



HAL
open science

Elastic full waveform inversion in anisotropic media : methodology and application to ocean bottom node data from the North Sea

Aline Robin

► **To cite this version:**

Aline Robin. Elastic full waveform inversion in anisotropic media : methodology and application to ocean bottom node data from the North Sea. Earth Sciences. Université Paris Cité, 2021. English. NNT : 2021UNIP7023 . tel-03516351

HAL Id: tel-03516351

<https://theses.hal.science/tel-03516351>

Submitted on 7 Jan 2022

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.



Thèse préparée à l'institut de physique du globe de Paris
Université de Paris

École doctorale Sciences de la Terre et de l'environnement
et physique de l'Univers STEP'UP n°560

Équipe de Géosciences marines

Elastic full waveform inversion in anisotropic media

Methodology and Application to ocean bottom node data from the North Sea

par Aline ROBIN

Thèse de doctorat de Sciences de la Terre et de l'environnement

- dirigée par
Satish SINGH
- présentée et soutenue publiquement le
4 février 2021
- devant un jury composé de
Thomas BOHLEN, Professeur, Geophysikalisches Institut (Karlsruhe, Allemagne), Rapporteur
Romain BROSSIER, Maître de conférence, Isterre (Grenoble, France), Rapporteur
Jean-Paul MONTAGNER, Professeur, IPGP (Paris, France), Examinateur et président du jury
Phuong-Thu TRINH, Chercheuse géophysicienne, TOTAL (Pau, France), Examinatrice
Paul WILLIAMSON, Chercheur géophysicien, TOTAL (Houston, USA), Invité
Mark NOBLE, Maître de recherche, MINES ParisTech (Fontainebleau, France), Co-encadrant
Satish SINGH, Physicien, IPGP (Paris, France), Directeur de thèse

Acknowledgements

I would first like to thank Satish SINGH and Mark NOBLE for their role of Ph.D. director and co-supervisor respectively. Thank you for giving me my chance despite your busy schedules. I learned numerous things by your side. If I had to summarize the dynamics of our group, I would say that Satish, you made me aim for the stars, while Mark, you helped me achieving the small steps of this long journey. I finally only reached the moon, but I am still aiming for the stars.

I would also like to thank Paul WILLIAMSON for his advisor role all along this thesis. Whenever I needed it, you offered me your help, whether it be for theoretical developments, synthetic/real applications, or general discussions. You also helped me obtaining my next job, I am deeply touched by the trust you show me.

Thank you to Thomas BOHLEN and Romain BROSSIER for their role as thesis referees, and thank you to Jean-Paul MONTAGNER and Phuong-Thu TRINH for their role as thesis examiners. Each one of you means something special to me. Consequently, you telling me that I deserve the Ph.D. title is a great honor.

In order to be able to defend my thesis, my Ph.D. thesis committee members have also been of great help: Jean-Pierre VILOTTE (first year), Gillian ROYLE (first to last year) and Nobuaki “Nobu” FUJI (first to last year). I would particularly like to single out Nobu and Gillian. Nobu, while it was not your role, you decided to take a number of hours of your time to help me, at a moment where I highly needed it: reciprocity, we are coming! Gillian, just like Paul, you encouraged me to take a break during my Ph.D. to perform an internship in industry, and it was excellent advice.

Likewise, I would like to thank Colin THOMSON for his great contribution in the reflection coefficients computation and analysis. I admire your scientific and human skills. You have a way to make us, young researchers, feel worthy and become better.

I would also like to acknowledge TOTAL E&P Research and Development for their funding, and TOTAL E&P UK Limited for their data, without which this project would not have been possible. Similarly, data use would not have been achievable without PANGEA (TOTAL, Pau, France) & S-CAPAD (IPGP, Paris, France) platforms and teams, on and with which the numerical computations were performed. A special thank to Geneviève MOGUILNY and Mathieu APPE, always here to come to my rescue when I got lost inside S-CAPAD and PANGEA respectively.

In fact, thank you to all IPGP, TOTAL, STEP’UP, University of Paris and other entity members who participated in this Ph.D., including for emotional support and administrative aspects. To cite a few, thank you Dmitry BORISOV for introducing

me to the 3D FWI code, when I was assessing the feasibility of performing 3D inversions of Alwyn data. Thank you Shihao YUAN, especially for explaining me localized FWI, when I was appraising the achievable resolution one can locally get with a smooth baseline model, in a cost reduction context. Similarly, thank you Weiguang HE for initiating me to elastic VTI FWI and the corresponding in-house code, among other things. Thank you Aurélie COUDURIER CURVEUR McCALLUM for your great help to write the introduction. Thank you Nassima BENBELKACEM and Emma GREGORY for helping me with the abstract. Thank you to both of you and to Pranav AUDKHASI, Hélène CARTON, Matilde FERRANTE, Milena MARJANOVIC, Venkata VADDINENI, Mathurin WAMBA and Zhikai WANG for your assistance in the defense presentation preparation. I would also like to thank my wonderful officemates and/or occasional/regular running partners: Gaëlle BENATRE, Manon BICKERT, Pauline BONNET, Stephan BORENSZTAJN, Jie CHEN, Hanjin CHOE, Océane FOIX, Ssu-Ting LAI, Muriel LAURENCIN, Aude LAVAYSSIERE, Lianjun LI, Ekeabino MOMOH, Chloé SEIBERT, Benjamin WHEELER, Xiaodong YANG, Fei ZHOU and again, Pranav, Nassima, Aurélie, Venkata and Zhikai. Thank you Mihai ROHARIK for your IT support and much more, and thank you Marie-Dominique ROCHERON, Martine PRUDENT and Florence ROUILLE for your administrative support and sympathetic ear. Actually, I have a special thought for the members of the marine geoscience team and the seismology team. I enjoyed very much my time with you, I will miss our scientific and non scientific discussions.

In addition, I would like to thank all previous colleagues/teachers who paved my road. In particular, thank you to my TOTAL Houston's colleagues, especially Fuchun GAO, my former supervisor, and Bertrand DUQUET, my former boss, with who the idea of doing a Ph.D. emerged in my head. Together with Jean-Luc PIAZZA, they also layed the first bricks of this Ph.D..

Finally, I would like to thank my parents. You have done so much for me that it is impossible to summarize it in a few sentences. I will therefore simply thank you for everything. Thank you to my brother and his family. You have been an example to me on so many aspects. Finally, borrowing a concept from my dear colleague Milena, I would like to award the title of Ph.D. Assistant (aka Ph.D.A.) to my significant other. I had wonderful years by your side, and I can't wait to share the next ones with you, thank you.

Abstract Seismology and seismic imaging methods hold a prominent place in the discipline of Earth Sciences. Through its use of the entire seismic data (i.e. the whole waveform), Full Waveform Inversion (**FWI**) has emerged as a powerful class of algorithms to retrieve quantitative and high resolution models of mechanical properties of the sub-surface, such as density or (visco-)elastic parameters. However, **FWI** suffers from high computing costs and from the intrinsic limitations of the embedded line-search inversion algorithm, which provides a local solution instead of the global one. Focusing on Earth local imaging in marine environments, these limitations led to the initial restriction of **FWI** to a fluid characterization of the Earth, where only the P-wave velocity was updated. Benefiting from increased computing resources, recent efforts produced a successful and much needed 2D inversion strategy for both P- and S-wave velocities in soft seabed environments. In those environments, the design of such a strategy was complicated by a poorer S-wave information data content, and consequently to a difficult S-wave velocities retrieval. The strategy was developed assuming an isotropic sub-surface and using ocean-bottom cable data. In order to meet modern applications needs, my work focuses on more contemporary recording devices, and partially relaxes the isotropic limitation.

Over the years, ocean-bottom cables are indeed replaced by ocean-bottom seismometers/nodes (**OBS/OBN**). Without the use of the reciprocity principle, where the roles of the airgun sources and the **OBS/OBN** receivers are interchanged, the increased computing cost would be prohibitive. After gathering the reciprocity relations needed for data modeling, I extend the theoretical developments to the justification of the equality between inversions with and without reciprocity. A demonstration of the modeling reciprocity relations allows me to emphasize their limitations and to derive a few extensions. Synthetic modeling and inversions highlight the role of reciprocity as a control tool, to obtain in two different ways the same results with the same algorithm, even for non-simple media. Overall, I provide an extensive guide on reciprocity in elastic anisotropic 2D/3D **FWI**, from theory to implementation, and from data modeling to inversion.

Furthermore, the available soft seabed strategy was developed in isotropic environments, while the real Earth is anisotropic. I assess the performance of the strategy for vertical velocity retrieval when slightly inaccurate remaining anisotropy parameters are used, as usual in real data applications. Those remaining anisotropy parameters are not updated and I focus on a specific kind of anisotropy called “vertical transverse isotropy”. After adapting the parameterization, reflection coefficient computations and synthetic **FWI** applications are used for the evaluation, which concludes that the inversion is perturbed but the strategy can still remain overall successful. I also show that the low-frequency content of modern data and the use of the increasing frequency multiscale strategy is able to substitute the essential intermediate-scale S-wave velocity features inversion stage, which uses the amplitudes of wide-aperture P-P reflections. This stage may fail in real data applications, which, for example, may not consider the existing attenuation. This variant, however, shows an increased sensitivity to an erroneous anisotropy.

I support my synthetic and theoretical findings with a real 2D **OBN** dataset from the North Sea Alwyn North hydrocarbon field. This **FWI** application benefits from high quality inputs and excellent control tools (e.g., well data) that allow for a true validation of the success of the new strategy.

A natural perspective of this study is to work towards a fully anisotropic **FWI**, where anisotropy parameters are also inverted. Such an inversion is a problem that remains unsolved, despite the great efforts of the community. In pursuit of such a goal, I notably propose to question the design of future acquisition systems.

Résumé Les méthodes de sismologie et d'imagerie sismique occupent une place considérable au sein des sciences de la Terre. Parmi ces méthodes, l'inversion de forme d'onde (**FWI**) est apparue comme une formidable classe d'algorithmes pour obtenir des modèles quantitatifs et haute résolution des propriétés mécaniques du sous-sol (par exemple la densité ou les paramètres (visco-)élastiques). Le succès de la **FWI** provient de son utilisation de l'ensemble des données sismiques (c'est-à-dire de la forme d'onde complète). Cependant, la **FWI** souffre de coûts de calculs élevés et des limitations inhérentes à l'algorithme d'inversion de recherche linéaire utilisé. En effet, la recherche linéaire fournit une solution locale et non une solution globale. En ce qui concerne l'imagerie de la Terre en environnement marin à l'échelle locale, ces limitations ont conduit à la restriction initiale de la **FWI** à une représentation fluide de la Terre, où seule la vitesse des ondes P est mise à jour. L'augmentation des ressources de calcul a récemment permis d'aboutir à une stratégie d'inversion 2D tant des vitesses des ondes P que celles des ondes S, pour les environnements présentant de faibles vitesses S en fond de mer. Dans ce type d'environnement, la conception d'une telle stratégie était entravée par l'amointrissement des données correspondantes en information S, et donc à une inversion plus délicate des vitesses S. La stratégie a été élaborée en approximant le sous-sol par un milieu isotrope et en utilisant des données de câbles de fond de mer (**OBC**). Cette thèse se concentre sur des dispositifs d'enregistrement plus contemporains et assouplit partiellement la limitation isotrope, afin de répondre aux besoins des applications modernes.

Au fil des ans, les **OBC** sont en effet remplacés par des sismomètres/noeuds de fond de mer (**OBS/OBN**). Sans l'utilisation du principe de réciprocité, où les rôles des sources (canons à air) et des récepteurs sont interchangeables, le coût de calcul accru de ces types d'acquisition serait prohibitif. Après avoir rassemblé les relations de réciprocité nécessaires à la modélisation des données, j'étends les développements théoriques à l'explication de l'égalité entre inversions avec et sans réciprocité. Une démonstration des relations de modélisation me permet de souligner leurs limites et d'en déduire des extensions. Des modélisations et inversions synthétiques mettent en évidence le rôle de la réciprocité comme un outil de contrôle, pour obtenir de deux manières différentes les mêmes résultats avec le même algorithme, même pour des milieux complexes. Au final, je fournis un guide complet et sans précédent sur la réciprocité pour la **FWI** 2D/3D en milieux élastiques anisotropes : de la théorie à l'implémentation et de la modélisation à l'inversion.

Par ailleurs, la stratégie d'inversion des vitesses des ondes P et S étudiée a été développée dans des environnements isotropes, alors que la Terre est anisotrope. J'évalue la performance de la stratégie pour l'inversion des vitesses verticales, lorsque les autres paramètres d'anisotropie sont légèrement erronés, comme c'est le cas pour des applications réelles. Ces paramètres ne sont pas mis à jour et je me concentre sur un type d'anisotropie spécifique nommée « isotropie transverse verticale » (**VTI**). Après adaptation de la paramétrisation, des calculs de coefficients de réflexion et des applications synthétiques **FWI** sont utilisés pour l'évaluation. Cette évaluation conclut à une inversion perturbée mais à une stratégie qui peut rester globalement probante. Je montre également que le contenu basse fréquence des données modernes, combiné à l'utilisation de la stratégie d'inversion par accroissement des fréquences, est capable de remplacer l'étape d'inversion des détails de taille intermédiaire des vitesses d'ondes S, qui utilise les amplitudes des réflexions P-P grand angle. Cette étape pourrait échouer lors d'applications à des données réelles, qui, par exemple, ne prennent généralement pas en compte l'atténuation de la Terre. En revanche, la variante proposée montre une sensibilité accrue à des paramètres d'anisotropie erronés.

Je confirme finalement les conclusions synthétiques et théoriques avec un jeu de données réelles **OBN** 2D, acquis au dessus du champ d'hydrocarbures Alwyn North en mer du Nord. Cette application bénéficie d'entrées de haute qualité et de données permettant une véritable validation du succès de la nouvelle stratégie, grâce à d'excellents outils de contrôle (par exemple des données de puits).

Une perspective naturelle de cette étude est de travailler en direction d'une **FWI** où les paramètres d'anisotropie seraient également mis à jour. Une telle inversion est un problème qui reste non résolu, malgré les efforts de la communauté. Dans la poursuite d'un tel objectif, je propose notamment de nous interroger sur la conception des futurs systèmes d'acquisition.

Contents

Introduction	1
1 Full waveform inversion: overview, theory and implementation	9
1.1 Introduction	10
1.2 Forward problem	11
1.2.1 A choice of wave equation: a first-order, time-space domain, 2D elastic vertical transverse wave equation	11
1.2.2 A choice of solving method: a finite-difference method	13
1.3 Inverse problem	20
1.3.1 A choice of objective function: an ℓ_2 -norm cost function	20
1.3.2 A choice of inversion algorithm: a conjugate-gradient and a BFGS algorithm	22
1.3.3 The gradient computation	26
1.4 Inversion strategy	28
1.4.1 Resolution aspects	28
1.4.2 Driving full waveform inversion towards the intended optimum	28
1.4.3 About multiparameter inversion	30
1.5 Discussion and conclusions	31
2 Reciprocity for fast full waveform inversion on airgun-ocean bottom node or seismometer data in elastic vertical transverse isotropic media	33
2.1 Introduction	34
2.2 A few necessary ingredients...	37
2.2.1 The used time-domain, elastic vertical transverse isotropic wave equation	37
2.2.2 From convolutional type reciprocity relation to Green's functions and Green's functions reciprocity	37
2.2.3 Displacement and pressure expressions for an airgun source using Green's functions	38
2.3 Reciprocity for pressure and velocity data modeling	39
2.3.1 Reciprocity relation for pressure data	40
2.3.2 Reciprocity relation for velocity data	42
2.3.3 Synthetic data applications	44
2.4 Reciprocity for full waveform inversion	49
2.4.1 A change of point of view	49
2.4.2 Synthetic applications	50
2.5 Discussion and conclusions	52

3	Full waveform inversion to retrieve P-and S-wave vertical velocities in soft seabed vertical transverse isotropic media: inversion strategy, parameterization and sensitivity study	55
3.1	Introduction	56
3.2	The soft seabed strategy	58
3.3	A choice of parameterization	62
3.4	Amplitude sensitivity of wide-aperture P-P reflections to elastic vertical transverse isotropic parameters	68
3.5	Synthetic application of the soft seabed strategy with slightly erroneous anisotropy parameters	70
3.6	A variant to the soft seabed strategy	71
3.7	Discussion and conclusions	73
4	Alwyn real data: data and full waveform inversion inputs analysis	75
4.1	Introduction	75
4.2	Geology of Alwyn North field	76
4.3	Ocean bottom node data description and analysis	80
4.3.1	Objectives of the acquisition	80
4.3.2	Description of the acquisition	80
4.3.3	Seismic data processing	84
4.3.4	Data analysis	85
4.4	Discussion and conclusions	101
5	Alwyn real data: full waveform inversion application	105
5.1	Introduction	106
5.2	Inputs preparation	108
5.2.1	Initial model	108
5.2.2	Seismic data	110
5.2.3	Acquisition geometry	114
5.2.4	Source function estimation	115
5.3	Application of the variant of the soft seabed strategy	117
5.4	Application of the original soft seabed strategy	127
5.5	Original strategy with a modified stage order	131
5.6	Inversions with Scholte waves	135
5.7	Discussion and conclusions	141
	Conclusion	145
	A Details on Newton’s method	149
	B Details on the computation of the opposite direction of update: conjugate-gradient and L-BFGS-B	151
	C Steplength computation: details on Pica et al.’s and subspace methods	153
	D Details on the adjoint-state method for 2D elastic vertical transverse isotropic full waveform inversion	155
D.1	Adjoint-state formulation	155
D.2	Adjoint operators	157

D.2.1	First adjoint operator	157
D.2.2	Second adjoint operator	158
D.3	Computation of \mathbf{d}^\dagger	159
D.4	Gradient computation	160
D.4.1	Case $C_{44}(\mathbf{x}) \neq 0$ and $C_{13}(\mathbf{x})^2 - C_{11}(\mathbf{x})C_{33}(\mathbf{x}) \neq 0$	160
D.4.2	Case $C_{44}(\mathbf{x}) = 0$	161
D.4.3	Case $C_{13}(\mathbf{x})^2 - C_{11}(\mathbf{x})C_{33}(\mathbf{x}) = 0$	161
E	Convolutional type reciprocity relation derivation in elastic anisotropic media	163
F	A few elastic vertical transverse isotropic parameters	165
G	Alwyn North field history: an educational case	167
H	Source time function inversion: theory and synthetic applications	171
H.1	Theory	172
H.1.1	2D elastic vertical transverse isotropic wave equation	172
H.1.2	Source time function inversion	172
H.2	Synthetic applications on a modified Marmousi2 model	173
H.2.1	An ideal application	173
H.2.2	A more realistic application	175
I	Alwyn real data: variant soft seabed strategy additional displays	177
I.1	Variant strategy	178
I.2	Stage 1a	179
I.3	Stage 1b	184
I.4	Stage 3'	189
J	Alwyn real data: original strategy additional displays	197
J.1	Stage 2	197
J.2	Stage 3 - first frequency range	203
K	Alwyn real data: original strategy with a modified stage order additional displays	211
K.1	Stage 2'	211
K.2	Stage 1b' - first frequency range	217
	Bibliography	223

List of Figures

- 1 Schematic description of Full Waveform Inversion (FWI) algorithms where a single type of mechanical property is updated (P-wave vertical velocity here). In practice, it could be any number of property type. Note the iterative aspect of these algorithms. A single trace is represented but a multiplicity of them is actually used. 2
- 2 Synthetic applications illustrating the sensitivity of Full Waveform Inversion (FWI) to the initial model. Same inversion strategy is applied in both cases. 3
- 3 Description of the usual source system (airgun) and the three usual recording systems (streamer, Ocean Bottom Cable and Node/Seismometer (OBC and OBN/OBS)) in the context of local marine acquisitions. The sketch is not up-to-scale but aims at highlighting the different scales of these acquisition systems, as well are their location properties. Typically, OBNs/OBSs are spaced by a few hundreds meters and cover areas of a few dozens of kilometers. 5
- 4 Vertical Transverse Isotropic (VTI) medium: the velocity of the wave in red depends only on the angle θ with the vertical axis z , and not the angle φ with the horizontal axis x 6
- 1.1 General structure of Full Waveform Inversion FWI algorithms. 10
- 1.2 Implemented staggered grid. Δh is the spatial spacing. 14
- 1.3 Illustration of the functioning and user computing time reduction of shot parallelization. 18
- 1.4 Example of domain decomposition for a spatial domain of computation divided into 2 horizontal subdomains and a Finite Difference (FD) method using 2 points on each side of the computed point (4^{th} order centered spatial Taylor expansion). For 8^{th} order centered spatial Taylor expansions, 4 points are needed. 18
- 1.5 Example of speed-up of our domain decomposition implementation. 19
- 1.6 Illustration of the organization of a node and its memory. The memory is shared by all the cores of the node. 20
- 1.7 Visual description of the different kinds of stationary points on a virtual objective function which depends on a single parameter. 23
- 1.8 Visual description of the effect of a too big steplength on a virtual objective function which depends on a single parameter: the direction of update is the descent direction but the steplength is too big leading to an increase of the cost function. The blue cross indicates the initial model and the red cross indicates the desired minimum. The arrows represent the successive updates. 25

1.9	Visual description of the effect of an initial model closer to the aimed minimum, as well as the effect wider strictly convex neighborhood around the aimed minimum, with two virtual cost functions which depend on a single parameter. The blue cross indicates the initial model and the red cross indicates the desired minimum. The arrows show the expected obtained model for each cost function and initial model.	26
1.10	Steps of the gradient computation in our Full Waveform Inversion FWI algorithm. The inputs of each step are indicated in orange.	27
1.11	“Schematic of cycle-skipping artifacts in FWI. The solid black line represents a monochromatic seismogram of period T as a function of time. The upper dashed line represents the modeled monochromatic seismograms with a time delay greater than $T/2$. In this case, FWI will update the model such that the $n+1^{th}$ cycle of the modeled seismograms will match the n^{th} cycle of the observed seismogram, leading to an erroneous model. In the bottom example, FWI will update the model such that the modeled and recorded n^{th} cycle are in-phase because the time delay is less than $T/2$.” Figure and title from Virieux and Operto (2009) (212).	29
2.1	Schematic description of reciprocity relations, i.e.: relations linking two seismic quantities (e.g., pressure, displacement) from two source-receiver pairs with interchanged locations. As represented, both sources, receivers and quantities can be of different types.	34
2.2	Illustration of reciprocity for a dataset containing seven airguns and two Ocean Bottom Nodes (OBNs). When the acquisition contains seven sources, the reciprocal simulations would ideally contain only two virtual sources instead of seven simulations, reducing the computing cost.	40
2.3	Reciprocal experiment for pressure component.	41
2.4	Reciprocal experiment for velocity components. VTI = Vertical Transverse Isotropic. f_p , with $p = x, y, z$, are the components of the body force source of equation 2.1	43
2.5	P-wave vertical velocity model used to test reciprocity in modeling. Velocities in m/s.	45
2.6	Data obtained with (blue) and without (red) reciprocity. Only the blue color is visible because of an excellent match of the data obtained with and without reciprocity. From top to bottom: horizontal velocity component, vertical velocity component and hydrophone component. From left to right: 0, 3750 and 7500 m of offset.	46
2.7	Illustration of the limitation of the used reciprocity relations. Data obtained with (blue) and without (red) reciprocity in settings where the Kaiser window of hydrophone or source reaches the seabed. (a) pressure data obtained with a hydrophone whose interpolation window reaches the seabed; (b) pressure, (c) horizontal and (d) vertical velocity obtained with a source whose interpolation window reaches the seabed. 3750 m of offset.	47

2.8	Illustration of the limitation of the used pressure reciprocity relation with a fictional test, where the hydrophone is not located in the same isotropic acoustic medium as the airgun as requested by the pressure reciprocity relation, but inside the elastic Vertical Transverse Isotropic (VTI) sub-surface. Pressure trace without reciprocity is in red and the pressure trace obtained with interchanged airgun and hydrophone locations without any other modification is in black.	48
2.9	Illustration with a fictional test of the proposed pressure reciprocity relation for an airgun and a hydrophone in different isotropic acoustic media. The hydrophone is located inside the isotropic acoustic sub-surface while the airgun is inside the water layer. Pressure trace obtained with reciprocity is in dashed blue, pressure trace obtained without reciprocity in red. The trace obtained with interchanged airgun and hydrophone locations without any other modification is in black.	48
2.10	P-wave vertical velocity models (in in m/s): (a) true, (b) initial, (c) inverted with pressure component without reciprocity, (d) inverted with vertical velocity component without reciprocity, (e) inverted with pressure component and reciprocity, (f) inverted with vertical velocity component and reciprocity, (g) e-c ,(h) f-d.	51
3.1	Inversion results for a (a) strong seabed environment (S-wave seabed velocity = 1430 m/s), (b) soft seabed environment typical from a clastic sedimentary basin (S-wave seabed velocity = 300 m/s). The same simple inversion strategy, realistic source and acquisition geometry are used in both cases. Figures from Sears et al. (2008) (164).	59
3.2	Modeled data from strong (top: a-b-c) and soft (bottom: d-e-f) seabed models. Left (a-d): pressure component; Middle (b-e): horizontal velocity component; Right (c-f): vertical velocity component. Red (resp. green) polygon: S-wave-related (resp. wide-aperture) data approximate selection. Figure modified from Sears et al. (2008) (164).	60
3.3	Inversion results of the last stage of the soft seabed strategy applied to the same initial model and data of figure 3.1. Figure modified from Sears et al. (2008) (164).	61
3.4	P- and SV-wave velocities at all propagation angles with (dashed line - equation F.5) and without (solid line - equation F.1) the weak anisotropy approximation. Note the differences with the isotropic case, where the velocity would be the same for all angles (use the thin circles for reference). The medium parameters are those of the bottom layer from table 3.2. The angle is measured from the vertical axis.	62
3.5	Normalized P-P and P-SV reflection radiation patterns for P- and SV-wave velocities with the (V_p, V_s, ρ) parameterization. The angle corresponds to the incident angle. V_p, V_s and ρ are respectively set equal to the P- and SV-wave vertical velocities and density of the bottom layer from table 3.2. Computed with formula from Stolt and Weglein (2012) (183).	63

3.6	Normalized P-P and P-SV reflection radiation patterns for P- and SV-wave velocities with the $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$ parameterization. The angle corresponds to the incident angle. The medium parameters are those of the bottom layer from table 3.2. The patterns are computed with formula from He and Plessix (2016) (78).	64
3.7	Comparison of non weak (solid lines) and weak (crosses) approximation for normalized P-P and P-SV reflection radiation patterns for P- and SV-wave velocities with the $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$ parameterization. The angle corresponds to the incident angle. The medium parameters are those of the bottom layer from table 3.2. The patterns are computed with formula from He and Plessix (2016) (78).	64
3.8	Used source time function and its frequency spectrum.	65
3.9	Soft seabed synthetic inversion results, in an environment whose anisotropy parameters are known. Inversion performed with the logarithm of the parameterization $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$.	66
3.10	Normalized P-P and P-S reflection radiation patterns for P- and SV-wave velocities with the $(V_{pv}, V_{pn}, V_{ph}, V_{sv}, \rho)$ parameterization. The angle corresponds to the incident angle. The medium parameters are those of the bottom layer from table 3.2.	67
3.11	Soft seabed synthetic stage 1 inversion results obtained with the successful selected parameterization applied to all data (blue line) and narrow-aperture data only (cyan line), as well as results obtained with a failing parameterization applied to all data (magenta line). The anisotropy parameters are known. The used parameterizations are the logarithm of the parameterizations indicated in the legend.	67
3.12	P-P reflection coefficients for a downgoing plane-wave hitting a horizontal interface between two half-spaces. Computations performed for various interfaces and incident total velocity angles. The angle is measured from the vertical axis. Each subfigure corresponds to the perturbation of one parameter of the base model of table 3.2.	69
3.13	True and perturbed used anisotropy parameters.	70
3.14	Soft seabed synthetic inversion results, in an environment whose anisotropy parameters are slightly erroneous. Inversion performed with the logarithm of the $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$ parameterization.	71
3.15	“Reflection coefficients as a function of incidence angle with varying values of Qp in the layer above the reflector. Other parameters are kept constant ($V_{p1}=2500$ m/s, $V_{s1}=1200$ m/s, $Q_{s1}=20$, $V_{p2}=3000$ m/s, $Q_{p2}=50$, $V_{s2}=1400$ m/s, $Q_{s2}=40$ where 1 denotes the medium above and 2 the medium below the reflector) [...]”. Figure and title from Belahi et al., 2015 (17).	72
3.16	Comparison of soft seabed synthetic inversion results with and without the inversion of intermediate scale SV-wave velocities update using the amplitude of P-P wide-aperture reflections (i.e., stage 2), and with the effect of the increasing frequency multiscale strategy (Bunks et al., 1995 (35)) as a substitute for stage 2.	72

3.17 Comparison of soft seabed synthetic inversion results with and without the inversion of intermediate scale SV-wave velocities update using the amplitude of P-P wide-aperture reflections (i.e., stage 2), and with the effect of the increasing frequency multiscale strategy (Bunks et al., 1995 (35)) as a substitute for stage 2. 73

4.1 Alwyn North field location. Grabens' location from Harker et al. (2003) (77). 76

4.2 Approximate lateral extension of Alwyn North field reservoirs, and acquisition used in this study. Petroleum data and location of the grabens from Harker et al. (2003) (77). OBN=Ocean Bottom Node. The blue rectangle on the regional map indicates the extent of the reservoirs map. 77

4.3 Migrated image below the source line used in this study, without (a) and with (b) interpreted horizons and faults (long dashed red lines), and approximate locations of reservoirs. The blurring in the middle of the shallow part comes from a lack of data due to a production platform. No depth exaggeration. Horizontal distance is measured along the source line, from the western most source. The reservoir locations are from Harker et al. (2003) (77). 78

4.4 Available seismic depth horizons, from top to bottom. BCU = Base Cretaceous Unconformity. The location of the source line used in this study is indicated in red. 79

4.5 Aspect of TrilobitsTM receivers (from SBGS trilobit specification sheet 2014). 81

4.6 a): Galperin configuration. b): Definition of radial and transverse components. G1, G2, G3 are the three geophone components (orange arrows). Receiver is on the seabed (green triangle) and source (red star) is on the surface. Blue arrows indicate the radial and transverse components in the horizontal plane. 81

4.7 Location of the acquisition area and planned acquisition geometry. Red dots are receivers. Note the two different colors used for the source lines (blue and grey colors inside the blue polygon) corresponding to a change of source line separation from 37.5 m to 50 m. Pipelines are annotated in green, with a 200 m safety limit annotated in purple. The black and white circles represent well heads. Source: modified from CGG internal communications (160919 Alwyn processing report draft). 82

4.8 Depth of the receivers along the survey. The receivers of the available swath are the receivers inside the blue rectangle. The green line correspond to the profile for this study. The color indicates the depth of the receivers from 122 m (blue) to 138 m (red). Source: modified from a figure of CGG (160919 Alwyn processing report draft). 83

4.9	CGG processing workflow applied on the available seismic data. The components to which each processing step is applied are specified: G1, G2, G3 are the three geophone components; P is the hydrophone component; Hx, Hy and Vz are respectively two orthogonal horizontal and the vertical downward velocity components; N, E, R and T are respectively the North, East, Radial and Transverse horizontal velocity components (as defined in figure 4.6b). DC = Direct Current. LC = Low-Cut. The question mark indicates that the processing step may not have been performed, or at least not at this place in the workflow.	84
4.10	Western most receiver gather of the profile and corresponding frequency spectra. Pink arrows: waterlayer multiples and guided waves; Red arrows: wrap-around noise; Blue arrow: short-offset noise; Green arrows: Scholte waves; Purple arrow: P-P reflection; Yellow arrow: P-S reflection; Orange arrows: inverse filter noise. Dashed green line: picked Scholte wave velocity. The same color scale is used for all velocity components. Extreme values of color scale are divided by 1000 to improve global visibility of events.	86
4.11	Low frequency panel for the western most pressure component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.	87
4.12	Low frequency panel for the western most vertical component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.	88
4.13	Low frequency panel for the western most radial component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.	89
4.14	Low frequency panel for the western most transverse component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.	90
4.15	Real data superimposed with the computed direct arrival times. Vertical velocity component of the western most receiver gather of the profile. Amplitudes have been clipped. Red line (resp. blue): times computed with 1478 m/s (resp. 1500 m/s).	91
4.16	Computed seabed S-wave velocities (in m/s) for three Scholte wave velocities, from left to right: lowest picked velocity (375 m/s), velocity from figure 4.10 (400 m/s), highest picked velocity (415 m/s). Various density and P-wave velocity seabed values are tested.	92
4.17	Initial model below the profile. The horizontal distance is computed along the profile from the western most source.	93
4.18	Seabed along the profile from: the provided initial 2D velocity model (blue), the depth of the receivers (green) and a seabed horizon (black). Horizontal distance is computed along the profile from the western most source.	94

4.19 Map of wells and sonic logs locations with respect to the profile (black line). The red stars correspond to the sources and the green markers correspond to the receivers. Each well path is represented by a different color and line style combination, as described in the legend. The solid circles correspond to the head wells location and the thick solid lines overlaid on the well paths correspond to the extent of the sonic logs measurements. 94

4.20 Wells and sonic logs locations on the vintage migrated section below the profile. Each well path is represented by a different color and line style combination, as described in the legend. The thick solid lines overlaid on the well paths correspond to the extent of the sonic logs measurements. Horizontal distance is computed along the profile from the western most source. 95

4.21 Smoothed density and P-wave velocities from a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The medium thick line corresponds to the initial S-wave vertical velocity model and the thick line corresponds to the initial P-wave vertical velocity model along the horizontal projection of the well on the profile. The colors of the line correspond to the horizontal distance between the well and the profile. 96

4.22 Smoothed density, P- and S-wave velocities (when available) from a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The medium thick line corresponds to the initial S-wave vertical velocity model and the thick line corresponds to the initial P-wave vertical velocity model along the horizontal projection of the well on the profile. The colors of both lines correspond to the horizontal distance between the well and the profile. 97

4.23 Smoothed density, P- and S-wave velocities (when available) from a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The medium thick line corresponds to the initial S-wave vertical velocity model and the thick line corresponds to the initial P-wave vertical velocity model along the horizontal projection of the well on the profile. The colors of both lines correspond to the horizontal distance between the well and the profile. 98

4.24 Western most receiver gather of the profile without (top) and with (bottom) the lines chosen to fit the headwaves. The solid lines correspond to the headwaves and the dashed lines correspond to the solid lines extension for all offsets. Vertical velocity component. Amplitudes are clipped for visualization purposes. 99

4.25 Sketch illustrating shear-wave splitting. Wave behavior without (left) and with (right) shear-wave splitting. Figure from Bale et al. (2013) (15). HTI = Horizontal Transverse Isotropy. 100

5.1	Description of the original soft seabed strategy and its variant, applied in a Vertical Transverse Isotropic (VTI) context to obtain P- and S-wave vertical velocities in soft seabed, the remaining anisotropy parameters not being updated. Note that the original soft seabed strategy contains three successive stages (i.e., stages 1, 2 and 3), while the variant contains only two stages (i.e., stages 1 and 3'). Indeed, the intermediate scale S-wave velocity update of stage 2 is replaced by the application of the increasing frequency multiscale strategy (Bunks et al., 1995 (35)) during stage 3'. Consequently, the variant strategy entirely relies on the low frequency data content, instead of on the amplitude of wide-aperture P-P reflections. The wide-aperture and S-wave related data are approximately selected with time-offset mutes. Note that it is advised to use the increasing frequency multiscale strategy at all stages, as is usually done for real data applications (e.g., Sears et al. (2010) (165)).	107
5.2	Vertical profiles in the middle of the smoothed (black line) and non-smoothed (colored line) initial models. (a) P-wave vertical velocity model; (b) S-wave vertical velocity model; (c) $\sqrt{1 + 2\delta}$ model; (d) $\sqrt{1 + 2\eta}$ model. Below the targets (which are between 3 and 5 kilometer depth), the deepest part of the models have been replaced by a vertically homogeneous layer for the absorbing boundary, and will not be shown in the next figures.	108
5.3	Density from the initial model (blue line) compared with smoothed density from well data (magenta line) for a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line indicates the horizontal distance between the well and the profile.	109
5.4	High-cut filter application on the western most pressure receiver gather. (a) Response amplitude frequency spectrum of the high-cut filter; (b) Initial data in the time-offset domain and corresponding amplitude frequency spectrum; (c) High-cut filtered data in the time-offset domain and corresponding amplitude frequency spectrum. Extreme values of color scales are divided by 1000 to improve global visibility of events.	110
5.5	“Analytical 2D and 3D time-domain Green’s functions convolved with a 50-Hz dominant-frequency Ricker wavelet, displayed for a single receiver which is 100 m away from a source, both sitting in a homogeneous full space. The 2D wavelet is larger in amplitude, exhibits a long tail, and is phase shifted by $\frac{\pi}{4}$ with respect to the 3D wavelet.” Figure and title from Auer et al. (2013) (14).	111
5.6	Geometrical spreading correction and low-cut filter application on the western most pressure receiver gather. (a) High-cut filtered data after geometrical spreading correction in the time-offset domain and corresponding amplitude frequency spectrum; (b) Data from a after low-cut filter; (c) Difference between the low-cut filtered and unfiltered data when no geometrical spreading correction is applied; (d) Response amplitude frequency spectrum of the used low-cut filter. Extreme values of color scales are divided by 1000 to improve global visibility of events.	112

5.7	FK application on the western most pressure receiver gather. (a) FK filtered data in the time-offset domain and corresponding amplitude frequency spectrum; (b) Difference between the FK filtered and unfiltered data (see figure 5.6b). Extreme values of color scales are divided by 1000 to improve global visibility of events.	113
5.8	Despiking on the seventh western most horizontal velocity component receiver gather. (a) Before despiking; (b) After despiking; (c) a-b; (d) Amplitude frequency spectra before (blue) after (orange) despiking. Extreme values of color scales are divided by 10 to improve global visibility of events.	114
5.9	Source time function inversion on Alwyn prepared data. (a) and (b) Western most observed and modeled pressure receiver gather respectively, the red polygons delineate the data used for the source inversion; (c) and (d) Initial, inverted and final source time function in the time and frequency domains respectively. Extreme values of color scales are divided by 10 to improve global visibility of events. The initial source time function is scaled for visualization purpose.	116
5.10	Normalized individual source time functions and corresponding amplitude frequency spectra for the western most pressure receiver gather. Before normalization, the difference between the highest and the lowest maximum value in the time domain is of a factor of 15.	117
5.11	Data from the western most pressure receiver gather selected for the source estimation, compared with modeled data using the inverted source function. Time-offset domain and frequency domain representations.	117
5.12	Updated P-wave vertical velocity model from the first stage, without gradient smoothing. The strong stripes are interpreted as the imprint of the sparse receiver geometry. The horizontal distance is computed along the profile from the western most source.	118
5.13	Variant strategy. Normalized cost function for each frequency range of each stage. (a) Stage 1a; (b) Stage 1b; (c) Stage 3'. The normalization is performed for visualization purpose, and the non-normalized cost functions are provided in Appendix I.	120
5.14	Variant strategy. Observed (O.) and initial modeled (M.) data represented in an alternate manner, for the first frequency range of each stage. (a) Stage 1a (pressure component); (b) Stage 1b (pressure component); (c) Stage 3' (horizontal velocity component). Western most receiver gather. The corresponding trace normalized wiggle plots are provided in Appendix I.	122
5.15	Variant strategy. Observed (O.) and modeled (M.) data represented in an alternate manner, at the beginning and end of each stage. (a) Stage 1a (pressure component); (b) Stage 1b (pressure component); (c) Stage 3' (horizontal velocity component). Western most receiver gather. The corresponding trace normalized wiggle and residual plots are provided in Appendix I.	123

5.16	Variant strategy. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above b. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The horizontal distance is computed along the profile from the western most source.	124
5.17	Variant strategy. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above b. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The horizontal distance is computed along the profile from the western most source.	125
5.18	Variant strategy. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	126
5.19	Variant strategy. (a) Initial P- over S-wave vertical velocity model, for comparison purpose; (b) P- over S-wave vertical velocity model; (c) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above b. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The color scale is clipped above 4.8 to increase visibility of reasonable ratios, maximum ratio is ~ 10.85 . The horizontal distance is computed along the profile from the western most source.	127
5.20	Original strategy. Normalized cost function for each frequency range of each stage. (a) Stage 2; (b) Stage 3. The normalization is performed for visualization purpose, and the non-normalized cost functions are provided in Appendix J.	128
5.21	Original strategy. Observed (O.) and modeled (M.) data fit represented in an alternate manner, at the beginning and end of each stage. (a) Stage 2 (pressure component); (b) Stage 3, first frequency range (horizontal velocity component, first frequency range only). Western most receiver gather. The corresponding trace normalized wiggle and residual plots are provided in Appendix J.	129
5.22	Original strategy. P-wave vertical velocity model inversion results. (a) Stage 2 initial model, provided for comparison purpose; (b) Stage 2 final model; (c) Stage 3 final model. The horizontal distance is computed along the profile from the western most source.	130
5.23	Original strategy. S-wave vertical velocity model inversion results. (a) Stage 2 initial model, provided for comparison purpose; (b) Stage 2 final model; (c) Stage 3 final model. The horizontal distance is computed along the profile from the western most source.	131

5.24	Original strategy with a modified stage order. Normalized cost function for each frequency range of each stage. (a) Stage 2'; (b) Stage 1b'. The normalization is performed for visualization purpose, and the non-normalized cost functions are provided in Appendix K	132
5.25	Original strategy with a modified stage order. Observed (O.) and modeled (M.) data fit represented in an alternate manner, at the beginning and end of each stage. (a) Stage 2' (pressure component); (b) Stage 3', first frequency range (horizontal velocity component, first frequency range only). Western most receiver gather. The corresponding trace normalized wiggle and residual plots are provided in Appendix K.	133
5.26	Original strategy with modified stage order. P-wave vertical velocity model inversion results. (a) Stage 2' initial model, for comparison purpose; (b) Stage 2' final model; (c) Stage 1b' final model. The horizontal distance is computed along the profile from the western most source.	134
5.27	Original strategy with modified stage order. S-wave vertical velocity model inversion results. (a) Stage 2' initial model, for comparison purpose; (b) Stage 2' final model. The horizontal distance is computed along the profile from the western most source.	135
5.28	Stage 1b first frequency range. P-wave vertical velocity model inversion results obtained with (a), and without (b) Scholte waves. The horizontal distance is computed along the profile from the western most source. . . .	136
5.29	Stage 1b first frequency range. final P- and S-wave vertical velocity models obtained with (thin red line) and without (thin blue line) Scholte waves compared with smoothed velocities from a selection of wells (thin black line for P-wave velocities and thin grey line for S-wave velocity). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	137
5.30	Stage 3' first frequency range. P-wave vertical velocity model inversion results obtained with (a), and without (b) Scholte waves; S-wave vertical velocity model inversion results obtained with (c), and without (d) Scholte waves. The horizontal distance is computed along the profile from the western most source.	138
5.31	Stage 3' first frequency range. Final P- and S-wave vertical velocity models obtained with (thin red line) and without (thin blue line) Scholte waves compared with smoothed velocities from a selection of wells (thin black line for P-wave velocities and thin grey line for S-wave velocity). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	139
5.32	Stage 1b first frequency range with Scholte waves. (a) Observed (O.) and initial modeled (M.) data represented in an alternate manner; (b) Final (F.) and initial (L.) modeled data represented in an alternate manner. Color scale is saturated for visualization purpose. Western most pressure receiver gather.	140

5.33	Stage 3' first frequency range with Scholte waves. (a) Observed (O.) and initial modeled (M.) data represented in an alternate manner; (b) Observed (O.) and final modeled (M.) data represented in an alternate manner. Color scale is saturated for visualization purpose. Western most horizontal velocity component receiver gather.	140
5.34	Variant strategy stopped after the first frequency range of the last stage, i.e. stage 3'. (a) Final P- over S-wave vertical velocity model; (b) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above a. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The color scale is clipped above 4.8 to increase visibility of reasonable ratios, maximum ratio is ~ 5.8 . The horizontal distance is computed along the profile from the western most source.	142
G.1	Alwyn North field in the Great Alwyn area. Source: TOTAL public communication (oil & gas UK share fair 31 october 2018).	167
H.1	P-wave vertical velocity models used for the synthetic source inversion tests. (a) True model; (b) Smooth model. Velocities in m/s.	175
H.2	Data used for the synthetic source inversion tests. Pressure receiver gather. (a) Observed data; (b) Modeled data with smooth model; (c) Modeled data with true model; (d) The used subset of modeled data.	176
H.3	Source time function inversion results. (a) and (c) Inversion performed with the true model only; (b) and (d) More realistic inversion results with the smooth model.	176
I.1	Variant strategy. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	178
I.2	Variant strategy. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	179
I.3	Stage 1a. Non-normalized cost function for each frequency range.	179
I.4	Stage 1a. Data inversion results. (a) Initial data fit restricted to the first frequency range [0-3 Hz]; (b) Initial data fit; (c) Initial residual; (d) Final data fit; (e) Final residual. For a, b and d, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.	180
I.5	Stage 1a. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	181

I.6	Stage 1a. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	182
I.7	Stage 1a. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	183
I.8	Stage 1a. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	184
I.9	Stage 1b: non-normalized cost function for each frequency range.	184
I.10	Stage 1b. Data inversion results. (a) Initial data fit restricted to [0-3 Hz]; (b) Observed data; (c) Initial data fit; (d) Initial residual; (e) Final data fit; (f) Final residual. For a, c and e, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.	185
I.11	stage 1b. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	186
I.12	Stage 1b. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	187
I.13	Stage 1b. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	188
I.14	Stage 1b. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	189
I.15	Stage 3'. Non-normalized cost function for each frequency range.	189

I.16	Stage 3'. Data inversion results. (a) Initial data fit restricted to the first frequency range [0-3 Hz]; (b) Initial data fit; (c) Initial residual; (d) Final data fit; (e) Final residual. For a, b and d, the traces are normalized and the observed data are in black while the modeled data are in red. Western most horizontal velocity component receiver gather.	190
I.17	Stage 3'. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	191
I.18	Stage 3'. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	192
I.19	Stage 3'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	193
I.20	Stage 3'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	194
I.21	Stage 3'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	195
J.1	Stage 2. Non-normalized cost function for each frequency range.	197
J.2	Stage 2. Data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.	198
J.3	Stage 2. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	199
J.4	Stage 2. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	200
J.5	Stage 2. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	201

J.6	Stage 2. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	202
J.7	Stage 2. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	203
J.8	Stage 3. Non-normalized cost function for the first frequency range.	203
J.9	Stage 3 first frequency range. Data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most horizontal velocity component receiver gather.	204
J.10	Stage 3 first frequency range. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	205
J.11	Stage 3 first frequency range. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	206
J.12	Stage 3 first frequency range. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	207
J.13	Stage 3 first frequency range. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	208
J.14	Stage 3 first frequency range. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	209
K.1	Stage 2?. Non-normalized cost function for each frequency range.	211
K.2	Stage 2?. Data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.	212

K.3	Stage 2'. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b. . . .	213
K.4	Stage 2'. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b. . . .	214
K.5	Stage 2'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	215
K.6	Stage 2'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	216
K.7	Stage 2'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	217
K.8	Stage 1b'. Non-normalized cost function for the first frequency range. . . .	217
K.9	Stage 1b' first frequency range. data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.	218
K.10	Stage 1b' first frequency range. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.	219
K.11	Stage 1b' first frequency range. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	220
K.12	Stage 1b' first frequency range. Initial (thin blue line) and final (thin red line) P-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.	221
K.13	Stage 1b' first frequency range. Initial (thin blue line) and final (thin red line) P-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations	222

List of Tables

- 1 Generation of S-waves depending on the environment, for local acquisitions. 4
- 1.1 Substitutions describing Voigt notation. Display inspired from Thomsen (1986) (*192*). 12
- 3.1 True model parameters. Linear gradients are defined by their extreme values. 65
- 3.2 Base model parameters. 68
- 4.1 Acquisition Geometry. 83
- 4.2 Time acquisition parameters. 83
- 4.3 Estimated thickness and P-wave horizontal velocity for each layer, as well as the time at zero offset of the line fitting the headwave traveling in the layer. The colors refer to the lines of figure 4.24. n.a. = not applicable. . . 99
- 4.4 Estimated zero-offset arrival times for P-P and P-S reflections on three horizons. 100
- G.1 Prior seismic imaging efforts made above Alwyn North field. PSTM = Pre-Stack Time Migration. PSDM = Pre-Stack Depth Migration. 168

Acronyms

- BCU** Base Cretaceous Unconformity. [xv](#), [xx](#), [xxii](#), [77](#), [79](#), [100](#), [101](#), [124](#), [125](#), [127](#), [142](#)
- BFGS** Broyden-Fletcher-Goldfarb-Shanno algorithm. [vii](#), [xxix](#), [9](#), [20](#), [22–24](#)
- CPML** Convolutional Perfectly Matched Layer. [16](#)
- DC** Direct Current. [xvi](#), [84](#), [85](#)
- FD** Finite-Differences. [xi](#), [13](#), [14](#), [16](#), [18](#)
- FE** Finite-Elements. [13](#)
- FK** Frequency-Wavenumber. [xix](#), [91](#), [92](#), [102](#), [113](#), [115](#), [119](#)
- FWI** Full Waveform Inversion, also called Full Wavefield Inversion. [iii–v](#), [xi](#), [xii](#), [2–7](#), [10–31](#), [34–36](#), [39](#), [41](#), [43](#), [49](#), [50](#), [52](#), [53](#), [56–58](#), [62–64](#), [66](#), [73–76](#), [83](#), [85](#), [92](#), [95](#), [101–103](#), [105–107](#), [110–115](#), [117](#), [119](#), [128](#), [131](#), [139–143](#), [145–147](#), [155](#), [161](#), [162](#)
- HC** HydroCarbon. [167](#)
- HTI** Horizontal Transverse Isotropy/Isotropic. [xvii](#), [100](#)
- L-BFGS** Limited memory BFGS. [24](#), [53](#), [151](#), [152](#)
- L-BFGS-B** Limited memory BFGS bounded. [viii](#), [22](#), [24](#), [25](#), [28](#), [151](#), [152](#)
- LC** Low-Cut. [xvi](#), [84](#)
- MPI** Message Passing Interface. [17](#), [19](#)
- OBC** Ocean Bottom Cable. [iv](#), [xi](#), [4](#), [5](#), [34](#), [145](#)
- OBN** Ocean Bottom Node. [iii–v](#), [xi](#), [xii](#), [xv](#), [4–7](#), [34–36](#), [40](#), [52](#), [77](#), [80](#), [81](#), [145](#), [146](#)
- OBS** Ocean Bottom Seismometer. [iii](#), [iv](#), [xi](#), [4–7](#), [34–36](#), [52](#), [145](#), [146](#)
- PML** Perfectly Matched Layer. [16](#)
- PR-CG** Polak-Ribière Conjugate Gradient. [24](#), [151](#)
- PSDM** Pre-Stack Depth Migration. [xxvii](#), [168](#)
- PSTM** Pre-Stack Time Migration. [xxvii](#), [168](#)

QC Quality Control. 76, 101–103

RF Recovery Factor. 80

SBGS Seabed Geosolutions. xv, 75, 76, 81, 85

SWS Shear wave Splitting, also called birefringence. 100, 101

TTI Tilted Transverse/Transversely Isotropy/Isotropic. 77, 101

VTI Vertical Transverse/Transversely Isotropy/Isotropic. iv, xi–xiii, xviii, 6, 7, 11–13, 23, 27, 30, 36, 37, 39, 43, 47–49, 52, 58, 62, 63, 68, 71, 73, 77, 93, 101, 106, 107, 141, 145, 146, 155, 157, 160, 165, 166, 171, 172

Introduction

Getting information about the composition, structure, and behavior of the Earth, as well as of other planets, has been a long standing focus. Generations of scientists have developed foremost techniques to image the sub-surface, constantly aiming for higher resolution and greater depths, for example. One of the information type the most investigated is the mechanical properties of the medium. Mechanical properties can be seen as the properties that govern the seismic waves propagation. In other words, a mechanical property is any part of a property set that allows to recover all of the following elements: the medium density, the velocities at which the medium allows the seismic P- and S-body waves to propagate, and the medium capacity to attenuate those seismic waves. Because mechanical properties show a high variability depending on the Earth media type (e.g., Schön, 2015 ([163](#))), their knowledge has proven very useful in various domains of geosciences such as lithology identification (e.g., Miller and Stewart, 1990 ([124](#))), fluid content and porosity characterization (e.g., Berryman et al., 2002 ([18](#))), or structural analysis (e.g., Audhkhasi and Singh, 2019 ([13](#))).

Among the methods developed to provide images of mechanical properties of the sub-surface, one can cite gravimetry. Gravimetry offers density models of deep and large portions of the Earth sub-surface from surface gravitational field measurements. However, as any potential method, the obtained models tend to show considerable ambiguities between the depth and the magnitude of the retrieved physical property (e.g., Skeels, 1947 ([177](#)); Johnson and van Klinken, 1979 ([87](#))). As a consequence, gravimetry is usually not self-sufficient, hence commonly used as a supplementary approach to other techniques, which often include seismological/seismic imaging methods (e.g., Arecco et al., 2016 ([7](#))).

Seismic data in themselves can provide information about the sub-surface, but they need processing (i.e., seismic imaging) to extract the most of them. Seismic migration is a major seismic imaging method, where the recorded reflections are re-located to their reflection point location, thereby creating a high-resolution, deep and large reflectivity image of the sub-surface. The reflection point location can be the zero-offset travel time (time migration) or the veritable spatial location (depth migration). Both types of migration only provide a qualitative image of a single mechanical property. In time migrations, the obtained images are distorted versions of the subsurface, in opposition with depth migrations. However, the success of the depth migration process is highly dependent on the quality of the used velocity models. Those velocity models can be obtained with the travel time tomography technique for instance.

Travel time tomography provides deep and large scale wave velocity models of the sub-surface. Starting from an initial model, travel time tomography iteratively minimizes a travel time misfit between observed and modeled seismic data. The diversity of used seismic data reduces ambiguities (e.g., depth and amplitude ambiguity) compared with gravimetry (e.g., Bickel, 1990 ([24](#)); Korenaga et al., 2000 ([101](#))). This data diversity

notably lies in a large range of offsets (i.e., the source-receiver horizontal distance), azimuths (i.e., the horizontal orientation of the source-receiver line), and in the use of both reflections and refractions. The restriction of the misfit to travel times explains why travel time tomography can only provide velocity models. Moreover, the implied travel times picking, through its manual nature, is time consuming and difficult. As a consequence, only a few arrivals are used leading to a limited amount of details in the retrieved models.

Full Waveform Inversion (also known as Full Wavefield Inversion or **FWI**) differs from travel time tomography by the involvement of the whole data waveform in the misfit (i.e., the travel time, the amplitude and the phase (i.e., shape)), see figure 1. As a consequence, all arrivals can be used and no travel time picking is needed, which leads to (1) a better illumination of the sub-surface, (2) even more reduced ambiguities, and (3) a better time management. Moreover, **FWI** is theoretically able to provide all mechanical properties, instead of velocities only, as well as higher resolution models (e.g., Arnulf et al., 2012 (10)). **FWI** even has the theoretical potential of reaching migration and even well logs resolution (e.g., Virieux and Operto, 2009 (212)).

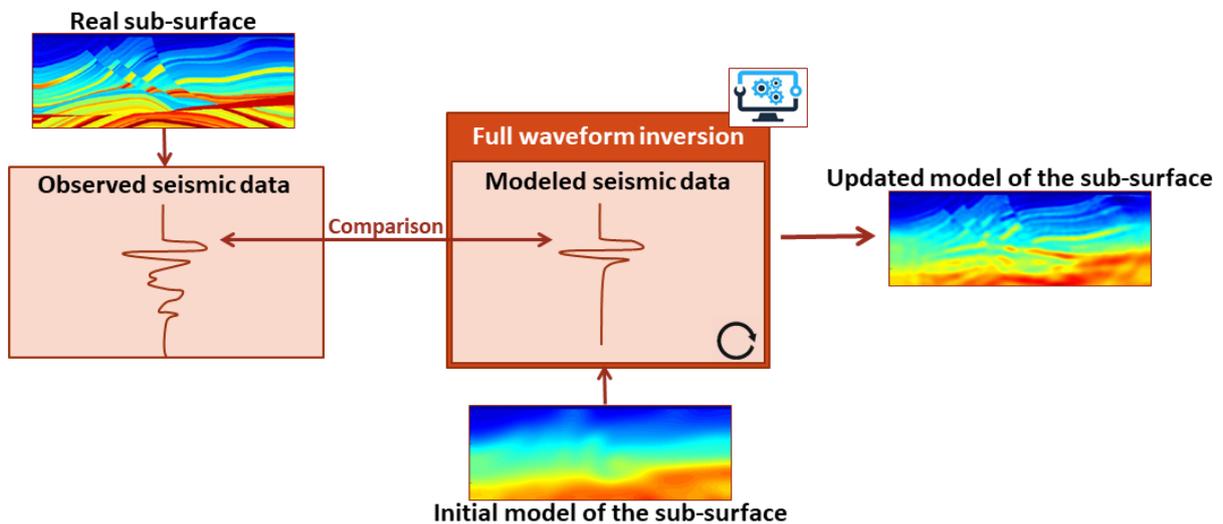


Figure 1: Schematic description of Full Waveform Inversion (**FWI**) algorithms where a single type of mechanical property is updated (P-wave vertical velocity here). In practice, it could be any number of property type. Note the iterative aspect of these algorithms. A single trace is represented but a multiplicity of them is actually used.

Well data can offer very high resolution densities and velocities through the use of gamma-rays and seismic waves induced by piezoelectric transmitters, or through the use of laboratory measurements on core samples. Those data can be used for applied investigations such as natural resources extraction as well as for pure research studies such as projects led in the framework of the International Ocean Discovery Program (IODP). However, these well data are confined along the well trajectory. Consequently, well data can be helpful to ground truth or constrain the results from **FWI** for instance (e.g., Reilly et al., 1984 (154); Oh et al., 2018 (140)), which offers a good balance between the size of the covered area and the obtained resolution.

The above-described advantages of **FWI** over comparable methods explain why **FWI** became an essential method in Earth sciences. However, even though **FWI** proved its usefulness, that method did not reach yet its full potential.

Indeed, the embedded large and numerous data modeling entail high computing costs. These high computing costs lead to the use of a local inversion algorithm for the **FWI** model update computation, which provides a local solution instead of the desired global one. Consequently, since the updated model is searched in the vicinity of the initial model, the resulting local solution is problematic for inversions with poor initial models (figure 2). The first step to overcome this issue consists in using good initial models. Such models are achievable to a limited extent through the preliminary use of more robust techniques, such as travel time tomography. The second step requires the design of advanced inversion strategies to compensate for the remaining insufficiency of the initial models. As a matter of fact, the higher the number of updated parameter types, the more a priori knowledge is needed in the initial model and the harder the design of a suitable inversion strategy. Despite the use of this local inversion algorithm, **FWI** is still limited by its computing needs. These computing costs increase with the choice of realism, especially for the elastic choice of realism.

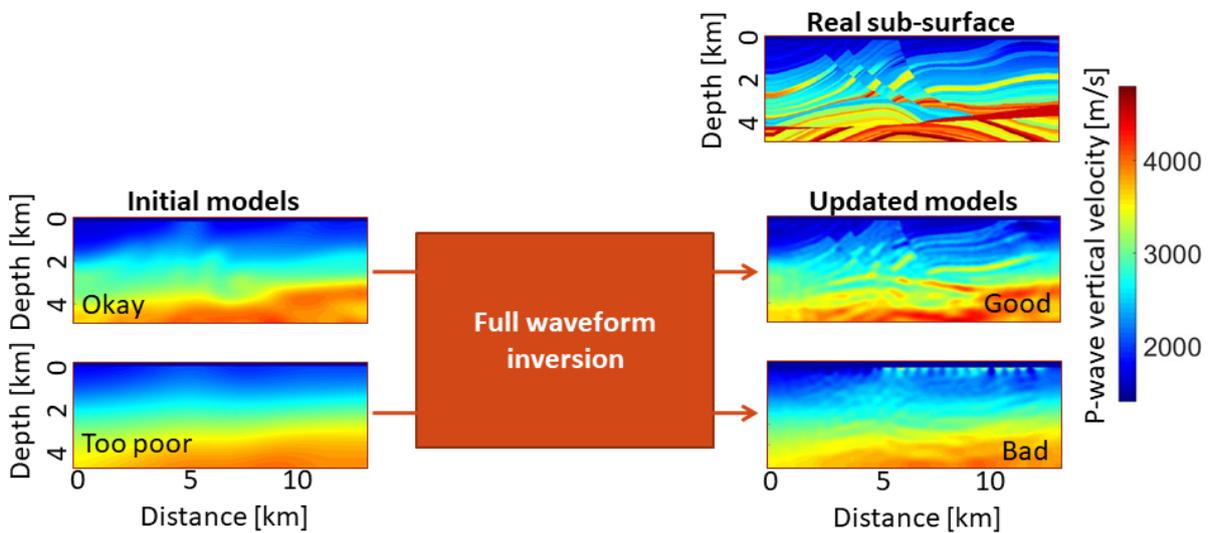


Figure 2: Synthetic applications illustrating the sensitivity of Full Waveform Inversion (**FWI**) to the initial model. Same inversion strategy is applied in both cases.

To summarize the above-described limitations, an increased choice of realism generates an increased number of parameter types and potentially an increased number of updated parameter types, both leading to increased (and sometimes prohibitively high) computing needs. Similarly, the increased number of updated parameter types leads to harder (and sometimes not yet achieved) suitable inversion strategy design.

As a consequence of these limitations, even though **FWI** was introduced for general media by Tarantola in 1988 (189), the modeling methods for local **FWI** applications (i.e., explored area with dimensions of a few kilometers (up to a dozen)) were initially mainly acoustic, with a single inverted parameter type: the P-wave velocity. This implies that the Earth is considered as a fluid (i.e., acoustic medium), which prevents the generation and propagation of S-waves, and not a solid (i.e., elastic medium), which allows for all kinds of waves. The data modeling differences between acoustic and elastic media not only consist in missing waves, but also in phase differences (e.g., Chapman et al., 2014 (40)), and in erroneous amplitudes. Indeed, the energy that should be accounted for the missing waves is conveyed to the non-missing waves.

Because of its partial solid character, the elastic nature of the Earth is indisputable and the proofs of S-waves are multiple (e.g., Knopoff, 1999 (100); Brossier et al., 2009 (31)). The missing waves and erroneous phases and amplitudes can lead to minimal to significant inversion errors, depending notably on the medium and acquisition properties (e.g., Barnes and Charara., 2009 (16); Marelli et al., 2012 (117)). Besides, S-wave velocity models also have the potential to notably provide improved lithology characterization by offering a complementary perspective on the medium (e.g., Szymańska-Majysa and Dubiel, 2019 (185)).

As recalled in recent articles (e.g., Kayum, 2020 (91)), extensive efforts have been made to increase computing capacities in general and to decrease FWI computing needs. Benefiting from this progress, the acoustic approximation is more and more replaced by the elastic one, in order to improve the P-wave velocity update and obtain meaningful S-wave velocity models. Several successful P- and S-wave FWI applications have been made on land (e.g., Shi et al., 2007 (173); Brossier et al., 2009 (31); Vassiliou et al., 2017 (205); Ren et al., 2019 (155)). In marine environments however, the inversion suffers from a poorer S-wave velocity information content, because of a decreased amount of S-waves (table 1). Indeed, a land vibrator, for example, allows for the generation of both P- and S-waves, while the usual airgun source used in marine environments cannot lead to the generation of S-waves, since this source acts on fluid media. Consequently, in marine environments, the inversion of S-wave velocities mainly relies on P to S conversions, notably at the seabed (table 1). Those conversions allow for successful inversion results in strong seabed environments (i.e., seabed S-wave velocities > 1000 m/s) (e.g., Choi and Shin, 2007 (42); Mora and Wu, 2018 (132)). For soft seabeds environments (i.e., seabed S-wave velocities < 300 m/s, typical of siliclastic seabed sediments, see Hamilton, 1976 (74), and Castagna et al., 1985 (39)), however, P to S seabed conversions are less strong and the corresponding S-waves arrive at later times (Sears et al., 2008 (164); Barnes and Charara, 2009 (16)), see table 1. This lower amount of information explains the failing of usual inversion strategies. Yet, Sears et al. (2008) (164) recently proposed the only published successful inversion strategy in soft seabed context (hereafter dubbed the “soft seabed strategy”), which was applied on both synthetic and real 2D data (Brossier et al. 2009 (32); Brossier et al. 2010 (33); and Sears et al. 2010 (165); Oh et al. 2018 (140) respectively). However, this strategy suffers from limitations.

S-wave source \ Environment	Land	Marine	
		Strong seabed	Soft seabed
Seismic source device	Yes	No	No
P to S conversions at the seabed	Not applicable	Yes	Yes but poor
Other P to S conversions	Yes	Yes	Yes

Table 1: Generation of S-waves depending on the environment, for local acquisitions.

A first limitation is related to the fact that the soft seabed strategy relies on multi-component ocean bottom data availability. Ocean bottom recording systems (i.e., the Ocean Bottom Cables and Nodes/Seismometers, or OBCs, and OBNs/OBSs respectively) are becoming more and more popular, replacing streamers which only record pressure in water below the sea-surface. Indeed, ocean bottom systems allow for the recording of both pressure in water and velocity components at the seabed, including the direct recording of S-waves, which is essential for the success of the soft seabed strategy. This strategy was illustrated on OBC acquisitions, which are nowadays increasingly replaced by OBNs/OBSs

acquisitions. Indeed, the independence of **OBNs/OBSs** receivers compared with **OBCs**' (see figure 3) allows for the needed large offsets, wide azimuth, and near platform acquisitions for example. However, **OBN/OBS** acquisitions contain many more sources than **OBC** acquisitions, but also many more sources than receivers. The traditional **FWI** computing cost increases with the number of sources and is mostly independent on the number of receivers. This leads to increased and prohibitive computing costs for **FWI** applications with **OBN/OBS** data, compared with **OBC** data. To decrease those computing costs, the principle of reciprocity is usually applied. Reciprocity allows to interchange the role of receivers and sources in modeling, and to benefit from the lower number of receivers. However, the available documentation about such reciprocity in **FWI** needs to be improved.

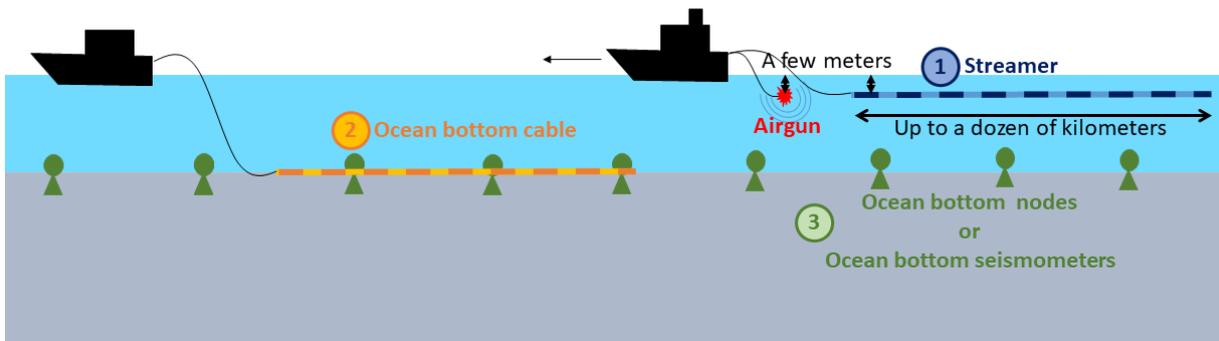


Figure 3: Description of the usual source system (airgun) and the three usual recording systems (streamer, Ocean Bottom Cable and Node/Seismometer (**OBC** and **OBN/OBS**)) in the context of local marine acquisitions. The sketch is not up-to-scale but aims at highlighting the different scales of these acquisition systems, as well as their location properties. Typically, **OBNs/OBSs** are spaced by a few hundreds meters and cover areas of a few dozens of kilometers.

A second limitation lies with the fact that the soft seabed strategy was developed assuming an isotropic medium (i.e., a medium exhibiting an invariance of the waves propagation velocity with the direction of propagation, in opposition to an anisotropic medium). However, the isotropic modeling of the Earth is mostly inadequate (e.g., Levin and Park, 1998 (109); Thomas and Kendall, 2002 (190); Ferreira et al., 2019 (56); Mattesini et al., 2010 (120); Song, 1997 (181); Tromp, 1993 (196)). Anisotropy can be due to the preferred orientation of intrinsically anisotropic constituents (olivine for instance, the primary mineral of Earth's upper mantle (e.g., Ismaïl and Mainprice, 1998 (85))), thus called intrinsic anisotropy. Anisotropy can also be due to extrinsic phenomena, such as scale phenomena as emphasized by Wang et al. (2013) (217) for example. In all cases, a wrong anisotropy can have a significant impact on the wave travel times. It also has an impact on amplitudes and phases. Besides, since increasingly wider offsets and/or azimuth data are acquired to diversify information, the data include an increasingly wider range of wave propagation directions. This urges the use of anisotropy in **FWI** (e.g., Prioux et al., 2011 (146)). In the interest of accuracy, let us mention that in anisotropic media, the properties of body waves are slightly different from isotropic media. Body waves are then called qP, qSV and qSH waves, where q means quasi. In this thesis, I omit q to simplify the notations.

As described above, **FWI** stands out through both its current capabilities and potential, among the methods that provide quantitative images of the sub-surface. In order to reach the maximum potential of this method and meet modern application needs, several limitations have to be overcome. This thesis aims at reducing two of these limitations: (1) extend the application of the soft seabed strategy to the more contemporary **OBN/OBS** recording devices, and (2) partially expand the strategy to anisotropic media.

To do so, I first propose to build an extensive guide of the above-described use of reciprocity in **FWI**, in order to allow for the use of **OBN/OBS**. This guide aims at compensating the scattered, non-unified and limited aspect of the current available documentation, which notably poorly emphasizes the reciprocity limitations, the used thought process, and the justification of the theoretical equality between **FWI** with and without reciprocity. For this reciprocity study, I limit the study to a specific kind of anisotropic environment: the Vertical Transverse Isotropic (**VTI**, also called vertical transversely isotropic) environments. **VTI** is an intermediate stage between isotropic and fully anisotropic media, where wave velocities present a symmetry around the vertical axis, as described in figure 4. This kind of anisotropy is an ubiquitous reality, especially in siliclastic sedimentary basins with horizontal shale layers, or fractured, finely layered, or clay bearing horizontal sand layers for example (e.g., Wang, 2002 (223); Cholach and Schmitt, 2003 (44); Kendall et al., 2007 (93)).

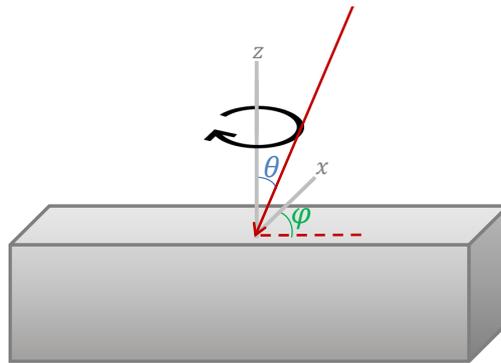


Figure 4: Vertical Transverse Isotropic (**VTI**) medium: the velocity of the wave in red depends only on the angle θ with the vertical axis z , and not the angle φ with the horizontal axis x .

Second, I propose to adapt and assess the behavior of the soft seabed strategy when applied to those elastic **VTI** environments. To be more precise, the performance of the soft seabed strategy to retrieve the vertical velocities will be assessed when slightly inaccurate remaining anisotropy parameters are used, these parameters not being updated. Such a scenario is common for real case applications.

The synthetic and theoretical findings are finally put into practice and tested with a real 2D application from the Alwyn North hydrocarbon field (North Sea). The need for a confirmation on real data is dictated by numerous aspects that were not taken into account during the synthetic and theoretical studies. Among those aspects, one can cite noisy data, poorly known sources/acquisition geometry or physical approximations.

This thesis is articulated in 5 chapters: **Chapter 1** provides the theoretical framework, implementation and application details about **FWI**, which will be necessary for deeper

understanding of the rest of this research project. I explain in greater details the main current **FWI** limitations, being the limiting computing costs and the need for advanced inversion strategies. In this chapter, I also emphasize the major role of **FWI** inputs and the resultant need for: (1) their prior analysis and (2) a thorough **FWI** results assessment.

Chapter 2 covers the extension of the soft seabed strategy to **OBN/OBS** data using reciprocity. After explaining the need for reciprocity, I present a review of related publications. Then, I provide insights to such a reciprocity application, focusing on elastic **VTI** media, and using theoretical developments and synthetic tests. To be more precise, limitations and some extensions of the gathered pre-existing needed reciprocity relations are derived, and the thought process consisting in using reciprocity in **FWI** model update computations is explained, as well as its causes and consequences on said update. In this chapter, I also highlight the role of reciprocity as a debugging/testing technique for modeling and **FWI** algorithms.

Chapter 3 covers the partial expansion of the soft seabed strategy to **VTI** media. First, I detail and explain the strategy of Sears et al. (2008) (164) to retrieve both P- and SV-waves velocities in soft seabed environments. At the same time, I emphasize the need for such strategy, provide a review of related publications, and highlight the limitations of the context in which the strategy was developed (especially the isotropic aspect). Then, I discuss the choice of parameterization, choice imposed by the change of choice of realism. The parameterization selection is based on analytically computed reflection radiation patterns. I subsequently synthetically illustrates the behavior of the soft seabed strategy in elastic **VTI** media to retrieve the vertical velocities, when slightly inaccurate remaining anisotropy parameters are used, these parameters not being updated. The illustration is preceded by a partial prediction of the behavior through reflection coefficients computations. Finally, an extension of the strategy for applications where the amplitudes cannot be properly modeled (in attenuating media for example) is also proposed.

Chapter 4 and **Chapter 5** are, respectively, the first and second step of the application of 2D elastic **FWI** in **VTI** environment on Alwyn real data. Indeed, this real application is divided into two parts. First, I provide an analysis and description of the available data and **FWI** inputs. Then the **FWI** application is detailed, putting into practice and testing the synthetic and theoretical findings from previous chapters, with a thorough control of the inversion results.

Finally, in the last part of this manuscript, I summarize the project and propose directions for future work.

Chapter 1

Full waveform inversion: overview, theory and implementation

Chapter's objectives and highlights

- Explain the existence of the two main limitations of full waveform inversion, i.e.: limiting computing costs and need for specific inversion strategies.
 - Emphasize the major role of the full waveform inversion inputs and the resultant need for their prior analysis and for a thorough full waveform inversion results verification.
 - Provide a consistent, broad and detailed description of full waveform inversion that would allow for the deep understanding of the rest of the thesis.
-

Contents

1.1	Introduction	10
1.2	Forward problem	11
1.2.1	A choice of wave equation: a first-order, time-space domain, 2D elastic vertical transverse wave equation	11
1.2.2	A choice of solving method: a finite-difference method	13
1.3	Inverse problem	20
1.3.1	A choice of objective function: an ℓ_2 -norm cost function	20
1.3.2	A choice of inversion algorithm: a conjugate-gradient and a BFGS algorithm	22
1.3.3	The gradient computation	26
1.4	Inversion strategy	28
1.4.1	Resolution aspects	28
1.4.2	Driving full waveform inversion towards the intended optimum	28
1.4.3	About multiparameter inversion	30
1.5	Discussion and conclusions	31

1.1 Introduction

Full Waveform Inversion (also known as full wavefield inversion or FWI) is a class of algorithms for obtaining models of mechanical parameters of the sub-surface from seismic data. They iteratively update an initial model by minimizing a misfit between observed and modeled seismic data. The misfit is quantified by an objective function and involves the whole waveform of the seismic data, instead of the travel-time only as in travel-time tomography algorithms, explaining the origin of the word “full waveform inversion”. Figure 1.1 describes the general structure of FWI algorithms. This figure emphasizes the fact that FWI can be divided into two parts: the forward problem which provides modeled data from a model, an acquisition geometry and a source function and the inverse problem which updates a model using a misfit between modeled and observed data.

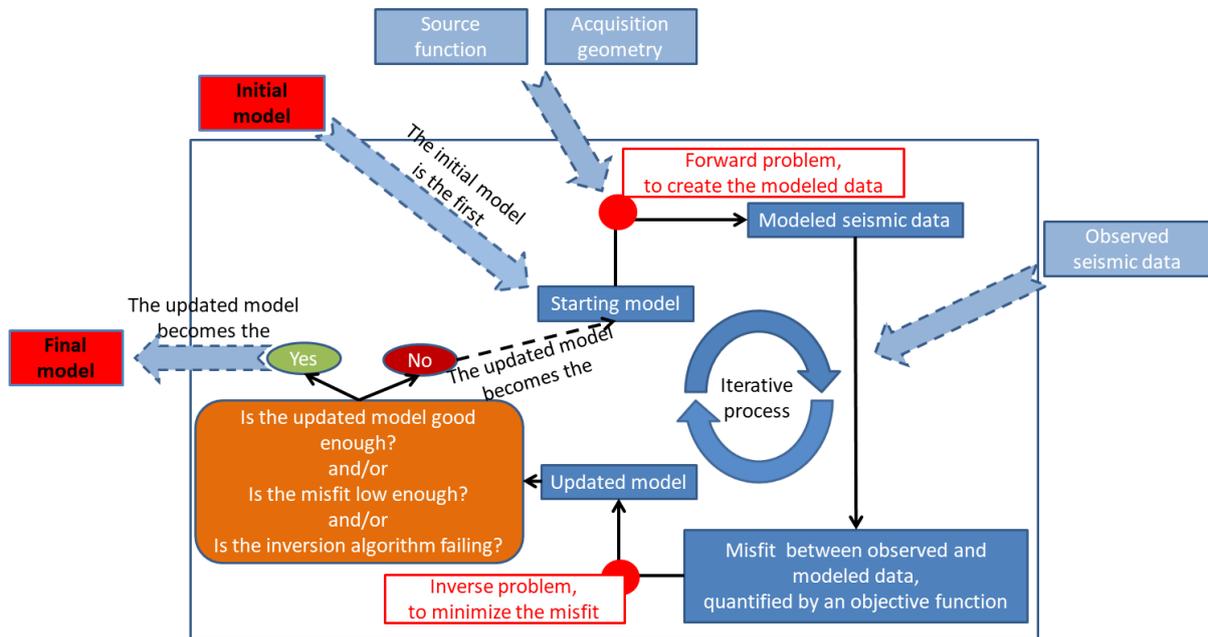


Figure 1.1: General structure of Full Waveform Inversion FWI algorithms.

FWI was introduced by Tarantola in 1984 (187) and has been used for several decades in different situations. For example, Rickers et al. (2013) (156) applied FWI on earthquake data and obtained a regional S-wave velocity model indicating the presence of a plume system in the North Atlantic region. Marjanović et al. (2019) (118) applied 3D FWI on controlled airgun sources data to obtain a local model of P-wave velocities at the East Pacific Rise. Wang et al. (2019) (220) used 2D FWI for tunnel detection from land velocity model. Ratcliffe et al. (2011) (149) used 3D FWI to improve a marine migrated image for oil and gas exploration and production, by offering a better velocity model. Hicks et al. (2016) (82) applied 4D FWI (also called time-lapse FWI) for reservoir monitoring, where data are repeatedly acquired and inverted over time to monitor the evolution of the sub-surface during production. Karakostas et al. (2019) (89) even performed synthetic FWI tests on Rayleigh waves for future applications on Mars planet. As illustrated by these applications, these decades of application led to different flavors of FWI. The inversion process usually involves several calls to these flavors. The choice of the strategy can consequently be considered as a third aspect of FWI.

The objectives of this chapter are multiple: from justifying the FWI limiting comput-

ing costs and the need for specific inversion strategy, to providing the necessary details for a deep understanding of subsequent chapters as well as an FWI overview, via emphasizing the major role of the FWI inputs and the resultant need for their prior analysis and for a thorough FWI results verification. The objectives are fulfilled by providing the physical, mathematical and implementation details of the FWI algorithm used in this Ph.D. thesis, while setting these implementation choices into a bigger picture of FWI, and mentioning some major inversion strategy aspects. One can refer to Virieux and Operto (2009) (212) or Tromp (2020) (195) for more detailed overviews of FWI. The FWI algorithm used in this work has initially been developed by Weiguang He as part of his Ph.D. thesis (He et al., 2016 (78); He, 2017 (80); He et al., 2018 (79)), and has been further improved during this Ph.D. thesis, as specified later in this chapter.

I first address the forward problem, then, I focus on the inverse problem, to finally describe some inversion strategies.

1.2 Forward problem

As mentioned in the introduction of this chapter, FWI requires solving a forward problem in order to provide modeled data from a model, an acquisition geometry and a source function. It implies the choice of a wave equation and a solving method. I detail here the first-order, time-space domain, 2D elastic and Vertical Transverse Isotropic (VTI) wave equation used in our FWI algorithm, as well as the finite-difference method used to solve it. In order to add more perspective, some other possible choices are mentioned: I notably derive the chosen equation from a more general wave equation.

1.2.1 A choice of wave equation: a first-order, time-space domain, 2D elastic vertical transverse wave equation

The propagation of seismic waves in a medium can be described using a wave equation with appropriate initial and boundary conditions. Assuming an inertial frame of reference, from Newton's second law and generalized Hooke's law, the second-order, time-space domain, 3D elastic anisotropic wave equation for a heterogeneous non-attenuating medium can be written as:

$$\begin{cases} \rho(\mathbf{x}, t) \dot{v}_i(\mathbf{x}, t) &= \sigma_{ij,j}(\mathbf{x}, t) + f_i(\mathbf{x}, t), \\ \dot{\sigma}_{ij}(\mathbf{x}, t) &= C_{ijkl}(\mathbf{x}) v_{k,l}(\mathbf{x}, t) + \dot{m}_{ij}(\mathbf{x}, t), \end{cases} \quad (1.1)$$

where the Einstein convention of summing on repeated indices is used. The possible values for each subscripts i, j, k and l are $\{x, y, z\}$, with x and y representing the horizontal axes and z representing the vertical axis. Hereafter, the Einstein convention will not apply to the explicit indices x, y , and z . t is the time and \mathbf{x} is a position vector made up of three components. \cdot_j indicates a partial spatial differentiation along j and $\dot{\cdot}$ indicates a first order partial time derivative. v_i is the i^{th} component of the velocity field (unit of length per unit of time) and σ_{ij} is the ij^{th} component of the stress tensor (force per unit of surface). ρ is the density (mass per unit of volume) and C_{ijkl} are the components of the 4th order elasticity tensor (i.e., the elastic moduli) -or stiffness tensor- introduced in the generalized Hooke's law (mass per unit of length and per squared unit of time). Together with density, the elastic moduli describe the mechanical properties of the sub-surface. Note that the use of Hooke's law implies linear elasticity, experimentally valid for small deformations. The components of the elasticity tensor are positive, exhibit symmetries

($C_{jikl} = C_{ijkl} = C_{ijlk} = C_{jilk}$) (see Aki and Richards, 2002 (3) for instance) and can be represented more compactly by a 6*6 matrix using Voigt notation, regrouping the first and second pair of indices as explained in table 1.1. f_i and m_{ij} are the components of external forces. f_i is the i^{th} component of an external body force source (force per unit of volume). m_{ij} are the components of an external moment density source (force per unit of surface, i.e. stress). Note that mathematically, a moment density source can be replaced by a body force.

11	22	33	23=32	13=31	12=21
↓	↓	↓	↓	↓	↓
1	2	3	4	5	6

Table 1.1: Substitutions describing Voigt notation. Display inspired from Thomsen (1986) (192).

Equation 1.1 can be written in various domains, e.g.: the time-space, frequency-space, time-wave number or frequency-wave number domains. Moreover, equation 1.1 is said to be a first-order wave equation because it involves first order partial time derivatives. Other formulations are possible, such as the displacement-stress formulation (in such a case, the equation involves second order partial time derivatives and is then a second-order wave equation). Our FWI algorithm uses the time-space domain with the velocity-stress formulation (introduced in seismic imaging in 1976 by Madariaga (114) to model fault-rupture dynamic). Consequently, the description provided here is restrained to this domain and formulation.

While equation 1.1 corresponds to very general media (heterogeneous elastic anisotropic media), as explained in the introductory chapter, I focus in this Ph.D. on the elastic VTI choice of realism. I recall that VTI media are characterized by wave velocities having a symmetry around the vertical axis (Thomsen, 1986 (192)). It is therefore no surprise that the less general elastic VTI wave equation was used, by using the following simplified elasticity tensor:

$$\begin{bmatrix} C_{11}(\mathbf{x}) & C_{11}(\mathbf{x}) - 2C_{66}(\mathbf{x}) & C_{13}(\mathbf{x}) & 0 & 0 & 0 \\ C_{11}(\mathbf{x}) - 2C_{66}(\mathbf{x}) & C_{11}(\mathbf{x}) & C_{13}(\mathbf{x}) & 0 & 0 & 0 \\ C_{13}(\mathbf{x}) & C_{13}(\mathbf{x}) & C_{33}(\mathbf{x}) & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44}(\mathbf{x}) & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44}(\mathbf{x}) & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66}(\mathbf{x}) \end{bmatrix}. \quad (1.2)$$

The real world being 3D, the general wave equation 1.1 describes the propagation of waves in 3D media. However, the computing resources are limited and the computing cost of the forward problem usually decreases with the dimensionality. The acquisition geometry and the heterogeneity of the medium can allow for the use of a lower dimensionality. For example, in a 3D VTI medium with enough homogeneity along one horizontal dimension and an airgun acquisition along the perpendicular horizontal dimension, the wave propagation can be described by a 2D wave equation, provided that a geometrical spreading correction can be applied on the observed data at the preparation step. As the waves move away from the source, the area that the wave energy covers becomes larger and wave amplitude decreases. This is referred to as “geometrical spreading”. For the same source function, the geometrical spreading of a 3D (resp. 2D) algorithm and a 2D

(resp. 1D) algorithm is not the same, as illustrated by Auer et al. (2013) (14) for instance. Hence the need for a correction. Other dimensionalities are possible: 1D, 1.5D, and 2.5D. 2.5D (resp. 1.5D) wave-equations reproduce the 3D (resp. 2D) geometrical spreading. It is more costly than the inferior dimension, but requires no geometrical spreading correction, which is usually approximate. In our case, the computing resources are seriously limited by the elastic choice of realism (because of modeling stability criteria as explained later in this section), while Alwyn's environment and acquisition seems to allow for 2D modeling (see Chapter 4). It is therefore no surprise that a 2D wave equation is used.

Integrating in equation 1.1 all the mentioned modeling choices, the chosen wave equation (i.e., the wave equation used in our FWI algorithm) is the following first-order, time-space domain, 2D elastic VTI wave equation:

$$\begin{cases} \rho(\mathbf{x})\dot{v}_x(\mathbf{x}, t) &= \sigma_{xx,x}(\mathbf{x}, t) + \sigma_{xz,z}(\mathbf{x}, t) + f_x(\mathbf{x}, t), \\ \rho(\mathbf{x})\dot{v}_z(\mathbf{x}, t) &= \sigma_{xz,x}(\mathbf{x}, t) + \sigma_{zz,z}(\mathbf{x}, t) + f_z(\mathbf{x}, t), \\ \dot{\sigma}_{xx}(\mathbf{x}, t) &= C_{11}(\mathbf{x})v_{x,x}(\mathbf{x}, t) + C_{13}(\mathbf{x})v_{z,z}(\mathbf{x}, t) + \dot{m}_{xx}(\mathbf{x}, t), \\ \dot{\sigma}_{zz}(\mathbf{x}, t) &= C_{13}(\mathbf{x})v_{x,x}(\mathbf{x}, t) + C_{33}(\mathbf{x})v_{z,z}(\mathbf{x}, t) + \dot{m}_{zz}(\mathbf{x}, t), \\ \dot{\sigma}_{xz}(\mathbf{x}, t) &= C_{44}(\mathbf{x})[v_{x,z}(\mathbf{x}, t) + v_{z,x}(\mathbf{x}, t)] + \dot{m}_{xz}(\mathbf{x}, t). \end{cases} \quad (1.3)$$

v_i is the i^{th} component of the velocity field (unit of length per unit of time), and σ_{xx} and σ_{zz} (resp. σ_{xz} and σ_{zx}) are the normal (resp. shear) components of stress. Note that our modeling algorithm also imposes $\dot{m}_{xz}(\mathbf{x}, t) = \dot{m}_{zx}(\mathbf{x}, t)$.

1.2.2 A choice of solving method: a finite-difference method

For simple cases - such as a point body force source (i.e., a body force source localized in a single point) in a homogeneous unbounded medium - an analytical solution of equation 1.3 can be retrieved. Otherwise, a solution can be provided by various classes of numerical methods. The most common are the Finite-Difference and Finite-Element methods (resp. FD and FE). In comparison with FE methods, FD ones can be simpler to understand and to implement, but can show limitations with rough seabeds or strong interfaces in general. The seabed above Alwyn North field being almost flat (see Chapter 4), FD methods seem suitable. However, since those methods cannot reproduce sharp interfaces, their accuracy is questionable there, particularly for seabed data. This question is left for another study. One can refer to Virieux et al. (2012) (213) for more details about these two kinds of methods.

In FD methods, the seismic quantities (here the stress and velocity fields) and medium parameters are spatially and time discretized and the partial derivatives are approximated with the help of Taylor expansions. The time discretization used in our FWI algorithm is a leap-frog scheme between stress and velocity components (i.e., interleaved time points for stress and velocity components), with a spatial and a time spacing of Δh and Δt respectively. The spatial discretization used in the algorithm corresponds to the rectangular staggered grid described in figure 1.2. Other kinds of discretizations are possible. For example, when Kelly et al. (1976) (92) introduced FD methods for a wave equation, a second-order wave equation was used, all displacement components were computed at the same locations and the stress components were not computed. Later, Virieux (1984 and 1986) (214) (215) proposed a full staggered grid to solve the first-order wave equation. This modeling method allows for easier free-surface boundaries and stress sources implementation, as well as a better stability in the presence of fluid-solid interfaces, which are

needed for marine data (such as Alwyn real data, see Chapter 4). Note that the spatial step does not have to be the same for all directions, and that more complex FD scheme with variable grid (along one or more direction) are also possible (see for instance Wang et al., 1996 (222)).

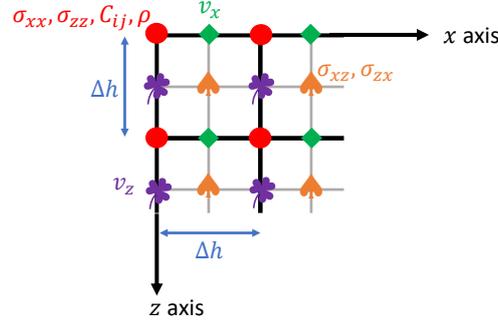


Figure 1.2: Implemented staggered grid. Δh is the spatial spacing.

The algorithm uses a 2^d -order centered Taylor expansion in time and a 4^{th} or a 8^{th} -order centered Taylor expansion in space. Other kinds of Taylor expansions are possible. For example, Virieux (1984 and 1986) (214) (215)) used a 2^d -order time and spatial Taylor expansions. Dablain (1986) (48) and Levander (1988) (110) used higher order spatial approximations (and sometimes even higher order time expansions). These higher order approximations minimized dispersion errors for large simulation times without the need for very fine modeling grids, decreasing the memory needs without increasing the computation time. The 2^d order centered Taylor expansion in time of a function $g(t)$ is:

$$\dot{g}(t) \approx \frac{g(t + \frac{\Delta t}{2}) - g(t - \frac{\Delta t}{2})}{\Delta t}. \quad (1.4)$$

Let us rewrite vector \mathbf{x} with (x, z) when necessary. The 4^{th} and 8^{th} -order centered Taylor expansion in space of a function $g(x, z)$ are:

$$g_{,x}(\mathbf{x}) \approx \sum_{k=0}^K \frac{c_k}{\Delta h} \left[g\left(x + \Delta h\left(k + \frac{1}{2}\right), z\right) - g\left(x - \Delta h\left(k + \frac{1}{2}\right), z\right) \right], \quad (1.5)$$

with $K = 1$, $c_0 = \frac{9}{8}$ and $c_1 = -\frac{1}{24}$ for the 4^{th} -order Taylor expansion, and $K = 3$, $c_0 = \frac{1225}{1024}$, $c_1 = -\frac{245}{3072}$, $c_2 = \frac{49}{5120}$ and $c_3 = -\frac{5}{7168}$ for the 8^{th} -order Taylor expansion. A similar equation can be obtained for $g_{,z}$.

By combining approximate time and spatial derivatives 1.4 and 1.5 with the chosen

staggered grid, wave equation 1.3 becomes:

$$\left\{ \begin{array}{l}
 v_x(x + \Delta h, z, t + \frac{\Delta t}{2}) \approx v_x(x + \Delta h, z, t - \frac{\Delta t}{2}) \\
 \quad + \frac{\Delta t}{\rho(x+\Delta h, z)} \left(\sigma_{xx, x}(x + \Delta h, z, t) + \sigma_{xz, z}(x + \Delta h, z, t) \right. \\
 \quad \left. + f_x(x + \Delta h, z, t) \right), \\
 v_z(x, z + \Delta h, t + \frac{\Delta t}{2}) \approx v_z(x, z + \Delta h, t - \frac{\Delta t}{2}) \\
 \quad + \frac{\Delta t}{\rho(x, z+\Delta h)} \left(\sigma_{xz, x}(x, z + \Delta h, t) + \sigma_{zz, z}(x, z + \Delta h, t) \right. \\
 \quad \left. + f_z(x, z + \Delta h, t) \right), \\
 \sigma_{xx}(\mathbf{x}, t + \Delta t) \approx \sigma_{xx}(\mathbf{x}, t) \\
 \quad + \Delta t \left(C_{11}(\mathbf{x})v_{x, x}(\mathbf{x}, t + \frac{\Delta t}{2}) + C_{13}(\mathbf{x})v_{z, z}(\mathbf{x}, t + \frac{\Delta t}{2}) \right. \\
 \quad \left. + \dot{m}_{xx}(\mathbf{x}, t + \frac{\Delta t}{2}) \right), \\
 \sigma_{zz}(\mathbf{x}, t + \Delta t) \approx \sigma_{zz}(\mathbf{x}, t) \\
 \quad + \Delta t \left(C_{13}(\mathbf{x})v_{x, x}(\mathbf{x}, t + \frac{\Delta t}{2}) + C_{33}(\mathbf{x})v_{z, z}(\mathbf{x}, t + \frac{\Delta t}{2}) \right. \\
 \quad \left. + \dot{m}_{zz}(\mathbf{x}, t + \frac{\Delta t}{2}) \right), \\
 \sigma_{xz}(x + \Delta h, z + \Delta h, t + \Delta t) = \sigma_{zx}(x + \Delta h, z + \Delta h, t + \Delta t) \\
 \approx \sigma_{xz}(x + \Delta h, z + \Delta h, t) \\
 \quad + \Delta t \left(C_{44}(x + \Delta h, z + \Delta h)[v_{x, z}(x + \Delta h, z + \Delta h, t + \frac{\Delta t}{2}) \right. \\
 \quad \left. + v_{z, x}(x + \Delta h, z + \Delta h, t + \frac{\Delta t}{2}) \right] \\
 \quad \left. + \dot{m}_{xz}(x + \Delta h, z + \Delta h, t + \frac{\Delta t}{2}) \right).
 \end{array} \right. \tag{1.6}$$

Linear interpolation of medium parameters is used when necessary. Other kinds of interpolation are possible. To ensure accuracy and stability, arithmetic averaging of density and harmonic averaging of elastic moduli may be required in elastic modeling (Moczo et al., 2002 (127)).

