastic velocities (V_P and V_S) and quality factors (Q_P and Q_S), as well as the mean density ρ, are the data of our inverse problem and are classically estimated by seismic imaging methods. These effective properties (or seismic attributes) describe the behavior of the medium at the macroscale (wavelength).

SEMI-GLOBAL OPTIMISATION METHOD

Seismic imaging methods by FWI or first arrival-time tomography are based upon linearized local optimization methods which use local properties of the misfit function to determine the perturbation direction (see Menke (1984); Tarantola (1987); Brossier (2009) for complete theory). These methods run smoothly to find a local minimum in the close neighboring of the starting model. The global optimization methods look for global minimum of the misfit function on the whole model domain, avoiding the convergence towards a local minimum. These methods consist in the exploration of the whole model space (inverted parameters)

only limited by minimal and maximal values, to find the lowest misfit function. We use an oriented Monte-Carlo method (semi-global explorations) which is based upon random exploration but guided towards best models in order to reduce computational costs. The Neighborhood Algorithm (NA) is introduced by Sambridge (1999a,b). The main advantage of this method (compared to simulated annealing or genetic algorithms methods) is that it needs only two control parameters.

The inverse problem consists in the extraction from data (seismic attributes deduced from seismograms) of models (poroelastic parameters) and is formulated as

$$\boldsymbol{d} = g(\boldsymbol{m}) \ , \tag{16}$$

where $\boldsymbol{d}$ is the data vector, $\boldsymbol{m}$ is the model vector and g is a non-linear function linking models and data.

In our approach, this function g contains the analytical Biot-Gassmann relations which allow to compute P-, S- and Biot velocities and attenuations and the mean density with respect to the microscale porous parameters. This function is non-linear and the inverse of g can not be computed. The solving of the system $\boldsymbol{m} = g^{-1}(\boldsymbol{d})$ has to be done by optimization methods. The system linearization can be done by local optimization methods which require to have a good a priori starting model, close to the true model to avoid local minima. This method is the only one for big problems with a high number of unknowns and/or a large computational time for the direct problem (for instance, the solving of PDE for FWI). In our case, the direct model computation is very fast (computation of tens of analytical relations) and the number of parameters/models is low.

The optimization aims to minimize a scalar function (misfit function) describing the

discrepancy between observed data d_{obs} and calculated data $g(\boldsymbol{m})$ (by forward modeling). We use a L_2 norm to compute the misfit as

$$misfit = \left[(\boldsymbol{d_{obs}} - g(\boldsymbol{m}))^T \ (\boldsymbol{d_{obs}} - g(\boldsymbol{m})) \right]^{1/2} \ . \tag{17}$$

For the inversion, we use the neighboring algorithm of Sambridge (1999a). The algorithm principle consists in dividing the model space in Voronoi cells to focus the resampling with respect to the lower misfit function values. For each iteration, new cells are defined with respect to previous sample misfit value. The resample is then focused in the lower misfit areas. In the examples given in this paper, we compute 1000 iterations and we choose a resample factor equal to 10, so 10.000 models are generated by each inversion.

After each sampling step, Sambridge (1999b) propose a probabilistic estimation step for the resolution of the models obtained in the first step. By the definition of density probability functions, we can quantify the errors and the resolution obtained by the first sampling step. This bayesian analysis step is not used in this paper.

SENSITIVITY ANALYSIS

We determine the sensitivity of poroelastic parameters with respect to various classes of macroscale data parametrizations: only P-wave propagation velocity (V_P), P- and S-waves propagation velocities (V_P, V_S), P- and S-waves propagation velocities and mean density (V_P, V_S, ρ), P-wave propagation velocity and quality factor and mean density (V_P, Q_P, ρ), P- and S-waves propagation velocities and quality factors (V_P, Q_P, V_S, Q_S) and P- and S-waves propagation velocities and quality factors and mean density (V_P, Q_P, V_S, Q_S, ρ).

Various parametrizations are studied for saturated and unsaturated media:

- saturated media, inversion of all parameters of Biot-Gassmann model: fluid phase (K_f, η and ρ_f), solid phase (K_s, G_s and ρ_s) and skeleton (m, ϕ and cs),
- saturated media, inversion of skeleton parameters: porosity ϕ and consolidation parameter cs or bulk and shear effective moduli K_D and G_D,
- saturated media, inversion of fluid phase parameters: bulk modulus K_f, viscosity η and density ρ_f,
- partially saturated media, inversion of water saturation V_1 and skeleton parameters (ϕ and cs).

Inversion of all poroelastic parameters (saturated media)

We consider a consolidated medium saturated with water. The microscale parameters which describe the medium and the macroscale parameters computed by the Biot-Gassmann relations are given in table 1. We invert all the parameters of the solid phase (K_s, G_s and ρ_s), of the fluid phase (K_f, η and ρ_f) and of the skeleton (m, cs and ϕ) from V_P, V_S, Q_P, Q_S and ρ data (figure 1). Due to the high under-determination of the system (9 parameters and 1 to 5 data), we are not showing the inversion results from less data.

In figure 1, we give 2D sections of the model space (G_s with K_s, η with K_f, ρ_s with ρ_f, cs with ϕ and m with ϕ). The color of each dot (representing each computed model) corresponds to the misfit value. More the number of computed models is high, more we converge towards the one or several best values thanks to the optimization process that guides the resampling of medium by Voronoi cells. We can then interpret the misfit function

shape for pairs of inverted parameters.

Microscale parameters	K_s	(GPa)	40
	G_s	(GPa)	10
	ρ_s	(kg/m^3)	2700
	K_f	(GPa)	2.2
	ρ_f	(kg/m^3)	1000
	η	$(Pa.s)$	0.001
	m		1
	ϕ		0.4
	k_0	(m^2)	10^{-11}
	cs		5
Macroscale parameters	V_P	(m/s)	2570
	V_S	(m/s)	862
	Q_P		1190
	Q_S		161
	ρ	(kg/m^3)	2010

Table 1: Microscale and macroscale parameters for the saturated medium. Velocities and quality factor are computed at 200 Hz.

The bulk and shear moduli of grains K_s and G_s are roughly well estimated. Even if the whole model space is investigated, there is only one main minimum, located in the true value range. The sensitivities of the fluid bulk modulus K_f and the fluid viscosity η are very low, with no clear minimum in the misfit function. The densities, fluid ρ_f and solid ρ_s, are more or less well recovered, but the misfit function shows two local minima, even if the lowest misfit value is located in the right minimum. The skeleton properties, cs and ϕ, are well estimated, with only one minimum in the model space. However, the cementation factor m is not recovered. It is important to note that we use all the data available for this test, the results obtained with less data are not shown here but are strongly worst.