Numerical dispersion and stability

Because of the discretization, the spatial and time steps have to satisfy stability criteria. These criteria are upper bounds for Δt and Δh , specific to the employed method, and are more restrictive than the Nyquist-Shannon sampling theorem (e.g., Whittaker, 1915 (229); Nyquist, 1928 (138); Shannon, 1949 (168); Kotelnikov, 1933 (102)). This theorem states that the minimum sampling frequency of a signal should be equal to twice the signal highest frequency component, to not distort the signal information.

In order to minimize the effect of the approximate spatial derivatives, spatial discretization has to satisfy a minimum number of points N per wavelength. If this criterion is not met, the affected wavefield appears as a ringing event called numerical (or grid) spatial dispersion. It is due to the phase speed becoming too dependent on the discretization step (see for instance Dablain, 1986 (48)). Generally, N decreases with the order of the used Taylor expansion. For example, Dablain (1986) (48) (resp. Levander, 1988 (110)) determined with numerical experiences that with a second order Taylor expansion, the body waves requires 8 (resp. 10) points per wavelength while only 4 (resp. 5) are

required for a 4th order Taylor expansion. This stability condition translates into:

$$\Delta h \leq \frac{\min(v)}{N \max(f)}, \quad (1.7)$$

$\max(f)$ being the maximum modeled frequency, and $\min(v)$ being the minimum non-null velocity of any sort. A higher frequency and a lower velocity lead to a lower limit for Δh and thus to a higher computing cost. A consequence is that when performing elastic instead of acoustic modeling, the upper bound for Δh decreases significantly, increasing drastically the computing costs. Note that accurate modeling of Scholte waves can require higher spatial discretization (van Vossen et al., 2002 (199)).

Time discretization also has to satisfy a minimum number of points per period. This condition is called the Courant-Friedrichs-Lewy condition, and aims at ensuring that the scheme can access the information required to form the solution. The theoretical analysis assuming homogeneous medium applied to our modeling algorithm gives (Courant et al., 1967 (45); Virieux, 1986 (215); Levander, 1988 (110); Saenger and Bohlen, 2004 (160); Martin et al., 2008 (116); He, 2017 (80)):

$$\Delta t \leq \frac{\sqrt{2\Delta h^2}}{2 \max(v) \sum_{k=0}^K |c_k|}, \quad (1.8)$$

$\max(v)$ being the maximum velocity of any sort, and c_k being the same FD coefficients as in equation 1.5. As a consequence, when the grid point spacing is reduced, the upper limit for the time step also decreases. This explains the particularly high computing cost of elastic or high frequency modeling.

Initial and boundary conditions

The chosen FD method computes the wavefield at a specific spatial point using the wavefield at neighboring points. Boundary conditions are consequently needed. Two kinds of boundaries are implemented: (1) the free surface boundary across the air-solid or air-water interface (the standard atmosphere pressure is of ~ 10 kPa, which is negligible in comparison with the solid Earth or the water); (2) the absorbing boundary to mimic an infinite space. The absorbing boundary cannot mimic waves that come back from outside the model. If their amplitudes are small enough, the errors should have a limited impact. Note that no explicit condition is applied for the fluid-solid boundary (i.e., the seafloor).

Various absorbing boundaries/layers implementation are possible, among them, the Perfectly Matched Layers (PML) and the Convolutional Perfectly Matched Layers (CPML). PML were originally formulated by Berenger (1994) (21), and CPML was initially formulated by Kuzuoglu and Mittra (1996) (104) and Roden and Gedney (2000) (159), both for use with Maxwell's equations. See Virieux et al. (2012) (213) for a review of these techniques. Our FWI algorithm uses the CPML technique. CPML needs the addition of a layer at the boundary of the model which will dampen the wavefields.

A free surface is defined by vanishing normal components of stress at the surface. For example, for our top horizontal free surface (located at $z = 0$):

$$\sigma_{zz}(x, z = 0, t) = \sigma_{xz}(x, z = 0, t) = \sigma_{zx}(x, z = 0, t) = 0. \quad (1.9)$$

The numerical implementation requires some care for our resolution algorithm since the vertical normal stress σ_{zz} is defined at the free surface, while the shear stress σ_{xz} is half-grid point below the free surface. For this configuration, the mirror technique can be used

(Robertsson, 1996 (158)). Virieux et al. (2012) (213) further suggested to extrapolate the velocity wavefields as even functions across the free surface, leading to:

$$\begin{cases} \sigma_{zz}(x, -n\Delta h, t) = -\sigma_{zz}(x, n\Delta h, t) = 0, \\ \sigma_{zz}(x, z = 0, t) = 0, \\ \sigma_{xz}(x, -m\frac{\Delta h}{2}, t) = -\sigma_{xz}(x, m\frac{\Delta h}{2}, t), \\ \sigma_{zx}(x, -m\frac{\Delta h}{2}, t) = -\sigma_{zx}(x, m\frac{\Delta h}{2}, t), \\ v_x(x, -n\Delta h, t) = v_x(x, n\Delta h, t), \\ v_z(x, -m\frac{\Delta h}{2}, t) = v_z(x, m\frac{\Delta h}{2}, t), \end{cases} \quad (1.10)$$

with $n \in [1, K + 1]$ and $m \in [0, K + 1]$. One can refer to Virieux et al. (2012) (213) for a review of free-surface implementations.

The wavefield at a specific time step is computed using the wavefield at an anterior time step. Consequently, initial conditions are needed. The algorithm assumes zero initial conditions for stress, velocities and their time derivatives, in agreement with the causality of seismic wave creation.

Sources and receivers

Our modeling algorithm allows the sources and the receivers to be located in between computation grid points. An interpolation algorithm is then required. Our modeling algorithm uses Kaiser's interpolation, as implemented by Hicks (2002) (83).

The chosen wave-equation (i.e., a stress-velocity formulation, in contrast to a displacement formulation) allows for an easy insertion of both body force and moment density sources.

Managing computing resources: shot parallelization, domain decomposition and basic algorithm optimization techniques

Solving the wave equation with numerical methods requires a considerable amount of computing resources (time and memory). Because FWI needs numerous simulations, this initially limited the application of FWI to very small and simplified cases. The increase of computing resources (together with the development of techniques to benefit from it) and the development of optimization techniques allowed FWI to be applied to more complex and larger cases, making it a major tool for the Earth exploration field. Two usual and vital techniques are available to benefit from the increased computing resources: the shot parallelization technique and the domain decomposition technique. Both techniques are implemented in our algorithm (using Message Passing Interface standard (MPI)), along with basic algorithm optimization rules. Domain decomposition was implemented as part of this Ph.D. thesis to allow for FWI applications on Alwyn real data. Because of the constant need for better choices of realism (such as the elastic one), higher frequency modeling and/or larger models and acquisitions, current FWI applications are still very time and memory consuming and limited by the available computing resources: efforts still need to be made.

As illustrated in figure 1.3, the shot parallelization technique allows to perform different simulations at the same time (i.e., in parallel). Each core performs one modeling. The user computing time (i.e., the time that the user has to wait) can then be divided by the number of available cores. The total computing time (i.e., the sum of the computing

times of each core) stays almost the same: the difference is due to the setup time of the parallel computations.

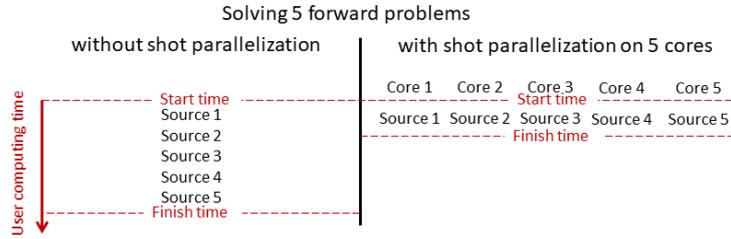


Figure 1.3: Illustration of the functioning and user computing time reduction of shot parallelization.

As illustrated in figure 1.4, the domain decomposition technique allows to spread a simulation over different cores. For each modeling, the spatial domain of computation is divided into subdomains and each core computes the wavefield in a different subdomain. To compute the wavefield at a specific point and at a time t , one needs the wavefield at a previous time in the neighboring points (potentially computed by another core). Consequently, each subdomain has to communicate with the neighboring subdomains. The information exchanged between subdomains is stored inside shadow subdomains. In FD methods, this is due to the Taylor expansions and the width of these shadow zones depends on the order of the Taylor expansion.

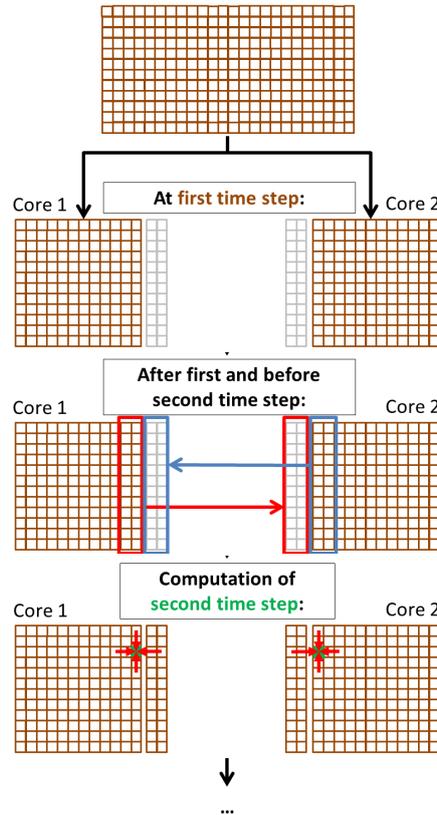


Figure 1.4: Example of domain decomposition for a spatial domain of computation divided into 2 horizontal subdomains and a Finite Difference (FD) method using 2 points on each side of the computed point (4^{th} order centered spatial Taylor expansion). For 8^{th} order centered spatial Taylor expansions, 4 points are needed.

Figure 1.5 provides an example of the honorable speed-up of our domain decomposition implementation. The speed-up is the ratio between the user computing time with and without domain decomposition, for different numbers of cores. Ideally, the speed-up should be a linear function. This ideal is impossible to reach because of the communication time needed between the cores (in addition to the MPI setup time): the more subdomains, the more communication. This explains the increasing (but small) differences between the ideal and observed speed-up curves. If the number of subdomains increases too much, domain decomposition can even be detrimental. Shot parallelization has much less communication. That is why, when facing only user computing time problems, if the number of available cores is lower than the number of simulations, shot parallelization should be preferred over domain decomposition. To illustrate the need for domain decomposition to allow for FWI applications on Alwyn real data, domain decomposition allowed us to use 16 cores per shot, reducing the user computing time from ~ 6 hours to ~ 25 minutes for a single forward modeling.

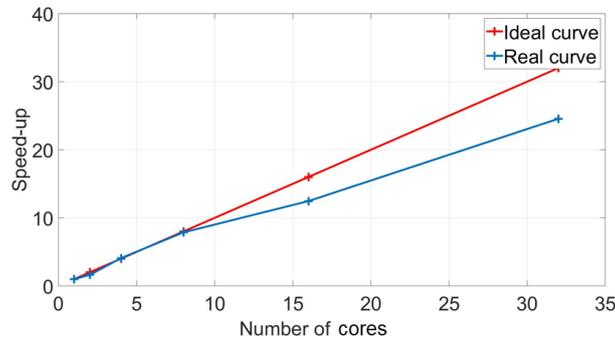


Figure 1.5: Example of speed-up of our domain decomposition implementation.

Domain decomposition can also be useful when dealing with memory problems. Indeed, as illustrated in figure 1.6, cores are organized by nodes and each core uses the shared memory of its node to perform the computations. With MPI, each core works with its own copy of the necessary variables (such as the models). For large simulations, the memory of the node can be insufficient to store the variables for all the cores. In such cases, only part of the cores of the node are used. The computing capacities of the other cores are then wasted. For even larger simulations, the memory of the node can be insufficient to store the necessary variables of a single core. In such cases the modeling is impossible. When using domain decomposition, only a fraction of the variables (such as the corresponding domain of the models and the shadow zone) is needed and stored by each core. The memory needs of each core is then reduced, preventing the waste of computing resources and allowing some very memory-demanding applications. To further illustrate the need for domain decomposition to allow for FWI applications on Alwyn real data, without domain decomposition, a core requires ~ 77 Gb of memory, preventing the application on the 64 Gb-nodes and leading to the waste of all the other cores (usually 15 or 23) on the 128 Gb-nodes, the two kinds of available nodes. Note that domain decomposition implemented with MPI increases the total memory need of memory. This effect could be partially mitigated by using OpenMP (alongside MPI), where the processes do not use their own copy of the variables, but shared variables on the node memory.

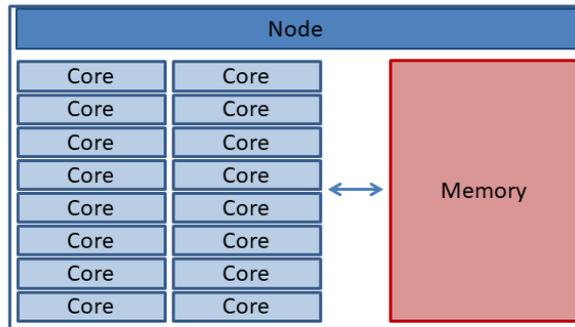


Figure 1.6: Illustration of the organization of a node and its memory. The memory is shared by all the cores of the node.

Basic algorithm optimization rules must be applied in order to reduce the number of operations and increase the speed of modeling. The acceleration due to these basic optimization rules is not negligible. The main rule concerns the organization of imbricated loops over arrays, which have to take advantage of the storage order of the array by using the proper loop nesting. This storage order depends on the programming language. In addition to this optimization technique (which was already implemented by Weiguang He in our FWI algorithm), additional algorithm optimizations were performed, as part of this Ph.D. thesis. Detection of bottlenecks with a profiler (here Intel[®] VTune[™]) led to the removal of unnecessary allocations and initialization, replacement of divisions by multiplications and simplification of imbricated loops and if conditions. All these additional algorithm optimizations led to a division by up to two of the computing time.

1.3 Inverse problem

As mentioned in the introduction, the inverse problem consists in the minimization over a model space of a misfit between the observed and modeled seismic data, the forward problem providing the modeled data. The misfit is quantified by an objective function. In most cases, the objective function is a cost function, and the minimization of the misfit is performed through the minimization of this cost function (or loss function). In a few cases, the minimization of the misfit is performed through the maximization of a reward function instead. In both cases, an inversion (or optimization) algorithm is needed. Various objective functions and inversion algorithms are possible. After providing the objective function available in our FWI algorithm (i.e., an ℓ_2 -norm cost function), I detail the implemented two inversion algorithms: a conjugate-gradient and a BFGS algorithm. As for previous section, in order to add more perspective, some other possible choices are mentioned. Throughout this section, the proofs are made in the non-discretized or discretized space, depending on which space brings the most simple demonstration.

1.3.1 A choice of objective function: an ℓ_2 -norm cost function

Depending on the objective, the data and the a priori knowledge of the sub-surface, more or less sophisticated (even under development) objective functions can be chosen. A single objective function is available in our FWI algorithm. It is the objective function that dominates the FWI world - maybe because of its simplicity and good performance

on a wide range of problems - i.e. the following ℓ_2 -norm (or least-square) cost function:

$$\begin{aligned} \mathcal{O}: \mathbb{R}^N &\rightarrow \mathbb{R} \\ \mathbf{m} &\mapsto \frac{1}{2} \sum_{\text{src}} \iint_S \int_{t_1}^{t_2} \sum_{\text{cmp}} \left[w_{\text{cmp,src}}(t, \mathbf{x}) \left(d_{\text{cmp,src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - d_{\text{cmp,src}}^{\text{obs}}(t, \mathbf{x}) \right)^2 \right] dt dS, \end{aligned} \quad (1.11)$$

where:

$$\begin{aligned} \sum_{\text{src}} &= \text{sum over the sources src,} \\ \iint_S &= \text{space integral (replaced by a triple integral in 3D),} \\ \int_{t_1}^{t_2} &= \text{time integral between the start } t_1 \text{ and stop } t_2 \text{ recording times,} \\ \sum_{\text{cmp}} &= \text{sum over the components cmp: the pressure } P \text{ (i.e., mean of stresses) and the velocity components } v_x \text{ and } v_z, \\ d_{\text{cmp,src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) &= \text{modeled wavefield at time } t \text{ and location } \mathbf{x} \text{ for the component cmp, the source src, and the model } \mathbf{m}, \\ d_{\text{cmp,src}}^{\text{obs}}(t, \mathbf{x}) &= \text{observed wavefield at time } t \text{ and location } \mathbf{x} \text{ for the component cmp, the source src, and the model } \mathbf{m}, \\ w_{\text{cmp,src}}(t, \mathbf{x}) &= \text{weight introduced to:} \\ &\quad - \text{account for the space and time discrete aspect of seismic data,} \\ &\quad - \text{represent the (tapered) time/receiver selection,} \\ &\quad - \text{select the component(s). Here, } P, v_x, v_z \text{ or } (v_x, v_z). \end{aligned}$$

In practice, in our FWI algorithm, the time integration of the computation of the cost function is performed using the rectangle rule. $w_{\text{src,cmp}}$ could also be used to balance the different data, such as the data coming from different components.

Some more recently introduced objective functions bring hope in the fight against some specific limitations of this ℓ_2 -norm cost function, among them:

- ℓ_1 -norm or mixed ℓ_1 -, ℓ_2 -norm cost functions. Because of the (partial) suppression of the square exponent, these cost functions tend to put less weight on extreme data values and tend to perform better than ℓ_2 -norm cost functions with data contaminated with high amplitude noises (e.g., Crase et al., 1990 (47)).
- Envelope cost functions. Because of the replacement of the data by their envelope, these cost functions tend to allow for poorer initial models, but tend to provide less detailed final models (e.g., Wu et al., 2014 (232)).
- Trace-normalized cost functions. Because of the use of normalized traces (i.e., the data coming from a single source-receiver pair and component are normalized separately), these cost functions put more weight on the arrival times, and less weight on the amplitude, notably the amplitude variation with offset. These cost functions can show advantages when the observed data amplitudes cannot be reproduced (e.g., Louboutin et al., 2017 (113)).
- Optimum transport cost functions, which also tend to allow for poorer initial models (e.g., Métivier et al., 2016 (123)). In practice, Métivier et al. (2016) minimized their cost function by maximizing an associated reward function.
- Cost functions involving constraints on the model parameters. Among the different model regularizations, one can cite total variation regularization (e.g., Aghamiry et al., 2019 (1)) and Tikhonov regularizations of different order (e.g., Asnaashari et al., 2013 (12)). Model regularizations can force the model to be close to a priori model, to well data and/or to have limited spatial variations for example.

- Travel-time based cost functions. These cost functions also tend to allow for poorer initial models but tend to provide less detailed final models (see Wang et al., 2015 (218) for instance, who used cross-correlation in their wave-equation tomography).

The value of allowing for poorer initial models or of applying model constraints is emphasized later in this chapter.

1.3.2 A choice of inversion algorithm: a conjugate-gradient and a BFGS algorithm

From now on, because the chosen objective function is a cost function, I focus on minimization problems. A similar reasoning can be done for maximization problems.

The two most usual inversion algorithms in FWI have been implemented in our FWI algorithm: Nocedal’s Limited-memory Broyden-Fletcher-Goldfarb-Shanno Bounded L-BFGS-B algorithm (Morales and Nocedal, 2011 (131)) and the Polak-Ribière conjugate-gradient algorithm (Polak and Ribière, 1969 (145)). Both are local inversion algorithms of the line search type.

The need for local inversion algorithms

Because of the high number of model parameters and of the high computing cost of the forward problem, looking for global minima in the whole model space is too costly. Instead, a local solution in the neighborhood of an initial model \mathbf{m}_0 is sought, using line search algorithms. As a consequence of the use of a local algorithm, FWI can lead to a minimum different from the desired one (another global minimum or a local minimum).

With notably the increase of computing resources and the growth of machine learning, more and more other inversion methods are tested to overcome the limitations of local inversion methods. In particular, one can cite (partial) global inversion techniques such as Monte-Carlo-based techniques (e.g., Zhao and Sen, 2019 (237); Visser et al., 2019 (216)), or genetic algorithms (e.g., Mazzotti et al., 2016 (121); Sajeve et al., 2017 (161); Guo et al., 2020 (72)). Global inversion techniques explore the whole model space and try to retrieve the global minima.

Line search algorithms

Line search algorithms propose to iteratively find a local optimum of an objective function in the neighborhood of \mathbf{m}_0 with:

$$\begin{aligned}\mathbf{m}_1 &= \mathbf{m}_0 + \beta_0 \mathbf{p}_0, \\ \mathbf{m}_2 &= \mathbf{m}_1 + \beta_1 \mathbf{p}_1, \\ &\dots\end{aligned}\tag{1.12}$$

Or more concisely written:

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \beta_n \mathbf{p}_n, n \in \mathbb{N}^+,$$

with \mathbf{p}_n an update direction (hopefully a descent direction), and β_n a steplength set to have the largest decrease along the given direction. Several methods co-exist to set \mathbf{p}_n , among them: BFGS methods and conjugate gradient methods.

BFGS methods

Broyden-Fletcher-Goldfarb-Shanno **BFGS** methods are quasi-Newton's methods (initially proposed in 1970 by Broyden (34), Fletcher (59), Goldfarb (67), and Shanno (169)). Let us recall that Newton's method proposes to iteratively find a stationary point (figure 1.7 visually describes the different kinds of stationary points) of an objective function \mathcal{O} with:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mathbf{H}_n^{-1} \nabla \mathcal{O}_n, n \in \mathbb{N}^+. \quad (1.13)$$

$\nabla \mathcal{O}_n$, \mathbf{H}_n and \mathcal{O}_n are respectively the gradient of \mathcal{O} , the Hessian (i.e., the second derivatives) of \mathcal{O} and the objective function \mathcal{O} evaluated at \mathbf{m}_n , with:

$$\nabla \mathcal{O}_n = \begin{pmatrix} \frac{\partial \mathcal{O}}{\partial m_{n1}} \\ \vdots \\ \frac{\partial \mathcal{O}}{\partial m_{nN}} \end{pmatrix} \text{ and } \mathbf{H} = \begin{pmatrix} \frac{\partial^2 \mathcal{O}}{\partial m_{n1}^2} & \cdots & \frac{\partial^2 \mathcal{O}}{\partial m_{n1} \partial m_{nN}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{O}}{\partial m_{nN} \partial m_{n1}} & \cdots & \frac{\partial^2 \mathcal{O}}{\partial m_{nN}^2} \end{pmatrix},$$

where \mathbf{m}_n is a column vector containing the N model parameters m_{ni} , with $i \in \llbracket 1, N \rrbracket$. For example, for a rectangular model containing Nx points along the horizontal direction, Nz points along the vertical direction and 5 parameter types (such as in 2D VTI models), $N = Nx * Nz * 5$. Appendix A provides the background and limitations of Newton's method.

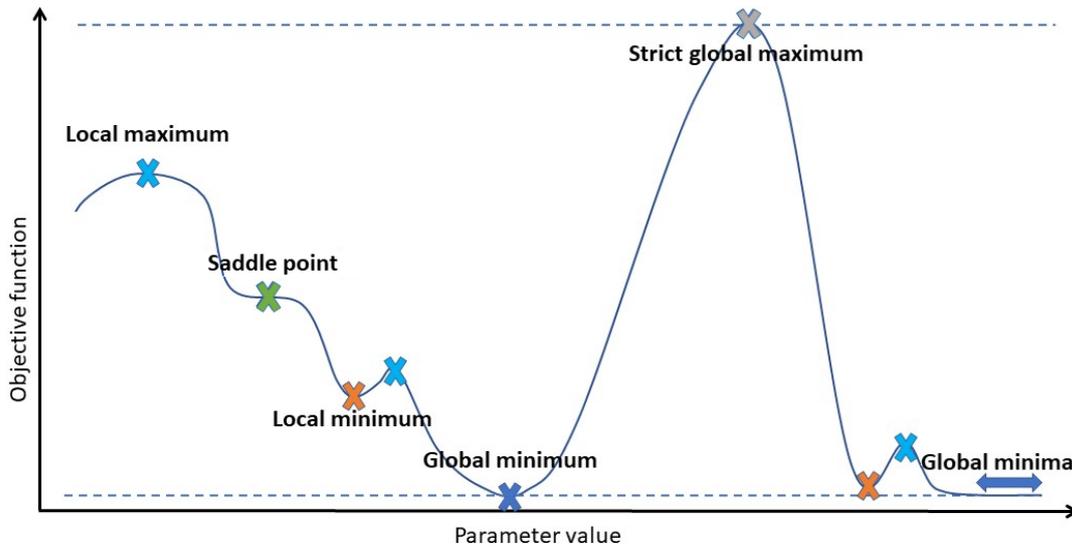


Figure 1.7: Visual description of the different kinds of stationary points on a virtual objective function which depends on a single parameter.

If one wants to compute the Hessian with simulations, it would require $N \times N$ simulations. In usual **FWI** contexts, I cannot afford that many simulations. Instead, the Hessian is approximated. The steepest descent-direction can be seen as the harshest approximation of the Newton's method, where the Hessian is replaced by the identity matrix. The model minimizing the cost function \mathcal{O} is then obtained by iterative update of an initial model \mathbf{m}_0 in the descent direction (i.e., the opposite of the gradient direction). The distance traveled along the descent direction is specified by the steplength $\beta_n \in \mathbb{R}^{+*}$. The steepest-descent then has the form:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \beta_n \nabla \mathcal{O}_n. \quad (1.14)$$

The steepest-descent method does not take into account the curvature of the objective function. Consequently, they might require a larger number of iterations compared with Newton’s method. That is why methods that better approximate the Hessian (or directly the product $\mathbf{H}_n^{-1}\nabla\mathcal{O}_n$) are preferred over the steepest-descent method. Among them, the implemented **L-BFGS-B** method is the most common quasi-Newton’s method in **FWI**. **L-BFGS-B** is an extended version of Limited memory Broyden-Fletcher-Goldfarb-Shanno **L-BFGS** methods where the updated model can be bounded (i.e., extreme values can be imposed). **L-BFGS** is a limited memory version of the **BFGS** algorithm. See Nocedal and Wright (2006) (136) and Métivier et al. (2014) (122) for background and details (such as advantages and disadvantages) about this method and other possible Hessian approximations. Details on the computation of \mathbf{p}_n for **L-BFGS** method are provided in **Appendix B**. **L-BFGS-B** algorithm was made available in our **FWI** algorithm by Weiguang He, by introducing version 3.0 of an open-source **L-BFGS-B** algorithm: Morales and Nocedal (2011) (131).

Conjugate-gradient methods

Conjugate-gradients methods are another type of line search methods. They were originally proposed by Hestenes and Stiefel (1952) (81) to find the global optimum of quadratic functions in at most n iterations, n being the numbers of variables. They were extended by Fletcher and Reeves (1964) (58) for general non-quadratic functions. Among conjugate-gradient methods, the implemented Polak-Ribière variant (Polak and Ribière, 1969 (145)) (**PR-CG**) is the most common conjugate-gradient method in **FWI**. See Nocedal and Wright (2006) (136) for background and details (such as advantages and disadvantages) about this method and other possible conjugate-gradient methods. Details on the computation of \mathbf{p}_n for **PR-CG** method are provided in **Appendix B**. This inversion algorithm has been added to our **FWI** algorithm as part of this Ph.D. thesis in order to offer an easy-to-master alternative to the Morales and Nocedal’s **L-BFGS-B** implementation. Some of their principal advantages are their simplicity and small memory requirements compared with (L-)BFGS methods.

Comments on both methods

Besides the intrinsic limitation of local optimization methods, which are providing only a local minimum, **L-BFGS-B** and **PR-CG** have their own shortcomings, for instance:

- Unlike the steepest descent method, **L-BFGS-B** and **PR-CG** do not ensure a descent direction, although an ascent direction rarely happens in practice. Other variants of the Newton’s method and conjugate-gradient methods that ensure a descent direction exist (Nocedal and Wright, 2006 (136));
- Since all three methods are gradient-based, they can provide a saddle point instead of a minimum.

Comparing **BFGS** and conjugate gradient methods, **BFGS** algorithms tend to show a higher rate of convergence (need for less iterations), less simulations and better results than conjugate-gradient (Métivier et al., 2014 (122)). However, they have higher memory requirements than conjugate gradient methods, and even more so than the steepest descent. In this thesis, for simplicity reasons, the conjugate gradient method is selected.

Steplength computation

The choice of the steplength β_n is critical, for instance: small steplengths increase the number of iterations and big steplengths can miss the optimum we are aiming at, by moving off this optima or even moving in the neighborhood of another optimum, as illustrated in figure 1.8.

In our FWI algorithm, L-BFGS-B and conjugate-gradient inversion methods have separate steplength computation implementations. I refer to Morales and Nocedal (2011) (131) for details about the steplength computation for the L-BFGS-B algorithm. Our conjugate-gradient can be performed with two usual steplength computation methods: the steplength computation as described by Pica et al. (1990) (143) and the subspace method as described by Kennett et al. (1988) (94).

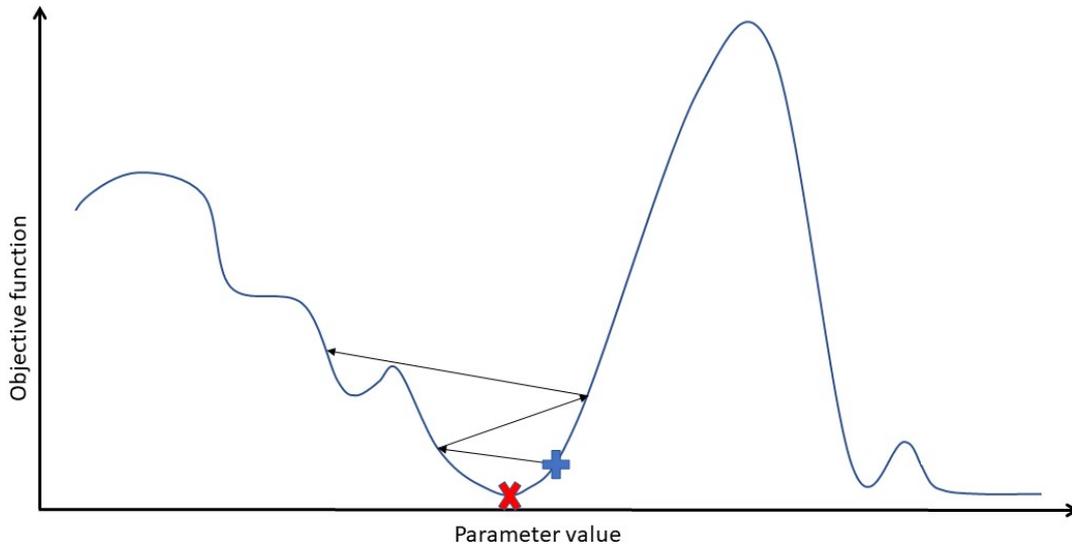


Figure 1.8: Visual description of the effect of a too big steplength on a virtual objective function which depends on a single parameter: the direction of update is the descent direction but the steplength is too big leading to an increase of the cost function. The blue cross indicates the initial model and the red cross indicates the desired minimum. The arrows represent the successive updates.

Details of these methods applied to our cost function are provided in Appendix C. Both methods are equivalent when inverting for a single parameter type (e.g., P-wave vertical velocity) and differ for multi-parameter inversions: while Pica et al.'s method provides a single steplength, subspace method allows for a different steplength per parameter kind. For our cost function, at each iteration and for each source, to compute the steplength(s), Pica et al.'s method requires a single additional modeling, while the subspace method requires as many additional simulations as inverted parameter kinds. Subspace method is therefore more costly in multiparameter inversions, but was shown to better take into account the involvement of different parameter classes (Kennett et al., 1988 (94)). This last method was consequently selected in this thesis applications. They both assume that the objective function is locally quadratic in the direction(s) of update.

Choice of initial model and cost function

Let us focus on a cost function whose minimum of interest is in a strictly convex neighborhood. Assuming that \mathbf{m}_0 is inside this neighborhood, the cost function should decrease at each iteration then stabilize (“converge”) at the desired minimum. The inversion is then successful. The way up to the result highly depends on the objective function and on the choice of the initial model \mathbf{m}_0 . In practice, one has a poor idea of the behavior of the objective function but knows it has multiple stationary points. This explains the need for imposing model constraints, to guide the inversion algorithm towards the desired minimum. Another consequence is the need for a \mathbf{m}_0 as close as possible to the (unknown) desired minimum, in order to increase the probability of success (figure 1.9). This explains the utility of cost functions allowing for poorer initial models. These cost functions are built to exhibit a wider strictly convex neighborhood around the aimed minima and/or less disturbing stationary points, increasing the probability of success (figure 1.9).

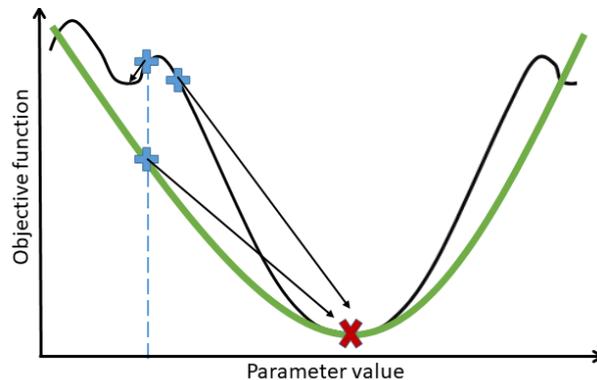


Figure 1.9: Visual description of the effect of an initial model closer to the aimed minimum, as well as the effect wider strictly convex neighborhood around the aimed minimum, with two virtual cost functions which depend on a single parameter. The blue cross indicates the initial model and the red cross indicates the desired minimum. The arrows show the expected obtained model for each cost function and initial model.

1.3.3 The gradient computation

The gradient computation with the adjoint-state method

Whatever line search method, the computation of the gradient (and of the steplength) is a central part of the inverse problem and is explained below.

If one wants to estimate the gradient with simulations, it would require N additional simulations per source. In usual FWI contexts, we cannot afford that many simulations. Instead, FWI algorithms (including our algorithm) typically compute the gradient with the adjoint-state method, introduced in FWI by Tarantola (1984) (187).

Appendix D provides details about the application of the adjoint-state method for the cost function and wave-equation used in our FWI algorithm. It provides the gradient in terms of the elastic moduli and the density $(C_{11}, C_{13}, C_{33}, C_{44}, \rho)$. Similar applications can be done for other objective functions and wave-equations. See Plessix (2006) (144) and Fichtner (2011) (57) for more details about the adjoint-state method in FWI. Details of the implementation can also be found in Weiguang He’s Ph.D. thesis (2017) (80). In opposition with Appendix D which uses the same framework as Fichtner, He used the Lagrangian framework to explain the adjoint-state method. Figure 1.10 describes

the different steps of the gradient computation in our FWI algorithm. With the adjoint state method, a single additional modeling per source is needed (for the adjoint wavefield computation).

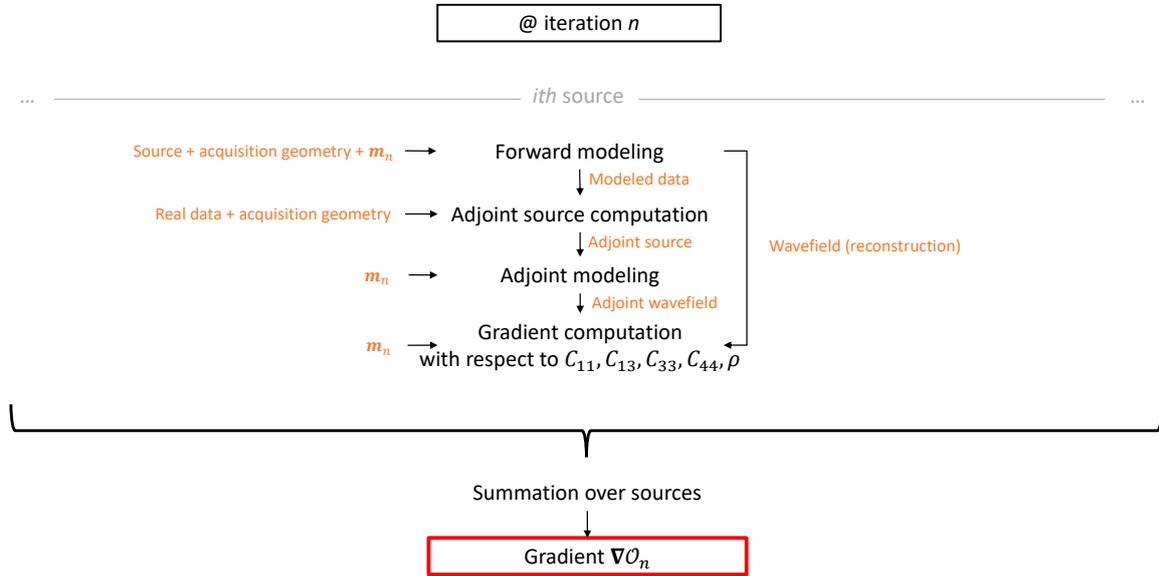


Figure 1.10: Steps of the gradient computation in our Full Waveform Inversion FWI algorithm. The inputs of each step are indicated in orange.

The gradient can be computed in terms of any other parameterization (b_1, b_2, b_3, b_4, b_5) with the chain rule:

$$\frac{\partial \mathcal{O}}{\partial b_i} = \frac{\partial \mathcal{O}}{\partial C_{11}} \frac{\partial C_{11}}{\partial b_i} + \frac{\partial \mathcal{O}}{\partial C_{13}} \frac{\partial C_{13}}{\partial b_i} + \frac{\partial \mathcal{O}}{\partial C_{33}} \frac{\partial C_{33}}{\partial b_i} + \frac{\partial \mathcal{O}}{\partial C_{44}} \frac{\partial C_{44}}{\partial b_i} + \frac{\partial \mathcal{O}}{\partial \rho} \frac{\partial \rho}{\partial b_i}. \quad (1.15)$$

Indeed, whatever the choice of realism, an infinity of parameterizations (i.e., the parameter types on which the gradients is computed) can describe the same medium. For example $C_{11}, C_{13}, C_{33}, C_{44}, \rho$ and $Vpv, Vph, Vsv, \delta, \rho$ are two parameterizations that can be used to describe an elastic VTI 2D medium, Vpv being the P-wave vertical velocity, Vph being the P-wave horizontal velocity, Vsv being the SV-wave vertical velocity, and δ being a parameter allowing the full description of the variation of velocities with the direction, given Vpv, Vph and Vsv (Thomsen, 1986 (192)).

Gradient pre-conditioning

Pre-conditioning can be applied to the gradient, and aims at constraining the inversion through a modification of the computed gradient. For example, our FWI algorithm allows notably for: smoothing (useful to remove the imprint of a sparse acquisition for instance), masking (usually applied around sources and receivers, or to select a (subset of) updated parameter type(s)), updated model clipping (which clips values higher or lower than defined bounds).

Aside from masking (if constant over iterations and if no taper is applied between the masked and the unmasked area), the other pre-conditionings are not mathematically justifiable and can perturb the inversion algorithm. In practice, being supported by a priori knowledge, these pre-conditionings can lead the inversion algorithm towards a

better results. Clipping is implemented for the conjugate-gradient only, which does not have a built-in feature to bound the model parameters values (like L-BFGS-B has).

1.4 Inversion strategy

As explained in introduction and as shown in the two previous sections, different flavors of FWI exist. In addition, the inversion process usually involves several calls to these different flavors. The choice of the flavors and the choice of the order in which they are called is called the inversion strategy. This inversion strategy highly depends on the objectives, the geological environment (including a priori knowledge) and on the available FWI inputs (model, data, source function and acquisition geometry). The number of inversion strategies is infinite. In this section, I aim at providing additional crucial (but non-exhaustive) knowledge for inversion strategy design.

One of the main objective of the inversion strategy is to deal with one of the limitations of the used line search methods: the risk to obtain an optimum different from the desired one. After reminding a useful fact (when designing an inversion strategy) about the resolution of FWI, I describe three usual strategies to lessen this problem. I finally focus on the related multiparameter inversion subject. In both above cases, the inversion strategy is in fact designed to direct FWI towards an updated model that the geoscientist deems better than other possible updated models.

1.4.1 Resolution aspects

As reminded by Virieux and Operto (2009) (212), the resolution of the updates brought to the initial model depends on the angle between the forward waves and adjoint waves, as well as the frequency of these waves and the velocity at which they travel. For waves traveling with frequency f at a velocity of c_0 and with the angle θ , the maximum resolution of the update is in the angle bisector direction, along which the thinnest update has a width w such that:

$$w = \frac{c_0}{2f \cos(\frac{\theta}{2})}. \quad (1.16)$$

Therefore, a more detailed update is provided by the use of higher frequencies, smaller apertures and/or lower velocities. As a consequence, for the same frequency and angle, due to their lower velocities, S-waves provide more detailed S-wave velocity model updates than the P-wave velocity model updates obtained with P-waves. As another consequence, the choice of the maximum inverted frequency is a balance between the needed resolution and the available computing resources (equation 1.7). This resolution property also highlights the impact of the acquisition geometry which, given a sub-surface, controls the direction of the bisectors and consequently the direction in which the update can be resolved.

1.4.2 Driving full waveform inversion towards the intended optimum

Beydoun and Tarantola (1988) (23) notably explained that, in the context of the first Born approximation, the initial model should allow the fitting of observed traveltimes within an error of half the period. Otherwise, the algorithm could match the data by fitting

the wrong oscillations together, leading to a wrong model (i.e., wrong minimum). This problem is called the “cycle-skipping problem” and is due to the oscillatory nature of the data. Figure 7 of Virieux and Operto (2009) (212) (provided in figure 1.11) illustrates the cycle-skipping problem. The same kind of reasoning can be performed on wavefield fitting and wavelength (i.e., the starting model should allow the fitting of observed wavefield within an error of half the wavelength).

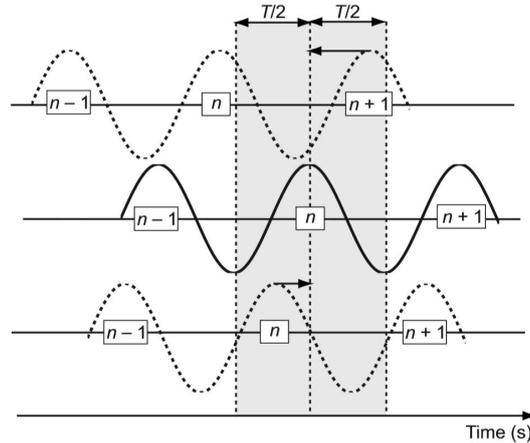


Figure 1.11: “Schematic of cycle-skipping artifacts in FWI. The solid black line represents a monochromatic seismicogram of period T as a function of time. The upper dashed line represents the modeled monochromatic seismicograms with a time delay greater than $T/2$. In this case, FWI will update the model such that the $n+1^{th}$ cycle of the modeled seismicograms will match the n^{th} cycle of the observed seismicogram, leading to an erroneous model. In the bottom example, FWI will update the model such that the modeled and recorded n^{th} cycle are in-phase because the time delay is less than $T/2$.” Figure and title from Virieux and Operto (2009) (212).

In order to obtain a good initial model, more robust model building methods can be used, such as: reflection and/or refraction travel-time tomography (e.g., in Arnulf et al., 2011 and 2012 (9) (10)), multiphysics (e.g., gravimetry in Sirtori et al., 2019 (176) or well data in Reilly, 1984 (154)) and regional/global knowledge incorporation. These methods usually lack of detailed updates, provide too local information or partial information (e.g., density information only), which is a reason to employ FWI.

Once the initial model is built, to avoid the local minimum problem, FWI can be driven by model regularization, or by using as a first step, more robust objective functions such as the above introduced envelope, optimum transport or travel-time based objective functions.

To lessen even more the cycle-skipping problem, Bunks et al. (1995) (35) proposed a multiscale strategy where progressively higher frequencies are inverted (hereafter called “increasing frequency multiscale strategy”). The period of low frequencies being higher, the chances for the initial model to allow the fitting of observed traveltimes within an error of half the period are better. In our FWI algorithm, the multiscale strategy is applied by filtering the observed and modeled data (i.e., by modifying the cost function). Various frequency selection strategies are possible. Sirgue and Pratt (2004) (175), based on wave propagation in homogeneous models, proposed to select frequencies by making use of the redundancy between frequency and aperture for a given the resolution (equation 1.16). Cho et al. (2008) (43) and Kim et al. (2011) (96) compared frequency selection strategies

in the context of frequency-domain FWI. This multiscale strategy relies on the low frequency content of the data and explains the effort made to provide lower and better low frequencies (e.g., through low frequency sources as the sources proposed by Landrø and Amundsen, 2014 (107) or Chelminsky et al., 2019 (41); through denoising techniques or low-frequency recording systems). In the absence of low frequencies, strategies to infer them are under investigation (e.g., Li and Demanet, 2016 (112)).

Another multiscale strategy was proposed by Shipp and Singh (2002) (172) (hereafter called “decreasing aperture multiscale strategy”). They inverted progressively narrower aperture data (using time and offset selection). This strategy follows the logic of equation 1.16: the wider the aperture, the larger the wavelength appears, which lowers the chances of cycle-skipping. This strategy relies on wide-aperture data, explaining the development of its acquisition (such as ocean bottom seismometer/node).

The frequency multiscale strategy is usually nested within the aperture multiscale strategy (e.g., Sears et al., 2010 (165)). These strategies emphasize the need for frequencies and apertures providing a continuity in the wavelength and period space, and a continuity between the lowest resolution achievable with FWI and the resolution of the initial model.

The last inversion strategy I would like to mention is the layer stripping strategy where the shallowest parts of the model are updated first. The deeper the wave travels, the more cycle-skipped it can be, due to accumulation of model errors. By fitting first the shallowest data (usually not cycle skipped), the misfit between deep modeled and observed data is reduced. For instance, Masoni et al. (2016) (119) applied this strategy by offset and frequency selection, for surface wave inversion. For other kinds of wave, layer stripping seems to be naturally done: ℓ_2 -norm FWI seems to focus first on the stronger events, usually waves with a shorter path (i.e., shallower) (Wang et al., 2020 (221)).

1.4.3 About multiparameter inversion

Even if FWI can theoretically provide all the medium parameters involved in the forward problem, multiparameter inversion (i.e., inversion of more than one parameter type) is still a challenge that needs to be solved. As a consequence, only a subset of the parameter types is usually updated during inversion. The choice of the updated parameters and their order and combination of update turn out to be critical.

This challenge partly comes from crosstalks (or trade-offs) and data insensitivity problems. Parameters are crosstalked when a variation of one can be compensated by a variation of the others. For example, let us consider a horizontal wave traveling in a VTI medium, where horizontal velocities are higher than vertical velocities. The propagation of the wave can also be explained by an isotropic medium with higher P-wave velocities.

Data insensitivity problems can occur when the used wave types are not sensitive (enough) to the desired model parameters. For instance, post-critical P-P reflections and density (see an illustration of the insensitivity in Sears Ph.D. thesis, 2007 (166) or in Chapter 3).

In both cases, a multitude of models equally explains the observed data and there is no strict global minimum. Therefore, it might be worth keeping some parameters constant over iterations and allow for FWI updates of the other parameter(s) only, provided the parameterization allows it.

For each parameter type of each parameterization, FWI has a different wave propagation direction sensitivity, and the choice of the parameterization is critical for a successful

FWI (see for instance He et al., 2018 (79), who show inversion results obtained with different parameterizations). The parameterization should be chosen such that FWI shows a sensitivity for the desired parameter type and for the direction of waves used in the objective function. A strategy to avoid crosstalk-related problems is to use a parameterization in which the desired parameters have a non overlapping range of propagation direction sensitivity. This strategy relies on the presence of a variety of wave directions and raises the question of how much this strategy relies on a low frequency content (because the multiscaling strategy from wide-aperture to narrow-aperture is not possible anymore).

The importance of the update order is notably illustrated by Sears et al. (2008) (164). They showed that - in the case of elastic FWI in a marine airgun acquisition in a soft seabed environment - inverting first for P-wave and then for P- and S-wave velocities leads to better results than inverting for all velocities at once. This strategy will be explained in more detail in Chapter 3.

In the absence of a good multiparameter inversion strategy, one can use passive updates of some parameters (usually density and S-wave velocity). For example, P-wave velocity is updated using a line search algorithm, and density is updated at each iteration based on the new P-wave velocity model using an empirical relationship (for example, from Gardner et al., 1974 (65) and Hamilton, 1978 (75) or Castagna et al., 1985 (39) for S-wave velocity update). Despite not being mathematically justifiable, it has a physical justification and can provide better results. Additionally, multiparameter inversion offers the possibility to use one of the parameter type as a garbage collector: it will be wrongly updated, allowing for a better update of other parameter types, by including effects (such as noise or too simple choice of realism) that cannot be handled by the desired parameter types (see for instance Marjanović et al., 2019 (118)). However, because of crosstalks, one might take the risk of putting in this garbage collector effects that are due to the desired parameter types.

1.5 Discussion and conclusions

In this chapter, the high computational needs of the forward problem has been emphasized, especially for elastic modeling. Examples of time and memory needs were provided from the application on Alwyn real data. Moreover, I showed that the inverse problem also requires solutions of the forward problem. For example, the use of conjugate-gradient with the adjoint-state method and Pica et al.'s steplength brings to three the number of simulations per source and per iteration. Combined with the high number of iterations, FWI is therefore onerous in terms of computing resources and limited by this aspect.

Without going into the details of the kind of inversion this thesis is focusing at, I explained the main reasons why multi-parameter inversions are a challenging task and why specific inversion strategies need to be designed.

Throughout this chapter, I insisted on the limitation of the local inversion methods used in FWI. These limitations, especially the possibility to find a local minimum, call for the need for quality control tools for FWI outputs. The crucial role of the acquisition (as the seismic data provider) and of a good starting model has also been highlighted: the seismic data and input model quality largely determine the choice of the inversion strategy and the achievable results. This explains why the application on real data (Chapter 5) will be preceded by an analysis of the FWI inputs that will focus on these aspects (Chapter 4): inputs preparation and improvement, assessment of the achievable results, choice of the inversion strategy, and FWI outputs quality control tools identification.

Chapter 2

Reciprocity for fast full waveform inversion on airgun-ocean bottom node or seismometer data in elastic vertical transverse isotropic media

Chapter's objectives and highlights

- Explains the need for reciprocity in the studied full waveform inversion context, with notably a realistic example of computing time reduction.
- Propose an extensive guide of the use of reciprocity in the studied full waveform inversion context, from theory to implementation, and from data modeling to model update computation, including synthetic tests. In particular, I provide:
 - the thought process of using reciprocity in full waveform inversion model update computation, as well as its causes and consequences on said update.
 - the extension of already known data modeling reciprocity relations to more general environments. The extension and the proof of the modeling reciprocity relations are used to highlight the limitations of the pre-existing relations.
- Emphasize the usefulness of reciprocity as a debugging/testing technique for modeling and full waveform inversion algorithms. Indeed, it allows to obtain the same result in two different ways but with the same algorithm, offering a unique comparison method for non-simple applications.

Contents

2.1	Introduction	34
2.2	A few necessary ingredients...	37
2.2.1	The used time-domain, elastic vertical transverse isotropic wave equation	37
2.2.2	From convolutional type reciprocity relation to Green's functions and Green's functions reciprocity	37
2.2.3	Displacement and pressure expressions for an airgun source using Green's functions	38
2.3	Reciprocity for pressure and velocity data modeling	39

2.3.1	Reciprocity relation for pressure data	40
2.3.2	Reciprocity relation for velocity data	42
2.3.3	Synthetic data applications	44
2.4	Reciprocity for full waveform inversion	49
2.4.1	A change of point of view	49
2.4.2	Synthetic applications	50
2.5	Discussion and conclusions	52

2.1 Introduction

As explained in [Chapter 1](#), Full Waveform Inversion ([FWI](#)) is a powerful method for obtaining models of mechanical parameters from seismic data. It iteratively updates an initial model by minimizing the misfit between observed and modeled seismic data. Consequently, [FWI](#) requires multiple data and wavefield modeling, and the more simulations, the higher the computing cost. The number of modeling is usually proportional to the number of sources, and can usually be considered as independent from the number of receivers. I focus here on Earth local exploration, as explained in the introductory chapter. In such a context, in traditional acquisitions such as streamer or Ocean Bottom Cable ([OBC](#)) acquisitions, the number of sources is reasonable and the dense coverage is provided by a high number of receivers, allowing the use of [FWI](#). Conversely, ocean bottom recordings such as Ocean Bottom Node ([OBN](#)) or Ocean Bottom Seismometer ([OBS](#)) acquisitions can contain much more sources, and much more sources than receivers, dissuading the use of [FWI](#). [OBN/OBS](#) acquisitions consist of receivers laying on the seabed, with hydrophone recording the pressure field just above the seabed, and geophones recording the three components of the velocity fields at the seabed. It has various advantages for seismic imaging, such as allowing for wide-azimuth/offset/aperture data, deployment near production platforms or inversion of S-wave velocity models (Sears et al., 2008 ([164](#))). In fact, [OBN/OBS](#) acquisitions are increasingly used. The summary of the 2000 SEG/EAGE Summer Research Workshop on multi-component technology (Gaiser et al., 2001 ([62](#))) represents well the use and usefulness of [OBN/OBS](#).

Seismic reciprocity relations link seismic quantities with interchanged source and receiver locations, as illustrated in [2.1](#). The term “quantity” is used in this chapter as a generic term to call the wavefield quantities, e.g.: pressure, displacements, dilatation or velocities. In opposition with reciprocity of travel-times, reciprocity of the whole waveform of these wavefield quantities is not direct, hence the existence of the seismic reciprocity field, especially in domains such as [FWI](#).

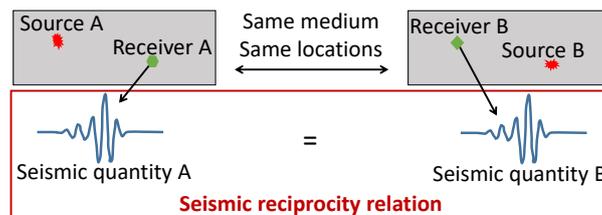


Figure 2.1: Schematic description of reciprocity relations, i.e.: relations linking two seismic quantities (e.g., pressure, displacement) from two source-receiver pairs with interchanged locations. As represented, both sources, receivers and quantities can be of different types.

The broad bibliography transcribed hereafter indicates that seismic reciprocity relations have been extensively used for several decades. The first seismic reciprocity relations are attributed to Betti for elastostatics (1872) (22) and to Rayleigh for elastodynamics (1873, 1878) (151) (152). Depending on the medium, its boundary conditions, the sources, the recorded quantities or the domain of analysis, etc., the reciprocity can be expressed differently. For instance, Fokkema and van den Berg (1993) (61) have provided reciprocity relations using convolution and cross-correlation (and the equivalent local and Laplace-domain relations) from acoustic wave equation in heterogeneous media. De Hoop (1966) (49) and Gangi (2000) (63) derived similar reciprocity relations of convolution type (and the equivalent local and Laplace-domain relations) for visco-elastic, anisotropic and heterogeneous media, while Wapenaar et al. (2004) (224) targeted media with imperfect interfaces for acoustic waves in fluids, elastodynamic waves in solids, and poroelastic waves in porous solids. From these general reciprocity relations, more specific relations can be derived. For instance, Knopoff and Gangi (1959) (99), Nowack and Chen (1999) (137), Arntsen and Carcione (2000) (8), Carcione (2014) (38) and Wang et al. (2015) (219) provided very specific reciprocity relations for various types of point sources (forces, monopoles, dipoles and quadrupoles), and various recorded quantities in inhomogeneous, (visco-)elastic, anisotropic media. Wapenaar and Grimbergen (1996) (69) and Wapenaar (1998) (227) even derived reciprocity relations for one-way wavefields and propagators. These reciprocity approaches could be verified using synthetic data computations or real data: Fenati and Rocca (1984) (55) and Katou et al. (2017) (90) illustrated reciprocity relations using real data and showed that even if the assumptions behind a reciprocity relation are not respected, reasonable reciprocity may still be observed. Mittet and Hokstad (1995) (125), Moghaddam et al. (2012) (128), and Ung et al. (2015) (198) verified reciprocity relations with finite difference modeling results. Ung et al. (2015) (198) provided an example where reciprocity highlighted an incorrect boundary implementation. Razavy and Leonach (1986) (153) pointed out that reciprocity does not necessarily hold if asymptotic ray theory, fourth (or higher) order differencing methods, or analytical approximations are used. Seismic reciprocity relations have been used for various purposes including radiation patterns computation (White, 1960 (230)), representation theorem derivation -which can be considered as a reciprocity relation- (Gangi, 1970 (64)), mitigating geophone-ground coupling problems (Tan, 1987 (186)), source and/or receiver amplitude equalization (Kmenbach, 1994 (98), and van Vossen et al., 2005 (200), 2006a (203), 2006b (202)), symmetric sampling (Vermeer, 1991 and 1998 (206) (207)), double plane-wave migration (Seifoullaev et al., 2005 (167), and Zhao et al., 2017 (235)), redatuming (Berryhill, 1984 (20), and Mulder, 2005 (133)), interferometry (van Manen et al., 2006 (201), and Wapenaar et al., 2010 (225)), wavefield decomposition (van Borselen et al., 2013 (204)), or passive monitoring (Vidal et al., 2019 (208)).

In particular, reciprocity relations have been used to decrease the computing cost of **FWI** for **OBN/OBS** acquisitions where the same receiver geometry is used for all sources (namely a stationary receiver geometry). To be more precise, the role of sources and receivers are interchanged in the **FWI** modeling, to make use of the lower number of receivers. Several authors used this strategy in **FWI** (e.g., Operto et al., 2006 (141); Operto et al., 2015 (142); Amestoy et al., 2015 (6); Alves, 2017 (5) used reciprocity for velocity components in an acoustic and elastic isotropic setting respectively). Sheen et al. (2004 and 2006) (170) (171) provided details for the use of reciprocity in the inversion scheme of **FWI**, but the scenario is different and is not specific to **OBN/OBS** acquisitions: they used reciprocity to make the expensive computation of partial derivative wavefields

of Gauss-Newton method more affordable.

Despite this extensive use of reciprocity, the need for the research work described in this chapter was dictated by the scattered, non-unified and limited aspect of the current available documentation. For instance, this documentation poorly emphasizes the reciprocity limitations, and includes no phrased explanation regarding the application of reciprocity in [FWI](#) for the model update computation. Indeed, not only data needs to be modeled, but also wavefields, for which the usual data reciprocity relations are not sufficient, as highlighted later in this chapter. This chapter aims at explaining the thought process consisting in using reciprocity in [FWI](#) model update computations, as well as its causes and consequences on said update. This explanation is built on a detailed derivation of the reciprocity relations used for data modeling in [FWI](#). This derivation also allows to highlight the limitations of the pre-existing relations.

In this chapter, I describe the application of reciprocity in the specific case of [FWI](#) applied on data coming from a stationary [OBN/OBS](#) acquisition with airgun (or airgun array) sources. Following the choice of realism explained in the introductory chapter, I limit the study to the elastic Vertical Transverse Isotropic [VTI](#) context, without intrinsic attenuation. When modeling the source, I assume that the airgun (or airgun array) sources are point moment density sources with no directivity (see for instance Arntsen et al., 2000 ([8](#))). A point source being a source located in a single space point, it is described through its location and its time dependence: the source time function. In the rest of the document, the word “airgun” will indifferently refer to an airgun or an airgun array. Following assumptions are also made when building the model used to propagate the waves:

- The water layer (or in fact the layer in which the sources and the hydrophones are, which in practice is a water layer) is a homogeneous acoustic isotropic medium. The acoustic and isotropic assumptions are supported by the fact that water is intrinsically an isotropic liquid. The homogeneous assumption is reasonable at the first order, especially when modeling shallow water environments such as our real data case Alwyn, where little variation of water properties with depth is expected.
- The source time function is the same for all sources, which is reasonable at the first order, given the fact that the same type of equipment is used for all sources, and with similar local surroundings.

I first specify a few necessary ingredients: the wave equation used in this document, the corresponding convolutional type reciprocity relation, Green’s functions and subsequent basic Green’s functions reciprocity relationship, as well as the used analytical description of an airgun source and the resulting expressions for the displacement and the pressure. I then provide a time-domain derivation of the usual reciprocity relations for pressure and velocity data modeling, and compare a set of them with finite difference modeling results. I finally focus on reciprocity in [FWI](#), notably through synthetic applications. The mathematical derivations are made in 3D while the applications are in 2D. The necessary adaptations for 2D reciprocity are provided. The synthetic applications used a modified Marmousi2 model (Martin et al., 2006 ([115](#))).

2.2 A few necessary ingredients...

2.2.1 The used time-domain, elastic vertical transverse isotropic wave equation

The propagation of seismic waves in an elastic **VTI** medium can be modeled using the following wave equation, with appropriate initial and boundary conditions:

$$\begin{cases}
 \rho(\mathbf{x})\ddot{u}_x(\mathbf{x}, t) & = \sigma_{xx,x}(\mathbf{x}, t) + \sigma_{xy,y}(\mathbf{x}, t) + \sigma_{xz,z}(\mathbf{x}, t) + f_x(\mathbf{x}, t), \\
 \rho(\mathbf{x})\ddot{u}_y(\mathbf{x}, t) & = \sigma_{yx,x}(\mathbf{x}, t) + \sigma_{yy,y}(\mathbf{x}, t) + \sigma_{yz,z}(\mathbf{x}, t) + f_y(\mathbf{x}, t), \\
 \rho(\mathbf{x})\ddot{u}_z(\mathbf{x}, t) & = \sigma_{zx,x}(\mathbf{x}, t) + \sigma_{zy,y}(\mathbf{x}, t) + \sigma_{zz,z}(\mathbf{x}, t) + f_z(\mathbf{x}, t), \\
 \sigma_{xx}(\mathbf{x}, t) & = C_{11}(\mathbf{x})u_{x,x}(\mathbf{x}, t) + [C_{11}(\mathbf{x}) - 2C_{66}(\mathbf{x})]u_{y,y}(\mathbf{x}, t) + C_{13}(\mathbf{x})u_{z,z}(\mathbf{x}, t), \\
 \sigma_{yy}(\mathbf{x}, t) & = [C_{11}(\mathbf{x}) - 2C_{66}(\mathbf{x})]u_{x,x}(\mathbf{x}, t) + C_{11}(\mathbf{x})u_{y,y}(\mathbf{x}, t) + C_{13}(\mathbf{x})u_{z,z}(\mathbf{x}, t), \\
 \sigma_{zz}(\mathbf{x}, t) & = C_{13}(\mathbf{x})[u_{x,x}(\mathbf{x}, t) + u_{y,y}(\mathbf{x}, t)] + C_{33}(\mathbf{x})u_{z,z}(\mathbf{x}, t), \\
 \sigma_{xy}(\mathbf{x}, t) = \sigma_{yx}(\mathbf{x}, t) & = C_{66}(\mathbf{x})[u_{x,y}(\mathbf{x}, t) + u_{y,x}(\mathbf{x}, t)], \\
 \sigma_{xz}(\mathbf{x}, t) = \sigma_{zx}(\mathbf{x}, t) & = C_{44}(\mathbf{x})[u_{x,z}(\mathbf{x}, t) + u_{z,x}(\mathbf{x}, t)], \\
 \sigma_{yz}(\mathbf{x}, t) = \sigma_{zy}(\mathbf{x}, t) & = C_{44}(\mathbf{x})[u_{y,z}(\mathbf{x}, t) + u_{z,y}(\mathbf{x}, t)],
 \end{cases} \tag{2.1}$$

with x and y representing the horizontal axes and z representing the vertical axis. t is the time and \mathbf{x} is a position vector made up of three components. ∂_j indicates a partial spatial differentiation along j and $\ddot{\cdot}$ indicates a second order partial time derivative. u_i is the i^{th} component of the displacement field (unit of length). σ_{xx}, σ_{yy} and σ_{zz} (resp. $\sigma_{xy}, \sigma_{yx}, \sigma_{xz}, \sigma_{zx}, \sigma_{yz}$ and σ_{zy}) are the normal (resp. shear) components of stress (force per unit of surface). ρ is the density (mass per unit of volume) and C_{ij} are the components of the 4th order elasticity tensor -or stiffness tensor- introduced in the generalized Hooke's law (mass per unit of length and per squared unit of time). Together with density, the elastic moduli describe the mechanical properties of the sub-surface. f_i is the i^{th} component of an external body force source (force per unit of volume). The velocity of waves in all directions can be computed from the density and these elastic moduli, especially the P-wave vertical velocity V_{pv} and the SV-wave vertical velocity V_{sv} (Thomsen, 1986 (192)). Note that this wave equation does not model attenuation. Also note that the use of Hooke's law implies linear elasticity, experimentally valid for small deformations. Equation 2.1 corresponds to the 3D equivalent of equation 1.3, without the moment density sources. One can then find more details about this equation in Chapter 1.

2.2.2 From convolutional type reciprocity relation to Green's functions and Green's functions reciprocity

I recall here the main steps to obtain the needed Green's function reciprocity relations from the convolutional type reciprocity relation.

Let us consider two body force sources \mathbf{f}^A and \mathbf{f}^B leading respectively to the seismic quantities $\{\mathbf{u}^A, \boldsymbol{\sigma}^A\}$ and $\{\mathbf{u}^B, \boldsymbol{\sigma}^B\}$. Let us consider a volume V bounded by a surface S . The convolutional type reciprocity relation (also known as Betti-Rayleigh relation) for such elastic, anisotropic equation for heterogeneous media (see Knopoff and Gangi, 1959

(99) or Aki and Richards, 2002 (3)) is:

$$\iiint_V \left(f_i^B(\mathbf{x}, t) *_t u_i^A(\mathbf{x}, t) - f_i^A(\mathbf{x}, t) *_t u_i^B(\mathbf{x}, t) \right) dV = \iint_S \left(\sigma_{ij}^A(\mathbf{x}, t) *_t u_i^B(\mathbf{x}, t) n_j(\mathbf{x}) - \sigma_{ij}^B(\mathbf{x}, t) *_t u_i^A(\mathbf{x}, t) n_j(\mathbf{x}) \right) dS, \quad (2.2)$$

where the Einstein convention of summing on repeated indices is used. This convention is used hereafter, except on the explicit indices x , y and z . \mathbf{n} is the normal to the surface, pointing outwards. $*_t$ represents the time convolution. This equation assumes zero initial and/or final conditions for the displacement and its time derivative, and integrability of the different quantities. The derivation of this equation is also available in appendix E. By assuming homogeneous boundary conditions (valid for a horizontal free-surface or absorbing boundaries for instance), the surface integral becomes null and one can obtain:

$$\iiint_V \left(f_i^B(\mathbf{x}, t) *_t u_i^A(\mathbf{x}, t) - f_i^A(\mathbf{x}, t) *_t u_i^B(\mathbf{x}, t) \right) dV = 0. \quad (2.3)$$

Let us focus on a receiver located at \mathbf{x}_{rcv} . We assume that the Green's function $G_{iq}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}, 0)$ exists. $G_{iq}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}, 0)$ is the value at \mathbf{x}_{rcv} and at time t of the i^{th} Green's function of wave equation 2.1 for a unit body force source oriented along the q^{th} axis, shot at time 0 and located at \mathbf{x} . It can be seen as the i^{th} component of the displacement at \mathbf{x}_{rcv} and at time t due to a unit body force source oriented along the q^{th} axis, shot at time 0 and located at \mathbf{x} . Equation 2.3 can be used to derive the very well known Green's functions reciprocity relation (Aki and Richards, 2002 (3)):

$$G_{iq}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}, 0) = G_{qi}(\mathbf{x}, t; \mathbf{x}_{\text{rcv}}, 0). \quad (2.4)$$

2.2.3 Displacement and pressure expressions for an airgun source using Green's functions

Equation 2.3 can also be used to express the displacement due to any kind of source (Aki and Richards, 2002 (3)):

$$u_i^{\text{force}}(\mathbf{x}_{\text{rcv}}, t) = \iiint_V f_q(\mathbf{x}, t) *_t G_{iq}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}, 0) dV. \quad (2.5)$$

We use the superscript force to specify that the kind of source that generated the seismic quantity is a body force source: for example u_x^{force} refers to a horizontal displacement due to a body force source. A recorded data can be seen as the combination of the source, the receiver and the propagation effects between the source and the receiver. From equation 2.5, one can understand that these Green's functions are containers of the propagation effects.

An airgun source at \mathbf{x}_{src} is modeled by a point moment density source with no directivity, which means in terms of force source (see for instance Arntsen et al., 2000 (8)):

$$f_q(\mathbf{x}, t) = \left(-p(t)\delta(\mathbf{x} - \mathbf{x}_{\text{src}}) \right)_{,q} \text{ for } q = x, y, z, \quad (2.6)$$

where δ is the Dirac delta function, and $p(t)$ is the source time function (unit of stress, i.e. force per unit of surface).

Inserting 2.6 into 2.5 and using a property of Dirac delta function, the i^{th} component of displacement for an airgun source is then:

$$u_i^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) = -p(t) * G_{iq,q}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}_{\text{src}}, 0), \quad (2.7)$$

where the superscript $\text{airgun}(\mathbf{x}_{\text{src}}, p)$ specifies that the quantity is due to an airgun source located at \mathbf{x}_{src} with source time function p .

The pressure $P(\mathbf{x}_{\text{rcv}}, t)$ at \mathbf{x}_{rcv} and at time t is defined by:

$$P(\mathbf{x}_{\text{rcv}}, t) = \frac{\sigma_{xx}(\mathbf{x}_{\text{rcv}}, t) + \sigma_{yy}(\mathbf{x}_{\text{rcv}}, t) + \sigma_{zz}(\mathbf{x}_{\text{rcv}}, t)}{3}, \quad (2.8)$$

which gives in our elastic **VTI** case (i.e., using equation 2.1):

$$\begin{aligned} P(\mathbf{x}_{\text{rcv}}, t) &= \frac{2C_{11}(\mathbf{x}_{\text{rcv}}) + C_{13}(\mathbf{x}_{\text{rcv}}) - 2C_{66}(\mathbf{x}_{\text{rcv}})}{3} u_{x,x}(\mathbf{x}_{\text{rcv}}, t) \\ &+ \frac{2C_{11}(\mathbf{x}_{\text{rcv}}) + C_{13}(\mathbf{x}_{\text{rcv}}) - 2C_{66}(\mathbf{x}_{\text{rcv}})}{3} u_{y,y}(\mathbf{x}_{\text{rcv}}, t) \\ &+ \frac{2C_{13}(\mathbf{x}_{\text{rcv}}) + C_{33}(\mathbf{x}_{\text{rcv}})}{3} u_{z,z}(\mathbf{x}_{\text{rcv}}, t). \end{aligned} \quad (2.9)$$

By inserting 2.7 into 2.9, we obtain the pressure due an airgun source:

$$\begin{aligned} P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) &= - \frac{2C_{11}(\mathbf{x}_{\text{rcv}}) + C_{13}(\mathbf{x}_{\text{rcv}}) - 2C_{66}(\mathbf{x}_{\text{rcv}})}{3} \left(p(t) * G_{xq,qx}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}_{\text{src}}, 0) \right) \\ &- \frac{2C_{11}(\mathbf{x}_{\text{rcv}}) + C_{13}(\mathbf{x}_{\text{rcv}}) - 2C_{66}(\mathbf{x}_{\text{rcv}})}{3} \left(p(t) * G_{yq,qy}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}_{\text{src}}, 0) \right) \\ &- \frac{2C_{13}(\mathbf{x}_{\text{rcv}}) + C_{33}(\mathbf{x}_{\text{rcv}})}{3} \left(p(t) * G_{zq,qz}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}_{\text{src}}, 0) \right). \end{aligned} \quad (2.10)$$

2.3 Reciprocity for pressure and velocity data modeling

As mentioned in the introduction of this chapter, **FWI** requires seismic data modeling (pressure or velocity data). In order to reduce the computing cost of modeled data simulation, it is preferable to model by receiver gather instead of source gather, due to the much smaller number of receivers to be considered (figure 2.2). We are therefore interested in reciprocity relations which link the pressure or velocity fields recorded at the receiver with another seismic quantity generated by a source at the receiver location. We suppose that all the airguns have the same source time function. Let us call this quantity the “reciprocal seismic quantity” and the corresponding source and receiver the “virtual source” and the “virtual receiver”.

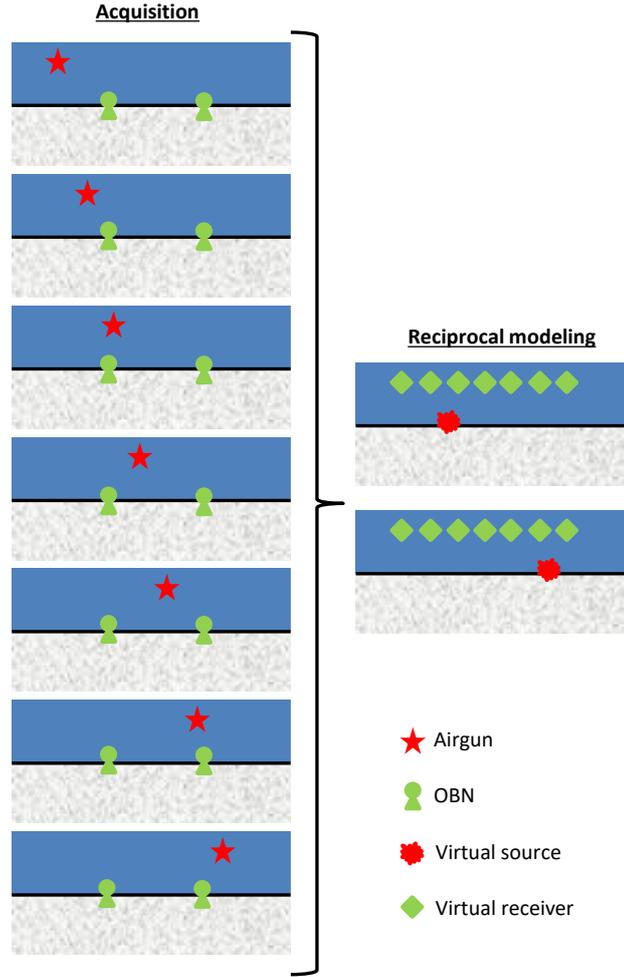


Figure 2.2: Illustration of reciprocity for a dataset containing seven airguns and two Ocean Bottom Nodes (OBNs). When the acquisition contains seven sources, the reciprocal simulations would ideally contain only two virtual sources instead of seven simulations, reducing the computing cost.

2.3.1 Reciprocity relation for pressure data

From reciprocity relation 2.4, the pressure $P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t)$ of equation 2.10 can be written:

$$\begin{aligned}
 P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) = & - \frac{2C_{11}(\mathbf{x}_{\text{rcv}}) + C_{13}(\mathbf{x}_{\text{rcv}}) - 2C_{66}(\mathbf{x}_{\text{rcv}})}{3} \left[p(t) \ast_t G_{qx, xq}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0) \right] \\
 & - \frac{2C_{11}(\mathbf{x}_{\text{rcv}}) + C_{13}(\mathbf{x}_{\text{rcv}}) - 2C_{66}(\mathbf{x}_{\text{rcv}})}{3} \left[p(t) \ast_t G_{qy, yq}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0) \right] \\
 & - \frac{2C_{13}(\mathbf{x}_{\text{rcv}}) + C_{33}(\mathbf{x}_{\text{rcv}})}{3} \left[p(t) \ast_t G_{qz, zq}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0) \right].
 \end{aligned} \tag{2.11}$$

This is a first reciprocity relation for pressure data, where the reciprocal seismic quantity involves the Green's functions. If this reciprocity relation was used, reciprocity would be applied by computing the necessary spatially differentiated Green's functions, and then applying equation 2.11. This equation essentially requires Green's function computation

for three single force directions. Consequently, reciprocity would become advantageous when the number of sources would be more than three times the number of receivers.

In the vast majority of cases, the source and the receiver are in the the water layer, considered as a homogeneous isotropic acoustic medium, leading to:

$$\begin{aligned} C_{11} - 2C_{66} &= C_{13}, \\ C_{11} &= C_{33}, \\ C_{13}(\mathbf{x}_{\text{rcv}}) &= C_{13}(\mathbf{x}_{\text{src}}), \\ C_{11}(\mathbf{x}_{\text{rcv}}) &= C_{11}(\mathbf{x}_{\text{src}}). \end{aligned} \tag{2.12}$$

From reciprocity relation 2.4, equation 2.10 and equations 2.12, we obtain:

$$\boxed{P^{\text{airgun}}(\mathbf{x}_{\text{src}}, p)(\mathbf{x}_{\text{rcv}}, t) = P^{\text{airgun}}(\mathbf{x}_{\text{rcv}}, p)(\mathbf{x}_{\text{src}}, t)}. \tag{2.13}$$

This is another reciprocity relation for pressure data, where the reciprocal seismic quantity is a pressure. As shown in figure 2.3, with this reciprocity relation, reciprocity is applied by computing the pressure at the source location due to an airgun shot at the receiver location with source time function p . In the studied case, reciprocity for pressure component is then direct, and both phase and amplitude are conserved when interchanging the source and the receiver locations. This second reciprocity relation (equation 2.13) has a more restricted range of application but is preferred over relation 2.11. Indeed, this new relation uses only one simulation per receiver location. Consequently, reciprocity becomes advantageous when the number of sources is higher than the number of receivers. In addition to be less costly than relation 2.11, this relation avoids the computation of the spatially differentiated Green's functions, and requires minimal modifications of the FWI code.

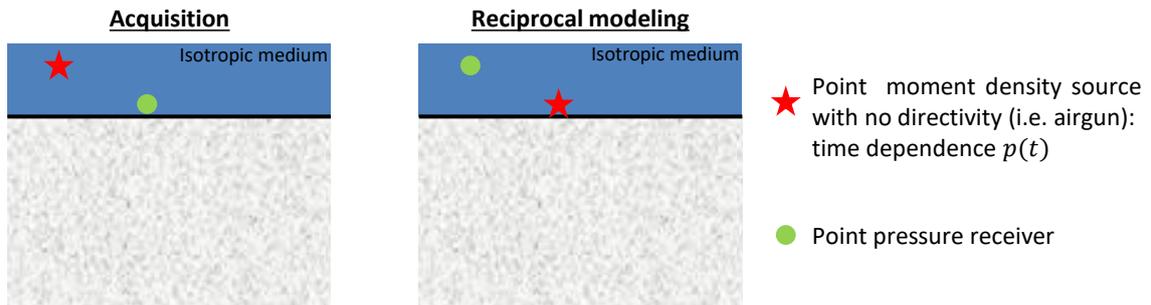


Figure 2.3: Reciprocal experiment for pressure component.

Equations 2.9 and 2.11 seem overly complicated given the fact that the water layer is considered to be a homogeneous isotropic acoustic medium. However, these equations notably allow to show that reciprocity for pressure component is not straightforward if the medium at the receiver locations does not exhibit the same isotropic properties as the one at the source locations. Moreover, those equations allow the extension to another useful case. Indeed, for an heterogeneous isotropic acoustic water layer (useful in deep water environments for instance, when medium parameters can differ between the source and the receiver locations, (Han et al., 2012 (76))), a simple reciprocity relation can be

obtained (from relation 2.4, equation 2.10 and applying $C_{11} - 2C_{66} = C_{13}$, and $C_{11} = C_{33}$):

$$P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) = \frac{1}{C_{11}(\mathbf{x}_{\text{src}})} P^{\text{airgun}(\mathbf{x}_{\text{rcv}}, C_{11}(\mathbf{x}_{\text{rcv}})p)}(\mathbf{x}_{\text{src}}, t). \quad (2.14)$$

For such a case, the reciprocity use is not as straightforward as one could think (e.g., Mittet and Hokstad, 1985 (125); Alves, 2007 (5)). Indeed, if interchanging the source and receiver locations has no phase effect, it still has a scaling effect, the scaling effect being the consequence of local source and receiver properties. This point highlights the importance of carefully deriving these reciprocity relations, as done in this chapter.

Equation 2.14 was also derived from a dilatation reciprocity relationship by Wang et al. (2015) (219). Arntsen and Carcione (2000) (8) and Carcione (2014) (38) derived an equation equivalent to the dilatation reciprocity relationship of Wang et al.. Note that equation 2.14 corresponds to the acoustic isotropic case and should thus be obtainable from the isotropic acoustic scalar wave equation.

2.3.2 Reciprocity relation for velocity data

Let us apply reciprocity relation 2.4 to equation 2.7:

$$u_i^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) = -p(t) * \dot{G}_{qi, q}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0). \quad (2.15)$$

By differentiating with respect to time t :

$$\dot{u}_i^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) = \dot{u}_i^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) = -p(t) * \dot{G}_{qi, q}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0), \quad (2.16)$$

where $(\dot{})$ indicates the partial time derivative. This is a first reciprocity relation for velocity data, where the reciprocal seismic quantity involves Green's functions. If this reciprocity relation was used, reciprocity would be applied by computing the necessary time and spatially differentiated Green's functions, and then applying equation 2.16. One simulation per receiver location and component would be necessary (to compute the time and spatially differentiated Green's functions). Consequently, reciprocity would become advantageous when the number of sources N_{src} is larger than the number of receivers N_{rcv} multiplied by the number of desired velocity components N_{comp} : $N_{\text{src}} > N_{\text{rcv}} * N_{\text{comp}}$. Note that the Green's function involved in this relation may not be the same as the one involved in equation 2.11, since the geophone may be located at the seabed while the corresponding hydrophone may be located right above the seabed. Derivations of equation 2.15 are also available in Nowack and Chen (1999) (137), and Eisner and Clayton (2001) (52). Moghaddam et al. (2012) (128) and Alves (2017) (5) derived equation 2.16 after finite difference discretizations and approximations application. Alves (2017) kept the Green's function in its final derived velocity reciprocity relations.

Using equations 2.9 and 2.5, the pressure $P^{\text{force}_i(\mathbf{x}_{\text{rcv}}, p)}(\mathbf{x}_{\text{src}}, t)$ recorded at the source location and due to a point body force source with source time function $p(t)$ shot at the receiver location on the i^{th} component is:

$$\begin{aligned} P^{\text{force}_i(\mathbf{x}_{\text{rcv}}, p)}(\mathbf{x}_{\text{src}}, t) &= \frac{2C_{11}(\mathbf{x}_{\text{src}}) + C_{13}(\mathbf{x}_{\text{src}}) - 2C_{66}(\mathbf{x}_{\text{src}})}{3} \left[p(t) * \dot{G}_{xi, x}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0) \right] \\ &+ \frac{2C_{11}(\mathbf{x}_{\text{src}}) + C_{13}(\mathbf{x}_{\text{src}}) - 2C_{66}(\mathbf{x}_{\text{src}})}{3} \left[p(t) * \dot{G}_{yi, y}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0) \right] \\ &+ \frac{2C_{13}(\mathbf{x}_{\text{src}}) + C_{33}(\mathbf{x}_{\text{src}})}{3} \left[p(t) * \dot{G}_{zi, z}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0) \right]. \end{aligned} \quad (2.17)$$

Since the source is in a homogeneous isotropic acoustic water layer:

$$P^{\text{force}_i(\mathbf{x}_{\text{rcv}}, p)}(\mathbf{x}_{\text{src}}, t) = C_{11}^{\text{water}} \left[p(t) *_t G_{qi, q}(\mathbf{x}_{\text{src}}, t; \mathbf{x}_{\text{rcv}}, 0) \right], \quad (2.18)$$

the superscript ^{water} referring to the properties of the water layer. Comparing equations 2.18 and 2.15 and differentiating with respect to time leads to:

$$\begin{aligned} v_i^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) &= -\frac{1}{C_{11}^{\text{water}}} P^{\text{force}_i(\mathbf{x}_{\text{rcv}}, \dot{p})}(\mathbf{x}_{\text{src}}, t) \\ &= P^{\text{force}_i(\mathbf{x}_{\text{rcv}}, -\frac{1}{C_{11}^{\text{water}}} \dot{p})}(\mathbf{x}_{\text{src}}, t). \end{aligned} \quad (2.19)$$

This is another reciprocity relation for velocity data, where the reciprocal seismic quantity is a pressure. As shown in figure 2.4, when this reciprocity relation is used, reciprocity for i^{th} component is applied by recording the pressure at the source location due to a point body force source along i^{th} axis with source time function $-\frac{1}{C_{11}^{\text{water}}} \dot{p}$ applied at the receiver location. Note that interchanging the source and receiver locations has only a scaling effect, the phase being conserved. The scaling effect is the consequence of local source effects. This second reciprocity relation (equation 2.19) is preferred over relation 2.16. Indeed, the computing cost of this reciprocity relation is the same as the previous one, but no time and spatially differentiated Green's functions are needed. Consequently, this reciprocity relation requires minimal modifications of the FWI code.

As for equation 2.14, we can extend equation 2.19 to a heterogeneous isotropic acoustic water layer:

$$\begin{aligned} v_i^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) &= -\frac{1}{C_{11}(\mathbf{x}_{\text{src}})} P^{\text{force}_i(\mathbf{x}_{\text{rcv}}, \dot{p})}(\mathbf{x}_{\text{src}}, t) \\ &= P^{\text{force}_i(\mathbf{x}_{\text{rcv}}, -\frac{1}{C_{11}(\mathbf{x}_{\text{src}})} \dot{p})}(\mathbf{x}_{\text{src}}, t), \end{aligned} \quad (2.20)$$

where interchanging the source and receiver locations has also only a scaling effect, the phase being conserved, and the scaling effect being the consequence of local source effects.

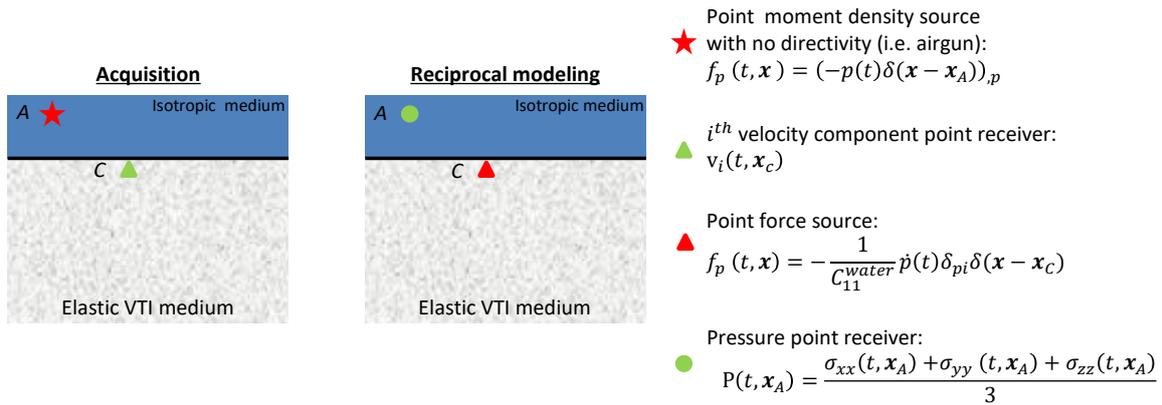


Figure 2.4: Reciprocal experiment for velocity components. VTI = Vertical Transverse Isotropic. f_p , with $p = x, y, z$, are the components of the body force source of equation 2.1

Equations 2.19 was derived by Operto et al. (2006) (141) and can also be derived from the dilatation-displacement reciprocity relationship of Wang et al. (2015) (219), which is

the equivalent equation but expressed in terms of displacement, and without assuming the source time function of force source and airgun is the same. Arntsen and Carcione (2000) (8), Carcione (2014) (38) and Nowack and Chen (1999) (137) also derived an equation equivalent to the dilatation-displacement reciprocity relationship of Wang et al..

2.3.3 Synthetic data applications

The selected reciprocity relations (framed equations: equation 2.13 for pressure component and equation 2.19 for velocity components) are tested for 2D synthetic models. These equations have been derived for a 3D model but the restriction to a 2D model is straightforward and the 2D equations are equal to the 3D ones. Data with and without reciprocity are computed for a modified Marmousi2 model: the P-wave vertical velocity Vpv model corresponds to the central part of the P-wave velocity Marmousi2 model (Martin et al., 2006 (115)) without the horizontal layer just below the seabed (in order to complicate the seabed transition) and with 126 m of isotropic homogeneous water layer on the top. SV-wave vertical velocity model Vsv and density models are generated by applying an experimental relationship on Vpv model. The missing SV-wave vertical velocity model Vsv is computed from the initial P-wave vertical velocity model Vpv , with (Castagna et al., 1985 (39); Flidner and White, 2001 (60); Shipp and Singh, 2002 (172) and Sears et al., 2010 (165)):

$$\begin{cases} Vsv = 0 \text{ m/s, in water,} \\ Vsv = \frac{Vpv-1360}{1.16} \text{ m/s, if } Vpv < 3500 \text{ m/s,} \\ Vsv = 0.53Vpv \text{ m/s, if } Vpv \geq 3500 \text{ m/s,} \end{cases} \quad (2.21)$$

where Vpv is in m/s. The missing density model ρ is computed from the initial P-wave vertical velocity model Vpv , with:

$$\begin{cases} \rho = 1030 \text{ kg/m}^3 \text{ in water,} \\ \rho = 2351 - 7497 * (Vpv * 10^{-3})^{-4.656} \text{ kg/m}^3 \text{ if } Vpv < 2000 \text{ m/s,} \\ \rho = x_a + \frac{Vpv-2000}{150} * (x_c - x_a) \text{ kg/m}^3 \text{ if } Vpv \geq 2000 \text{ m/s and } Vpv \leq 2150 \text{ m/s,} \\ \quad \text{with } x_a = 2351 - 7497 * (2000 * 10^{-3})^{-4.656} \text{ and } x_c = 1741 * (2150 * 10^{-3})^{0.25}, \\ \rho = 1741 * (Vpv * 10^{-3})^{0.25} \text{ kg/m}^3 \text{ if } Vpv > 2150 \text{ m/s,} \end{cases} \quad (2.22)$$

where Vpv is in m/s. The empirical expression of Hamilton (1978) (75) is used for the shallow seabed poorly consolidated sediments ($Vpv < 2000$ m/s), and the empirical expression of Gardner et al. (1974) (65) is used for the deeper consolidated sediments ($Vpv > 2150$ m/s). A linear transition between the two empirical relations is applied. Note that these two empirical relationships were derived assuming isotropic media and are here applied in anisotropic media by replacing the P-wave velocity of the initial values by the P-wave vertical velocities Vpv . Vpv model is shown on figure 2.5. Inspired by Alwyn's initial model, anisotropy is added by setting Thomsen's anisotropy parameters δ and ϵ to 0.105 and 0.232 respectively (Thomsen, 1986 (192)). A free surface is used on the top of the model. The source time function $\dot{p}(t)$ is a 7.5 Hz central frequency first derivative of a Gaussian. Source is 12 m deep and 180 m away from the left boundary of the model, geophones are on the first grid point below the seabed and hydrophones are on the first gridpoint above the seabed. The modeling code is a 8th order in space and 2nd order in time finite difference code using a staggered grid. The diagonal components of stress and the elastic parameters are on grid points, the other stress and velocity components are

in between gridpoints. 8 s of data are generated with a time step of 0.9 ms. The spatial step is of 12 m.

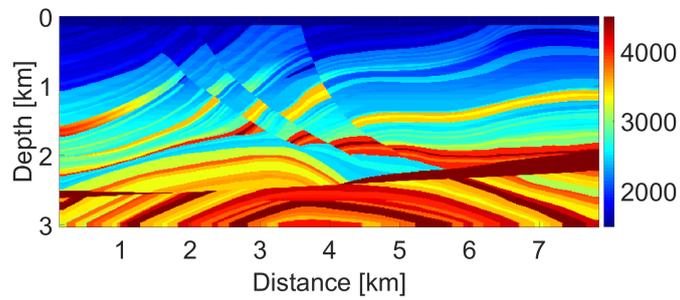


Figure 2.5: P-wave vertical velocity model used to test reciprocity in modeling. Velocities in m/s.

Figure 2.6 shows the data modeled with and without reciprocity at 0, 3250 and 7500 m of offset. The match between the data with and without reciprocity is excellent (the maximum difference is at least $\sim 10^5$ smaller than the maximum modeled value), validating the derived reciprocity relations and their implementation.

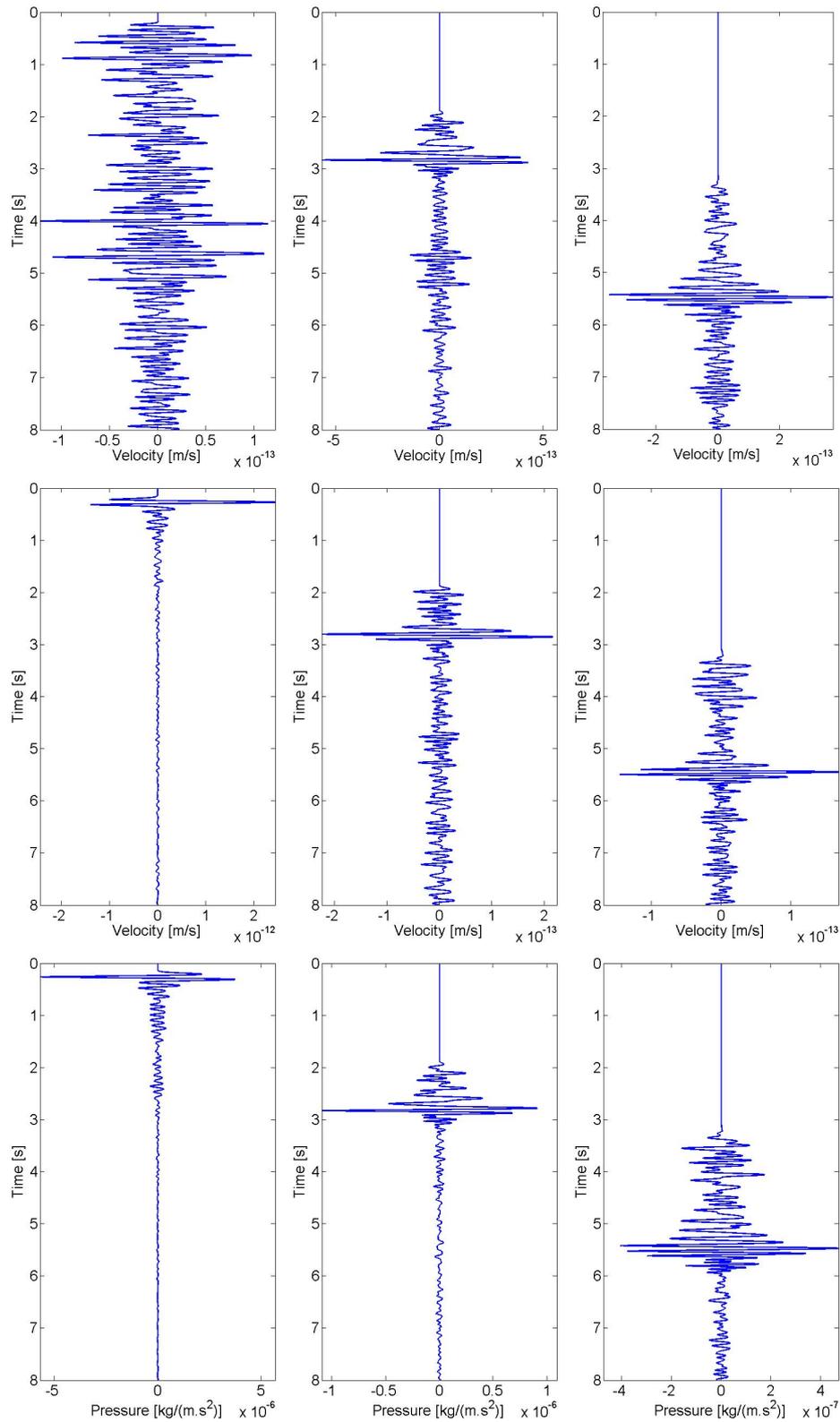


Figure 2.6: Data obtained with (blue) and without (red) reciprocity. Only the blue color is visible because of an excellent match of the data obtained with and without reciprocity. From top to bottom: horizontal velocity component, vertical velocity component and hydrophone component. From left to right: 0, 3750 and 7500 m of offset.