The inversion of all the poroelastic parameters is then possible when we have no a priori information about the medium characteristics. Nevertheless, some parameters are

badly estimated and mostly, to invert correctly some of the parameters, we need to use all available data. In most classic real cases, we have a priori knowledge of several parameters (especially for grains and fluid phase), so we can have more targeted inversion on a selection of parameters.

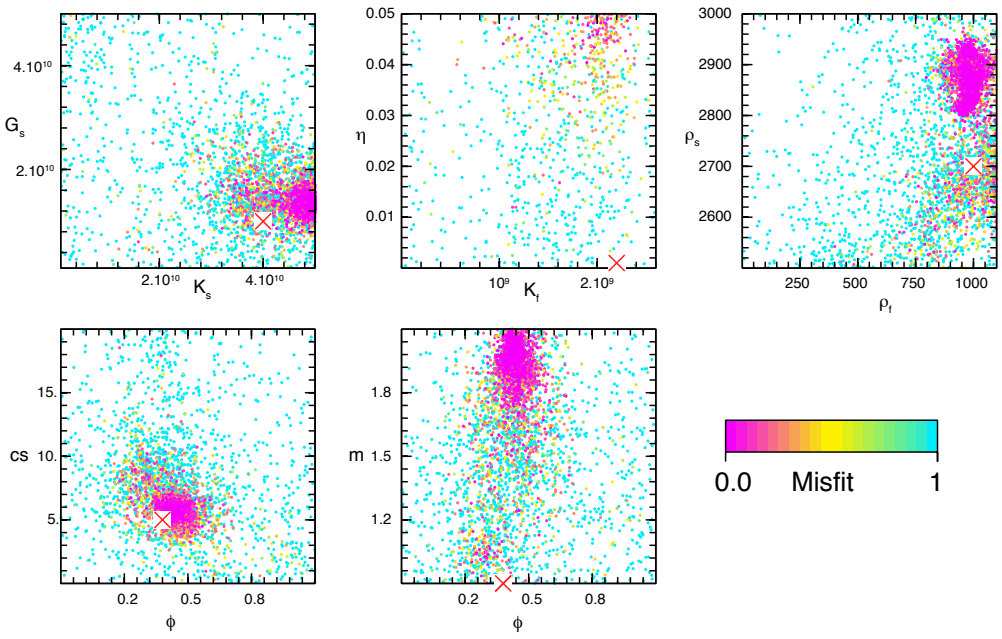


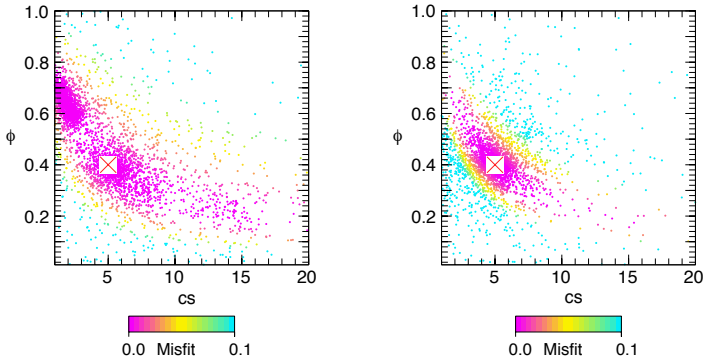
Figure 1: Inversion of poroelastic parameters (grains shear modulus G_s (Pa), grains bulk modulus K_s (Pa), fluid viscosity η $(Pa.s)$, fluid bulk modulus K_f (Pa), grains density ρ_s (kg/m^3), fluid density ρ_f (kg/m^3), consolidation parameter cs, cementation factor m and porosity ϕ from V_P, V_S, Q_P, Q_S and ρ data. The true model is represented by the red cross. Each plot shows the values of computed models in the parameter space. The dot color depends on the misfit value.

Skeleton parameters inversion in saturated media

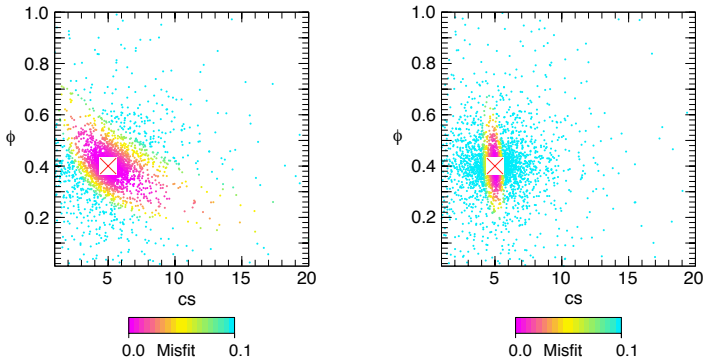
In this second sensitivity test, we assume that the fluid and solid phases parameters are known as prior information (see table 1) and we invert only the porosity ϕ and the skeleton properties (consolidation parameter cs or bulk and shear effective moduli K_D and G_D). We give the results as 2D sections of the model space for the parametrization (ϕ,cs) in figure 2 and for the parametrization (ϕ,K_D,G_D) in figure 3.

It is quite simple to analyze the results: as soon as the inversion system is no more under-determined, the skeleton properties are very well recovered (less than 1% of error). That means that ϕ and cs can be well estimated with V_P and V_S data (figure 2(b)) but not with only V_P (figure 2(a), about 50% of error). We clearly see on the 2D sections of the model space that the area of low misfit is considerably larger on figure 2(a) than on figure 2(b), with several local minima while the low misfit zone is unique and become more and more reduced when we add more data (figures 2(c), 2(d), 2(e) and 2(f)), it means the inversion becomes more and more well-posed. To assess the robustness of the inversion, several tests have been done using erroneous a priori data. When we consider an error of 5% on a prior parameter (K_f or K_s), the discrepancy between true and estimated values increases and we need more data to obtain a good estimation.

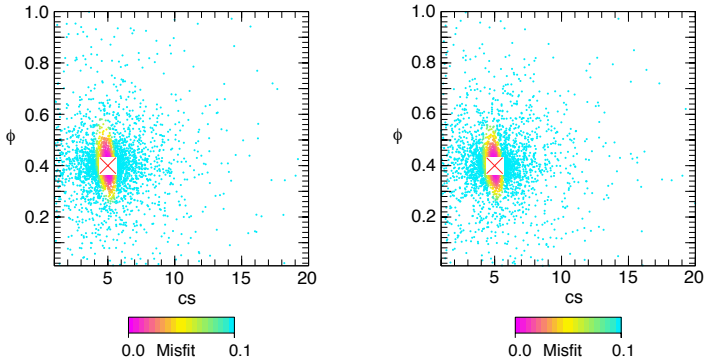
For the (ϕ,K_D,G_D) parametrization (figure 3), we observe similar patterns with a good estimation of the three parameters as soon as we have three types of input data (figures 3(c), 3(d), 3(e) and 3(f)). When we use only V_P data (figure 3(a)), the local minima are numerous and the parameters are not well estimated. Using V_P and V_S data (figure 3(b)), we reduce the low misfit zone but we still have between 20% and 60% of discrepancy between true and estimated values.



(a) $Data = (V_P)$ (b) $Data = (V_P, V_S)$

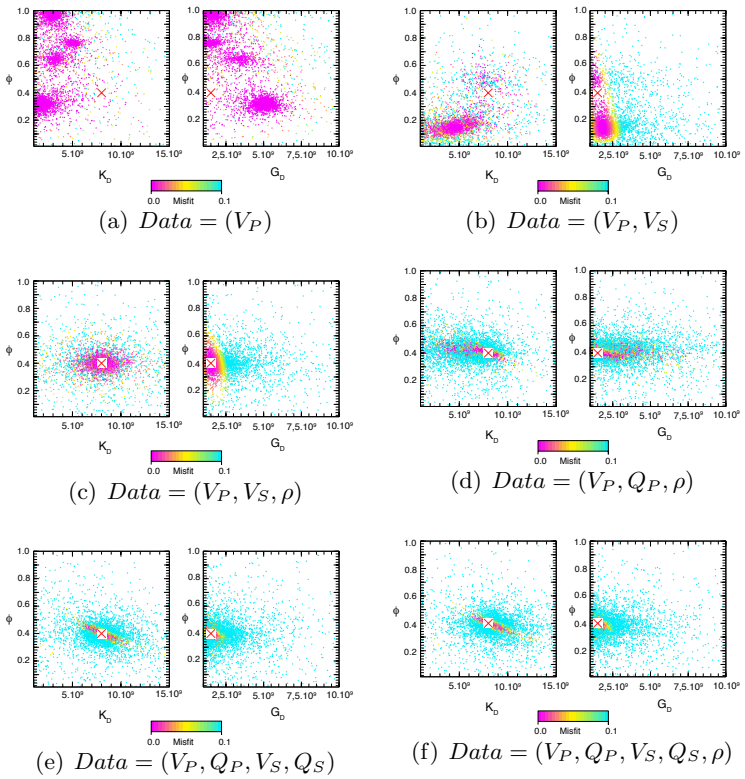


(c) $Data = (V_P, V_S, \rho)$ (d) $Data = (V_P, Q_P, \rho)$



(e) $Data = (V_P, Q_P, V_S, Q_S)$ (f) $Data = (V_P, Q_P, V_S, Q_S, \rho)$

Figure 2: Inversion of skeleton parameters (porosity ϕ and consolidation parameter cs) from (a) V_P, (b) V_P, V_S, (c) V_P, V_S, ρ, (d) V_P, Q_P, ρ, (e) V_P, Q_P, V_S, Q_S et (f) V_P, Q_P, V_S, Q_S, ρ. The true model is represented by the red cross. Each plot shows the values of computed models in the parameter space. The dot color depends on the misfit value.



(a) $Data = (V_P)$ (b) $Data = (V_P, V_S)$

(c) $Data = (V_P, V_S, \rho)$ (d) $Data = (V_P, Q_P, \rho)$

(e) $Data = (V_P, Q_P, V_S, Q_S)$ (f) $Data = (V_P, Q_P, V_S, Q_S, \rho)$

Figure 3: Inversion of skeleton parameters (porosity ϕ, bulk drained modulus K_D and shear drained modulus G_D) from (a) V_P, (b) V_P, V_S, (c) V_P, V_S, ρ, (d) V_P, Q_P, ρ, (e) V_P, Q_P, V_S, Q_S et (f) V_P, Q_P, V_S, Q_S, ρ. The true model is represented by the red cross. Each plot shows the values of computed models in the parameter space. The dot color depends on the misfit value.

Fluid phase properties inversion (saturated media)

For time-lapse applications in oil, gas or water reservoirs (pumping or storing), it can be interesting to know which fluid is saturating the rock. As an example, when we inject steam at high temperature in an oil reservoir, the geometrical extension of oil, heated oil and steam in the geological layer can be determined by the inversion of macroscale data to recover the fluid parameters: K_f, η et ρ_f. Here, we consider consolidated sands (Dai et al., 1995) where the steam is injected and we want to determine which fluid is saturating the reservoir layer. The physical parameters are given in table 2 for several fluids filling up the porous medium (oil, heated oil, steam or water).

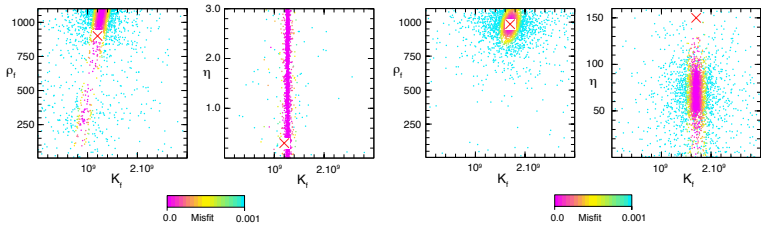
Figure 4 shows 2D sections of the model space ((ρ_f, K_f) and (η, K_f)) for each fluid inverted from V_P data or from V_P, V_S and ρ data. Oil, heated oil and water have properties in the same order of magnitude and show similar results. Using only V_P (figures 4(a), 4(c) and 4(e)), only the bulk modulus K_f is well estimated, ρ_f and η has a large spreading of the low misfit zone. Adding new data (figures 4(b), 4(d) and 4(f)) allow to better constrain the inversion and to recover the fluid density. Nevertheless, the fluid viscosity is never well estimated, probably due to its low sensitivity (even using additional Q_P and Q_S data, not shown here).