Additional tests using source and/or receiver locations in between gridpoints high-

lighted a limitation of Kaiser interpolation near the seabed. Indeed, when the airgun or the hydrophone is in between grid points and close to the seabed, an interpolation is needed, such as Kaiser interpolation (Hicks, 2002 (83)). If the source or the hydrophone needs an interpolation and if the interpolation window reaches the (anisotropic and/or elastic) seabed, the criteria of applicability of the reciprocity relations are not met anymore and the match is not excellent anymore. Indeed, nothing ensures that the airgun and the hydrophone are completely in the water layer in terms of finite difference modeling. This might be one more reason to modify the receivers and sources locations so that they are on a grid point. Figure 2.7 shows data modeled with and without reciprocity, using the problematic configurations. This mismatching is not observed when geophone's interpolation window reaches the seabed because the geophone is not subject to the same applicability criteria. Note that if one wants reciprocity to reproduce these inexact behaviors of Kaiser interpolation of non-reciprocal modeling, one can use equation 2.14 in isotropic acoustic media and the equivalent equation in isotropic elastic media (not provided here). For example, if a recorded pressure is simulated by interpolating the stress field across six weighted surrounding points where five are effectively in the water and one in the solid part of the model, then the reciprocal experiment would involve spreading the pressure source back across the six points with the same weights, simulating six source positions being excited at once, with the appropriate reciprocity relation applied at each location. This is equivalent to doing six experiments but effectively combining them on the receiver side. If a solution is proposed for elastic seabeds, no simple solution is available for anisotropic media.

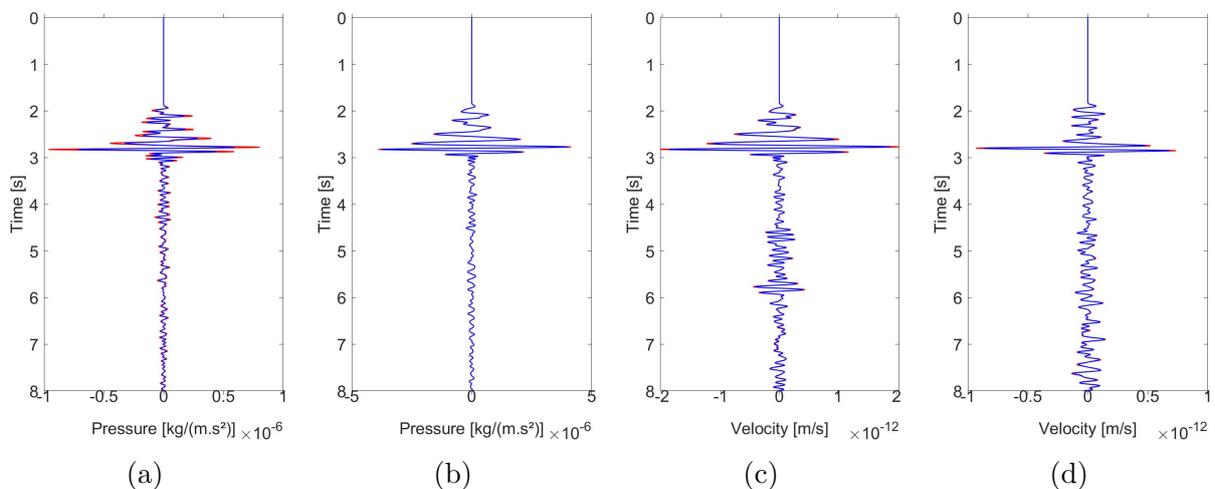


Figure 2.7: Illustration of the limitation of the used reciprocity relations. Data obtained with (blue) and without (red) reciprocity in settings where the Kaiser window of hydrophone or source reaches the seabed. (a) pressure data obtained with a hydrophone whose interpolation window reaches the seabed; (b) pressure, (c) horizontal and (d) vertical velocity obtained with a source whose interpolation window reaches the seabed. 3750 m of offset.

To illustrate the limitation of the used pressure reciprocity relation (i.e., sources and receivers have to be in the same homogeneous isotropic acoustic medium) that is causing the non-reciprocal modeling with Kaiser's interpolation near the seabed, I conducted a fictional similar test where the hydrophone is located inside the elastic VTI sub-surface, at a depth of 1506 m and an offset of 3750 m. Figure 2.8 shows the corresponding trace

together with the trace obtained with interchanged source and receiver locations without any other modification. The fit between the two traces is not good anymore due to the incorrect environment of application. The differences between the two traces not only correspond to a variable amplitude misfit but also to a phase misfit, as visible around time=2 s for instance.

To illustrate the application of the reciprocity relation 2.14 proposed for sources and receivers in different isotropic acoustic media, the same fictional test was performed but setting S-wave velocity V_{sv} , and δ and ϵ anisotropy parameters to zero, creating this way a heterogeneous isotropic acoustic medium. Again, equation 2.14 has been derived for a 3D model while our test model is 2D. Consequently, the used equation is in fact the 2D version of equation 2.14 which happens to be the same as the 3D one. Figure 2.9 shows the successful application of this equation.

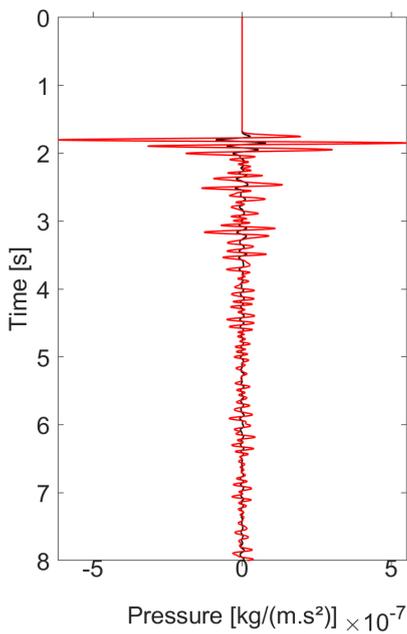


Figure 2.8: Illustration of the limitation of the used pressure reciprocity relation with a fictional test, where the hydrophone is not located in the same isotropic acoustic medium as the airgun as requested by the pressure reciprocity relation, but inside the elastic Vertical Transverse Isotropic (VTI) sub-surface. Pressure trace without reciprocity is in red and the pressure trace obtained with interchanged airgun and hydrophone locations without any other modification is in black.

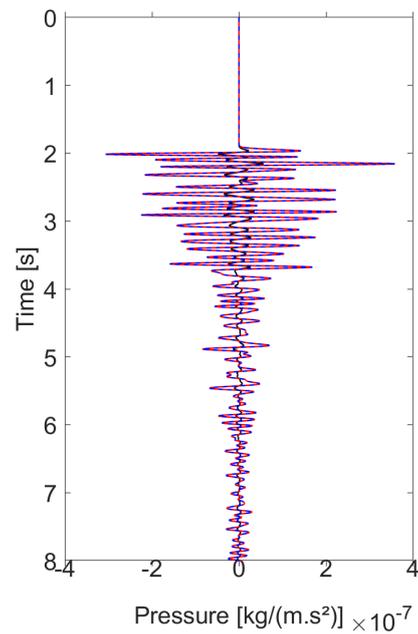


Figure 2.9: Illustration with a fictional test of the proposed pressure reciprocity relation for an airgun and a hydrophone in different isotropic acoustic media. The hydrophone is located inside the isotropic acoustic sub-surface while the airgun is inside the water layer. Pressure trace obtained with reciprocity is in dashed blue, pressure trace obtained without reciprocity in red. The trace obtained with interchanged airgun and hydrophone locations without any other modification is in black.

2.4 Reciprocity for full waveform inversion

2.4.1 A change of point of view

As mentioned in the introduction of this chapter, **FWI** iteratively updates an initial model by minimizing the misfit between observed and modeled seismic data. This misfit is quantified by an objective function. Pressure and velocity data simulations are needed for the objective function computation and the previous section shows us how to apply reciprocity for these data simulations. The objective function computation is not the only aspect of **FWI** that requires modeling. In particular, the minimization of the misfit needs the computation of the gradient of the objective function with respect to the model parameters. This gradient computation also requires simulations: the modeling of two wavefields: the direct and adjoint wavefields (see [Chapter 1](#)). One could try to reproduce these wavefields with reciprocity. The required wavefields are in fact data measurements for all model points. Computing these wavefields with reciprocity would not be a straightforward task since this would require the derivation of additional reciprocity relationships (for diagonal stress relationships for anisotropic sub-surface for example). Instead, a change of point of view is preferred. In this new point of view, the velocity and pressure data are processed as if they were pressure data coming from the adequate sources shot from the receivers locations. Let us take the example of the following ℓ_2 -norm objective function \mathcal{O} , which is the 3D, space discretized and expanded version of equation 1.11 ([Chapter 1](#)):

$$\begin{aligned} \mathcal{O}(\mathbf{m}) = \frac{1}{2} \sum_{\substack{\text{src} \\ \text{rcv}}} \int_{t_1}^{t_2} & \left[w_{v_x, \text{src}}(t, \text{rcv}) \left(v_x^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{mod}}}(\mathbf{m}; t, \mathbf{x}_{\text{rcv}}) - v_x^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \right. \\ & + w_{v_y, \text{src}}(t, \text{rcv}) \left(v_y^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{mod}}}(\mathbf{m}; t, \mathbf{x}_{\text{rcv}}) - v_y^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \\ & + w_{v_z, \text{src}}(t, \text{rcv}) \left(v_z^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{mod}}}(\mathbf{m}; t, \mathbf{x}_{\text{rcv}}) - v_z^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \\ & \left. + w_{P, \text{src}}(t, \text{rcv}) \left(P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{mod}}}(\mathbf{m}; t, \mathbf{x}_{\text{rcv}}) - P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \right] dt. \end{aligned} \quad (2.23)$$

The external sum is over the sources (src) and the receivers (rcv). The time integral is between the start recording time t_1 and the stop recording time t_2 . \mathbf{m} is a column vector containing the N model parameters. For example, for a rectangular model containing Nx points along the horizontal direction, Nz points along the vertical direction and 5 parameter types (such as in 2D **VTI** models), $N = Nx * Nz * 5$. Superscript ^{obs} is used for observed data, and superscript ^{mod} is used for modeled data. $w_{\text{src}, \text{cmp}}(t, \text{rcv})$, with $\text{cmp} \in [v_x, v_y, v_z, P]$, is introduced to:

- represent the (tapered) time selection and (tapered) receiver selection (e.g., offset selection);
- represent the component (cmp) selection, e.g.: pressure component, both velocity components or one of the velocity components.

$w_{\text{src}, \text{cmp}}$ could also be used to balance the contribution of the different components. The gradient for each source and each component is computed separately before being added to the other gradients.

The observed data being processed as if they were the corresponding reciprocal pressure data, the cost function with reciprocity can be written as:

$$\begin{aligned} \mathcal{O}(\mathbf{m}) = \frac{1}{2} \sum_{\substack{\text{src} \\ \text{rcv}}} \int_{t_1}^{t_2} & \left[w_{v_x, \text{src}}(t, \text{rcv}) \left(P^{\text{force}_x(\mathbf{x}_{\text{rcv}}, -\frac{1}{C_{11}^{\text{water}} \dot{p}})}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}_{\text{src}}) - v_x^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \right. \\ & + w_{v_y, \text{src}}(t, \text{rcv}) \left(P^{\text{force}_y(\mathbf{x}_{\text{rcv}}, -\frac{1}{C_{11}^{\text{water}} \dot{p}})}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}_{\text{src}}) - v_y^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \\ & + w_{v_z, \text{src}}(t, \text{rcv}) \left(P^{\text{force}_z(\mathbf{x}_{\text{rcv}}, -\frac{1}{C_{11}^{\text{water}} \dot{p}})}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}_{\text{src}}) - v_z^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \\ & \left. + w_{P, \text{src}}(t, \text{rcv}) \left(P^{\text{airgun}(\mathbf{x}_{\text{rcv}}, p)^{\text{mod}}}(\mathbf{m}; t, \mathbf{x}_{\text{src}}) - P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)^{\text{obs}}}(t, \mathbf{x}_{\text{rcv}}) \right)^2 \right] dt. \end{aligned} \quad (2.24)$$

The gradient for each receiver and each ‘‘component’’ is computed separately before being added to the other gradients. This method requires only minimal modifications of the [FWI](#) code, inversion with pressure data being already available.

Let us recall that: (1) [FWI](#) is in fact a walk along an objective function, from an initial model to a final model; (2) for a given local optimization algorithm, this walk depends only on the objective function and the starting model, aside from numerical errors. As a consequence, because the values of the objective functions are exactly the same in the reciprocal (equation 2.24) and non-reciprocal (equation 2.23) cases, and because the initial models are the same, the inversion results (i.e., the updated models) are exactly the same, apart from numerical errors.

2.4.2 Synthetic applications

I compare inversion results with and without reciprocity. Previous 2D application on modified Marmousi2 model is recycled, however, the central frequency of $\dot{p}(t)$ is set to 5.5 Hz, and more sources and receivers are used: 21 sources with a spacing of 350 m and 219 receivers with a spacing of 35 m. The first source and the first receiver are 180 m away from the left boundary of the model. The initial P-wave vertical velocity Vpv model is a smoothed version of the true Vpv model. The initial SV-wave vertical velocity Vsv and density ρ models are generated by applying the same experimental relationships as previously on the initial Vpv model. The initial anisotropy parameter models are equal to the true models and the true Vsv model is set as equal to the initial Vsv model. I invert for Vpv and I passively update the density using the same empirical relationship as previously. The other models are not updated. The whole time-offset data domain with a increasing frequency multiscale strategy is used (Bunks, 1995 (35)): 0-4 Hz, 0-5.5 Hz, 0-7 Hz and then 0-9 Hz. Each stage contains 15 iterations. No preconditioning is applied to the gradient aside from a slight horizontal and vertical smoothing. Conjugate gradient of Polak-Ribière (Polak en Ribière, 1969 (145)), step-length of Pica et al. (1990) (143) and ℓ_2 -norm cost function are used. To ensure a sensitivity of [FWI](#) to Vpv perturbations in all wave propagation, the selected parameterization is the following: Vpv , Vsv , $\sqrt{1+2\delta}$, $\sqrt{1+2\eta}$ and density, with $\eta = \frac{\epsilon-\delta}{1+2\delta}$ (He et al., 2016 (78)) To be more precise, the logarithm of these parameters is used (Tromp et al., 2005 (194)). I conducted two tests: the first one using the pressure component only, the second one using the vertical velocity component only.

Figure 2.10 shows the true, initial and inverted P-wave vertical velocity V_{pv} models, as well as the differences between the two inverted models. In both cases, the inverted V_{pv} models with and without reciprocity are similar but not equal. The differences could be attributed to numerical errors (such as round-off or truncation errors) or a negligible anomaly in the modeling code.

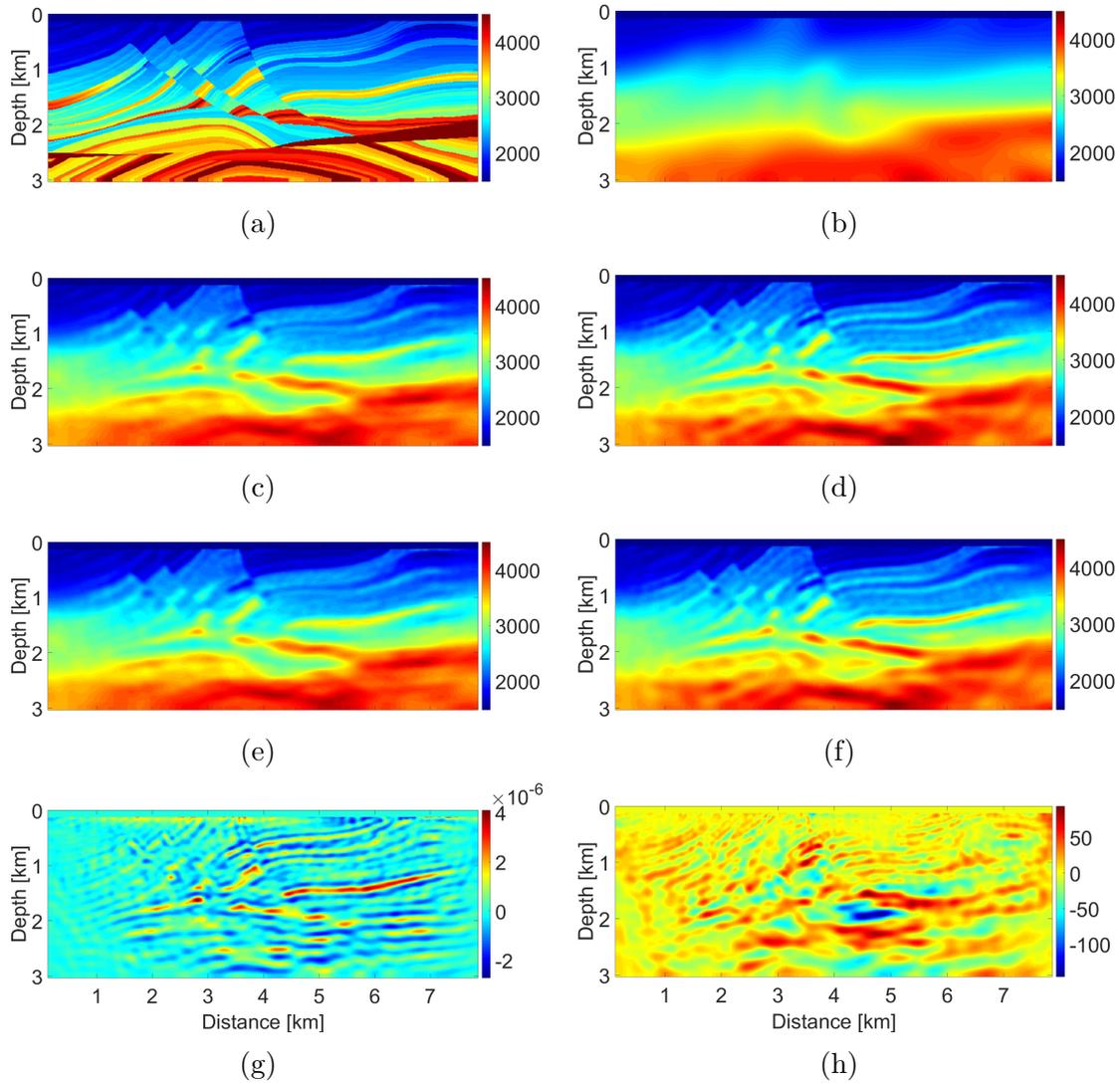


Figure 2.10: P-wave vertical velocity models (in in m/s): (a) true, (b) initial, (c) inverted with pressure component without reciprocity, (d) inverted with vertical velocity component without reciprocity, (e) inverted with pressure component and reciprocity, (f) inverted with vertical velocity component and reciprocity, (g) e-c ,(h) f-d.

For these realistically sized synthetic tests, the inversions with reciprocity lasted ~ 13.25 hours using 21 cores (1 per receiver) while the inversions without reciprocity lasted ~ 40.5 h using 80 cores. All cores were used for shot parallelization and all the computations were made on the same kind of nodes (2xE5-2690 16 core @2.9 GHz). The reciprocity allowed here for a drastic reduction of the computing times: the user computing time (i.e., the time that the user has to wait) was reduced by a factor of ~ 3 , and the total computing time (i.e., the sum of the computing times of each core) was reduced by a factor ~ 10 , as expected.

2.5 Discussion and conclusions

I explained and illustrated how I implemented reciprocity for elastic **VTI FWI** for stationary **OBN/OBS** acquisitions with airgun (or airgun array) sources. This method consists in the replacement of each data component by a single reciprocal pressure, followed by a change of point of view for gradient computation. I would like to point out that the derivations are true for airgun (or airgun arrays) modeled by point moment density sources with no directivity, and data such as: (1) hydrophones recording the pressure in a homogeneous isotropic water layer, and/or (2) geophones recording one or several components of velocity at the seabed. All sources must have the same source time function. I showed that within the area of applicability, the reciprocity relations for data modeling work perfectly, and I emphasized that the change of point of view involves different but reasonable inversion results compared with inversion without reciprocity. Because I proved that, aside from drastically decreased computing costs with reciprocity, the inversion results are theoretically exactly the same with and without reciprocity, these differences are attributed to numerical errors or a negligible anomaly in the modeling code.

Possible solutions for applications which do not meet the required applicability criteria are suggested. I specifically discussed an extension for hydrophones in an isotropic acoustic medium with properties different from the properties at the airguns (or airgun arrays) locations. This case corresponds to the acoustic wave equation, but the derived reciprocity relation is different from the corresponding acoustic reciprocity relation. This is probably due to the fact that acoustic and elastic equations are differently constituted and this highlights the necessity of a thorough derivation of reciprocity relations, as done in this chapter. Additionally, from the provided demonstrations, it is also clear that an equivalent reciprocity implementation for elastic/acoustic isotropic or more general anisotropic seabeds would not be different than the elastic **VTI** one discussed in this chapter (e.g., Operto et al., 2015 (142) for acoustic isotropic seabeds).

All in all, reciprocity is advantageous when the number of sources N_{src} is higher than the number of required components N_{comp} used at once in a particular inversion, multiplied by the number of receivers N_{rcv} :

$$N_{\text{comp}} * N_{\text{rcv}} < N_{\text{src}}, \quad (2.25)$$

At one extremity, if all four components (pressure and the three velocity components) are required, reciprocity is advantageous when the number of sources is more than four times the number of receivers, which is common in **OBN/OBS** acquisitions. Reciprocity becomes more and more advantageous when the number of required components decreases, which is also usually the case, all components being usually not exploited together. Compared with the objective described in figure 2.2, it has not been possible to model the four components of the receiver gather with a single simulation, but we are still able to drastically decrease the computing cost. This application of reciprocity in **FWI** can allow for a drastic reduction of the computing cost, broadening the range of possibilities.

Reciprocity can be used in addition to most of the other (sometimes under development) techniques that make **FWI** more affordable: parallel computing methods (such as domain decomposition or shot parallelization), program optimization methods (such as proper loop nesting for memory access optimization), better but cheap Hessian approximations such as L-BFGS, localized **FWI** or down-sampling during frequency hierarchical strategy. As explained in Chapter 1, parallelization does not decrease the computing needs of **FWI** but allows the use of more computing resources. Proper loop nesting for

memory access optimization reduces the memory access latency, thus the computing time of **FWI** applications. Because it better approximates the Hessian in the approximate Newton's methods used in **FWI**, of **L-BFGS** can reduce the number of iterations, without increasing too much the number of modeling per iteration or the memory requirements, hence reducing the computing time of **FWI** applications (see Métivier et al., 2014 (122) for instance). This is not a characteristic shared by all improved Hessian approximations. When the difference between the initial model and the updated model is localized, which might be the case for time-lapse applications for instance, localized **FWI** can be used to save computing resources: it virtually relocates the source and the receivers in a smaller model around the target thus avoiding the computation of the wavefield where it is not interesting (see for instance Yuan et al., 2017 (233)). Down-sampling during frequency hierarchical strategy consists in the re-sampling with larger steps of models and seismic data, allowed by the filtering of the source time function before modeling for the inversion of the lowest data frequencies (e.g., Bunks, 1995 (35)). Indeed, as explained in chapter [Chapter 1](#), for stability reasons, the sampling steps decrease with the highest used frequency (see for instance Dablain, 1986 (48), Levander, 1988 (110) for the spatial sampling criterion, and Courant et al., 1967 (45), Virieux, 1986 (215), Levander, 1988 (110), Saenger and Bohlen, 2004 (160), and Martin et al., 2008 (116), for the time sampling criterion). Yet, down-sampling strategy might cause problems in shallow water environments, where a large spatial step might interfere with a correct seabed and ocean bottom receivers modeling. However, even though theoretically possible, it might not be advantageous to use reciprocity in combination with source subsampling (also called shot decimation) or an encoded source technique. This question should be the next focus. Shot decimation consists in using only part of the shots at each iteration, and eventually all of them through all iterations. It can be performed through a predefined selection or rotation, or through random rotation as done by Díaz and Guitton (2011) (50). Encoded sources principle is to blend the data from all sources into one or only a few gathers and model all sources at the same time with a single or a few modeling (see for instance Moghaddam et al., 2013 (129)). Both shot decimation and encoded sources techniques reduce the number of simulations per iteration but tends to need more iterations to balance artifacts, unless more advanced techniques are used.

This work highlights two additional points:

- Reciprocity can be used as a method to check a modeling code. Indeed, reciprocity allows for obtaining the same results using two different paths, including for complex simulations. For instance, our reciprocity implementation highlighted a limitation of Kaiser's interpolation method near the seabed.
- Pressure and vertical velocity components behave differently in **FWI** (see the different inversion results in figure 2.10), which participates in the attractiveness of multi-component seabed data, and needs further investigations. The differences can have multiple causes, among them: (1) the velocity components are recordings of the wavefield at the seabed, while the pressure component is the recording of the wavefield in the water. Consequently, these two kinds of components do not record the same waves; (2) In the same way as the velocity components are higher frequencies than the displacement components due to the time derivative, the velocity components may, by essence, not have the same frequency content as the pressure component; (3) the pressure recording is a non-directional measurement. The relative weights of the waves in a recording are consequently different from the weights from the vertical velocity component.

Chapter 3

Full waveform inversion to retrieve P- and S-wave vertical velocities in soft seabed vertical transverse isotropic media: inversion strategy, parameterization and sensitivity study

Chapter's objectives and highlights

- Recall the “soft seabed strategy” of Sears et al. (2008) ([164](#)), which was designed to retrieve P- and SV-waves velocities in soft seabed environments. Emphasize the limitations of the context in which the strategy was developed, especially the isotropic aspect.
- Propose a suitable full waveform inversion parameterization, for the application of the strategy in vertical transverse isotropic media. The parameterization selection is based on analytically computed patterns. A synthetic illustration of the effect of a poor choice of parameterization is also provided.
- Appraise the performance of the strategy to retrieve P- and S-wave vertical velocities in vertical transverse isotropic contexts, when the remaining anisotropy parameters are not updated and slightly erroneous, as occurs in real data applications. A partial preliminary assessment is made through reflection coefficient computations. The appraisal is finally performed with synthetic inversions and concludes to a perturbed inversion but that can still be successful.
- Describe a variant of the soft seabed strategy where the increasing frequency multiscale strategy of Bunks et al. (1995) ([35](#)) is combined with modern data low-frequency content, and substitutes the essential intermediate-scale S-wave velocity features inversion stage. Indeed, this stage could fail due to wrong attenuation model for instance. This variant, however, shows an increased sensitivity to wrong anisotropy parameters.

Contents

3.1	Introduction	56
3.2	The soft seabed strategy	58

3.3	A choice of parameterization	62
3.4	Amplitude sensitivity of wide-aperture P-P reflections to elastic vertical transverse isotropic parameters	68
3.5	Synthetic application of the soft seabed strategy with slightly erroneous anisotropy parameters	70
3.6	A variant to the soft seabed strategy	71
3.7	Discussion and conclusions	73

3.1 Introduction

In the introductory chapter, I explained how Full Waveform Inversion (**FWI**) became a reputed method to obtain quantitative models of mechanical parameters of the subsurface, from seismic data. Introduced by Tarantola in 1984 for acoustic media (*187*) then for general media in 1988 (*189*), **FWI** showed its value through multiple and diverse applications (**Chapter 1**). However, limited computing resources and lack of good inversion strategies led to approximate choices of realism (e.g., acoustic instead of elastic, isotropic instead anisotropic, and/or elastic instead visco-elastic modeling). In this introduction, I hereafter describe the evolution of **FWI** choices of realism, from the first one to the elastic anisotropic one explored in this thesis. For the completeness of this chapter, I recall pieces of information provided in the introductory chapter, then go into more details.

As a matter of fact, in the Earth local exploration context (i.e., dimensions of a few dozens of kilometers), **FWI** algorithms were initially mainly limited to the acoustic isotropic choice of realism, and a single parameter was inverted: the P-wave velocity. Indeed, this parameter is then the dominant one. However, the Earth being partially solid, its elastic nature is unequivocal. At the global scale, the most telling proof could be the propagation and disappearance of earthquake SV-waves, which highlighted the core-mantle boundary between the solid lower mantle and the liquid outer core. This boundary (also called Gutenberg, Oldham-Gutenberg, or Wiechert-Gutenberg discontinuity) is located at approximately 2900 km depth beneath the Earth's surface, and was discovered by Gutenberg in 1912, as recounted by Knopoff (1999) (*100*). At a more local scale (i.e., the focus of this Ph.D. thesis), the examples are also numerous, as evident from the here-under publications (to cite a few: Brossier et al., 2009 (*31*); Sears et al., 2010 (*165*); Vassiliou et al., 2017 (*205*)).

I remind that the data modeling differences between an acoustic and an elastic medium not only consist in missing waves, but also in erroneous phase and amplitude, as illustrated by Chapman et al. (2018) (*40*). I also remind that in practice, the strength of the errors coming from an acoustic inversion of elastic data ranges from minimal to significant, depending notably on the medium and acquisition properties (e.g., Barnes and Charara., 2008 (*16*); Vigh et al., 2009 (*209*); Marelli et al., 2012 (*117*); Solano et al., 2013 (*178*); Raknes et al., 2015 (*148*); Agudo et al., 2018 (*2*) and Mora and Wu, 2018 (*132*) for synthetic tests; Arnulf et al., 2014 (*11*) and Marjanović et al., 2019 (*118*) for real data inversions). In the context of the use of the whole data waveform, in comparison with travel time tomography, these modeling errors are expected to lead to inversion errors. As a general conclusion from these research works, if the S-waves imprint on the data is strong, the inversion errors are expected to be significant: for instance, strong velocity contrasts increase the imprint of S-waves.

To overcome the approximate modeling, Agudo et al. (2018) (2) proposed to pre-process the data to remove the elastic effects, and reduced the inversion errors in their inverted P-wave velocity models. A data selection of less affected P-waves can also allow for better results (e.g., Arnulf et al., 2014 (11)), but is not always possible for acquisition and physical reasons (e.g., Marelli et al., 2012 (117) and Marjanović et al., 2019 (118)). Another strategy to improve the P-wave velocity models is to replace the acoustic approximation by the elastic one. This strategy was allowed by the growth of computing resources. Barnes and Charara (2009) (16) showed that even an elastic FWI where only the P-wave velocity is updated could lead to improved P-wave velocity models. The SV-wave velocity model is usually passively updated based on an empirical relationship with P-wave velocities (e.g., Raknes et al., 2015 (148) and Huot and Singh, 2018 (84)). Arnulf et al., (2014) (11) and Marjanović et al. (2019) (118) actively updated both P- and SV-wave velocities, but the obtained SV-wave velocity models were regarded as not reliable. All three strategies aim at providing better inverted P-wave velocity models than simple acoustic inversion, through a more accurate data modeling.

Adequate empirical relationships are not always available and never perfect. Moreover, empirical relationships can show a large variability (e.g., Mondol et al., 2010 (130)) and detrimental effects when poorly chosen (e.g., Raknes et al., 2015 (148)). Consequently, the need for meaningful S-wave velocity models inversion is apparent. Besides allowing for better P-wave velocity model updates, such S-wave velocity models is likely to, among other things: provide better imaging through gas-filled sediments (e.g., Granli et al., 1999 (68)); higher resolution images (e.g., Borisov and Singh, 2015 (29)); improve lithology, porosity or fluid characterization by providing a complementary point of view on the media (e.g., Li and Zhang, 2011 (111); Szymańska-Małysa and Dubiel, 2019 (185)); and characterize fractures when combined with azimuthal anisotropy (e.g., Zhang et al., 2017 (234)). Other examples are provided by Stewart et al. (2003) (182). As a consequence, more and more efforts are made to invert meaningful both P- and S-wave velocities.

Several successful P- and S-wave FWI applications have been published. In land environments, one can cite Brossier et al. (2009) (31), Zhang et al. (2017) (234), Zhang et al. (2019) (236) and Ren et al. (2019) (155) for synthetic applications; Shi et al. (2007) (173) for both synthetic and real applications, and Vassiliou et al. (2017) (205) for real applications. In marine environments, Sears et al. (2008) (164) defined strong (or high) and soft (or low) seabeds using the seabed S-wave velocities. Seabeds with S-wave velocities above 1000 m/s are considered as strong seabeds. For example, those seabeds can be found near mid-ocean ridges, where the magmatic rocks are free of sediments (e.g., Mitchell, 1998 (126)). Indeed, magmatic rocks show high S-wave velocities (e.g., Tsuji and Iturrino, 2008 (197)). Seabeds with S-wave velocities typical of siliclastic seabed sediments (around or below 300 m/s, see Hamilton, 1976 (74), and Castagna et al., 1985 (39)) are considered as soft seabeds. Choi and Shin (2007) (42), Barnes and Charara (2009) (16), Borisov and Singh (2015) (29) and Mora and Wu (2018) (132) published successful results on synthetic strong seabed applications. For soft seabeds however, the inversion suffers from a poorer S-wave velocity information content because of a decreased S-wave data content (Sears et al., 2008 (164) and Barnes and Charara, 2009 (16)), explaining the failing of usual inversion strategies. In that context, Sears et al. (2008) (164) proposed a specific 2D strategy based on seabed receivers. This strategy is hereafter called the “soft seabed strategy”. This strategy is, to my knowledge, the only published successful strategy, and was applied on real data by Sears et al. (2010) (165). This strategy was developed in isotropic media.

However, as written and justified by Levin and Park (1998) (109) in their introduction, “[...] seismic isotropy [...] may actually be a rarity rather than the rule in the shallow earth”. The shallow earth means here the lithosphere, i.e. the outermost shell of the Earth composed of the crust, which can reach a thickness of a few dozens of kilometers, and the upper mantle, which goes at depths up to 700 km. Consequently, the upper mantle corresponds to more than 25 % of the Earth’s volume. Not only the shallow Earth is anisotropic, but also the lower mantle (e.g., Thomas and Kendall, 2002 (190); Ferreira et al., 2019 (56)) and the inner core (e.g., Mattesini et al., 2010 (120); Song, 1997 (181); Tromp, 1993 (196)). Anisotropy causes are provided in the introductory chapter. This anisotropic nature of Earth, combined with the progressively wider range of wave propagation directions in data, explains why the anisotropy approximation is increasingly occurring in FWI (e.g., Prioux et al., 2011 (146); Bergslid et al., 2015 (19); Gholami et al., 2013 (66); Kamath et al., 2017 (88); He et al., 2018 (79)).

This chapter aims at making another step towards a more realistic choice of realism, by illustrating the performance of the soft seabed strategy in anisotropic media. To be more precise, I focus on inversions where only P- and SV-wave vertical velocities are inverted and where the remaining anisotropy parameters are slightly erroneous, as expected in real data applications. To simplify, those remaining anisotropy parameters are hereafter called “anisotropy parameters” or “remaining anisotropy parameters”. Slightly erroneous models of the anisotropy parameters are models which allow for an approximate but not perfect arrival-time and phase fitting of both wide- and narrow-aperture P arrivals, without cycle-skipping, allowing for a reasonable P-wave velocity inversion. Following the choice of realism explained in the introductory chapter, I limit the study to the Vertical Transverse Isotropic (VTI) context, where the wave velocities depend on the angle of the wave direction of propagation, measured from the vertical direction (e.g., Thomsen, 1986 (192)).

I first recall the need, reasoning and context behind the soft seabed strategy. I then justify the choice of the FWI parameterization. This step is imposed by the change of choice of realism: Sears et al.’s (2008) (164) parameterization cannot be used because it cannot describe an elastic VTI media. I explain and illustrate why the parameterization should be carefully selected. I finally illustrate the performance of the soft seabed strategy with a slightly erroneous anisotropy model, using synthetic inversions of horizontally layered models with sharp interfaces. A preliminary partial assessment is made through reflection coefficient computations. Before concluding, I propose a simple evolution of the strategy, that shows potentials in other circumstances.

3.2 The soft seabed strategy

The strategy developed by Sears et al. (2008) (164) to invert both P- and SV-wave velocities in soft seabed environments uses multi-component seabed data, and was developed using 1- and 2D isotropic synthetic tests.

Sears et al. (2008) (164) illustrated the need for a dedicated inversion strategy. Figure 3.1 provides this illustration. They compared the application of the usual inversion strategy where the whole data and whole frequencies are inverted at once, on two 2D synthetic datasets mimicking a strong and a soft seabed environment. Both P- and SV-wave velocities are inverted in an alternate manner over iterations. Density is updated using the same empirical relationship as the one used to generate the true and initial density models. A free surface and a typical airgun-seabed acquisition are used. In the strong

seabed environment, both P- and SV-wave velocities are well retrieved. However, in the soft seabed case, poorer P-wave and bad SV-wave velocities are obtained, explaining the need for an advanced inversion strategy.

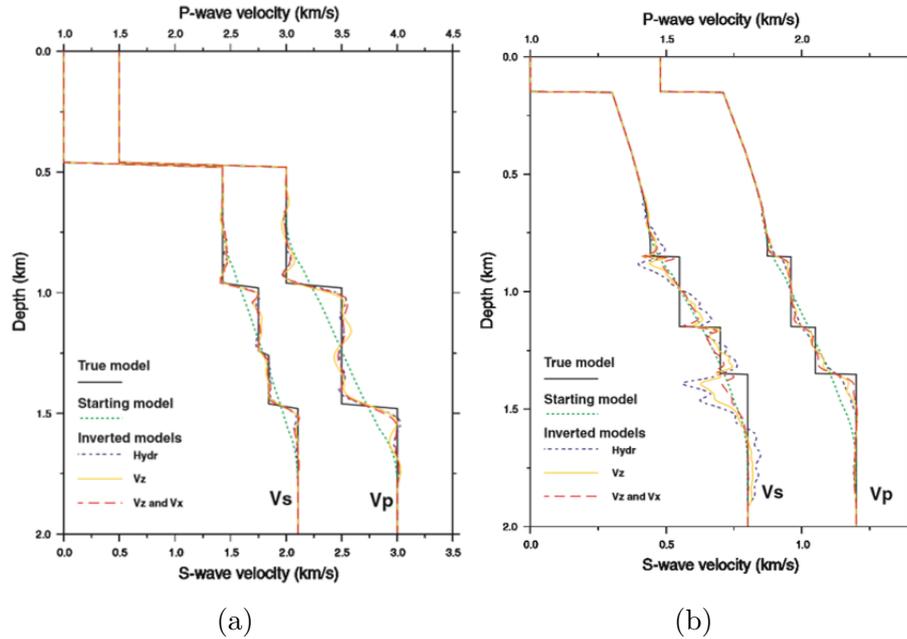


Figure 3.1: Inversion results for a (a) strong seabed environment (S-wave seabed velocity = 1430 m/s), (b) soft seabed environment typical from a clastic sedimentary basin (S-wave seabed velocity = 300 m/s). The same simple inversion strategy, realistic source and acquisition geometry are used in both cases. Figures from Sears et al. (2008) (164).

Sears et al. (2008) (164) explained these different results by a poorer amount of S-wave velocity information in the data for the soft seabed case, due to a decreased amount of S-waves. Indeed, with airgun acquisitions, due to the fluid nature of water, no S-waves are generated by the source. Consequently, the inversion entirely relies on the SV-waves generated by P- to SV-wave conversions in the sub-surface. Compared with strong seabeds, poor P to S conversions (i.e., transmission) occur at soft seabeds, which is usually a major place for conversions. Moreover, if any occur, they arrive at much later times, as visible on the corresponding modeled data (see figure 3.2).

Tarantola (1986) (188) suggested that the most dominant parameter should be inverted first. Following this recommendation, the soft seabed strategy consists notably in the inversion of P-wave velocities followed by the inversion of SV-wave velocities. In order for the P-wave velocities to benefit from the improved SV-wave velocities and the added data information, the soft seabed strategy allows P-wave velocities to be updated during the SV-wave velocity inversion stage. Indeed, as explained in the introduction, the P-wave velocity updates can be affected by the SV-wave velocities.

Moreover, based on data observation (see figure 3.2), Sears et al. (2008) (164) selected the data for each stage such that the weight of the desired velocity is increased: pressure or vertical velocity component is used for the first stage (where P-waves dominate), horizontal velocity component is used for the last phase (where P-S reflections and the little P to S conversions at the seabed dominate), eventually combined with vertical velocity component. The selection of the S-wave-related data is approximately performed through time-offset selection (figure 3.2).

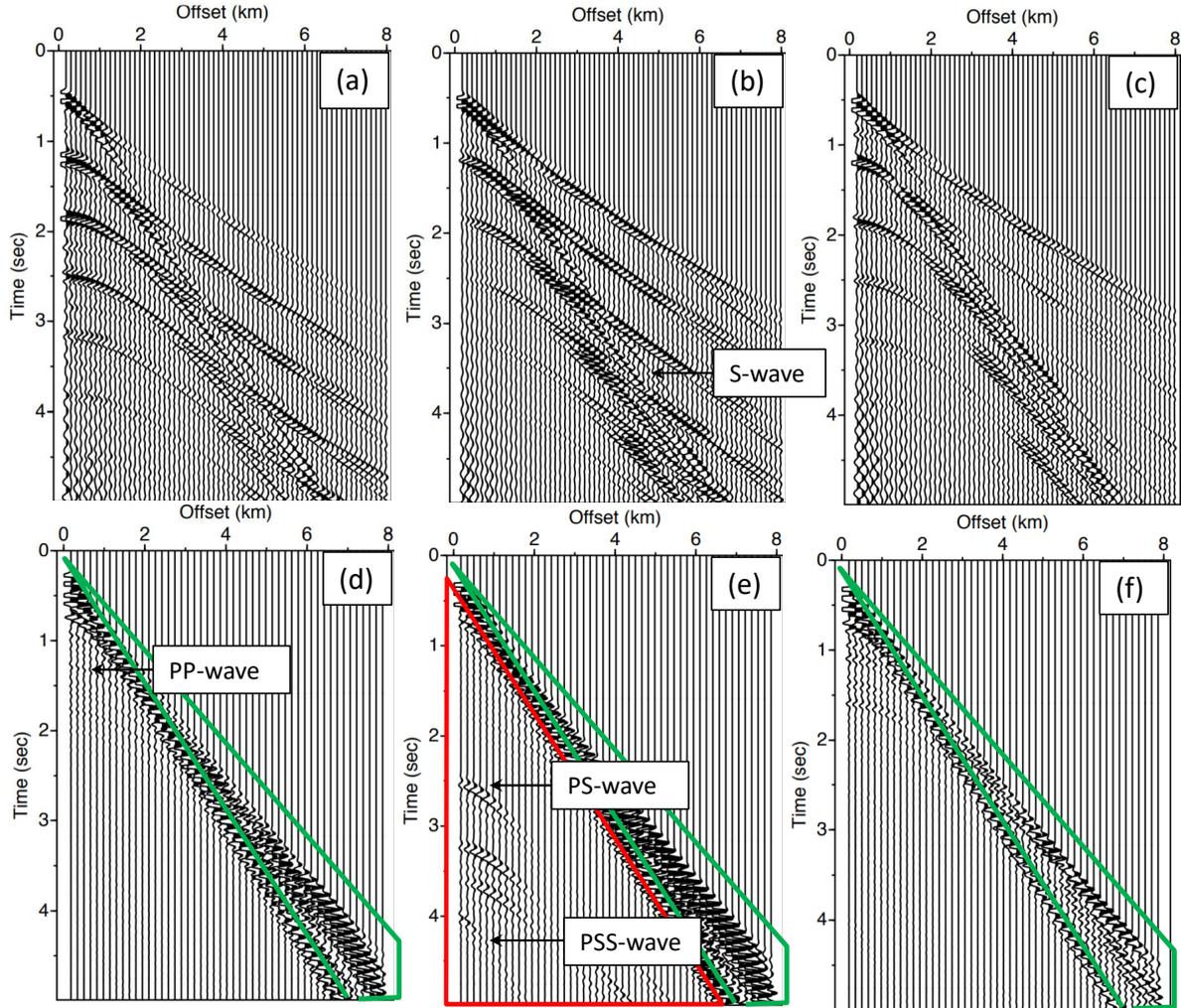


Figure 3.2: Modeled data from strong (top: a-b-c) and soft (bottom: d-e-f) seabed models. Left (a-d): pressure component; Middle (b-e): horizontal velocity component; Right (c-f): vertical velocity component. Red (resp. green) polygon: S-wave-related (resp. wide-aperture) data approximate selection. Figure modified from Sears et al. (2008) (164).

Sears (2007) (166) mentioned that the inverted SV-wave velocities were lacking of long scale updates. To overcome this problem, inspired by the decreasing aperture multiscale strategy of Shipp and Singh (2002) (172), Sears (2007) (166) and Sears et al. (2008) (164) proposed to add an intermediate inversion stage. This additional stage uses the amplitude effect of SV-wave velocities on wide-aperture P-P reflection data. Indeed, synthetic tests performed by Sears (2007) (166) and Sears et al. (2008) (164) showed that inversion with these waves are able to retrieve intermediate scale SV-wave velocity features. Moreover, the wide-aperture P-P reflection amplitudes show no sensitivity to densities (Sears, 2007 (166)). Consequently, in an isotropic case, assuming the effects - including the amplitude effect - from P-wave velocities have been taken into account during the first stage, the only remaining parameter affecting these amplitudes are the SV-wave velocities. This parameter can then be updated. The selection of the wide-aperture P-P reflection data is also approximately performed through time-offset selection on the pressure or vertical velocity component (figure 3.2).

The need for an additional phase dedicated to intermediate scale updates can also be encountered during the P-wave velocities inversion. However, the need is increased with

SV-wave velocities since the size of the updates depends on the velocities, higher velocities giving larger updates (following the same reasoning as for equation 1.16 of Chapter 1). SV-wave velocities being smaller than P-wave velocities, the size of the largest achievable update from P-S reflection is much higher than for P-P reflections.

To summarize, the soft seabed strategy identifies different sources of information that can be isolated, and use them in a sequential manner following their domination and resolution power:

1. Intermediate and small scale P-wave velocities inversion from pressure or vertical velocity component;
2. Assuming a correct (enough) P-wave velocity model, intermediate scale SV-wave velocity features retrieval from their amplitude effect on P-P wide-aperture reflections. Those reflections are approximately selected with time-offset mute on the pressure or the vertical velocity components. The P-wave velocities are allowed to be updated too, to benefit from the improved SV-wave velocity model;
3. Small scale SV-wave velocity features inversion from P-S reflections (and probably the poor P to S conversion at the seabed). Those reflections are approximately selected with time-offset mute on the horizontal velocity component (and eventually the vertical velocity component). The P-wave velocities are allowed to be updated too, to benefit from the improved SV-wave velocities and the added data information.

Note that the strategy uses multi-component seabed data with (1) a P-P and P-S reflections separation over components, as expected in horizontally or sub-horizontally layered media, and (2) both wide- and narrow-aperture waves recording.

Sears et al. (2008) (164) applied their strategy on the velocity components of the soft seabed case of figure 3.1. Figure 3.3, modified from Sears et al. (2008) (164), illustrates the success of such strategy. Similar but slightly poorer models were obtained using pressure instead of vertical velocity component (not shown). The differences were interpreted as the consequence of stronger water reverberations on the pressure component.

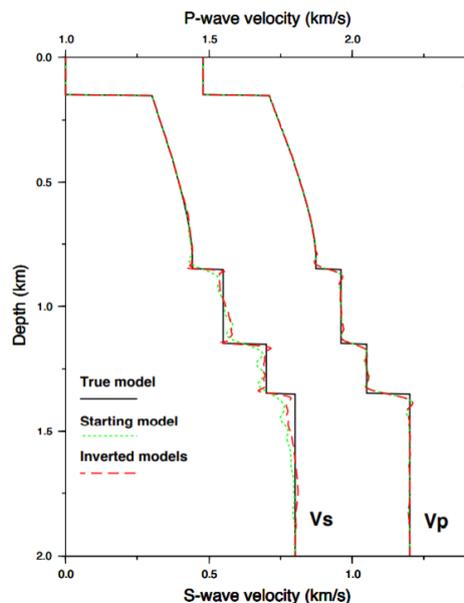


Figure 3.3: Inversion results of the last stage of the soft seabed strategy applied to the same initial model and data of figure 3.1. Figure modified from Sears et al. (2008) (164).

3.3 A choice of parameterization

N.B.: A 2D (resp. 3D) elastic **VTI** medium is described by five (resp. six) mechanical parameters, for example the density (mass per unit of volume), and the following four (resp. five) elastic moduli (force per unit of surface), as usually done in wave equations for general media: $C_{11}, C_{13}, C_{33}, C_{44}, \rho$ (resp. $C_{11}, C_{13}, C_{33}, C_{44}, C_{66}\rho$). Other parameterizations are possible, as long as they can be expressed using density and all above elastic moduli only. This chapter uses the following parameters: P- and SV-wave vertical velocities V_{pv} and V_{sv} , the normal move-out P-wave velocity V_{pm} , Thomsen parameter γ , $\sqrt{1+2\epsilon}$ and $\sqrt{1+2\delta}$ modified Thomsen parameters, $\sqrt{1+2\eta}$ and density ρ (Thomsen, 1986 (192)). The meaning of these parameters is provided in [Appendix F](#). Even if part of the used parameters are defined within the weak anisotropy approximation or the short offset reflection domain, they can also be used outside of it, as done in this chapter. This questioning is caused by the fact that the anisotropy parameters of Alwyn-like environments do not meet the weak anisotropy criteria. [Figure 3.4](#) illustrates the difference in P- and SV-wave velocities at all propagation angles, with and without the weak anisotropy approximation.

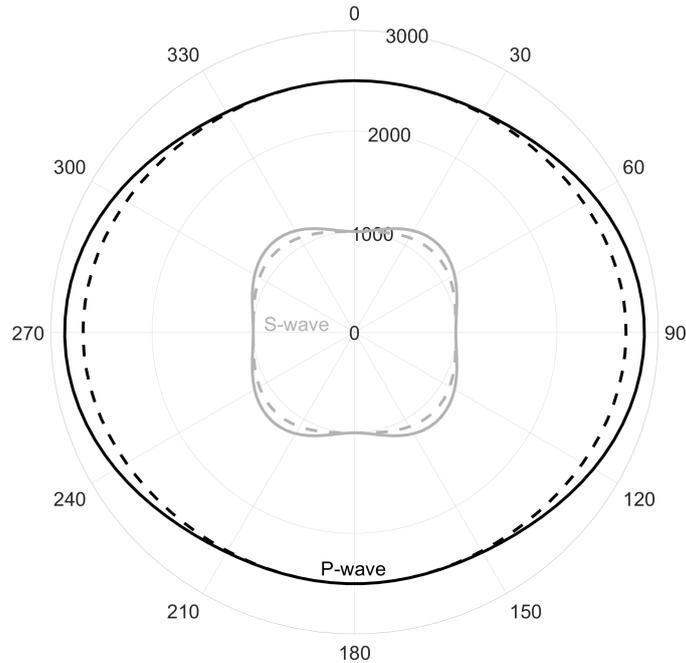


Figure 3.4: P- and SV-wave velocities at all propagation angles with (dashed line - equation [F.5](#)) and without (solid line - equation [F.1](#)) the weak anisotropy approximation. Note the differences with the isotropic case, where the velocity would be the same for all angles (use the thin circles for reference). The medium parameters are those of the bottom layer from [table 3.2](#). The angle is measured from the vertical axis.

As mentioned in [Chapter 1](#), **FWI** results depend on the parameterization. The choice of the parameterization is critical for a successful **FWI**. For instance, He et al. (2018) ([79](#)) showed significantly distinct inversion results obtained with different parameterizations. Indeed, **FWI** sensitivity to a parameter perturbation depends on the used parameterization, but also and the used wave type(s) (i.e., P-P reflections or P-S reflections for example) and the directions of wave propagation.

The soft seabed strategy uses the P- and SV-wave velocities and density parameteri-

zation (V_p, V_s, ρ) , and relies mainly on P-P and P-S wide- to narrow-aperture reflections, to update P- and SV-wave velocities. To reproduce Sears et al.'s (164) inversion results with an elastic VTI FWI algorithm, the chosen parameterization(s) should exhibit(s) a sensitivity similar to that of the (V_p, V_s, ρ) parameterization, and that for the inverted velocity parameters and for P-P and P-S wide- to narrow-aperture reflections. To a certain extent, the sensitivity can be provided by the reflection radiation patterns, which are analytically computed using the high frequency approximation and the Born approximation. Born approximation is valid for weak perturbations. The P-P and P-S reflection radiation patterns for the P- and SV-wave velocities with the (V_p, V_s, ρ) parameterization as computed by Stolt and Weglein (2012) (183) are shown in figure 3.5. These patterns indicate a sensitivity to P-wave velocity perturbations at all apertures for P-P reflections, and a sensitivity to SV-wave velocity perturbation mostly around intermediate apertures for P-P and P-SV reflections.

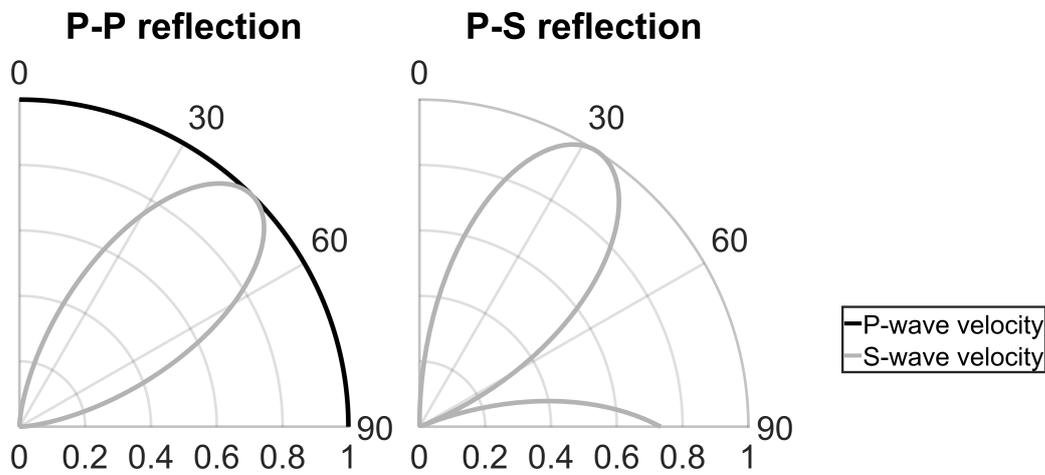


Figure 3.5: Normalized P-P and P-SV reflection radiation patterns for P- and SV-wave velocities with the (V_p, V_s, ρ) parameterization. The angle corresponds to the incident angle. V_p , V_s and ρ are respectively set equal to the P- and SV-wave vertical velocities and density of the bottom layer from table 3.2. Computed with formula from Stolt and Weglein (2012) (183).

Following He and Plessix (2016) (78) but without the weak anisotropy approximation, the P-P and P-SV reflection radiation patterns for various elastic VTI parameterizations have been computed. Among them, the $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$ parameterization shows the desired sensitivity (see figure 3.6). Note that the weak anisotropy approximation has a poor effect on the computed patterns, as qualitatively shown in figure 3.7, or as quantified by He and Plessix (2016) (78).

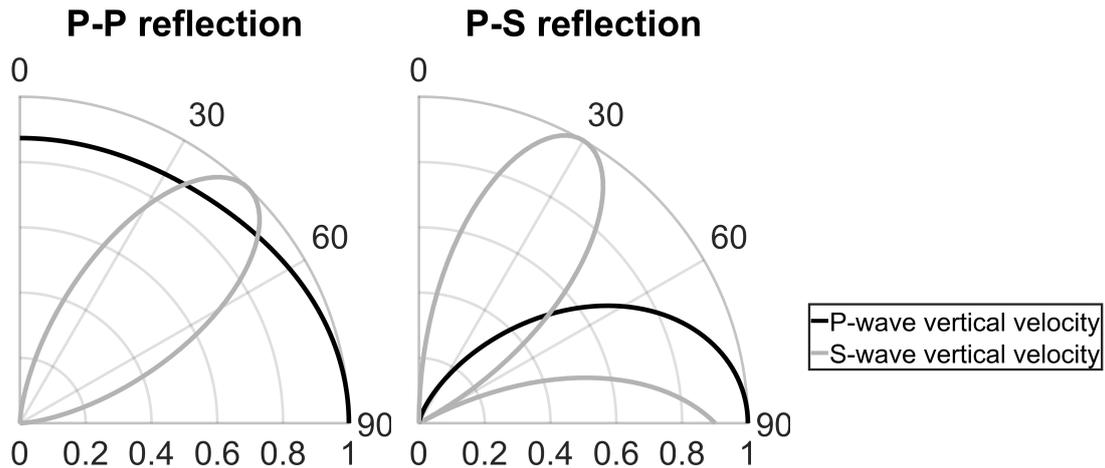


Figure 3.6: Normalized P-P and P-SV reflection radiation patterns for P- and SV-wave velocities with the $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$ parameterization. The angle corresponds to the incident angle. The medium parameters are those of the bottom layer from table 3.2. The patterns are computed with formula from He and Plessix (2016) (78).

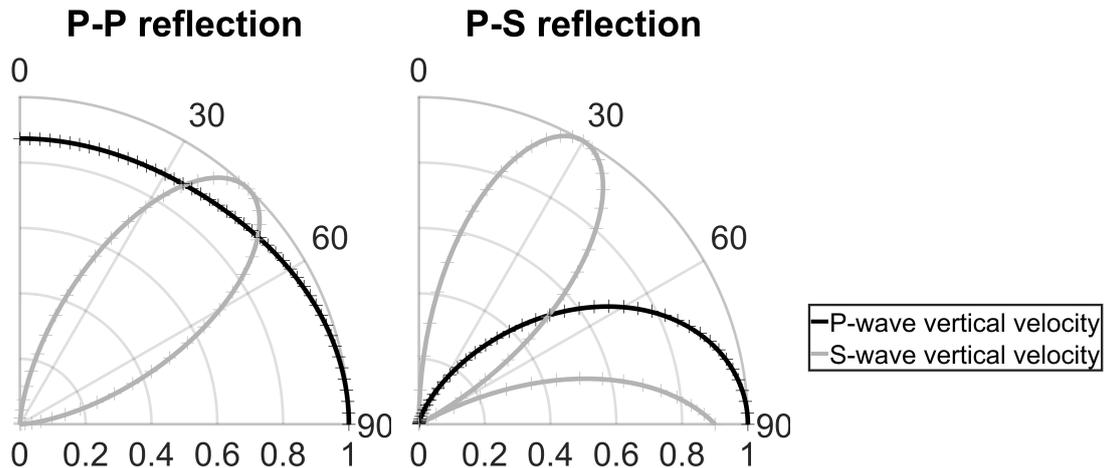


Figure 3.7: Comparison of non weak (solid lines) and weak (crosses) approximation for normalized P-P and P-SV reflection radiation patterns for P- and SV-wave velocities with the $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$ parameterization. The angle corresponds to the incident angle. The medium parameters are those of the bottom layer from table 3.2. The patterns are computed with formula from He and Plessix (2016) (78).

Reflection radiation patterns allow for a quick parameterization pre-selection but suffers from limitations. For instance, in addition to the Born approximation, those patterns do not take into account propagation effects and other types of waves. The pre-selected parameterization should thus been validated with synthetic inversion tests. Consequently, this parameterization is tested on a synthetic soft seabed 2D FWI application, inspired from the gradient layer low velocity model application of Sears et al. (2008) (164), and using anisotropy parameters estimations of Alwyn's area. The true model is made of horizontal layers whose properties are provided in table 3.1. The model contains a few complexities: (1) anti-correlated P- and S-wave vertical velocities in the fourth layer where the S-wave velocity updates cannot be due to a leakage from P-wave vertical velocity, (2) (partial) low-velocity zones in the fourth and sixth layers. The density model ρ is computed from the initial P-wave vertical velocity model V_{pv} with the empirical relationship

2.22. This equation is recalled hereafter for the chapter's completeness:

$$\left\{ \begin{array}{l} \rho = 1030 \text{ kg/m}^3 \text{ in water,} \\ \rho = 2351 - 7497 * (V_{pv} * 10^{-3})^{-4.656} \text{ kg/m}^3 \text{ if } V_{pv} < 2000 \text{ m/s,} \\ \rho = x_a + \frac{V_{pv}-2000}{150} * (x_c - x_a) \text{ kg/m}^3 \text{ if } V_{pv} \geq 2000 \text{ m/s and } V_{pv} \leq 2150 \text{ m/s,} \\ \text{with } x_a = 2351 - 7497 * (2000 * 10^{-3})^{-4.656} \text{ and } x_c = 1741 * (2150 * 10^{-3})^{0.25}, \\ \rho = 1741 * (V_{pv} * 10^{-3})^{0.25} \text{ kg/m}^3 \text{ if } V_{pv} > 2150 \text{ m/s,} \end{array} \right. \quad (3.1)$$

where the empirical expression of Hamilton (1978) (75) is used for the shallow seabed poorly consolidated sediments ($V_{pv} < 2000$ m/s), and the empirical expression of Gardner et al. (1974) (65) is used for the deeper consolidated sediments ($V_{pv} > 2150$ m/s). A linear transition between the two empirical relations is applied. Note that these two empirical relationships were derived assuming isotropic media and are here applied in anisotropic media using the P-wave vertical velocities V_{pv} in place of P-wave velocities. A free surface is set on top, and absorbing boundaries are set on the other model boundaries. The model is 9-km-wide and 2-km-deep.

Layer	V_{pv} [m/s]	V_{sv} [m/s]	$\sqrt{1 + 2\delta}$ [-]	$\sqrt{1 + 2\epsilon}$ [-]	Thickness [m]
1	1480	0	1	1	150
2	1710-1870	300-440	$\sqrt{1 + 2 * 0.0051}$	$\sqrt{1 + 2 * 0.0051}$	550
3	1960-2050	550-600	$\sqrt{1 + 2 * 0.0253}$	$\sqrt{1 + 2 * 0.04551}$	200
4	1950-2050	700-750	$\sqrt{1 + 2 * 0.4}$	$\sqrt{1 + 2 * 0.05}$	300
5	2200-2300	900-1000	$\sqrt{1 + 2 * 0.1453}$	$\sqrt{1 + 2 * 0.1949}$	300
6	2200-2300	900-1000	$\sqrt{1 + 2 * 0.08049}$	$\sqrt{1 + 2 * 0.08049}$	500

Table 3.1: True model parameters. Linear gradients are defined by their extreme values.

The initial models are the true anisotropy parameters and smooth versions of the P- and S-wave vertical velocities. The initial density is generated by applying the same empirical relationship as the one use to generate the true density model. The sources are point moment density sources modeling airguns or airgun arrays (see for instance Arntsen et al., 2000 (8)). The source time function is estimated from Alwyn real data (figure 3.8). 400 sources at a depth of 12 m and with a spacing of 20 meters are used. 24 seabed receivers with a spacing of 350 m are used. The hydrophones are located on the first finite-difference grid point above the seabed and the geophones are located on the first grid point below the seabed. The maximum offset is of ~ 8 km.

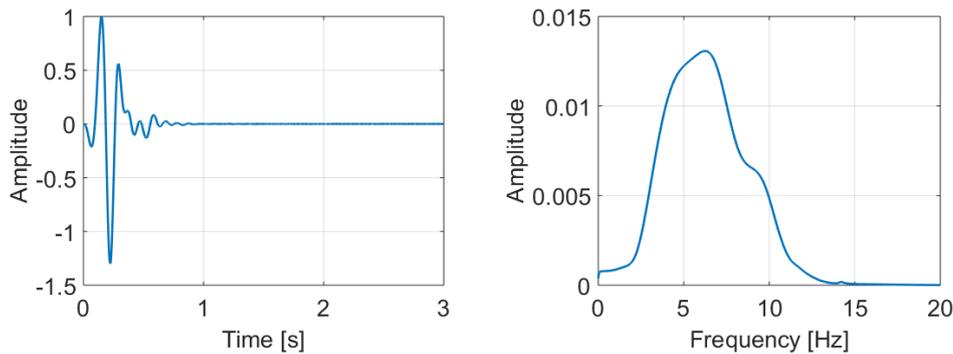


Figure 3.8: Used source time function and its frequency spectrum.

Inversion is performed with vertical and horizontal velocity components. Reciprocity as explained in [Chapter 2](#) is used to decrease the computing cost. The subspace method of Kennett et al. (1988) ([94](#)) for the steplength computation and the Polak-Ribière (1969) ([145](#)) conjugate gradient method are used. The cost function is an ℓ_2 -norm cost function. The source time function is assumed to be known. The density is passively updated applying the same empirical relationship as the one use to generate the true density model. In an attempt to balance the contribution to FWI of the various parameters, the logarithm of the selected parameterization is used (Tromp et al. (2005) ([194](#))), which does not affect previous conclusions. Following Sears et al. (2008) ([164](#)), the water layer and shallow sub-surface is assumed to be known or previously estimated. No other preconditioning is applied on the gradient except for a mute around the receivers, and a horizontal gaussian smoothing (whose $\sigma \sim 287$ m) to remove the imprint of the receivers sparsity. The results of the application of the soft seabed strategy are shown in [figure 3.9](#). These successful results, similar to the one obtained by Sears et al. (2008) ([164](#)), support the selected parameterization.

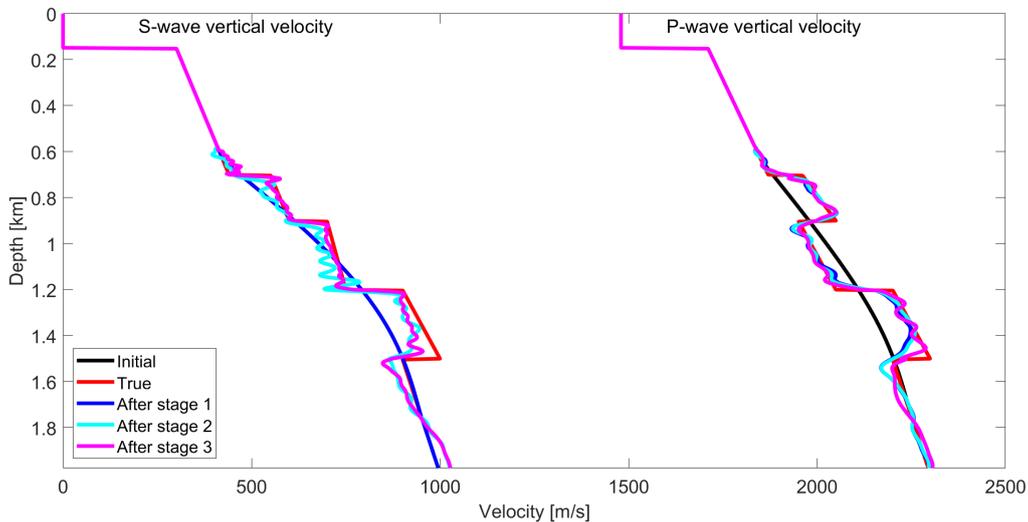


Figure 3.9: Soft seabed synthetic inversion results, in an environment whose anisotropy parameters are known. Inversion performed with the logarithm of the parameterization $(V_{pv}, V_{sv}, \sqrt{1 + 2\eta}, \sqrt{1 + 2\delta}, \rho)$.

To emphasize the need for a careful selection of the parameterization, the same application is performed but using the logarithm of the $(V_{pv}, V_{pn}, V_{ph}, V_{sv}, \rho)$ parameterization. This parameterization stands out in He and Plessix (2016) ([78](#)) through its decoupling of P-wave velocity-related parameters and the similar amplitudes of all its parameters and was in fact the parameterization available in our FWI code. However, this parameterization shows a lack of sensitivity to P-wave vertical velocity from P-P wide-aperture reflections ([figure 3.10](#)). The stage 1 inversion results are shown in [figure 3.11](#) and are compared with the selected parameterization results. As expected, the inversion fails and shows a lack of intermediate scale update, similar to the results obtained with the successful parameterization, but with narrow-aperture data only ([figure 3.11](#)). Indeed, wide-aperture data allows to reconstruct intermediate scale perturbations (Shipp and Singh, 2002 ([172](#)))

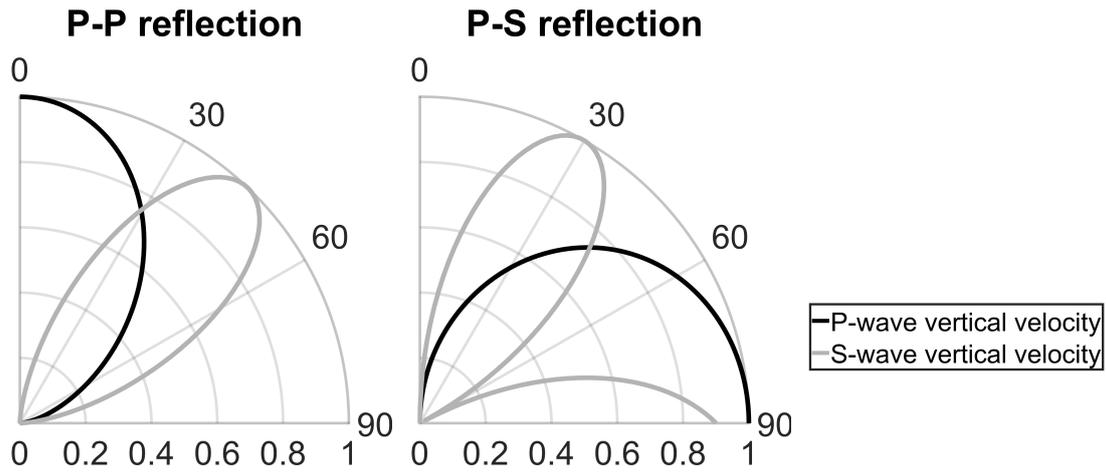


Figure 3.10: Normalized P-P and P-S reflection radiation patterns for P- and SV-wave velocities with the $(V_{pv}, V_{pn}, V_{ph}, V_{sv}, \rho)$ parameterization. The angle corresponds to the incident angle. The medium parameters are those of the bottom layer from table 3.2.

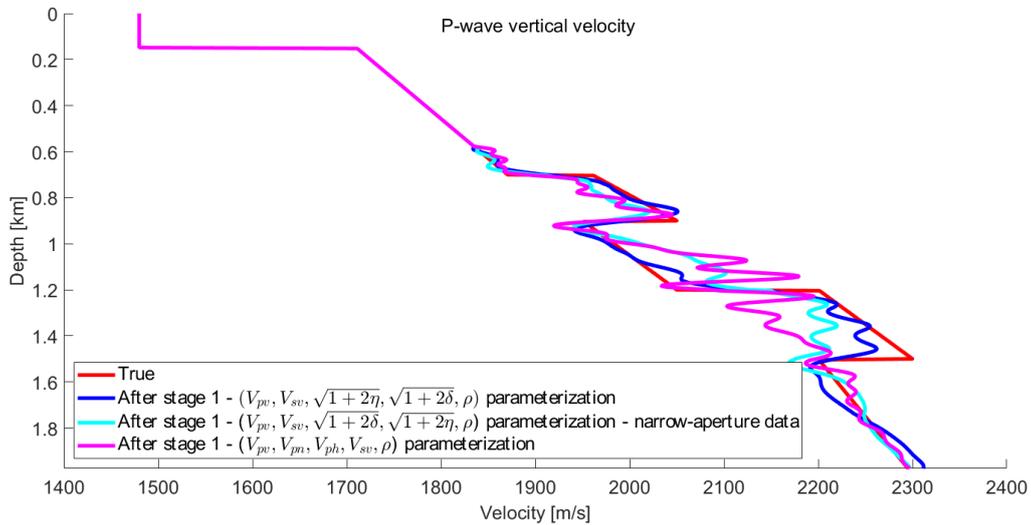


Figure 3.11: Soft seabed synthetic stage 1 inversion results obtained with the successful selected parameterization applied to all data (blue line) and narrow-aperture data only (cyan line), as well as results obtained with a failing parameterization applied to all data (magenta line). The anisotropy parameters are known. The used parameterizations are the logarithm of the parameterizations indicated in the legend.

3.4 Amplitude sensitivity of wide-aperture P-P reflections to elastic vertical transverse isotropic parameters

As explained in [Section 3.2](#), the second inversion stage - to obtain the intermediate scale SV-wave velocity variations - assumes that SV-waves velocities are the main remaining unknowns affecting the amplitude of wide-aperture P-P reflections. When extending the strategy to P- and S-wave vertical velocity inversions with slightly erroneous remaining anisotropy models, one needs to verify that the amplitude sensitivity of these waves to these erroneous remaining anisotropy parameters is minimal.

Following [Sears \(2007\) \(166\)](#), the amplitude sensitivity of P-P wide-aperture reflections to elastic VTI parameters is assessed using an analytical solution of the wave equation for mono-frequency plane waves reflecting on a horizontal surface separating two homogeneous half-spaces (as done by [Kennett, 1983 \(95\)](#)). The computations do not depend on the used frequency. The codes used in this section are provided by Colin Thomson. The computations are performed in 3D but the conclusion are expected to be valid in 2D too. Moreover, the computations are exact and valid for elastic generally-anisotropic media. In particular, it does not use [Thomsen's \(1986\) \(192\)](#) weak anisotropy approximation.

A realistic base model is built, inspired from Alwyn real data initial model (the parameters around top Dornoch horizon are used). The model parameters are provided in [table 3.2](#). Each one of the 10 model parameters is perturbed separately. The corresponding modulus of the P-P reflection coefficient of a down-going plane-wave reflecting on the interface is computed. This modulus ranges from 0 to 1, and indicates how much of the P incident wave goes into the P reflected wave. No extrinsic or intrinsic attenuation occurs because of the plane-property of the wave and the non-attenuating model. Consequently, the amplitude of the incident and reflected wave does not depend on the distance to the interface. Therefore, the reflection coefficient directly provides the amplitude of the reflected wave by multiplication with the amplitude of the incident wave. This computation is performed for different incident plane-wave total velocity angles and different perturbations. The anisotropy parameters perturbations are chosen to range from no anisotropy to the extreme values of the provided initial model for Alwyn real data. The other parameters are chosen to cover a wide range of values. [Figure 3.12](#) provides the modulus of the reflection coefficients for the various tested interfaces and incident angles. The results for the perturbation of γ are not shown because no sensitivity to that parameter is observed. Indeed, no conversion to SH-wave can occur in such setting, while γ is a parameter solely impacting SH-waves.

	P-wave vertical velocity [m/s]	Density [kg/m ³]	SV-wave vertical velocity [m/s]	$\sqrt{1+2\epsilon}$ [-]	$\sqrt{1+2\delta}$ [-]	γ [-]
Top layer	2100	2075	750	1.05	1.03	0
Bottom layer	2500	2200	1000	1.07	1	0

Table 3.2: Base model parameters.

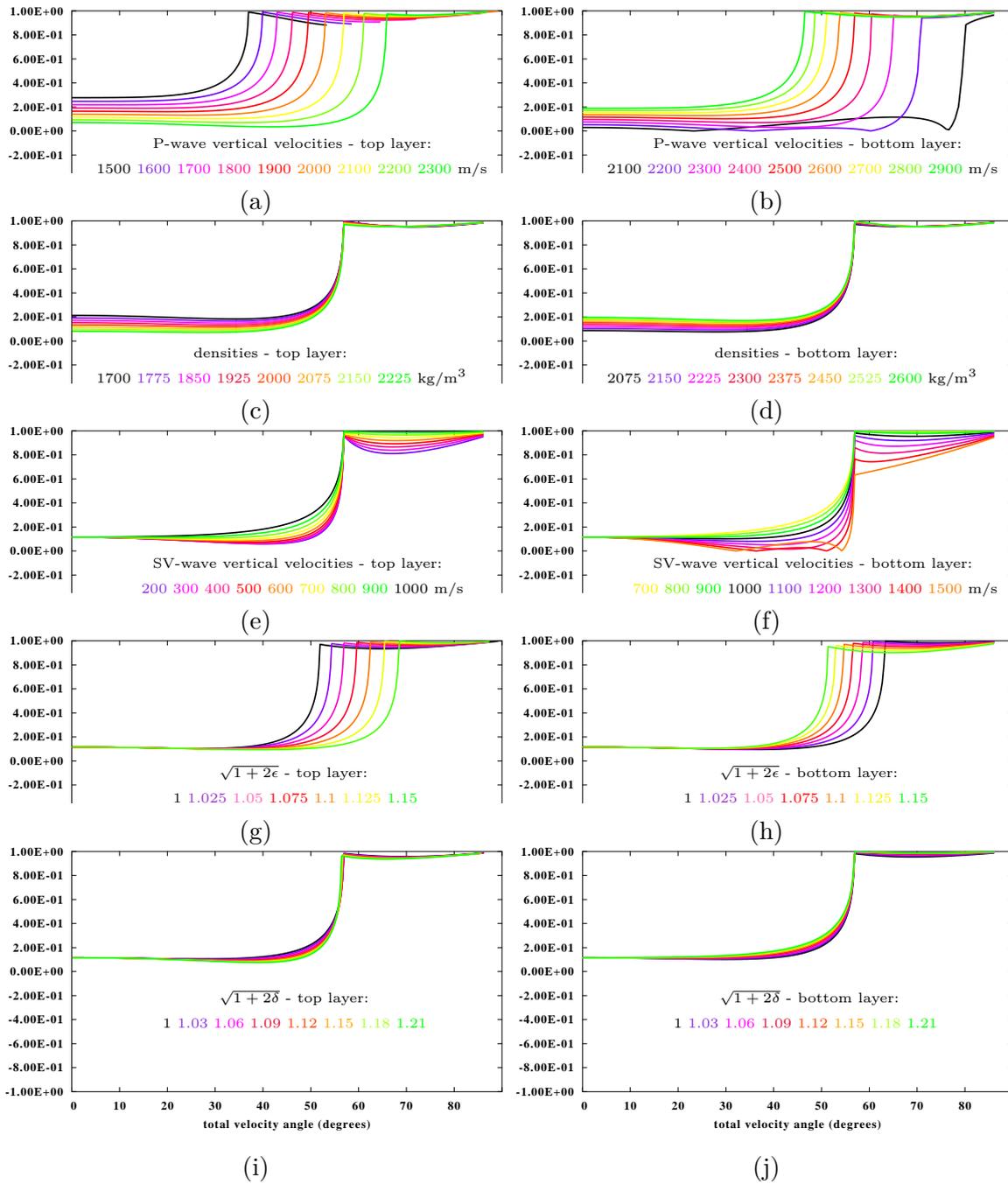


Figure 3.12: P-P reflection coefficients for a downgoing plane-wave hitting a horizontal interface between two half-spaces. Computations performed for various interfaces and incident total velocity angles. The angle is measured from the vertical axis. Each subfigure corresponds to the perturbation of one parameter of the base model of table 3.2.

The behavior of the reflection coefficients around and after the critical angle is analyzed, assuming it corresponds to the approximately selected wide-aperture data. The critical angle is the angle from which the whole incident wave is reflected (i.e., reflection coefficient of 1, the total reflection). In practice, this angle is located at the first discontinuity on the P-P reflection coefficients as a function of the angle of figure 3.12.

The effect of a change of density and SV-wave vertical velocity is as expected (figures 3.12c, 3.12d, 3.12e and 3.12f): their perturbation do not modify the critical angle, and even a small SV-wave velocity perturbation has significant impact on the wide-aperture

reflection coefficient, while even a strong density perturbation has almost no impact on the same coefficient. P-wave vertical velocity also impacts wave amplitude (figures 3.12a and 3.12b), but is supposed to be obtained beforehand during the first inversion stage.

While $\sqrt{1+2\delta}$ does not have much effect (figures 3.12i and 3.12j), even for strong perturbations and for pre-critical events, $\sqrt{1+2\epsilon}$ impacts the reflection coefficient (figures 3.12g and 3.12h). From these variations, the impact of erroneous anisotropy parameters on the performance of the soft seabed strategy is ambiguous. Indeed, on one hand, the post critical amplitude is moderately impacted by $\sqrt{1+2\epsilon}$ while the critical angle is significantly impacted by the same parameter. It indicates that, assuming that the anisotropy parameters allow for a rough P-wave fitting, this might impose small enough errors in the anisotropy parameters model, and consequently a small enough amplitude effect of those errors on the amplitudes. On the other hand, the effect on the second stage inversion results of the high amplitude variation near critical angle (given an incident angle) due to the modification of critical angle is difficult to assess. Indeed, which incident angle fits a certain arrival at the end of the first inversion stage is unsure.

3.5 Synthetic application of the soft seabed strategy with slightly erroneous anisotropy parameters

To appraise in a more complete way the effect of slightly erroneous anisotropy parameters on the strategy, the previous inversion of Section 3.3 is performed with perturbed anisotropy parameters instead. Figure 3.13 provides the used anisotropy parameters, compared with the true ones. To build the perturbed models, the $\sqrt{1+2\delta}$ and $\sqrt{1+2\eta}$ values have been randomly perturbed by 5 %, the interface between layers 3 and 4 has been misplaced 20 m too deep, and the interface between layers 5 and 6 has been misplaced 60 m too shallow. Those perturbations try to reproduce the errors of Alwyn anisotropy parameters models, by mimicking the strength of the generated P-wave wide-aperture/headwave travel-time errors. The perturbation of layer six seems to be overestimated. Finally, the anisotropy parameters models have been vertically smoothed.

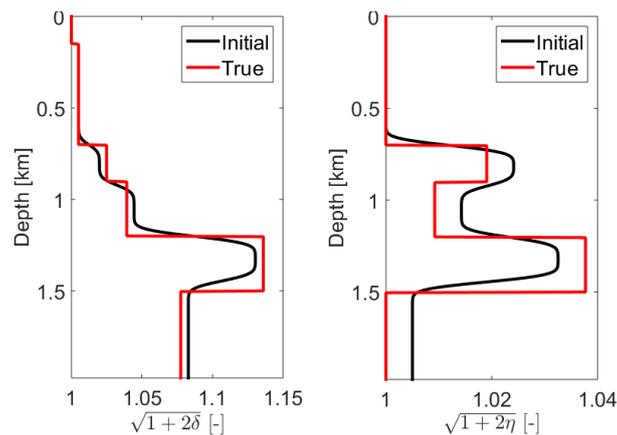


Figure 3.13: True and perturbed used anisotropy parameters.

Figure 3.14 provides the inversion results for the three inversion stages. The erroneous anisotropy parameters allow to reasonably recover P-wave vertical velocities from wide- and narrow-aperture P-wave data, confirming that the tested erroneous parameters fall within the selected scope of the study. The inversion results seem to indicate that the soft seabed strategy is able to manage reasonable anisotropy errors and provide perturbed, but

reasonable results. The differences with the non-perturbed application might be explained by the range of uncertainty of the inversion scheme. However, in this example, it could also be interpreted as a compensation process of the inversion. Indeed, the velocity inversion errors are (at least qualitatively) correlated with the anisotropy parameters error: one can notice that an underestimation of velocities from anisotropy parameters corresponds to an overestimation of inverted vertical velocities, and vice versa. On figure 3.14, this compensation is not visible for the S-wave vertical velocities of layer 5, unless one compares with the inverted models with correct anisotropy parameters, where the velocities are underestimated (figure 3.9).

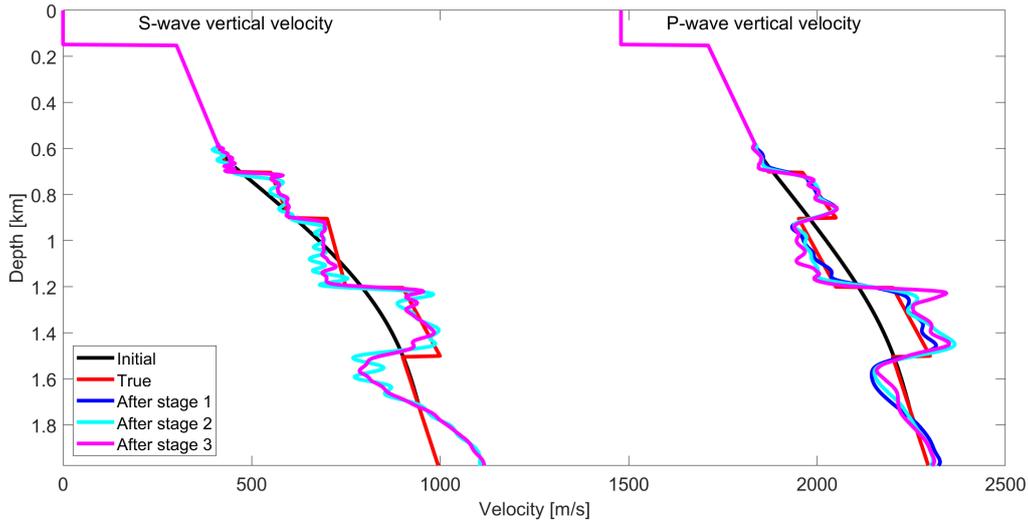


Figure 3.14: Soft seabed synthetic inversion results, in an environment whose anisotropy parameters are slightly erroneous. Inversion performed with the logarithm of the $(V_{pv}, V_{sv}, \sqrt{1 + 2\eta}, \sqrt{1 + 2\delta}, \rho)$ parameterization.

3.6 A variant to the soft seabed strategy

Aside from VTI, other physical phenomena can show an effect on the amplitudes of P-P wide-aperture reflection, and could hamper the results of the second stage of the soft seabed strategy, for example:

- Intrinsic attenuation notably affects the amplitude of the waves during propagation and reflections, as illustrated by figure 1 of Belahi et al. (2015) (17), provided in figure 3.15. This attenuation is due to internal grain frictions that result in seismic energy being dissipated by heat during wave propagation. Intrinsic attenuation acts on amplitudes, and high-frequencies suffer more from attenuation than low-frequencies. Intrinsic attenuation also acts on phase.
- Poor geometrical spreading corrections for 2D tests also affect wave amplitudes (see for instance Auer et al., 2013 (14)).

In real data applications, all of these effects are expected and usually not modeled for practical reasons. These applications would assume that: (1) because only low frequencies are used, the attenuation has a small effect; (2) because P-waves only are used, usual approximations for geometrical spreading correction have a small impact.

Additionally, a correct P-P post critical reflection phase and travel time fit is also expected to be critical for the success of the second inversion stage. Both attenuation

and geometrical spreading phenomena (as well as incorrect shallow sub-surface, sources or acquisition geometry for instance), can lead to inadequate P-P post critical reflection phase and travel time fit. This inadequate fit could hamper the results of the second stage of the soft seabed strategy.

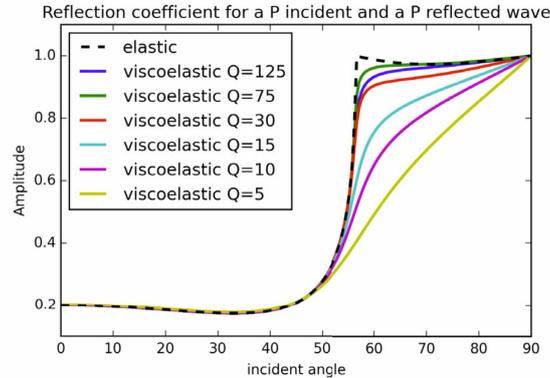


Figure 3.15: “Reflection coefficients as a function of incidence angle with varying values of Q_p in the layer above the reflector. Other parameters are kept constant ($V_{p1} = 2500$ m/s, $V_{s1} = 1200$ m/s, $Q_{s1} = 20$, $V_{p2} = 3000$ m/s, $Q_{p2} = 50$, $V_{s2} = 1400$ m/s, $Q_{s2} = 40$ where 1 denotes the medium above and 2 the medium below the reflector) [...]”. Figure and title from Belahi et al., 2015 (17).

I notice that with today’s data low frequency content, the second stage can be replaced by the application of the increasing frequency multiscale strategy of Bunks et al. (1995) (35) during the third stage (figure 3.16): the increasing frequency multiscale strategy builds the intermediate scale SV-wave velocities features before the small scale features, through a gradual increase of used data frequencies. This explain why this variant strategy was successfully applied by Brossier et al. (2009) (32) and Brossier et al. (2010) (33) on synthetic data. This variant of the soft seabed strategy could be a solution for a failing of the second stage due to the above mentioned reasons.

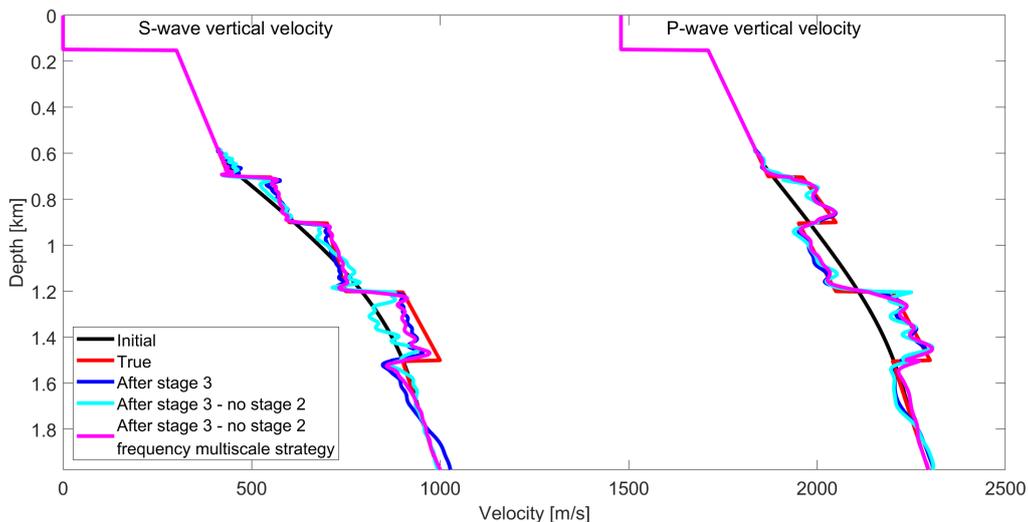


Figure 3.16: Comparison of soft seabed synthetic inversion results with and without the inversion of intermediate scale SV-wave velocities update using the amplitude of P-P wide-aperture reflections (i.e., stage 2), and with the effect of the increasing frequency multiscale strategy (Bunks et al., 1995 (35)) as a substitute for stage 2.

However, this variant strategy seems to be more sensitive to erroneous anisotropy

parameters, as illustrated in figure 3.17. This figure shows the results of the variant strategy, compared with the results of the original soft seabed strategy when the second stage is skipped, using the erroneous anisotropy parameters instead. The results still are comparable for the layer for which the anisotropy is the lowest (i.e., layer 4). The decreased weight of amplitudes in this inversion leads to an increased weight of travel-times. If those travel-times show more sensitivity to erroneous anisotropy parameters, this could explain the increased sensitivity of this variant strategy to such inaccurate parameters.

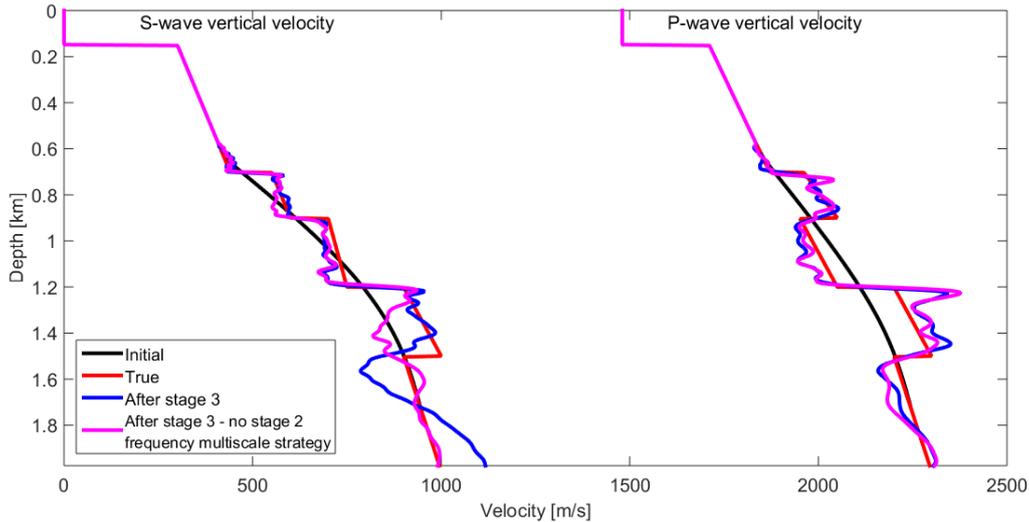


Figure 3.17: Comparison of soft seabed synthetic inversion results with and without the inversion of intermediate scale SV-wave velocities update using the amplitude of P-P wide-aperture reflections (i.e., stage 2), and with the effect of the increasing frequency multiscale strategy (Bunks et al., 1995 (35)) as a substitute for stage 2.

3.7 Discussion and conclusions

In this chapter, I recalled the need and the reasoning behind the soft seabed strategy, which inverts for both P- and SV-wave velocities in soft seabed environments. I especially emphasized the isotropic approximation within which the strategy was designed.

In order to apply the soft seabed strategy in VTI media to update P- and S-wave vertical velocities, the $(V_{pv}, V_{sv}, \sqrt{1 + 2\eta}, \sqrt{1 + 2\delta}, \rho)$ FWI parameterization was selected, based on analytically computed reflection radiation patterns. This parameterization choice was confirmed with a synthetic inversion test. The importance of the parameterization was also emphasized with an inversion test that used a wrong parameterization, selected with the same procedure.

After this adaptation, on one hand I showed with a synthetic application that the strategy could still remain successful when slightly erroneous VTI parameters are used, those parameters not being updated. Moreover, the inversion errors could be used to qualitatively estimate the anisotropy parameters errors, with an overestimation of velocities from anisotropy parameters compensated by a commensurate underestimation of vertical velocities. On the other hand, to overcome the anticipated failure of the second stage of the strategy in attenuating media for instance, I noticed that this second stage could be substituted by the introduction of Bunks et al.'s (1995) (35) increasing frequency multiscale strategy in the third stage of the soft seabed strategy. Indeed, current data

low frequency content seems to allow for the application and success of such combination. This substitution strategy, however, seems to be more sensitive to erroneous anisotropy parameters. In real data applications, the choice of the strategy between the soft seabed strategy and its variant would thus depend on the side on which the scale tilts: wronger P-P wide-aperture fit or wronger anisotropy parameters?