When the filling fluid is steam, we have 3 orders of magnitude between moduli for oil and steam, 2 orders between densities, 7 orders between viscosities. The inversion is then really worst for steam properties, the low misfit zone is broader whatever the data and only the density is well estimated (lowest difference in value) using V_P, V_S and ρ data (figure 4(h)). However, even if several fluid parameters are difficult to invert (mainly the viscosity), we can even though determine the kind of fluid using the information obtained by the K_f

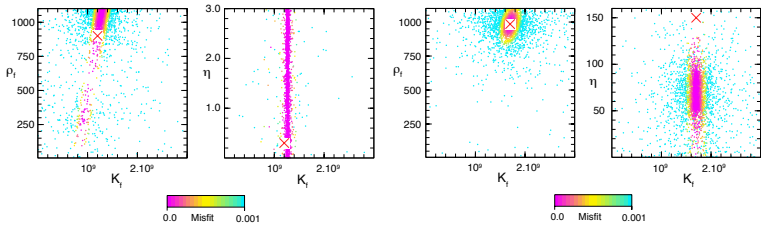
and ρ_f estimation from V_P and V_S velocities and from the density ρ, in order to determine the nature of the saturating fluid.

Microscale parameters	K_s	(GPa)	37			
	G_s	(GPa)	4.4			
	ρ_s	(kg/m^3)	2650			
	m		1.5			
	ϕ		0.33			
	k_0	(m^2)	10^{-12}			
	cs		20			
			Oil	Heated oil	Steam	Water
	K_f	(GPa)	1.7	1.2	$1.4\ 10^{-3}$	2.5
	ρ_f	(kg/m^3)	985	900	10	1040
	η	$(Pa.s)$	150	0.3	$2.2\ 10^{-5}$	0.001
Macroscale parameters	V_P	(m/s)	1900	1769	1428	2090
	V_S	(m/s)	359	361	390	357
	Q_P		$+\infty$	$+\infty$	$+\infty$	1967
	Q_S		$2.6\ 10^9$	$6.1\ 10^6$	$3.1\ 10^6$	15588
	ρ	(kg/m^3)	2101	2073	1779	2119

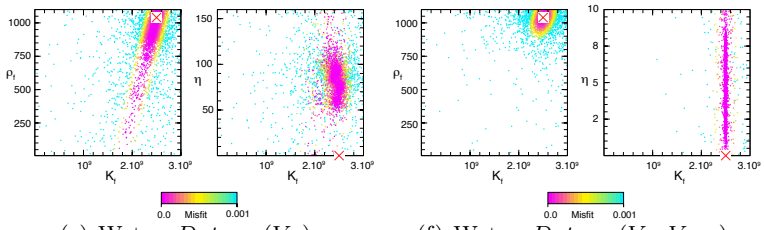
Table 2: Microscale and macroscale parameters for the Dai model. Velocities and quality factor are computed at 20 Hz.



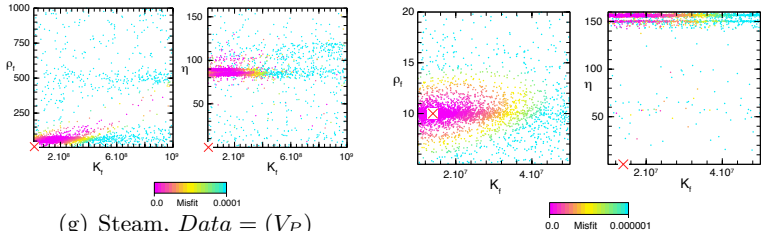
(a) Heated oil, $Data = (V_P)$ (b) Heated oil, $Data = (V_P, V_S, \rho)$



(c) Oil, $Data = (V_P)$ (d) Oil, $Data = (V_P, V_S, \rho)$



(e) Water, $Data = (V_P)$ (f) Water, $Data = (V_P, V_S, \rho)$



(g) Steam, $Data = (V_P)$ (h) Steam, $Data = (V_P, V_S, \rho)$

Figure 4: Inversion of fluid phase parameters: bulk modulus K_f (Pa), viscosity η $(Pa.s)$ and density ρ_f (kg/m^3) from (a,c,e,g) V_P and from (b,d,f,h) V_P, V_S, ρ data. The true model is represented by the red cross. The saturating fluid are (a,b) heated oil, (c,d) oil, (e,f) water and (g,h) steam. The maximum value of the misfit for the steam results (i,j) is lower (0.0001 or 0.000001 instead of 0.001). Each plot shows the values of computed models in the parameter space. The dot color depends on the misfit value.

Saturation and skeleton parameters inversion (unsaturated media)

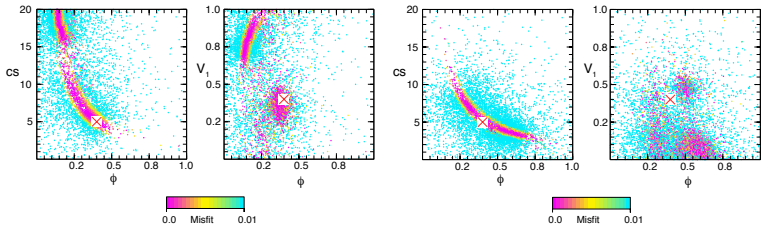
In this last sensitivity test, we use a Biot-Gassmann model for a partially saturated medium including an effective fluid phase computed by simple averages. The geological layer is composed of sands partially saturated with air and water. The parameters values are given in table 3. We invert jointly the skeleton parameters ϕ and cs and the water saturation V_1.

Microscale parameters	K_s	(GPa)	40
	G_s	(GPa)	10
	ρ_s	(kg/m^3)	2700
	V_1		0.4
	K_f (water)	(GPa)	2.2
	K_f (air)	(GPa)	$1.5\ 10^{-4}$
	K_f (effective)	(GPa)	0.0227
	ρ_f (water)	(kg/m^3)	1000
	ρ_f (air)	(kg/m^3)	1.2
	ρ_f (effective)	(kg/m^3)	400.72
	η (water)	$(Pa.s)$	0.001
	η (air)	$(Pa.s)$	$1.8\ 10^{-5}$
	η (effective)	$(Pa.s)$	$2\ 10^{-4}$
	m		1
	ϕ		0.4
	k_0	(m^2)	10^{-11}
	cs		5
Macroscale parameters	V_P	(m/s)	2377
	V_S	(m/s)	919
	Q_P		85
	Q_S		82
	ρ	(kg/m^3)	1780

Table 3: Microscale and macroscale parameters for the unsaturated medium. Velocities and quality factor are computed at 200 Hz.