That being said, in a context where poor SV-wave velocity information is contained within data, it would be a waste not to use the information contained in amplitudes (and thus to use the variant of the soft seabed strategy), especially given their highlighted low resolution power and possible lower sensitivity to erroneous anisotropy parameters. This supports the need for attenuating modeling and correct 3D simulations for instance. The development of amplitude-focused cost functions might also be a solution to consider. For instance, in the fictional case where the phase of P-P wide-aperture reflections is expected to be incorrectly modeled, in opposition with the amplitude and travel-time, the use of a cost function such as the envelope cost function that minimizes the role of the phase might solve the problem (see Wu et al., 2014 (232) for instance).

In addition to its need for multicomponent ocean bottom data, the soft seabed strategy and its variant also rely on (1) wide- and narrow-aperture data availability, and (2) a P- and S-wave separation between the vertical or pressure components and the horizontal component. This second need is met for roughly horizontally layered media. Another strategy probably needs to be designed for applications that do not meet those two criteria. Similarly, some adjustments for 3D applications are needed, and other strategies should be used for poorer initial models (including vertical velocity models) or poor low-frequency data content.

This study provided a physical insight on different strategies. Whatever the strategy, a quantification of the amount and type of errors in the updated models in function of the amount and type of errors in the anisotropy model (and vice versa) could be a very welcome contribution, together with extensions to other kinds of anisotropy (e.g., tilted transverse isotropy or azimuthal anisotropy). Regarding the type of errors in the anisotropy model, one could for instance study the update of horizontal velocities with erroneous remaining anisotropy parameters that allow for correct horizontal velocities, or include the effect of wrong seabed values or at least updated seabed values. More significantly, this chapter illustrating the need for good anisotropy models, determining a suitable FWI inversion strategy for all velocity parameters (including anisotropy parameters) should be the next methodological focus. On this specific point, a side learning from the differences of evolution with angle between P- and SV-wave velocities is that, SV-waves having a different sensitivity to anisotropy (figure 3.4), the use of SV-waves for anisotropy parameters inversions should be considered.

This work highlights an additional point: the power of the increasing frequency multiscale strategy, which allows the introduced variant of the soft seabed strategy to provide better results than the original strategy when the anisotropy parameters are perfectly known (figure 3.16). It is therefore recommended that this frequency multiscale strategy be used at any stage of any inversion strategy, as is usually done for real data applications (e.g., Sears et al. (2010) (165)).

Chapter 4

Alwyn real data: data and full waveform inversion inputs analysis

Chapter's objectives and highlights

- Show that Alwyn real data fit the criteria for the application of the inversion strategy of [Chapter 3](#).
 - Analyze the available full waveform inversion inputs to design their preparation workflow, and assess the achievable inversion results. This analysis is done with respect to the acquisition objectives and the geology. The inputs hint promising inversion results.
 - Present the available quality control tools to check the inversion results, which transform Alwyn real data into an interesting dataset for full waveform inversion methodological studies (i.e., inversion strategy).
-

Contents

4.1 Introduction	75
4.2 Geology of Alwyn North field	76
4.3 Ocean bottom node data description and analysis	80
4.3.1 Objectives of the acquisition	80
4.3.2 Description of the acquisition	80
4.3.3 Seismic data processing	84
4.3.4 Data analysis	85
4.4 Discussion and conclusions	101

4.1 Introduction

Alwyn real data were acquired in 2014 in the Northern North Sea basin, above the British Alwyn North hydrocarbon field (figure [4.1](#)). These seismic data were acquired by Seabed Geosolutions ([SBGS](#)), and processed by CGG, on behalf of TOTAL. The processing workflow included travel-time tomography, acoustic Full Waveform Inversion ([FWI](#)), and P-P and P-S migrations.

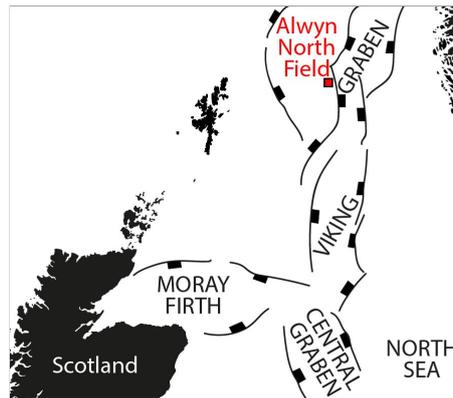


Figure 4.1: Alwyn North field location. Grabens' location from Harker et al. (2003) (77).

Alwyn North field is a mature field discovered in 1975, brought on stream in 1987, and still producing. Alwyn North field has a long history, summarized in [Appendix G](#). This summary also shows the role of seismic imaging in the life of a hydrocarbon field, and hints at how this mature field can warrant a new seismic acquisition and further production efforts. As a consequence from the field history, Alwyn real data is accompanied by a large quantity of various pieces of information: 3D models of elastic parameters of the sub-surface, well data, a vintage time- and a vintage depth-migrated volume, seismic horizons, and numerous documents from [SBGS](#), [CGG](#) and [TOTAL](#) describing their analysis and processing of these data. These pieces of information turn the Alwyn real dataset into a good candidate for [FWI](#) methodological (i.e., inversion strategy) studies. Indeed, they allow for a better understanding of the data, of the sub-surface, and of the geological targets.

The aim of this chapter is to contribute to the three critical facets of [FWI](#) stated in the conclusion of [Chapter 1](#), i.e., to:

- select/design a suitable [FWI](#) inversion strategy.
- constrain (pre-process) and Quality Control ([QC](#)) the [FWI](#) inputs.
- identify [QC](#) tools for [FWI](#) outputs.

The (petroleum) geology of the field and the targets of the acquisition are detailed first, followed by a description and analysis of the seismic data. The other pieces of information (models, well data, migrated images and horizons) are introduced and analyzed throughout this chapter, whenever needed. Finally, I summarize the results of the analysis in terms of the three critical facets listed above, for the 2D [FWI](#) application of [chapter Chapter 5](#).

4.2 Geology of Alwyn North field

The structure of the Northern North sea basin is principally the consequence of an Upper Jurassic/Lower Cretaceous rifting followed by sediment deposition (Rathey and Hayward, 1993 (150)). This rifting led to the current Viking Graben, as represented in [figure 4.1](#). Near Alwyn North field, the geometry of the Viking Graben suggests a relative invariance of the sub-surface geometry along a \sim N-S axis. I call this geometry a “ \sim N-S 2.5D geometry”.

Alwyn North field is constituted by three reservoirs located on top of each other, within a tilted block of the western side of the Viking graben: Brent (oil), Statfjord (gas)

and Triassic (gas and condensate). Figure 4.2 provides the approximate lateral extension of the reservoirs, whose elongated direction suggests a \sim N-S 2.5D geometry. The location of the source and the Ocean Bottom Node (OBN) receiver lines used in this study is also indicated.

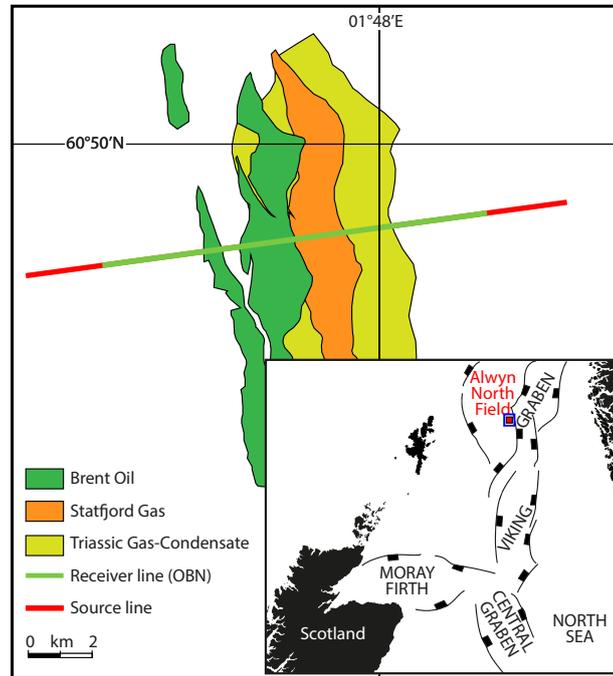


Figure 4.2: Approximate lateral extension of Alwyn North field reservoirs, and acquisition used in this study. Petroleum data and location of the grabens from Harker et al. (2003) (77). OBN=Ocean Bottom Node. The blue rectangle on the regional map indicates the extent of the reservoirs map.

The vintage time- and depth-migrated volumes and the seismic horizons (both obtained before Alwyn real data acquisition) provide more details about the sub-surface. Figure 4.3 displays the depth migrated image below the source line used in this study, together with some clearly identifiable horizons and some interpreted faults. This figure shows the expected tilted block separated by tilted faults in the older units (such as Brent, Dunlin, Statfjord or Lomvi), covered by sub-horizontal younger units (such as Skade, Larke or Dornoch). Both parts are separated by the Base Cretaceous Unconformity (Base Cretaceous Unconformity (BCU)), and can be divided into several seismic units made of sandstones, siltstones, shalestones, mudstones, and/or occasionally claystones (British Geological Survey website), and separated by strong reflectors. The tilted and faulted block are associated with dipping layers. Because of the type of rocks and the orientation of layers, Tilted Transverse Isotropy (TTI, also known as tilted transversely isotropy) can be expected, especially Vertical Transverse Isotropy (VTI) above BCU. A TTI medium is a VTI medium where the axis of symmetry is tilted. Non economic gas and oil were observed inside Dornoch formation, and between Dornoch and Lark formations (gas).

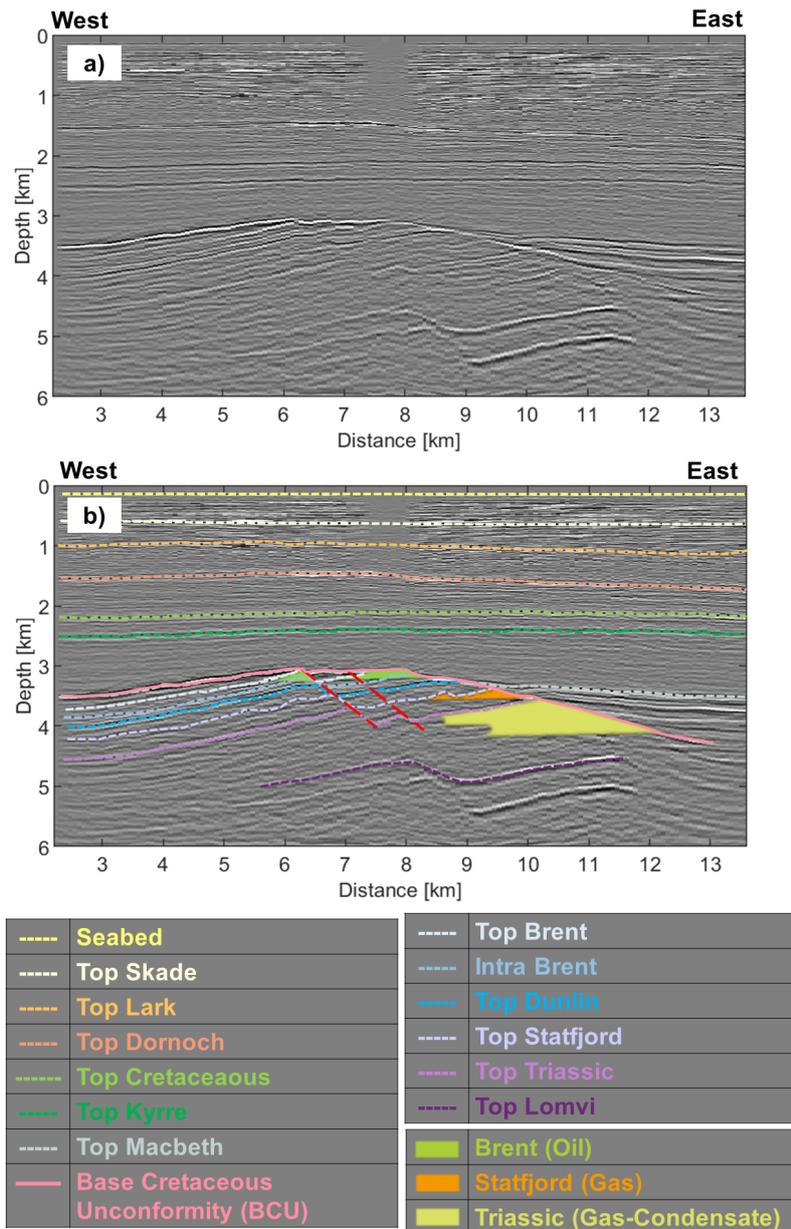


Figure 4.3: Migrated image below the source line used in this study, without (a) and with (b) interpreted horizons and faults (long dashed red lines), and approximate locations of reservoirs. The blurring in the middle of the shallow part comes from a lack of data due to a production platform. No depth exaggeration. Horizontal distance is measured along the source line, from the western most source. The reservoir locations are from Harker et al. (2003) (77).

Different versions of each horizon are available, all of them agree with the migrated images. Figure 4.4 displays a depth version of each horizon and confirms the \sim N-S 2.5D geometry. The seabed horizon indicates a flat and shallow seabed located around 150 m.

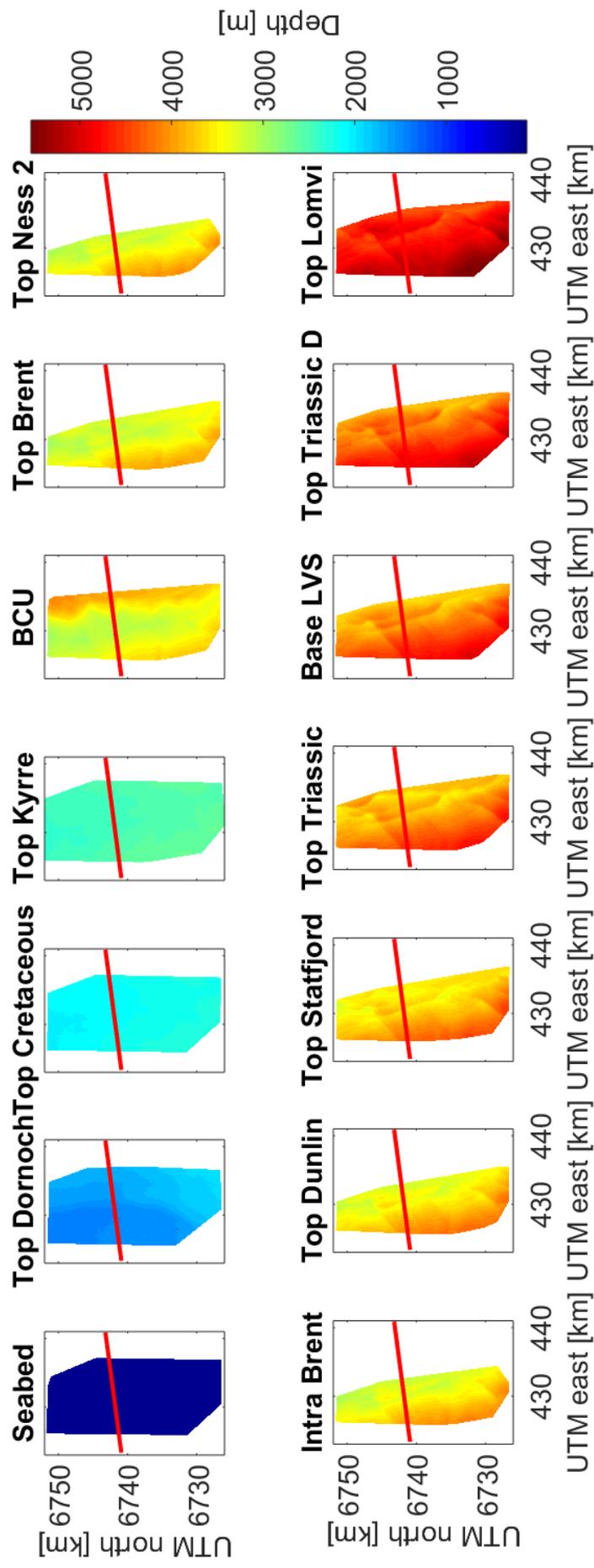


Figure 4.4: Available seismic depth horizons, from top to bottom. **BCU** = Base Cretaceous Unconformity. The location of the source line used in this study is indicated in red.

As shown on figure 4.3, all reservoirs sit within a tilted block of the Viking graben, between 3 and 5 kilometers depth. Each reservoir is split into several pools by tilted block faults. The division of Triassic reservoir into pools is also due to the structure of the formation: interbedded sequence of thin sandstones (2 to 10-meter-thick), and sealing stratigraphic intervals. The Recovery Factors (RFs, which are the proportion of the hydrocarbon that can be recovered) before Alwyn real data acquisition were: 49% for Brent, 60% for Statfjord and 15% for the Triassic reservoir. While RFs for Brent and Statfjord reservoirs are normal, the RF of Triassic reservoir is low. Indeed, ~ 5 to 80% of the hydrocarbon is usually recovered, and RFs for gas tend to be higher.

4.3 Ocean bottom node data description and analysis

4.3.1 Objectives of the acquisition

As a consequence from the low (Triassic reservoir) RF, the reasons for this acquisition include:

- **Main objective:** Better delineation of the faults and thin sand layers within and bounding reservoirs, through an increase of signal to noise ratio, event continuity and resolution of the migrated seismic volumes. The pools of the reservoirs should then be better identifiable, and the RFs should increase, especially for the Triassic reservoir.
- **Research related reasons:** Improving the understanding and the use, and assess the usefulness of: (1) multi-component and ocean bottom seismic data, and (2) elastic wave propagation; when addressing the above main objective. Indeed, I remind that ocean bottom seismic and multi-component data notably allow for S-wave velocity model building in soft seabed environment, through P-S reflections and wide-aperture data recording (Sears et al., 2008 (164)). Among other things, S-wave imaging can provide a better imaging below gas (e.g., Granli et al., 1999 (68)) and a better resolution (e.g., Borisov and Singh, 2015 (29)). I refer to the introductions of Chapter 2 and Chapter 3 for more examples of the use of ocean bottom and multi-component seismic data, and elastic wave propagation respectively.

4.3.2 Description of the acquisition

The acquisition area is above Alwyn North field. The acquisition was performed from December 2013 to June 2014. Sources are airgun Bolt arrays, with a volume per source of 4990 cu.in and an operating pressure of 2000 PSI. Receivers are TrilobitsTM. A TrilobitTM is a type of multi-component Ocean Bottom Node (OBN). It records pressure above the seabed using a hydrophone (P), and the three components of the velocity vector at the seabed using three geophones. The receivers were deployed using remotely operated vehicles, ensuring their correct location and orientation (see figure 4.5). In the design of such receivers, special attention was paid to coupling, vector fidelity and noise reduction. Alwyn's shallow water and flat seabed environment is ideal for ocean bottom node acquisitions.



Figure 4.5: Aspect of TrilobitsTM receivers (from SBGS trilobit specification sheet 2014).

In this particular type of multi-component OBN, the configuration of the geophones is a Galperin configuration. In the Galperin configuration (also known as symmetric triaxial, homogeneous triaxial or 54° geometry), three single-component geophones are mounted at an angle of 35.3° to the horizontal and at an angle of 120° relative to each other (figure 4.6a). It means that the three sensors are positioned orthogonally. Obtained data must be rotated into an Earth referenced cardinal X, Y, Z coordinate system for most analyses. Because of uncertainty in tilt measurements, residual tilt computations might still be needed.

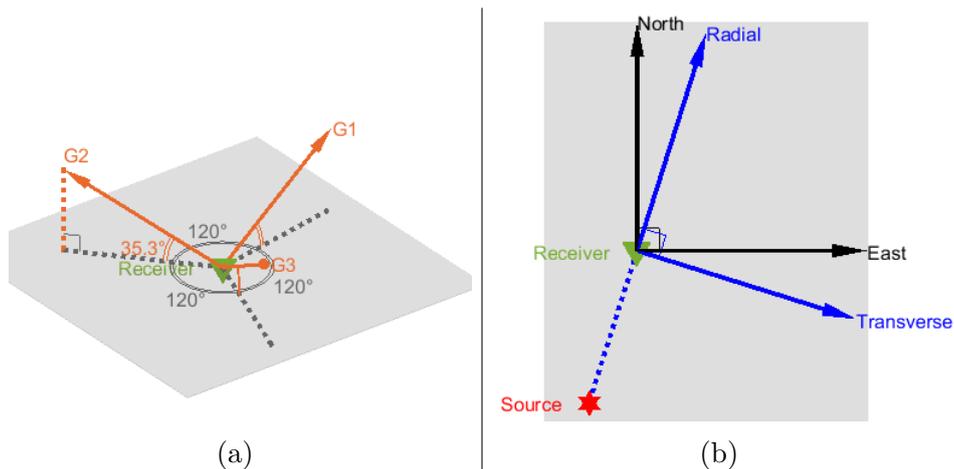


Figure 4.6: a): Galperin configuration. b): Definition of radial and transverse components. G1, G2, G3 are the three geophone components (orange arrows). Receiver is on the seabed (green triangle) and source (red star) is on the surface. Blue arrows indicate the radial and transverse components in the horizontal plane.

The acquisition geometry is described in figure 4.7 and in table 4.1. The total source coverage is not the same as the source coverage per receiver. Each receiver has its own source coverage. The source coverage for one receiver typically fits into a $\sim 12 \times 20$ km rectangle (\sim W-E orientation), but the variability of this coverage is high. Unlike what was planned before the acquisition, source lines are strongly deviated around wells. A few sources are missing for certain receivers, and one receiver is missing on the South-West part of the survey. More details about the depth of the receivers is provided in figure 4.8. This figure again shows that the seabed is almost flat, which leads to a low variability of the receiver depths (from 122 m to 138 m with a mean of 129 m). Time acquisition parameters are available in table 4.2.

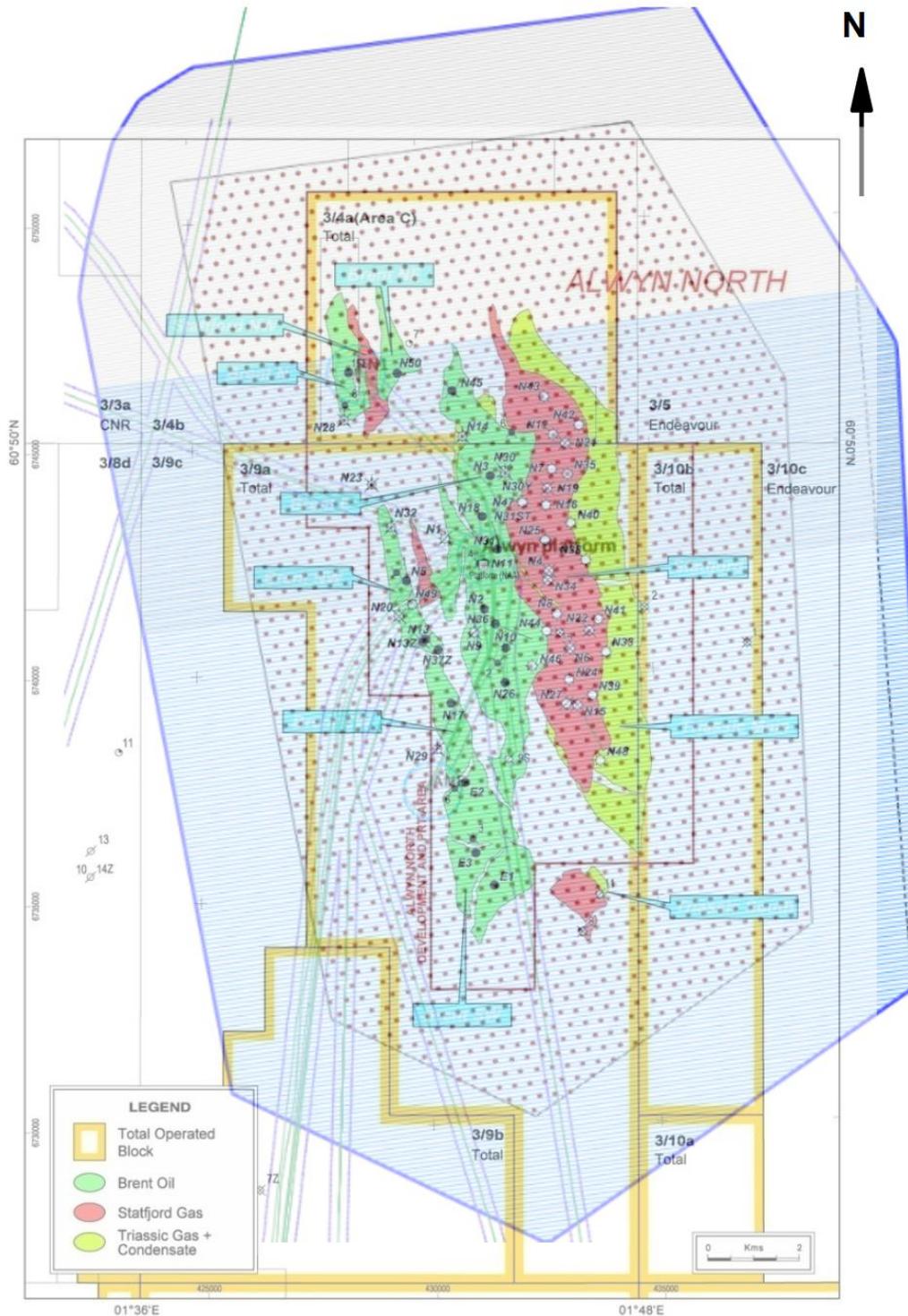


Figure 4.7: Location of the acquisition area and planned acquisition geometry. Red dots are receivers. Note the two different colors used for the source lines (blue and grey colors inside the blue polygon) corresponding to a change of source line separation from 37.5 m to 50 m. Pipelines are annotated in green, with a 200 m safety limit annotated in purple. The black and white circles represent well heads. Source: modified from CGG internal communications (160919 Alwyn processing report draft).

Source area size	$\sim 404 \text{ km}^2$
Source line orientation	82.56°N
Distance between consecutive source lines	$\sim 37.5 \text{ m}$ or $\sim 50 \text{ m}$
Distance between consecutive sources inside a source line	$\sim 37.5 \text{ m}$
Source depth	$\sim 9 \text{ m}$
Receiver area size	$\sim 228 \text{ km}^2$
Distance between consecutive receiver lines	$\sim 300 \text{ m}$
Distance between consecutive receivers inside a receiver line	$\sim 346 \text{ m}$

Table 4.1: Acquisition Geometry.

Sampling interval	2 ms
Recording length	7998 ms
Number of time steps	4000

Table 4.2: Time acquisition parameters.

Data are available at various processing stages (see [subsection 4.3.3](#)). Except for raw data, only a swath of 11 receiver lines and all corresponding sources are available (blue rectangle on figure 4.8). This study focuses on 2D FWI performed on the line indicated in figure 4.8. This line was chosen such that:

- It takes advantage of the 2.5D geometry.
- Sources and receivers are as dense as possible on this line (i.e., follow acquisition lines).
- It is close enough to the wells to benefit from well logs, but not too close to the platform to avoid platform-related noise and the deviated source lines.
- It is contained in the initial models and in the swath of data.
- It is above the geological targets.

The chosen line, hereafter called “the profile”, is the straight line that best fits the chosen receiver line (35 receivers). The chosen source line is the closest source line (456 sources). The maximum offset is $\sim 14 \text{ km}$. The sources and the receivers are in close proximity to the profile: the maximum distance for the receivers (resp. sources) is $\sim 5 \text{ m}$ (resp. $\sim 27 \text{ m}$).

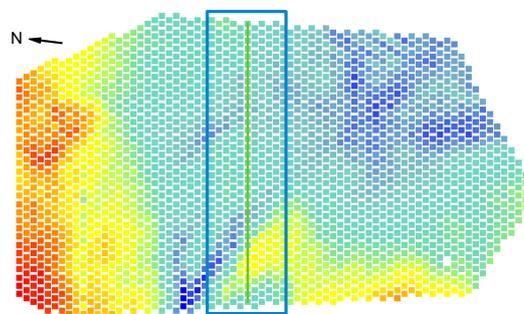


Figure 4.8: Depth of the receivers along the survey. The receivers of the available swath are the receivers inside the blue rectangle. The green line correspond to the profile for this study. The color indicates the depth of the receivers from 122 m (blue) to 138 m (red). Source: modified from a figure of CGG (160919 Alwyn processing report draft).

4.3.3 Seismic data processing

Onshore seismic data processing steps were performed between August 2014 and April 2016. Refer to figure 4.9 and below comments for details about applied processing step.

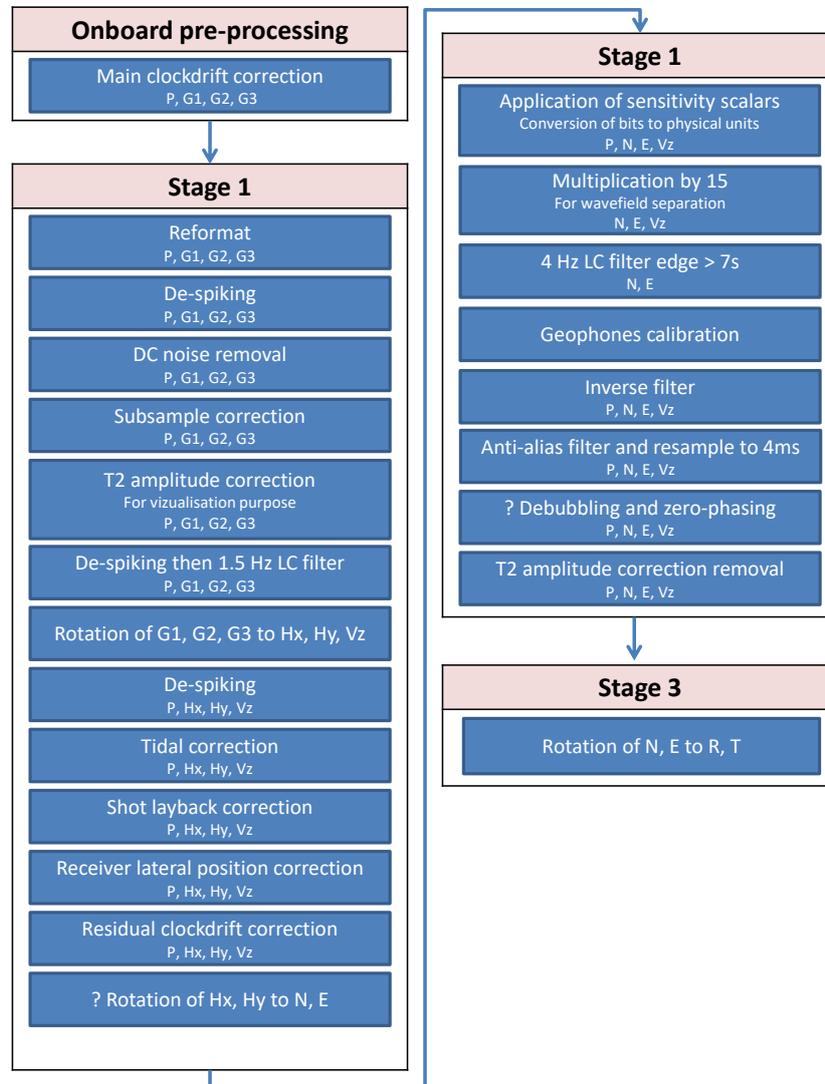


Figure 4.9: CGG processing workflow applied on the available seismic data. The components to which each processing step is applied are specified: G1, G2, G3 are the three geophone components; P is the hydrophone component; Hx, Hy and Vz are respectively two orthogonal horizontal and the vertical downward velocity components; N, E, R and T are respectively the North, East, Radial and Transverse horizontal velocity components (as defined in figure 4.6b). DC = Direct Current. LC = Low-Cut. The question mark indicates that the processing step may not have been performed, or at least not at this place in the workflow.

Three de-spiking passes were performed. The first pass consisted of zeroing the anomalous samples (using threshold), followed by an interpolation using good surrounding traces. The second pass aimed to remove residual spikes not affected by the first pass. It consisted of the application of a threshold in the frequency-offset domain. The last pass used the same process as the second pass, and aimed to remove spikes that appeared with the rotation of the data.

DC (Direct Current) noise (an almost constant noise) was removed using a 1.5 Hz low-cut filter, which did not remove the useful signal, even at low frequencies.

The effects of the recording system and of the analogue-to-digital conversion were removed using a deterministic inverse filter derived from the instrument response provided by SBGS. The inverse filter increased artifacts in the data, as will be discussed later in this chapter. To decrease this effect, a 4 Hz low-cut filter was applied to the North and East components, before calibration and for times higher than 7 s.

Two de-signature processing steps were applied: de-bubbling and zero-phasing. They have two objectives, respectively: (1) to remove the effect of bubble reverberations that follow the main energy peak (low frequency effect associated with the release of energy by the airgun); (2) to convert the sequence primary-source ghost into a symmetric (or zero-phase) signal, easier to interpret. The de-bubbling and zero-phasing operators were designed from a far-field signature, obtained by stacking near-offset data recording direct arrivals.

4.3.4 Data analysis

I discuss here different aspects of the provided velocity models and well data, as well as the seismic data, such as the noise and frequency content.

Data A being the dataset used in this study, only the chosen receiver and source lines of this dataset are analyzed. All four components of a typical receiver gather from dataset A and the corresponding amplitude frequency spectra are displayed in figure 4.10. Figures 4.11, 4.12, 4.13 and 4.14 show corresponding low frequency panels. Only the usual FWI frequency range is used for the panels.

Noise content: Although rig reflections, seismic interferences from neighboring acquisitions and bad weather related noises were expected, only wraparound from previous sources (red arrows on figure 4.10; mainly guided waves and Scholte waves), inverse filter noise (orange arrows), short-offset noise (blue arrow), and high amplitude low frequency spikes on velocity components (red arrows on figure 4.13) were observed.

Frequency content: The highest frequency is around 115 Hz, which fits with the anti-aliasing filter applied to the data (100 Hz high cut & 120 dB/oct roll off). The small notch around 80 Hz in the frequency spectra fits with the location of the expected source ghost notch at zero offset (using formula from Landrø and Amundsen, 2010 (105)). It is attenuated by the fact that several traces with various offsets were used to produce these frequency spectra. Figures 4.11, 4.12, 4.13 and 4.14 show that information below 3 Hz is available.

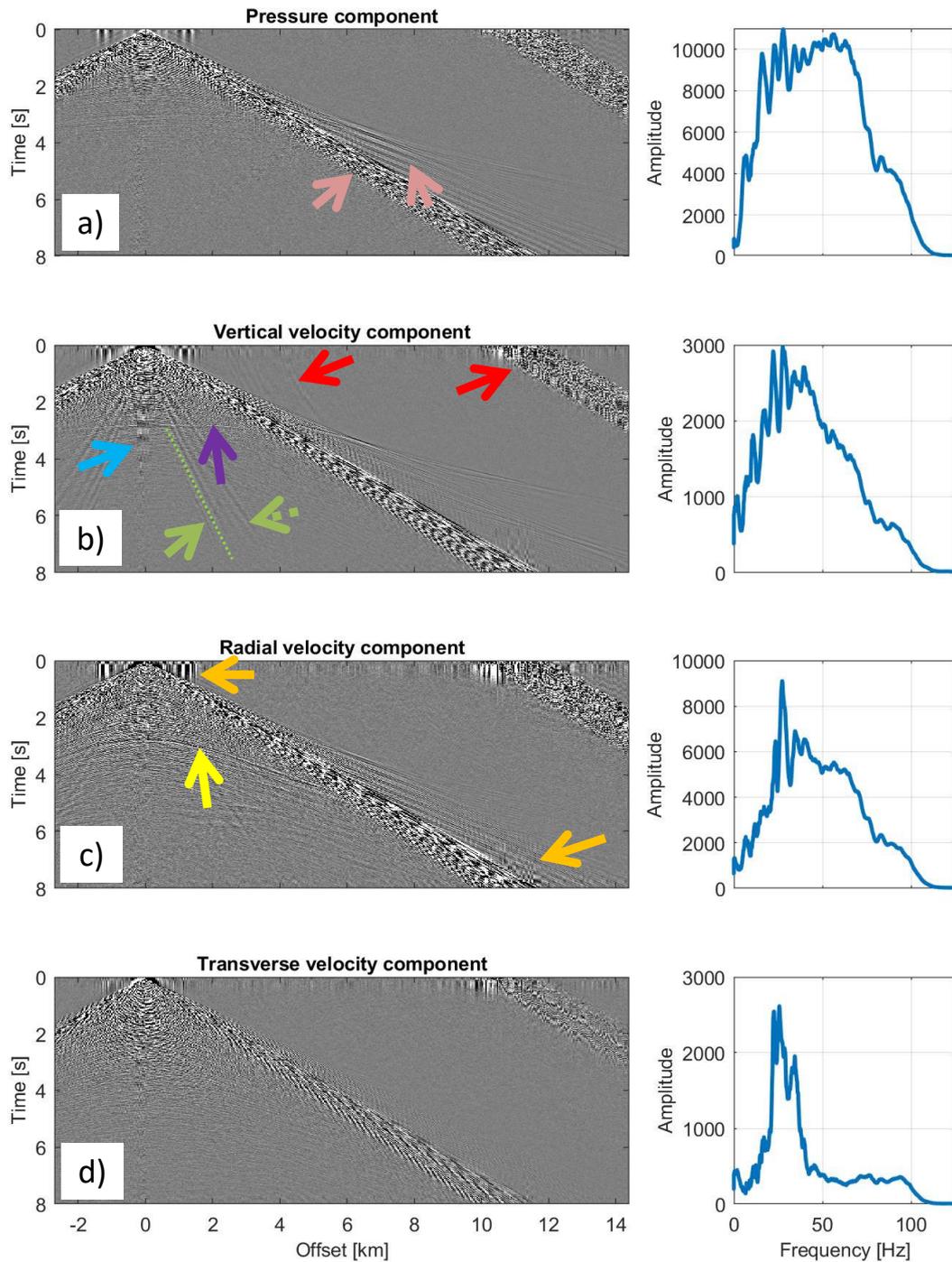


Figure 4.10: Western most receiver gather of the profile and corresponding frequency spectra. Pink arrows: waterlayer multiples and guided waves; Red arrows: wrap-around noise; Blue arrow: short-offset noise; Green arrows: Scholte waves; Purple arrow: P-P reflection; Yellow arrow: P-S reflection; Orange arrows: inverse filter noise. Dashed green line: picked Scholte wave velocity. The same color scale is used for all velocity components. Extreme values of color scale are divided by 1000 to improve global visibility of events.

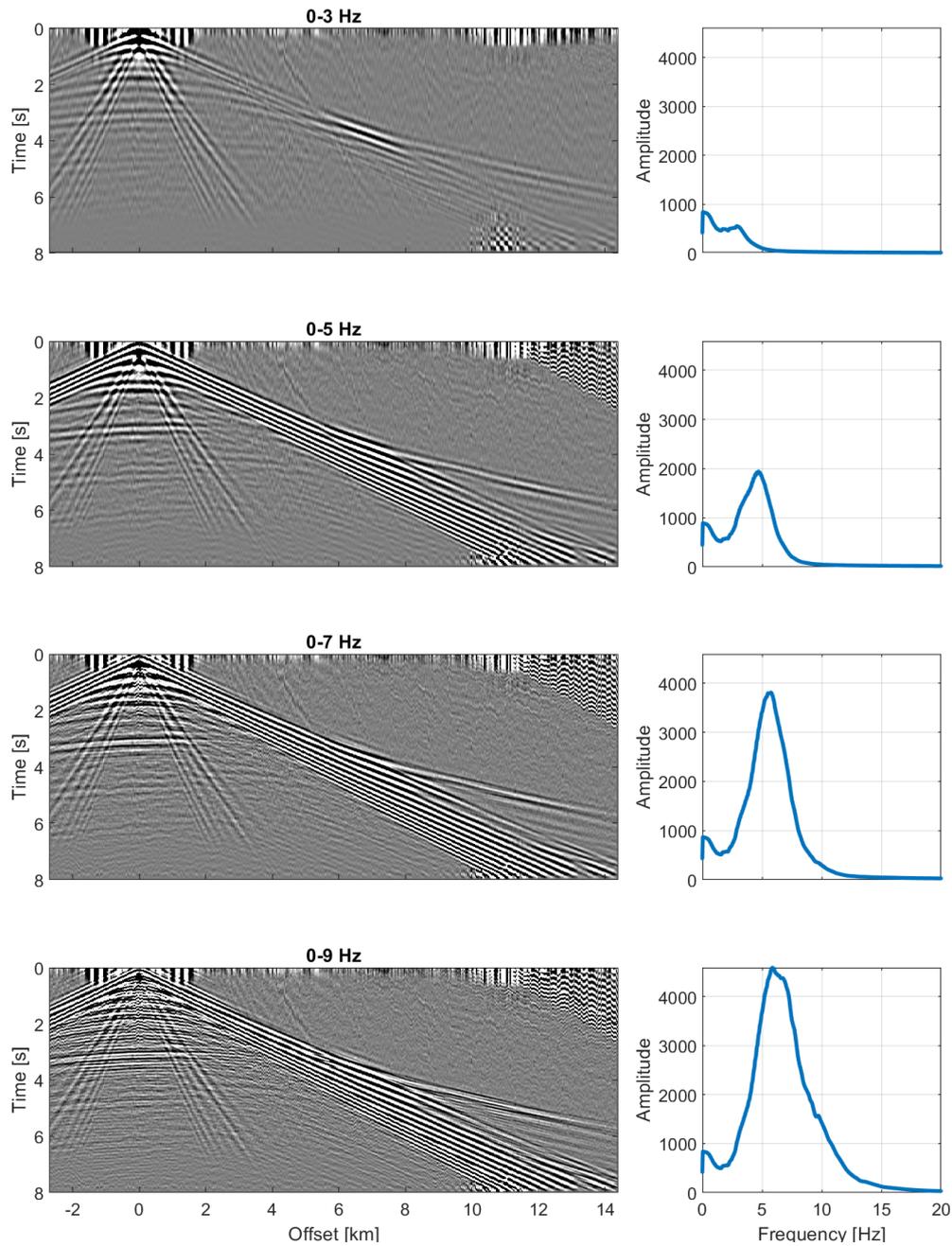


Figure 4.11: Low frequency panel for the western most pressure component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.

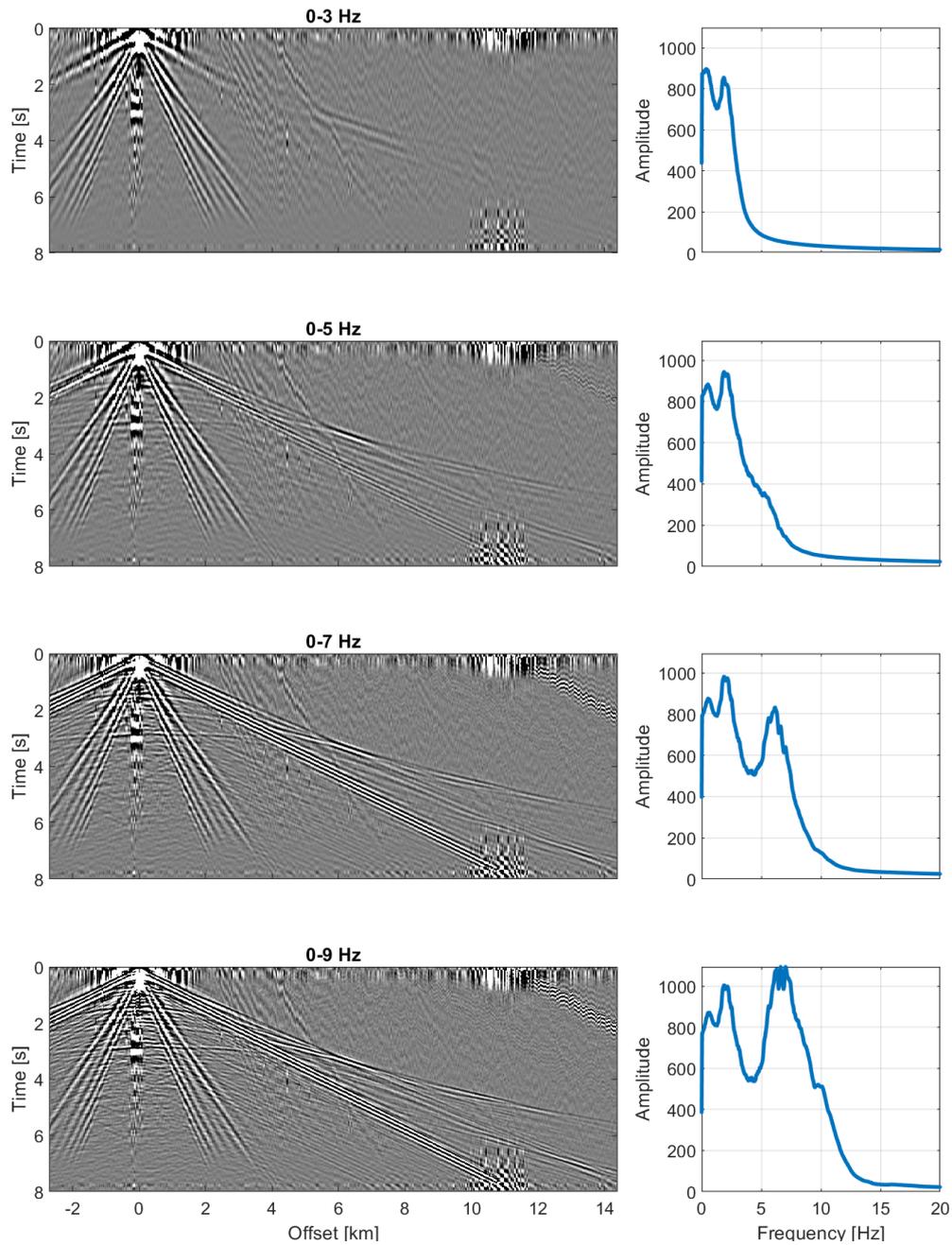


Figure 4.12: Low frequency panel for the western most vertical component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.

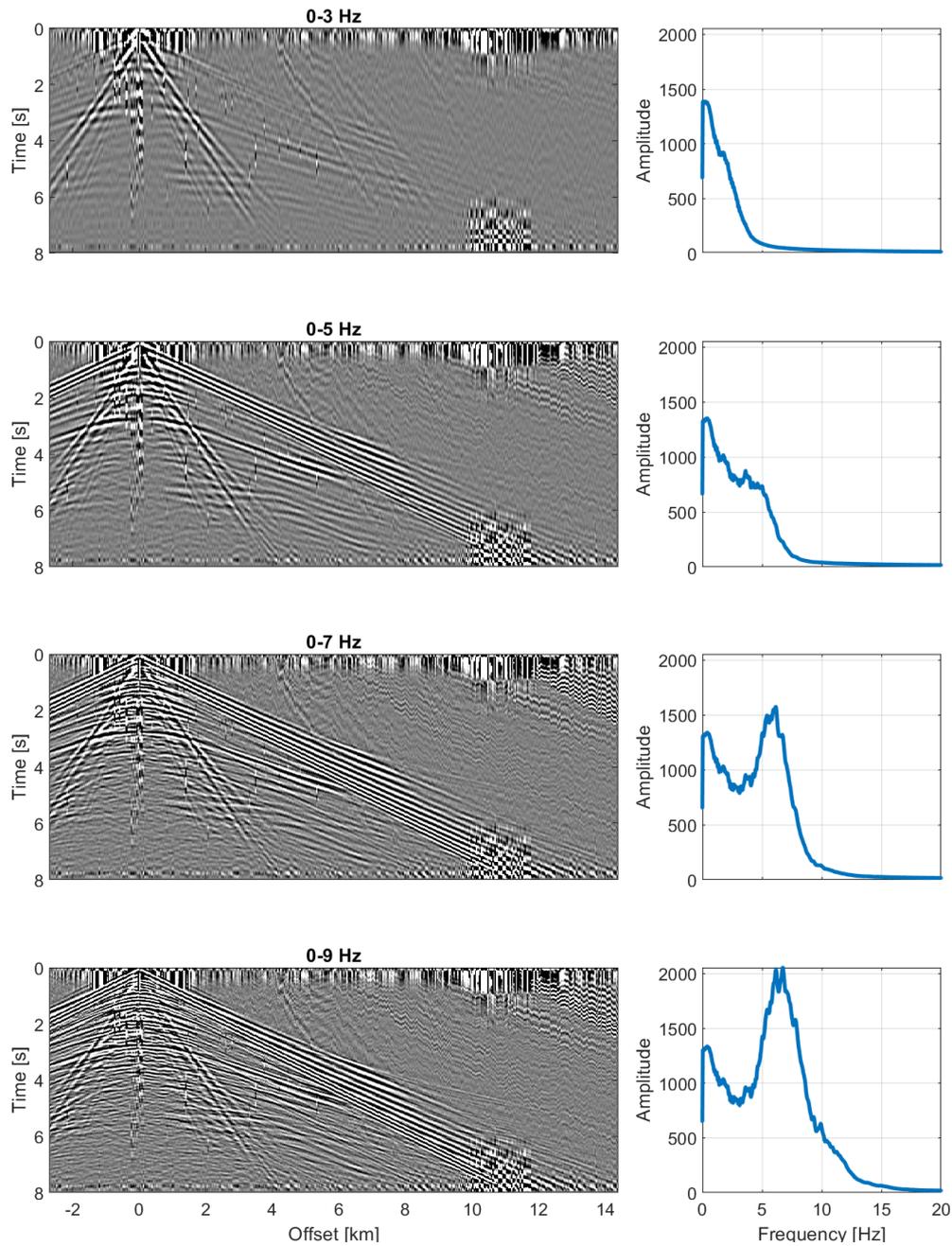


Figure 4.13: Low frequency panel for the western most radial component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.

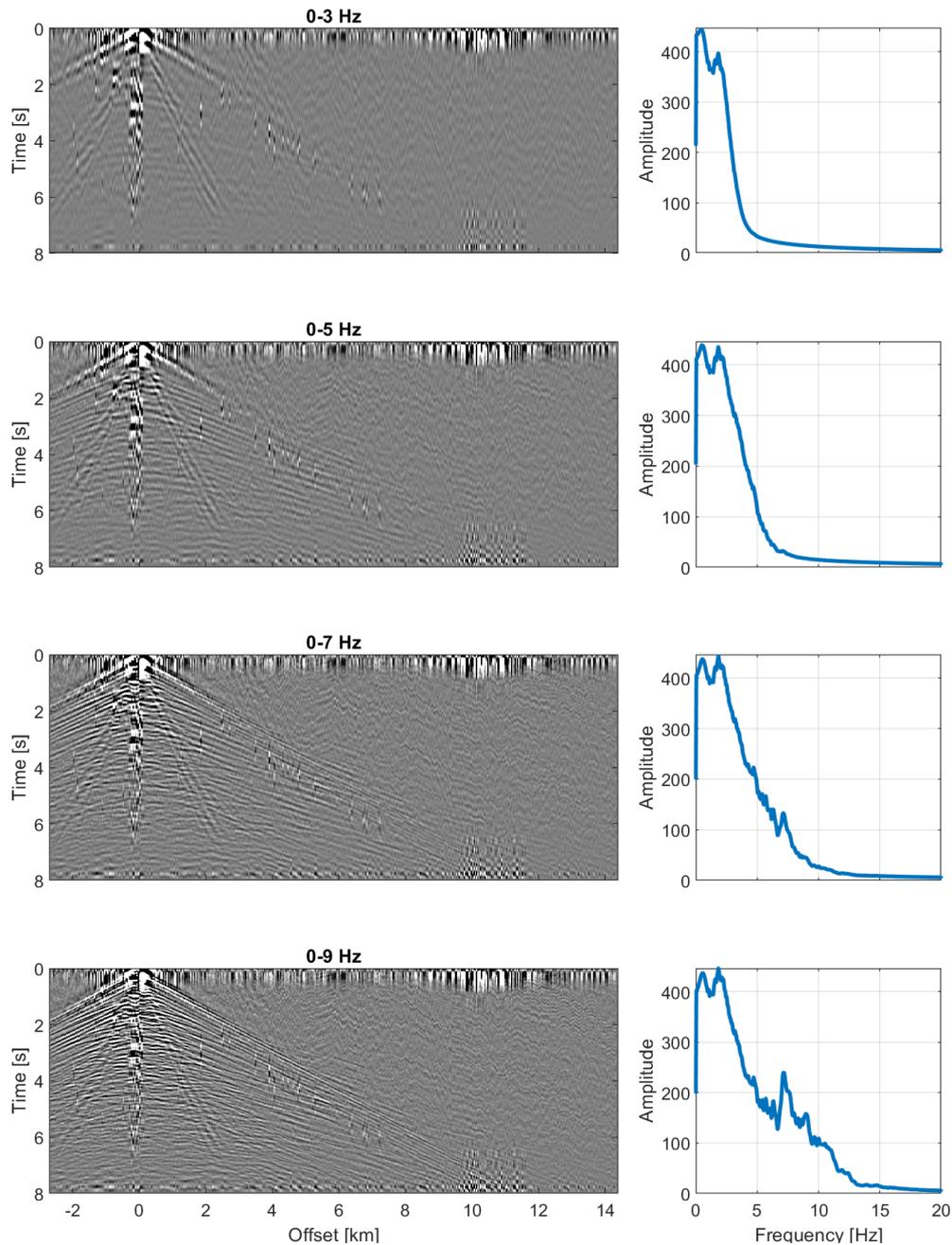


Figure 4.14: Low frequency panel for the western most transverse component receiver gather of the profile. The same color scale is used for all frequency bands and for all components. Extreme values of color scale are divided by 1000 to improve global visibility of events.

Direct arrivals, waterlayer multiples and guided waves: P-wave water velocity is usually assumed as constant and isotropic (esp. in shallow waters). It can be estimated from direct waves. The theoretical arrival times of direct waves from a point Dirac source shot at time 0 were computed and superimposed with the real data. Figure 4.15 shows the western most vertical velocity component receiver gather of the profile and the arrival times computed with: (1) the obtained best fit velocity value of 1478 m/s (manual fitting of the shape only of the arrival times), and (2) the usual 1500 m/s value (as in the provided model). Note that the picked direct wave is mixed with a seabed reflection since

the hydrophone is just above the seabed. CGG also estimated various water velocities for this dataset. The final water velocity value used by CGG in their velocity models is of 1478 m/s, agreeing with my water velocity estimation.

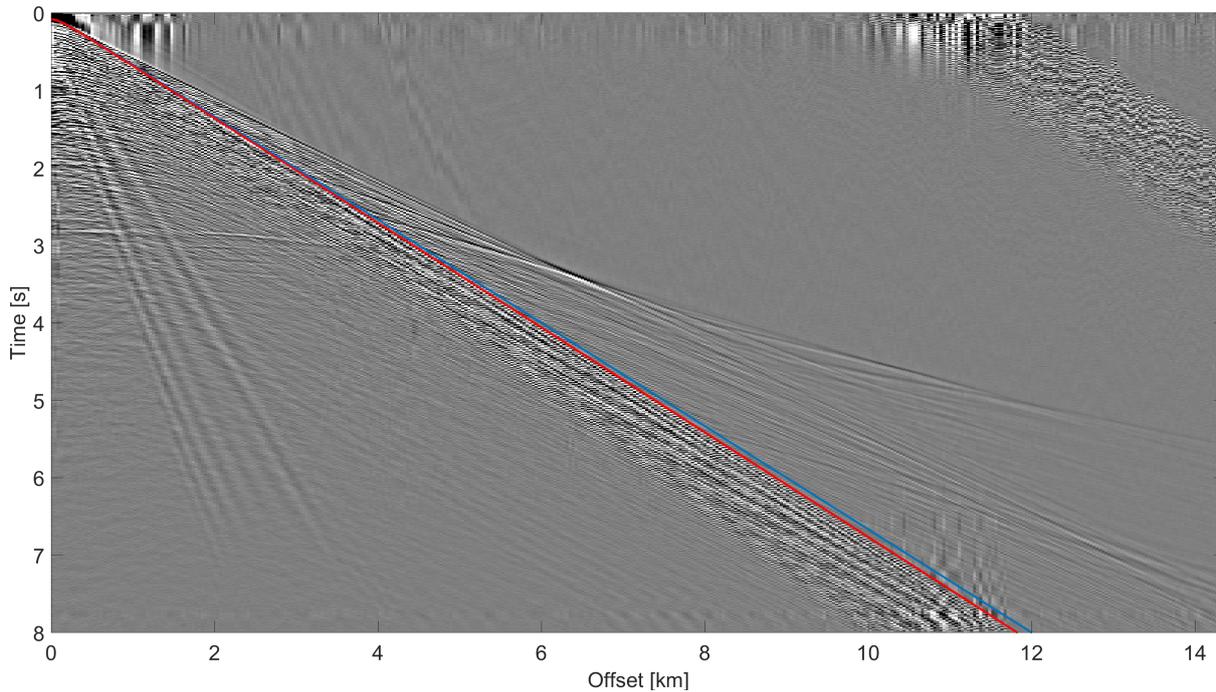


Figure 4.15: Real data superimposed with the computed direct arrival times. Vertical velocity component of the western most receiver gather of the profile. Amplitudes have been clipped. Red line (resp. blue): times computed with 1478 m/s (resp. 1500 m/s).

The shallow water environment generates strong and clustered waterlayer multiples (pink arrows in figure 4.10). These multiples generate guided waves (also called reverberations, ringing waves, leaky mode/waves, or ultra low frequency waves). Guided waves are the constructive interferences of multiply reflected waves trapped in the water layer, as explained by Nicolas et al. (2002) (135) for instance. With their high amplitudes, high frequencies and dispersive aspect, guided waves are visible in the data, as shown by the solid pink arrow in figure 4.10. Both waterlayer multiples and guided waves interfere with useful reflections. The removal of guided waves is a challenging issue: because their frequency and moveout approach those of reflections, their removal is not possible with usual F-K or τ -p filters, but can be done with (adaptive) subtraction of predicted guided waves (e.g., Ernst et al., 1998 and 2002(54) (53) and Boiero et al., 2013 and 2013b (26) (27)). Ogarague (2008) (139) succeeded in removing guided waves with adaptive subtraction of extracted guided waves with FK (Frequency-Wavenumber) filtering after normal move out correction.

Scholte waves: These fluid-solid interface waves are identifiable with their low frequency content, dispersive aspect (for heterogeneous sub-surfaces), high amplitudes and low apparent velocities (green arrows in figure 4.10). Scholte waves propagate in the solid, radiate in the fluid, and depend notably on the fluid and shallow solid properties. The lower the frequency, the deeper the Scholte wave propagates. If the S-wave velocity increases with depth, the lower the frequency, the faster the Scholte wave propagates.

Trial and error simulations can be performed to invert for these properties, assuming that the modeling algorithm correctly handles the seabed interface. More automatic inversions were performed through dispersion analysis of Scholte waves and possibly guided waves (e.g., Klein et al., 2005 (97); Kugler et al., 2005 (103); Strobbia et al., 2010 (184); Boiero et al., 2013 and 2014 (27) (28) and Tomar et al., 2015 (193)). Inversion of S-wave velocities from Scholte waves with FWI is a challenging task due to the high risk of cycle-skipping and the amplitude modeling difficulties. Borisov et al. (2019) (30) showed a successful inversion using layer stripping and an envelope-based cost function. Scholte waves might need to be removed. It can be done by FK filtering, τ -p filtering, (adaptive) subtraction of predicted Scholte waves, or both (e.g., Le Meur et al., 2010 (108); Strobbia et al., 2010 (184); Boiero et al., 2013 (26) and Ishiyama, 2013 (86)).

Vinh (2013) (211) derived the relation between the Scholte wave velocity c , the density ρ and the P- and S-wave velocities v_p and v_s for the fluid fl. and the solid so.:

$$0 = \frac{c^4 \rho^{\text{fl.}}}{v_s^{\text{so.4}} \rho^{\text{so.}}} \sqrt{1 - \frac{c^2}{v_p^{\text{so.2}}}} + \left(2 - \frac{c^2}{v_s^{\text{so.2}}}\right)^2 \sqrt{1 - \frac{c^2}{v_p^{\text{fl.2}}}} - 4 \sqrt{1 - \frac{c^2}{v_s^{\text{so.2}}}} \sqrt{1 - \frac{c^2}{v_p^{\text{so.2}}}} \sqrt{1 - \frac{c^2}{v_p^{\text{fl.2}}}} \quad (4.1)$$

This equation is valid for homogeneous isotropic fluid and solid media. To estimate the seabed S-wave velocities, equation 4.1 was applied to the highest frequency (and slower) surface wave arrivals (solid green arrow on figure 4.10). The computations were performed for three values of Scholte wave velocities: the lowest and highest values picked along the profile, and the value picked on the gathers of figure 4.10 (dashed green line). The water properties are assumed to be known ($\rho^{\text{fl.}} = 1030 \text{ kg/m}^3$ and $v_p^{\text{fl.}} = 1478 \text{ m/s}$) and a large range of P-wave velocity and density values were tried. Figure 4.16 shows the results, which suggest a poor sensitivity to P-wave velocities and densities of the seabed, as well as a small lateral variation of the seabed S-wave velocity between 400 and 520 m/s, implying a soft seabed. These values might be slightly overvalued: the measured velocity might correspond to a kind of mean of the sub-surface covered by the used surface wave arrivals, while the shallow sub-surface probably has velocities increasing with depth.

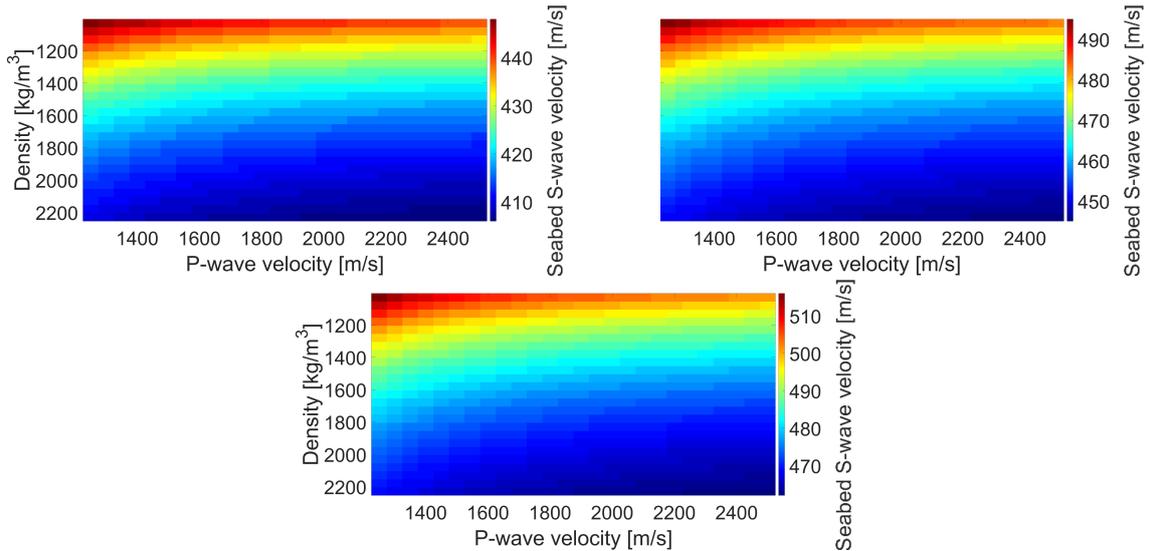


Figure 4.16: Computed seabed S-wave velocities (in m/s) for three Scholte wave velocities, from left to right: lowest picked velocity (375 m/s), velocity from figure 4.10 (400 m/s), highest picked velocity (415 m/s). Various density and P-wave velocity seabed values are tested.

This analysis assumes an isotropic shallow sub-surface, which seems to be adequate based on the invariance of modeled Scholte waves with and without anisotropy. For the anisotropic modeling, the seabed anisotropy values from the provided initial model are used.

Velocity model analysis: A VTI model coming from previous surveys is used as the initial model for this study. It contains P-wave vertical velocity, and Thomsen epsilon ϵ and delta δ parameters (Thomsen, 1986 (192)). No density model is available. This model probably includes information from well data, regional geological knowledge, horizon interpretation on migrated volumes, as well as travel-time tomography. A S-wave vertical velocity model was computed from the initial P-wave vertical velocity model, using an S- over P-wave vertical velocity ratio derived by CGG. This ratio was notably derived from horizon registration (top Skade, Lark, Dornoch, Cretaceous and Kyrre) on one of their updated model. Horizon registration consists in the inversion of S-wave velocities such that consistent horizon locations are obtained with P-P and P-S migrations. A 2D model is extracted from this initial model along the profile, using a nearest neighbor interpolation method. The approach is acceptable due to the $\sim 2.5D$ geometry and the maximum distance of 17 m between the furthest nearest points. The 2D model is shown in figure 4.17. It exhibits the expected geometry from the vintage migrated image (figure 4.3), with horizontal layers in the upper part, a tilted block in the deeper part, and strong interfaces. A comparison of the locations of the strong interfaces with the vintage migrated image (not shown) concludes to an excellent fit for the anisotropy parameters and possible mislocations for the velocity parameters. This model suggests strong VTI layers with an anisotropy higher than 5% and horizontal velocities higher than vertical velocities, in agreement with sedimentology at least for upper part. The velocity of the water layer is constant and equals to 1500 m/s. The seabed P- and S-wave vertical velocities are resp. ~ 1850 m/s and ~ 480 m/s, in agreement with the estimated values from Scholte wave analysis. The S-wave velocity model is less deep than the other, with a maximum depth of 6 km. The values of the deepest part of the other models seem to be poorly constrained.

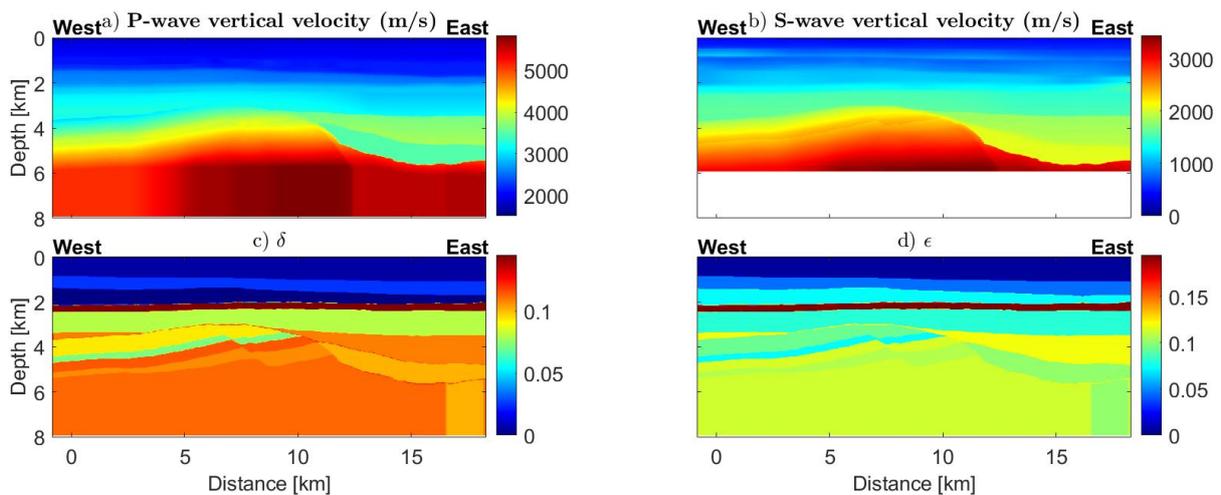


Figure 4.17: Initial model below the profile. The horizontal distance is computed along the profile from the western most source.

Figure 4.18 shows the seabed along the profile from: (1) the initial 2D velocity model,

(2) the receivers and (3) a seabed horizon. The fit between the depth from the receivers and the horizon is excellent, with an almost flat seabed at mean depth of 128 m. The depth of the seabed in the initial model could be improved.

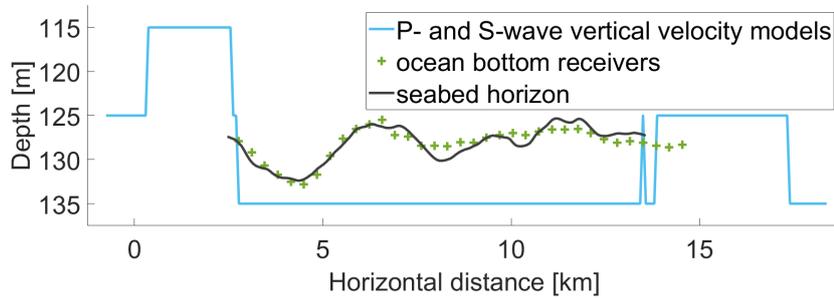


Figure 4.18: Seabed along the profile from: the provided initial 2D velocity model (blue), the depth of the receivers (green) and a seabed horizon (black). Horizontal distance is computed along the profile from the western most source.

Well data versus initial model: Well logs from 37 wells in the vicinity of the profile are available. Most of them are deviated production wells. 6 wells contain shear sonic, and only at the reservoir level. Most of the sonic logs are at the reservoir level, and there is no sonic log information above 500-meter-depth. 5 sonic logs were corrected with checkshots. Figures 4.19 and 4.20 show the location of a selection of wells (i.e., wells with shear sonic logs, checkshots corrected sonic logs, shallow sonic log information, or non deviated wells).

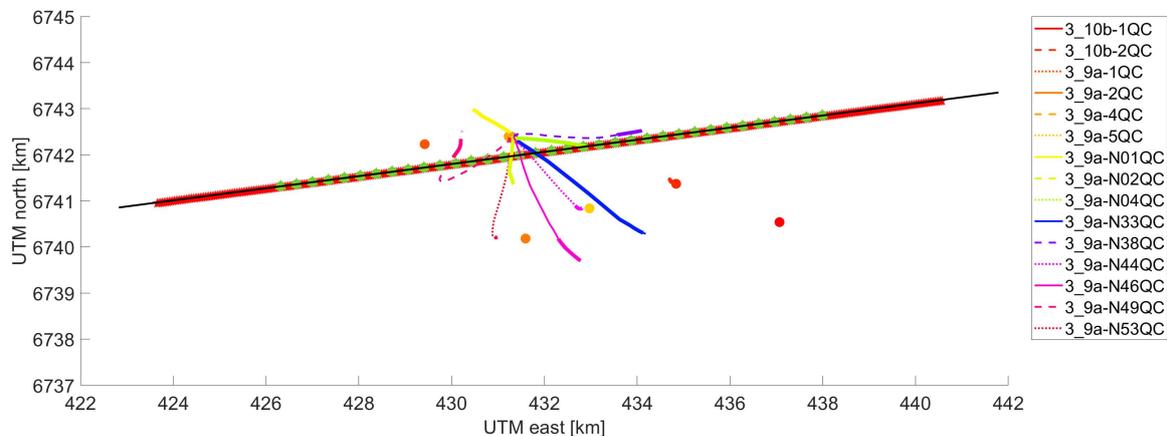


Figure 4.19: Map of wells and sonic logs locations with respect to the profile (black line). The red stars correspond to the sources and the green markers correspond to the receivers. Each well path is represented by a different color and line style combination, as described in the legend. The solid circles correspond to the head wells location and the thick solid lines overlaid on the well paths correspond to the extent of the sonic logs measurements.

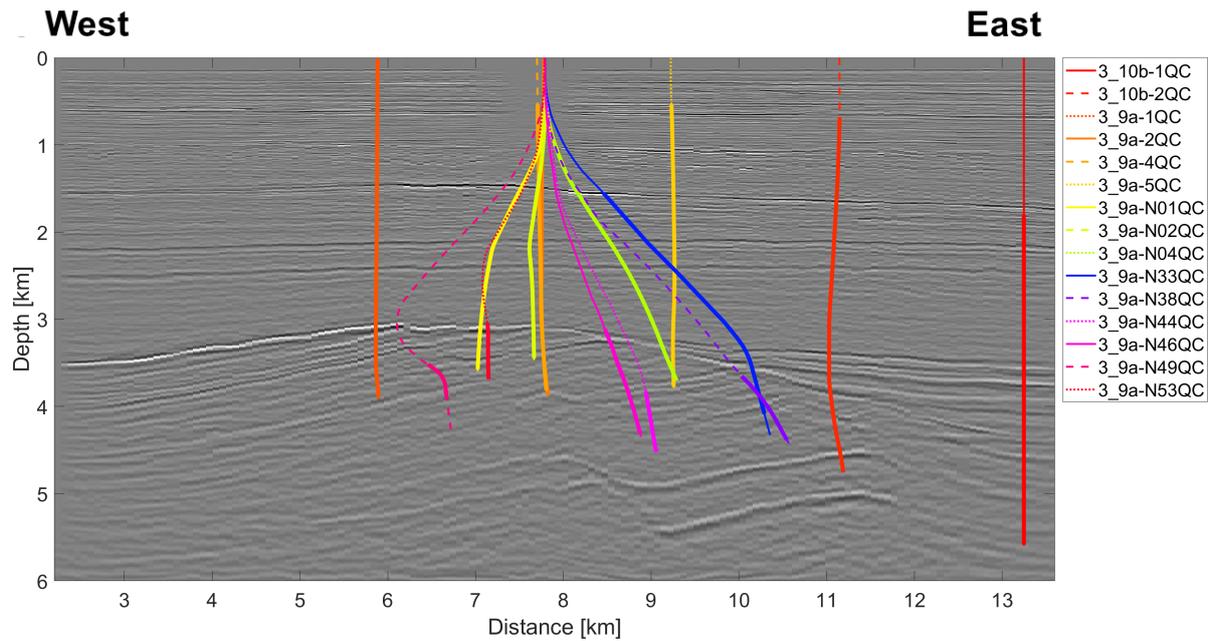


Figure 4.20: Wells and sonic logs locations on the vintage migrated section below the profile. Each well path is represented by a different color and line style combination, as described in the legend. The thick solid lines overlaid on the well paths correspond to the extent of the sonic logs measurements. Horizontal distance is computed along the profile from the western most source.

Figures 4.21, 4.22 and 4.23 show the corresponding smoothed P- and S-wave velocities and densities. Even though the well data have to be analyzed carefully due to their uncorrected and deviated nature (in an anisotropic environment), they indicate sub-horizontal layers in the upper part, strong interfaces and suggest a soft seabed from shallowest P-wave velocities combined with lithology. The fit between velocity logs and initial model is promising for FWI, but the strong interfaces in the initial model are not always correctly depth located.

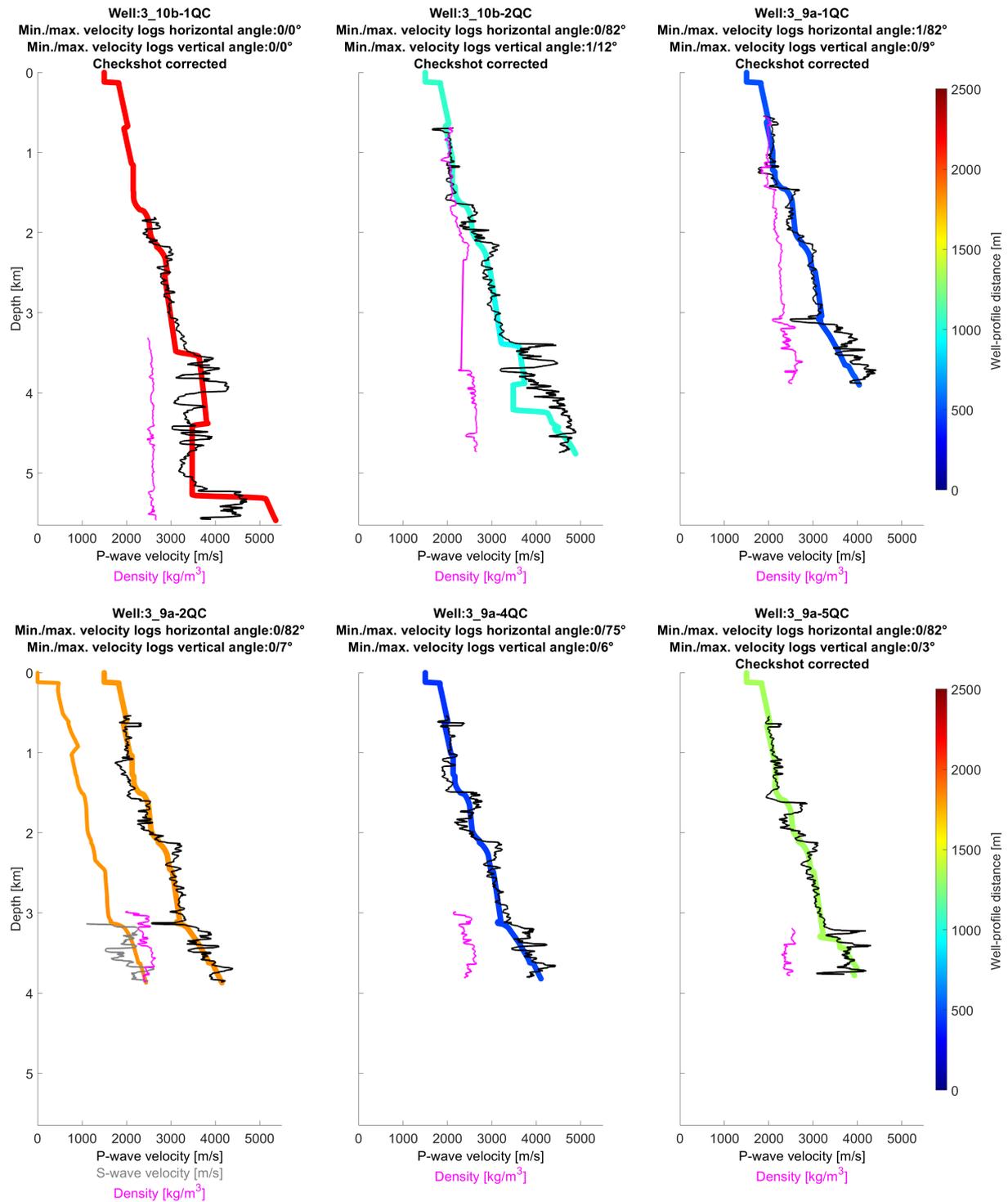


Figure 4.21: Smoothed density and P-wave velocities from a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The medium thick line corresponds to the initial S-wave vertical velocity model and the thick line corresponds to the initial P-wave vertical velocity model along the horizontal projection of the well on the profile. The colors of the line correspond to the horizontal distance between the well and the profile.

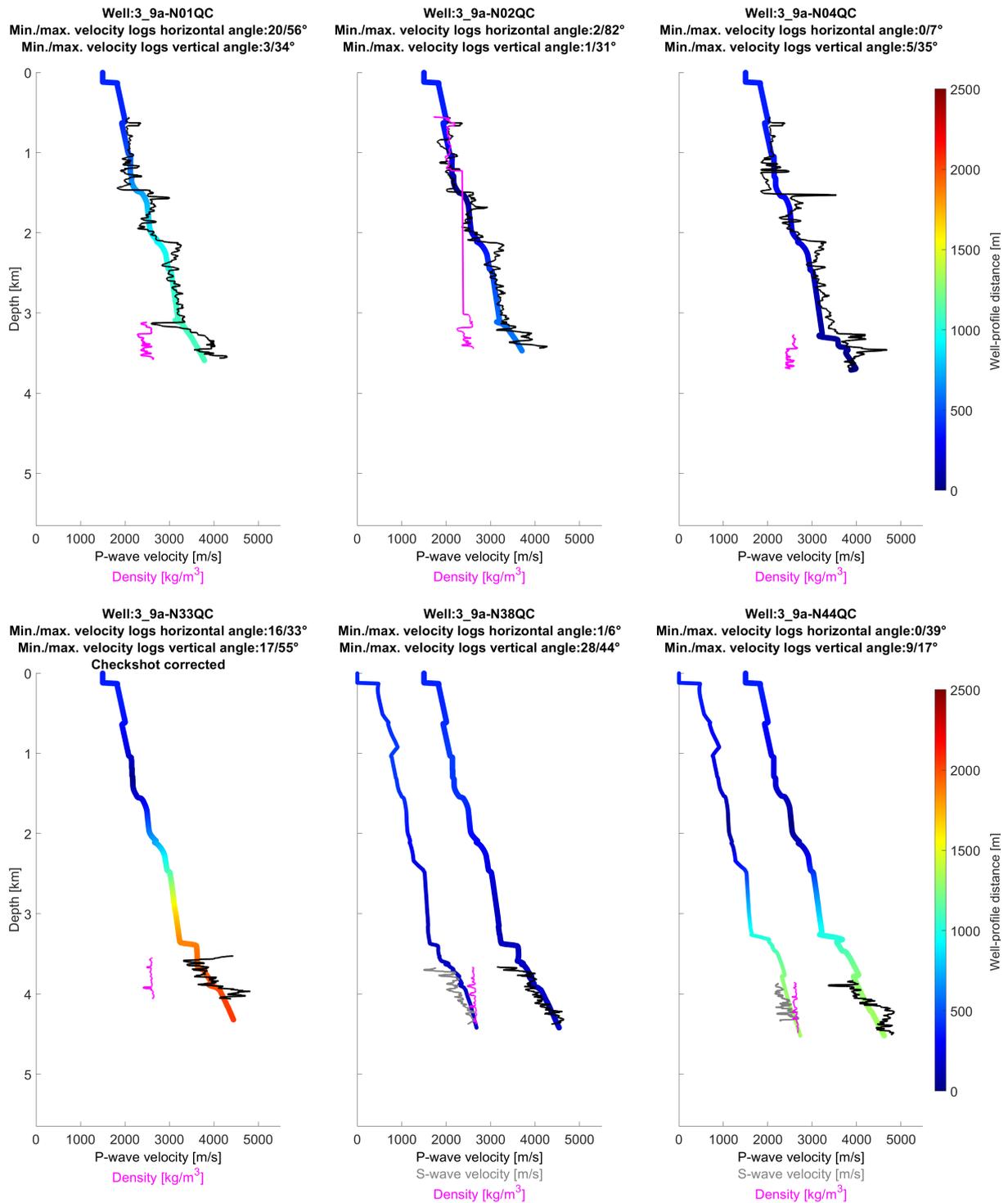


Figure 4.22: Smoothed density, P- and S-wave velocities (when available) from a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The medium thick line corresponds to the initial S-wave vertical velocity model and the thick line corresponds to the initial P-wave vertical velocity model along the horizontal projection of the well on the profile. The colors of both lines correspond to the horizontal distance between the well and the profile.

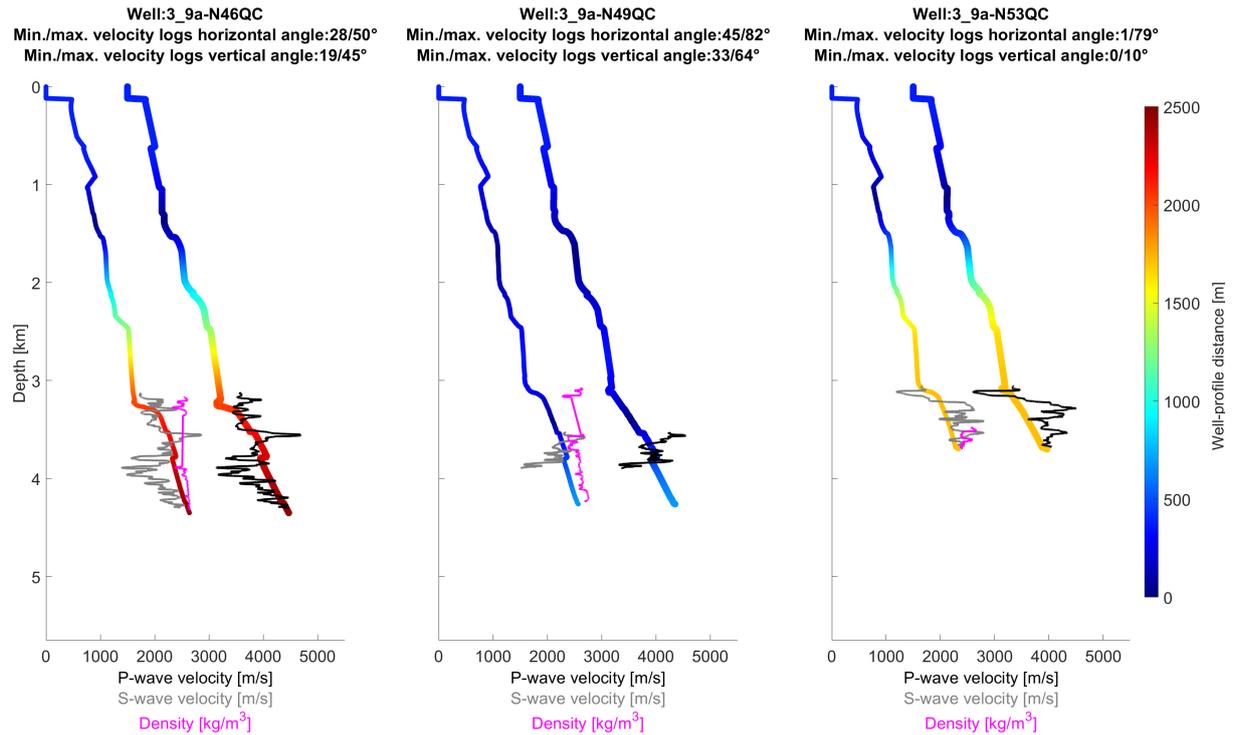


Figure 4.23: Smoothed density, P- and S-wave velocities (when available) from a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The medium thick line corresponds to the initial S-wave vertical velocity model and the thick line corresponds to the initial P-wave vertical velocity model along the horizontal projection of the well on the profile. The colors of both lines correspond to the horizontal distance between the well and the profile.

Headwaves: Compressive headwaves (also called diving or turning waves) are visible at far offsets, especially on the vertical velocity component where the waterlayer multiples are less pronounced. They can be used to determine the velocity and thickness of the layers of a layercake sub-surface with increasing isotropic velocity with depth, using the following equation:

$$t = \frac{x}{v_n} + \frac{h_0 \cos i_0}{v_0} + 2 \sum_{j=1}^{n-1} \frac{h_j \cos i_j}{v_j}, \quad \text{with } i_j = \arcsin \frac{v_j}{v_n}, \quad (4.2)$$

t being the time of the line fitting the headwave traveling within the n^{th} layer, x being the offset between the source and the receiver, h_j (resp. v_j) being the thickness (resp. velocity) of the j^{th} layer, starting from 0, the 0^{th} layer being the water layer. These equations are valid for seabed recordings with sources at the free-surface. The vintage migrated volumes, local geology knowledge and the straight, length and invariant appearance among gathers of the headwaves indicate a relative layercake-like sub-surface along the profile, at least for the upper part (< 3 km). As shown in figure 4.24, 5 headwaves were identified (i.e., 6 layers, including the water layer).

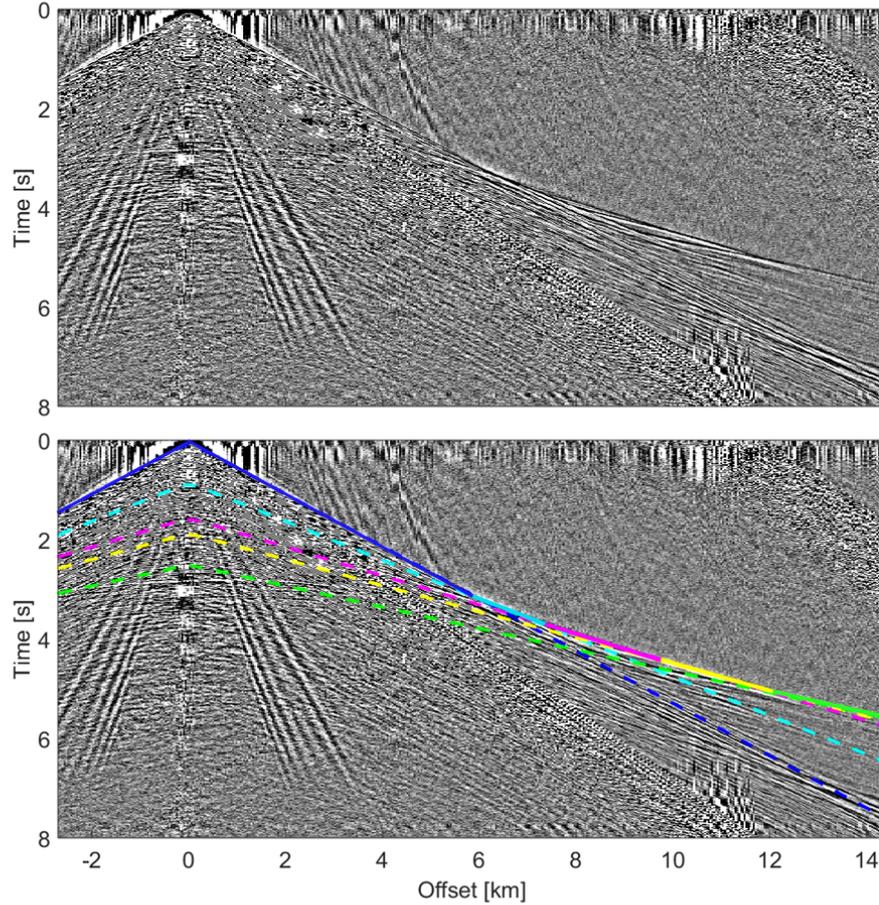


Figure 4.24: Western most receiver gather of the profile without (top) and with (bottom) the lines chosen to fit the headwaves. The solid lines correspond to the headwaves and the dashed lines correspond to the solid lines extension for all offsets. Vertical velocity component. Amplitudes are clipped for visualization purposes.

Applying equation 4.2, table 4.3 summarizes the computed thickness and P-wave velocity of each layer. The headwaves suggest strong interfaces as well as a soft seabed when combined with lithology.

Layer	P-wave horizontal velocity (m/s)	Time at zero-offset (s)	Thickness (m)
0	~ 1478 (from direct waves)	n.a	~ 120
1	~ 1900 (dark blue)	~ 0.05	~ 1110
2	~ 2560 (light blue)	~ 0.85	~ 970
3	~ 3430 (pink)	~ 1.5	~ 890
4	~ 3800 (yellow)	~ 1.87	~ 1710
5	~ 4660 (green)	~ 2.5	n.a

Table 4.3: Estimated thickness and P-wave horizontal velocity for each layer, as well as the time at zero offset of the line fitting the headwave traveling in the layer. The colors refer to the lines of figure 4.24. n.a. = not applicable.

The computed velocities from headwaves corresponds well to the well data, with slight differences which might come from anisotropy: the estimated velocities from the headwaves are horizontal velocities. The thicknesses do not fit very well with well data, mi-

grated volumes and horizons, especially for the deepest layers. I assume that it is mainly due to the non-precise fitting, to the velocity which might not be always increasing with depth, and to the fact that errors of fitting add up for the computation of the layer thickness. Based on a comparison with well log data, I find that headwaves penetrate down to 4 km depth in the central part of the model.

P-P and P-S reflections: Reflections are clearly visible on figures 4.10 to 4.14. The strong reflections on the hydrophone and the vertical velocity components (e.g., purple arrow) are interpreted as P-P reflections. The strong reflections on the radial velocity component (e.g., yellow arrow) are interpreted as P-S reflections. P-S reflections are identified by their low apparent velocity and decreasing amplitudes at near offset. Note that the amplitude decrease is not strong and the disappearance is not visible due to the clipped color scale, non-existing zero offset data, and not perfectly horizontal layers. The initial 2D model is used to predict the zero offset arrival times of P-P and P-S events coming from BCU, top Lomvi and a virtual 6-km-deep horizon. The computations assume a layer-cake model and uses the central part of the model, where the targets are. The results are gathered in table 4.4. The presence of P-P (resp. P-S) reflection around 4.1 s (resp 6 s) indicates a penetration of these waves up to 6 km depth.

Horizon	P-P reflection	P-S reflection
Base Cretaceous Unconformity (BCU)	~ 2.6 s	~ 4 s
top Lomvi	~ 3.7 s	~ 5.4 s
6-km-deep horizon	~ 4.1 s	~ 6 s

Table 4.4: Estimated zero-offset arrival times for P-P and P-S reflections on three horizons.

Shear-wave splitting: Also called birefringence, shear-wave splitting (SWS) coming from a layer right above top Skade horizon was observed by CGG. SWS is the phenomenon by which an incident S-wave, when entering into a specific anisotropic medium, splits into two S-waves traveling at different velocities and polarization directions (figure 4.25). Velocities and polarization directions depend on the medium parameters.

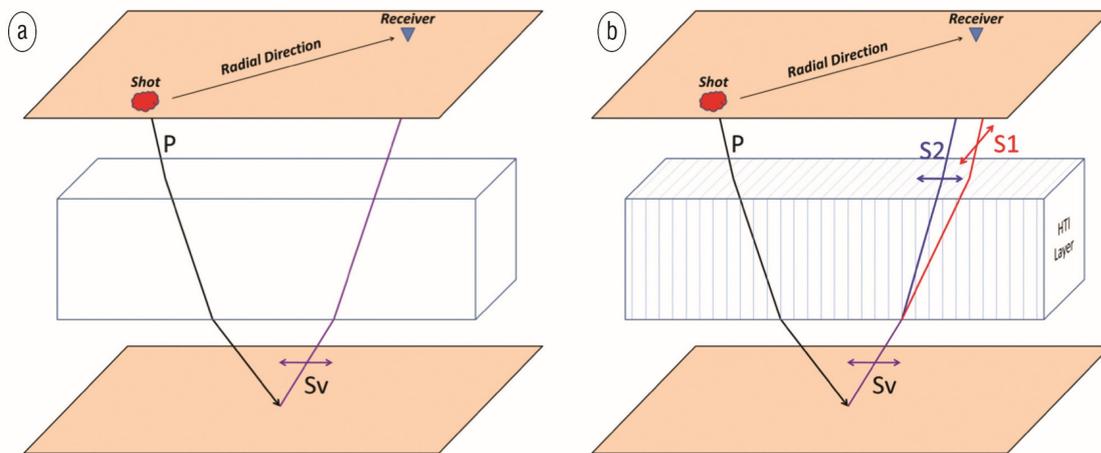


Figure 4.25: Sketch illustrating shear-wave splitting. Wave behavior without (left) and with (right) shear-wave splitting. Figure from Bale et al. (2013) (15). HTI = Horizontal Transverse Isotropy.

This phenomenon is not in contradiction with the anisotropy of the **VTI** kind, but adds up to it. In such a case, the **VTI** assumption remains valid, better than the isotropic assumption, but not as good as the combination of both kinds of anisotropy. Such a combination would be the next step, once fast elastic 3D inversions become possible.

SWS can be due to fractures or differential stress for instance (e.g., Crampin et al., 1980 (46); Thomsen et al., 1999 (191); Dillen, 2000 (51); Gumble and Gaiser, 2006 (70); Simmons, 2009 (174); Whale et al., 2009 (228); Sayers, 2010 (162); Cary et al., 2010 (37) or Wikel et al., 2012 (231) for instance). For data within 2-km-offset, CGG determined a delay time between the fast and the slow waves up to 25 ms, and a polarization for the fast waves roughly along the E-W direction.

4.4 Discussion and conclusions

This chapter illustrated how useful pieces of information can be obtained from a simple analysis (sometimes using simplified physics), and contribute to the three critical facets of **FWI** stated in the conclusion of **Chapter 1**, i.e., to:

- select/design a suitable **FWI** inversion strategy;
- constrain (pre-process) and Quality Control (**QC**) the **FWI** inputs;
- identify **QC** tools for **FWI** outputs.

The suitability of the soft seabed strategy, as described in **Chapter 3**, is justified by the soft seabed environment, the availability of wide-aperture P-P reflections data and multi-component seabed data, as well as the separation over components of P-P and P-S reflections. The soft seabed was mainly confirmed by the Scholte wave analysis, and is in agreement with both the lithology (combined with the water depth, and the headwaves and sonic well data analysis) and the provided initial model. The separation of P-P and P-S reflections over components is made possible by the sub-horizontal geology along the profile, especially in the shallow sub-surface below 3-km-depth. This geometry is visible on the vintage depth-migrated image and its horizons, on the well data, and on the appearance of the reflections and headwaves, and is in agreement with the bathymetry and the provided initial model.

I showed that the sub-surface was approximated by the **VTI** elastic choice of realism, which is supported by the combination of the lithology and the horizontal geological layering. This choice of realism justifies the need for the analysis of the performance of the soft seabed strategy with slightly erroneous **VTI** models, as done in **Chapter 3**. Because of the tilting layers, a **TTI** approximation could have been a better choice below the **BCU**.

Motivated by the limited computing resources, the choice to perform a 2D inversion of 3D data is supported by:

- the \sim N-S 2.5D geometry perpendicular to the profile, as expected from the Viking graben orientation and confirmed by the seismic horizons, and in agreement with the bathymetry, the lateral extension of the reservoirs and the shallow well data;
- the small distances between the profile and the sources/receivers;
- the orientation of one of the **SWS** axis roughly along the profile direction, which should minimize the impact of the impossibility to model **SWS** with the 2D simulations.

Note that this choice of dimensionality implies a geometrical spreading correction to be applied to the seismic data as a preparation step, as explained in [Chapter 1](#), as well as a re-location on the profile of the used sources and receivers.

Aside from the density model and the source time functions which will need to be estimated, the provided [FWI](#) inputs are promising overall. In addition to the already mentioned agreement between the model structures and the migrated image, the analysis described in this chapter allowed to constrain the water P-wave velocity to 1478 m/s through a direct wave analysis, to constrain the seabed depth (from the seabed seismic horizon and the receivers depth), and to confirm the seabed S-wave vertical velocity of the provided velocity model with a Scholte wave analysis. Because of the seabed being almost flat and of the poor handling of topography by the used finite-difference modeling algorithm, I propose to set the seabed depth at the constant value of 128 m (the computed mean value along the profile). The initial P- and S-wave vertical velocity models are compatible with the well data, apart from slight depth mislocations of interfaces, which can be handled by vertical smoothing. The targets are between 4 and 5-km-depth. The poor quality of the P-wave vertical velocity and anisotropy models and the missing S-wave vertical velocity values below 6-km-depth can then be handled by cutting the models below 6-km-depth. This reduced model size also allows to decrease the computing needs.

The data analysis ensured that the targets were illuminated by the recorded P-P and P-S reflections, with estimated zero-offset arrival times of ~ 3.7 s and ~ 5.4 s respectively for top Lomvi horizon, located below the targets. I propose to mute the useless later arrivals (i.e times greater than 4.5 and 7 s respectively, taking into account the move-out and the lateral variability along the profile), which are noisy, damaged by the processing and inexplicable by the model. This reduced time could also allow to decrease the computing resources needs. In contrast, the wide-aperture data necessary to build the intermediate scale P- and S-wave models (Shipp and Singh, 2002 ([172](#)) and Sears et al. 2008 ([164](#)), see [Chapter 1](#) and [Chapter 3](#)) do not illuminate the deepest targets: the maximum penetration depth is around 4 km in the central part of the model, and less on the sides. This lack of illumination might be compensated by the high quality low-frequency data below 3 Hz, allowing to extensively benefit from the increasing frequency multiscale strategy (Bunks et al., 1995 ([35](#)), see [Chapter 1](#)). The regular acquisition should allow for a reasonable illumination of the sub-surface, but the sparse receivers are expected to leave an imprint in the updated models, which could be handled by spatial smoothing. Note that the fixed dense sources and sparser receivers acquisition is a typical geometry for which reciprocity as described in chapter [Chapter 2](#) is needed to decrease the computing resources need. Finally, the acquisition and processing allowed for a limited imprint of noise on the data, with short-offset and wrap-around noises that can be handled with time-offset muting and [FK](#) filtering, and spikes that can be handled by filtering, muting and interpolation.

This chapter's analysis highlighted some (additional) aspects that might reduce the success of the inversion. First, the surface waves and the guided waves (due to the flat and shallow seabed) might disturb the inversion. Indeed, because of their high amplitudes, difficult inversions, and shallow-only information, [FWI](#) might focus on these waves and not update the desired deeper parts of the model. Second, the physical meaning of the data may have been damaged by the unclear de-bubbling and zero-phasing processing steps. These two points will have to be addressed during inversion.

Finally, regarding the [QC](#) tools for [FWI](#) outputs, in addition to the usual methods (such as the analysis of the data fit evolution), Alwyn real data benefit from P- and S-wave velocity information from wells (> 500 meter-depth and 3 km-depth respectively),

structural information from a vintage depth-migrated image, and an estimation of seabed S-wave velocities from a Scholte wave analysis. These additional QC tools turn Alwyn real dataset into a good candidate for FWI methodological (i.e., inversion strategy) studies, especially the well data (even though they are from deviated and non-corrected logs in an anisotropic environment).

Chapter 5

Alwyn real data: full waveform inversion application

Chapter's objectives and highlights

- Assess the performance of the two 2D full waveform inversion strategies theoretically and synthetically discussed in [Chapter 3](#) (i.e., the original soft seabed strategy and its variant), on a real dataset which exhibits usual real data complexities. Indeed, those FWI applications notably reveal imperfectly modeled amplitudes and erroneous anisotropy parameters. I recall that those strategies aim at obtaining P- and S-wave vertical velocities in soft seabed vertical transverse isotropic environments, the remaining anisotropic parameters not being updated, and possibly slightly erroneous. The assessment concludes to the success and usefulness of the variant strategy, especially considering the observed lower performance of the original strategy. Based on the findings of [Chapter 3](#), this behavior of the original strategy is expected to be specific to the application, and the usefulness of such strategy is not questioned. For instance, the original strategy might remain preferable for applications with less amplitude and eventually more anisotropy modeling defects.
- Provide a real data example of full waveform inversion application to retrieve both P- and S-wave vertical velocities in soft seabed environment, with results supported in an extensive way by verification tools.
- Illustrate the need for a good initial S-wave vertical velocity model and/or suitable inversion strategy, for meaningful P- and S-wave vertical velocity model update.
- Highlight the fact that full waveform inversion can show robustness to relatively strong noises (here, inaccurately modeled Scholte waves). Some possible explanatory mechanisms are proposed.

Contents

5.1	Introduction	106
5.2	Inputs preparation	108
5.2.1	Initial model	108
5.2.2	Seismic data	110
5.2.3	Acquisition geometry	114
5.2.4	Source function estimation	115
5.3	Application of the variant of the soft seabed strategy	117

5.4	Application of the original soft seabed strategy	127
5.5	Original strategy with a modified stage order	131
5.6	Inversions with Scholte waves	135
5.7	Discussion and conclusions	141

5.1 Introduction

In [Chapter 3](#), a 2D Full Waveform Inversion (FWI) strategy (developed by Sears et al. (2008) ([164](#)) in an isotropic context) and a variant strategy were theoretically and synthetically explored, in order to obtain P- and S-wave vertical velocities in soft seabed Vertical Transverse Isotropic (VTI) environments, the remaining anisotropic parameters not being updated, and possibly slightly erroneous. Those remaining anisotropy parameters are hereafter called “remaining anisotropy parameters” or “anisotropy parameters”. To allow for an independent reading of this chapter, I recall in this paragraph the characteristics of those strategies, but I refer the reader to [Chapter 3](#) for more details. VTI environments are environments where the wave velocities depend on the angle of the wave direction of propagation, measured from the vertical direction (see Thomsen, 1986 ([192](#)) for instance). Soft (or low) seabed environments are characterized by seabeds with S-wave velocities typical of siliclastic seabed sediments (around or below 300 m/s, see Hamilton, 1976 ([74](#)), and Castagna et al., 1985 ([39](#))). Obtaining both P- and S-wave velocities in such environments was made difficult due to a poorer S-wave velocity information content in seismic data (Sears et al., 2008 ([164](#)); Barnes and Charara, 2009 ([16](#))), compared with land and other marine environments. Slightly erroneous models of the anisotropy parameters are models which allow for an approximate but not perfect fitting of both wide- and narrow-aperture P arrivals, without cycle-skipping, allowing for a reasonable P-wave velocity inversion. This situation is inspired from real-case applications. The original strategy is, to my knowledge, the only published successful strategy in soft seabed environments, and was applied on synthetic isotropic data (Sears et al., 2008 ([164](#)); Brossier et al., 2009 ([32](#)) and 2010 ([33](#))). To summarize the studied strategies, they use different sources of information that can be isolated, and use them in a sequential manner following their domination and resolution power, as detailed in [figure 5.1](#). The isolation of the information relies on multi-component ocean bottom data with a separation between the horizontal velocity component and the vertical velocity (or pressure) component of P-wave and S-wave-related arrivals. This separation occurs in (sub-)horizontal sub-surfaces. Simple synthetic tests concluded that the original strategy is expected to be more robust than its variant to errors in the remaining anisotropy parameters. In contrast, the variant is expected to suffer less from amplitude modeling approximations (e.g., from unmodeled intrinsic attenuation).

In this chapter, the strategies are put into practice on Alwyn real data, which exhibits the characteristics for which the approaches were designed (see [Chapter 4](#)). Strategies assessment on real data is dictated by the numerous aspects that are not taken into account during the synthetic and theoretical studies. Among those aspects, one can cite noisy data, poorly known sources/acquisition geometry or physical approximations. Assessment on real data is part of the last steps towards the industrial use of a strategy, and requires the availability of validation tools. The original strategy has already been tested on real data (Sears et al., 2010 ([165](#))), but Alwyn real data allows for an extensive

assessment of the original strategy and its variant, with the availability of lithology, P- and S-wave velocity information from well data, as well as a vintage migrated image.

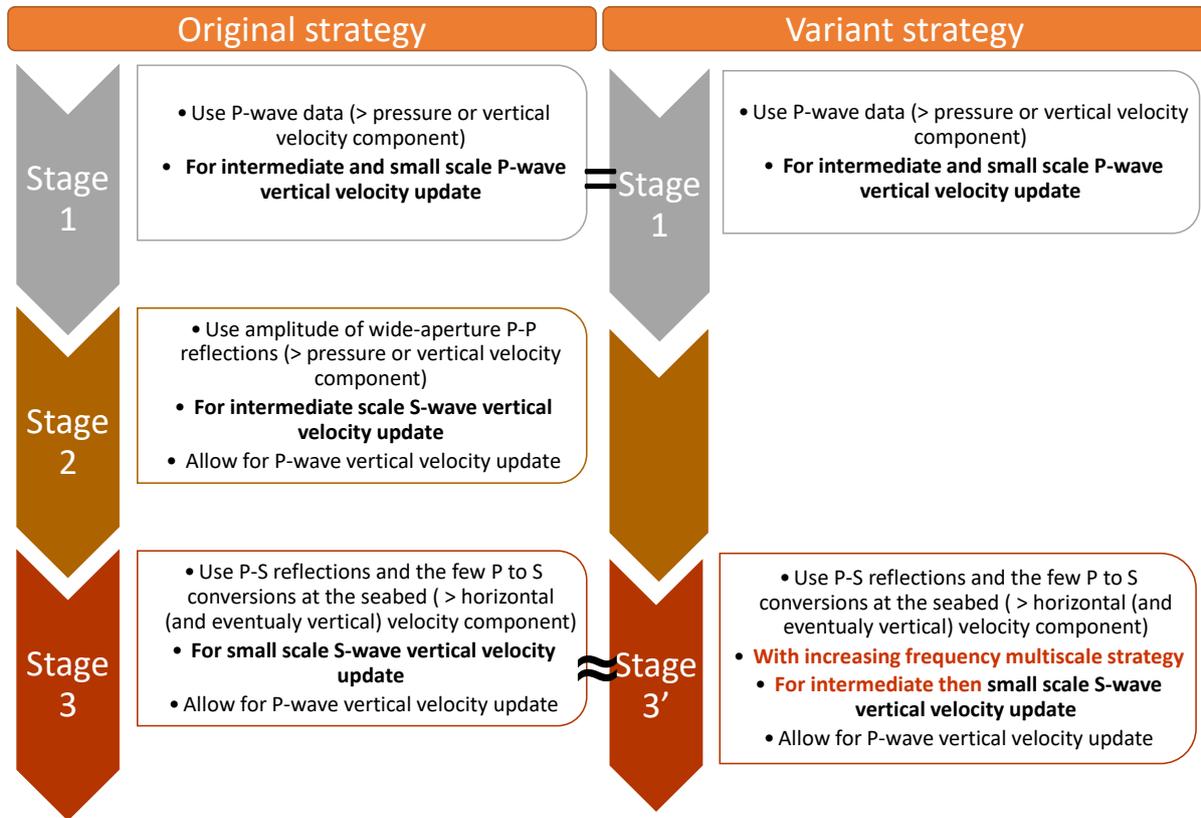


Figure 5.1: Description of the original soft seabed strategy and its variant, applied in a Vertical Transverse Isotropic (VTI) context to obtain P- and S-wave vertical velocities in soft seabed, the remaining anisotropy parameters not being updated. Note that the original soft seabed strategy contains three successive stages (i.e., stages 1, 2 and 3), while the variant contains only two stages (i.e., stages 1 and 3'). Indeed, the intermediate scale S-wave velocity update of stage 2 is replaced by the application of the increasing frequency multiscale strategy (Bunks et al., 1995 (35)) during stage 3'. Consequently, the variant strategy entirely relies on the low frequency data content, instead of on the amplitude of wide-aperture P-P reflections. The wide-aperture and S-wave related data are approximately selected with time-offset mutes. Note that it is advised to use the increasing frequency multiscale strategy at all stages, as is usually done for real data applications (e.g., Sears et al. (2010) (165)).

As explained in Chapter 1, in addition to the need for quality control tools for FWI outputs, FWI requires a number of inputs (an initial model and observed data with the adequate acquisition geometry and source functions). Consequently, this chapter starts with the description of the inputs preparation, before depicting the application of the strategies.

5.2 Inputs preparation

5.2.1 Initial model

Following Chapter 3's parameterization assessment, the initial model is converted into the $(V_{pv}, V_{sv}, \sqrt{1+2\eta}, \sqrt{1+2\delta}, \rho)$ parameterization, using equation F.7.

The initial model preparation also follows Chapter 4's assessment. The model is cut below 6-km-depth, the water velocity is set to 1478 m/s and the seabed interface is relocated at 128-meter-depth. Compared with the picked seabed horizon, the maximum vertical location error is around 5 m. In practice, with a spatial sampling of 20 m (i.e., the largest sampling used in this thesis) the seabed is located between 120 and 140 meter-depth. With a spatial sampling of 4 m (i.e., the smallest sampling used in this thesis) the seabed is located between 128 and 132 meter-depth. Each model parameter type is spatially smoothed. The smoothing is performed with a Gaussian window ($\sigma = 200$ m). To preserve the seabed properties which show plausible vertical P- and S-wave velocities, the size of the smoothing window is reduced when approaching the seabed, so that no smoothing is performed at and above this interface. Vertical profiles in the middle of the smoothed and non-smoothed models are provided in figure 5.2.

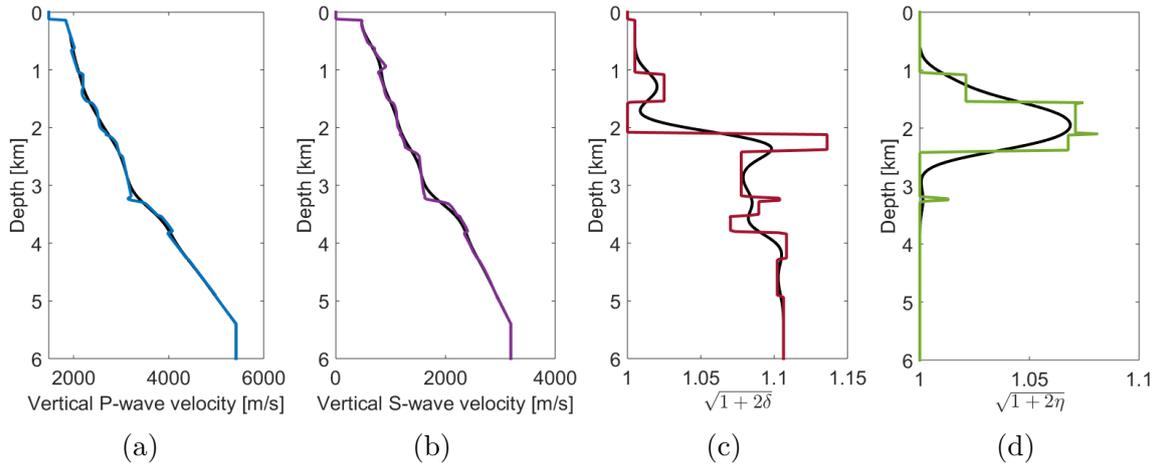


Figure 5.2: Vertical profiles in the middle of the smoothed (black line) and non-smoothed (colored line) initial models. (a) P-wave vertical velocity model; (b) S-wave vertical velocity model; (c) $\sqrt{1+2\delta}$ model; (d) $\sqrt{1+2\eta}$ model. Below the targets (which are between 3 and 5 kilometer depth), the deepest part of the models have been replaced by a vertically homogeneous layer for the absorbing boundary, and will not be shown in the next figures.

The missing density model ρ is computed from the initial P-wave vertical velocity V_{pv} model using the empirical relationship 2.22. This equation is recalled hereafter for the

chapter's completeness:

$$\left\{ \begin{array}{l} \rho = 1030 \text{ kg/m}^3 \text{ in water,} \\ \rho = 2351 - 7497 * (V_{pv} * 10^{-3})^{-4.656} \text{ kg/m}^3 \text{ if } V_{pv} < 2000 \text{ m/s,} \\ \rho = x_a + \frac{V_{pv}-2000}{150} * (x_c - x_a) \text{ kg/m}^3 \text{ if } V_{pv} \geq 2000 \text{ m/s and } V_{pv} \leq 2150 \text{ m/s,} \\ \quad \text{with } x_a = 2351 - 7497 * (2000 * 10^{-3})^{-4.656} \text{ and } x_c = 1741 * (2150 * 10^{-3})^{0.25}, \\ \rho = 1741 * (V_{pv} * 10^{-3})^{0.25} \text{ kg/m}^3 \text{ if } V_{pv} > 2150 \text{ m/s,} \end{array} \right. \quad (5.1)$$

where V_{pv} is in m/s. The empirical expression of Hamilton (1978) (75) is used for the shallow seabed poorly consolidated sediments ($V_{pv} < 2000$ m/s), and the empirical expression of Gardner et al. (1974) (65) is used for the deeper consolidated sediments ($V_{pv} > 2150$ m/s). A linear transition between the two empirical relations is applied. These relationships were derived from sediments and sedimentary rocks partially of the same type as Alwyn area ones. Note that these two empirical relationships were derived assuming isotropic media and are here applied in anisotropic media using the P-wave vertical velocities V_{pv} in place of P-wave velocities. Figure 5.3 provides a few comparisons between well density data and the obtained initial density model. Those comparisons show a reasonable fit at very large scale, but a questionable fit at smaller scale. Consequently, the use of a better relationship is to consider for future applications. In pursuit of such goal, the available well log data could be used to derive this relationship.

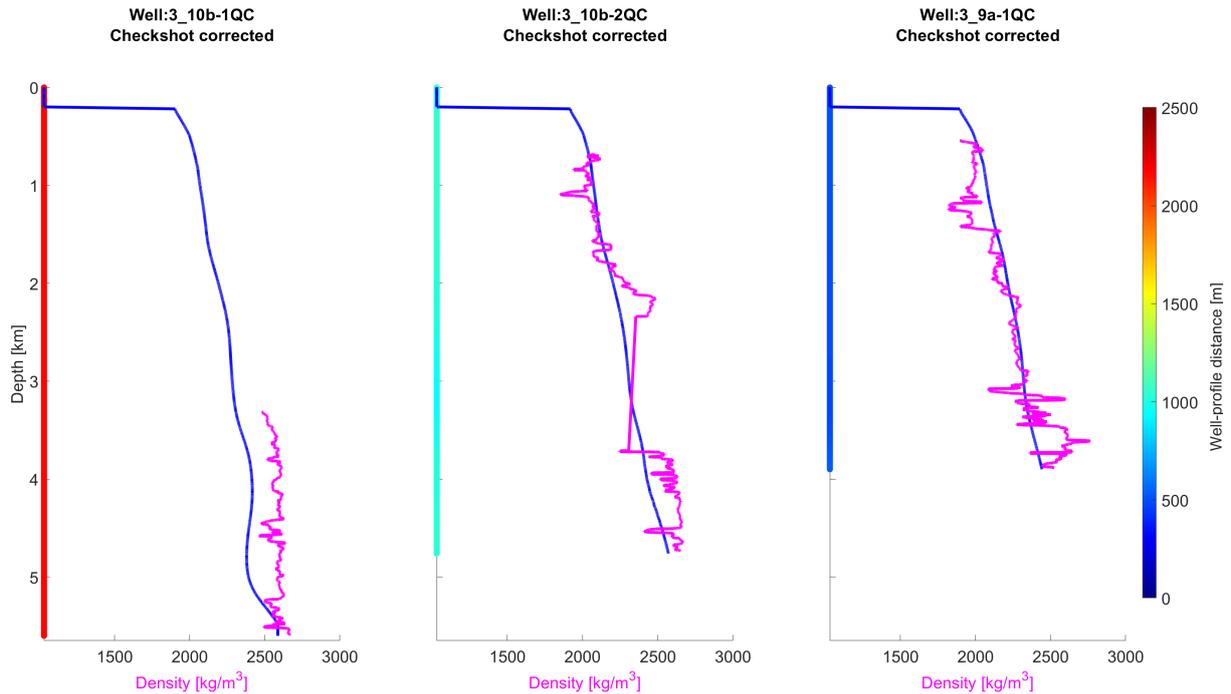


Figure 5.3: Density from the initial model (blue line) compared with smoothed density from well data (magenta line) for a selection of wells. The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line indicates the horizontal distance between the well and the profile.

5.2.2 Seismic data

Based on our analysis discussed in [Chapter 4](#), the seismic data preparation consists in:

1. **A high-cut filter.** Given the limited computing resources and the modeling stability criteria provided in equations 1.7 and 1.8, it is necessary to limit the highest modeled frequency. The highest modeled frequency is set here to ~ 9 Hz. In the interest of modeled and observed data comparability, a causal butterworth filter with a cut-off frequency of 10 Hz and an order of 14 is applied ([figure 5.4](#)).

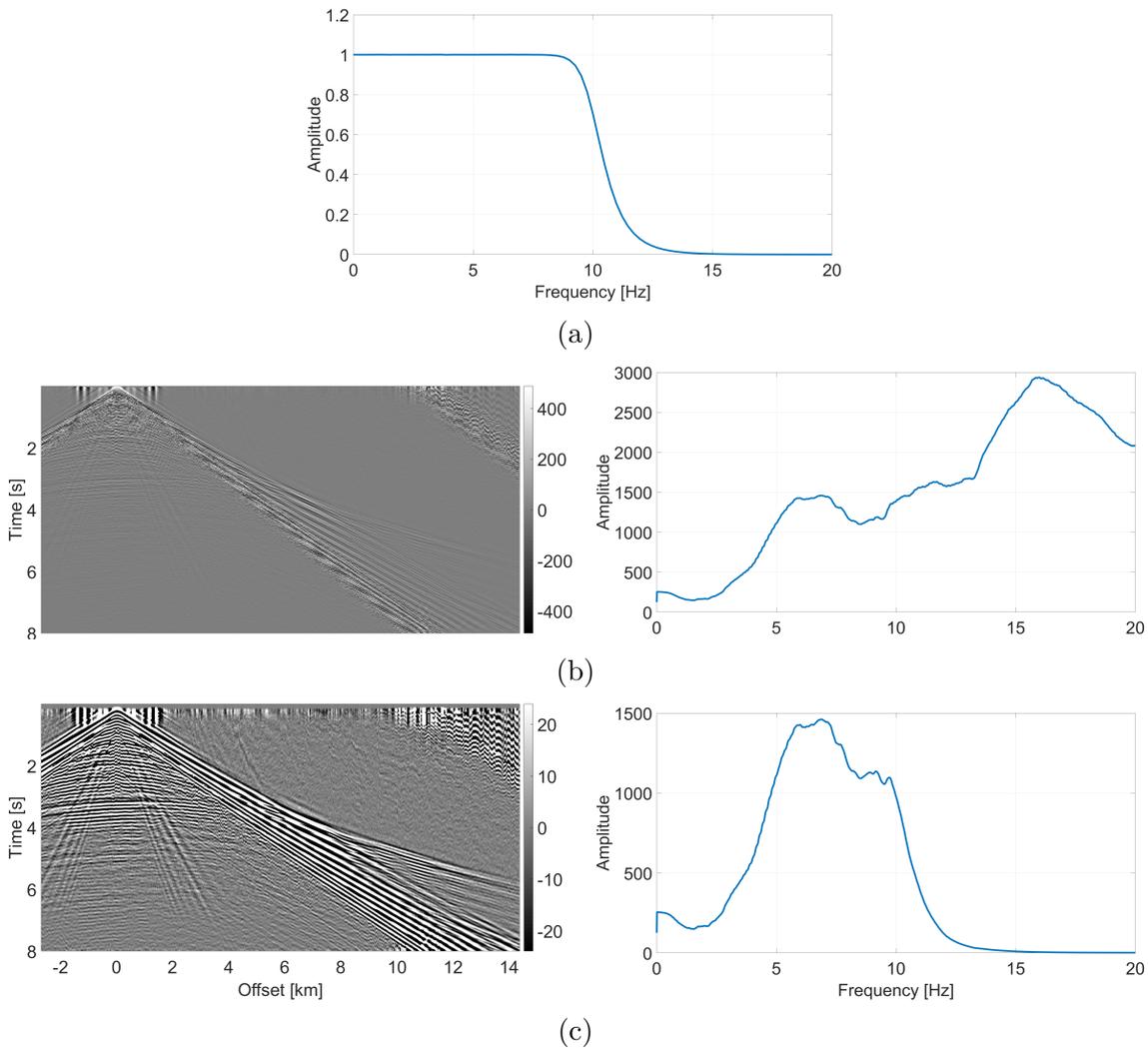


Figure 5.4: High-cut filter application on the western most pressure receiver gather. (a) Response amplitude frequency spectrum of the high-cut filter; (b) Initial data in the time-offset domain and corresponding amplitude frequency spectrum; (c) High-cut filtered data in the time-offset domain and corresponding amplitude frequency spectrum. Extreme values of color scales are divided by 1000 to improve global visibility of events.

2. **A geometrical spreading correction.** Let us consider a configuration where the receivers and the point pressure sources are in a same plane, and where the sub-surface is invariant along the straight lines perpendicular to this plane. This configuration is a necessary configuration when one performs 2D FWI with real data. Indeed, the symmetry of the sub-surface prevents the return in the acquisition plane

of rays that leaved this plane (i.e., the out of plane waves), these rays being non-reproducible with a 2D algorithm. Moreover, let us imagine that we model the data recorded at the receivers location using two different propagators: a 3D one and the corresponding 2D one. We use in both cases the same source function. Because the geometrical spreading is different in 2D and in 3D, the data will not be identical, as illustrated by Auer et al. (2013) (14) (see figure 5.5). Not only the amplitudes are different but also the phases. A consequence of all those observations is that 3D algorithms are more realistic than 2D algorithms.

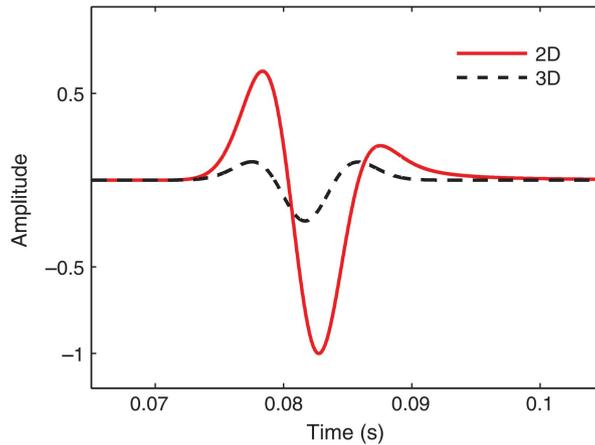


Figure 5.5: “Analytical 2D and 3D time-domain Green’s functions convolved with a 50-Hz dominant-frequency Ricker wavelet, displayed for a single receiver which is 100 m away from a source, both sitting in a homogeneous full space. The 2D wavelet is larger in amplitude, exhibits a long tail, and is phase shifted by $\frac{\pi}{4}$ with respect to the 3D wavelet.” Figure and title from Auer et al. (2013) (14).

Several techniques have been proposed to tackle this problem when one cannot afford applying 3D FWI, among them: (1) Song and Williamson (1995) (179) and Song et al. (1995) (180) used Fourier transforms with respect to the out-of-plane coordinate to reduce the problem of solving the acoustic wave equation in 3D to repeatedly solving a 2D equation. Because of these multiple simulations, this technique remains costly; (2) By assuming a laterally homogeneous medium, Wapenaar (1992) (226) proposed a technique to convert the 3D data to 2D data (or the opposite). This technique has the advantage of being very affordable and adequate for elastic wave equation and for multi-component data. The elastic amplitude variation with offset are then well preserved. Roberts (2005) (157) proposed a similar technique which uses the frequency-horizontal wave number domain instead. This technique was applied by Sears et al. (2010) (165); (3) From an acoustic wave equation, Bleistein (1986) (25) derived a correction that convert the 3D data to 2D data (or the opposite) in asymptotic conditions (i.e., when the source-receiver distance is large relative to the wavelength). This correction needs numerous ray tracings. A less realistic correction derived from Bleistein one’s consists in the convolution of each time domain trace with $1/\sqrt{t}$ followed by a multiplication by \sqrt{t} (in practice, the order does not matter). This correction is valid in a homogeneous sub-surface and when the source time function is small compared to the travel time. The corrected data needs to undergo a global scaling to make the amplitude fits the one of

the corresponding 2D data. This scaling can also be incorporated in the source function. Although very restrictive, this technique has been applied by various authors, even outside of its scope of applicability. For instance Shipp and Singh (2002) (172) applied it for an elastic FWI on hydrophone data coming from a non-homogeneous sub-surface. There is in particular no specific reason for this correction to work on (partial) S-waves arrivals. For the sake of simplicity, this last geometrical spreading correction is applied in this study (figure 5.6a). This correction drastically increases the low-frequency noise, which is removed with the next preparation step. Note that, as expected, the amplitude decay with propagation time is reduced.

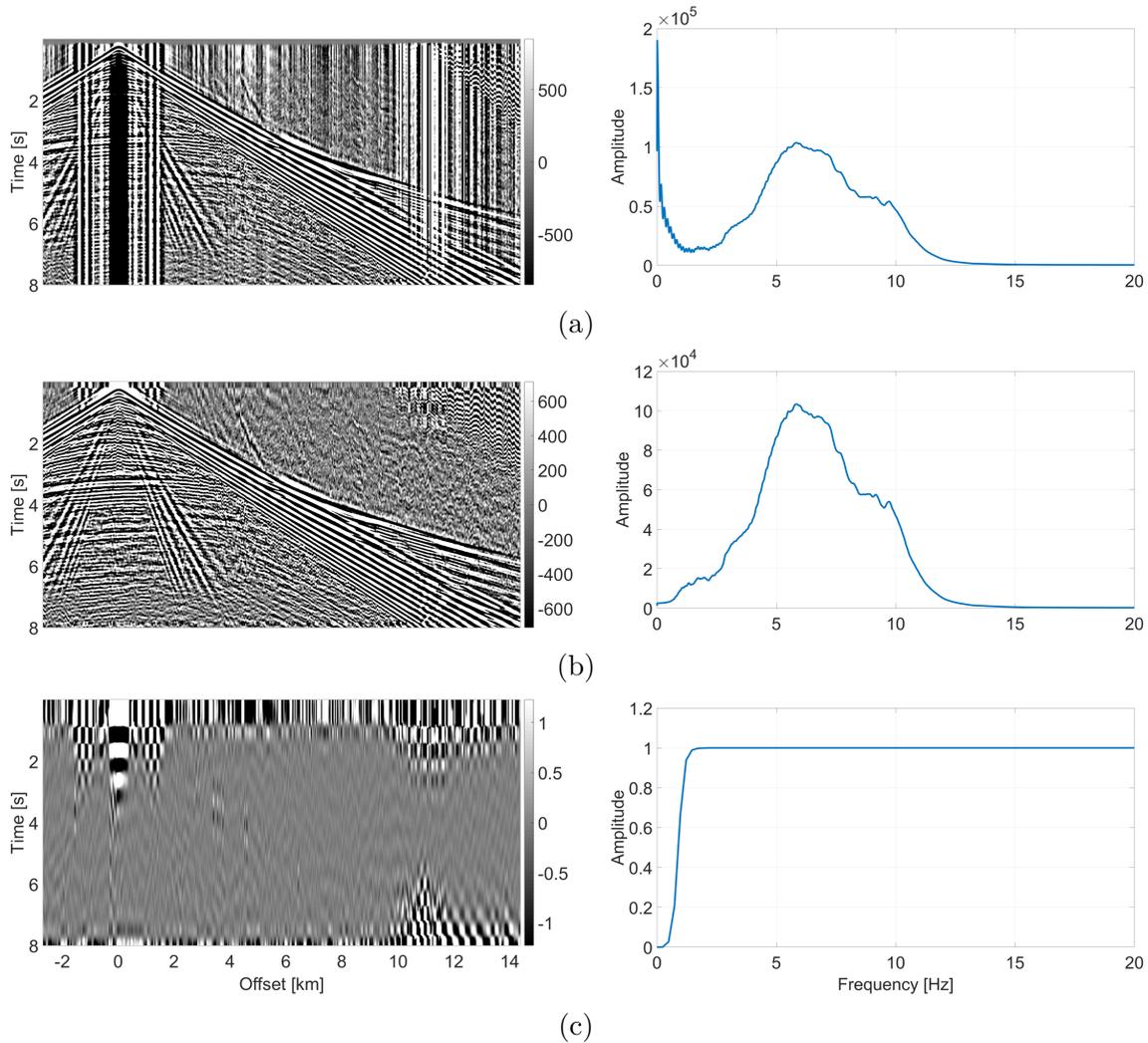


Figure 5.6: Geometrical spreading correction and low-cut filter application on the western most pressure receiver gather. (a) High-cut filtered data after geometrical spreading correction in the time-offset domain and corresponding amplitude frequency spectrum; (b) Data from a after low-cut filter; (c) Difference between the low-cut filtered and unfiltered data when no geometrical spreading correction is applied; (d) Response amplitude frequency spectrum of the used low-cut filter. Extreme values of color scales are divided by 1000 to improve global visibility of events.

3. **A low-cut filter.** To remove the low frequency noise (notably part of the spikes, the exacerbated inverse filter noise, and the geometrical spreading low-frequency

noise), a zero-phase butterworth filter with a cut-off frequency of 1 Hz and an order of 4 is applied (figures 5.6b and 5.6c). This figure also shows the difference between the filtered and unfiltered data when no geometrical spreading correction is applied. One can note that the low frequency noise decreased without removing useful low-frequency signal, critical for successful FWI (e.g., Bunks et al., 1995 (35)).

4. **An FK filter.** In order to remove the Scholte waves (see Section 5.3 for more details), including the wraparound noise coming from remaining Scholte waves, an FK (Frequency-Wavenumber) filter is applied (figure 5.7). The Scholte waves are drastically attenuated. However, the procedure is not perfect since the removal might not be complete and part of the P to S conversions at the seafloor might have been partially removed.

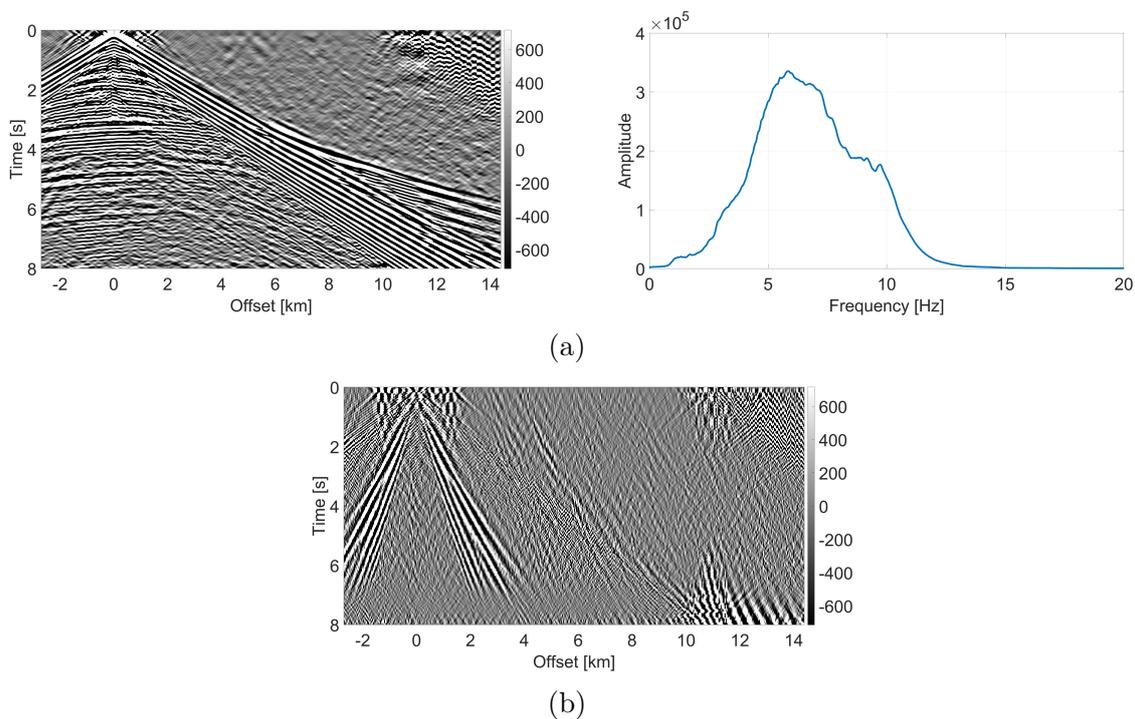


Figure 5.7: FK application on the western most pressure receiver gather. (a) FK filtered data in the time-offset domain and corresponding amplitude frequency spectrum; (b) Difference between the FK filtered and unfiltered data (see figure 5.6b). Extreme values of color scales are divided by 1000 to improve global visibility of events.

5. **A despiking for horizontal velocity component.** Velocity components are impacted by low-frequency spikes. Due to their high amplitudes, these noises are expected to cause problems during inversion (e.g., Crase et al., 1990 (47), who have stronger spikes in terms of time range imprint). Their imprint on the horizontal velocity component is therefore attenuated by manual selection of the major spikes, which are then replaced at low frequencies by interpolated values from neighboring traces. A low-cut filter is applied beforehand to decrease the number of spikes, without hampering the useful low-frequency data. This despiking routine is applied before the FK filter and its results are illustrated in figure 5.8. The despiking routine allows a drastic reduction of the spikes, even if interpolation of stack of spikes is not as successful as desired. More advanced interpolation tools such as structural tensor

based tools are to consider to further improve the despiking routine (e.g. Naeni and Hale, 2015 (134)). For large datasets, more automatic techniques should be used, such as an automatic spike detection using threshold values (as done during the data pre-processing). An ℓ_1 -norm or mixed ℓ_1 -, ℓ_2 -norm instead of an ℓ_2 -norm cost function could also be a possibility to handle these noises during inversion (e.g., Crase et al., 1990 (47)).

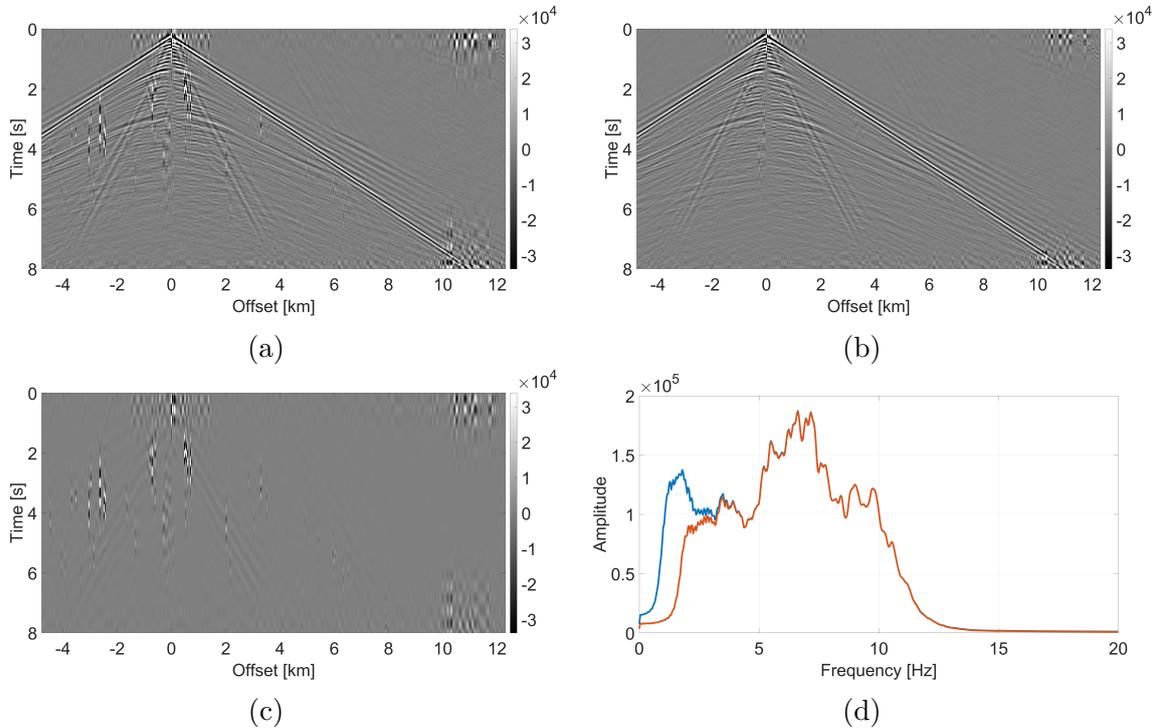


Figure 5.8: Despiking on the seventh western most horizontal velocity component receiver gather. (a) Before despiking; (b) After despiking; (c) a-b; (d) Amplitude frequency spectra before (blue) after (orange) despiking. Extreme values of color scales are divided by 10 to improve global visibility of events.

The remaining noises (such as the water layer guided waves or the short-offset noise) are handled with adequate muting in the time-offset domain during inversion. This muting is performed after the frequency filters embedded in the FWI algorithm, in order to ensure correct results after application of those filters. Note that since the model includes a free-surface, the multiples are kept.

5.2.3 Acquisition geometry

As explained in Chapter 4, the sources and the receivers are not exactly on the profile. The distance between the sources (resp. receivers) and the profile is of 27 (resp. 5) m maximum. Supported by those small distances as well as by the $\sim 2.5D$ geometry of the sub-surface, an orthographic projection of the receivers on the profile and a projection of the sources that conserves the offsets are performed. No data interpolation is performed. The transverse component of velocity is not used and the 2D inversion is performed with the radial velocity components as the horizontal component after adequate polarity changes.

The hydrophone (resp. geophone) depth is set as equal to the depth of the first grid point above (resp. below) the seabed. With the spatial sampling of 20 m, the maximum vertical location error is around 13 (resp. 15) m. With the spatial sampling of 4 m, the maximum vertical location error is around 5 (resp. 7) m.

5.2.4 Source function estimation

The last necessary **FWI** inputs are the source functions. Accurate source functions are necessary for satisfactory inversion results (e.g., Guo et al., 2019 (71)). An airgun-array type of source was used for Alwyn real data acquisition. This type of source is approximated by an isotropic point moment density source (e.g., Arntsen et al., 2000 (8)). The source function can then be divided into two parts: the location and the source time function. The location of the 456 sources of Alwyn real data acquisition is known but the time functions must be estimated. Various estimation techniques coexist. For example, Landrø et al. (2011) (106) showed a few examples of time function modeling based on the parameters of the airgun (array) (e.g., volume and spacing between airguns). The most straightforward approach is to extract (part of) the direct arrivals or seabed multiples and stack them as the source time function. However, because of propagation effects (including reflections at the free-surface or at the seabed), a direct or multiple arrival is not the true source time function. Consequently, the extracted arrivals require an adequate deghosting and/or scaling (e.g., Sears et al., 2010 (165)). Song et al. (1995) (180), during their frequency-domain **FWI** application, used an inversion technique that does not exhibit this problem. The time-domain equivalent of this strategy was selected for Alwyn real data application. In short, the selected technique uses the fact that the pressure data are the combination of the expression at the known receivers location of: (1) the source function of the airgun (array), and (2) the wave propagation through the sub-surface. The observed pressure data are modeled using an arbitrary source time function and the available model of the sub-surface. The expression of the wave propagation through the sub-surface (i.e., point (2)) is extracted from these modeled data. This extraction assumes a correct model of the sub-surface. Consequently, the only remaining unknown is the source time function, which is then estimated via mathematical inversion. The theoretical details of this estimation technique are presented in appendix H, alongside with synthetic applications. A more detailed review and comparison of source estimation and modeling technique would be a welcomed contribution for future applications.

Song et al. (1995) (180) inverted the source time function for each frequency and at each iteration of their **FWI** applications, using the whole data space. Consequently, the source time function includes the effects of model errors and behaves as a garbage collector. Instead, I invert the source time function for all used frequencies only once at the beginning, and using only the most trustable part of modeled data. The trustable part of the data is here the short offset direct arrivals, with a mute applied to the very short offsets due to low frequency noises on the seismic data and inadequate geometrical spreading correction and source relocation at short offset. This source time function inversion is performed on the prepared data. The inversion with non **FK** filtered observed data was also performed and led to similar results (not shown). In the context of inversion with reciprocity as explained in Chapter 2, I assume that all sources have the same source time function. Consequently, a single time function is inverted, as the mean of all individually inverted time functions. The initial source time function contains a wide range of frequency, ensuring the inversion of all frequencies. The inverted source time function

undergoes a high-cut filtering to remove unrealistic high-frequencies, and a tapering to ensure a zero value at time zero. Figure 5.9 shows the inverted and initial source time functions in the time and frequency domains, alongside with an illustration of the short offset direct arrivals selection.

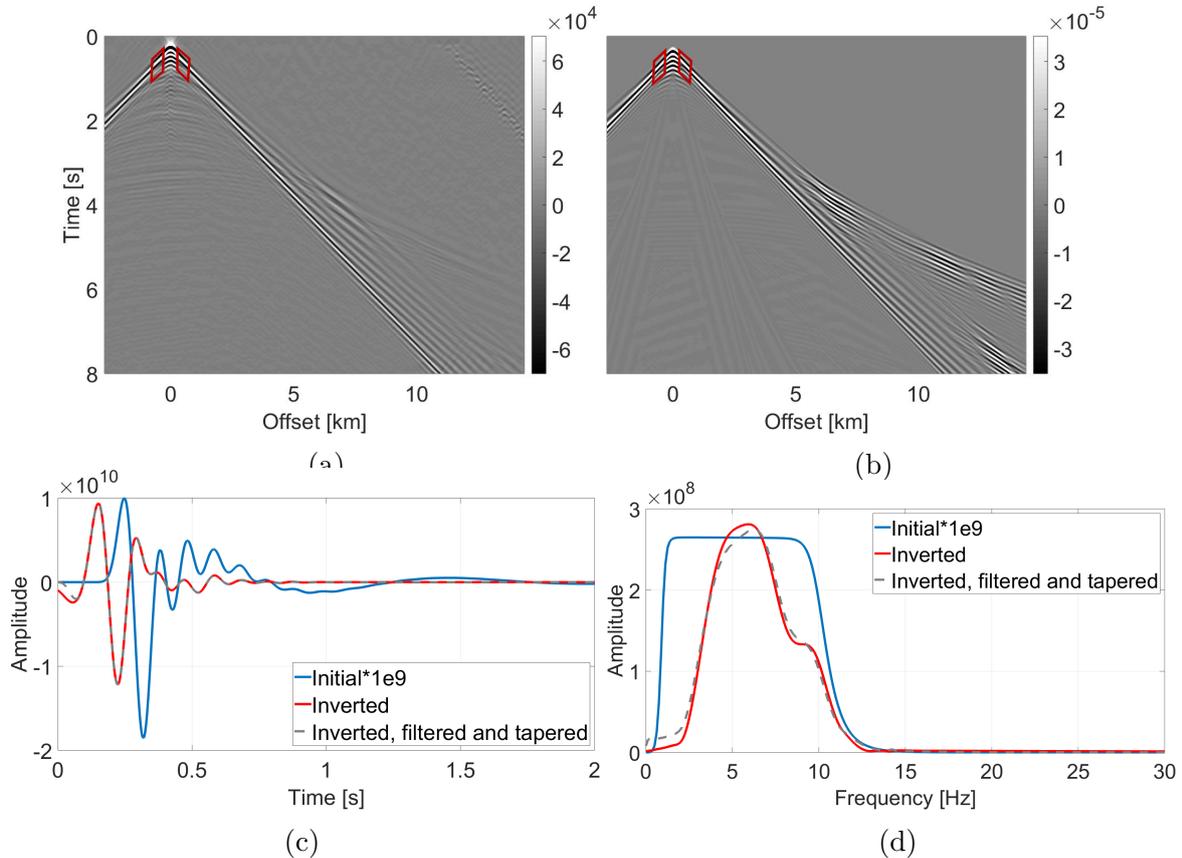


Figure 5.9: Source time function inversion on Alwyn prepared data. (a) and (b) Western most observed and modeled pressure receiver gather respectively, the red polygons delineate the data used for the source inversion; (c) and (d) Initial, inverted and final source time function in the time and frequency domains respectively. Extreme values of color scales are divided by 10 to improve global visibility of events. The initial source time function is scaled for visualization purpose.

The individually inverted time functions show similar general shape (i.e., same time location for highs and lows and same hierarchy between highs and lows), but different amplitudes (figure 5.10). These differences are attributed to the imperfection of modeled data: the seabed and receiver locations might not be perfectly represented for example. Most of the differences can be absorbed by a scaling, as shown for the western most pressure receiver gather. Those observations add confidence in the inverted source time function and its assumed invariability over sources. A receiver gather by receiver gather visual analysis of the coherency over traces of first arrivals confirms that there is no visible differences over sources (not shown).

After scaling of the source time function, this function is able to reproduce the used pressure data, but the corresponding amplitude frequency spectra shows that the fit varies with the frequency range, as illustrated with the western most receiver gather (figure 5.11). The source time function scaling is performed so that the mean of the receiver gather maximum absolute values between modeled and observed data matches, for the

subset of data used for the source estimation. Additional visualizations of modeled and observed used seismic data at different frequency ranges (not provided here) show that these frequency spectrum variations translate mainly into amplitude differences in the time domain, which differ from one frequency range to another. However, a frequency range by frequency range scaling of the source, still does not allow for the amplitude fitting of the wide-aperture pressure data used in the first step of the studied strategies (scaling and comparison not provided here). At each inverted frequency ranges, the maximum absolute value of the computed arrivals with the initial model is roughly two times higher than that of observed data. This high difference cannot be handled by the FWI process and could be due to an imperfect geometrical spreading correction and/or unmodeled intrinsic attenuation for example. In such a case, I decide to adjust the source amplitude at the beginning of each frequency range of each inversion stage, approaching Song et al.'s strategy (1995) (180), which should be examined as an improvement possibility for future applications.

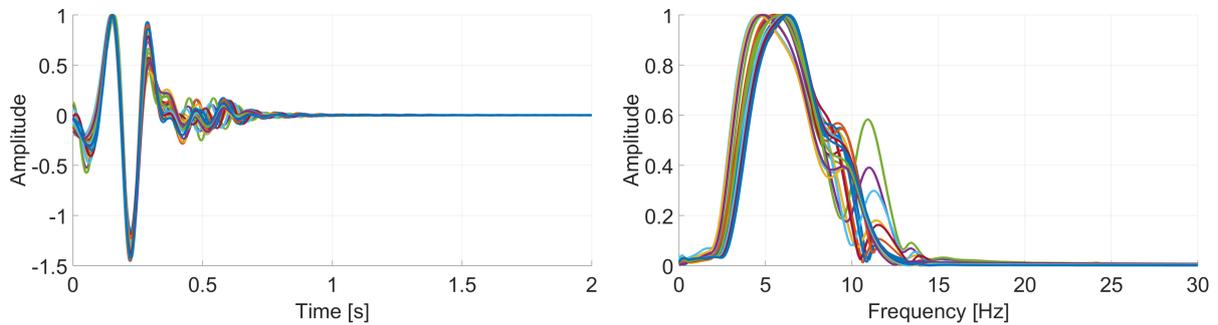


Figure 5.10: Normalized individual source time functions and corresponding amplitude frequency spectra for the western most pressure receiver gather. Before normalization, the difference between the highest and the lowest maximum value in the time domain is of a factor of 15.

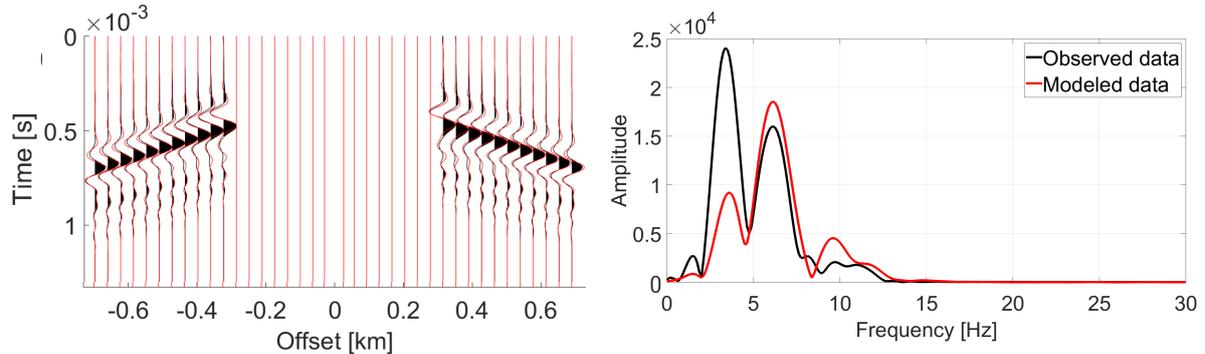


Figure 5.11: Data from the western most pressure receiver gather selected for the source estimation, compared with modeled data using the inverted source function. Time-offset domain and frequency domain representations.

5.3 Application of the variant of the soft seabed strategy

The best inversion results are obtained with the variant of the soft seabed strategy. The two successive stages (i.e., stage 1 and stage 3') of the strategy are applied. Moreover, in order to mitigate the FWI local optima problem, two usual multiscale strategies are

additionally used. First, stage 1 is divided into two substages, 1a and 1b, following the decreasing aperture multiscale strategy of Shipp and Singh (2002) (172). Second, the increasing frequency multiscale strategy of Bunks et al. (1995) (35) is applied at each (sub)stage, with following frequency ranges: [0-3 Hz, 0-5 Hz, 0-7 Hz, 0-9 Hz]. The choice of the maximum frequency is restricted by the available computing resources and lead to a limited maximum achievable resolution (see equation 1.16). To increase the benefit from this strategy, the first frequency range has been selected such as its upper frequency bound is as low as possible. The frequency step between each frequency range has been chosen lower than advised by Sirgue and Pratt (2004) (175). The subspace method of Kennett et al. (1988) (94) for the steplength computation and the Polak-Ribière conjugate-gradient method (Polak and Ribière, 1969 (145)) are used.

This inversion is made possible by the use of reciprocity, as explained in Chapter 2. Moreover, because of its lower spikes content (see Chapter 4), pressure component is preferred over vertical velocity component for stages 1a and 1b. For future applications, the use of the vertical velocity component after a proper denoising is to consider, since this component is expected to provide better inversion results (e.g., Chapter 2 and Sears, 2007 (166)). The noise-impacted short-offset data, the inverse filter noise and the guided-waves are muted in the time-offset domain. The data selection of each stage is also performed with mutes in the time-offset domain, and late arrivals muting after 4.5 s and 6 s are performed on pressure and horizontal velocity components respectively, following the recommendations of Chapter 4. The same mutes are applied on observed and modeled data.

Regarding the model updates, the water layer is not updated and the density is passively updated at each iteration using the same equation 5.1 as the one used to build the initial density model. The anisotropy models are not updated. For stage 3', the P-wave vertical velocities are clipped in order not to become lower than 1400 m/s. The gradient around the receivers is muted within a radius of 440 m (incl. a 140-meter-wide linear taper) and the gradients are horizontally smoothed with a Gaussian horizontal window ($\sigma \sim 287$ m). This smoothing aims at removing the 350-meter-spaced vertical stripes on inverted models (as illustrated on 5.12), due to the imprint of the acquisition geometry. The use of a gradient interpolation tool or a gradient smoothing that follows the structure, via the computation of structural tensors from the vintage migrated image for instance (see Hale, 2009 (73) for instance), should be contemplated in future acquisitions. Indeed the used horizontal smoothing might be destroying details, especially in and around the tilted block.

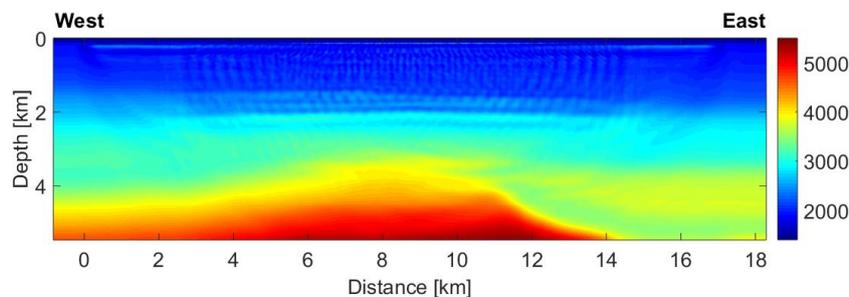


Figure 5.12: Updated P-wave vertical velocity model from the first stage, without gradient smoothing. The strong stripes are interpreted as the imprint of the sparse receiver geometry. The horizontal distance is computed along the profile from the western most source.

In order to save computing resources, the undersampling strategy used by Sears et al. (2010) (165) is applied during stage 1a. This strategy consists in ensuring the stability of modeled P-waves only. Although their modeled S-waves showed instabilities, Sears et al. (2010) (165) provided reasonable inversion results. This undersampling strategy is supported by the fact that stage 1a data contains poor S-waves related events (not shown), and by Levander (1988) (110) who noticed that his modeled P-waves data were not impacted by unstable S-waves modeling. The inversion is consequently performed with finite-difference spatial steps of 20 m and time steps of 1.2 ms, ensuring the stability of modeled P-waves for frequencies up to 14 Hz, in adequation with the maximum frequency of the prepared seismic data. For the other stages however, a fine sampling is used: the inversion is performed with spatial steps of 4 m and time steps of 0.25 ms, ensuring the stability of both P- and S-waves.

The source time function is scaled at the beginning of the first iteration of each frequency range, using the mean over receiver gathers of the ratio between the observed and modeled maximum absolute value, computed on a selection of data. The selection of data attempt to be as independent from the model as possible, but remains questionable. For the pressure data inversion stages (i.e., stages 1a and 1b), the data selection is the wide-aperture pressure data. For the horizontal velocity component inversion stage (i.e., stage 3'), the data selection is a subset of wide-aperture horizontal velocity component data, i.e.: the first arrival between 2 and 4 km of offset.

For stages 1b and 3', a Scholte wave free inversion is performed to avoid any problem due to wrong Scholte waves that could perturb the inversion through their cycle-skipped aspect combined with their high amplitudes. The Scholte waves propagating along the seabed, they offer information about the very shallow sub-surface only, mitigating the loss of information due to a Scholte wave free inversion strategy. This strategy consists in the removal of Scholte waves on both observed and modeled data, including in the cost function, and is in adequation with the adjoint-state method embedded in the FWI algorithm. This removal is performed with an FK filter on observed data and a subtraction of predicted Scholte waves on modeled data. The prediction of the Scholte waves is performed at the beginning of the stage only, those waves remaining highly unchanged during stage 1b due to their high insensitivity to vertical P-wave velocities or densities. During stage 3', their invariance is ensured by using a tapered mute of the shallow sub-surface (the S-wave velocity model in particular). The prediction is performed through a modeling followed by an FK filtering. A better Scholte wave free inversion should be considered for future inversions, with notably a mute applied on S-wave vertical velocity gradient only, a precise estimation of the Scholte waves penetration depth to limit the impact of the gradient mute, or the update of the predicted Scholte waves at each iteration. Ideally the use of Scholte waves (together with water layer guided waves) in inversion as described in Section 4.3.4 would be the best scenario.

The application of such variant strategy exhibits the required global decrease of cost functions over iterations (figure 5.13). For stages 1a, 1b and the first frequency range of stage 3', the cost function shows a stabilization, in accordance with the updated models stabilization. The ups and downs of the cost functions from stage 1a are mostly due to the passive density update (figure 5.13a). Gradient smoothing might also be playing a secondary role. The ups and downs of the cost functions from the next inversion stages are assumed to have the same causes. From the second frequency range of stage 3', the update is stopped before stabilization. The choice of the number of iterations is arbitrarily chosen, and a lower number of iterations could have allowed for better results

for instance. Indeed, after a few beneficial iterations, the inversion leads to unreasonably strong update with improvement of amplitude data fit only, and finally to an increase of the cost function. The impact of this phenomenon seems to increase with the frequency range (not shown). These unreasonable updates are assumed to come from data overfitting of wrongly modeled amplitudes, combined with the abnormal reaching of clipping values. Reaching model clipping values might also be a cause for ups and downs in cost functions.

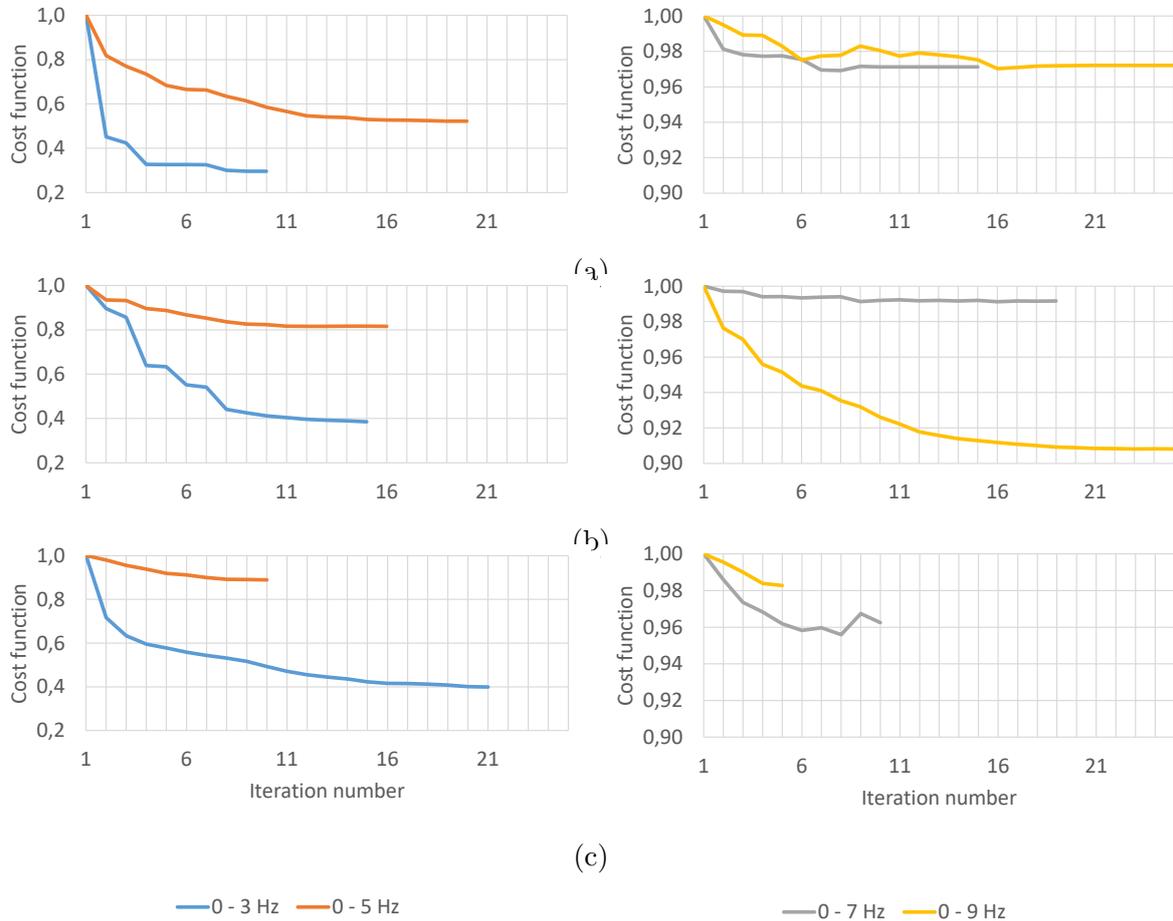


Figure 5.13: Variant strategy. Normalized cost function for each frequency range of each stage. (a) Stage 1a; (b) Stage 1b; (c) Stage 3'. The normalization is performed for visualization purpose, and the non-normalized cost functions are provided in [Appendix I](#).

The initial data fit of the first frequency range from each stage is promising (figure 5.14). Indeed it exhibits an initial prediction of the data with an error within half the period for the stronger arrivals, which could prohibit cycle skipping for all arrivals through natural layer stripping (Wang et al., 2020 (221)). Those good initial fits support the choice of source time function. The initial slight travel times misfit at zero-offset of stage 1b, especially the one from shallow reflectors, questions the correctness of the anisotropy model. Unsurprisingly, the obtained data fit is satisfying, with a drastic improvement during inversion, whether it be the travel time, the phase or, to a lesser extent, the amplitude, without obvious cycle-skipping (figure 5.15). Note that stage 1b uplift consists mainly in the building of high frequency data (figure 5.15b). The disappearance of high frequencies at late times and/or long offset of the observed data compared to the final modeled data from each stage questions the correctness of the non attenuating modeling.

Those encouraging modeled data improvements translate into exciting final P- and

S-wave vertical velocity models, with an increasing resolution and depth penetration over stages, as expected from the diving wave, P-P and P-S reflections penetration depth analysis (Chapter 4) as well as the theoretical resolution (1.16): P-P and P-S reflections penetration depth which goes up to the deepest targets is higher than that of diving waves, and the resolution of the updates increases with decreasing aperture. To illustrate those aspects, stage by stage model inversion results displays and well comparisons are available in Appendix I. Even the targets are illuminated (i.e., the reservoirs, sitting between 3 and 5 kilometer depths in the tilted block in the central part of the model). Most of the main structures of the vintage migrated image and the well data emerged, and the increased resolution enables the interpretation of the tilted block up to its main faults (figures 5.16 to 5.18, as well as Appendix I). Note that the well comparison is limited by the absence of shallow S-wave velocity well data, the deviated aspect of the wells in an anisotropic environment, the difference of resolution between well data and the models and the distance between the wells and the model. The effect of this distance (maximum 2.5 km) is limited by the ~ 2.5 subsurface geometry (see Chapter 4). Part of the over/underestimations of the vertical velocities could be a compensation of under/overestimation of velocities from the anisotropy parameters, as observed in Chapter 3, especially near the boundaries of the stronger anisotropic layers. Note that the anisotropy parameters model have been smoothed during the preparation step. It would be worth trying performing the inversion with less smoothed initial models. Another possible explanation could be an amplitude discrepancy between modeled and observed data (from unmodeled intrinsic attenuation, wrong density model or geometrical spreading correction correction or bad geophone-seabed coupling for instance). The variation of the ratio between P- and S-wave vertical velocities are also convincing, with an increase with depth and ratios around the reservoirs in agreement with lithology (e.g., Schön, 2015 (163)). Indeed, the ratios of the Dunlin group (mainly mudstones), are higher than the neighboring Brent and Statfjord groups (mainly sandstones), lithology from Harker et al., 2003 (77). Moreover, the ratios inside the tilted block range between ~ 1.51 and ~ 2.17 . The abnormal extreme ratio values locally observable can be due to the above-mentioned under/overestimated velocities as well as differences of resolution between both velocity models.

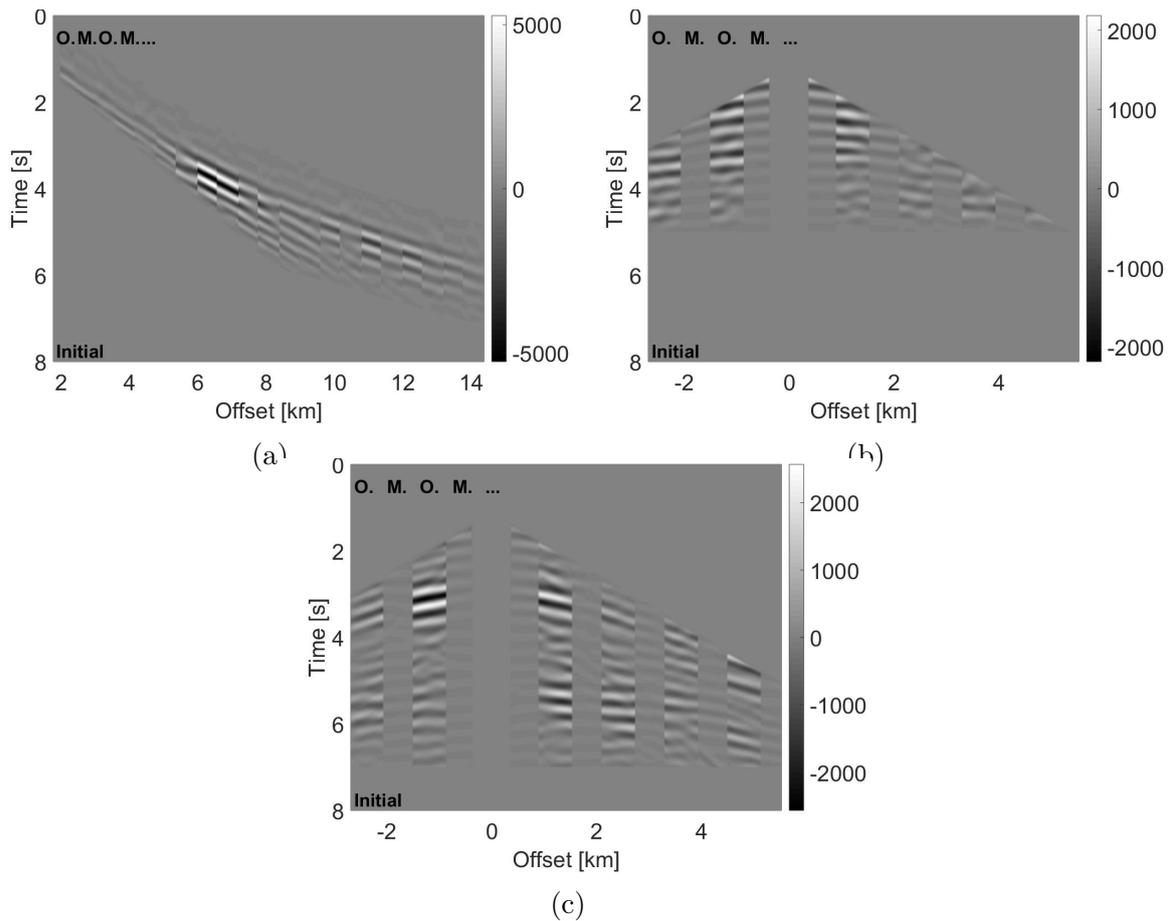


Figure 5.14: Variant strategy. Observed (O.) and initial modeled (M.) data represented in an alternate manner, for the first frequency range of each stage. (a) Stage 1a (pressure component); (b) Stage 1b (pressure component); (c) Stage 3' (horizontal velocity component). Western most receiver gather. The corresponding trace normalized wiggle plots are provided in [Appendix I](#).

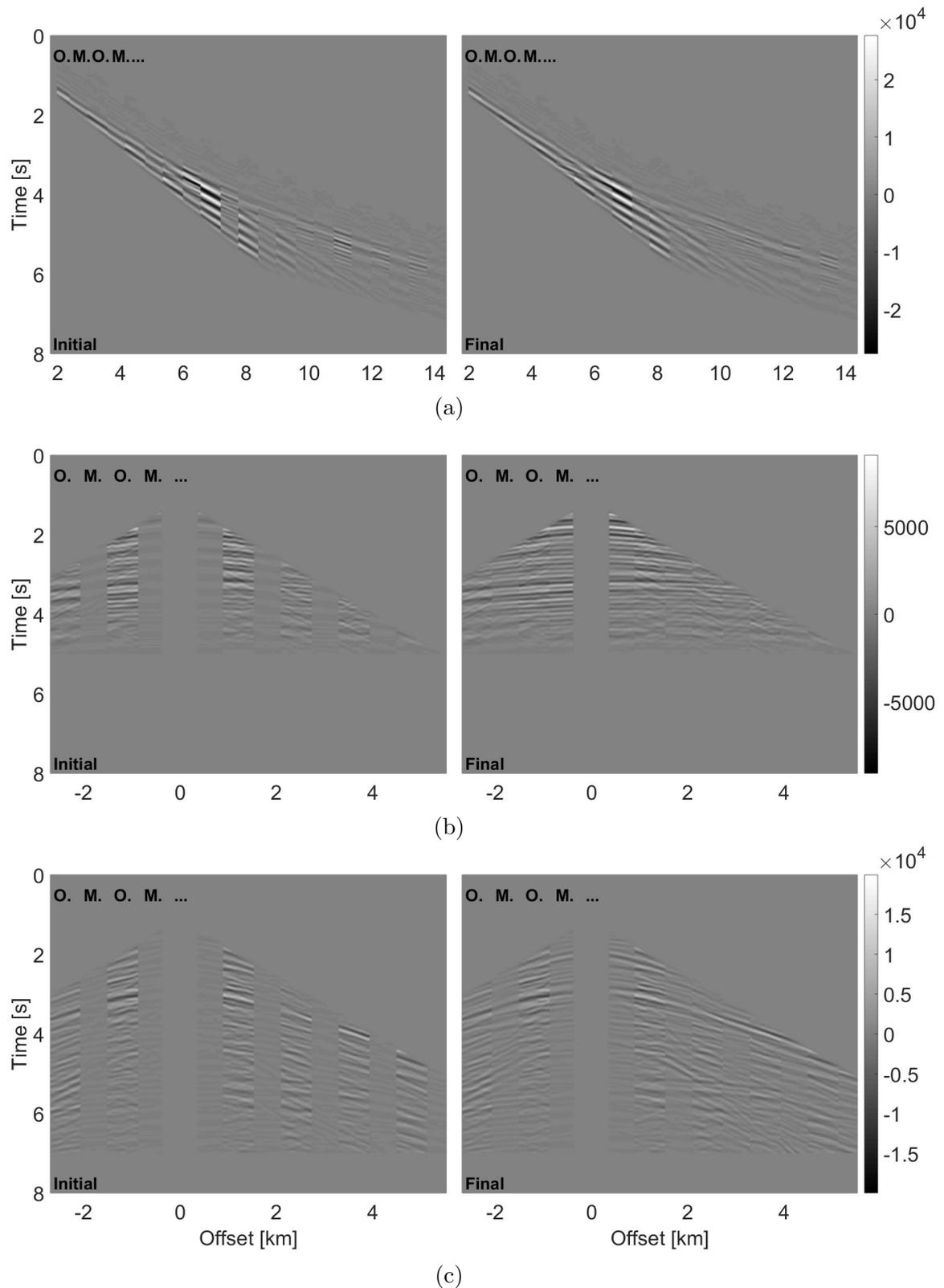


Figure 5.15: Variant strategy. Observed (O.) and modeled (M.) data represented in an alternate manner, at the beginning and end of each stage. (a) Stage 1a (pressure component); (b) Stage 1b (pressure component); (c) Stage 3' (horizontal velocity component). Western most receiver gather. The corresponding trace normalized wiggle and residual plots are provided in [Appendix I](#).

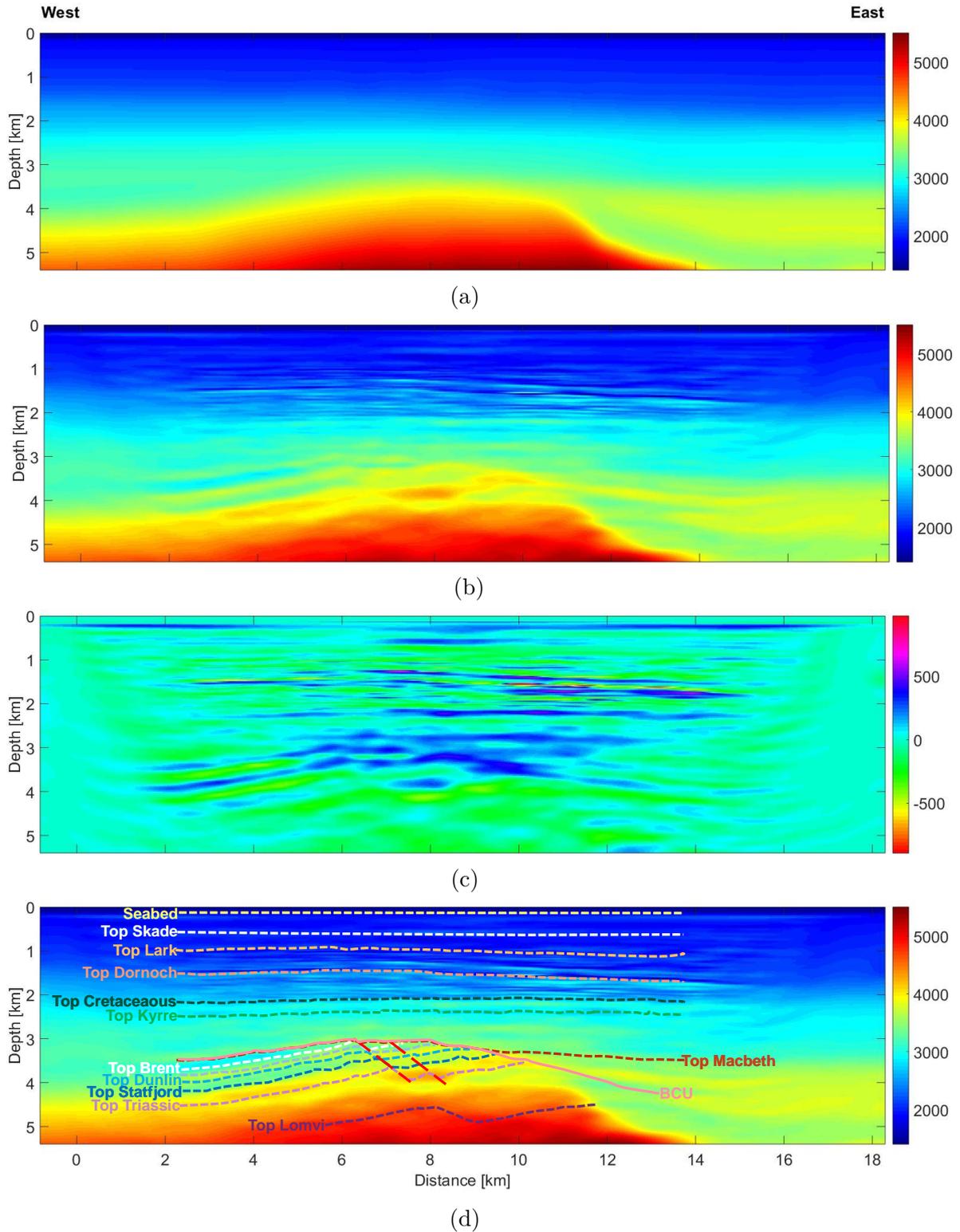


Figure 5.16: Variant strategy. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above b. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The horizontal distance is computed along the profile from the western most source.

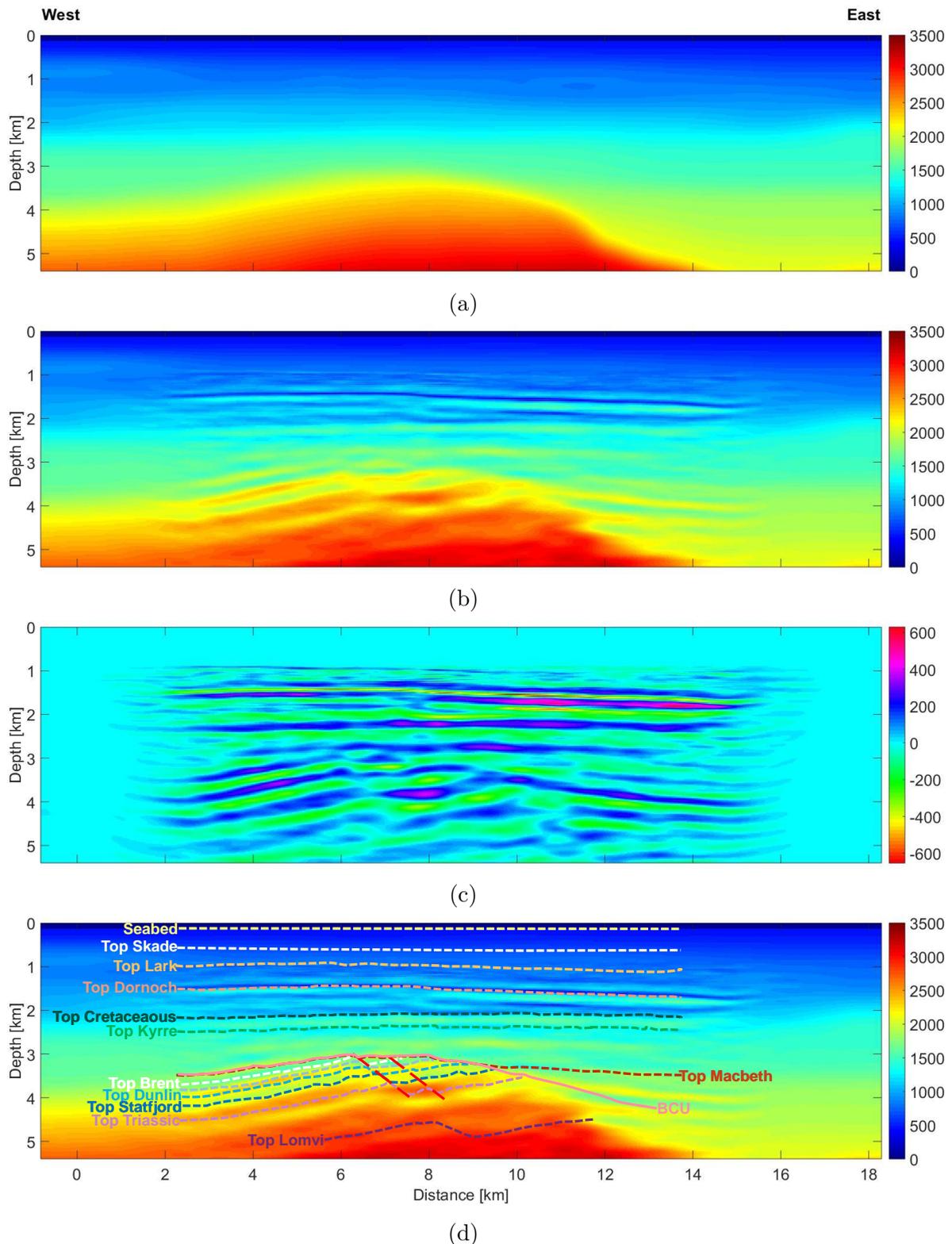


Figure 5.17: Variant strategy. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above b. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The horizontal distance is computed along the profile from the western most source.

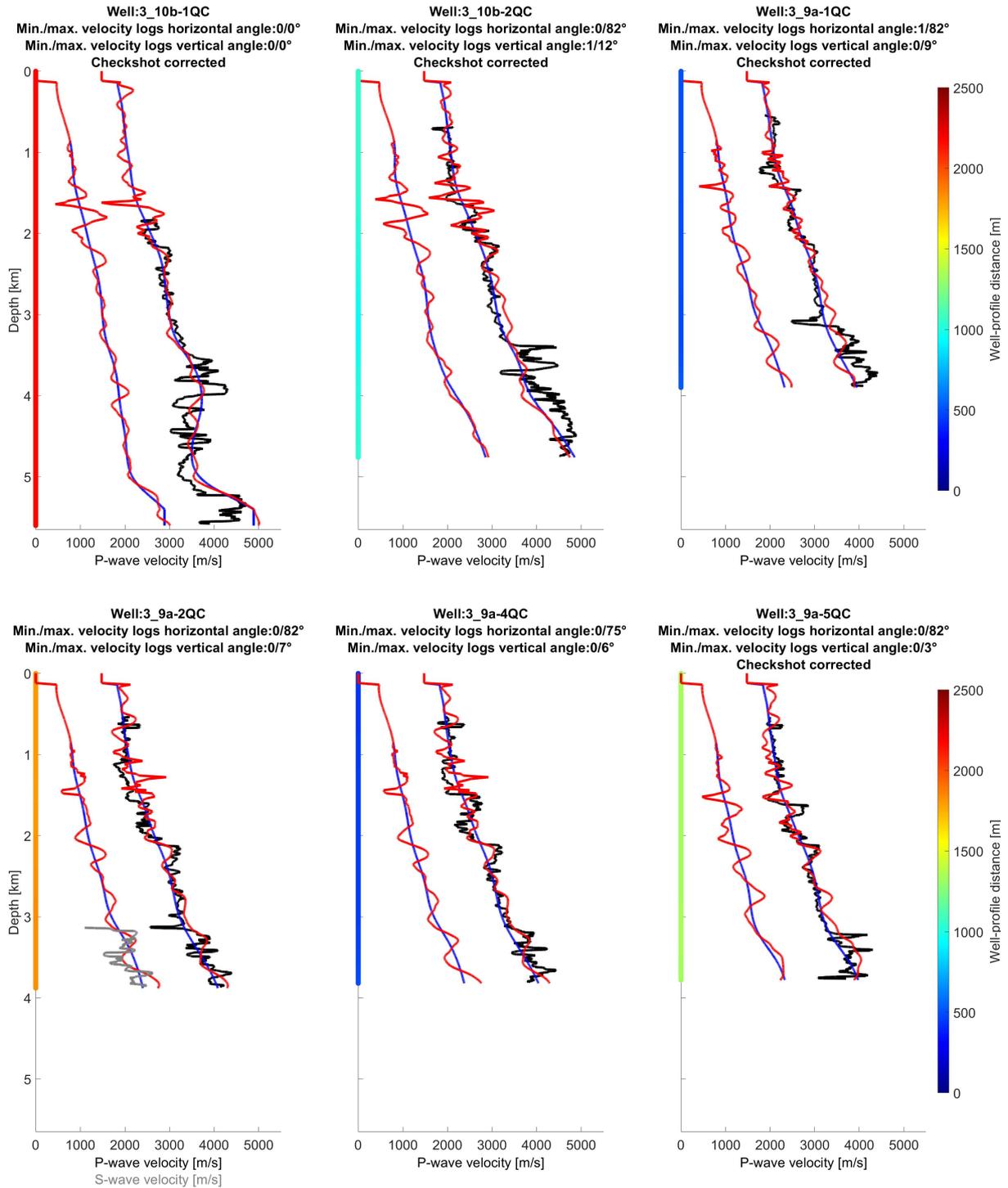


Figure 5.18: Variant strategy. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

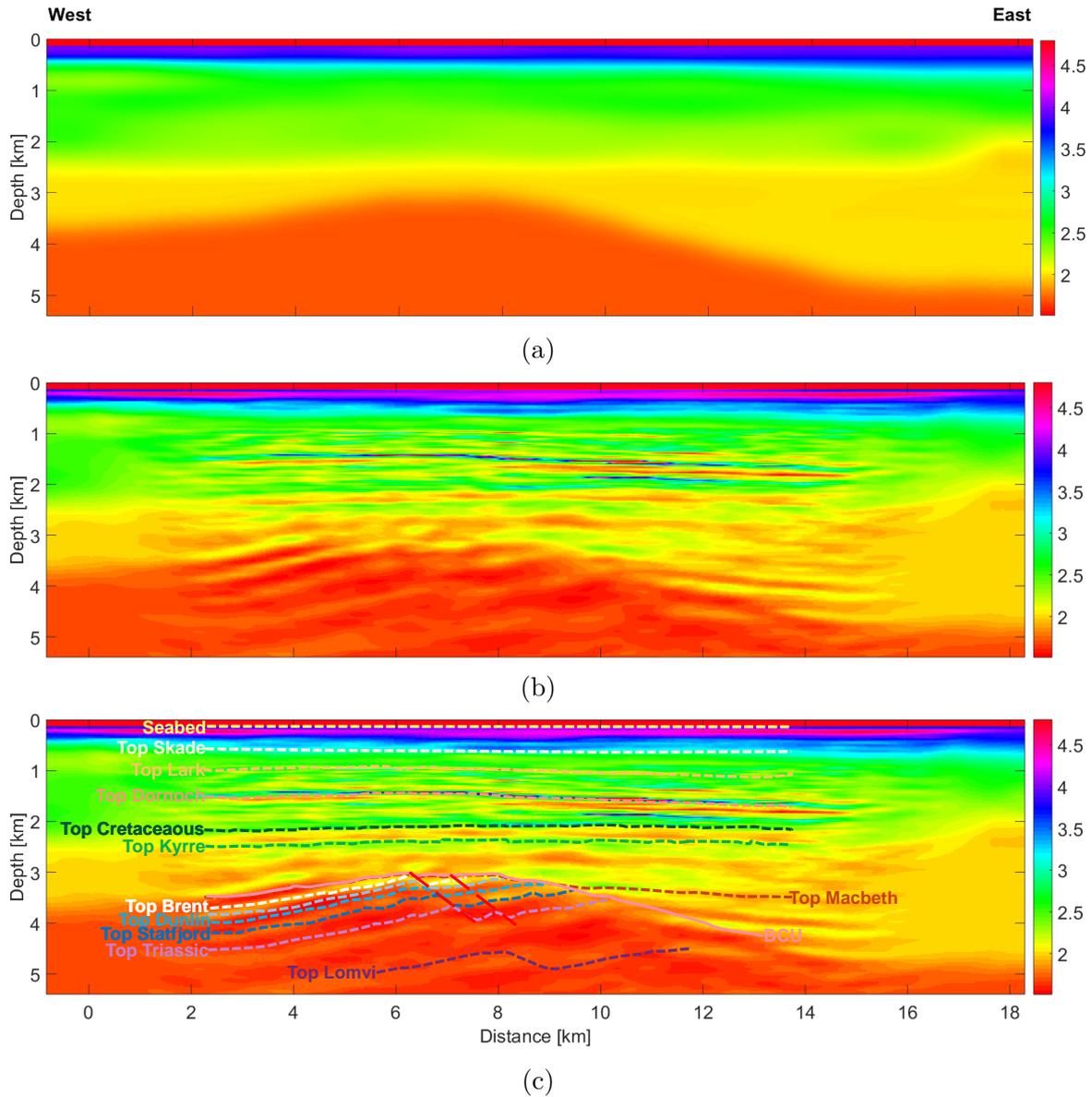


Figure 5.19: Variant strategy. (a) Initial P- over S-wave vertical velocity model, for comparison purpose; (b) P- over S-wave vertical velocity model; (c) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above b. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The color scale is clipped above 4.8 to increase visibility of reasonable ratios, maximum ratio is ~ 10.85 . The horizontal distance is computed along the profile from the western most source.

5.4 Application of the original soft seabed strategy

The original inversion strategy is not successful as successful as the variant strategy. As for the variant strategy, the first stage is divided into two sub-stages: stages 1a and 1b. The four successive stages (i.e., stages 1a, 1b, 2 and 3) of the strategy are here applied. Except for model clipping, the same inversion choices as for the variant strategy are selected. Notably, the undersampling strategy is applied during the inversion stages with

wide-aperture P-wave data (stages 1a and 2), and a fine sampling is applied during the other stages (i.e., stages 1b and 3). Similarly, the source scaling for the pressure data inversion stages (i.e., stages 1a, 1b and 2) uses the wide-aperture pressure data. For the horizontal velocity component inversion stage (i.e., stage 3), the data selection is the same subset of wide-aperture horizontal velocity component data as previously. Stages 1a and 1b being shared with the variant strategy, they are not recalled in this section. During stage 2, the S-wave vertical velocities are clipped in order not to increase above P-wave vertical velocities or to decrease below 200 m/s. Only the first frequency range of the last stage (i.e., stage 3) is performed for the hereafter-mentioned reasons.

Such FWI application also exhibits the required global decrease and stabilization of cost functions over iterations (figure 5.20), in accordance with the updated models stabilization.

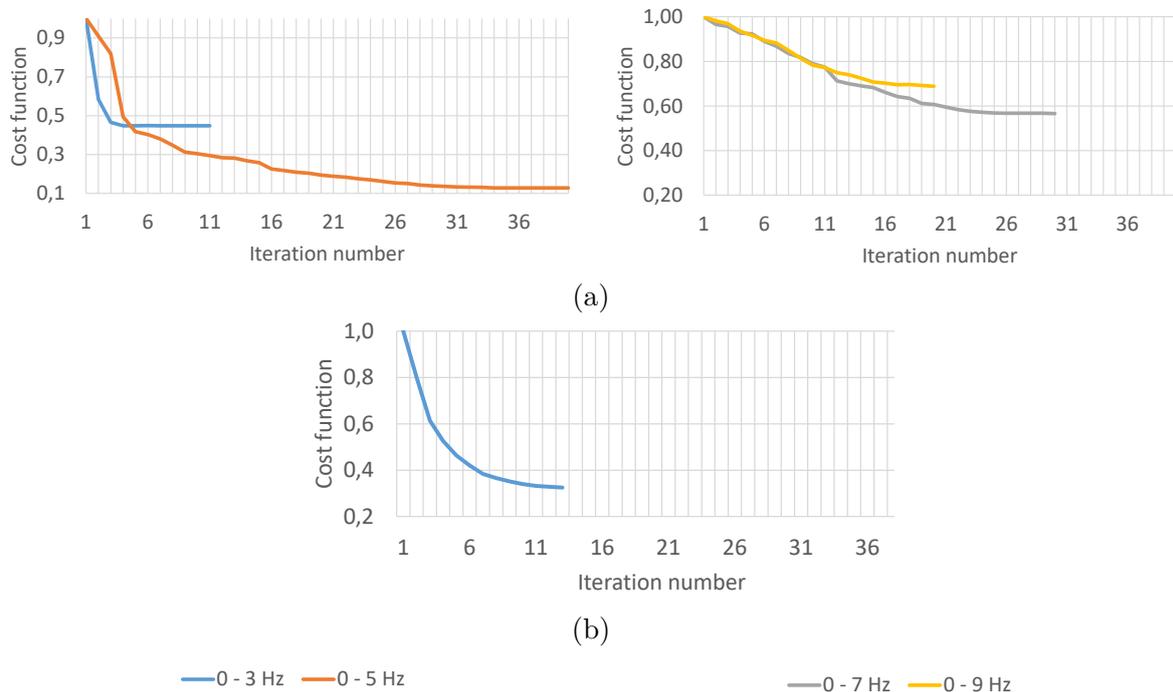


Figure 5.20: Original strategy. Normalized cost function for each frequency range of each stage. (a) Stage 2; (b) Stage 3. The normalization is performed for visualization purpose, and the non-normalized cost functions are provided in [Appendix J](#).