The results are shown in figure 5 for (V_P), (V_P, V_S), (V_P, V_S, ρ) and (V_P, Q_P, ρ) data. Similarly to previous results, when the system is under-determined (less data than inverted parameters), it is difficult to obtain a good estimation of the inverted parameters (by about

5 to 30% of error on ϕ and cs and between 30 to 80% of error on V_1). The misfit functions (figures 5(a), 5(b) and 5(c)) underline this, showing several local minima, especially for V_1. The shape of the misfit function for ϕ and cs is more uniform but still broad. Using (V_P, V_S, ρ) data, the skeleton parameters are not exactly estimated (12% of error on ϕ and 25% of error on cs) even if the misfit function has only one global minimum. In this case, the water saturation is badly recovered as well. Then, it is necessary to use amplitude data (Q_P, figure 5(d)) to obtain significant good results (less than 1% of error). The results for (V_P, Q_P, V_S, Q_S) and $(V_P, Q_P, V_S, Q_S, \rho)$ data are not given here but show similar results as (V_P, Q_P, ρ) data.



(a) $Data = (V_P)$ (b) $Data = (V_P, V_S)$

(c) $Data = (V_P, V_S, \rho)$ (d) $Data = (V_P, Q_P, \rho)$

Figure 5: Inversion of skeleton parameters (porosity ϕ and consolidation parameters cs) and water saturation V_1 from (a) V_P, (b) V_P, V_S, (c) V_P, V_S, ρ and (d) V_P, Q_P, ρ data. The true model is represented by the red cross. Each plot shows the values of computed models in the parameter space. The dot color depends on the misfit value.

SYNTHETIC TIMELAPSE EXAMPLES

We apply the downscaling method to realistic cases. We present two kinds of reservoir time-lapse studies, the first one based on Dai et al. (1995), steam injection into oil reservoir, and the second one based on a part of the Marmousi model. The approach is the following:

- computation of synthetic poroelastic data for a set of sources and receivers,
- inversion of macroscale parameter changes from poroelastic seismograms by time-lapse FWI (first, by acoustic FWI and then by elastic FWI),
- downscaling of poroelastic parameters by global optimization.

The work is still in progress and we are not able to show the results before completing the studies. Here, we briefly describe the Dai model in terms of microscale and macroscale parameters.

Fluid substitution in a layered reservoir

Model description

This realistic case is based on Dai et al. (1995) works and consists in a oil reservoir where steam is injected to improve the oil recovery. Eight horizontal sand layers, becoming more and more consolidated by increasing the depth, are defined by their poroelastic microscale parameters (see table 4). The fluid phase saturating the porous medium is constituted by water except for the reservoir layer (layer number 6), saturated with oil before the injection. During the steam injection, two concentric half-circles are made around the injection point (top of the reservoir layer) where the porous medium becomes saturated with steam and

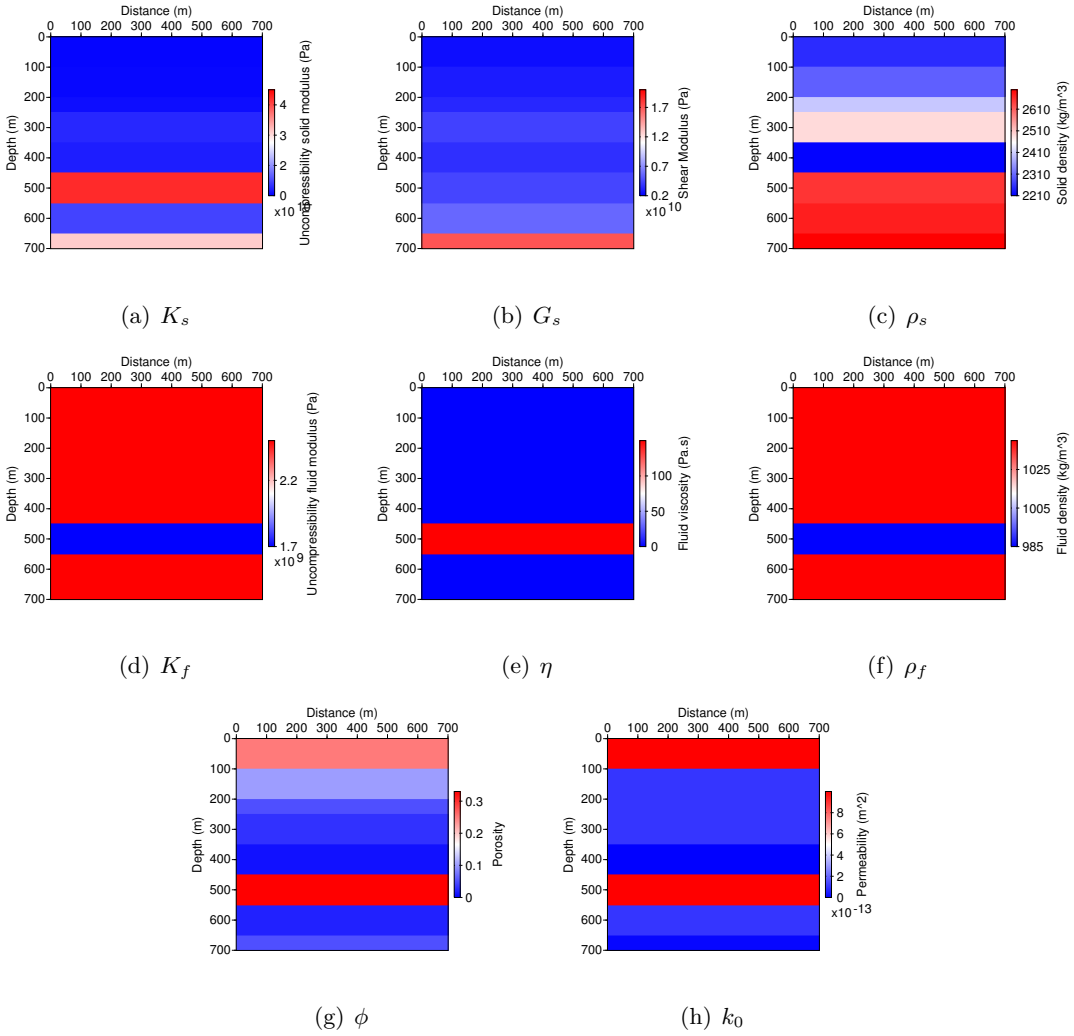
heated oil. These changes in the fluid properties affect the macroscale parameters (velocities and attenuations) (given in table 4).