The initial data fit of stage 2 indicates that the wide-aperture P-wave data time fit has been slightly damaged by the previous narrow-aperture P-wave data inversion stage (i.e., stage 1b), questioning again the correctness of the anisotropy model (compare the initial status of figure 5.21a with the final status of figure 5.15a). This initial fit allowed for the drastic improvement of the general data fit during the stage (figure 5.21a). The initial data fit of the first frequency range of stage 3 is poor compared with the corresponding one of the variant strategy with errors above half the period for the stronger arrivals (compare the initial status of figure 5.21b with figure 5.14c), hinting an unsuccessful stage 2. Consequently, even though the data fit greatly improved during stage 3, whether it be the travel time, the phase or, to a lesser extent, the amplitude, there is a high probability that the data are cycle-skipped.

Unsurprisingly, those disputable data fit improvements translate into unconvincing final models. During stage 2, the changes brought the P-wave vertical velocity model are

limited (figures 5.22a and 5.22b), as expected from synthetic tests (Chapter 3). However, the substantial updates brought to the S-wave vertical velocity model do not show the expected lateral continuity and the main structures of the migrated image (figures 5.23a and 5.23b), even though part of those structures can still be guessed. This failing is observed from the first frequency range inversion, and increases with frequencies (not shown). As expected, it translates into detrimental updates of the P-wave vertical velocities during the following stage 3, with a strong loss of lateral continuity and interfaces destruction (figure 5.22c). The S-wave vertical velocity model slightly improved with an increased continuity of the structures (figure 5.23c). Due to the failing of the first frequency range of this last stage, the inversions with higher frequencies is not performed. Indeed, from past experience, the higher frequency ranges are not expected to overcome this failing. Additional stage by stage model inversion results and well comparisons are available in Appendix J.

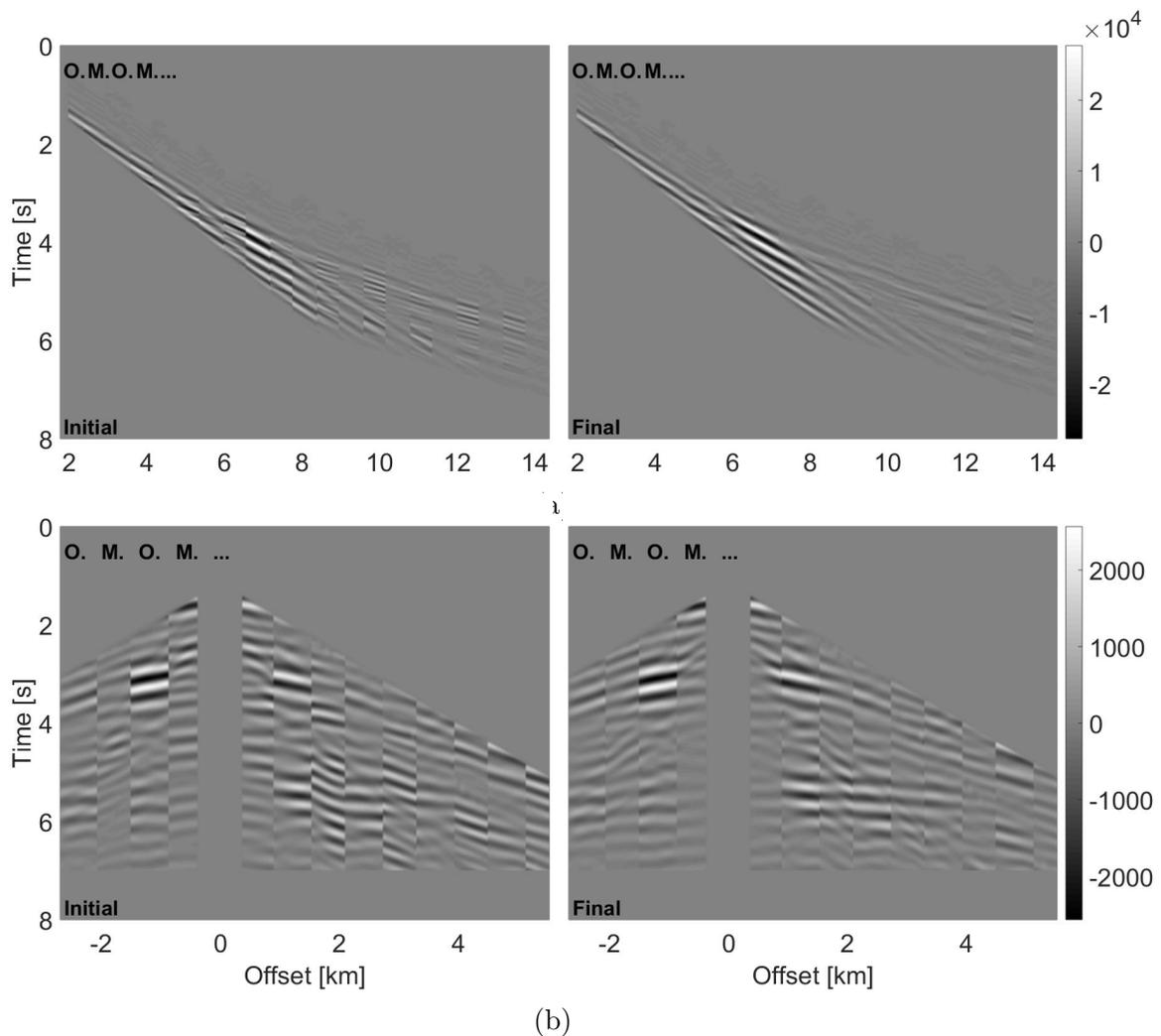


Figure 5.21: Original strategy. Observed (O.) and modeled (M.) data fit represented in an alternate manner, at the beginning and end of each stage. (a) Stage 2 (pressure component); (b) Stage 3, first frequency range (horizontal velocity component, first frequency range only). Western most receiver gather. The corresponding trace normalized wiggle and residual plots are provided in Appendix J.

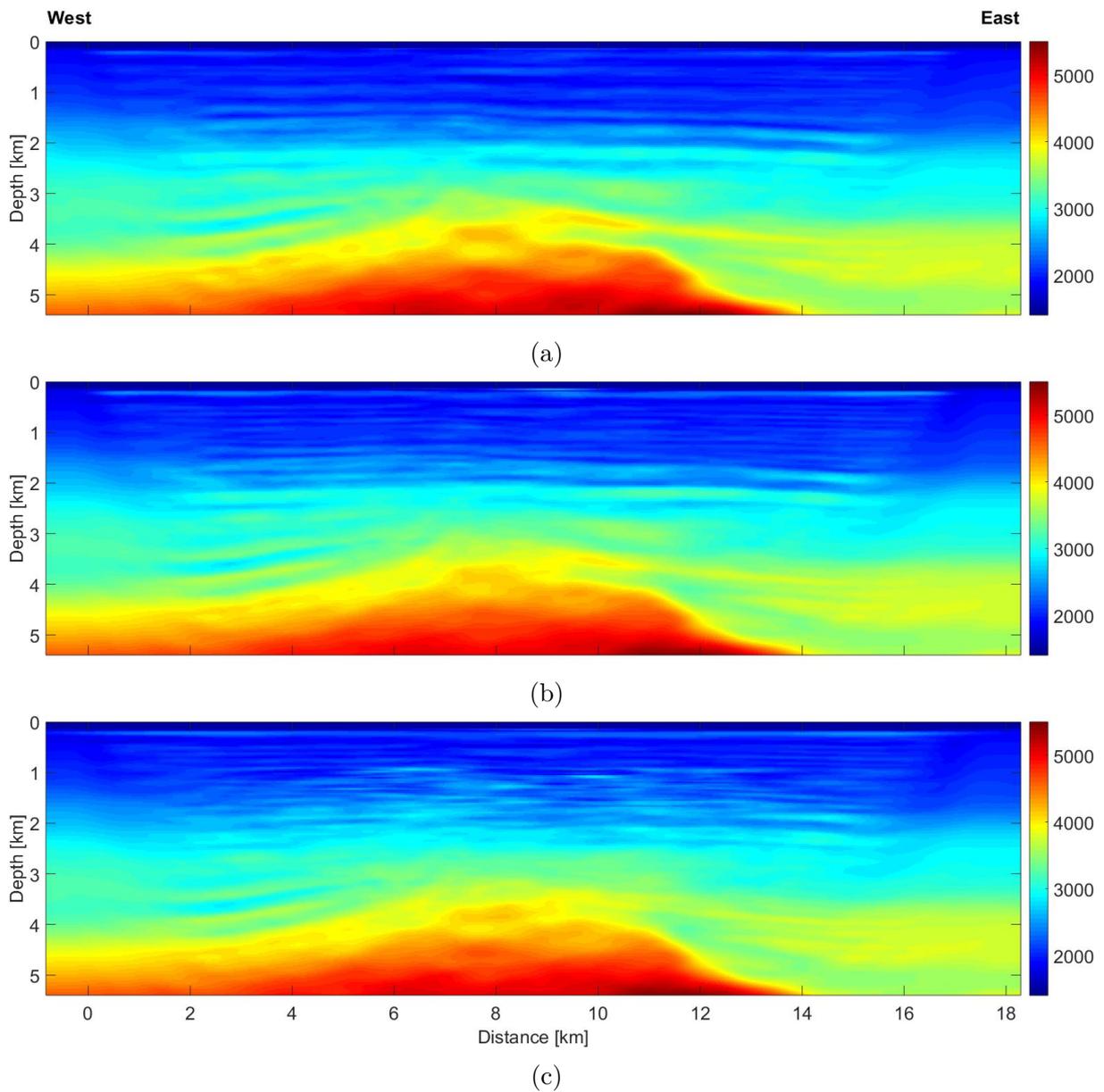


Figure 5.22: Original strategy. P-wave vertical velocity model inversion results. (a) Stage 2 initial model, provided for comparison purpose; (b) Stage 2 final model; (c) Stage 3 final model. The horizontal distance is computed along the profile from the western most source.

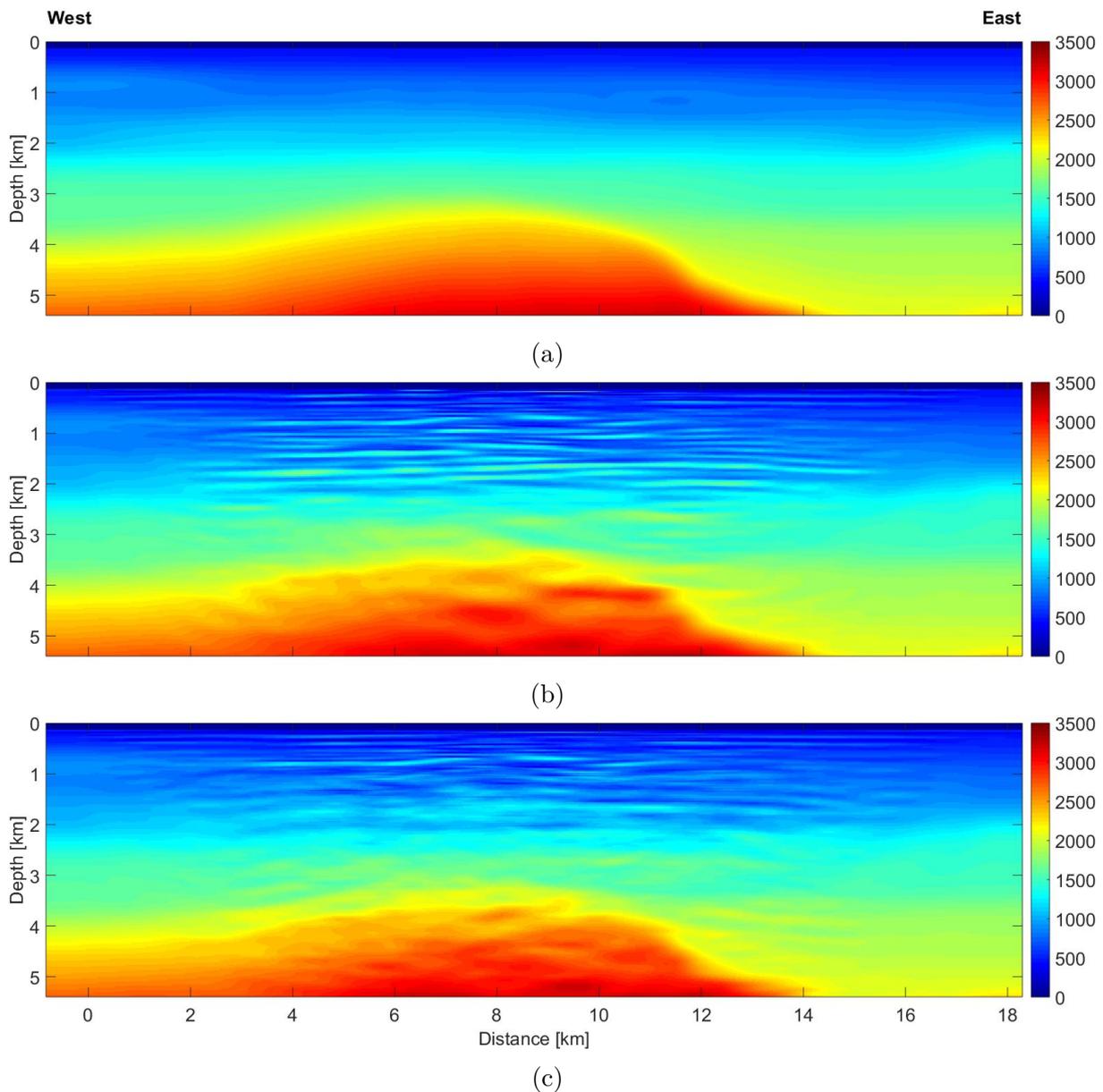


Figure 5.23: Original strategy. S-wave vertical velocity model inversion results. (a) Stage 2 initial model, provided for comparison purpose; (b) Stage 2 final model; (c) Stage 3 final model. The horizontal distance is computed along the profile from the western most source.

5.5 Original strategy with a modified stage order

As explained in [Section 5.3](#), the inversion with narrow-aperture data (stage 1b) damaged the wide-aperture travel time data fit, on which stage 2 is expected to rely ([Chapter 3](#)). Consequently, I decide to invert for the intermediate scale S-wave vertical velocity features (stage 2) before inverting for small scale P-wave vertical velocity features (stage 1b). Those stages are hereafter dubbed stage 2' and stage 1b' respectively. The S-wave vertical velocities are clipped to not be lower than 200 m/s. Only stage 2' and the first frequency range of the stage 1b' inversions are performed for the below-mentioned reasons.

Again, such [FWI](#) application exhibits the required global decrease and stabilization

of cost functions over iterations (figure 5.24), in accordance with the updated models stabilization.

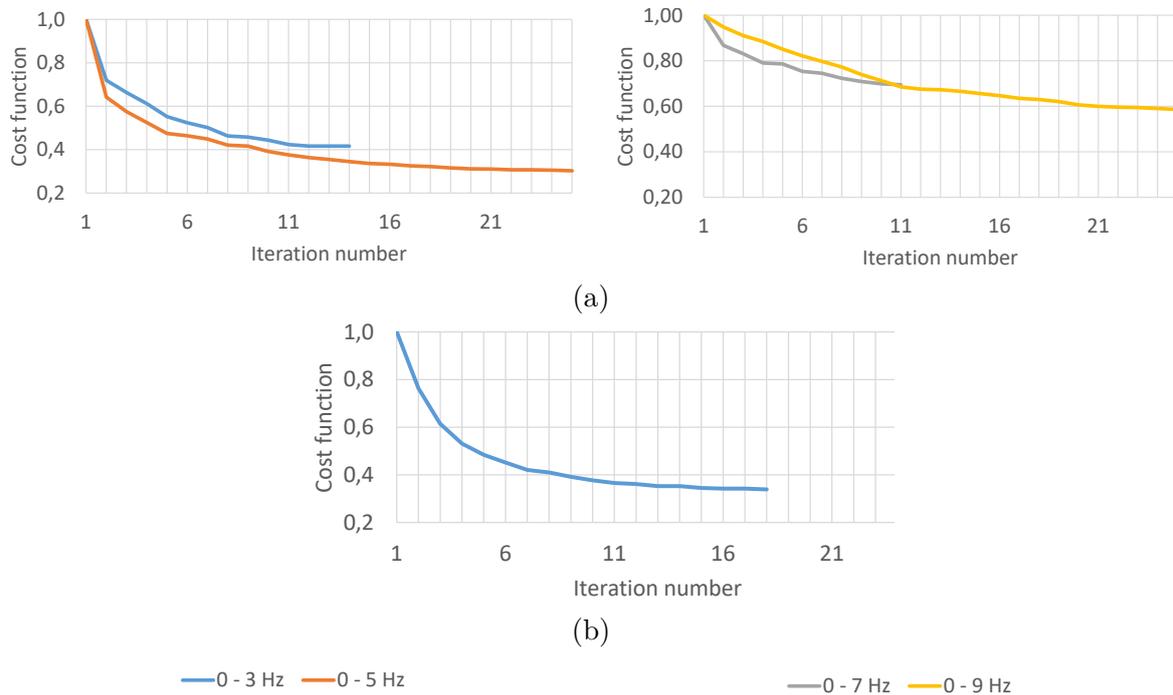


Figure 5.24: Original strategy with a modified stage order. Normalized cost function for each frequency range of each stage. (a) Stage 2'; (b) Stage 1b'. The normalization is performed for visualization purpose, and the non-normalized cost functions are provided in [Appendix K](#)

The data fit improved drastically during stage 2' (figure 5.25a), and mainly consists in an amplitude improvement, in agreement with the initial misfit being mainly due to amplitude disagreements. Aside from amplitudes, stage 1b' initial data fit is less good than with the original order (compare the initial status of figure 5.25b with figure 5.14b). This might be a hint from a failing stage 2' inversion. Stage 1b' inversion leads to an improved data fit, similar to the one of the original order inversion (compare the final status of figure 5.25b with the final status of figure 5.15b).

This improved data fit behavior, compared with the original order, helps producing better models. Those models however, remain disappointing. Indeed, if the P-wave vertical velocity model is only very slightly modified during stage 2' (figures 5.26a and 5.26b), in agreement with the synthetic application ([Chapter 3](#)), the meanwhile final S-wave vertical velocity model has more laterally continuous structures compared with the original order inversion, but it does not exhibit the expected structures (figures 5.27a and 5.27b). As a consequence, even though the first frequency range of stage 1b' brings structures in agreement with the migrated section in the deeper parts of the P-wave vertical velocity model, it strongly damages the lateral continuity of the shallower structures. Consequently, the next frequency range inversions are not performed. Additional stage by stage model inversion results and well comparisons are available in [Appendix K](#).

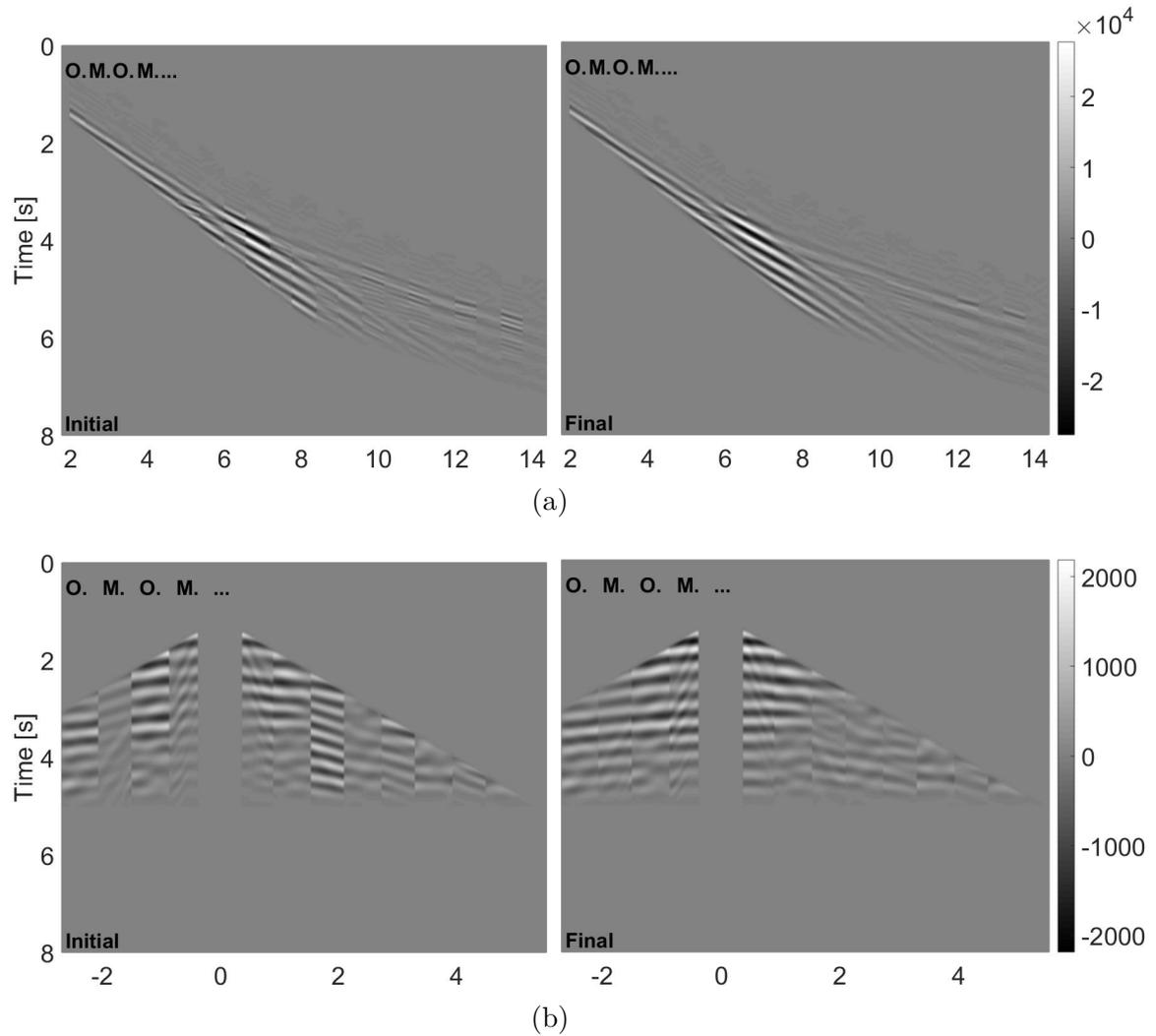


Figure 5.25: Original strategy with a modified stage order. Observed (O.) and modeled (M.) data fit represented in an alternate manner, at the beginning and end of each stage. (a) Stage 2' (pressure component); (b) Stage 3', first frequency range (horizontal velocity component, first frequency range only). Western most receiver gather. The corresponding trace normalized wiggle and residual plots are provided in [Appendix K](#).

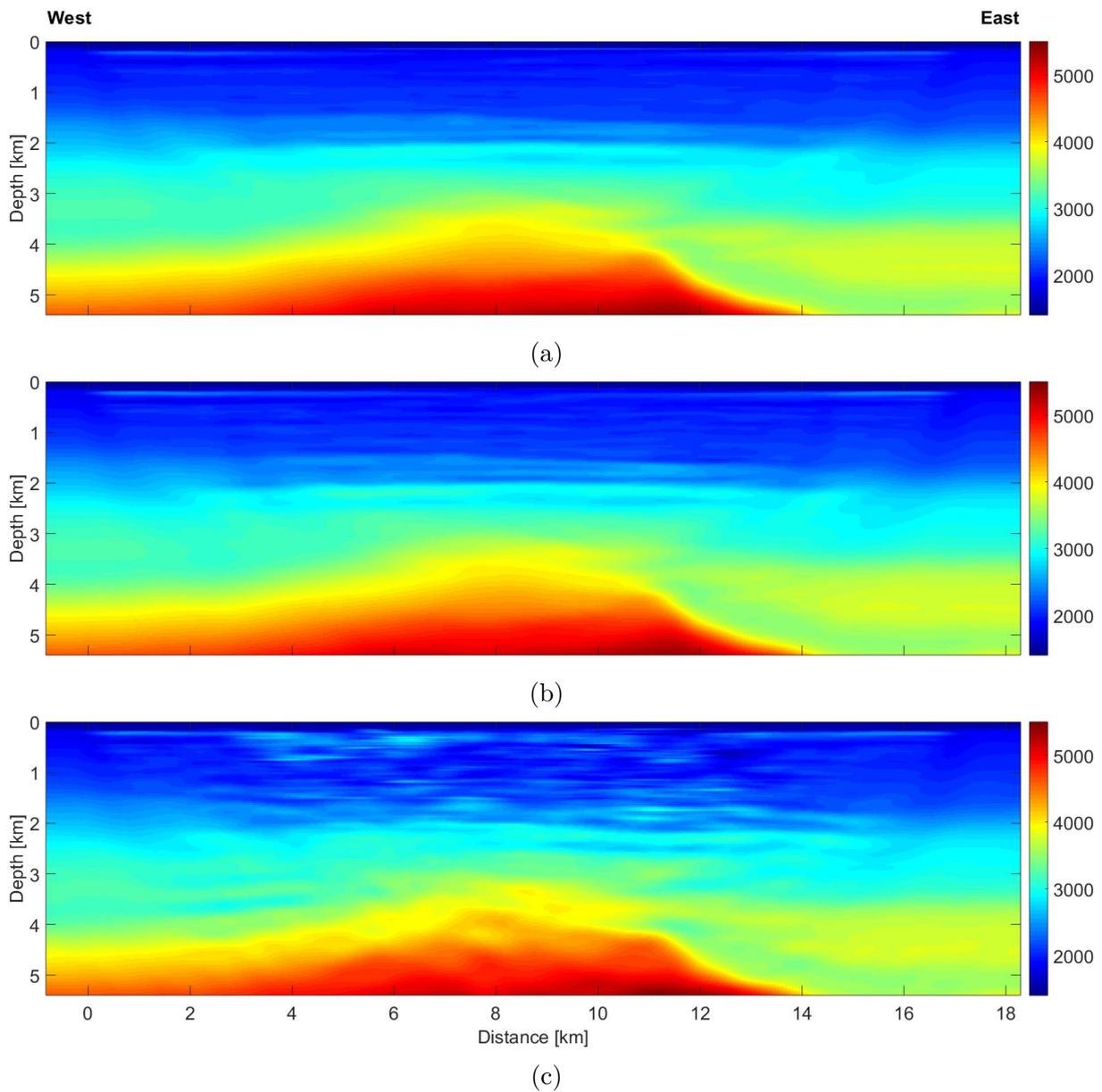


Figure 5.26: Original strategy with modified stage order. P-wave vertical velocity model inversion results. (a) Stage 2' initial model, for comparison purpose; (b) Stage 2' final model; (c) Stage 1b' final model. The horizontal distance is computed along the profile from the western most source.

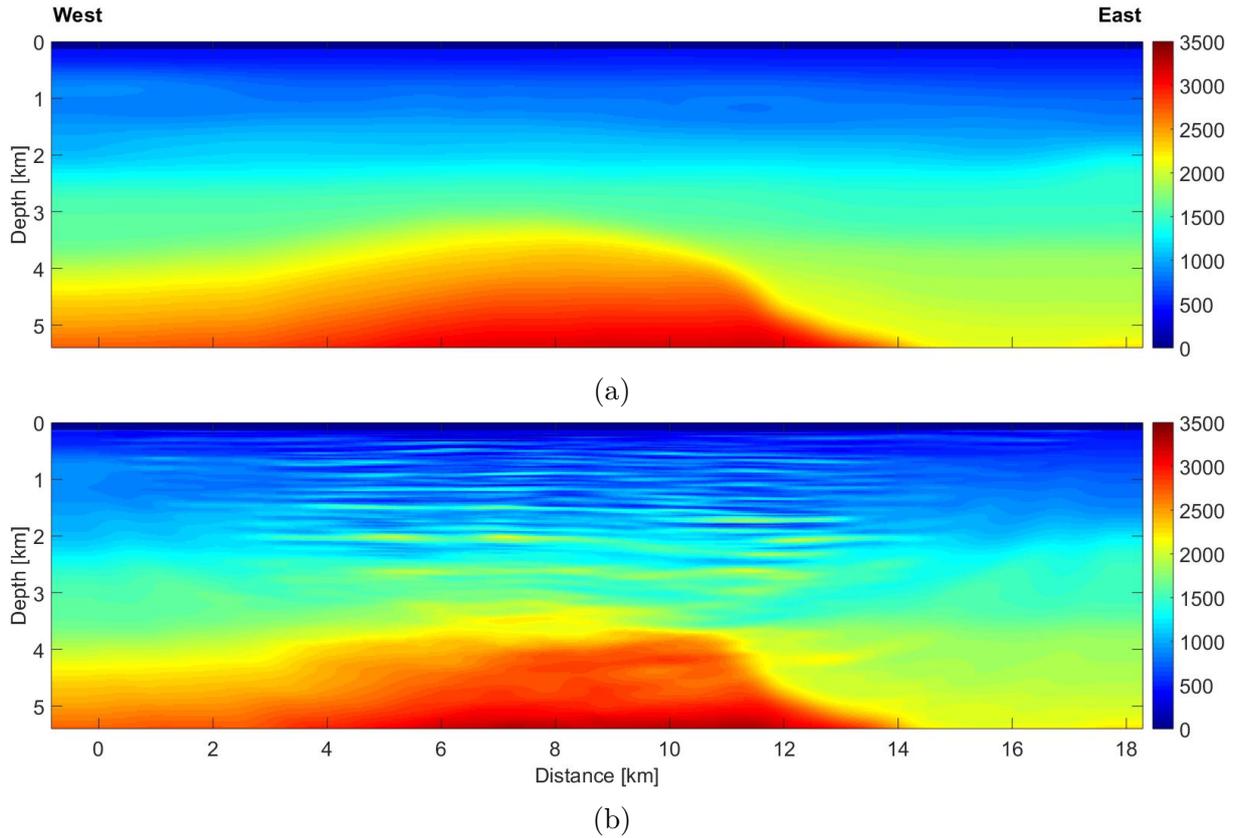


Figure 5.27: Original strategy with modified stage order. S-wave vertical velocity model inversion results. (a) Stage 2' initial model, for comparison purpose; (b) Stage 2' final model. The horizontal distance is computed along the profile from the western most source.

5.6 Inversions with Scholte waves

Performing a Scholte wave free inversion as in stages 1b, 1b', 3 and 3' is rationally defensible, but its benefit is not proven. Such strategy notably shows possible limitations coming from (1) the imperfect Scholte waves removal, and (2) the absence of shallow model update.

As a first step to evaluate such strategy, the previously successful stages 1b and 3' first frequency range inversions are performed omitting the Scholte wave removal strategy. The same source function scaling values as previously are used to ensure the results comparability between inversions with and without Scholte waves. The number of iterations is selected to allow for a stabilization of the cost functions and of the updated models.

For stage 1a, the obtained P-wave vertical velocity model is almost equal to the one obtained without Scholte waves (figures 5.28 and 5.29). For stage 3', the obtained vertical velocity models are similar to the one obtained without Scholte waves (figures 5.30 and 5.31) and the differences lie in the strength of update for both P- and S-wave vertical velocities, as well as slight depth differences for the S-wave vertical velocity models mainly, leading to a slightly better fit with well data. This improvement could be explained by a better shallow sub-surface whose update is then allowed, possibly combined with a slower convergence of the update.

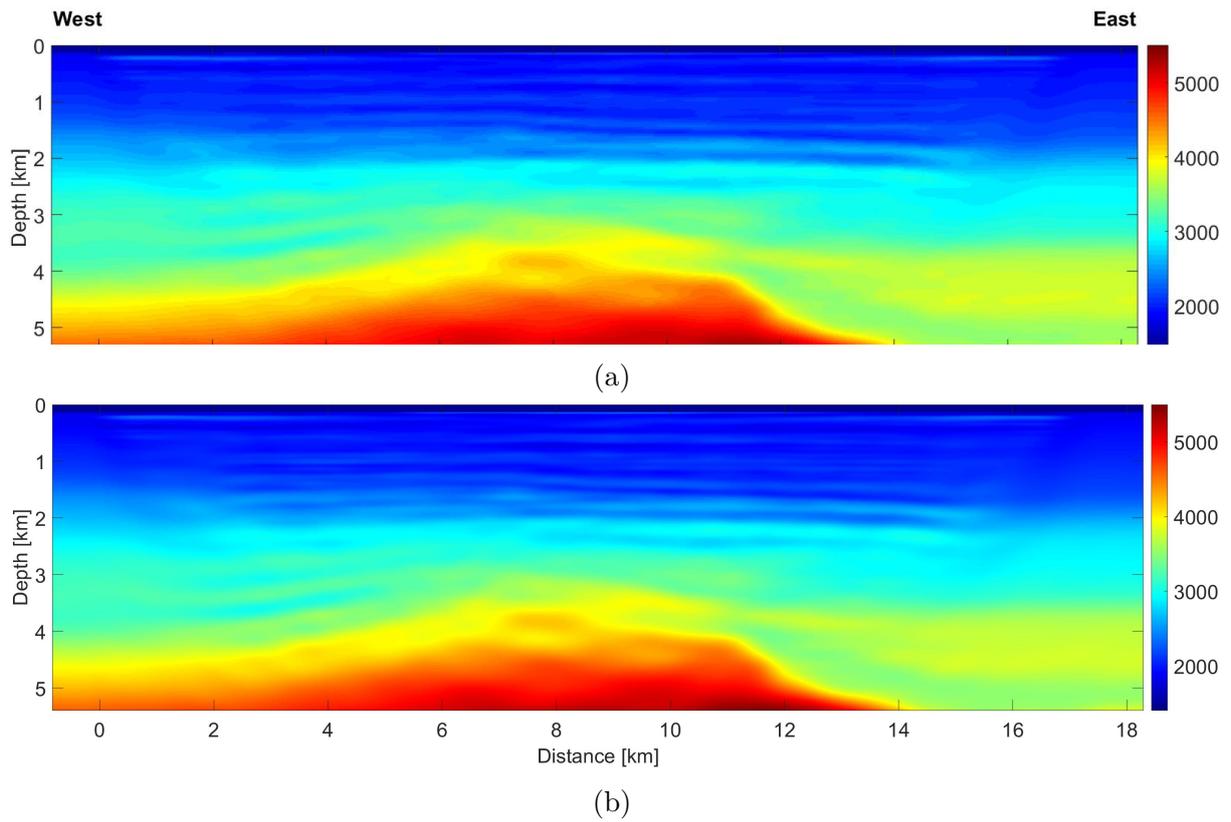


Figure 5.28: Stage 1b first frequency range. P-wave vertical velocity model inversion results obtained with (a), and without (b) Scholte waves. The horizontal distance is computed along the profile from the western most source.

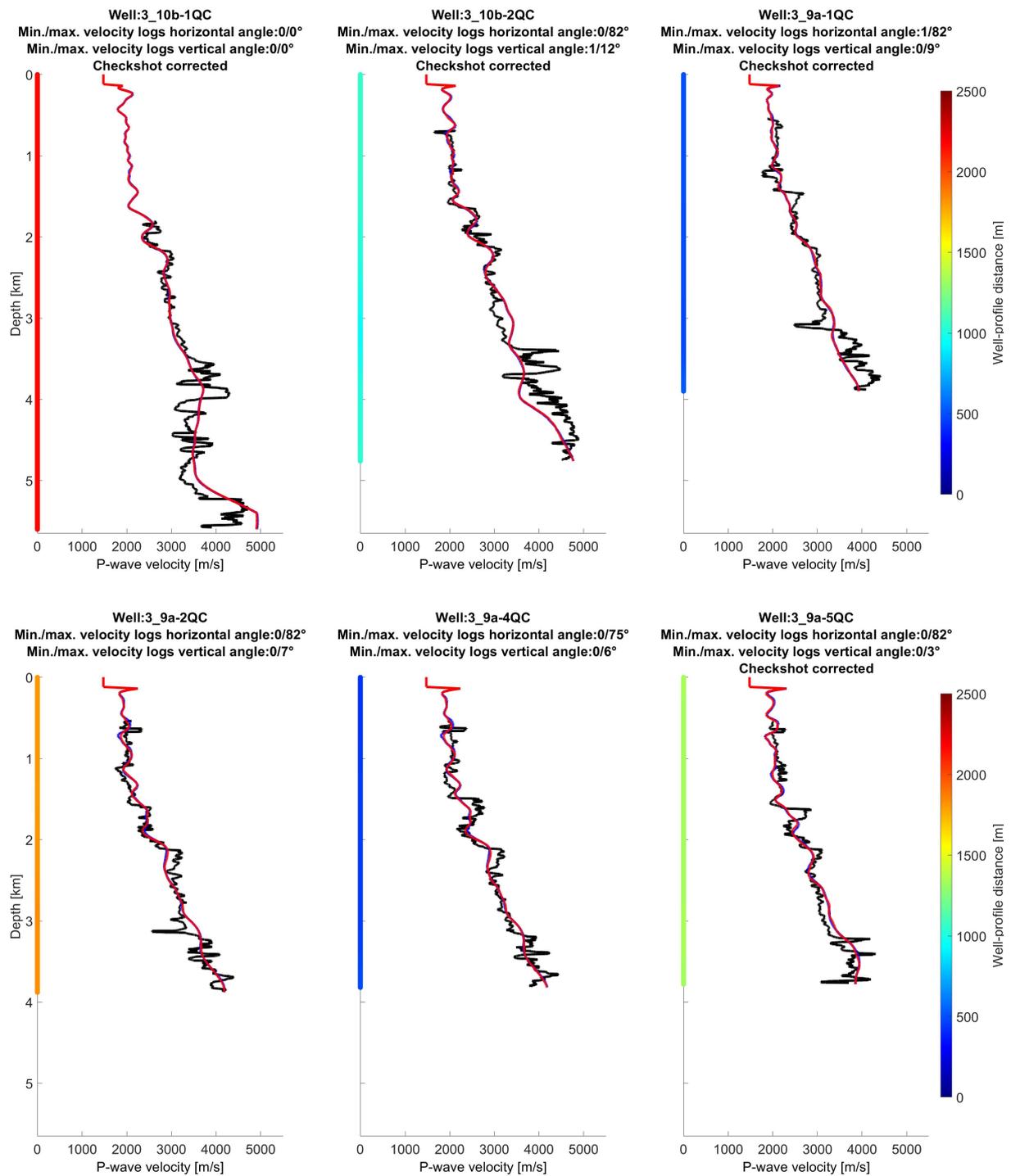


Figure 5.29: Stage 1b first frequency range. final P- and S-wave vertical velocity models obtained with (thin red line) and without (thin blue line) Scholte waves compared with smoothed velocities from a selection of wells (thin black line for P-wave velocities and thin grey line for S-wave velocity). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

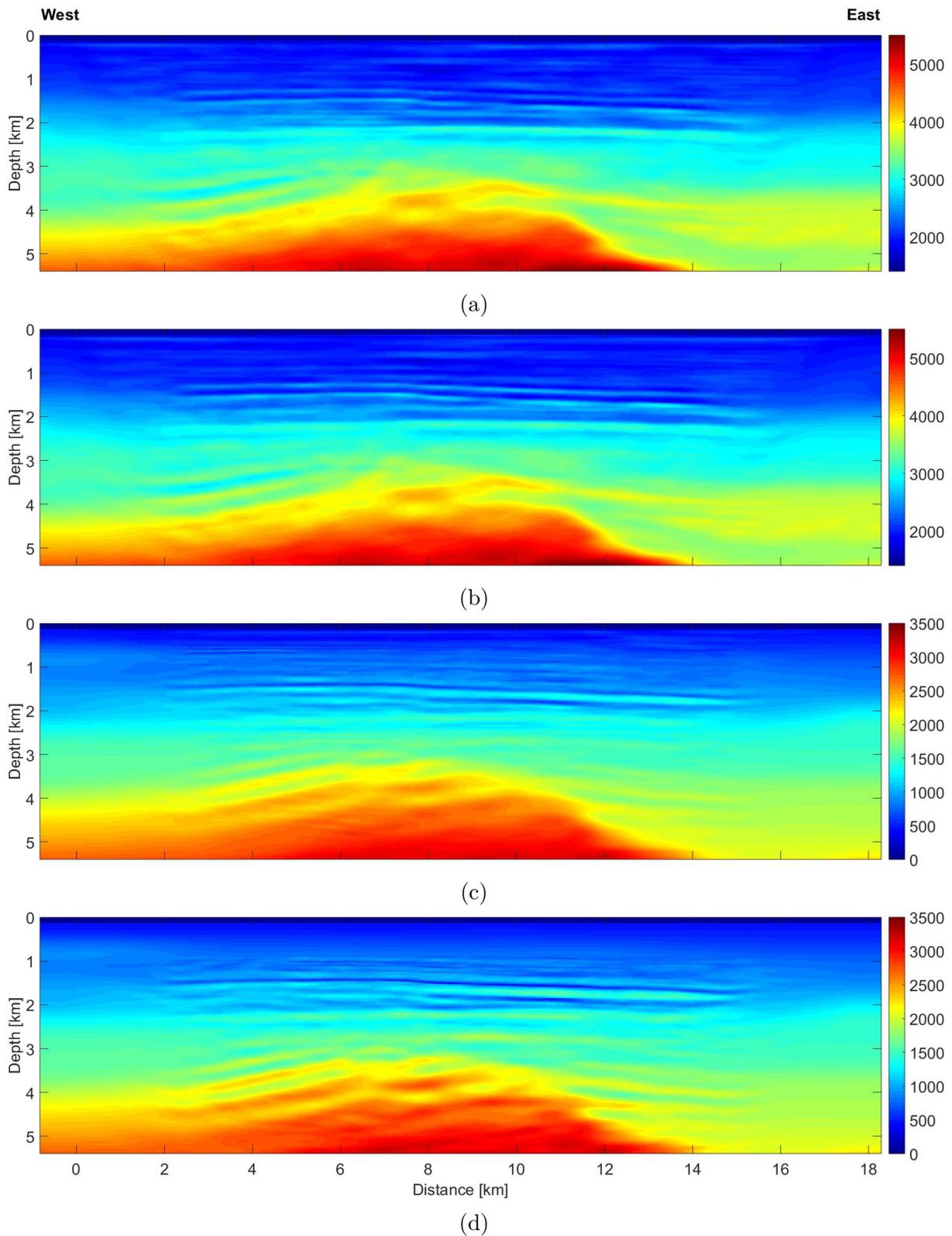


Figure 5.30: Stage 3' first frequency range. P-wave vertical velocity model inversion results obtained with (a), and without (b) Scholte waves; S-wave vertical velocity model inversion results obtained with (c), and without (d) Scholte waves. The horizontal distance is computed along the profile from the western most source.

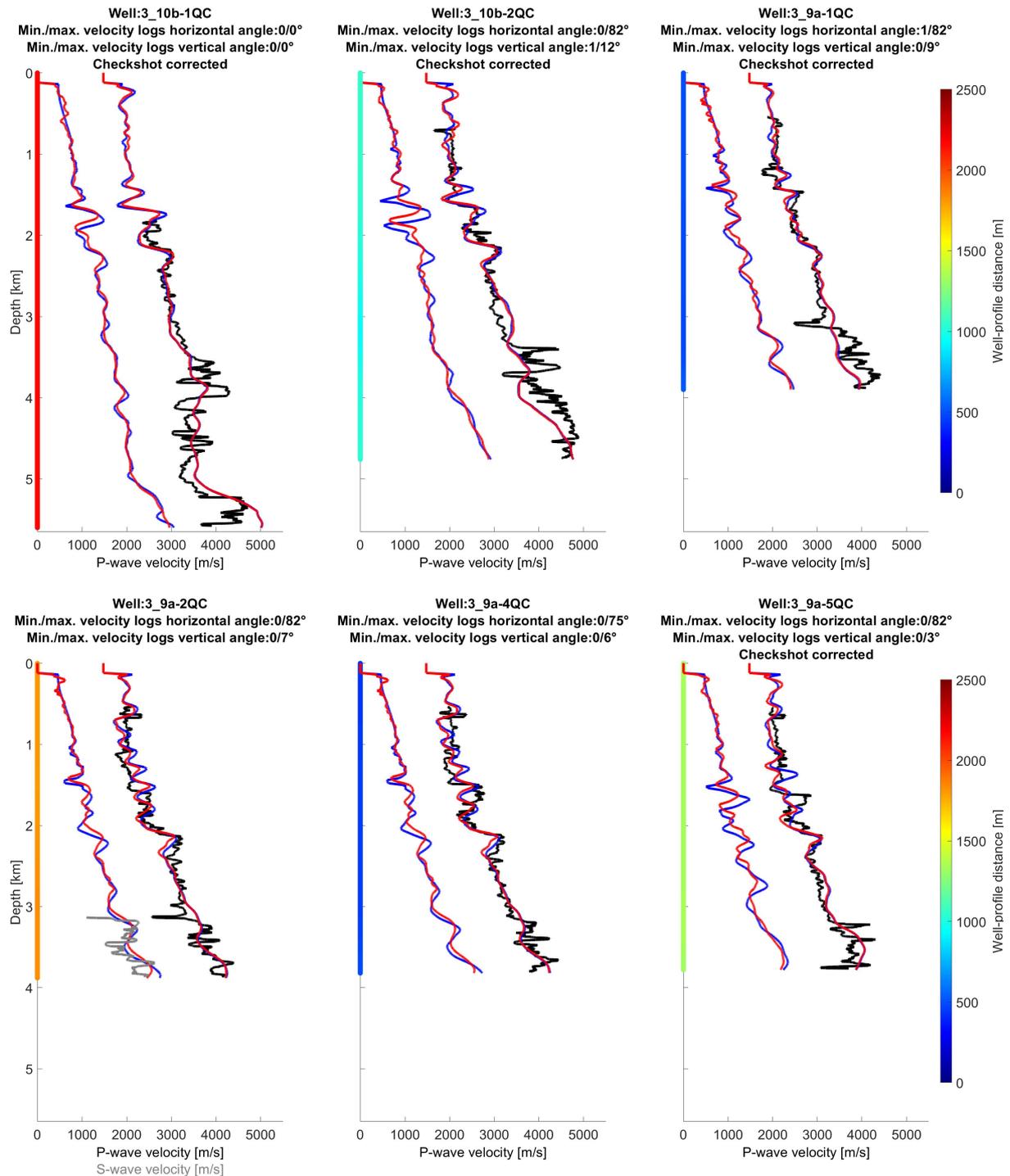


Figure 5.31: Stage 3' first frequency range. Final P- and S-wave vertical velocity models obtained with (thin red line) and without (thin blue line) Scholte waves compared with smoothed velocities from a selection of wells (thin black line for P-wave velocities and thin grey line for S-wave velocity). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

Unlike expected, FWI shows poor (stage 3') to no (stage 1b) sensitivity to relatively strong and wrong Scholte waves (figures 5.32a and 5.33a), with possibly slightly better results without the Scholte wave free inversion strategy, in a similar number of iterations.

Based on those tests, FWI seems to be able to show insensitivity to noise (here the Scholte waves). A deeper investigation is needed to confirm and explain the hypothetical improvement and determine the mechanisms of such robustness. For stage 1b, a possible mechanism could be the physical difficulty to modify the Scholte waves through the P-wave vertical model update (figure 5.32b). For stage 3', a possible mechanism could be link to the absence of strong constructive interference in the gradient construction, after fitting of surface waves with cycle skipping (figure 5.33). Note that the wrong initial Scholte wave fit indicates initial errors in the very shallow S-wave model, and the cycle skipping indicates that those errors increase during stage 3' inversion. Aside from allowing for a better visualization of the useful reflections, it might not be advantageous to adopt such Scholte wave free inversion strategy.

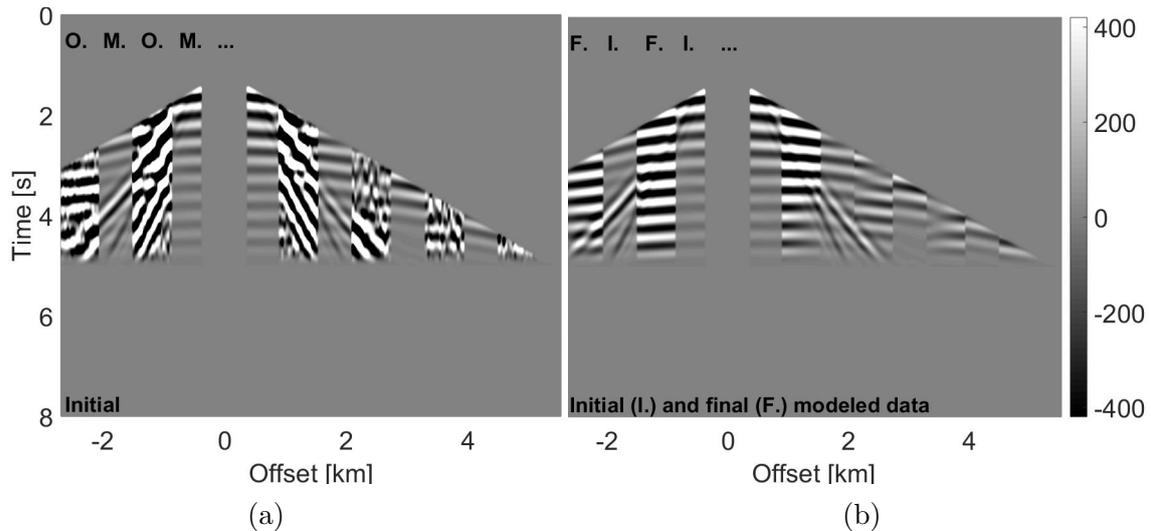


Figure 5.32: Stage 1b first frequency range with Scholte waves. (a) Observed (O.) and initial modeled (M.) data represented in an alternate manner; (b) Final (F.) and initial (I.) modeled data represented in an alternate manner. Color scale is saturated for visualization purpose. Western most pressure receiver gather.

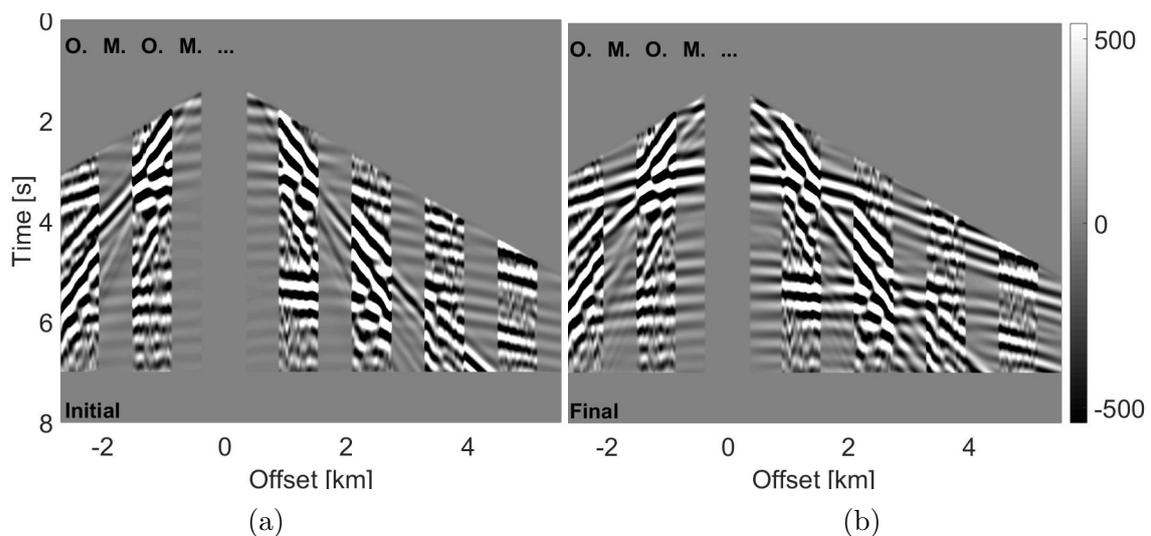


Figure 5.33: Stage 3' first frequency range with Scholte waves. (a) Observed (O.) and initial modeled (M.) data represented in an alternate manner; (b) Observed (O.) and final modeled (M.) data represented in an alternate manner. Color scale is saturated for visualization purpose. Western most horizontal velocity component receiver gather.

5.7 Discussion and conclusions

I applied on real data the two inversion strategies theoretically and synthetically explored in [Chapter 3](#). I recall that those strategies aim at obtaining P- and S-wave vertical velocities in soft seabed [VTI](#) environments, the remaining anisotropy parameters not being updated. Those applications confirm that the proposed variant (which entirely relies on the increasing frequency multiscale strategy of [Bunks et al. \(1995\)](#) ([35](#)) combined with the data low-frequency content) can by itself build the intermediate scale features of the S-wave vertical velocity model, and provide thrilling final inversion results. Those applications also corroborate the usefulness of such a variant strategy, with the unsuccessful application of the original inversion strategy. Indeed, a detrimental failing of the S-wave velocity model update from the amplitudes of P-wave data (stage 2 and 2') is observed, even when modifying the order of the stages. This modified order led nevertheless to improved inversion results, and might consequently be preferable over the initial order. It is very likely that this failing is linked to the recurring amplitude issue in the modeled data. This amplitude problem could for instance be due to the choice of geometrical spreading correction, the intrinsic attenuation, improvable source functions or even a bad coupling with the seabed for the inversion stage with the horizontal velocity component. Those assumptions are supported by (1) increasing difficulties with the horizontal velocity component inversion stage (stage 3 and 3'), which includes a larger amount of S-wave related events, and (2) better results at low frequencies. This last point indicates that a possible additional inversion strategy could be a mixed of the original strategy and its variant, with an application of stage 2 or 2' at low frequencies only. Based on the findings of [Chapter 3](#), this unsuccessful inversion with the original strategy is expected to be specific to the application, and the usefulness of the original strategy is not questioned. For instance, the original strategy might remain preferable for applications with less amplitude and eventually more anisotropy modeling defects.

Among the available flavors of [FWI](#), additional possibilities have been mentioned but not selected for the sake of simplicity. However, they should be examined to further improve the current inversion results. In addition to the already mentioned possible improvements, one could consider, during stage 3': (1) the use of the vertical velocity component in addition to the horizontal velocity component, (2) the update of S-wave velocities only, or (3) to not use the highest frequencies. Indeed, P-wave vertical velocity model stage 3' high frequency updates are questionable and would benefit additional constraints (compare the figures [5.30b](#), [5.30d](#) and [5.31](#) with the figures [5.16b](#), [5.17b](#) and [5.18](#)), as guessed and interpreted during the corresponding cost functions evolution analysis (see [Section 5.3](#)). In fact, the inverted models obtained without using the highest frequencies during stage 3' allow to obtain better P- over S-wave vertical velocity ratios, as shown on figure [5.34](#)). Indeed, this ratio model exhibits more reasonable extreme values. Naturally, the modeling of more complex kinds of anisotropy, the update of anisotropy parameters and the modeling/update of attenuation is also to consider to improve the results. In this research context, Alwyn real dataset, through its highly [VTI](#) anisotropic layers, its shear-wave splitting and the hints for inadequate [VTI](#) and attenuation parameters, could be an excellent test dataset to assess any synthetic finding.

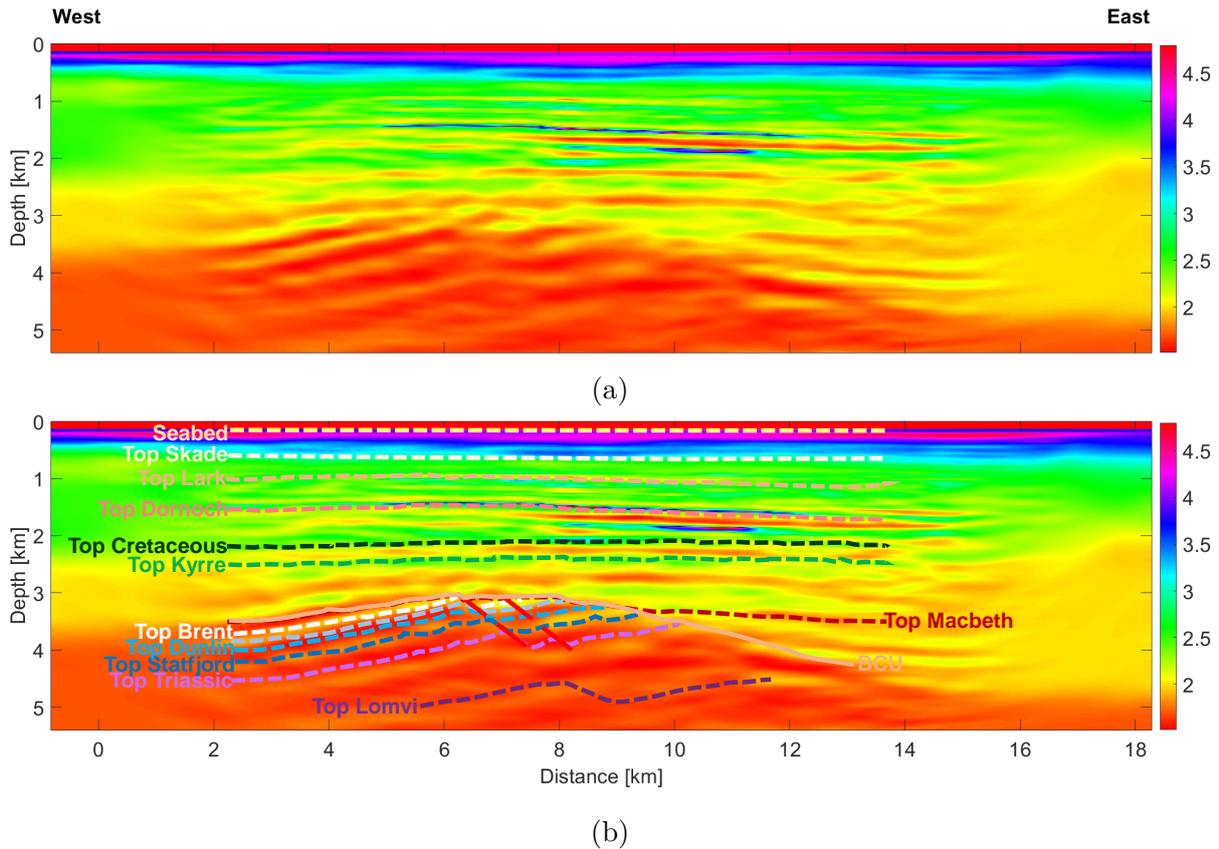


Figure 5.34: Variant strategy stopped after the first frequency range of the last stage, i.e. stage 3'. (a) Final P- over S-wave vertical velocity model; (b) Interpreted faults (thick dashed red lines) and horizons from the vintage migrated image, overlaid above a. To increase visibility, the color of the horizons has been modified in comparison with figure 4.3. The non-labeled horizon is an intra-Brent group horizon. BCU = Base Cretaceous Unconformity. The color scale is clipped above 4.8 to increase visibility of reasonable ratios, maximum ratio is ~ 5.8 . The horizontal distance is computed along the profile from the western most source.

Despite all approximations, the variant inversion strategy allowed for a drastic improvement of the model, qualified by reduced cost function values, and an improved fit with the observed data, migrated image, well logs and lithology. The inversion notably enabled the imaging of additional structures compared with the initial models. All in all, even though additional verification means could be used, this chapter provides a real data FWI application to retrieve both P- and S-wave vertical velocities in soft seabed environment, with results supported in an extensive way by verification tools. Among the reasons for P- and S-wave velocity retrieval cited in Chapter 3, those data were notably acquired to bring more continuity and more resolution to the reservoir structures (see Chapter 4), through high quality P-P and P-S migrations. Comparison and exploitation of P-P and P-S migrations between the initial and the final models would be the last step to evaluate the usefulness of the current FWI application in this context.

Possible additional verification means include: anisotropy corrected well velocity data; estimation of expected velocity variation with distance from well; migrations and pre-stack depth migration gathers; intermediate source inversion with used data; cost function evolution with depth and horizontal distance; P- over S-wave vertical velocity ratio computa-

tions with similar resolution for both models; and use of all components and all data (e.g. vertical velocity component for an inversion with pressure component; wide aperture data for an inversion with narrow aperture data; or low-frequency data for an inversion with high-frequency data). Generally speaking the design and use of more systematic (notably for lithology, including fluid content and depth effects), automatic and quantified verification means would be a useful contribution. Such an improved verification workflow would help assessing the validity of final models and making a better use of them. Additionally, such a verification workflow could help explaining (then avoiding or correcting) invalid updates.

This work highlights three additional points:

- Comparing the constructive updates of stage 3' (figures 5.16 and 5.17) with the destructive updates of stage 3 (figures 5.22c and 5.23c), stage 3 differing from stage 3' by the previous application of stage 2 that updates mainly the S-wave velocity model, this illustrates the fact that a poor S-wave velocity model can have a detrimental effect on both P- and S-wave velocity model updates. Consequently, those inversions highlighted the essential role of initial S-wave velocity models and inversion strategies.
- The density sensitivity of stage 1a cost functions (figure 5.13) confirms the predicted limit of the reflection coefficient based assessment of the sensitivity of the wide-aperture P-wave data to erroneous remaining anisotropy parameters (Section 3.4). Indeed, such assessment was supporting no sensitivity to density perturbations, in opposition with this real data application. This observation highlights the limitations of theoretical and synthetic studies and supports the need for real data tests when developing inversion strategies.
- FWI can show robustness to relatively strong noise (here wrong Scholte waves). Some possible explanatory mechanisms are proposed but not proven. This questions the need for a despiking routine as performed on the horizontal component: FWI may also show robustness to this kind of noise, leading to the uselessness of such a noise removal routine.

All the inversions of this chapter are notably allowed by the use of reciprocity as described in Chapter 2. For example, the application of the variant strategy requires more than 14 days of computation (stage 1a: ~ 3 hours and 54 minutes, stage 1b: ~ 196 hours and 4 minutes, and stage 3': ~ 138 hours and 44 minutes) with full shot parallelization and domain decomposition on 6 (resp. 16) domains for stage 1a (resp. stages 1b and 3') inversions. Without reciprocity, the same application would have required more than 5 months of computation, and years of computations when combined with no domain decomposition. Note that those computing times might be further reduced using one of the other techniques described in Chapter 2, and by the application of the undersampling strategy to stage 1b. Indeed, since Scholte waves-free inversion is performed, the imprint of unstable S-waves on the data might be minimal. Moreover, the computing cost could also have been reduced by slightly decreasing the number of iterations, as well as by modeling only the used data in stages 1b and 3' (i.e., up to 4.5 and 7 s respectively). Note that either or both the computing costs or the computing resources still have to be improved, in order to allow for 3D applications.

Conclusion

This research work is a consequence of the will to obtain both P- and S-wave velocity models of the sub-surface, for lithology identification purposes for instance. In pursuit of such a goal, the seismic imaging Full Waveform Inversion (**FWI**) class of algorithms appears to be the appropriate method. In these circumstances, a group of scientists (Sears et al., 2008 ([164](#))) previously developed an **FWI** strategy for soft seabed environments (i.e., environments with seabed S-wave velocities around or below 300 m/s), in the local Earth exploration context. Those soft seabed applications need to be investigated separately because the corresponding seismic data exhibit a lower S-wave velocity information content. Moreover, those soft seabed environments are typical of (unconsolidated) siliclastic sedimentary seabed environments, and consequently apply to a non-negligible fraction of the Earth sub-surface exploration projects.

However, the proposed strategy was developed assuming an isotropic sub-surface, while the Earth is anisotropic. This Ph.D. research work aimed at expanding the efforts of those scientists towards applications in Vertical Transverse Isotropic (**VTI**) environments, **VTI** being a common anisotropy reality in sedimentary basins. Specifically, inspired from real applications, the objective was to understand (1) how to apply such strategy to update both P- and S-wave vertical velocities, the remaining anisotropy parameters not being updated, and (2) what to expect if those remaining anisotropy parameters were to be slightly erroneous or if the data were to show other real data complexities such as attenuation.

Before being able to answer those questions, an additional piece of work was needed. Indeed, the proposed strategy, relying on multi-component ocean-bottom data, was developed and applied on Ocean Bottom Cable (**OBC**) data. However **OBC** are more and more replaced by Ocean Bottom Nodes/Seismometers (**OBNs/OBSs**). I showed that for computing resources reasons, this replacement beget the use of reciprocity, where the role of sources and receivers are interchanged during modeling and inversion. Although already applied in a few **FWI** applications, such a use of reciprocity was poorly documented.

As a consequence, I first provided the needed complete guide of such a use of reciprocity for 2D/3D **FWI** in elastic anisotropic media, including the limitations of this computing cost reduction strategy, the used thought process, extensions to other cases, as well as the justification of the required theoretical equality between **FWI** with and without reciprocity ([Chapter 2](#)).

This handbook allowed to tackle the initial question of the application of the proposed soft seabed inversion strategy to **VTI** environments ([Chapter 3](#)). An elastic **VTI** **FWI** inversion tool being already available, the adaptation of the strategy to such media consisted in the choice of the proper parameterization, based on reflection radiation patterns and synthetic inversion tests. Then, additional synthetic **FWI** applications with

erroneous anisotropy parameters (inspired from real data) allowed to obtain some physical insight about the sensitivity of the proposed strategy to erroneous anisotropy parameters: although perturbed, synthetic inversions indicate that the strategy can remain reasonably successful, with inversion errors that could be used to characterize the anisotropy parameter errors. Worried about other physical phenomena that are usually not taken into account (such as attenuation), I explored a variant that is expected to be more robust to aforementioned approximations, but which shows a higher sensitivity to erroneous anisotropy model, and relies on low-frequency data content of modern acquisitions instead of amplitudes of P-P wide-aperture reflections.

This variant strategy finally prove its usefulness during its application on the available (OBN) real data, on which this research work was founded (Chapter 4 and Chapter 5). The need for such a variant strategy was especially highlighted when compared to the poorer results obtained with the original strategy. Even though multiple possibilities of improvement have been proposed, this FWI application fulfilled its promises by providing: (1) P- and S-wave velocity models with an increased resolution allowing the interpretation of layers and other structures up to the main faults, and (2) quantitative models leading to P- over S-wave vertical velocity ratios in agreement with the known rock types. Closing the loop, this last point illustrates how this inversion strategy could participate in lithology identification.

All in all, I performed a successful real data application, which is part of the last steps before production. Indeed, a number of approximations usually performed in real data are not taken into account in the synthetic and theoretical developments. That is why validations on real data are always needed.

With this real data application and the underlying synthetic tests, I hope not only to provide an example of how to obtain P- and S-wave vertical velocity models in soft seabed VTI environments with OBN/OBS data, but also to clarify the reasoning behind this application. Indeed, I believe that such a detailed understanding is needed to allow for the informed adjustments that will undoubtedly be needed for other applications. In comparison with the few real applications publicly disclosed, this real application gathers in a single place the VTI aspect, results supported with a large number of quality control tools, and comparisons of different flavors of the soft seabed strategy (Sears et al., 2010 (165); Prioux et al., 2012 (147); Vigh et al., 2014 (210); Alves, 2017 (5)).

Regarding each studied aspect (i.e., the reciprocity, the synthetic and theoretical study of the extension of the soft seabed strategy to VTI media, and the real data application), a number of perspectives were mentioned in each chapter. Among them, my own priority would be the real data related perspectives: real data, by being the final objective, impose the needs and should therefore dictate the approach. In such a case, the main research lines would probably include: decreasing FWI computing needs and/or increasing computing capabilities, fully anisotropic FWI, and FWI with attenuation.

In addition to the real data oriented attitude, this research work helped me distinguish three other general research axis that seem to be primordial. The first one is related to the non-uniqueness and local aspect of FWI, that has been conspicuous all along this thesis. Therefore, any development of methods to quantify the inversion errors and statistically support the inversion results, together with the development of more global and robust inversion techniques is needed. The second research axis involves the ideal of an inversion strategy “one size fits all”. Indeed, the current FWI landscape is made of a multitude of

flavors with various advantages/disadvantages which are combined in a myriad of ways, depending on the application and the geophysicist. Working towards a single tool to facilitate applications would be an excellent contribution. At last, we should strive to determine what input is precisely required to achieve the desired results, in addition of exploiting what data are already available. This attitude already proved its usefulness in the past. For instance, it led to multicomponent ocean bottom acquisition systems, which are at heart of the strategies explored in this thesis, as well as other strategies. Regarding on-going projects of that kind, one can also cite the development of ultra-low frequency sources for initial model building with [FWI](#) (e.g., Chelminski et al., 2019 (41)).

Appendix A

Details on Newton's method

I describe here the theory behind Newton's method and the limitations of this method.

Let \mathcal{O} be a quadratic objective function of \mathbf{m} . Let us call m_i , with $i \in \llbracket 1, N \rrbracket$, the components of \mathbf{m} . i is a dummy variable and a similar notation is used for all vectors. Note that the subscript n is omitted in this appendix to lighten the equations. In the search for the model \mathbf{m}^* which optimizes \mathcal{O} , let us start from an initial model $\mathbf{m}_0 \in \mathbb{R}^N$. Whatever column vector $\delta\mathbf{m} \in \mathbb{R}^N$, the second order Taylor-Young expansion of \mathcal{O} at \mathbf{m}_0 is exact and is written as:

$$\mathcal{O}|_{\mathbf{m}_0+\delta\mathbf{m}} = \mathcal{O}|_{\mathbf{m}_0} + (\nabla\mathcal{O}|_{\mathbf{m}_0})^T \delta\mathbf{m} + \frac{1}{2}(\delta\mathbf{m})^T \mathbf{H}|_{\mathbf{m}_0} \delta\mathbf{m}. \quad (\text{A.1})$$

$\nabla\mathcal{O}$ and \mathbf{H} are respectively the gradient and the Hessian (i.e., the second derivative) of \mathcal{O} with respect to the model parameters, such that:

$$\nabla\mathcal{O} = \begin{pmatrix} \frac{\partial\mathcal{O}}{\partial m_1} \\ \vdots \\ \frac{\partial\mathcal{O}}{\partial m_N} \end{pmatrix} \text{ and } \mathbf{H} = \begin{pmatrix} \frac{\partial^2\mathcal{O}}{\partial m_1^2} & \cdots & \frac{\partial^2\mathcal{O}}{\partial m_1\partial m_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2\mathcal{O}}{\partial m_N\partial m_1} & \cdots & \frac{\partial^2\mathcal{O}}{\partial m_N^2} \end{pmatrix}.$$

Differentiation of equation A.1 with respect to any δm_l gives:

$$\frac{\partial\mathcal{O}}{\partial\delta m_l} \Big|_{\mathbf{m}_0+\delta\mathbf{m}} = \frac{\partial\mathcal{O}}{\partial m_l} \Big|_{\mathbf{m}_0} + \left(\frac{\partial\nabla\mathcal{O}}{\partial m_l} \Big|_{\mathbf{m}_0} \right)^T \delta\mathbf{m}. \quad (\text{A.2})$$

The model \mathbf{m}^* is reached when all $\frac{\partial\mathcal{O}}{\partial\delta m_l} \Big|_{\mathbf{m}_0+\delta\mathbf{m}}$, $l \in (\mathbb{N} \cap [1, N])$, vanish. Indeed, because \mathcal{O} is a quadratic function, it has a unique strict global optimum which is the optimum. It leads to a unique $\delta\mathbf{m}$ such that:

$$\delta\mathbf{m} = -(\mathbf{H}|_{\mathbf{m}_0})^{-1} \nabla\mathcal{O}|_{\mathbf{m}_0}. \quad (\text{A.3})$$

The model \mathbf{m}^* then becomes:

$$\mathbf{m}^* = \mathbf{m}_0 - (\mathbf{H}|_{\mathbf{m}_0})^{-1} \nabla\mathcal{O}|_{\mathbf{m}_0}. \quad (\text{A.4})$$

The objective function \mathcal{O} is not necessarily a quadratic function: it can have several stationary points, which can be saddle points, (strict) (local or global) maxima or minima.

Based on equation A.4, Newton's method proposes to iteratively find a stationary point with:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mathbf{H}_n^{-1} \nabla \mathcal{O}_n, n \in \mathbb{N}^+, \quad (\text{A.5})$$

$\nabla \mathcal{O}_n$ and \mathbf{H}_n are respectively the gradient and the Hessian of \mathcal{O} evaluated at \mathbf{m}_n . The method requires that \mathcal{O} be locally differentiable up to the second order and it locally approximates \mathcal{O} by a quadratic function, setting the updated model as equal to the unique strict global optimum of this quadratic function. The iterative aspect of this method aims at overcoming the fact that the objective function is not a quadratic function.

Besides the intrinsic limitation of local optimization methods; Newton's method has its own shortcomings. In a minimization context for instance:

- It is not a minimization method but an optimization method: it can lead to a maximum, even though we were aiming at a minimum.
- It can be stuck at a saddle point (because the gradient is null at this point).
- It can be stuck at a point where the objective function is locally linear (because the Hessian is null).
- It can fail at computing an update if the cost function is not differentiable up to the second order.

Appendix B

Details on the computation of the opposite direction of update: conjugate-gradient and **L-BFGS-B**

I describe here how to compute the opposite direction of update for Polak-Ribière conjugate-gradient (Polak and Ribière, 1969 ([145](#))) and Nocedal's **L-BFGS** (Nocedal and Wright, 2006 ([136](#))).

I remind that both inversion methods have the form:

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \beta_n \mathbf{p}_n, \tag{B.1}$$

where \mathbf{p}_n is the direction of update, and β_n is the steplength at iteration n . \mathbf{m} is the model.

Polak-Ribière conjugate gradient

For the Polak-Ribière conjugate-gradient (**PR-CG**), the opposite direction of update $\mathbf{p}_n^{\text{PR-CG}}$ has the form:

$$\mathbf{p}_n^{\text{PR-CG}} = -\nabla \mathcal{O}_n + \frac{(\nabla \mathcal{O}_n - \nabla \mathcal{O}_{n-1})^T \nabla \mathcal{O}_n}{\nabla \mathcal{O}_{n-1}^T \nabla \mathcal{O}_{n-1}} \mathbf{p}_{n-1}^{\text{PR-CG}}. \tag{B.2}$$

L-BFGS

In our full waveform inversion algorithm, the opposite direction of update for L-BFGS $\mathbf{p}_n^{\text{L-BFGS}}$ is computed as follows:

Pseudo-code: $\mathbf{p}_n^{\text{L-BFGS}}$ computation

```

 $\mathbf{q}_n = \nabla \mathcal{O}_n;$ 
for  $i=n-1 \rightarrow n-m$  do
   $\alpha_i = \frac{(\mathbf{m}_{i+1}-\mathbf{m}_i)^T \mathbf{q}_n}{\mathbf{y}_i^T (\mathbf{m}_{i+1}-\mathbf{m}_i)};$ 
   $\mathbf{q}_n = \mathbf{q}_n - \alpha_i \mathbf{y}_i;$ 
 $\mathbf{r}_n = \frac{\mathbf{y}_{n-1}^T (\mathbf{m}_n - \mathbf{m}_{n-1})}{\mathbf{y}_{n-1}^T \mathbf{y}_{n-1}} \mathbf{q}_n;$ 
for  $i=n-m \rightarrow n-1$  do
   $\mathbf{r}_n = \mathbf{r}_n + (\mathbf{m}_{i+1} - \mathbf{m}_i) \left( \alpha_i - \frac{\mathbf{y}_i^T \mathbf{r}_n}{\mathbf{y}_i^T (\mathbf{m}_{i+1} - \mathbf{m}_i)} \right);$ 
 $\mathbf{p}_n^{\text{L-BFGS}} = -\mathbf{r}_n;$ 

```

with $\mathbf{y}_i = \nabla \mathcal{O}_{i+1} - \nabla \mathcal{O}_i$. The recommended range of m is $[3, 20]$ (Byrd et al., 1995 (36) and Zhu et al., 1997 (238)).

The Polak-Ribière conjugate-gradient (resp. L-BFGS-B) needs the gradient and the direction of update (resp. gradient) from previous iterations. For the first iteration, these data are not available and both algorithms are replaced by the steepest-descent. For next iterations of L-BFGS-B, as long as necessary, m is set equal to the iteration number.

Appendix C

Steplength computation: details on Pica et al.'s and subspace methods

I describe here the theory behind the chosen steplengths: Pica et al.'s and subspace methods.

Pica et al. (1990) (143) assumes that the objective function is locally a quadratic function in the direction of update. It can therefore be written as a second order Taylor-Young expansion such that:

$$\mathcal{O}(\mathbf{m}_n + \beta_n \mathbf{p}_n) = \mathcal{O}(\mathbf{m}_n) - \beta_n \sum_{i=1}^N \frac{\partial \mathcal{O}}{\partial m_i} \Big|_{\mathbf{m}_n} p_i + \frac{\beta_n^2}{2} \sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2 \mathcal{O}}{\partial m_i \partial m_j} \Big|_{\mathbf{m}_n} p_i p_j. \quad (\text{C.1})$$

The steplength β_n is chosen such that it cancels the derivative of $\mathcal{O}(\mathbf{m}_n + \beta_n \mathbf{p}_n)$ with respect to β_n . By differentiating equation C.1 with respect to β_n and setting the derivative to zero, the steplength β_n can then be written as follows:

$$\beta_n = \frac{\sum_{i=1}^N \frac{\partial \mathcal{O}}{\partial m_i} \Big|_{\mathbf{m}_n} p_i}{\sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2 \mathcal{O}}{\partial m_i \partial m_j} \Big|_{\mathbf{m}_n} p_i p_j}. \quad (\text{C.2})$$

Let us focus on our ℓ_2 -norm cost function. Let $\mathbf{d}_n^{\text{mod}}$ be a vector containing all the modeled data from \mathbf{m}_n . Let us call \mathbf{J}_n^T the transposed Jacobian of the modeled data with respect to the model parameters, evaluated at \mathbf{m}_n . The steplength β_n can be written as:

$$\beta_n = \frac{\nabla \mathcal{O}_n \cdot \mathbf{p}_n}{(\mathbf{J}_n^T \mathbf{p}_n) \cdot (\mathbf{J}_n^T \mathbf{p}_n)}. \quad (\text{C.3})$$

The denominator is computed by slightly perturbing the model \mathbf{m}_n in the update direction \mathbf{p}_n , and modeling the corresponding perturbed data \mathbf{d}^{pert} . Let us call $\epsilon \in \mathbb{R}^{+*}$ the small scalar used to generate the perturbed model $\mathbf{m}_n^{\text{pert}} = \mathbf{m}_n + \epsilon \mathbf{p}_n$. The perturbed data depends on the model. Assuming it is locally a linear function:

$$\mathbf{d}_n^{\text{pert}} = \mathbf{d}(\mathbf{m}_n + \epsilon \mathbf{p}_n) = \mathbf{d}(\mathbf{m}_n) + \epsilon \mathbf{J}_n^T \mathbf{p}_n = \mathbf{d}_n^{\text{mod}} + \epsilon \mathbf{J}_n^T \mathbf{p}_n. \quad (\text{C.4})$$

It leads to:

$$\mathbf{J}_n^T \mathbf{p}_n = \frac{\mathbf{d}_n^{\text{pert}} - \mathbf{d}_n^{\text{mod}}}{\epsilon}, \quad (\text{C.5})$$

then:

$$\beta_n = \epsilon^2 \frac{\nabla \mathcal{O}_n \cdot \mathbf{p}_n}{(\mathbf{d}_n^{\text{pert}} - \mathbf{d}_n^{\text{mod}}) \cdot (\mathbf{d}_n^{\text{pert}} - \mathbf{d}_n^{\text{mod}})}. \quad (\text{C.6})$$

The subspace method of Kennett et al. (1988) (94) is similar to Pica et al.'s method, but provides a steplength per parameter kind. Let us take the example of a model containing two parameter types (e.g., an acoustic isotropic model with the P-wave velocity V_p and the density ρ) and N_p spatial model points. Explanation of this method will be performed for this specific case for more clarity, it can easily be extended to all kinds and numbers of parameters. \mathbf{m}_n is split into $\mathbf{m}_n^{V_p}$ and \mathbf{m}_n^ρ such that:

$$\text{for } \mathbf{m}_n = \begin{pmatrix} m_{n,1}^{V_p} \\ \vdots \\ m_{n,N_p}^{V_p} \\ m_{n,1}^\rho \\ \vdots \\ m_{n,N_p}^\rho \end{pmatrix}, \quad \mathbf{m}_n^{V_p} = \begin{pmatrix} m_{n,1}^{V_p} \\ \vdots \\ m_{n,N_p}^{V_p} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \text{and } \mathbf{m}_n^\rho = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ m_{n,1}^\rho \\ \vdots \\ m_{n,N_p}^\rho \end{pmatrix}.$$

Similarly, \mathbf{p}_n is also split into $\mathbf{p}_n^{V_p}$ and \mathbf{p}_n^ρ . Following the same strategy as for Pica et al.'s method, one can write:

$$\begin{aligned} \mathcal{O}(\mathbf{m}_n - \beta_n^{V_p} \mathbf{p}_n^{V_p} - \beta_n^\rho \mathbf{p}_n^\rho) &= \mathcal{O}(\mathbf{m}_n) - \beta_n^{V_p} (\nabla \mathcal{O}_n) \cdot (\mathbf{p}_n^{V_p}) - \beta_n^\rho (\nabla \mathcal{O}_n) \cdot (\mathbf{p}_n^\rho) \\ &+ \beta_n^{V_p} \beta_n^\rho (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho) + \frac{\beta_n^{V_p 2}}{2} (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) + \frac{\beta_n^{\rho 2}}{2} (\mathbf{J}_n^T \mathbf{p}_n^\rho) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho). \end{aligned} \quad (\text{C.7})$$

The steplengths are chosen such that they cancel the derivatives of $\mathcal{O}(\mathbf{m}_n - \beta_n^{V_p} \mathbf{p}_n^{V_p} - \beta_n^\rho \mathbf{p}_n^\rho)$ with respect to the each steplength. By differentiating equation C.7 with respect to $\beta_n^{V_p}$ and β_n^ρ and setting the derivatives to zero, the steplengths meet the following equations:

$$\begin{cases} -\nabla \mathcal{O}_n \cdot \mathbf{p}_n^{V_p} + \beta_n^\rho (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho) + \beta_n^{V_p} (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) = 0, \\ -\nabla \mathcal{O}_n \cdot \mathbf{p}_n^\rho + \beta_n^{V_p} (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho) + \beta_n^\rho (\mathbf{J}_n^T \mathbf{p}_n^\rho) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho) = 0. \end{cases} \quad (\text{C.8})$$

Let us define the 2*2 matrix \mathbf{A}_n and the two 2*1 vectors \mathbf{b}_n and \mathbf{B}_n such that:

$$\mathbf{A}_n = \begin{pmatrix} (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) & (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho) \\ (\mathbf{J}_n^T \mathbf{p}_n^{V_p}) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho) & (\mathbf{J}_n^T \mathbf{p}_n^\rho) \cdot (\mathbf{J}_n^T \mathbf{p}_n^\rho) \end{pmatrix}, \quad \mathbf{b}_n = \begin{pmatrix} \nabla \mathcal{O}_n \cdot \mathbf{p}_n^{V_p} \\ \nabla \mathcal{O}_n \cdot \mathbf{p}_n^\rho \end{pmatrix}, \quad \text{and } \mathbf{B}_n = \begin{pmatrix} \beta_n^{V_p} \\ \beta_n^\rho \end{pmatrix}.$$

Equation C.8 can be written:

$$\mathbf{A}_n \mathbf{B}_n = \mathbf{b}_n. \quad (\text{C.9})$$

The vector \mathbf{B}_n contains the steplength values and can be find as follows:

$$\mathbf{B}_n = \mathbf{A}_n^{-1} \mathbf{b}_n. \quad (\text{C.10})$$

The components of \mathbf{A}_n are computable as for Pica et al.'s method. For each source and at each iteration, subspace method requires as many perturbed data simulations as the number of inverted parameter kinds. The inverse of matrix \mathbf{A}_n is quickly computable due to its small size.

Appendix D

Details on the adjoint-state method for 2D elastic vertical transverse isotropic full waveform inversion

Contents

D.1	Adjoint-state formulation	155
D.2	Adjoint operators	157
D.2.1	First adjoint operator	157
D.2.2	Second adjoint operator	158
D.3	Computation of \mathbf{d}^\dagger	159
D.4	Gradient computation	160
D.4.1	Case $C_{44}(\mathbf{x}) \neq 0$ and $C_{13}(\mathbf{x})^2 - C_{11}(\mathbf{x})C_{33}(\mathbf{x}) \neq 0$	160
D.4.2	Case $C_{44}(\mathbf{x}) = 0$	161
D.4.3	Case $C_{13}(\mathbf{x})^2 - C_{11}(\mathbf{x})C_{33}(\mathbf{x}) = 0$	161

To lighten the equations, the subscript n is omitted.

I describe here the application of the adjoint-state formulation for the full waveform inversion (FWI) gradient. I focus on the 2D elastic Vertical Transverse Isotropic (VTI) FWI. The gradient computation of FWI is performed source by source and the final gradient is obtained by summation of all individual gradients. All below demonstrations are made for a single source.

D.1 Adjoint-state formulation

Let us write the objective function of FWI \mathcal{O} as follows:

$$\begin{aligned} \mathcal{O}: \mathbb{M} &\rightarrow \mathbb{R} \\ \mathbf{m} &\mapsto \iint_S \int_{t_1}^{t_2} \chi(\mathbf{d}(\mathbf{m}; t, \mathbf{x})) dt dS = \langle \chi(\mathbf{d}(\mathbf{m}; t, \mathbf{x})) \rangle. \end{aligned} \tag{D.1}$$

\mathbb{M} is the model space, \mathbf{d} is a function returning the velocity and stress components, or the displacement components, and $\langle \rangle$ denotes the space and time integrals. With the

adjoint-state method, we are looking for $\nabla \mathcal{O} \delta \mathbf{m}$: the derivative of \mathcal{O} with respect to the model \mathbf{m} evaluated at $\delta \mathbf{m} \in \mathbb{M}$.

Using the chain rule, let us start by differentiating the objective function [D.1](#) with respect to \mathbf{m} :

$$\nabla \mathcal{O} \delta \mathbf{m} = \langle \nabla_{\mathbf{d}} \chi \nabla_{\mathbf{m}} \mathbf{d} \delta \mathbf{m} \rangle, \quad (\text{D.2})$$

$\nabla_{\mathbf{d}} \chi$ and $\nabla_{\mathbf{m}} \mathbf{d}$ are respectively the derivative of χ with respect to \mathbf{d} and the derivative of \mathbf{d} with respect to \mathbf{m} .

In a general case, the computation of $\nabla_{\mathbf{m}} \mathbf{d}$ would require a prohibitive number of forward simulations. Instead, the adjoint-state method bypasses the computation of this term.

Let us write the wave equation as:

$$\mathbf{L}(\mathbf{d}(\mathbf{m}), \mathbf{m}) = \mathbf{s}, \quad (\text{D.3})$$

with \mathbf{s} the external source. Applying the chain rule to equation [D.3](#), one can obtain the derivative of \mathbf{L} with respect to \mathbf{m} at $\delta \mathbf{m}$:

$$\nabla_{\mathbf{d}} \mathbf{L} \nabla_{\mathbf{m}} \mathbf{d} \delta \mathbf{m} + \nabla_{\mathbf{m}} \mathbf{L} \delta \mathbf{m} = \mathbf{0}. \quad (\text{D.4})$$

By taking the dot product of an arbitrary test function \mathbf{d}^* with equation [D.4](#), and then applying the integrals $\langle \cdot \rangle$, one can obtain:

$$\langle \mathbf{d}^* \cdot (\nabla_{\mathbf{d}} \mathbf{L} \nabla_{\mathbf{m}} \mathbf{d} \delta \mathbf{m}) \rangle + \langle \mathbf{d}^* \cdot (\nabla_{\mathbf{m}} \mathbf{L} \delta \mathbf{m}) \rangle = 0. \quad (\text{D.5})$$

Summing equations [D.2](#) and [D.5](#) yields:

$$\nabla \mathcal{O} \delta \mathbf{m} = \langle \nabla_{\mathbf{d}} \chi \nabla_{\mathbf{m}} \mathbf{d} \delta \mathbf{m} \rangle + \langle \mathbf{d}^* \cdot (\nabla_{\mathbf{d}} \mathbf{L} \nabla_{\mathbf{m}} \mathbf{d} \delta \mathbf{m}) \rangle + \langle \mathbf{d}^* \cdot (\nabla_{\mathbf{m}} \mathbf{L} \delta \mathbf{m}) \rangle. \quad (\text{D.6})$$

Let us define $(\nabla_{\mathbf{d}} \mathbf{L})^\dagger$ and $(\nabla_{\mathbf{d}} \chi)^\dagger$, two adjoint operators such that:

$$\langle \mathbf{d}^\diamond \cdot ((\nabla_{\mathbf{d}} \mathbf{L})^\dagger \mathbf{d}^*) \rangle = \langle \mathbf{d}^* \cdot (\nabla_{\mathbf{d}} \mathbf{L} \mathbf{d}^\diamond) \rangle, \quad (\text{D.7})$$

and

$$\langle \mathbf{d}^\diamond \cdot (\nabla_{\mathbf{d}} \chi)^\dagger \rangle = \langle \nabla_{\mathbf{d}} \chi \mathbf{d}^\diamond \rangle, \quad (\text{D.8})$$

with the notation $\mathbf{d}^\diamond = \nabla_{\mathbf{m}} \mathbf{d} \delta \mathbf{m}$ introduced to lighten the equation. Inserting [D.8](#) and [D.7](#) inside [D.6](#) yields:

$$\nabla \mathcal{O} \delta \mathbf{m} = \langle (\nabla_{\mathbf{m}} \mathbf{d} \delta \mathbf{m}) \cdot [(\nabla_{\mathbf{d}} \chi)^\dagger + (\nabla_{\mathbf{d}} \mathbf{L})^\dagger \mathbf{d}^*] \rangle + \langle \mathbf{d}^* \cdot (\nabla_{\mathbf{m}} \mathbf{L} \delta \mathbf{m}) \rangle. \quad (\text{D.9})$$

If one can find the specific test function \mathbf{d}^\dagger such that:

$$(\nabla_{\mathbf{d}} \chi)^\dagger + (\nabla_{\mathbf{d}} \mathbf{L})^\dagger \mathbf{d}^\dagger = \mathbf{0}, \quad (\text{D.10})$$

one can bypass the computation of $\nabla_{\mathbf{m}} \mathbf{d}$, and compute $\nabla \mathcal{O} \delta \mathbf{m}$ with (from equation [D.6](#)):

$$\nabla \mathcal{O} \delta \mathbf{m} = \langle \mathbf{d}^\dagger \cdot (\nabla_{\mathbf{m}} \mathbf{L} \delta \mathbf{m}) \rangle. \quad (\text{D.11})$$

D.2 Adjoint operators

In our elastic 2D VTI wave equation, the operator $L(\mathbf{d}(\mathbf{m}), \mathbf{m})$ of equation D.3 is:

$$L(\mathbf{d}(\mathbf{m}), \mathbf{m}) = \begin{pmatrix} \rho(\mathbf{x})\dot{v}_x(\mathbf{x}, t) - \sigma_{xx,x}(\mathbf{x}, t) - \sigma_{xz,z}(\mathbf{x}, t) \\ \rho(\mathbf{x})\dot{v}_z(\mathbf{x}, t) - \sigma_{xz,x}(\mathbf{x}, t) - \sigma_{zz,z}(\mathbf{x}, t) \\ \dot{\sigma}_{xx}(\mathbf{x}, t) - C_{11}(\mathbf{x})v_{x,x}(\mathbf{x}, t) - C_{13}(\mathbf{x})v_{z,z}(\mathbf{x}, t) \\ \dot{\sigma}_{zz}(\mathbf{x}, t) - C_{13}(\mathbf{x})v_{x,x}(\mathbf{x}, t) - C_{33}(\mathbf{x})v_{z,z}(\mathbf{x}, t) \\ \dot{\sigma}_{xz}(\mathbf{x}, t) - C_{44}(\mathbf{x})[v_{x,z}(\mathbf{x}, t) + v_{z,x}(\mathbf{x}, t)] \end{pmatrix}. \quad (\text{D.12a})$$

The corresponding \mathbf{d} and \mathbf{s} are:

$$\mathbf{d} = \begin{pmatrix} v_x(\mathbf{x}, t) \\ v_z(\mathbf{x}, t) \\ \sigma_{xx}(\mathbf{x}, t) \\ \sigma_{zz}(\mathbf{x}, t) \\ \sigma_{xz}(\mathbf{x}, t) \end{pmatrix} \text{ and } \mathbf{s} = \begin{pmatrix} f_x(\mathbf{x}, t) \\ f_z(\mathbf{x}, t) \\ \dot{m}_{xx}(\mathbf{x}, t) \\ \dot{m}_{zz}(\mathbf{x}, t) \\ \dot{m}_{xz}(\mathbf{x}, t) \end{pmatrix}. \quad (\text{D.12b})$$

This operator is implemented with null initial conditions and absorbing or free surface boundary conditions. See Section 1.2 of Chapter 1 for more details about this equation.

D.2.1 First adjoint operator

To lighten the equations, the time and space dependencies are omitted. We are here looking for the adjoint operator $(\nabla_{\mathbf{d}}\mathbf{L})^\dagger$ as defined in equation D.7.

By using the linearity of \mathbf{L} in \mathbf{d} , one can write:

$$\langle \mathbf{d}^* \cdot (\nabla_{\mathbf{d}}\mathbf{L}\mathbf{d}^\diamond) \rangle = \langle \mathbf{d}^* \cdot \mathbf{L}(\mathbf{d}^\diamond) \rangle, \quad (\text{D.13})$$

which leads to, after development and regrouping:

$$\begin{aligned} \langle \mathbf{d}^* \cdot (\nabla_{\mathbf{d}}\mathbf{L}\mathbf{d}^\diamond) \rangle &= \langle v_x^* \rho \dot{v}_x^\diamond + v_z^* \rho \dot{v}_z^\diamond + \sigma_{xx}^* \dot{\sigma}_{xx}^\diamond + \sigma_{xz}^* \dot{\sigma}_{xz}^\diamond + \sigma_{zz}^* \dot{\sigma}_{zz}^\diamond \rangle \\ &\quad - \langle \sigma_{xx}^* (C_{11}v_{x,x}^\diamond + C_{13}v_{z,z}^\diamond) + \sigma_{xz}^* (C_{44}[v_{x,z}^\diamond + v_{z,x}^\diamond]) + \sigma_{zz}^* (C_{13}v_{x,x}^\diamond + C_{33}v_{z,z}^\diamond) \rangle \\ &\quad - \langle v_x^* (\sigma_{xx,x}^\diamond + \sigma_{xz,z}^\diamond) + v_z^* (\sigma_{xz,x}^\diamond + \sigma_{zz,z}^\diamond) \rangle. \end{aligned} \quad (\text{D.14})$$

Integrating by part with respect to time, using the null initial conditions of \mathbf{d}^\diamond and imposing null final conditions to \mathbf{d}^* (and therefore to \mathbf{d}^\dagger) yields:

$$\langle v_x^* \rho \dot{v}_x^\diamond + v_z^* \rho \dot{v}_z^\diamond + \sigma_{xx}^* \dot{\sigma}_{xx}^\diamond + \sigma_{xz}^* \dot{\sigma}_{xz}^\diamond + \sigma_{zz}^* \dot{\sigma}_{zz}^\diamond \rangle = - \langle v_x^\diamond \rho \dot{v}_x^* + v_z^\diamond \rho \dot{v}_z^* + \sigma_{xx}^\diamond \dot{\sigma}_{xx}^* + \sigma_{xz}^\diamond \dot{\sigma}_{xz}^* + \sigma_{zz}^\diamond \dot{\sigma}_{zz}^* \rangle. \quad (\text{D.15})$$

Let us remind the following corollary of the divergence theorem:

$$\iint_S [\mathbf{F} \cdot (\nabla g) + g(\nabla \cdot \mathbf{F})] dS = \oint_B g \mathbf{F} \cdot \mathbf{n} dB, \quad (\text{D.16})$$

with g a scalar function, and \mathbf{F} a 2D vector field, both continuously differentiable on the surface S bounded by the boundary B . \mathbf{n} are the unit normal vectors to B pointing outwards.

Applying twice the corollary D.16 with:

$$\mathbf{F} = \begin{pmatrix} C_{11}\sigma_{xx}^* + C_{13}\sigma_{zz}^* \\ C_{44}\sigma_{xz}^* \end{pmatrix} \text{ and } g = v_x^\diamond, \quad (\text{D.17})$$

and

$$\mathbf{F} = \begin{pmatrix} C_{44}\sigma_{xz}^* \\ C_{13}\sigma_{xx}^* + C_{33}\sigma_{zz}^* \end{pmatrix} \text{ and } g = v_z^\diamond, \quad (\text{D.18})$$

and imposing null boundary conditions to $\mathbf{F} \cdot \mathbf{n}$ (and therefore to the corresponding quantity with \mathbf{d}^\dagger) yields:

$$-\langle \sigma_{xx}^*(C_{11}v_{x,x}^\diamond + C_{13}v_{z,z}^\diamond) + \sigma_{xz}^*(C_{44}[v_{x,z}^\diamond + v_{z,x}^\diamond]) + \sigma_{zz}^*(C_{13}v_{x,x}^\diamond + C_{33}v_{z,z}^\diamond) \rangle = \langle v_x^\diamond \left((C_{11}\sigma_{xx}^* + C_{13}\sigma_{zz}^*)_{,x} + (C_{44}\sigma_{xz})_{,z} \right) + v_z^\diamond \left((C_{44}\sigma_{xz})_{,x} + (C_{13}\sigma_{xx}^* + C_{33}\sigma_{zz}^*)_{,z} \right) \rangle. \quad (\text{D.19})$$

Again, applying three times the corollary D.16 with:

$$\mathbf{F} = \begin{pmatrix} v_x^* \\ 0 \end{pmatrix} \text{ and } g = \sigma_{xx}^\diamond, \quad (\text{D.20})$$

$$\mathbf{F} = \begin{pmatrix} 0 \\ v_z^* \end{pmatrix} \text{ and } g = \sigma_{zz}^\diamond, \quad (\text{D.21})$$

and

$$\mathbf{F} = \begin{pmatrix} v_z^* \\ v_x^* \end{pmatrix} \text{ and } g = \sigma_{xz}^\diamond, \quad (\text{D.22})$$

and using the null boundary conditions of $\mathbf{F} \cdot \mathbf{n}$ (due to free surface and absorbing boundaries) yields:

$$-\langle v_x^*(\sigma_{xx,x}^\diamond + \sigma_{xz,z}^\diamond) + v_z^*(\sigma_{xz,x}^\diamond + \sigma_{zz,z}^\diamond) \rangle = \langle \sigma_{xx}^\diamond v_{x,x}^* + \sigma_{xz}^\diamond (v_{x,z}^* + v_{z,x}^*) + \sigma_{zz}^\diamond v_{z,z}^* \rangle. \quad (\text{D.23})$$

By inserting equations D.15, D.19, D.23 into D.14, one can identify $(\nabla_{\mathbf{d}}\mathbf{L})^\dagger$ as:

$$(\nabla_{\mathbf{d}}\mathbf{L})^\dagger \mathbf{d}^* = \begin{pmatrix} -\rho \dot{v}_x^* + \left((C_{11}\sigma_{xx}^* + C_{13}\sigma_{zz}^*)_{,x} + (C_{44}\sigma_{xz}^*)_{,z} \right) \\ -\rho \dot{v}_z^* + \left((C_{44}\sigma_{xz}^*)_{,x} + (C_{13}\sigma_{xx}^* + C_{33}\sigma_{zz}^*)_{,z} \right) \\ -\dot{\sigma}_{xx}^* + v_{x,x}^* \\ -\dot{\sigma}_{zz}^* + v_{z,z}^* \\ -\dot{\sigma}_{xz}^* + [v_{x,z}^* + v_{z,x}^*] \end{pmatrix}. \quad (\text{D.24})$$

A side conclusion is that the used wave equation operator D.12a is not self-adjoint, in opposition with the same wave equation operator but formulated with displacement only. Consequently, the self-adjointness of a wave equation operator depends notably on the used formulation, and not only on the physical phenomena it describes.

D.2.2 Second adjoint operator

We are looking for the adjoint operator $(\nabla_{\mathbf{d}}\chi)^\dagger$ as defined in equation D.8.

With our objective function (equation 1.11) and for a single source src, χ is:

$$\chi = \frac{1}{2} \sum_{\text{cmp}} \left[w_{\text{cmp},\text{src}}(t, \mathbf{x}) \left(d_{\text{cmp},\text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - d_{\text{cmp},\text{src}}^{\text{obs}}(t, \mathbf{x}) \right)^2 \right]. \quad (\text{D.25})$$

$d_{\text{cmp},\text{src}}^{\text{mod}}$ and $d_{\text{cmp},\text{src}}^{\text{obs}}$ are respectively the modeled and observed wavefields for the source src and the component cmp. The sum is over the components (cmp), the components being the pressure P (i.e., the mean of diagonal stresses: $\frac{\sigma_{xx} + \sigma_{zz}}{2}$) and the velocity components (v_x, v_z) . $w_{\text{cmp},\text{src}}(t, \mathbf{x})$ is introduced to:

- account for the fact that the observed (and therefore modeled) data are discrete in space and time;
- represent the (tapered) time selection and (tapered) receiver selection (e.g., offset selection);
- represent the component selection: our algorithm allows for the pressure component, both velocity components or one of the velocity components.

Note that $d_{\text{cmp,src}}^{\text{mod}}$ and $d_{\text{cmp,src}}^{\text{obs}}$ are not the same as \mathbf{d} .

One can write:

$$\begin{aligned}
\langle \nabla_{\mathbf{d}} \chi \mathbf{d}^\diamond \rangle &= w_{v_x, \text{src}}(t, \mathbf{x}) \left(v_{x, \text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - v_{x, \text{src}}^{\text{obs}}(t, \mathbf{x}) \right) v_x^\diamond(t, \mathbf{x}) \\
&+ w_{v_z, \text{src}}(t, \mathbf{x}) \left(v_{z, \text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - v_{z, \text{src}}^{\text{obs}}(t, \mathbf{x}) \right) v_z^\diamond(t, \mathbf{x}) \\
&+ \frac{w_{P, \text{src}}(t, \mathbf{x})}{2} \left(P_{\text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - P_{\text{src}}^{\text{obs}}(t, \mathbf{x}) \right) \sigma_{xx}^\diamond(t, \mathbf{x}) \\
&+ \frac{w_{P, \text{src}}(t, \mathbf{x})}{2} \left(P_{\text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - P_{\text{src}}^{\text{obs}}(t, \mathbf{x}) \right) \sigma_{zz}^\diamond(t, \mathbf{x}).
\end{aligned} \tag{D.26}$$

One can now identify $(\nabla_{\mathbf{d}} \chi)^\dagger$ as:

$$(\nabla_{\mathbf{d}} \chi)^\dagger = \begin{pmatrix} w_{v_x, \text{src}}(t, \mathbf{x}) \left(v_{x, \text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - v_{x, \text{src}}^{\text{obs}}(t, \mathbf{x}) \right) \\ w_{v_z, \text{src}}(t, \mathbf{x}) \left(v_{z, \text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - v_{z, \text{src}}^{\text{obs}}(t, \mathbf{x}) \right) \\ \frac{w_{P, \text{src}}(t, \mathbf{x})}{2} \left(P_{\text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - P_{\text{src}}^{\text{obs}}(t, \mathbf{x}) \right) \\ \frac{w_{P, \text{src}}(t, \mathbf{x})}{2} \left(P_{\text{src}}^{\text{mod}}(\mathbf{m}; t, \mathbf{x}) - P_{\text{src}}^{\text{obs}}(t, \mathbf{x}) \right) \\ 0 \end{pmatrix}. \tag{D.27}$$

D.3 Computation of \mathbf{d}^\dagger

To lighten the equations, the space and model dependencies are omitted. Computing \mathbf{d}^\dagger from equation D.10 with the derived $(\nabla_{\mathbf{d}} \mathbf{L})^\dagger$ and $(\nabla_{\mathbf{d}} \chi)^\dagger$ exhibits three difficulties:

- The implementation of the required null final conditions for \mathbf{d}^\dagger is not straightforward;
- The implementation of the required null boundary conditions for \mathbf{d}^\dagger is not straightforward;
- The structure of equation D.10 is nothing like something already implemented and consequently needs further implementation work.

To solve these difficulties, a new set of variables is introduced:

$$\begin{cases} \overline{v_x^\dagger}(t) = -v_x^\dagger(t_2 - t), \\ \overline{v_z^\dagger}(t) = -v_z^\dagger(t_2 - t), \\ \overline{\sigma_{xx}^\dagger}(t) = C_{11} \sigma_{xx}^\dagger(t_2 - t) + C_{13} \sigma_{zz}^\dagger(t_2 - t), \\ \overline{\sigma_{zz}^\dagger}(t) = C_{13} \sigma_{xx}^\dagger(t_2 - t) + C_{33} \sigma_{zz}^\dagger(t_2 - t), \\ \overline{\sigma_{xz}^\dagger}(t) = C_{44} \sigma_{xz}^\dagger(t_2 - t). \end{cases} \tag{D.28}$$

Inserting the two adjoint operators [D.24](#) and [D.27](#) into equation [D.10](#), and applying the change of variable yields:

$$\begin{cases} \dot{\overline{\rho v_x^\dagger}} = \overline{\sigma_{xx,x}^\dagger} + \overline{\sigma_{xz,z}^\dagger} + w_{v_x,\text{src}}(t, \mathbf{x}) \left(v_{x,\text{src}}^{\text{mod}}(t_2 - t) - v_{x,\text{src}}^{\text{obs}}(t_2 - t) \right), \\ \dot{\overline{\rho v_z^\dagger}} = \overline{\sigma_{xz,x}^\dagger} + \overline{\sigma_{zz,z}^\dagger} + w_{v_z,\text{src}}(t, \mathbf{x}) \left(v_{z,\text{src}}^{\text{mod}}(t_2 - t) - v_{z,\text{src}}^{\text{obs}}(t_2 - t) \right), \\ \dot{\overline{\sigma_{xx}^\dagger}} = C_{11} \overline{v_{x,x}^\dagger} + C_{13} \overline{v_{z,z}^\dagger} - \left(C_{11} + C_{13} \right) \frac{w_{P,\text{src}}(t_2-t)}{2} \left(P_{\text{src}}^{\text{mod}}(t_2 - t) - P_{\text{src}}^{\text{obs}}(t_2 - t) \right), \\ \dot{\overline{\sigma_{zz}^\dagger}} = C_{13} \overline{v_{x,x}^\dagger} + C_{33} \overline{v_{z,z}^\dagger} - \left(C_{13} + C_{33} \right) \frac{w_{P,\text{src}}(t_2-t)}{2} \left(P_{\text{src}}^{\text{mod}}(t_2 - t) - P_{\text{src}}^{\text{obs}}(t_2 - t) \right), \\ \dot{\overline{\sigma_{xz}^\dagger}} = C_{44} [\overline{v_{x,z}^\dagger} + \overline{v_{z,x}^\dagger}]. \end{cases} \quad (\text{D.29})$$

One can notice from equation [D.29](#) that $\overline{\mathbf{d}^\dagger} = (\overline{v_x^\dagger}, \overline{v_z^\dagger}, \overline{\sigma_{xx}^\dagger}, \overline{\sigma_{zz}^\dagger}, \overline{\sigma_{xz}^\dagger})$ (called the ‘‘adjoint wavefield’’) corresponds to the solution of the already implemented wave equation operator [D.12a](#) with a source \mathbf{s}^\dagger (called the ‘‘adjoint source’’) such that:

$$\mathbf{s}^\dagger = \begin{pmatrix} w_{v_x,\text{src}}(t, \mathbf{x}) \left(v_{x,\text{src}}^{\text{mod}}(t_2 - t) - v_{x,\text{src}}^{\text{obs}}(t_2 - t) \right) \\ w_{v_z,\text{src}}(t, \mathbf{x}) \left(v_{z,\text{src}}^{\text{mod}}(t_2 - t) - v_{z,\text{src}}^{\text{obs}}(t_2 - t) \right) \\ - \left(C_{11} + C_{13} \right) \frac{w_{P,\text{src}}(t_2-t)}{2} \left(P_{\text{src}}^{\text{mod}}(t_2 - t) - P_{\text{src}}^{\text{obs}}(t_2 - t) \right) \\ - \left(C_{13} + C_{33} \right) \frac{w_{P,\text{src}}(t_2-t)}{2} \left(P_{\text{src}}^{\text{mod}}(t_2 - t) - P_{\text{src}}^{\text{obs}}(t_2 - t) \right) \\ 0 \end{pmatrix}. \quad (\text{D.30})$$

Indeed, with this change of variable, the required null final and null boundary conditions translate into null initial and other null boundary conditions that are met with the already implemented wave equation, notably through the absorbing and free surface boundary conditions. $\overline{\mathbf{d}^\dagger}$ is then called the adjoint wavefield. Consequently, \mathbf{d}^\dagger can be computed by: first solving the wave equation [D.29](#) for $\overline{\mathbf{d}^\dagger}$; second applying the change of variable [D.28](#) (done in the next section).

D.4 Gradient computation

Let us consider a starting model $(C_{11}(\mathbf{x}), C_{13}(\mathbf{x}), C_{33}(\mathbf{x}), C_{44}(\mathbf{x}), \rho(\mathbf{x}))$. Let us call $(v_x(t, \mathbf{x}), v_z(t, \mathbf{x}), \sigma_{xx}(t, \mathbf{x}), \sigma_{zz}(t, \mathbf{x}), \sigma_{xz}(t, \mathbf{x}))$ the corresponding modeled forward wavefield. The gradient with respect to each model point and each parameter type is computed by using equation [D.11](#), where $\delta \mathbf{m}$ is chosen to focus on the selected model point and parameter type (hence the disappearance of the volume integrals in the following equations).

D.4.1 Case $C_{44}(\mathbf{x}) \neq 0$ and $C_{13}(\mathbf{x})^2 - C_{11}(\mathbf{x})C_{33}(\mathbf{x}) \neq 0$

Note that:

$$\begin{cases} v_{x,x}(t, \mathbf{x}) = \frac{C_{33}(\mathbf{x})\dot{\sigma}_{xx}(t,\mathbf{x}) - C_{13}(\mathbf{x})\dot{\sigma}_{zz}(t,\mathbf{x})}{C_{33}(\mathbf{x})C_{11}(\mathbf{x}) - C_{13}(\mathbf{x})^2}, \\ v_{z,z}(t, \mathbf{x}) = \frac{C_{11}(\mathbf{x})\dot{\sigma}_{zz}(t,\mathbf{x}) - C_{13}(\mathbf{x})\dot{\sigma}_{xx}(t,\mathbf{x})}{C_{33}(\mathbf{x})C_{11}(\mathbf{x}) - C_{13}(\mathbf{x})^2}, \\ v_{x,z}(t, \mathbf{x}) + v_{z,x}(t, \mathbf{x}) = \frac{\dot{\sigma}_{xz}(t,\mathbf{x})}{C_{44}(\mathbf{x})}, \end{cases} \quad (\text{D.31})$$

where dS is the area of the grid cell. Applying equation [D.11](#) for the 2D elastic VTI wave equation operator [D.12a](#) and using equations [D.31](#) to bypass the spatial derivative

computation, the gradient in terms of each elastic moduli and density point is:

$$\left\{ \begin{array}{l} \frac{\partial \theta}{\partial C_{11}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{(C_{33}(\mathbf{x})\dot{\sigma}_{xx}(t, \mathbf{x}) - C_{13}(\mathbf{x})\dot{\sigma}_{zz}(t, \mathbf{x}))\sigma_{xx}^\dagger(t, \mathbf{x})}{C_{33}(\mathbf{x})C_{11}(\mathbf{x}) - C_{13}(\mathbf{x})^2} dt, \\ \frac{\partial \theta}{\partial C_{13}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{(C_{11}(\mathbf{x})\dot{\sigma}_{zz}(t, \mathbf{x}) - C_{13}(\mathbf{x})\dot{\sigma}_{xx}(t, \mathbf{x}))\sigma_{xx}^\dagger(t, \mathbf{x}) + (C_{33}(\mathbf{x})\dot{\sigma}_{xx}(t, \mathbf{x}) - C_{13}(\mathbf{x})\dot{\sigma}_{zz}(t, \mathbf{x}))\sigma_{zz}^\dagger(t, \mathbf{x})}{C_{33}(\mathbf{x})C_{11}(\mathbf{x}) - C_{13}(\mathbf{x})^2} dt, \\ \frac{\partial \theta}{\partial C_{33}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{(C_{11}(\mathbf{x})\dot{\sigma}_{zz}(t, \mathbf{x}) - C_{13}(\mathbf{x})\dot{\sigma}_{xx}(t, \mathbf{x}))\sigma_{zz}^\dagger(t, \mathbf{x})}{C_{33}(\mathbf{x})C_{11}(\mathbf{x}) - C_{13}(\mathbf{x})^2} dt, \\ \frac{\partial \theta}{\partial C_{44}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{\dot{\sigma}_{xz}(t, \mathbf{x})\sigma_{xz}^\dagger(t, \mathbf{x})}{C_{44}(\mathbf{x})} dt, \\ \frac{\partial \theta}{\partial \rho(\mathbf{x})} = dS \int_{t_1}^{t_2} (\dot{v}_x(t, \mathbf{x})v_x^\dagger(t, \mathbf{x}) + \dot{v}_z(t, \mathbf{x})v_z^\dagger(t, \mathbf{x})) dt. \end{array} \right. \quad (\text{D.32})$$

To replace into D.32 the unavailable \mathbf{d}^\dagger by the available $\overline{\mathbf{d}^\dagger}$, equation D.28 is rearranged (omitting spatial dependencies):

$$\left\{ \begin{array}{l} \sigma_{xx}^\dagger(t) = \frac{C_{33}\overline{\sigma_{xx}^\dagger}(t_2-t) - C_{13}\overline{\sigma_{zz}^\dagger}(t_2-t)}{C_{33}C_{11} - C_{13}^2}, \\ \sigma_{zz}^\dagger(t) = \frac{C_{11}\overline{\sigma_{zz}^\dagger}(t_2-t) - C_{13}\overline{\sigma_{xx}^\dagger}(t_2-t)}{C_{33}C_{11} - C_{13}^2}, \\ \sigma_{xz}^\dagger(t) = \frac{\overline{\sigma_{xz}^\dagger}(t_2-t)}{C_{44}}, \\ v_x^\dagger(t) = -v_x^\dagger(t_2-t), \\ v_z^\dagger(t) = -v_z^\dagger(t_2-t). \end{array} \right. \quad (\text{D.33})$$

Inserting D.33 into D.32 yields (partially omitting spatial dependencies):

$$\left\{ \begin{array}{l} \frac{\partial \theta}{\partial C_{11}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{(C_{33}\dot{\sigma}_{xx}(t) - C_{13}\dot{\sigma}_{zz}(t))(C_{33}\overline{\sigma_{xx}^\dagger}(t_2-t) - C_{13}\overline{\sigma_{zz}^\dagger}(t_2-t))}{(C_{13}^2 - C_{11}C_{33})^2} dt, \\ \frac{\partial \theta}{\partial C_{13}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{(C_{13}^2 + C_{11}C_{33})(\dot{\sigma}_{xx}(t)\overline{\sigma_{zz}^\dagger}(t_2-t) + \dot{\sigma}_{zz}(t)\overline{\sigma_{xx}^\dagger}(t_2-t)) - 2C_{13}(C_{11}\dot{\sigma}_{zz}(t)\overline{\sigma_{zz}^\dagger}(t_2-t) + C_{33}\dot{\sigma}_{xx}(t)\overline{\sigma_{xx}^\dagger}(t_2-t))}{(C_{13}^2 - C_{11}C_{33})^2} dt, \\ \frac{\partial \theta}{\partial C_{33}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{(C_{11}\dot{\sigma}_{zz}(t) - C_{13}\dot{\sigma}_{xx}(t))(C_{11}\overline{\sigma_{zz}^\dagger}(t_2-t) - C_{13}\overline{\sigma_{xx}^\dagger}(t_2-t))}{(C_{13}^2 - C_{11}C_{33})^2} dt, \\ \frac{\partial \theta}{\partial C_{44}(\mathbf{x})} = -dS \int_{t_1}^{t_2} \frac{\dot{\sigma}_{xz}(t)\overline{\sigma_{xz}^\dagger}(t_2-t)}{C_{44}^2} dt, \\ \frac{\partial \theta}{\partial \rho(\mathbf{x})} = -dS \int_{t_1}^{t_2} (\dot{v}_x(t)v_x^\dagger(t_2-t) + \dot{v}_z(t)v_z^\dagger(t_2-t)) dt. \end{array} \right. \quad (\text{D.34})$$

D.4.2 Case $C_{44}(\mathbf{x}) = 0$

Because no division by 0 is possible, this case imposes a specific treatment for the computation of $\frac{\partial \theta}{\partial C_{44}(\mathbf{x})}$: the medium is then an acoustic medium and the FWI algorithm considers that no update should be performed on $C_{44}(\mathbf{x})$ (and consequently on S-wave velocities), leading to

$$\frac{\partial \theta}{\partial C_{44}(\mathbf{x})} = 0. \quad (\text{D.35})$$

D.4.3 Case $C_{13}(\mathbf{x})^2 - C_{11}(\mathbf{x})C_{33}(\mathbf{x}) = 0$

Similarly, in such a case, a specific treatment is required for the computation of $\frac{\partial \theta}{\partial C_{11}(\mathbf{x})}$, $\frac{\partial \theta}{\partial C_{13}(\mathbf{x})}$ and $\frac{\partial \theta}{\partial C_{33}(\mathbf{x})}$: the computations are performed replacing equations D.31 and D.33 by their limits when $C_{13}(\mathbf{x})^2 - C_{11}(\mathbf{x})C_{33}(\mathbf{x})$ tends towards 0.

In practice, the equivalent limits when $C_{13}(\mathbf{x})$ tends toward $\sqrt{C_{11}(\mathbf{x})C_{33}(\mathbf{x})}$ are computed instead ($C_{13} \geq 0$). Using L'Hôpital's rule and noticing that, in such a case and

in the absence of moment density sources at \mathbf{x} , $\dot{\sigma}_{zz} = \frac{\sqrt{C_{33}}}{\sqrt{C_{11}}} \dot{\sigma}_{xx}$ and $\overline{\sigma_{zz}^\dagger} = \frac{\sqrt{C_{33}}}{\sqrt{C_{11}}} \overline{\sigma_{xx}^\dagger}$, one can obtain:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial C_{11}(\mathbf{x})} &= -dS \int_{t_1}^{t_2} \frac{\dot{\sigma}_{xx} \overline{\sigma_{xx}^\dagger}}{4C_{11}C_{11}} dt, \\ \frac{\partial \mathcal{L}}{\partial C_{13}(\mathbf{x})} &= -dS \int_{t_1}^{t_2} \frac{\dot{\sigma}_{xx} \overline{\sigma_{xx}^\dagger}}{2C_{11}C_{13}} dt, \\ \frac{\partial \mathcal{L}}{\partial C_{33}(\mathbf{x})} &= -dS \int_{t_1}^{t_2} \frac{\dot{\sigma}_{xx} \overline{\sigma_{xx}^\dagger}}{4C_{11}C_{33}} dt. \end{cases} \quad (\text{D.36})$$

Regarding the practical implementation, whatever the case, our FWI algorithm uses Simpson's rule for the gradient time integrations. The gradients computations need the simultaneous availability of the time derivative of the forward and the adjoint wavefields at different time steps. To this end, two approaches are possible: either the forward wavefield is stored on disk during the prior modeling, or it is reconstructed on the fly during the adjoint wavefield computation through a reverse time approach. The first approach has a high memory cost. The second approach has consequently been implemented, following Tromp et al. (2005) (194).

Another way of explaining the gradient computation with the adjoint-state method and the implementation of our FWI algorithm can be found in He's Ph.D. thesis (2017) (80). Please, note that compared to He's Ph.D. thesis, the cost function has been modified: the inversion with both stress components has been replaced by an inversion with pressure component. For pressure measured in water (i.e., an isotropic acoustic medium), these two types of inversion are in practice equivalent, but theoretically slightly different. Pressure component is preferred since it directly corresponds to a recorded quantity.

Appendix E

Convolutional type reciprocity relation derivation in elastic anisotropic media

I derive here the convolutional type reciprocity relation for elastic anisotropic media.

Let us consider two body force sources \mathbf{f}^A and \mathbf{f}^B , respectively leading through wave equation 2.1 to the seismic quantities $\{\mathbf{u}^A, \boldsymbol{\sigma}^A\}$ and $\{\mathbf{u}^B, \boldsymbol{\sigma}^B\}$:

$$\begin{cases} \rho(\mathbf{x})\ddot{u}_i^A(\mathbf{x}, t) &= \sigma_{ij,j}^A(\mathbf{x}, t) + f_i^A(\mathbf{x}, t), \\ \rho(\mathbf{x})\ddot{u}_i^B(\mathbf{x}, t) &= \sigma_{ij,j}^B(\mathbf{x}, t) + f_i^B(\mathbf{x}, t). \end{cases} \quad (\text{E.1})$$

By convolving over time t first line of equation E.1 by $u_i^B(\mathbf{x}, t)$, second line by $u_i^A(\mathbf{x}, t)$, taking the difference and summing for all values of subscript i (using Einstein convention of summing on repeated indices), one can obtain:

$$\begin{aligned} \rho(\mathbf{x})[\ddot{u}_i^A(\mathbf{x}, t) *_{t} u_i^B(\mathbf{x}, t) - \ddot{u}_i^B(\mathbf{x}, t) *_{t} u_i^A(\mathbf{x}, t)] &= \sigma_{ij,j}^A(\mathbf{x}, t) *_{t} u_i^B(\mathbf{x}, t) - \sigma_{ij,j}^B(\mathbf{x}, t) *_{t} u_i^A(\mathbf{x}, t) \\ &\quad + f_i^A(\mathbf{x}, t) *_{t} u_i^B(\mathbf{x}, t) - f_i^B(\mathbf{x}, t) *_{t} u_i^A(\mathbf{x}, t). \end{aligned} \quad (\text{E.2})$$

Because of the initial and/or final conditions for displacement and its partial time derivative:

$$\begin{aligned} \ddot{u}_i^A(t) *_{t} u_i^B(t) - \ddot{u}_i^B(t) *_{t} u_i^A(t) &= \ddot{u}_i^A(t) *_{t} u_i^B(t) - u_i^B(t) *_{t} \ddot{u}_i^A(t) \\ &= 0. \end{aligned} \quad (\text{E.3})$$

By combining equations E.2 and E.3:

$$f_i^B(\mathbf{x}, t) *_{t} u_i^A(\mathbf{x}, t) - f_i^A(\mathbf{x}, t) *_{t} u_i^B(\mathbf{x}, t) = \sigma_{ij,j}^A(\mathbf{x}, t) *_{t} u_i^B(\mathbf{x}, t) - \sigma_{ij,j}^B(\mathbf{x}, t) *_{t} u_i^A(\mathbf{x}, t). \quad (\text{E.4})$$

Let us consider a volume V bounded by a surface S , by integrating equation E.4 over volume V :

$$\begin{aligned} \iiint_V \left(f_i^B(\mathbf{x}, t) *_{t} u_i^A(\mathbf{x}, t) - f_i^A(\mathbf{x}, t) *_{t} u_i^B(\mathbf{x}, t) \right) dV &= \\ \iiint_V \left(\sigma_{ij,j}^A(\mathbf{x}, t) *_{t} u_i^B(\mathbf{x}, t) - \sigma_{ij,j}^B(\mathbf{x}, t) *_{t} u_i^A(\mathbf{x}, t) \right) dV. \end{aligned} \quad (\text{E.5})$$

Due to the symmetry of the elasticity tensor $C_{ijkl} = C_{klij}$, the right hand side of equation E.5 is the divergence of a 3-component vector \mathbf{m} :

$$\sigma_{ij,j}^A(\mathbf{x}, t) *_{\dot{t}} u_i^B(\mathbf{x}, t) - \sigma_{ij,j}^B(\mathbf{x}, t) *_{\dot{t}} u_i^A(\mathbf{x}, t) = \text{div}(\mathbf{m}), \quad (\text{E.6})$$

with:

$$m_j = \sigma_{ij}^A(\mathbf{x}, t) *_{\dot{t}} u_i^B(\mathbf{x}, t) - \sigma_{ij}^B(\mathbf{x}, t) *_{\dot{t}} u_i^A(\mathbf{x}, t). \quad (\text{E.7})$$

By applying Gauss's theorem on the right hand side of equation E.5, we obtain the convolutional type reciprocity relation for elastic anisotropic media:

$$\begin{aligned} \iiint_V \left(f_i^B(\mathbf{x}, t) *_{\dot{t}} u_i^A(\mathbf{x}, t) - f_i^A(\mathbf{x}, t) *_{\dot{t}} u_i^B(\mathbf{x}, t) \right) dV = & \iint_S \left(\sigma_{ij}^A(\mathbf{x}, t) *_{\dot{t}} u_i^B(\mathbf{x}, t) n_j(\mathbf{x}) \right. \\ & \left. - \sigma_{ij}^B(\mathbf{x}, t) *_{\dot{t}} u_i^A(\mathbf{x}, t) n_j(\mathbf{x}) \right) dS, \end{aligned}$$

\mathbf{n} being the normal to the surface S , pointing outwards.

Appendix F

A few elastic vertical transverse isotropic parameters

The meaning of the parameters used in [Chapter 3](#) is provided below.

Within the framework of 3D Vertical Transverse Isotropy ([VTI](#)), Thomsen (1986) ([192](#)) defines new parameters, including: V_{pv} , V_{sv} , ϵ , δ^* and γ . V_{pv} and V_{sv} are the P- and SV-wave vertical velocities respectively (both SV and SH-waves have the same vertical velocity in [VTI](#) media). Let us call θ the angle of the direction of propagation measured from a vertical axis. ϵ , δ^* and γ describe the variation of P-, SV- and SH-waves velocities (V_p , V_{sV} , V_{sH} respectively) with the angle θ , given the vertical velocities (Thomsen, 1986 ([192](#))):

$$\begin{aligned} V_p^2(\theta) &= V_{pv}^2 \left[1 + \epsilon \sin^2 \theta + D^*(\theta) \right], \\ V_{sV}^2(\theta) &= V_{sv}^2 \left[1 + \frac{V_{pv}^2}{V_{sv}^2} \epsilon \sin^2(\theta) - \frac{V_{pv}^2}{V_{sv}^2} D^*(\theta) \right], \\ V_{sH}^2(\theta) &= V_{sv}^2 \left[1 + 2\gamma \sin^2(\theta) \right], \end{aligned} \quad (\text{F.1})$$

with

$$D^*(\theta) = \frac{1}{2} \left(1 - \frac{V_{sv}^2}{V_{pv}^2} \right) \left\{ \left[1 + \frac{4\delta^*}{\left(1 - \frac{V_{sv}^2}{V_{pv}^2} \right)^2} \sin^2(\theta) \cos^2(\theta) + \frac{4\epsilon \left(1 - \frac{V_{sv}^2}{V_{pv}^2} + \epsilon \right)}{\left(1 - \frac{V_{sv}^2}{V_{pv}^2} \right)^2} \sin^4(\theta) \right]^{0.5} - 1 \right\}. \quad (\text{F.2})$$

From equation [F.1](#), one can define the P- and S-wave horizontal velocities. The S-wave vertical and horizontal velocities are equal and the P-wave horizontal velocity V_{ph} is given by:

$$V_{ph}^2 = V_{pv}^2 (1 + 2\epsilon). \quad (\text{F.3})$$

Only ϵ controls the difference between the P-wave horizontal and vertical velocities, and a negative value leads to a P-wave horizontal velocity lower than the vertical one, and vice versa. Given P-wave vertical and horizontal velocities, δ^* controls the variation of P-wave velocity with angle. Both ϵ and δ^* control the variation of SV-wave velocity with angle, while only γ controls the variation of SH-wave velocity with angle. The relationships

between these parameters and the elastic moduli and density $C_{11}, C_{13}, C_{33}, C_{44}, C_{66}, \rho$ are (Thomsen, 1986 (192) and Alkhalifah and Tsvankin, 1995 (4)):

$$\begin{aligned}
V_{pv} &= \sqrt{\frac{C_{33}}{\rho}}, \\
V_{ph} &= \sqrt{\frac{C_{11}}{\rho}}, \\
V_{sv} &= \sqrt{\frac{C_{44}}{\rho}}, \\
\epsilon &= \frac{C_{11} - C_{33}}{2C_{33}}, \\
\delta^* &= \frac{1}{2C_{33}^2} \left[2(C_{13} + C_{44})^2 - (C_{33} - C_{44})(C_{11} + C_{33} - 2C_{44}) \right], \\
\delta &= \frac{(C_{13} + C_{44})^2 - (C_{33} + C_{44})^2}{2C_{33}(C_{33} - C_{44})}, \\
\gamma &= \frac{C_{66} - C_{44}}{2C_{44}}.
\end{aligned} \tag{F.4}$$

Within the framework of weak anisotropy approximation, the variation of the waves velocities with θ simplifies to (Thomsen, 1986 (192)):

$$\begin{aligned}
V_p(\theta) &\approx V_{pv} \left[1 + \delta \sin^2(\theta) \cos^2(\theta) + \epsilon \sin^4 \theta \right], \\
V_{sv}(\theta) &\approx V_{sv} \left[1 + \frac{V_{pv}^2}{V_{sv}^2} (\epsilon - \delta) \sin^2(\theta) \cos^2(\theta) \right], \\
V_{sH}(\theta) &\approx V_{sv} \left[1 + \gamma \sin^2(\theta) \right],
\end{aligned} \tag{F.5}$$

with

$$\delta = \frac{1}{2} \left[\epsilon + \frac{\delta^*}{1 - \frac{V_{sv}^2}{V_{pv}^2}} \right]. \tag{F.6}$$

Similarly to the non-weak equation F.1, given a P-wave vertical and horizontal velocity, δ controls the variation of P-wave velocity with angle.

Two other parameters, the normal move-out P-wave velocity V_{pn} , and the anelliptic parameter η , have been introduced based on the expression of the short offset reflection travel time in a horizontal, homogeneous, anisotropic layer (Thomsen, 1986 (192) and Alkhalifah and Tsvankin, 1995 (4)), with:

$$\begin{aligned}
V_{pn}^2 &= V_{pv}^2(1 + 2\delta), \\
\eta &= \frac{\epsilon - \delta}{1 + 2\delta}, \\
V_{ph}^2 &= V_{pn}^2(1 + 2\eta) = V_{pv}^2(1 + 2\epsilon).
\end{aligned} \tag{F.7}$$

Because η is a function of ϵ and δ , it is a parameter that participates in the control the variation of P- and SV-wave velocity with angle, given P- and SV-wave vertical velocities.

In order to have a more direct link between the various P-wave velocities, following parameters were defined by He and Plessix (2016) (78): $\sqrt{1 + 2\epsilon}$, $\sqrt{1 + 2\delta}$ and $\sqrt{1 + 2\eta}$.

Appendix G

Alwyn North field history: an educational case

Alwyn North field could be an educational case of the life of a hydrocarbon (HC) field. For the simplest cases, five main phases can be distinguished in the typical life of a HC field: exploration, evaluation of economic interest, potentially leading to development, production and abandonment. It can take 15-30 years from the beginning of production to abandonment, even more than 50 years for the largest deposits. The life of a HC field grows in complexity with the evolution of the environment, e.g.: new technologies, neighbouring discoveries or changing economical environment. Together with other disciplines, seismic imaging plays an important role in the life of a field. Alwyn North field history is a relevant example of a complex history and of the role of seismic imaging.

Alwyn North is owned and operated by TOTAL. Its infrastructure is the hub of the other Great Alwyn area fields operated by TOTAL. Exploration licenses were first granted in 1965, drilling began in 1971. Alwyn hub includes the following producing fields (map in figure G.1):

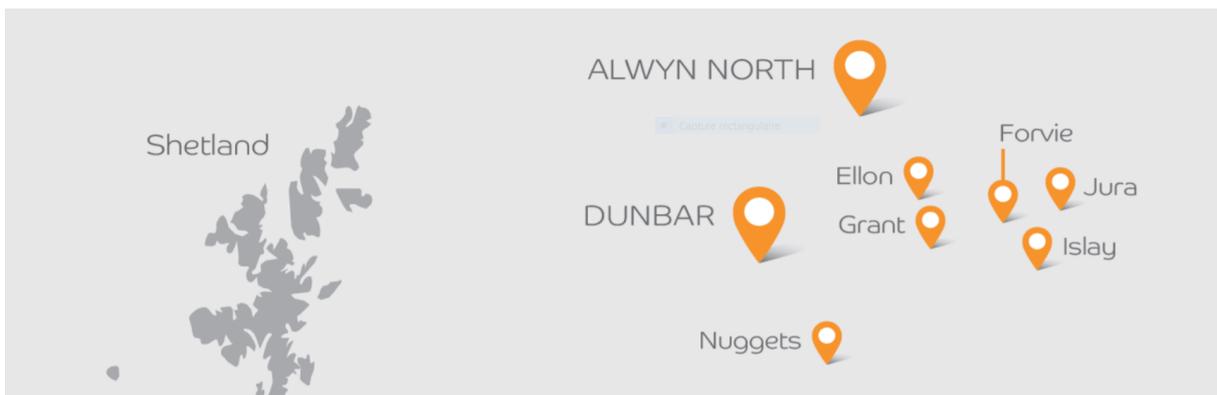


Figure G.1: Alwyn North field in the Great Alwyn area. Source: TOTAL public communication (oil & gas UK share fair 31 october 2018).

- **Dunbar and Ellon**

Those oil and gas fields were discovered in 1973, but were not economic at the time. Production started in 1994, thanks to Alwyn's infrastructure.

- **Alwyn North**

Discovered in 1975, it was more viable because of its greater size. However, because of its complex geological structure and despite considerable amounts of evaluation, a decision to develop could not be taken. By 1980, the arrival of new technologies facilitated a 3D seismic survey that provided the confidence to develop. Production from Statfjord and Brent reservoirs began in 1987. In 1993, a discovery was made in the deeper Triassic reservoir (gas). Triassic gas condensate started being recovered in 1995. Seismic reprocessing has enabled successful targeting of undeveloped portions of Triassic and Brent reservoirs. Table G.1 describes local seismic imaging efforts made before the acquisition of Alwyn real dataset.

Acquisition	Processing
1981 (3D)	2004: 4D processing
	1996/97
	1998
1995 (3D)	2004: 4D processing
	2002: PSTM
	2003: PSDM Kirchhoff
	2004: 4D processing + acoustic inversion
	2006: PSTM
	2007: PSDM Kirchhoff + acoustic inversion
	2008: acoustic inversion
2001 (3D)	2010: PSDM Beam

Table G.1: Prior seismic imaging efforts made above Alwyn North field. **PSTM** = Pre-Stack Time Migration. **PSDM** = Pre-Stack Depth Migration.

- **Grant**

Grant was discovered in 1977, but was not economic at the time. It started producing gas condensate in 1998, thanks to Dunbar infrastructure.

- **Nuggets**

Nuggets was discovered between 1972 and 1991, in four separate gas-bearing accumulations (N1, N2, N3 and N4), which were small in terms of recoverable reserves. Improvements to the gas-handling plant on Alwyn North in 1999 increased processing capacity, enabling the development of the Nuggets field. Production from N1, N2 and N3 began in November 2001. N4 came onstream in October 2003.

- **Forvie North**

Discovered in 2002, Forvie North is connected to Alwyn North infrastructure. Production started in 2005. Forvie North field has shown that considerable reserves remain to be discovered in the Alwyn area and have a significant part to play in extending the field life of the Alwyn North facilities.

- **Jura**

This gas and condensate field was discovered in 2006. Its discovery was exceptional in a mature zone and made possible by the use of state-of-the-art technology to interpret seismic data. Jura is tied back to Forvie. Production started in 2008.

- **Islay**

This gas and condensate field was discovered in 2008. Technological advancement helped to more accurately identify exploration targets. Isley is tied back to Forvie subsea manifold. Production started in 2012.

Designed with an original production life of up to 25 years, in 2014 the Alwyn platform was looking ahead to a further 20 years of production notably thanks to successful development of nearby fields in recent years. Alwyn North field's role as a hub hints how an old and mature field can still afford the luxury of new seismic acquisitions and production efforts, the production facilities being partly already available. Alwyn North field history also illustrates the key role of seismic imaging in the life of a field. This rich history results in a generous amount of knowledge and data about the area.

Appendix H

Source time function inversion: theory and synthetic applications

Contents

H.1 Theory	172
H.1.1 2D elastic vertical transverse isotropic wave equation	172
H.1.2 Source time function inversion	172
H.2 Synthetic applications on a modified Marmousi2 model . . .	173
H.2.1 An ideal application	173
H.2.2 A more realistic application	175

I describe here the theory of the used source time function estimation technique (Song et al., 1995 (*180*)), in the context of 2D Vertical Transverse Isotropic (VTI) wave propagation. The source time function is inverted from pressure data recorded in an acoustic isotropic medium and coming from an airgun (array) source modeled by a point moment density source with no directivity. I remind that a point source can be divided into two parts: the known location and the estimated source time function.

The estimation technique described here uses the fact that the pressure data are the combination of the expression at the known receivers location of: (1) the source function of the airgun (array), and (2) the wave propagation through the sub-surface. The observed pressure data are modeled using an arbitrary source time function and the model of the sub-surface. The expression of the wave propagation through the sub-surface (i.e., point (2)) is extracted from these modeled data. This extraction assumes a correct model of the sub-surface. The only remaining unknown ingredient of the observed pressure data is then the source time function, allowing for its estimation (via mathematical inversion).

After describing the theory, this technique is tested on synthetic applications using a modified Marmousi2 model (Martin et al., 2006 (*115*)), and an acquisition which mimics Alwyn real data acquisition. These tests show that this technique can lead to acceptable results even if one does not have a good model of the sub-surface, as long as at least one part of the data can be modeled correctly and separated from the rest of the data.

H.1 Theory

H.1.1 2D elastic vertical transverse isotropic wave equation

Let us recall that the 2D elastic **VTI** wave equation can be written:

$$\begin{cases} \rho(\mathbf{x})\ddot{u}_x(\mathbf{x}, t) &= \sigma_{xx,x}(\mathbf{x}, t) + \sigma_{xz,z}(\mathbf{x}, t) + f_x(\mathbf{x}, t), \\ \rho\ddot{u}_z(\mathbf{x}, t) &= \sigma_{xz,x}(\mathbf{x}, t) + \sigma_{zz,z}(\mathbf{x}, t) + f_z(\mathbf{x}, t), \\ \sigma_{xx}(\mathbf{x}, t) &= C_{11}(\mathbf{x})u_{x,x}(\mathbf{x}, t) + C_{13}(\mathbf{x})u_{z,z}(\mathbf{x}, t), \\ \sigma_{zz}(\mathbf{x}, t) &= C_{13}(\mathbf{x})u_{x,x}(\mathbf{x}, t) + C_{33}(\mathbf{x})u_{z,z}(\mathbf{x}, t), \\ \sigma_{xz}(\mathbf{x}, t) &= C_{44}(\mathbf{x})[u_{x,z}(\mathbf{x}, t) + u_{z,x}(\mathbf{x}, t)], \end{cases} \quad (\text{H.1})$$

where \cdot_j indicates a partial spatial derivative along j and $\ddot{\cdot}$ indicates a second order partial time derivative. u_i is the i^{th} component of the displacement field (unit of length), and σ_{xx} , σ_{zz} and σ_{xz} are the normal and shear components of stress respectively (force per unit of surface). ρ is the density (mass per unit of volume) and C_{ij} are the components of the 2D **VTI** elasticity tensor -or stiffness tensor- (mass per unit of length and per squared unit of time). f_i is the i^{th} component of an external body force source (force per unit of volume). t is the time, and \mathbf{x} is a position vector made up of two components. Note that there is no attenuation: the properties of the sub-surface (C_{ij} and ρ) do not depend on time. Equation [H.1](#) corresponds to the first order equivalent of equation [1.3](#), without the moment density sources.

H.1.2 Source time function inversion

An airgun (array) source in \mathbf{x}_{src} is modeled by a point moment density source with no directivity, which means in terms of force source (e.g., Arntsen et al., 2000 ([8](#))):

$$f_q(\mathbf{x}, t) = \left(-p(t)\delta(\mathbf{x} - \mathbf{x}_{\text{src}}) \right)_{,q}, \quad \text{for } q = x, z, \quad (\text{H.2})$$

where δ is the Dirac delta function, and $p(t)$ is the source time function (unit of stress, i.e., force per unit of surface). \mathbf{x}_{src} is known and $p(t)$ is the unknown.

In a fashion similar to the one used to obtain equation [2.11](#) of [Chapter 2](#), and using the same notations, the observed pressure recorded in the acoustic isotropic water layer in \mathbf{x}_{rcv} from the source of equation [H.2](#) can be written:

$$\begin{aligned} P^{\text{airgun}(\mathbf{x}_{\text{src}}, p)}(\mathbf{x}_{\text{rcv}}, t) &= p(t) *_{t} \left(-C_{11}(\mathbf{x}_{\text{rcv}}) \left(G_{xq,qx}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}_{\text{src}}, 0) + G_{zq,qz}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}_{\text{src}}, 0) \right) \right), \\ &= p(t) *_{t} \left(F(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}_{\text{src}}, 0) \right), \end{aligned} \quad (\text{H.3})$$

where $*_{t}$ indicates the time convolution, where Einstein convention of summing on repeated indices is used, and where $G_{iq}(\mathbf{x}_{\text{rcv}}, t; \mathbf{x}, 0)$ is the value in \mathbf{x}_{rcv} and at time t of the i^{th} Green's function of wave equation [H.1](#) for a unit body force source oriented along the q^{th} axis, shot at time 0 and located in \mathbf{x} . As explained in [Chapter 2](#), this Green's function can be seen as the i^{th} component of the displacement in \mathbf{x}_{rcv} and at time t due to a unit body force source oriented along the q^{th} axis, shot at time 0 and located in \mathbf{x} . The superscript

($\text{airgun}_{\mathbf{x}_{src}, p}$) specifies that the quantity is due to an airgun source located in \mathbf{x}_{src} , and whose source time function is $p(t)$. The quantity F is introduced to lighten the equations. Conversion of equation H.3 in the frequency domain gives:

$$\tilde{P}^{\text{airgun}(\mathbf{x}_{src}, p)} = \tilde{p}\tilde{F}, \quad (\text{H.4})$$

where the \sim denotes the Fourier transform.

Assuming \tilde{F} is non zero and known, one can then obtain \tilde{p} from the frequency-domain observed pressure data $\tilde{P}^{\text{airgun}(\mathbf{x}_{src}, p)}$ with:

$$\tilde{p} = \frac{\tilde{P}^{\text{airgun}(\mathbf{x}_{src}, p)}}{\tilde{F}}. \quad (\text{H.5})$$

The wanted source time function $p(t)$ is then obtained by applying an inverse Fourier transform on \tilde{p} . In practice, one cannot be sure that \tilde{F} is non zero. Consequently, the following approximation is used instead:

$$\tilde{p} \approx \frac{\tilde{P}^{\text{airgun}(\mathbf{x}_{src}, p)}\tilde{F}^*}{\tilde{F}\tilde{F}^* + \epsilon_1}, \quad (\text{H.6})$$

where ϵ_1 is a small real value and where $*$ refers to the complex conjugate.

The computation of \tilde{F} is performed by assuming that the properties of the sub-surface are known. The pressure $P^0(\mathbf{x}_{rcv}, t)$ coming from the H.2 source but with an arbitrary source time function $p^0(t)$ (called the “initial” source time function) is modeled. Using equation H.5 and the same stabilization technique as in equation H.6, one can obtain the missing \tilde{F} with:

$$\tilde{F} = \frac{\tilde{P}^0\tilde{p}^0^*}{\tilde{p}^0\tilde{p}^0^* + \epsilon_2}, \quad (\text{H.7})$$

where ϵ_2 is a small real value. The source time function $p(t)$ of the airgun (array) source of equation H.2 can then be estimated from a pressure recording, using equation H.6 and assuming a perfect knowledge of the sub-surface for \tilde{F} computation.

H.2 Synthetic applications on a modified Marmousi2 model

Above theory explains how to invert the source time function from a single pressure recording (i.e., one source-receiver pair). I present here synthetic applications of such theory. To reduce inversion errors, several recordings are used. Each source-receiver pair is used to invert an individual source time function. The final inverted source time function is the mean of all individual source time functions. From above theory, one can understand that the initial time function has to contain the frequencies of the true source time function, and can even contain additional frequencies.

H.2.1 An ideal application

An “observed” receiver pressure gather was obtained from a modified Marmousi2 model (the true model). The true P-wave vertical velocity model V_{pv} was obtained from V_p Marmousi2 model (Martin et al., 2006 (115)) whose horizontal layer just below the seabed

was removed and whose water layer was reduced to a 126-meter-thick isotropic homogeneous water layer. V_{pv} model is shown on figure H.1a. The density model ρ and a S-wave vertical velocity model V_{sv} were generated by applying experimental relationships on V_{pv} model. The missing SV-wave vertical velocity model V_{sv} is computed from the initial P-wave vertical velocity model V_{pv} , with (Castagna et al., 1985 (39); Flidner and White, 2001 (60); Shipp and Singh, 2002 (172) and Sears et al., 2010 (165)):

$$\begin{cases} V_{sv} &= 0 \text{ m/s, in water,} \\ V_{sv} &= \frac{V_{pv}-1360}{1.16} \text{ m/s, if } V_{pv} < 3500 \text{ m/s,} \\ V_{sv} &= 0.53V_{pv} \text{ m/s, if } V_{pv} \geq 3500 \text{ m/s,} \end{cases} \quad (\text{H.8})$$

where V_{pv} is in m/s. The missing density model ρ is computed from the initial P-wave vertical velocity model V_{pv} , with the empirical relationship 2.22. This equation is recalled hereafter for the chapter's completeness:

$$\begin{cases} \rho &= 1030 \text{ kg/m}^3 \text{ in water,} \\ \rho &= 2351 - 7497 * (V_{pv} * 10^{-3})^{-4.656} \text{ kg/m}^3 \text{ if } V_{pv} < 2000 \text{ m/s,} \\ \rho &= x_a + \frac{V_{pv}-2000}{150} * (x_c - x_a) \text{ kg/m}^3 \text{ if } V_{pv} \geq 2000 \text{ m/s and } V_{pv} \leq 2150 \text{ m/s,} \\ &\text{with } x_a = 2351 - 7497 * (2000 * 10^{-3})^{-4.656} \text{ and } x_c = 1741 * (2150 * 10^{-3})^{0.25}, \\ \rho &= 1741 * (V_{pv} * 10^{-3})^{0.25} \text{ kg/m}^3 \text{ if } V_{pv} > 2150 \text{ m/s,} \end{cases} \quad (\text{H.9})$$

where V_{pv} is in m/s. The empirical expression of Hamilton (1978) (75) is used for the shallow seabed poorly consolidated sediments ($V_{pv} < 2000$ m/s), and the empirical expression of Gardner et al. (1974) (65) is used for the deeper consolidated sediments ($V_{pv} > 2150$ m/s). A linear transition between the two empirical relations is applied. Note that these two empirical relationships were derived assuming isotropic media and are here applied in anisotropic media by replacing the P-wave velocity of the initial values by the P-wave vertical velocities. Anisotropy is added by setting $\sqrt{1+2\delta}$ and $\sqrt{1+2\eta}$ anisotropic parameters to 1.1 in the sediments (see equation F.7 of Chapter 3 for more details about these parameters). The receiver is located in the water layer, on the first gridpoint above the seabed, at a horizontal location of 180 meters. 219 airgun sources are used, located 12 meters below the free-surface and with a spacing of 35 meters. The left most source is also at a horizontal location of 180 meters. The airgun sources are modeled with equation H.2 and they all share the same source time function $p(t)$: a 7.5 Hz central frequency Ricker wavelet. This acquisition mimics Alwyn real data airgun-ocean bottom node acquisition. The modeling is performed using reciprocity equations of Chapter 2. The obtained ‘‘observed’’ gather is shown in figure H.2a.

The source time function $p(t)$ is first inverted using the whole ‘‘observed’’ and modeled data gather, and the true model. The initial source time function p^0 is shown in figures H.3a and H.3c. The inverted source function is shown in the same figure, alongside with the true source time function. One can conclude that the inversion performed well. The inverted source function contains unreal high frequencies that can be handled by high-cut filtering the inverted source time function, as shown in figures H.3a and H.3c. The low frequencies are slightly less well retrieved than the higher frequencies. These low frequencies and the removed unreal high frequencies are linked to the choice of ϵ_1 and ϵ_2 .

H.2.2 A more realistic application

The previous application is not realistic. Indeed in real applications, the parameters of the sub-surface are unknown. This application is performed using a smoothed version of the true model instead for the estimation. The smoothed V_{pv} model is shown on figure H.1b. The modeled data are shown in figure H.2b. Figure H.2c shows the modeled data obtained with the true model for comparison. The poor inversion results are showed in figures H.3b and H.3d. To overcome this problem, one should use only the trustable part of the modeled data (i.e., here, the direct arrivals extractable at short offsets as shown in figure H.2d). The corresponding inverted source time function is shown in figures H.3b and H.3d. This last application shows that this estimation technique can lead to acceptable results even if one does not have a good model of the sub-surface, as long as at least one part of the data can be modeled correctly and separated from the rest of the data. More realistic synthetic applications could be made by adding noise for instance.

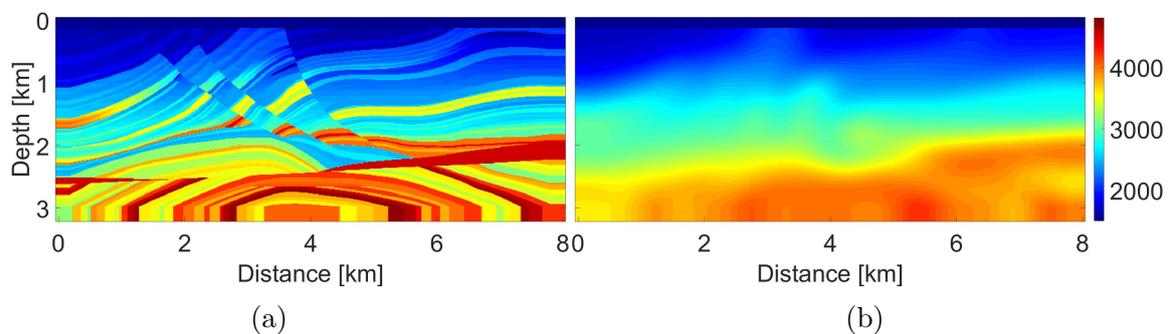


Figure H.1: P-wave vertical velocity models used for the synthetic source inversion tests. (a) True model; (b) Smooth model. Velocities in m/s.

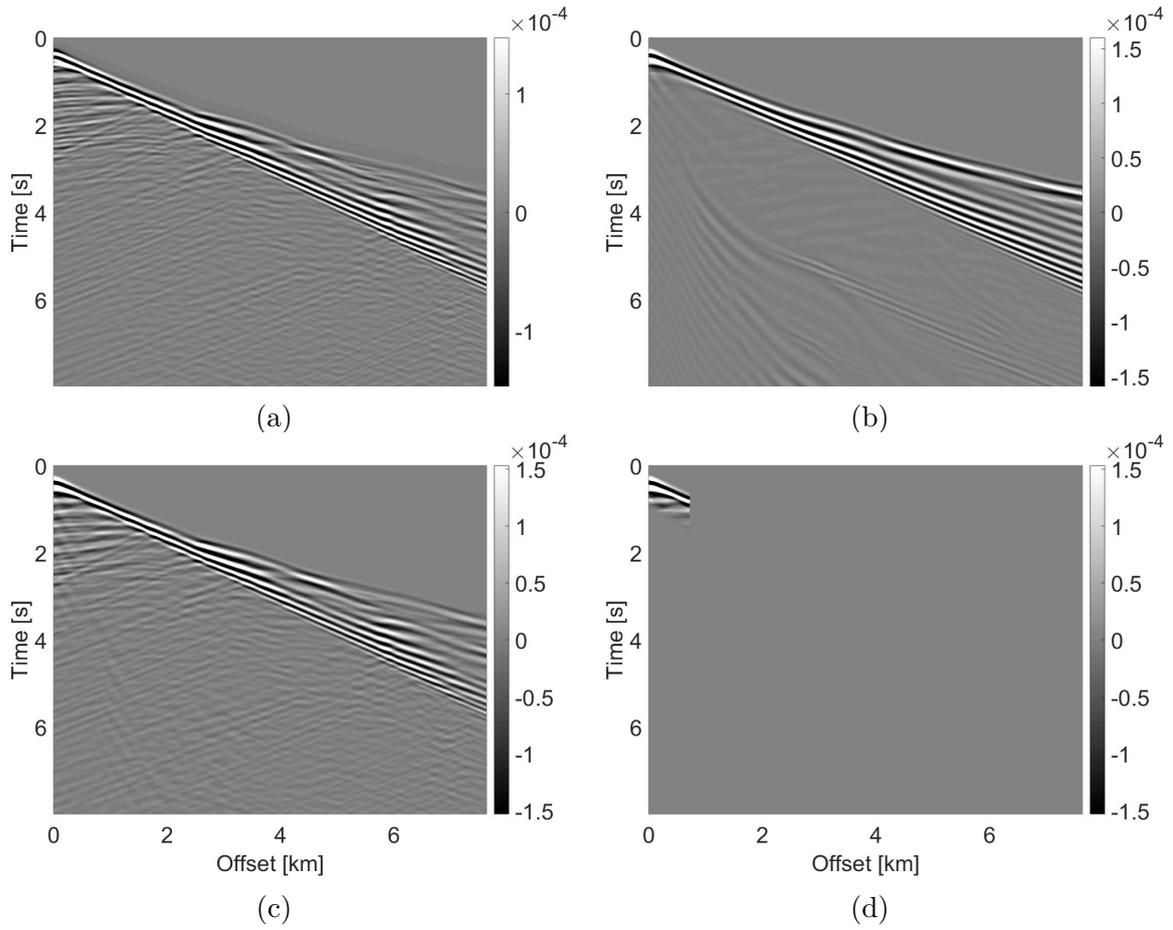


Figure H.2: Data used for the synthetic source inversion tests. Pressure receiver gather. (a) Observed data; (b) Modeled data with smooth model; (c) Modeled data with true model; (d) The used subset of modeled data.

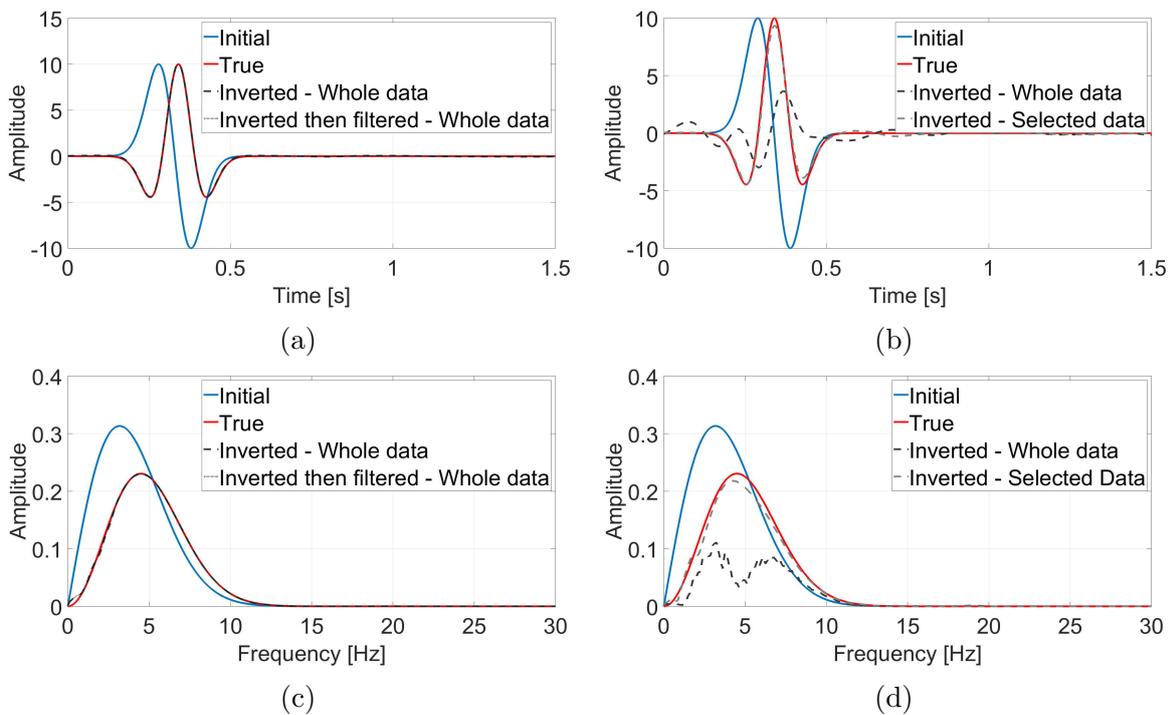


Figure H.3: Source time function inversion results. (a) and (c) Inversion performed with the true model only; (b) and (d) More realistic inversion results with the smooth model.

Appendix I

Alwyn real data: variant soft seabed strategy additional displays

Contents

I.1	Variant strategy	178
I.2	Stage 1a	179
I.3	Stage 1b	184
I.4	Stage 3'	189

Additional well comparisons with the initial and final vertical velocity models of the variant of the soft seabed strategy are available here. Stage by stage results are also provided: the non-normalized cost functions, the well comparisons, the models inversion results with comparison with the vintage migrated image, as well as observed and initial/final modeled data comparisons with trace by trace normalized displays, together with the corresponding residual displays (i.e., observed data - modeled data).

I.1 Variant strategy

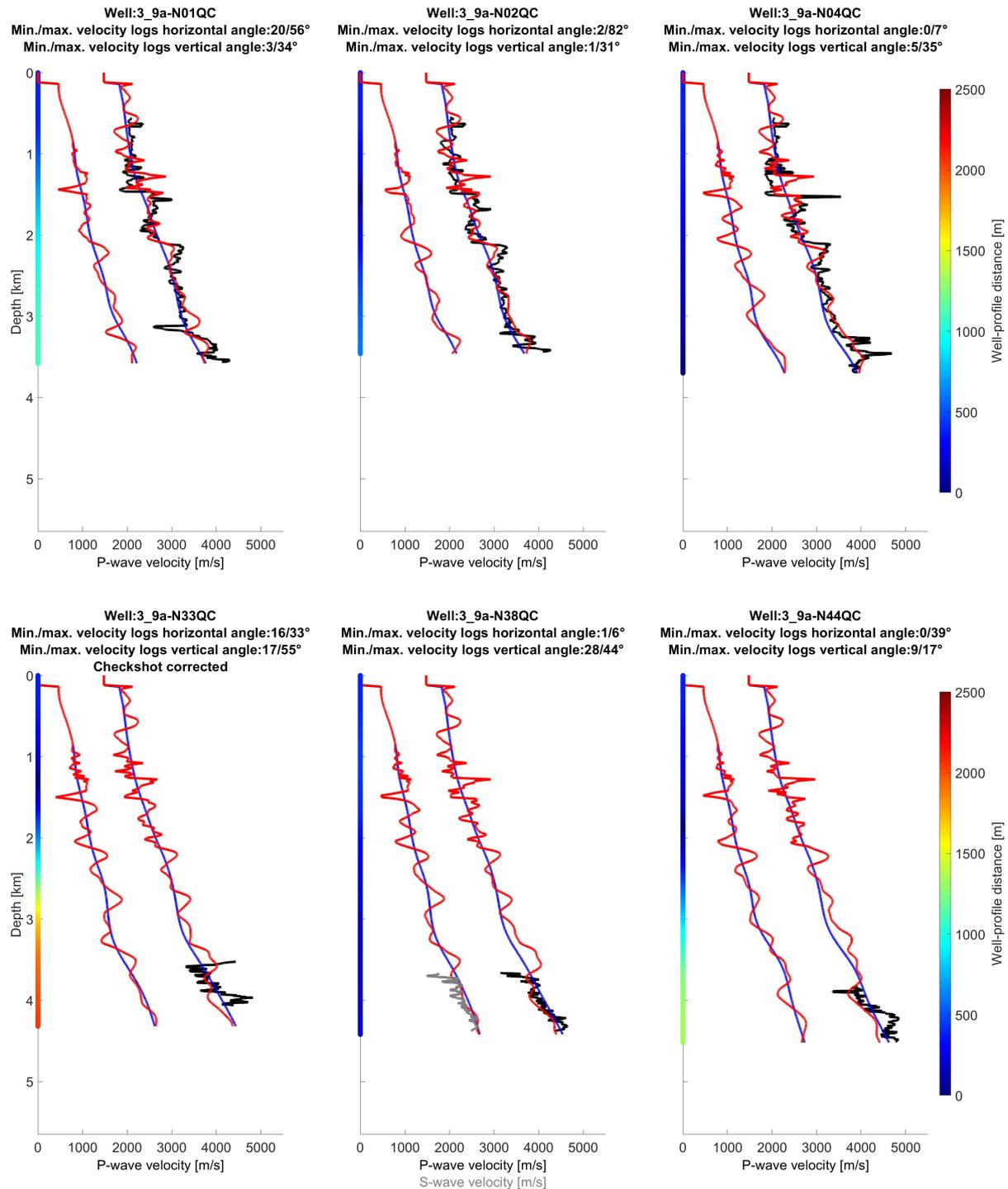


Figure I.1: Variant strategy. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

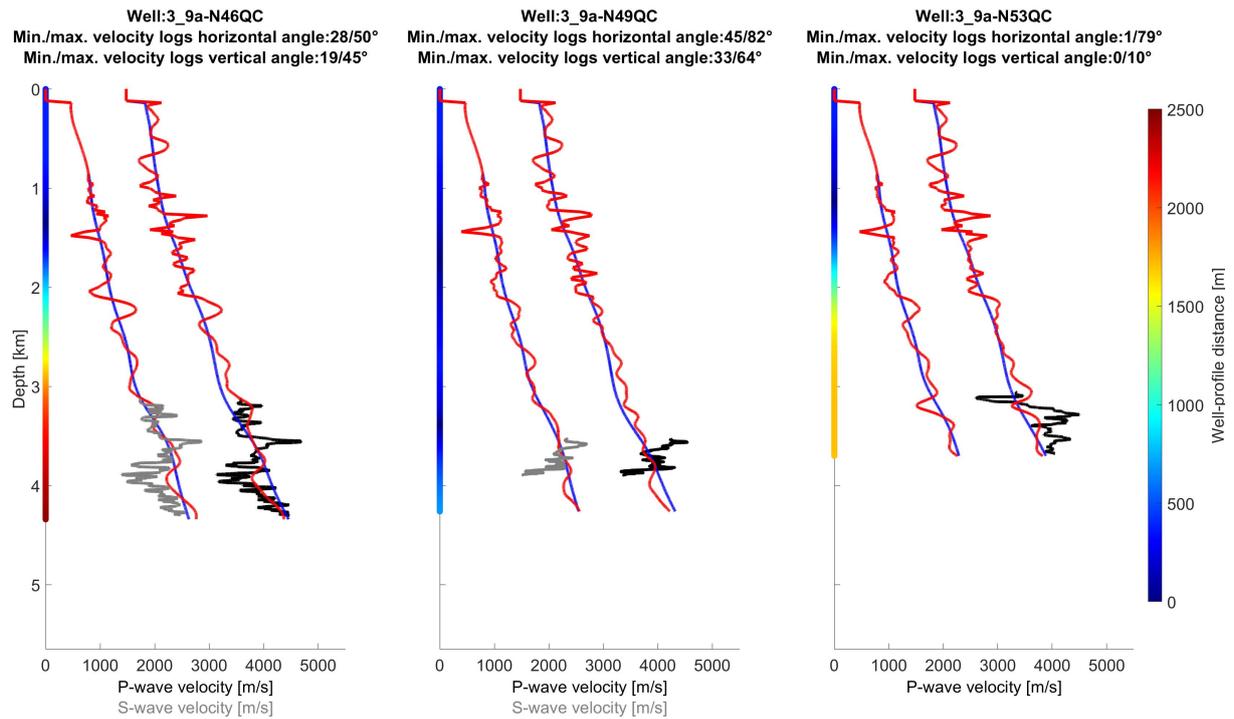


Figure I.2: Variant strategy. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

I.2 Stage 1a

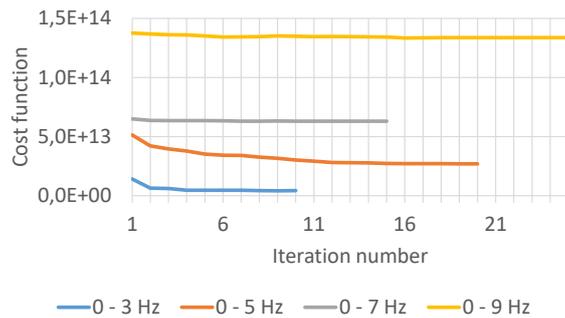


Figure I.3: Stage 1a. Non-normalized cost function for each frequency range.

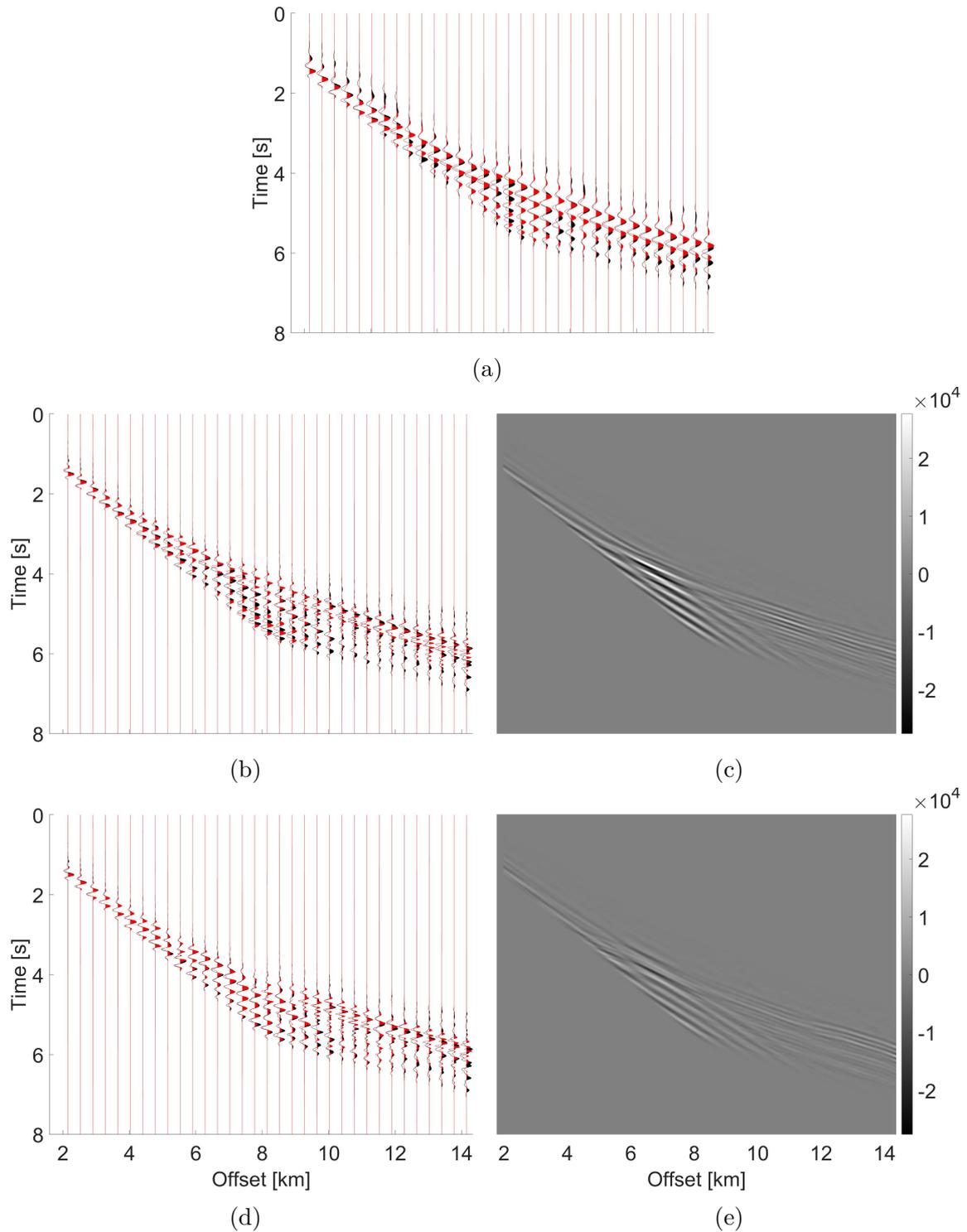


Figure I.4: Stage 1a. Data inversion results. (a) Initial data fit restricted to the first frequency range [0-3 Hz]; (b) Initial data fit; (c) Initial residual; (d) Final data fit; (e) Final residual. For a, b and d, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.

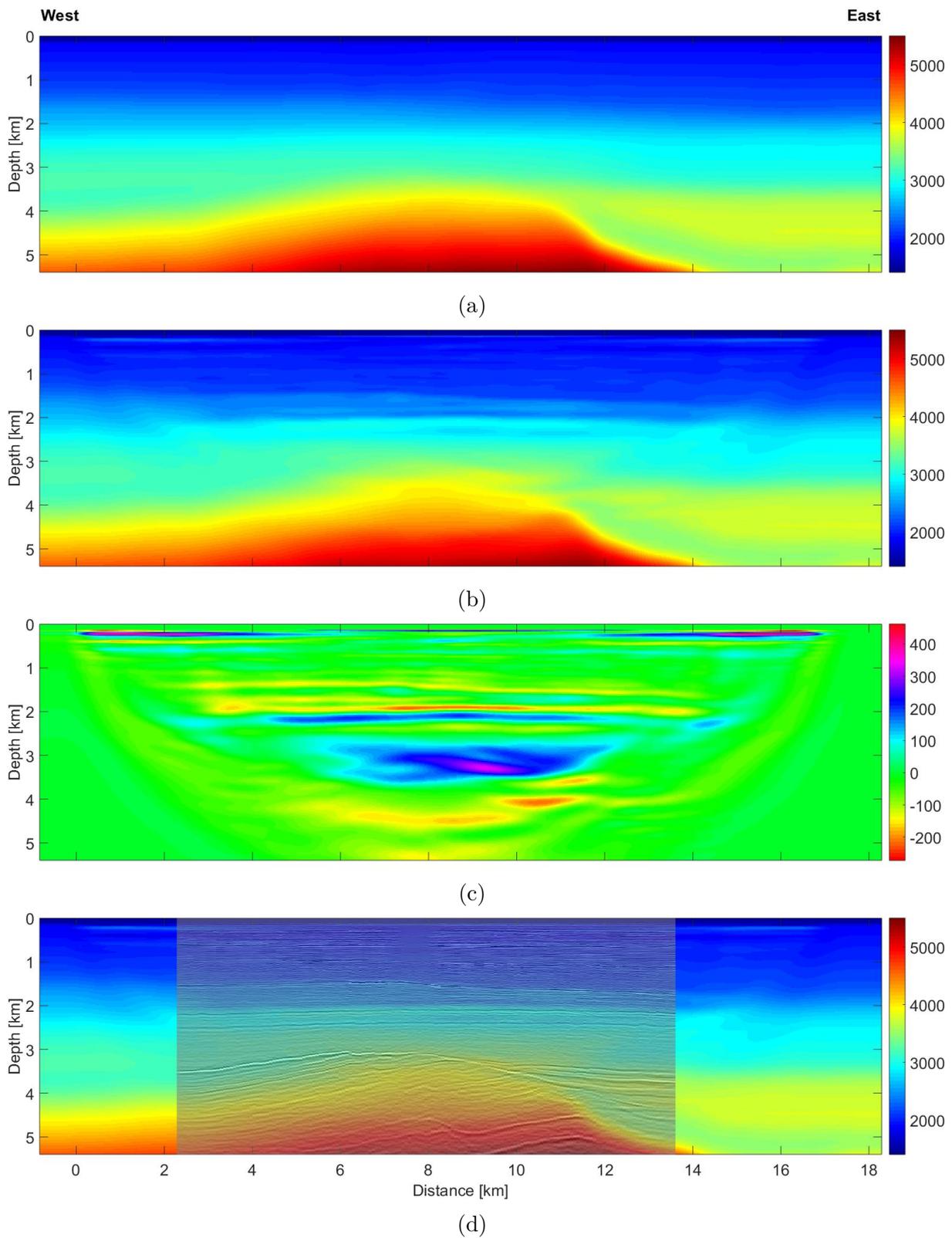


Figure I.5: Stage 1a. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

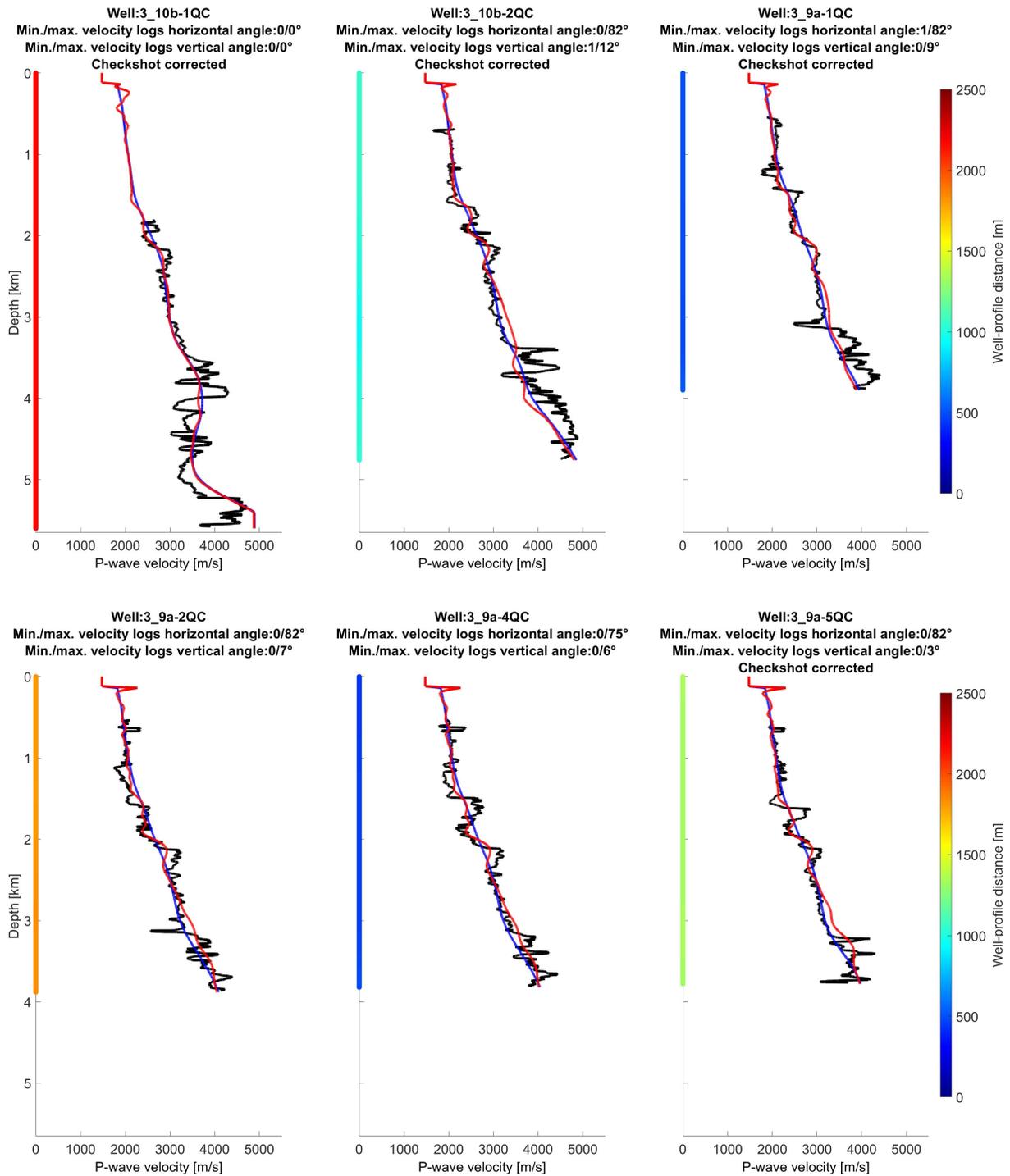


Figure I.6: Stage 1a. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

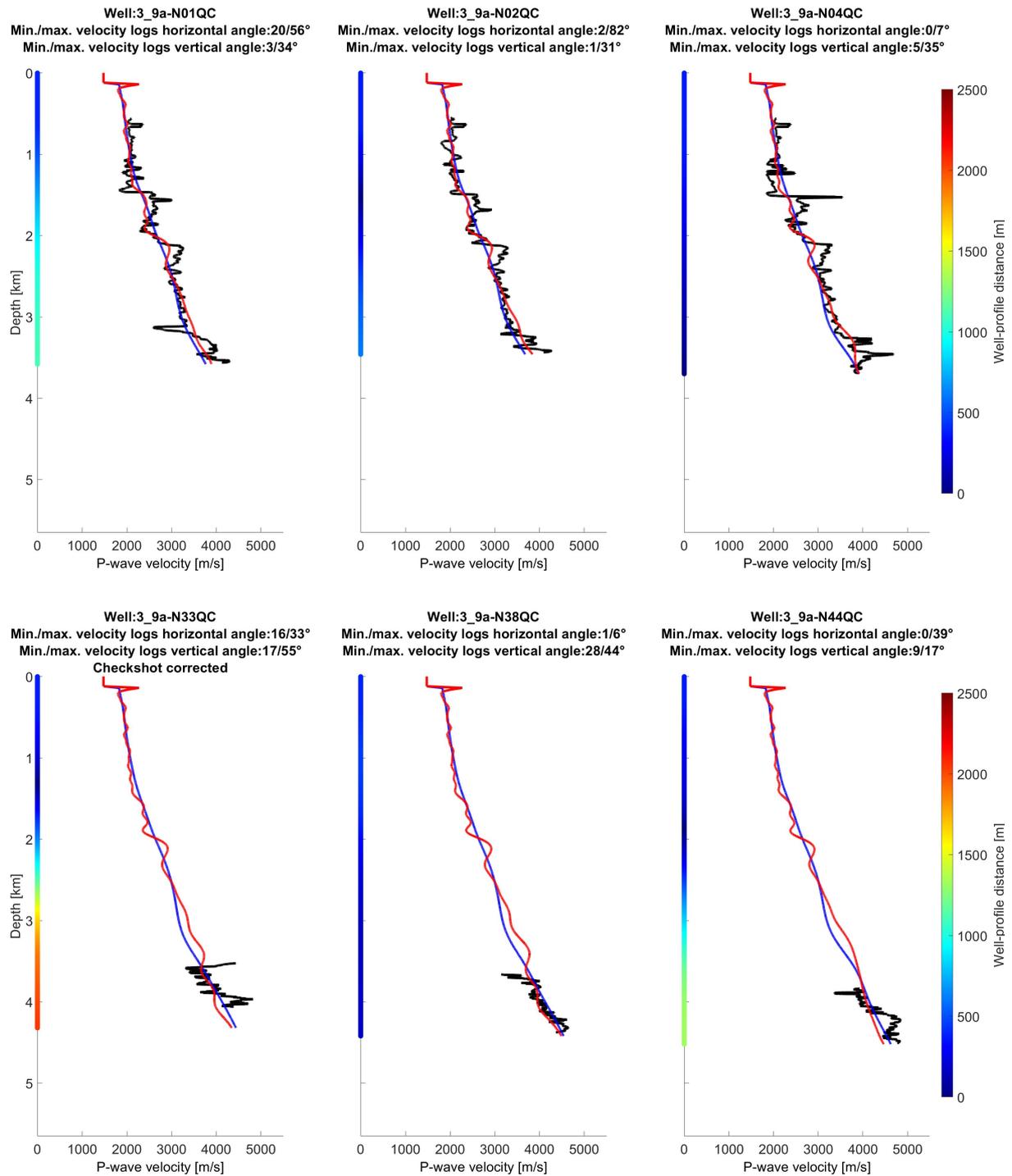


Figure I.7: Stage 1a. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

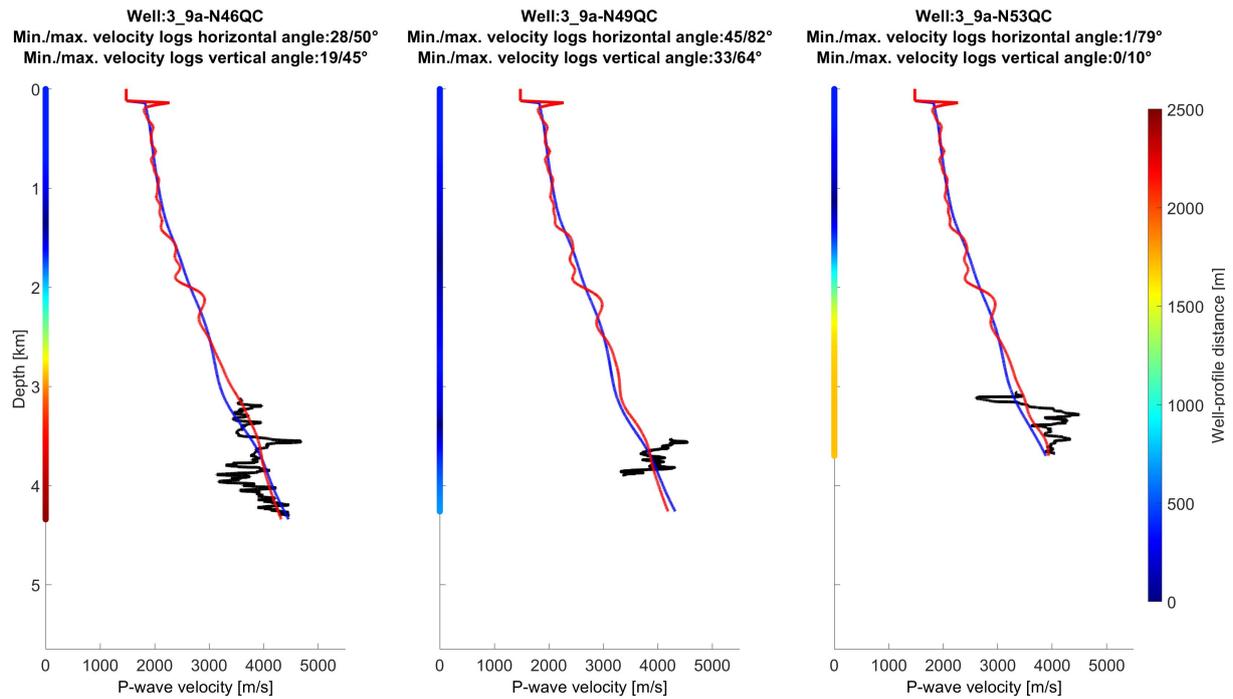


Figure I.8: Stage 1a. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

I.3 Stage 1b

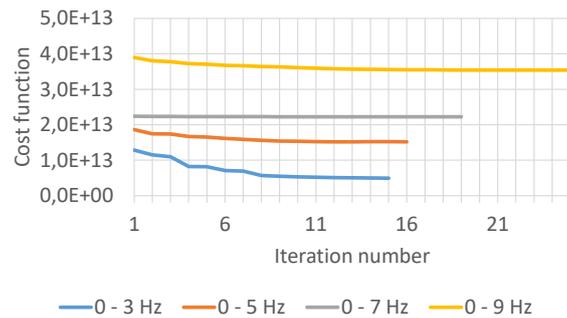


Figure I.9: Stage 1b: non-normalized cost function for each frequency range.

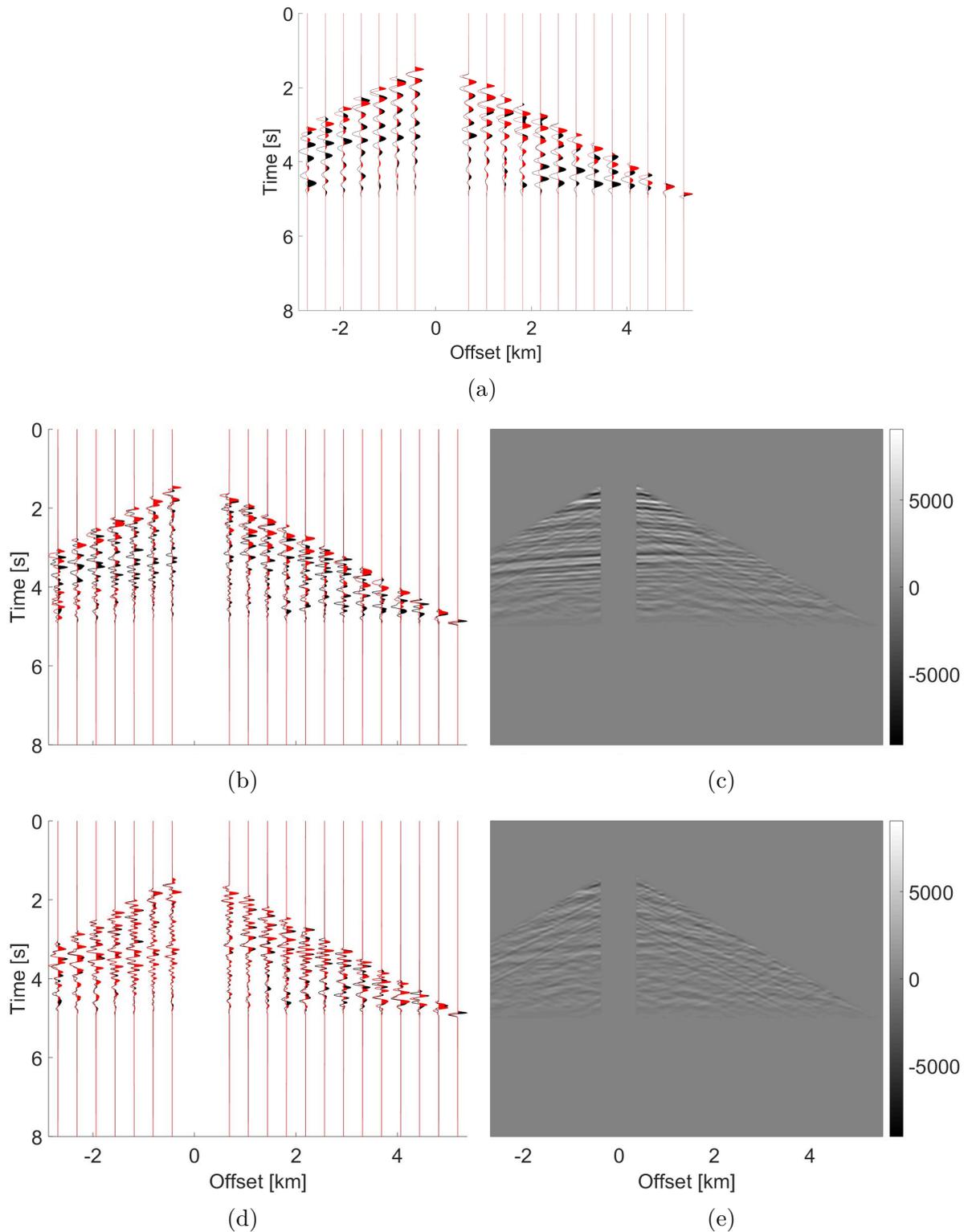


Figure I.10: Stage 1b. Data inversion results. (a) Initial data fit restricted to [0-3 Hz]; (b) Observed data; (c) Initial data fit; (d) Initial residual; (e) Final data fit; (f) Final residual. For a, c and e, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.