Figure 6 sums up the microscale parameters values (K_s, G_s, ρ_s, K_f, η, ρ_f, ϕ and k_0) for the baseline model and their geometrical repartition. The values of grains parameters (K_s, G_s and ρ_s) increase with depth, in the same way for skeleton parameters (ϕ and k_0), except in the reservoir layer which has high porosity and permeability. The parameters cs and m are constant. Indeed, in the model definition by Dai et al. (1995), the compacity of the solid phase is carried by the grain parameters K_S et G_S (which should be defined as K_D and G_D). The fluid phase parameters are constant in each layer (water) except in the sixth one, saturated with oil.

Figure 7 shows the values of the fluid parameters after injection. We can easily observe the two concentric half-circles for steam and heated oil phases in the reservoir layer. The macroscale parameters are given in figure 8 for the baseline model and in figure 9 for the monitor model. Globally, V_P, V_S and ρ increase with depth (related to the porosity and compacity values), except in the high porosity reservoir layer which presents velocities and density lower than overburden layers (which are above and below the reservoir). The attenuations are relatively low in near-surface layers. The fluid phase change in the reservoir has an important influence only on V_P and ρ values (figure 9).

		Sand layers								Steam	Heated oil
		1	2	3	4	5	6	7	8		
K_s	(GPa)	5.2	5.3	5.8	7.5	6.9	37	9.4	26	37	37
G_s	(GPa)	2.4	2.9	3.3	4.2	3.6	4.4	5.6	17	4.4	4.4
ρ_s	(kg/m^3)	2250	2300	2400	2490	2211	2650	2670	2700	2650	2650
K_f	(GPa)	2.5	2.5	2.5	2.5	2.5	1.7	2.5	2.5	0.0014	1.2
ρ_f	(kg/m^3)	1040	1040	1040	1040	1040	985	1040	1040	10	900
η	$(Pa.s)$	0.001	0.001	0.001	0.001	0.001	150	0.001	0.001	$2.2\ 10^{-5}$	0.3
m		1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
ϕ		0.25	0.1	0.05	0.03	0.01	0.33	0.02	0.05	0.33	0.33
k_0	(m^2)	10^{-12}	10^{-13}	10^{-13}	10^{-13}	10^{-16}	10^{-12}	10^{-13}	10^{-14}	10^{-12}	10^{-12}
cs		20	20	20	20	20	20	20	20	20	20
f_c	(kHz)	19	48	17	8	1530	$4.6\ 10^6$	4.3	171	66	10^4
Biot wave regime		Diffusive									
V_P	(m/s)	1505	1613	1749	2019	2179	1900	2265	3281	1428	1768
V_S	(m/s)	330	548	733	936	1116	359	1140	1571	390	361
V_{Biot}	(m/s)	7.8	4.2	6.2	8.8	0.4	0.03	11.7	3.5	3.5	0.6
Q_P		948	413	$+\infty$	1054	$+\infty$	$+\infty$	$+\infty$	$+\infty$	$+\infty$	$+\infty$
Q_S	$*10^3$	14.33	160	172	180	$1.62\ 10^5$	$2.58\ 10^6$	194	1925	3114	6108
Q_{Biot}		0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
ρ	(kg/m^3)	1948	2174	2332	2445	2200	2100	2637	2617	1779	2073

Table 4: Microscale and macroscale parameters of the Dai reservoir model (coming from Dai et al. (1995)). The velocities and the quality factors of P-, S- and Biot waves (V_P, V_S, V_{Biot}, Q_P, Q_S and Q_{Biot}) are computed for the central frequency of the source (20 Hz).

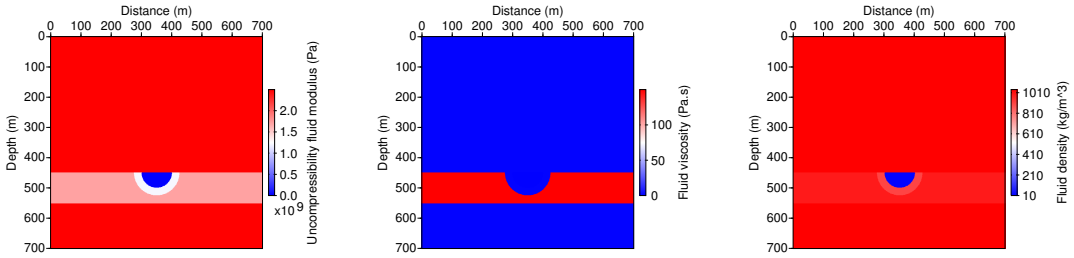


(a) K_s (b) G_s (c) ρ_s

(d) K_f (e) η (f) ρ_f

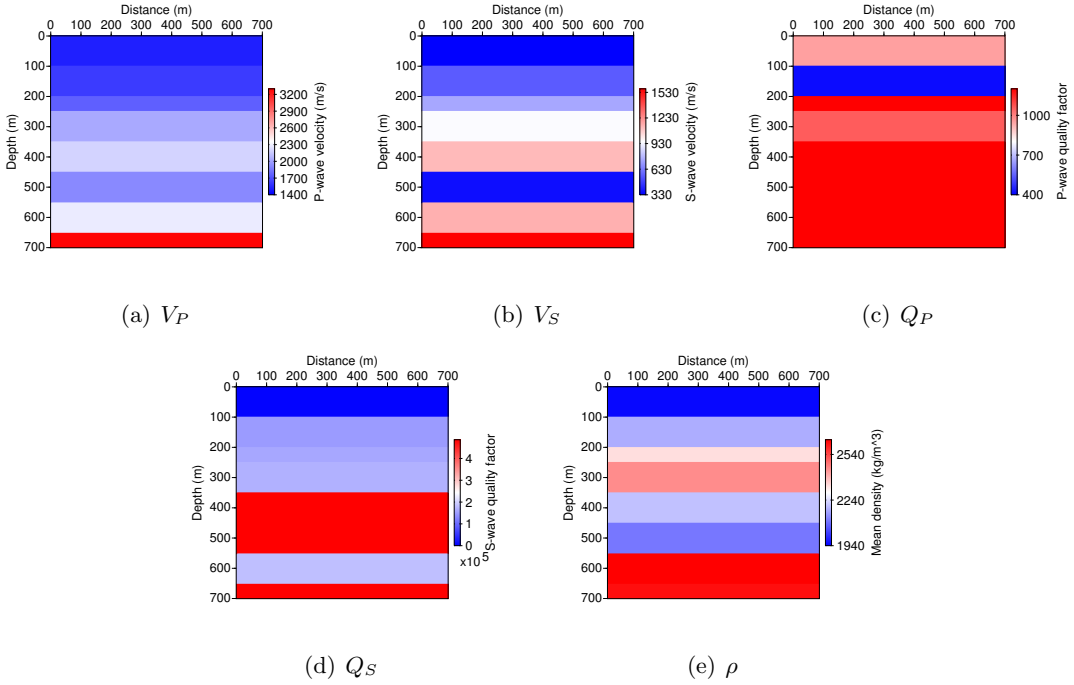
(g) ϕ (h) k_0

Figure 6: Microscale poroelastic parameters for the baseline model (before injection): (a) K_s, (b) G_s, (c) ρ_s, (d) K_f, (e) η, (f) ρ_f,(g) ϕ, (h) k_0. The exact values are given in table 4. The cs and m parameters are constant.



(a) K_f (b) η (c) ρ_f

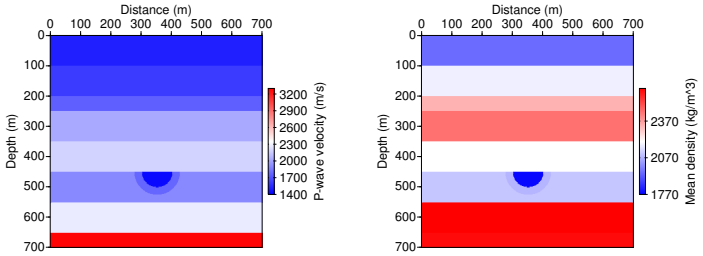
Figure 7: Microscale poroelastic parameters for the monitor model (after injection): (a) K_f, (b) η, (c) ρ_f. The exact values are given in table 4. The other parameters (grains, skeleton) have the same values for monitor and baseline models (see figure 6).



(a) V_P (b) V_S (c) Q_P

(d) Q_S (e) ρ

Figure 8: Macroscale visco-elastic parameters for the baseline model (computed at 20 Hz, before injection): (a) V_P, (b) V_S, (c) Q_P, (d) Q_S, (e) ρ. The exact values are given in table 4. The values of Q_P and Q_S that are higher than, respectively, 1200 et 500000, are not distinguished.



(a) V_P (b) ρ

Figure 9: Macroscale visco-elastic parameters for the monitor model (computed at 20 Hz, after injection): (a) V_P, (b) ρ. The exact values are given in table 4. The other parameters (V_S, Q_P et Q_S) have similar values than those for the baseline model (see figure 8).

CONCLUSIONS AND PERSPECTIVES

In this work, we have developed a complete workflow to estimate poroelastic parameters from seismograms data. We have used a two-step method, consisting of a first step of conventional seismic imaging (FWI) and a second step of downscaling. This paper is focused on the downscaling step which uses a semi-global optimization method. The forward model is constituted by analytic equations based on Biot-Gassmann theory for upscaling an effective porous medium. These equations can be frequency-dependent to take into account dispersion of propagation velocities and attenuations but these frequency effects are mainly observed for complex media (patchy saturation, double porosity). The inverse problem of downscaling is solved using a semi-global optimization method (NA algorithm) which is efficient for the fast and analytic forward model. Moreover, a probabilistic estimation can be useful to quantify the uncertainties (but it is not performed in this work).

A sensitivity study has been done first in order to find the best parameters sets to invert with respect to various macroscale data. Indeed, the jointly inversion of all poroelastic parameters is not reliable, we need to have some a priori assumptions to reduce the underdetermination of the inverse system. However, the inversion of frame/skeleton parameters is really good when the system is well determined (two data required for ϕ and cs set and three data required for ϕ, K_D and G_D set). The fluid properties are well estimated, especially the bulk modulus K_f (even with only V_P data). The sensitivity of fluid density ρ_f is lower but the inversion is correct using more data. Nevertheless, it is impossible to invert the viscosity η due to its very low sensitivity. Moreover, when the fluid has very low values (steam or gas), it is more difficult to estimate correct values because of high differences of order of magnitude in the parameters. In these cases, the inversion of the

logarithmic function of parameters could be a solution. But, even if we invert only K_f, we can determine the fluid type from it and then, recover the values of density and viscosity. The inversion of saturation and skeleton properties shows mainly that it is necessary to use Q_P to obtain a proper estimation of saturation.

To be continued ...

ACKNOWLEDGEMENTS

We would like to thank M. Sambridge for providing his neighborhood algorithm code used in this paper. Bastien Dupuy thank the ROSE project for financial support. The project was supported by the ANR "Captage et Stockage de CO_2" program (ANR-07-PCO2-002) and by national computer centers IDRIS and CINES through the project 046091. The LU factorization of the impedance matrix was performed with MUMPS available on http://graal.ens-lyon.fr/MUMPS/index.html. The mesh generation was performed with help of TRIANGLE, available on http://www.cs.cmu.edu/quake/triangle.html.

REFERENCES

Auriault, J.-L., Borne, L., and Chambon, R. (1985). Dynamics of porous saturated media, checking of the generalized law of Darcy. *Journal of Acoustical Society of America*, 77(5):1641–1650.

Avseth, P., Mukerji, T., and Mavko, G. (2005). *Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk*. Cambridge University Press, Cambridge, UK.

Bachrach, R. (2006). Joint estimation of porosity and saturation using stochastic rock-physics modeling. *Geophysics*, 71(5):53–63.

Batzle, M. and Wang, Z. (1992). Seismic properties of pore fluids. *Geophysics*, 57(11):1396–1408.

Berryman, J. (1995). *Mixture theories for rock properties in Rock Physics and Phase Relations, A Handbook of Physical Constants*. AGU Ref. Shelf, vol. 2, edited by Ahrens, T.J., AGU, Washington.

Berryman, J., Berge, P., and Bonner, B. (2000). Transformation of seismic velocity data to extract porosity and saturation values for rocks. *Journal of Acoustical Society of America*, 107(6):3018–3027.

Berryman, J., Berge, P., and Bonner, B. (2002). Estimating rock porosity and fluid saturation using only seismic velocities. *Geophysics*, 67:391–404.

Biot, M. (1956). Theory of propagation of elastic waves in a fluid-saturated porous solid. I. low-frequency range, II. higher frequency range. *Journal of Acoustical Society of America*, 28:168–191.

Bosch, M. (1999). Lithologic tomography : From plural geophysical data to lithology estimation. *Journal of Geophysical Research*, 104(B1):749–766.

Brie, A., Pampuri, F., Marsala, A., and Meazza, O. (1995). Shear sonic interpretation in gas-bearing sands. *SPE Annual Technical Conf. 30595*, pages 701–710.

Brossier, R. (2009). *Imagerie sismique à deux dimensions des milieux visco-élastiques par inversion des formes d'onde: développements méthodologiques et applications.* PhD thesis, Université de Nice-Sophia-Antipolis.

Brown, R. (1980). Connection between formation factor for electrical resistivity and fluid-solid coupling factor for Biot's equation in fluid filled porous media. *Geophysics*, 45(8):1269–1275.

Brown, R. and Korringa, J. (1975). On the dependence of the elastic properties of a porous rock on the compressibility of the pore fluid. *Geophysics*, 40:608–616.

Burridge, R. and Vargas, C. (1979). The fundamental solution in dynamic poroelasticity. *Geophysical Journal of the Royal Astronomical Society*, 58:61–90.

Carcione, J., Picotti, S., Gei, D., and Rossi, G. (2006). Physics and seismic modeling for monitoring CO_2 storage. *Pure and Applied Geophysics*, 163:175–207.

Castagna, J. P., Batzle, M. L., and Kan, T. K. (1993). Rock physics - the link between rock properties and avo response, in offset-dependent reflectivity-theory and practice of avo analysis, edited by Castagna J.P. and Backus M.M.

Chotiros, N. (2002). An inversion for Biot parameters in a water-saturated sand. *Journal of Acoustical Society of America*, 112(5):1853–1868.

Dai, N., Vafidis, A., and Kanasewich, E. (1995). Wave propagation in heteregeneous porous media: A velocity-stress, finite-difference method. *Geophysics*, 60(2):327–340.

Domenico, S. N. (1976). Effect of brine-gas mixture on velocity in an uncosolidated sand reservoir. *Geophysics*, 41:882–894.

Dupuy, B., Garambois, S., and Virieux, J. (2012). Wave propagation in multiphasic complex porous media formulated in the frequency-space domain. *Geophysics*, pages –.

Dutta, A. J. and Odé, H. (1979). Attenuation and dispersion of compressional waves in fluid-filled porous rocks with partial gas saturation (White model)-Part i: Biot theory. *Geophysics*, 44(11):1777–1788.

Gassmann, F. (1951). Über die elastizität poröser medien. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 96:1–23.

Gunning, J. and Glinsky, M. (2007). Detection of reservoir quality using bayesian seismic inversion. *Geophysics*, 72(3):37–49.

Han, D.-H., Nur, A., and Morgan, D. (1986). Effects of porosity and clay content on wave velocities in sandstones. *Geophysics*, 51(11):2093–2107.

Hill, R. (1952). The elastic behavior of a cristalline aggregate. *Proceedings of the Physical Society*, A65:349–354.

Johnson, D. (2001). Theory of frequency dependent acoustics in patchy-saturated porous media. *Journal of Acoustical Society of America*, 110(2):682–694.

Johnson, D., Koplik, J., and Dashen, R. (1987). Theory of dynamic permeability and tortuosity in fluid-saturated porous media. *Journal of Fluid Mechanics.*, 176:379–402.

Malinowski, M., Operto, S., and Ribodetti, A. (2011). High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion. *Geophysical Journal International*, 186(3):1179–1204.

Mavko, G., Mukerji, T., and Dvorkin, J. (2009). *The Rocks Physics Handbooks, Tools for Seismic Analysis in Porous Media; Second Edition.* Cambridge University Press, Cambridge, UK.

Menke, W. (1984). *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press, Inc., Orlando, USA.

Pride, S. (2005). *Hydrogeophysics*, pages 253–284. Water Science and Technology Library, Springer, The Netherlands.

Pride, S., Berryman, J., and Harris, J. (2004). Seismic attenuation due to wave-induced flow. *Journal of Geophysical Research*, 109(B01201):1–19.

Pride, S., Gangi, A., and Morgan, F. (1992). Deriving the equations of motion for porous isotropic media. *Journal of Acoustical Society of America*, 92(6):3278–3290.

Raiga-Clemenceau, J., Martin, J., and Nicoletis, S. (1988). The concept of acoustic formation factor for more accurate porosity determination from sonic transit time data. *Log Analyst*, 219:54–60.

Raymer, L., Hunt, E., and Gardner, J. (1980). An improved sonic transit time-to-porosity transform. In *Abstracts Soc Professional Well Log Analysis (SPWLA), 21st Ann. Logg. Symp., Paper P July.*

Reuss, A. (1929). Berechnung der fliebgrenze von mischkristallen auf grund der plastiz-

itatsbedingung fur einkristalle. *Zeitschrift fr Angewandte Mathematik und Mechanik / Journal of Applied Mathematics and Mechanics*, 9:49–58.

Rubino, J. and Velis, D. (2009). Thin-bed pre stack spectral inversion. *Geophysics*, 74(4):R49–R57.

Rubino, J. and Velis, D. (2011). Seismic characterization of thin beds containing patchy carbon dioxide-brine distributions: a study based on numerical simulations. *Geophysics*, 76(3):R57–R67.

Sambridge, M. S. (1999a). Geophysical inversion with a neighbourhood algorithm - I. searching a parameter space. *Geophysical Journal International*, 138:479–494.

Sambridge, M. S. (1999b). Geophysical inversion with a neighbourhood algorithm - II. appraising the ensemble. *Geophysical Journal International*, 138:727–746.

Tang, X. and Cheng, C. (1996). Fast inversion of formation permeability from Stoneley wave logs using a simplified Biot-Rosenbaum model. *Geophysics*, 61(3):639–645.

Tarantola, A. (1987). *Inverse problem theory: methods for data fitting and model parameter estimation*. Elsevier, New York.

Teja, A. and Rice, P. (1981). Generalized corresponding states method for viscosities of liquid mixtures. *Industrial and Engineering Chemistry Fundamentals*, 20:77–81.

van Dalen, K. N., Ghose, R., Drijkoningen, G., and Smeulders, D. (2010). In-situ permeability from integrated poroelastic reflection coefficients. *Geophysical Research Letters*, 37:L12303.

Voigt, W. (1889). Uber die beziehung zwischen den beiden elastizitatkonstanten isotroper körper. *Annalen der Physik*, 38:573–587.

Walton, K. (1987). The effective elastic moduli of a random packing of spheres. *Journal of the Mechanics and Physics of Solids*, 35:213–226.

White, J. E. (1975). Computed seismic speeds and attenuation in rocks with partial gas saturation. *Geophysics*, 40(2):224–232.

Wyllie, M., Gregory, J., and Gardner, L. (1956). Elastic wave velocities in heterogneous and porous media. *Geophysics*, 21:41–70.