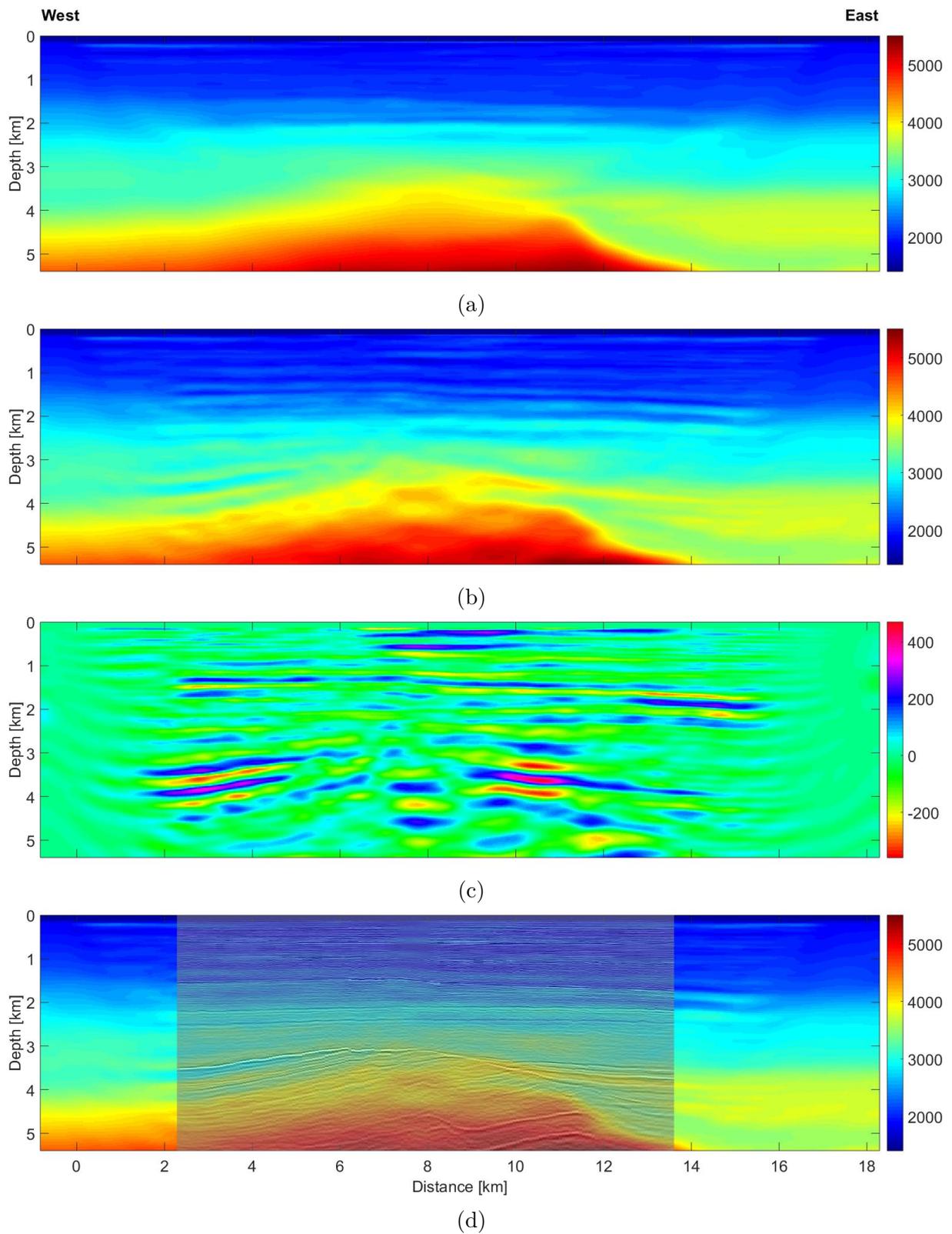


Figure I.11: stage 1b. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

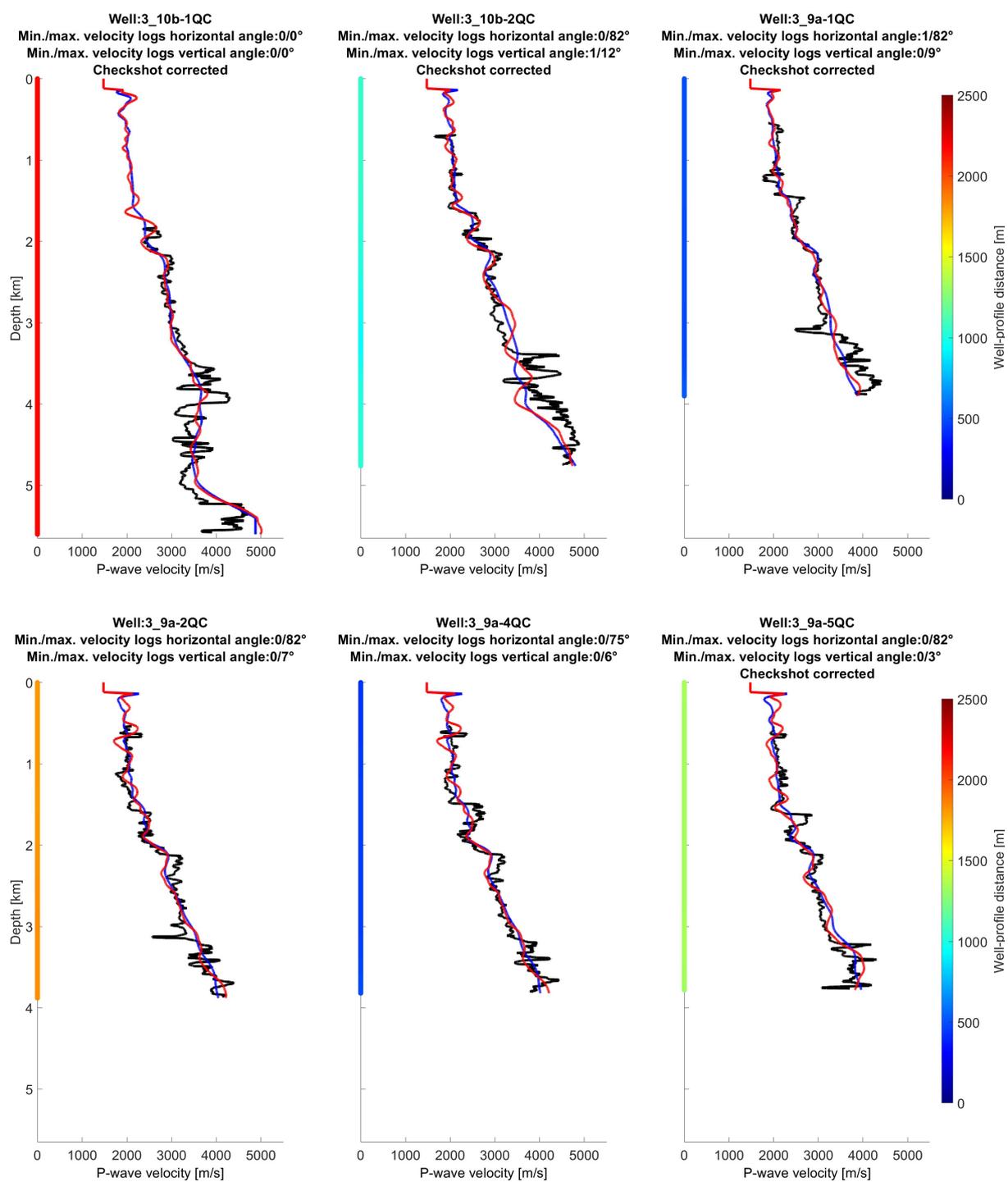


Figure I.12: Stage 1b. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

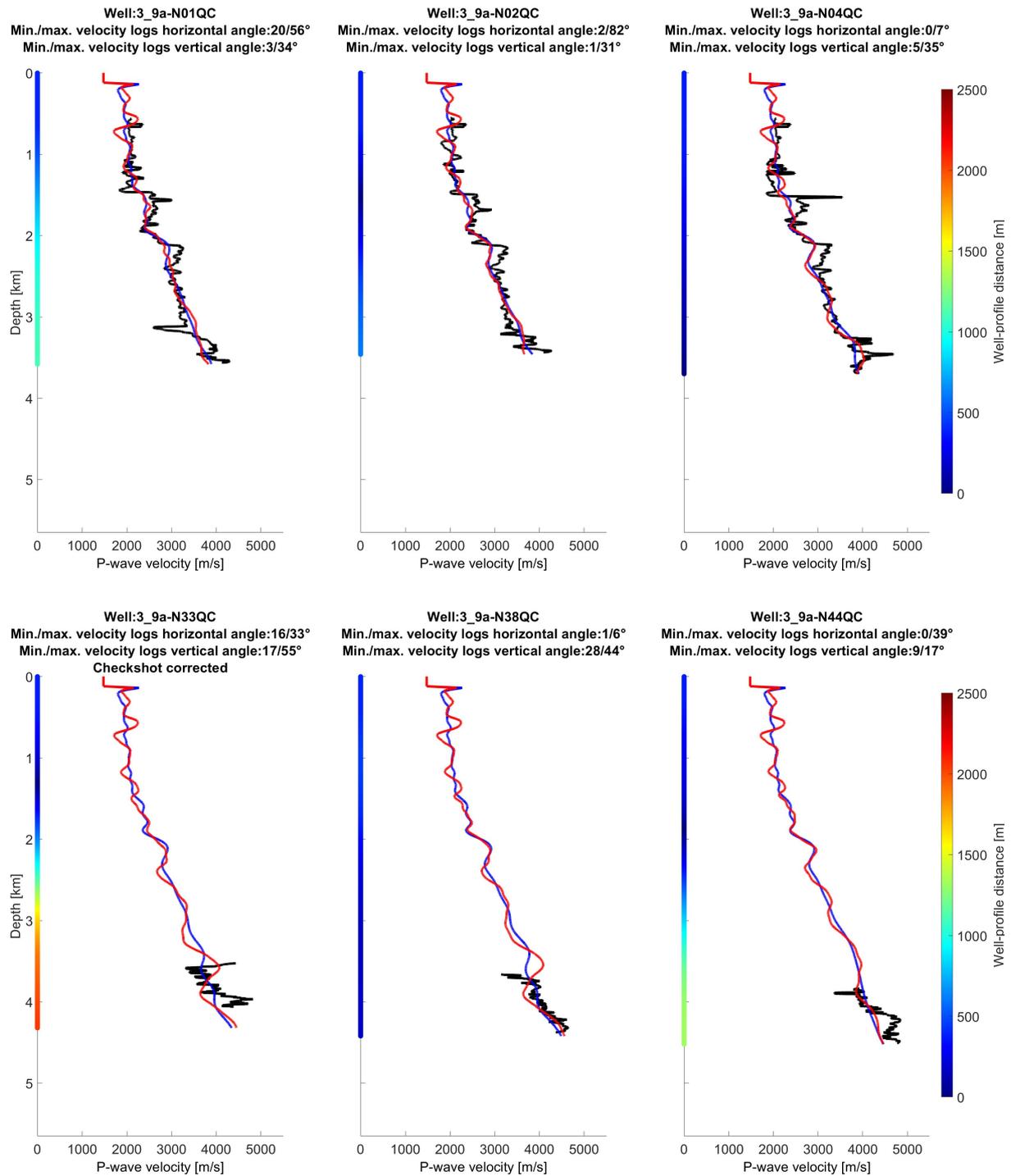


Figure I.13: Stage 1b. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

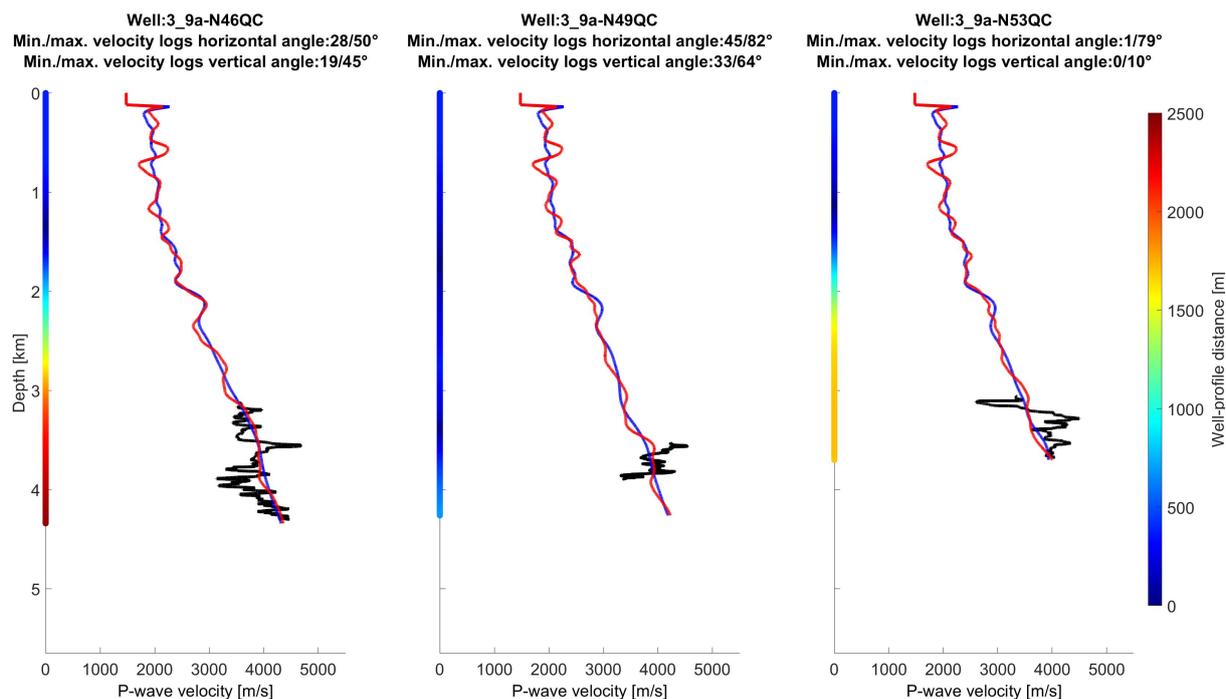


Figure I.14: Stage 1b. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed P-wave velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

I.4 Stage 3'

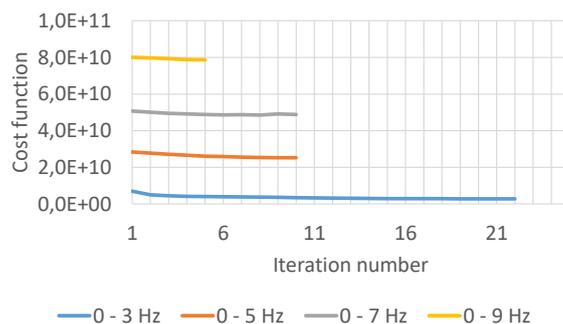


Figure I.15: Stage 3'. Non-normalized cost function for each frequency range.

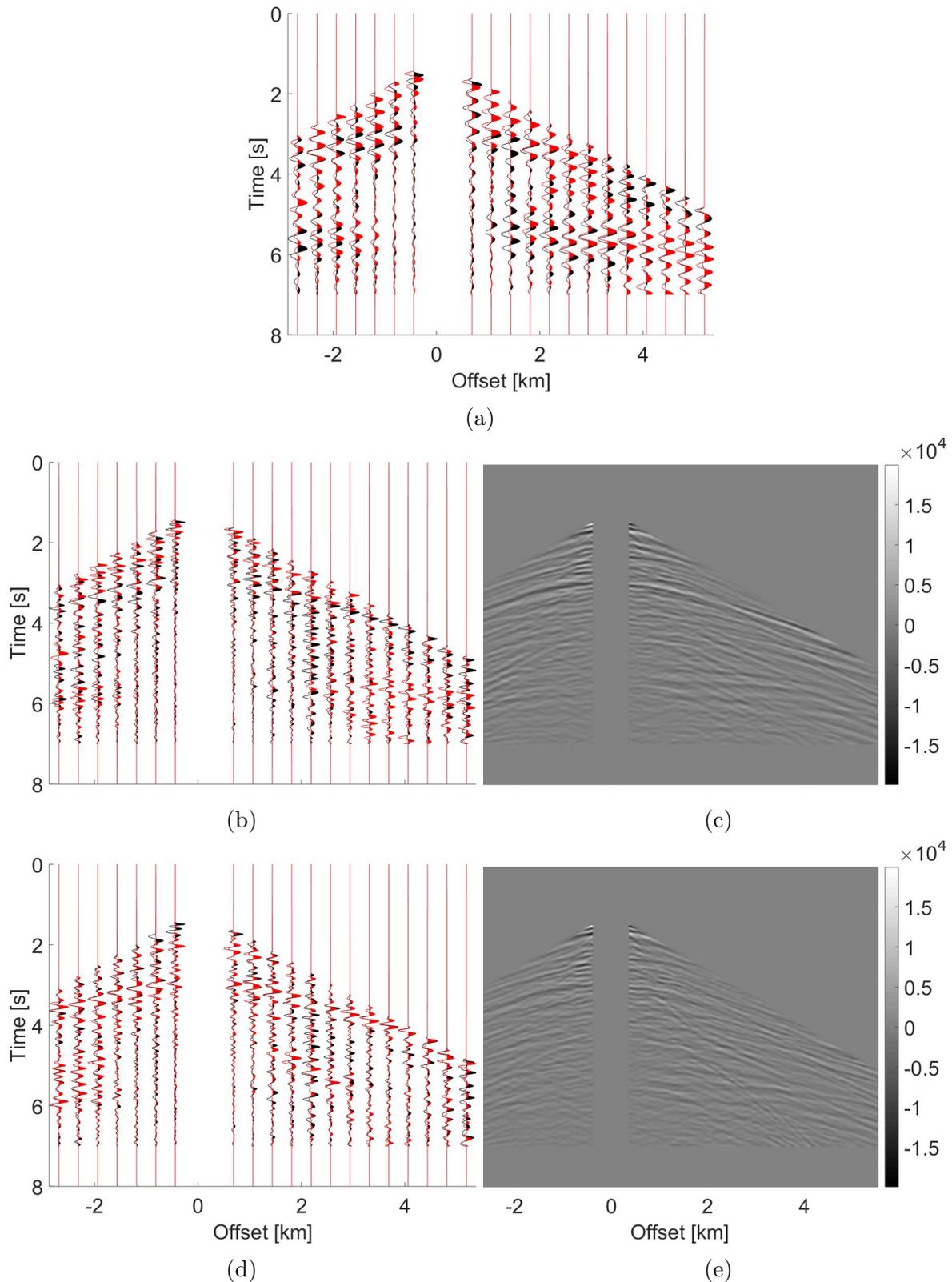


Figure I.16: Stage 3'. Data inversion results. (a) Initial data fit restricted to the first frequency range [0-3 Hz]; (b) Initial data fit; (c) Initial residual; (d) Final data fit; (e) Final residual. For a, b and d, the traces are normalized and the observed data are in black while the modeled data are in red. Western most horizontal velocity component receiver gather.

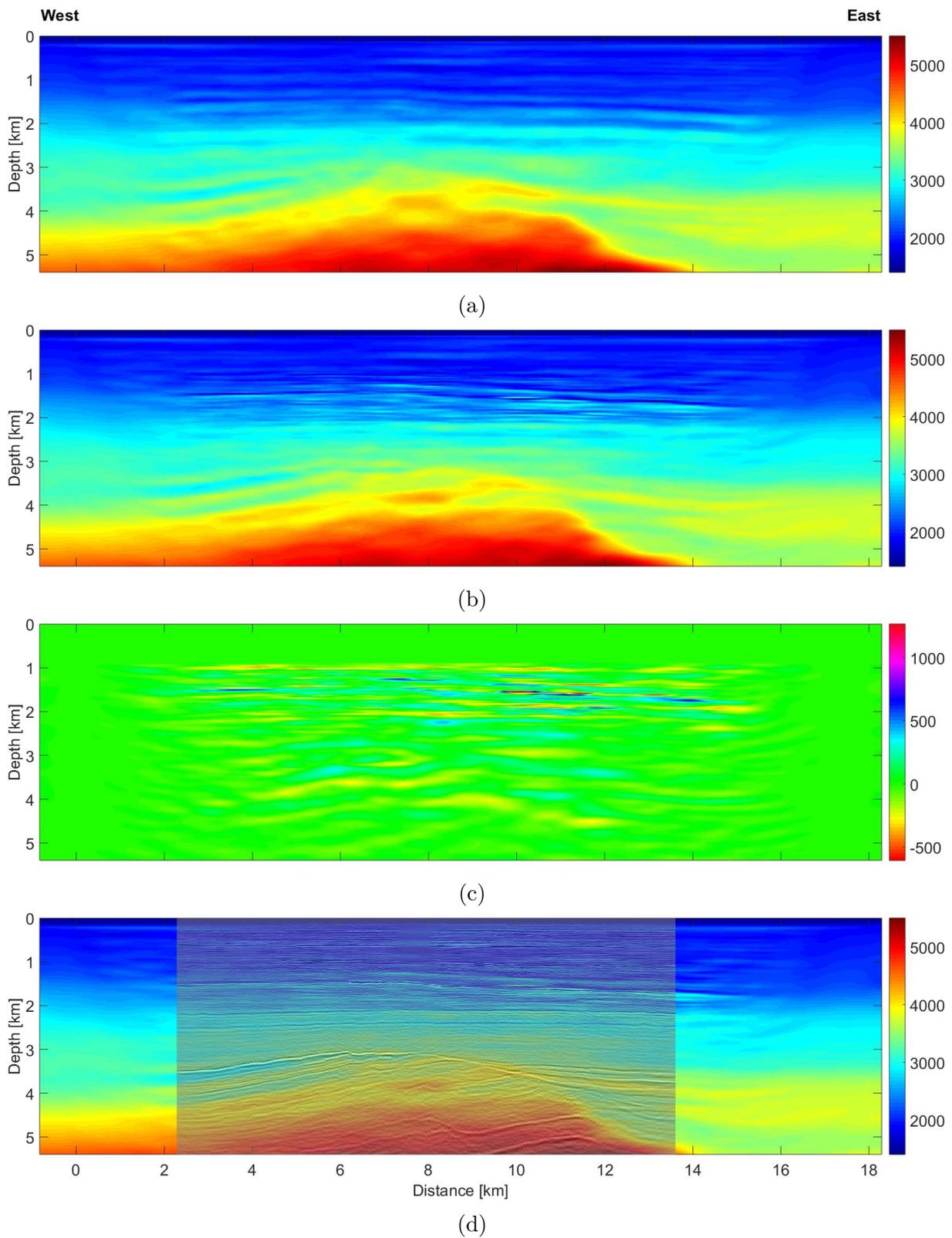


Figure I.17: Stage 3'. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

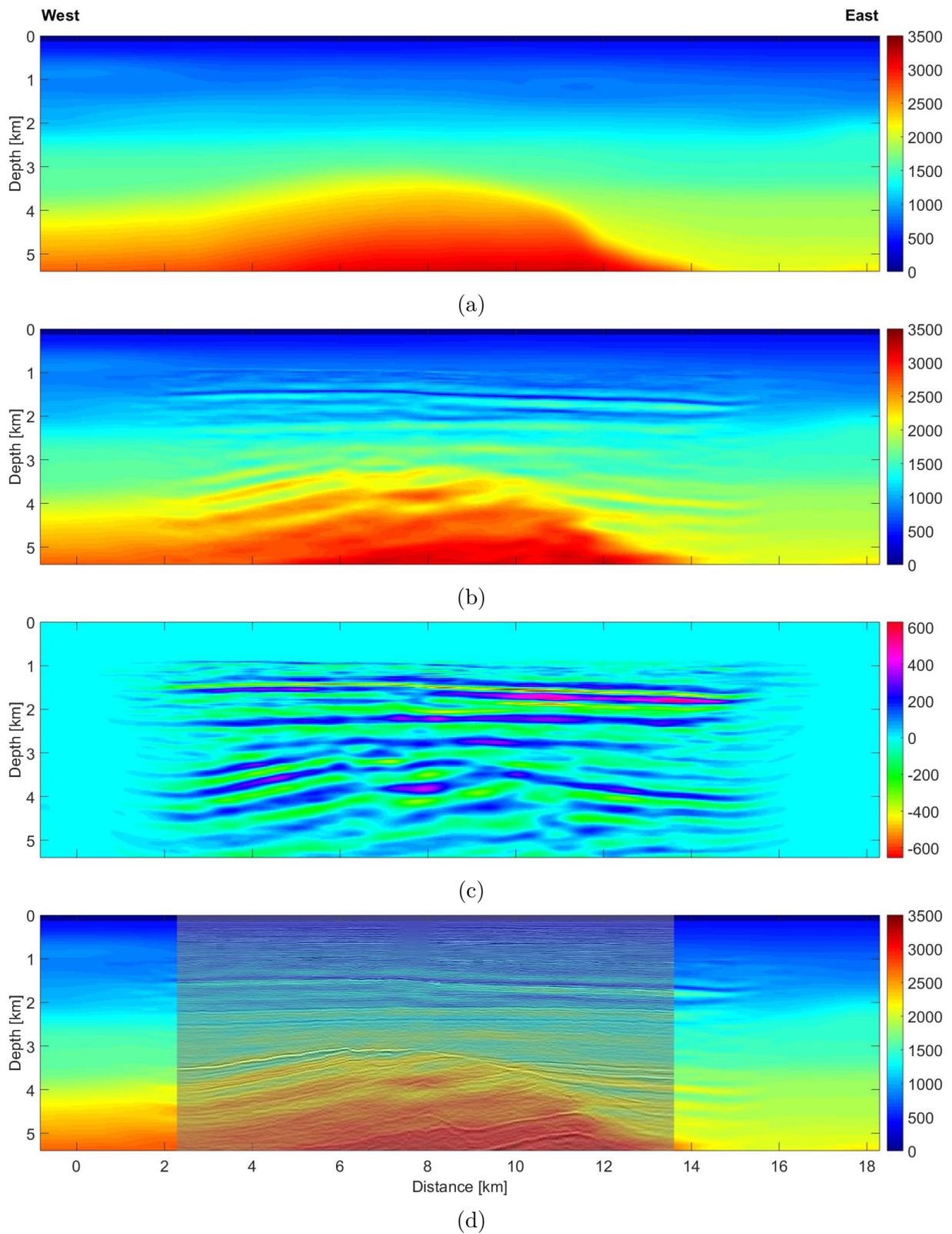


Figure I.18: Stage 3'. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

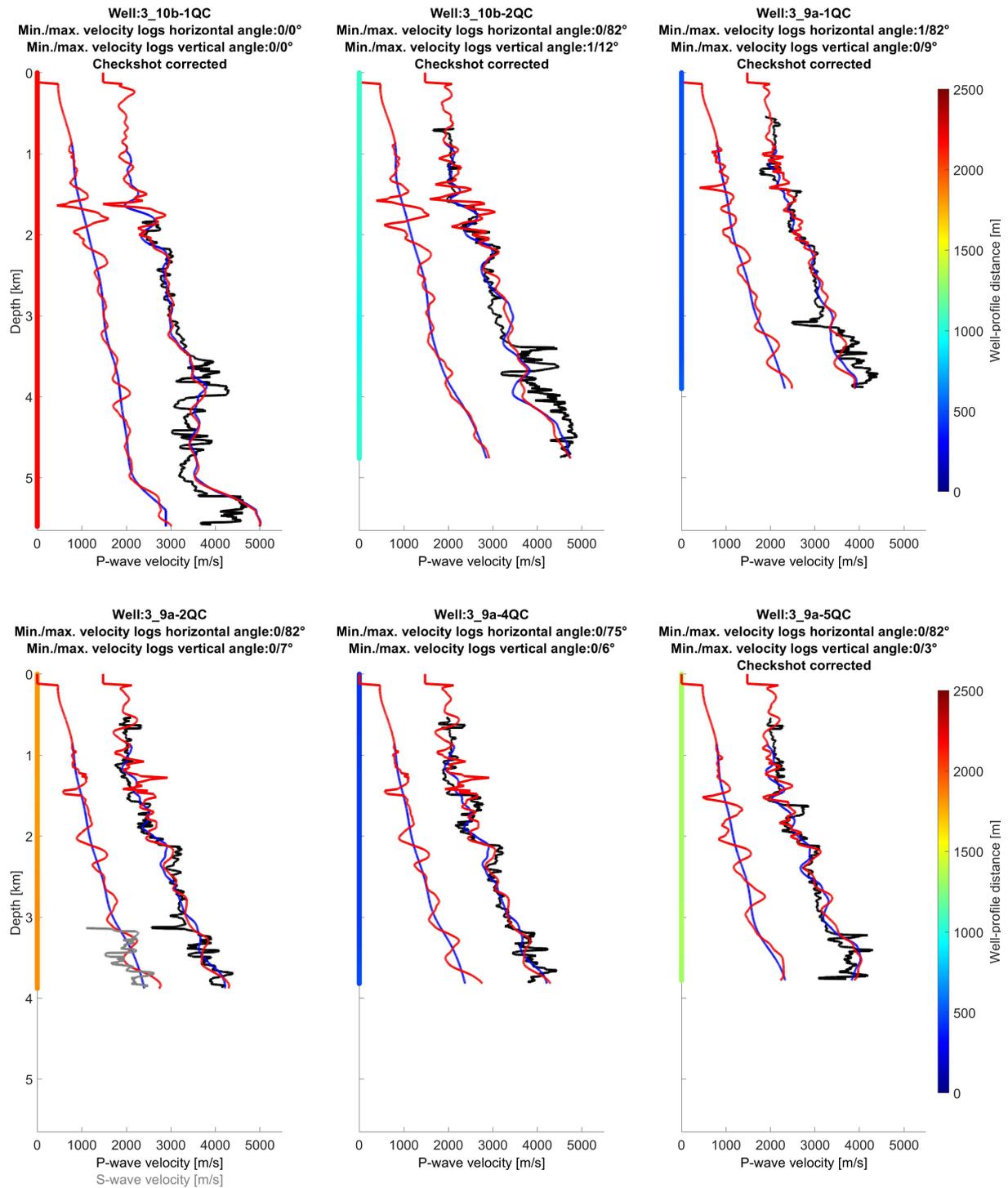


Figure I.19: Stage 3'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

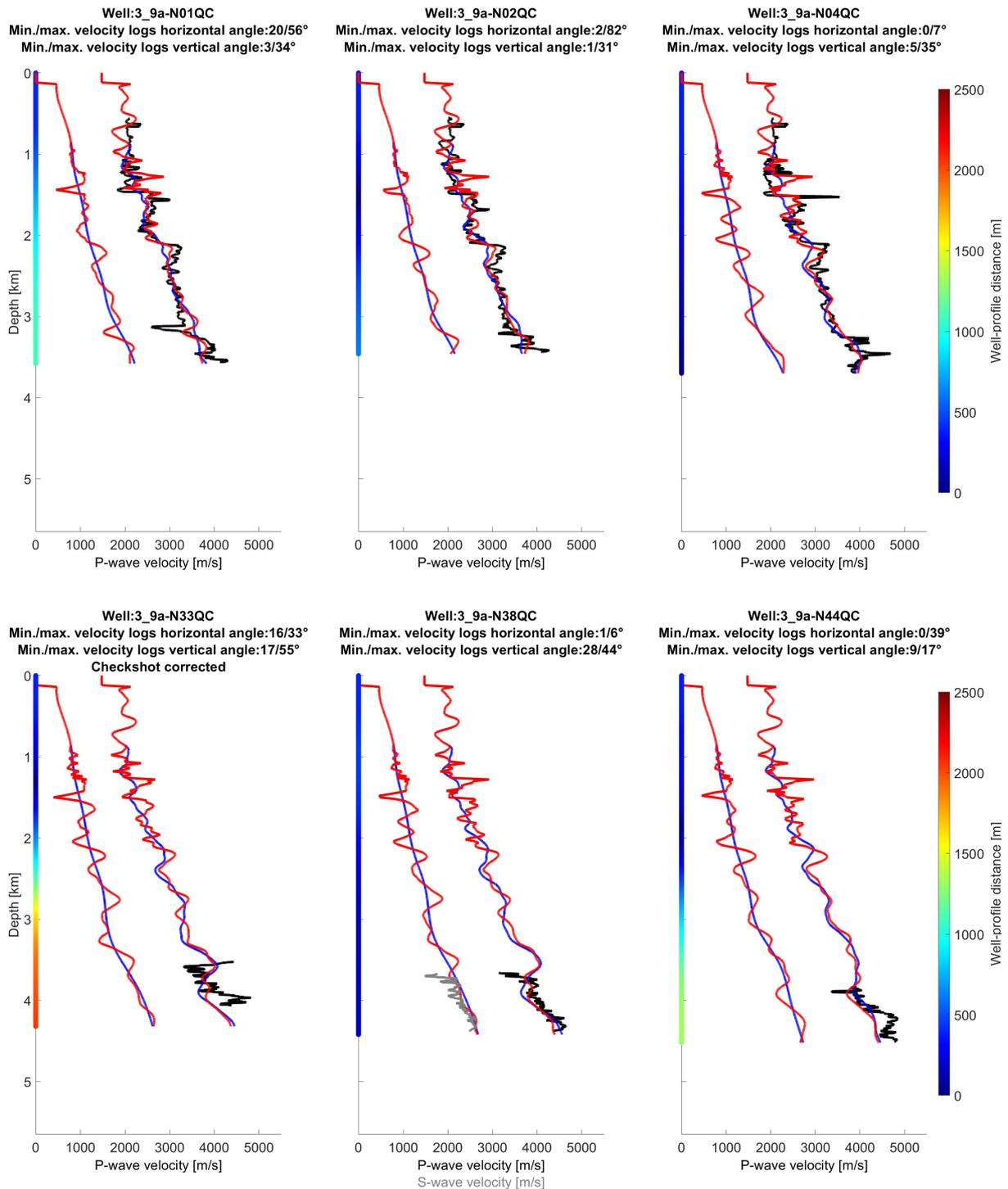


Figure I.20: Stage 3'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

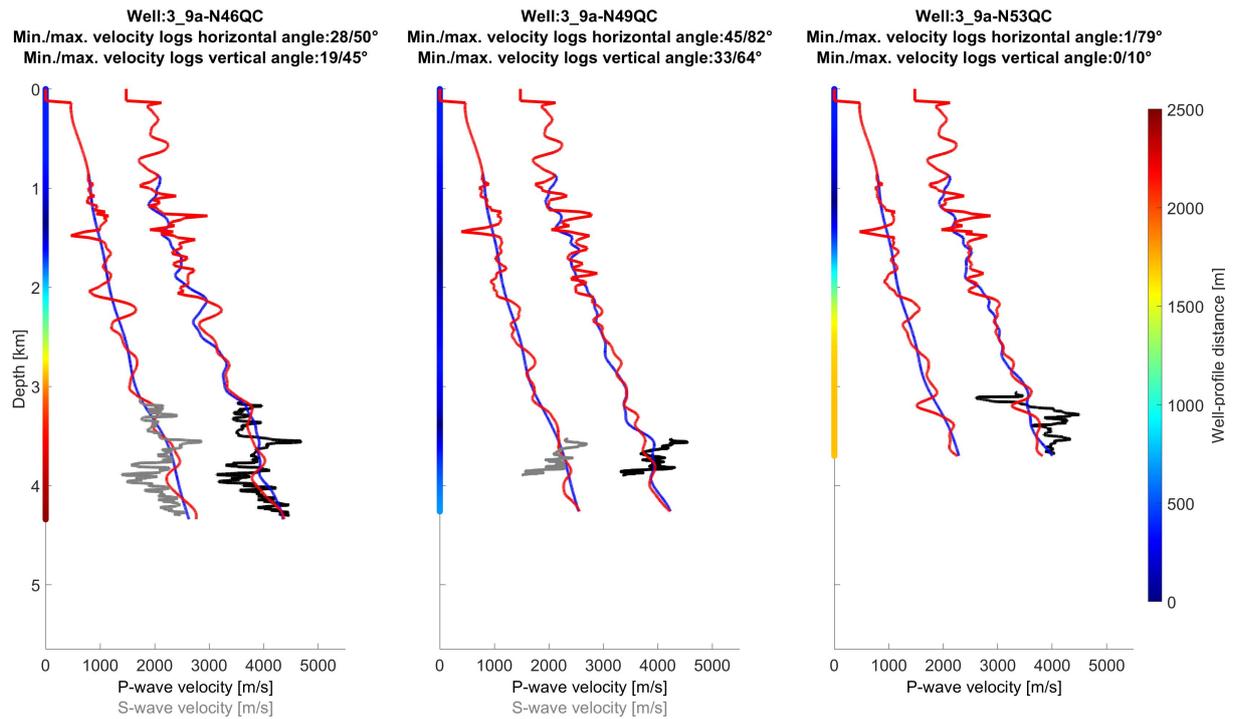


Figure I.21: Stage 3'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

Appendix J

Alwyn real data: original strategy additional displays

Contents

J.1 Stage 2	197
J.2 Stage 3 - first frequency range	203

Stage by stage results of the original strategy are provided here: the non-normalized cost functions, the well comparisons, the models inversion results with comparison with the vintage migrated image, as well as observed and initial/final modeled data comparisons with trace by trace normalized displays, together with the corresponding residual displays (i.e., observed data - modeled data).

J.1 Stage 2

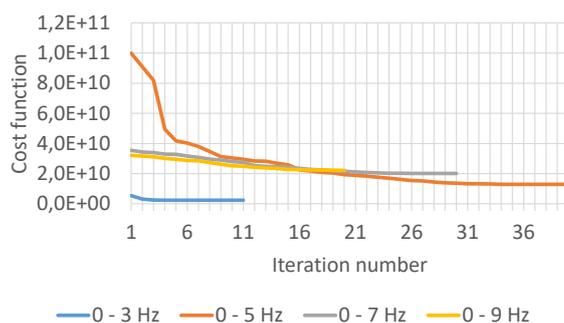


Figure J.1: Stage 2. Non-normalized cost function for each frequency range.

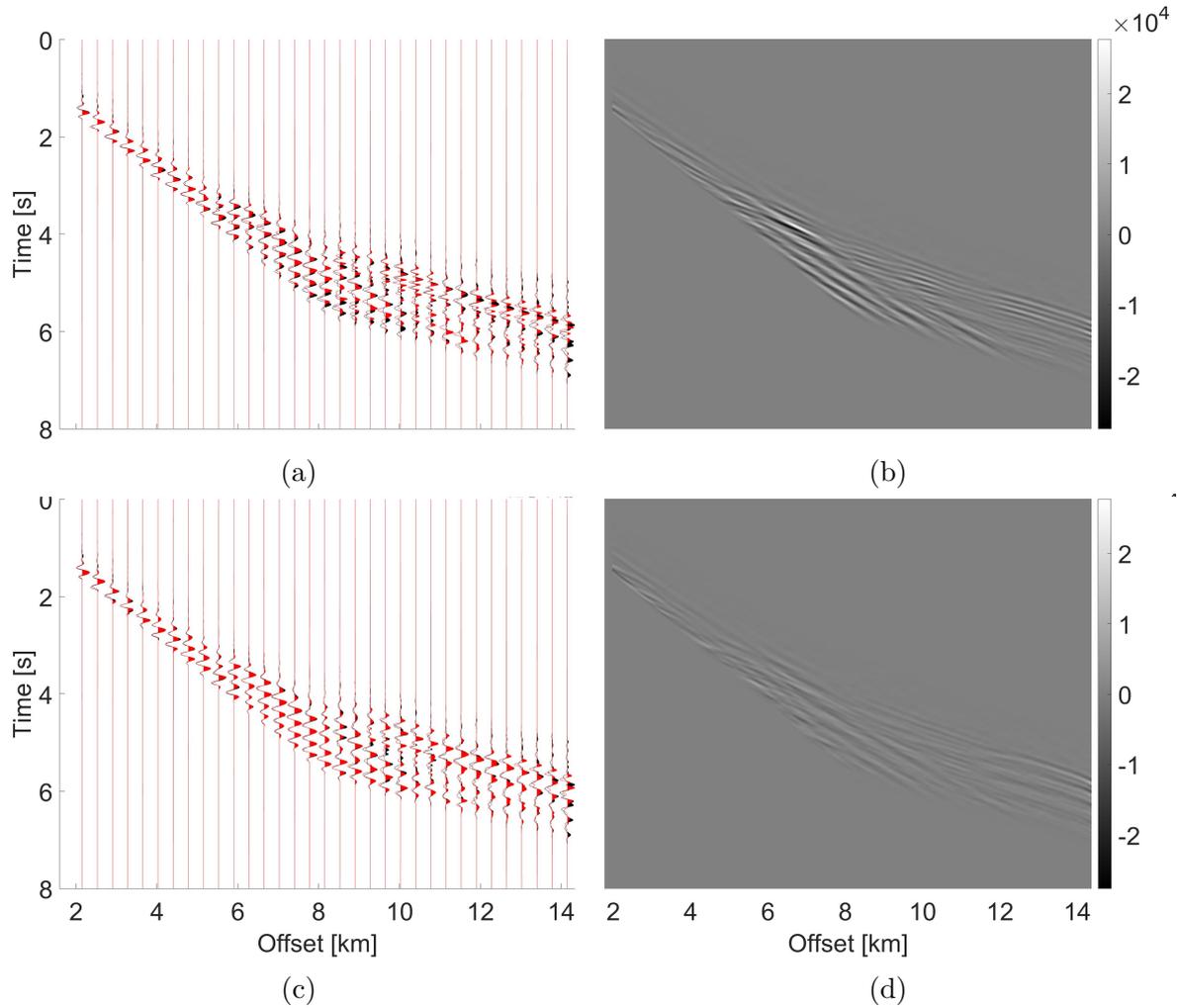


Figure J.2: Stage 2. Data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.

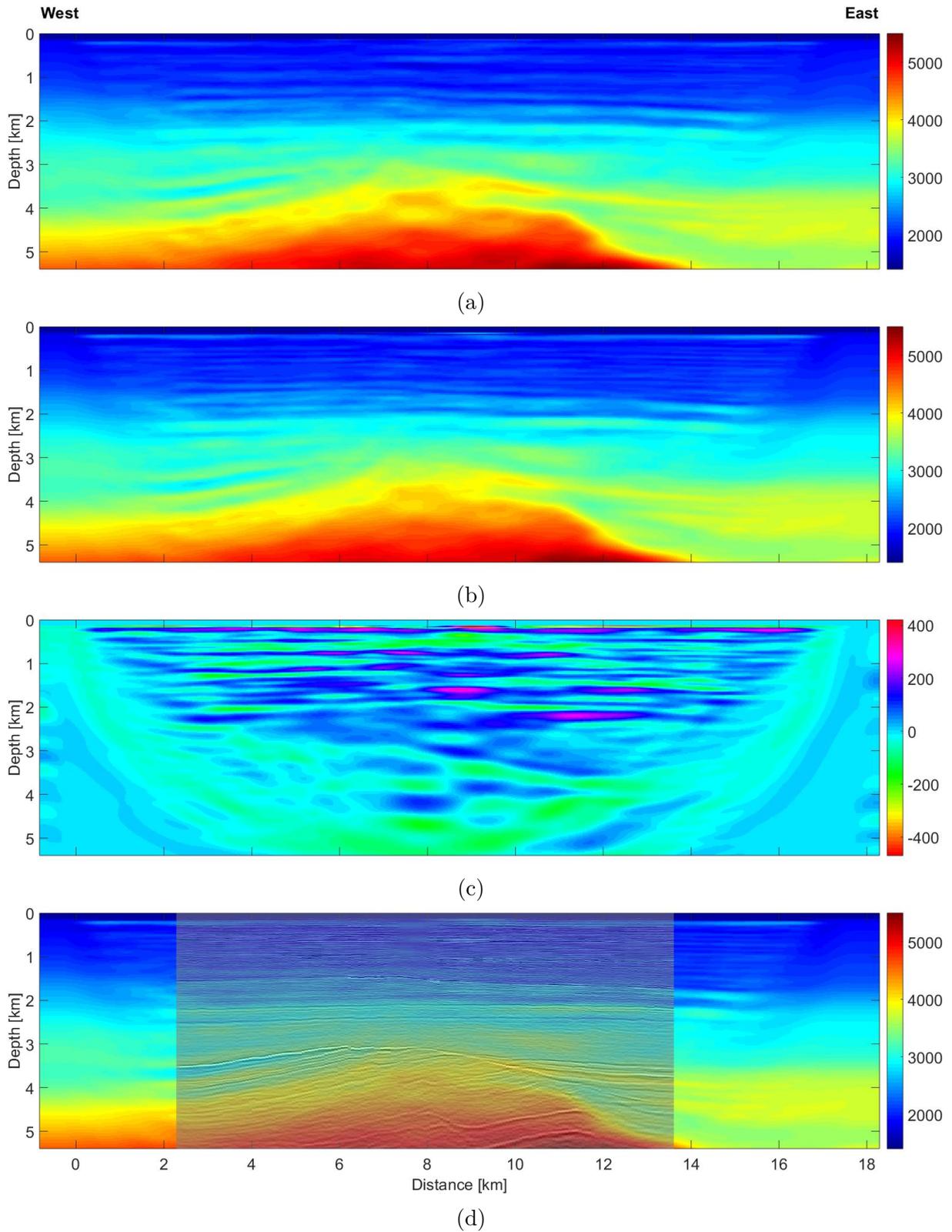


Figure J.3: Stage 2. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

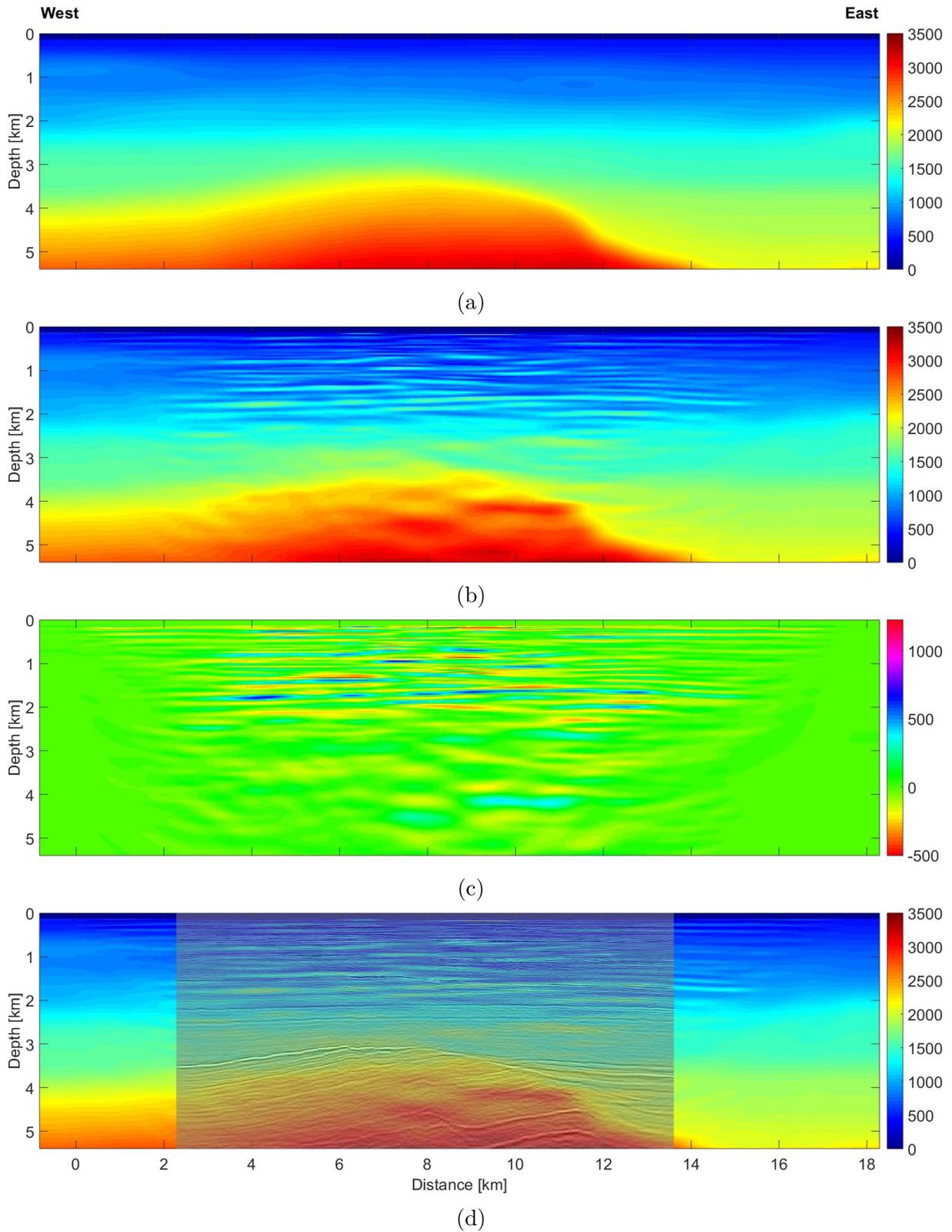


Figure J.4: Stage 2. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

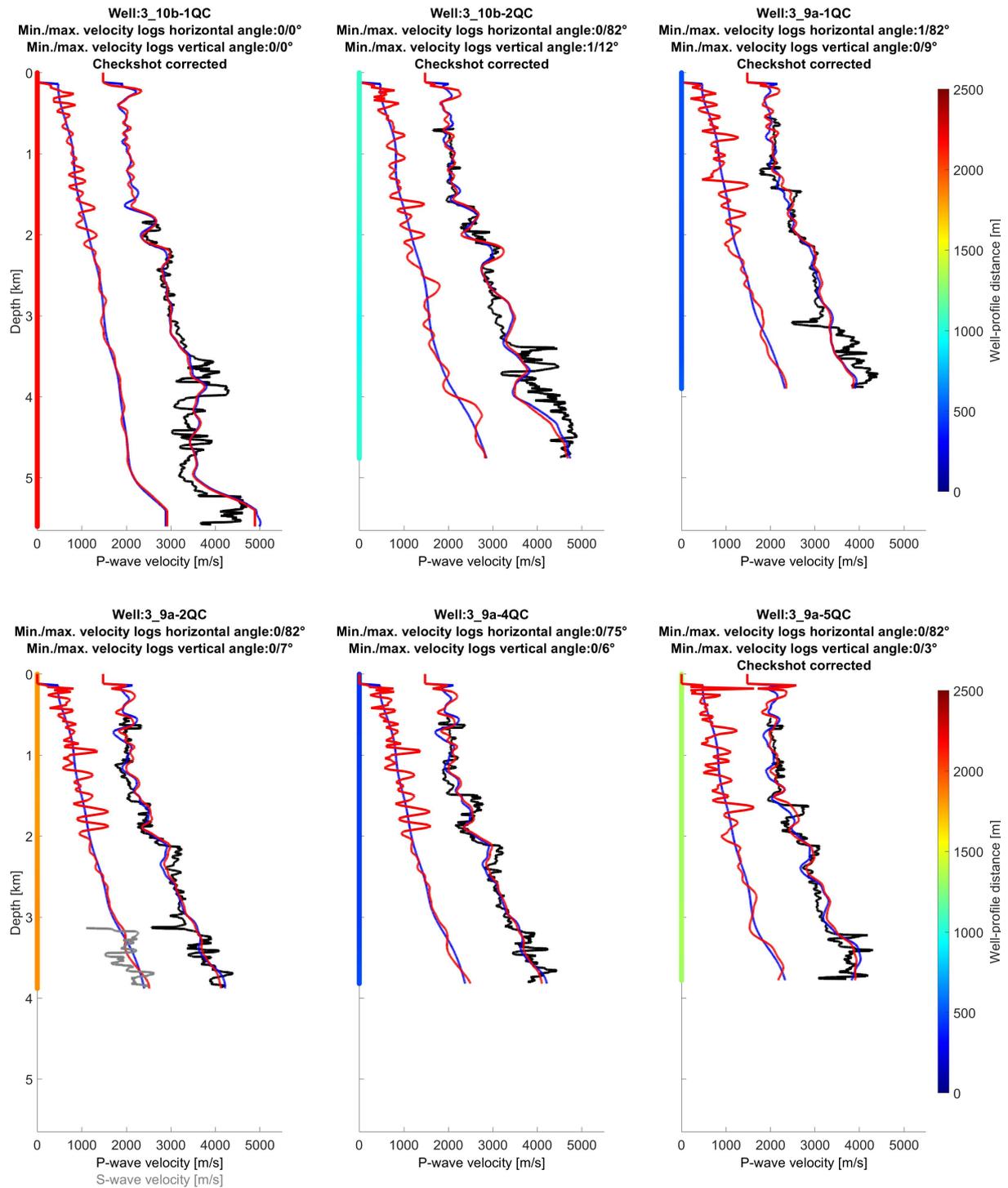


Figure J.5: Stage 2. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

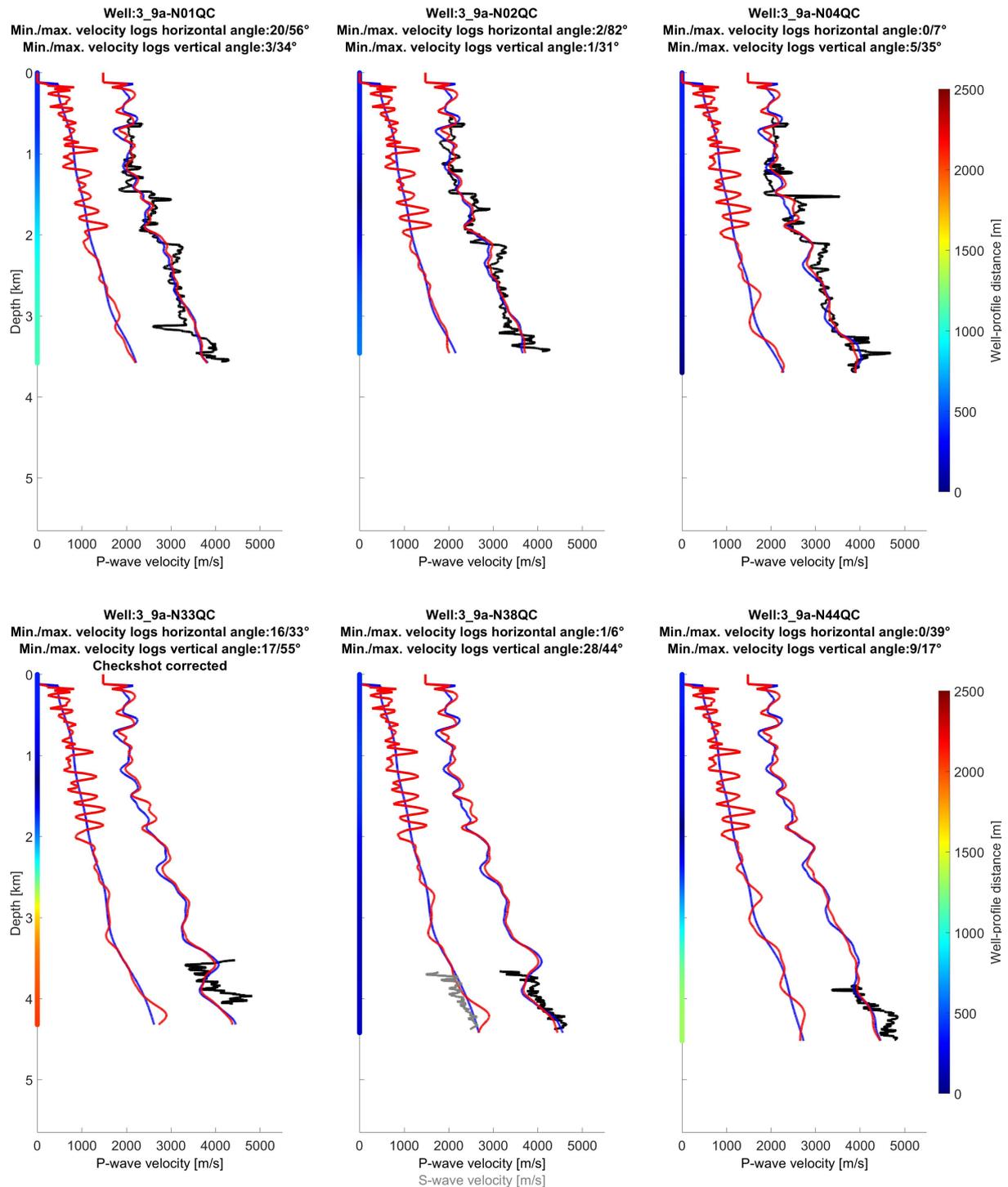


Figure J.6: Stage 2. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

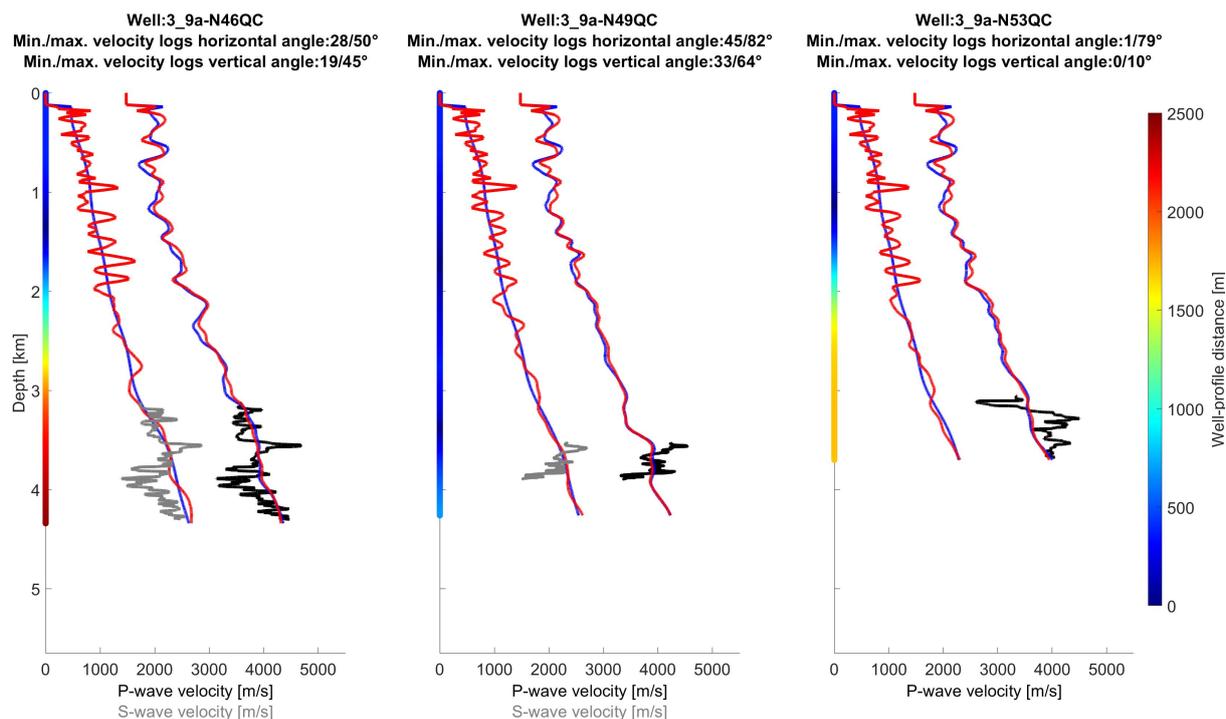


Figure J.7: Stage 2. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

J.2 Stage 3 - first frequency range

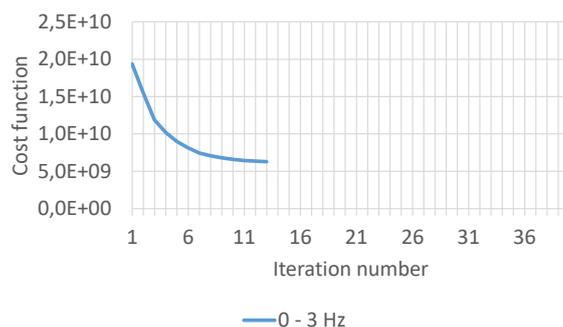


Figure J.8: Stage 3. Non-normalized cost function for the first frequency range.

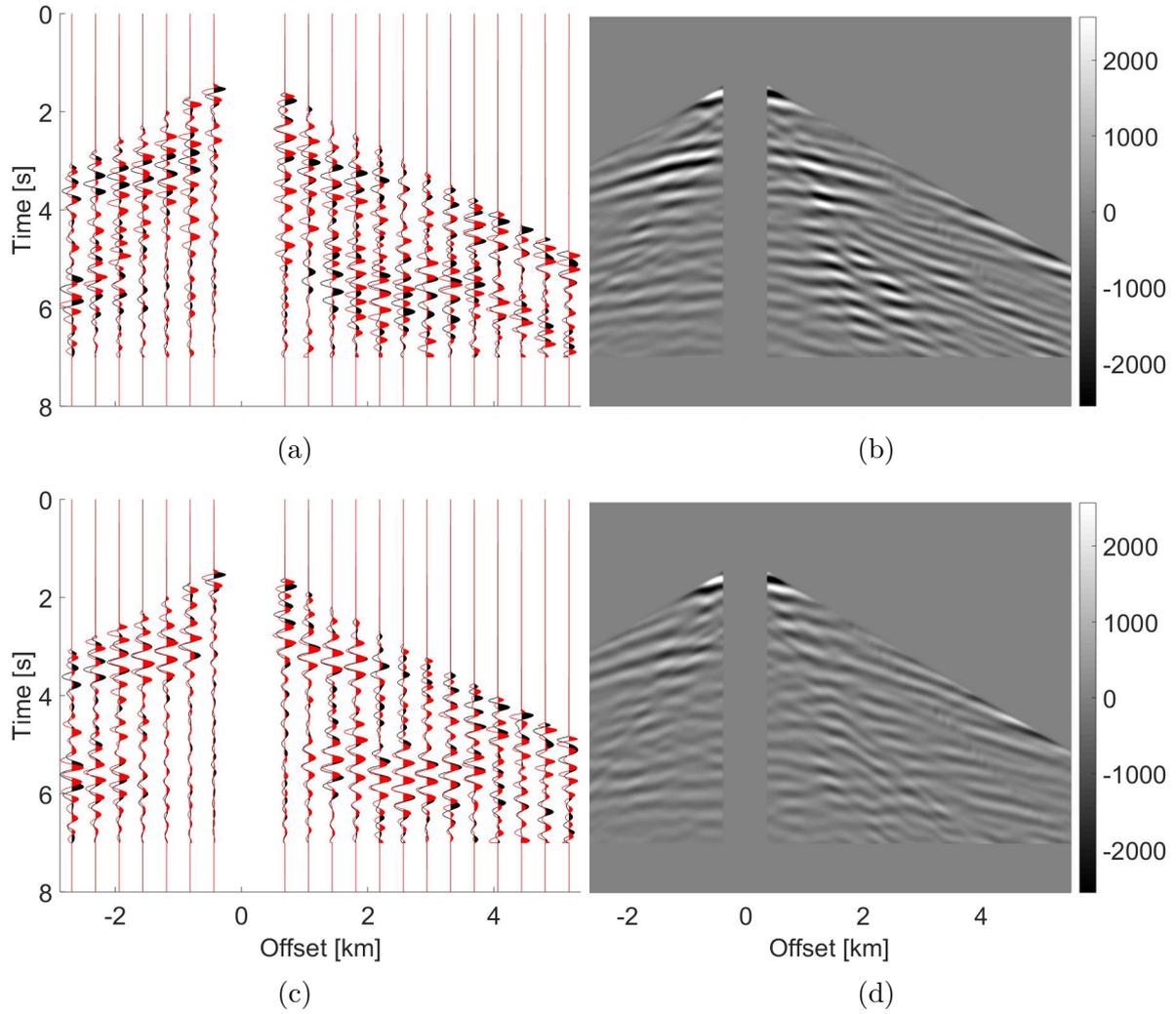


Figure J.9: Stage 3 first frequency range. Data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most horizontal velocity component receiver gather.

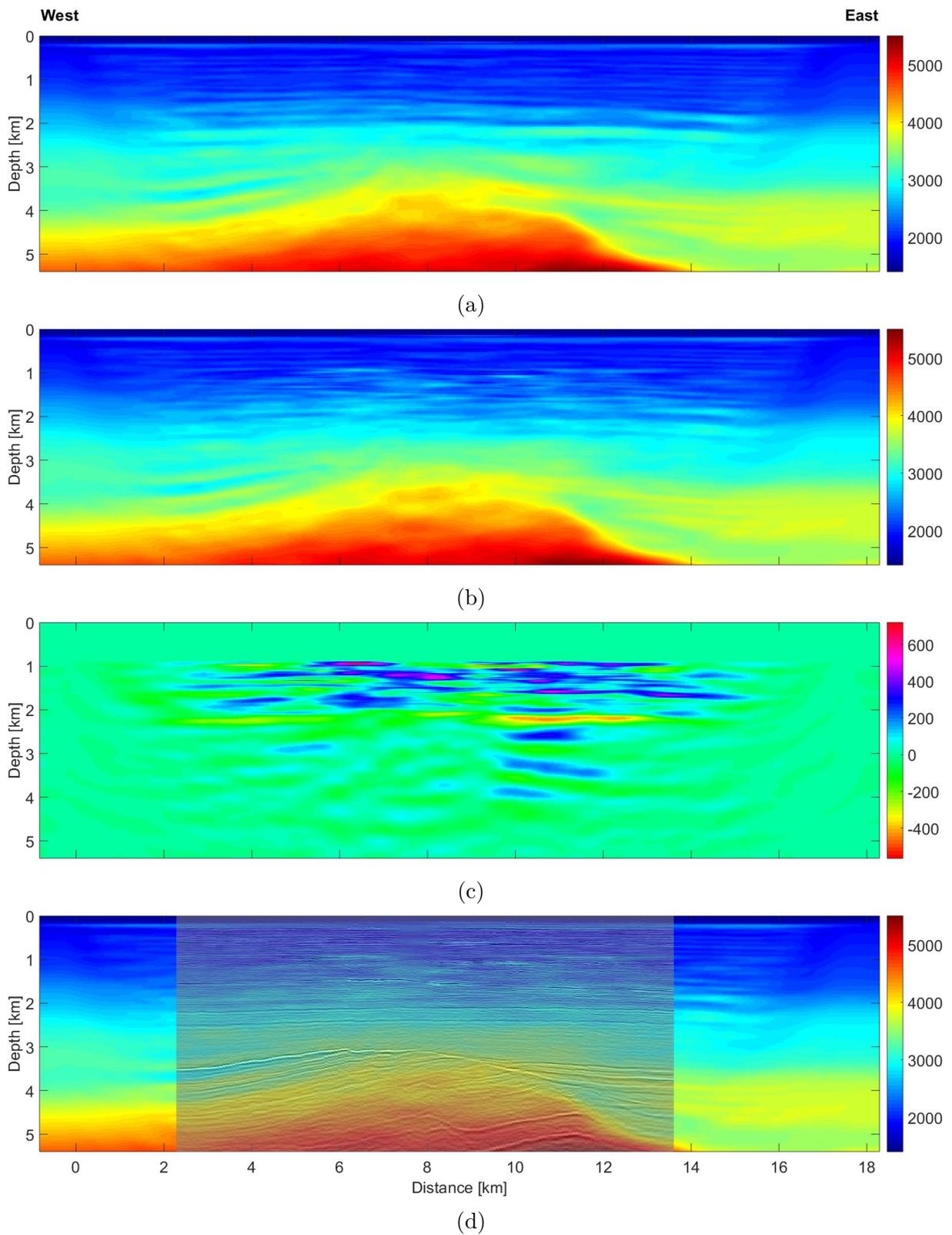


Figure J.10: Stage 3 first frequency range. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

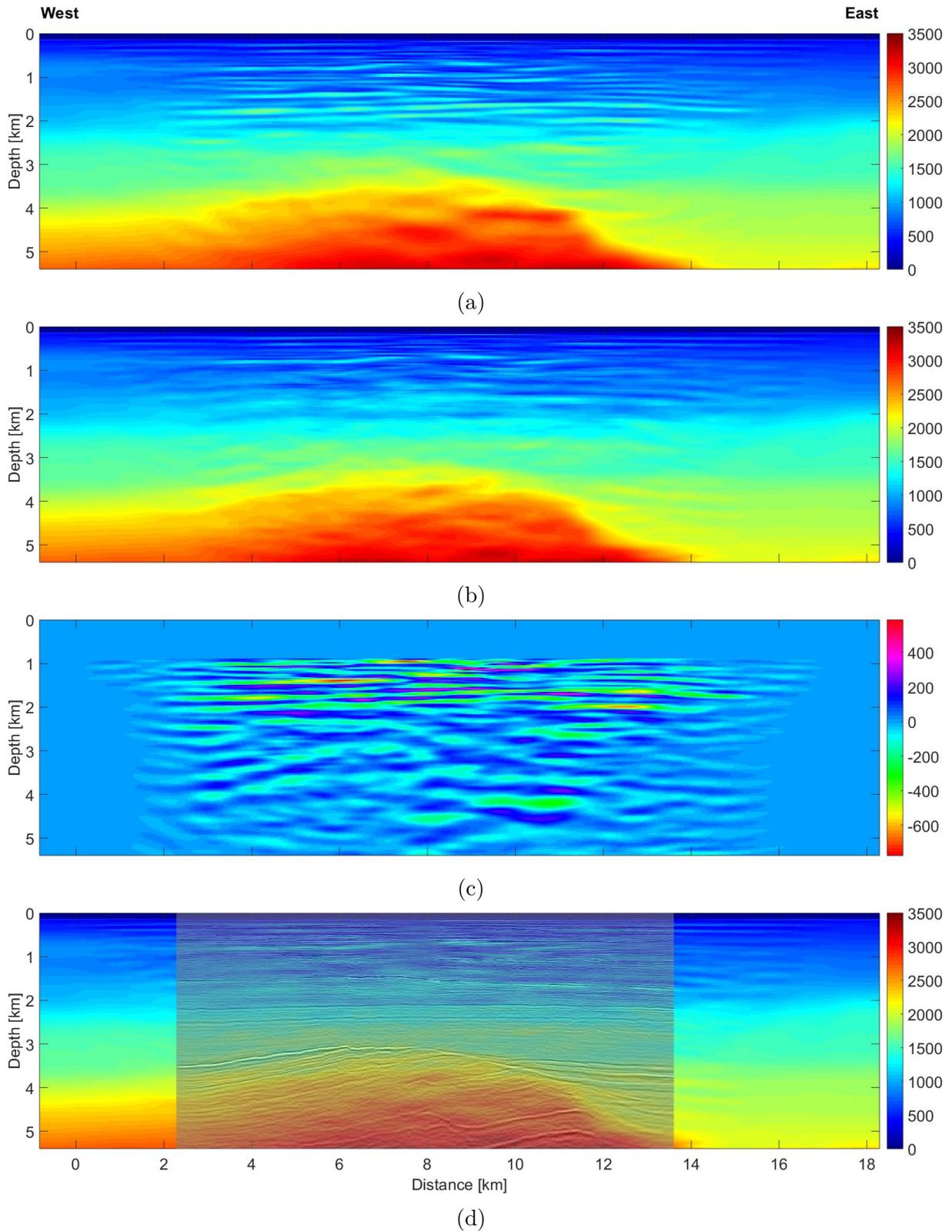


Figure J.11: Stage 3 first frequency range. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

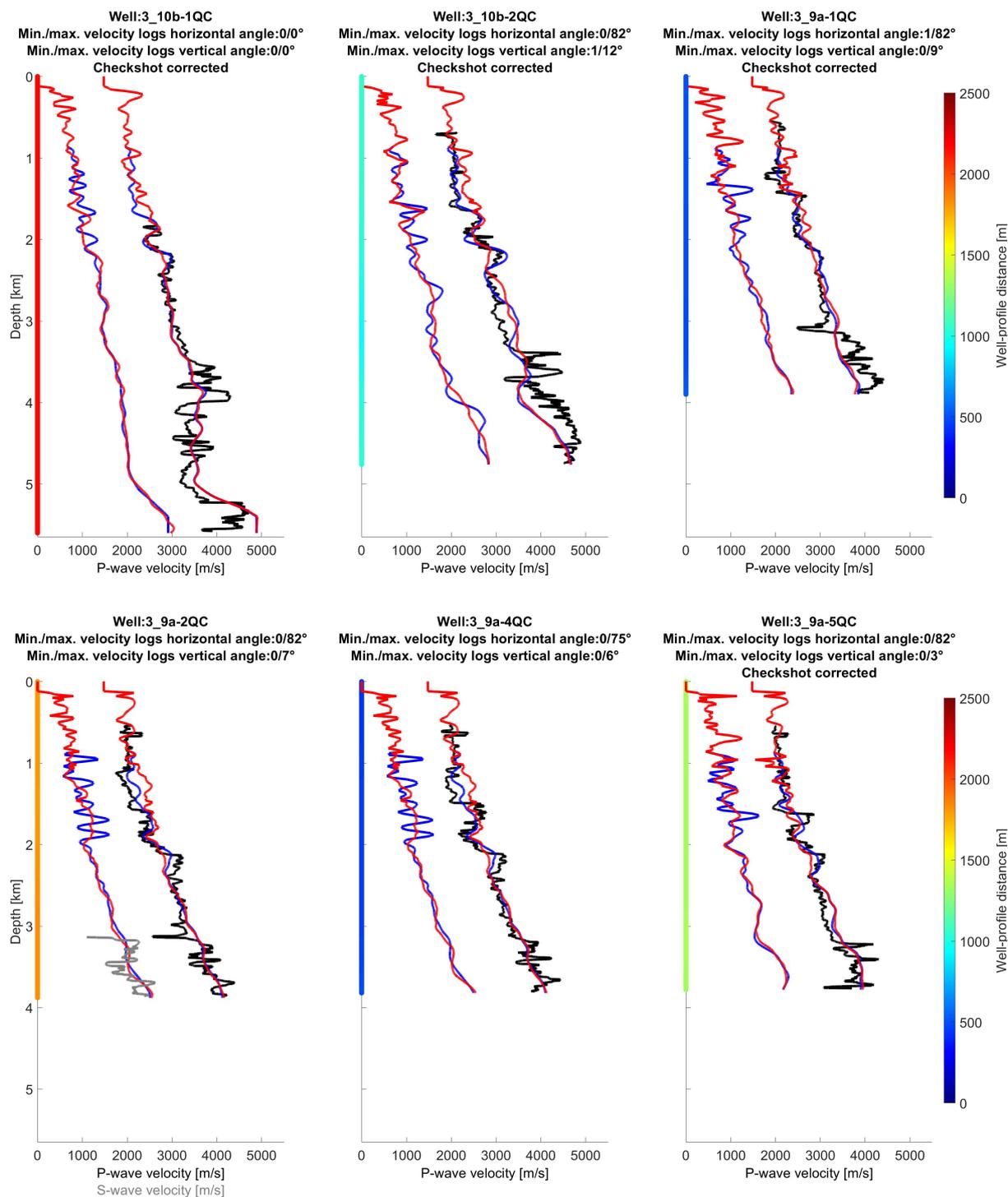


Figure J.12: Stage 3 first frequency range. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

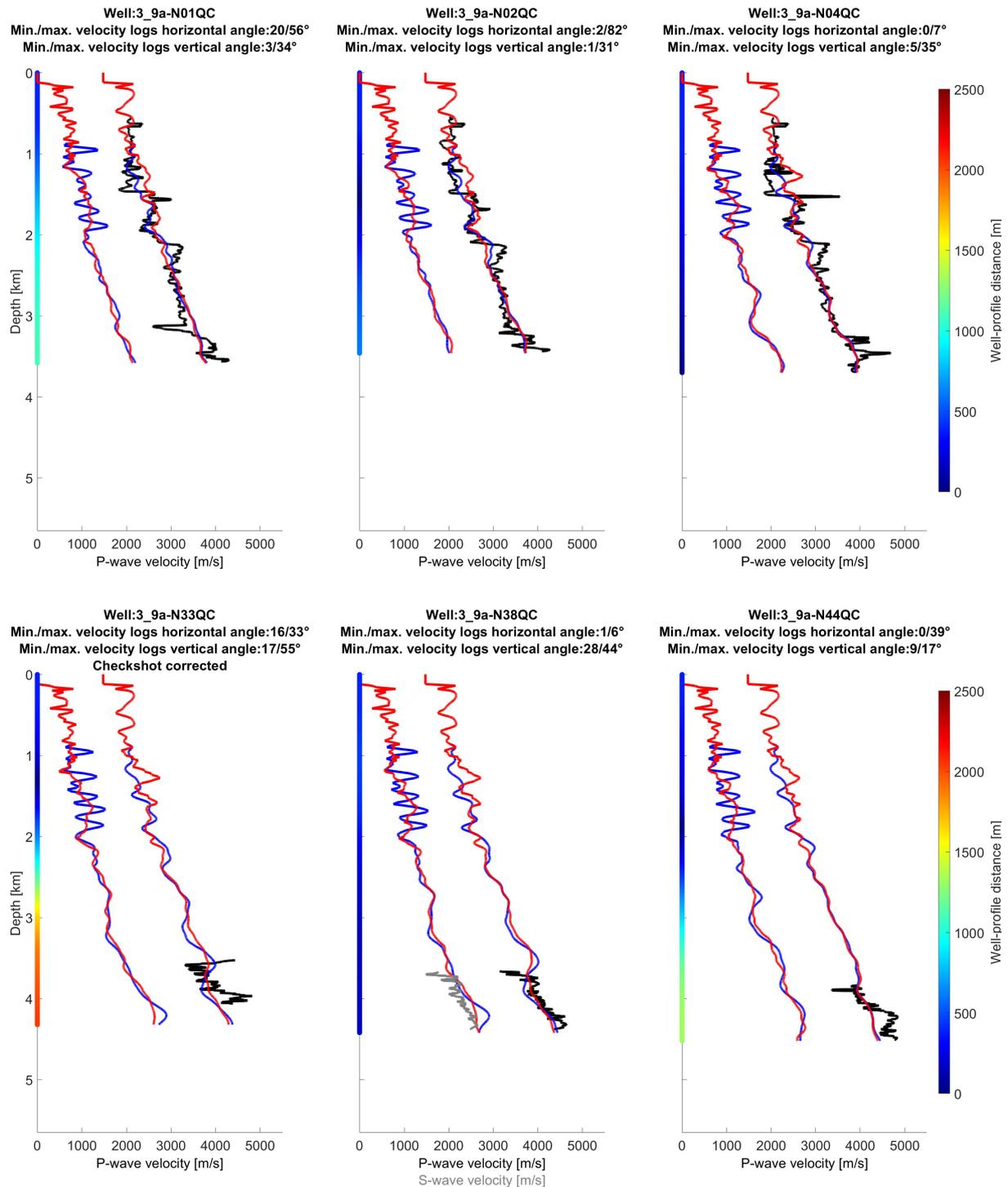


Figure J.13: Stage 3 first frequency range. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

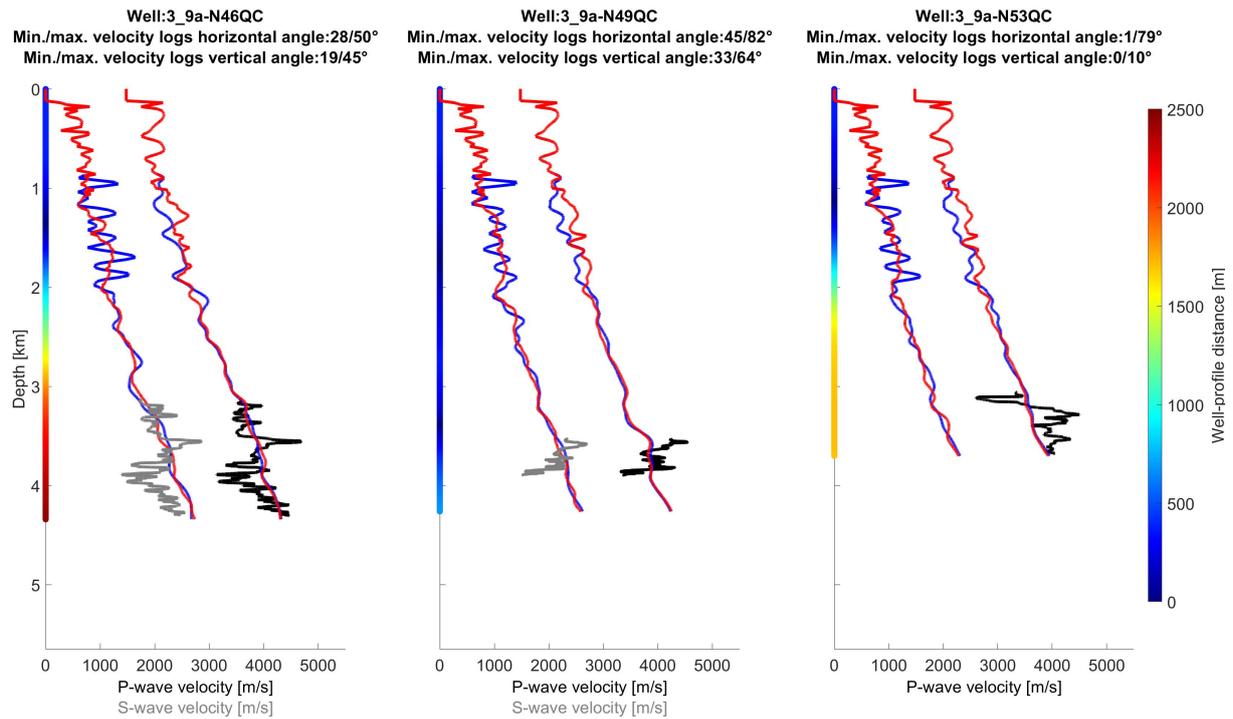


Figure J.14: Stage 3 first frequency range. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

Appendix K

Alwyn real data: original strategy with a modified stage order additional displays

Contents

K.1 Stage 2'	211
K.2 Stage 1b' - first frequency range	217

Stage by stage results of the original strategy with a modified stage order are provided here: the non-normalized cost functions, the well comparisons, the models inversion results with comparison with the vintage migrated image, as well as observed and initial/final modeled data comparisons with trace by trace normalized displays, together with the corresponding residual displays (i.e., observed data - modeled data).

K.1 Stage 2'

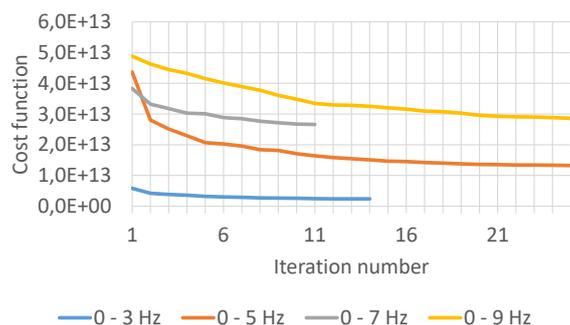


Figure K.1: Stage 2'. Non-normalized cost function for each frequency range.

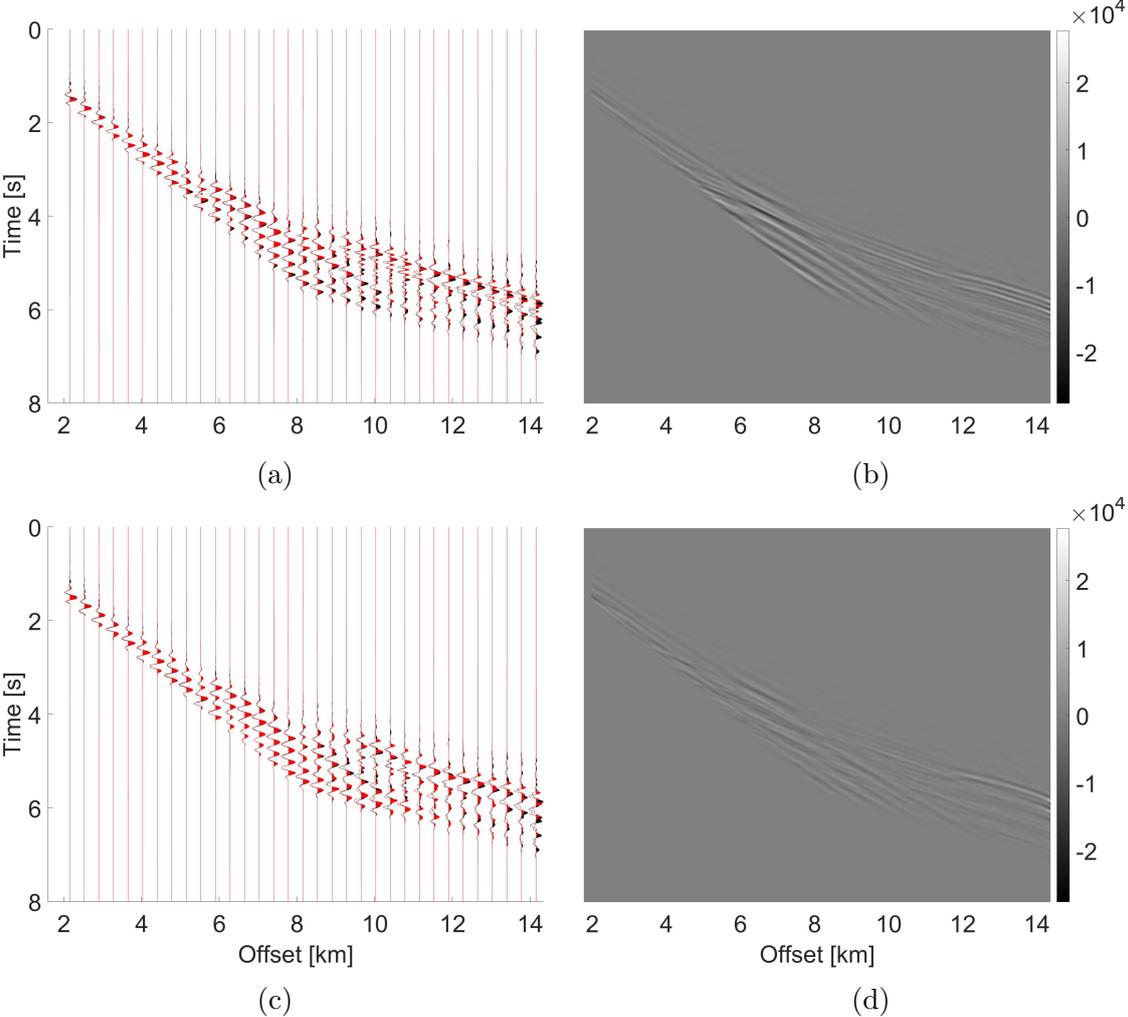


Figure K.2: Stage 2'. Data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.

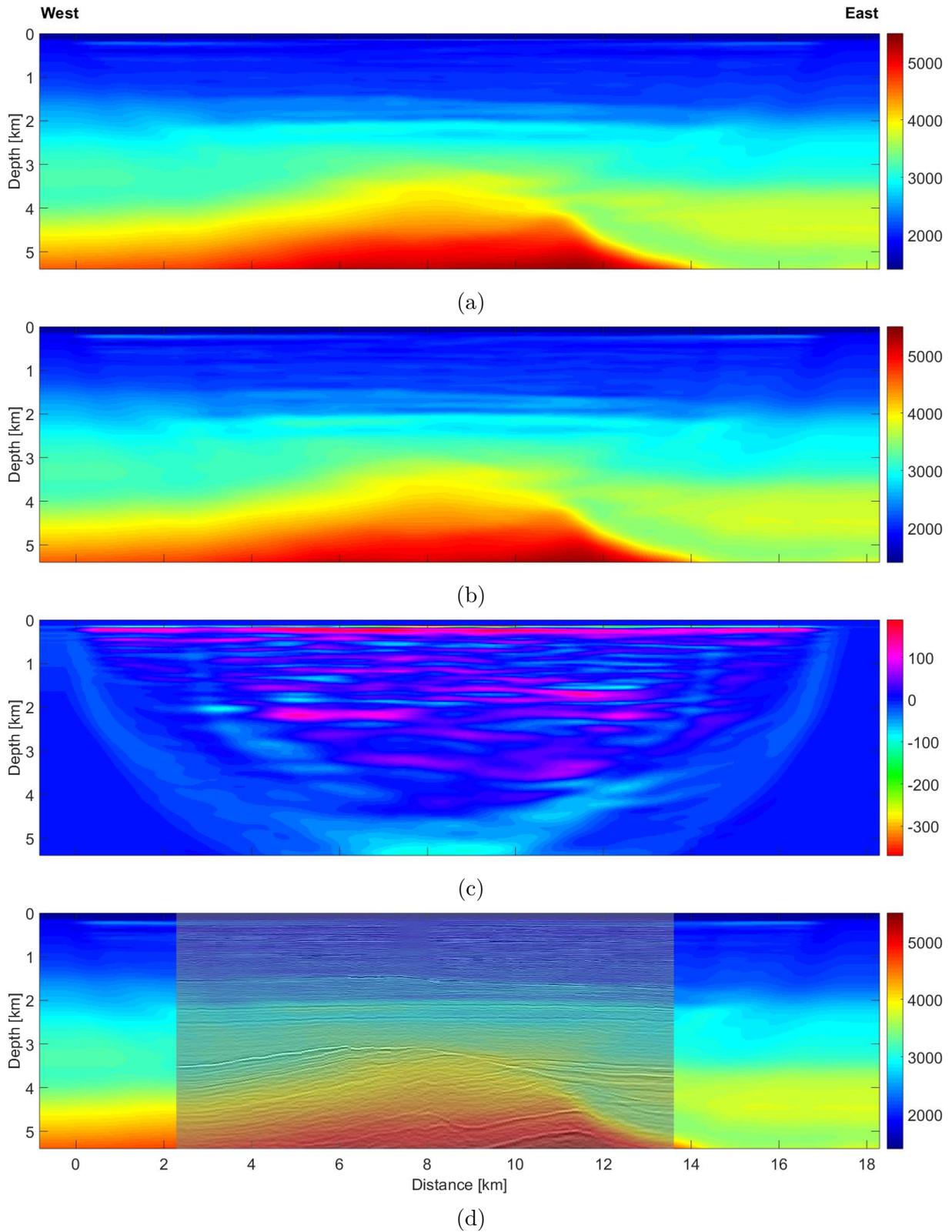


Figure K.3: Stage 2'. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

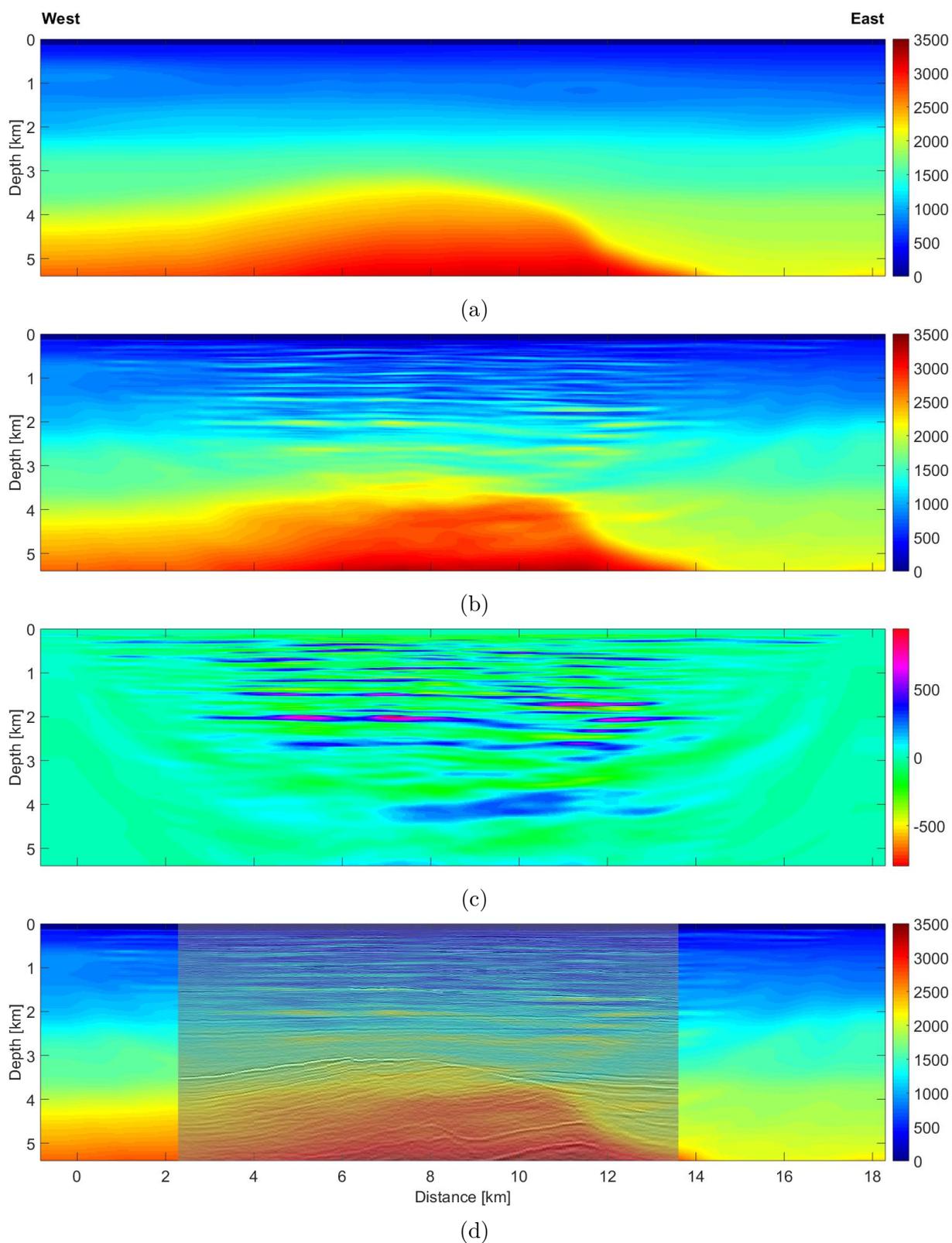


Figure K.4: Stage 2'. S-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

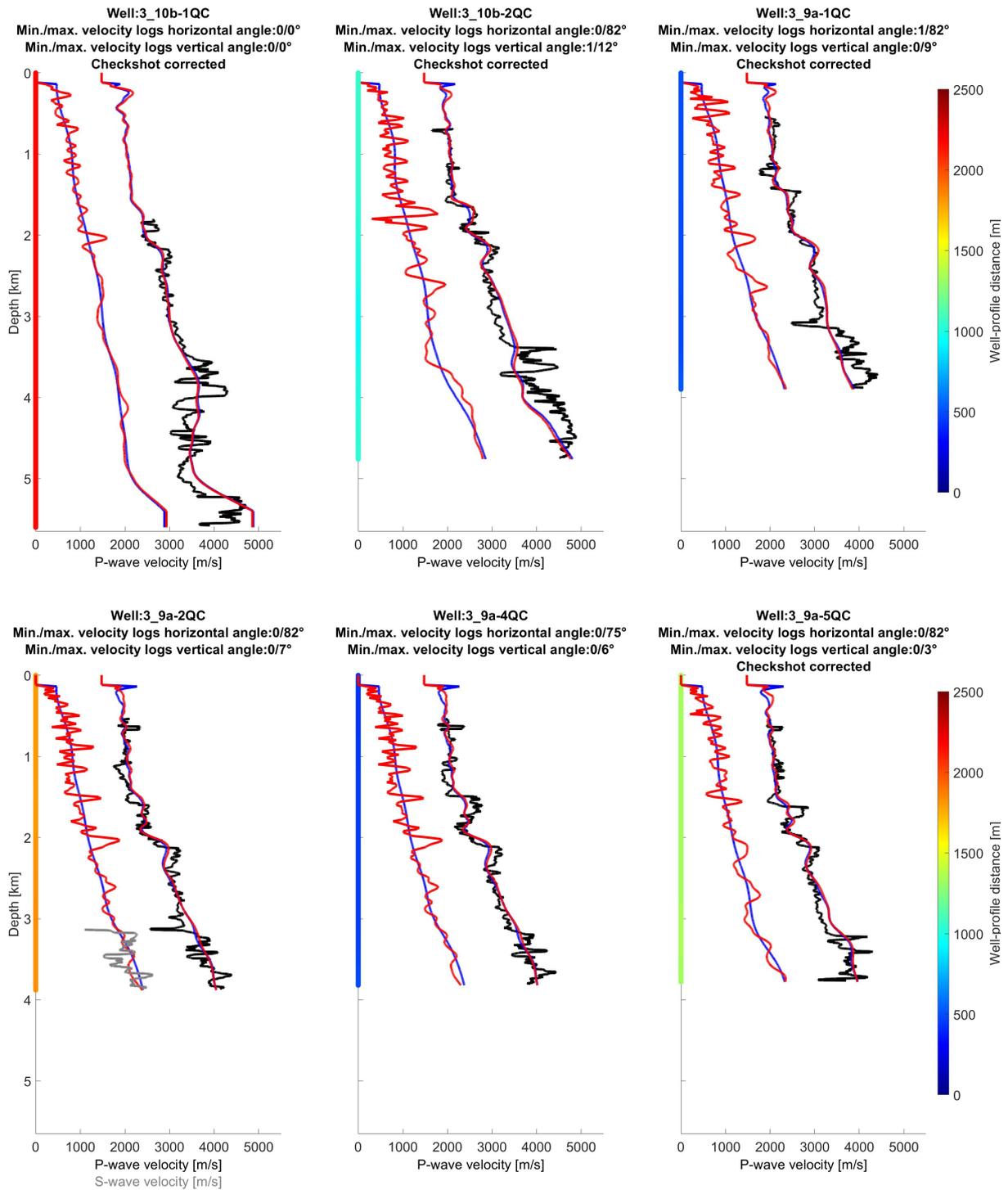


Figure K.5: Stage 2'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black and grey lines respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

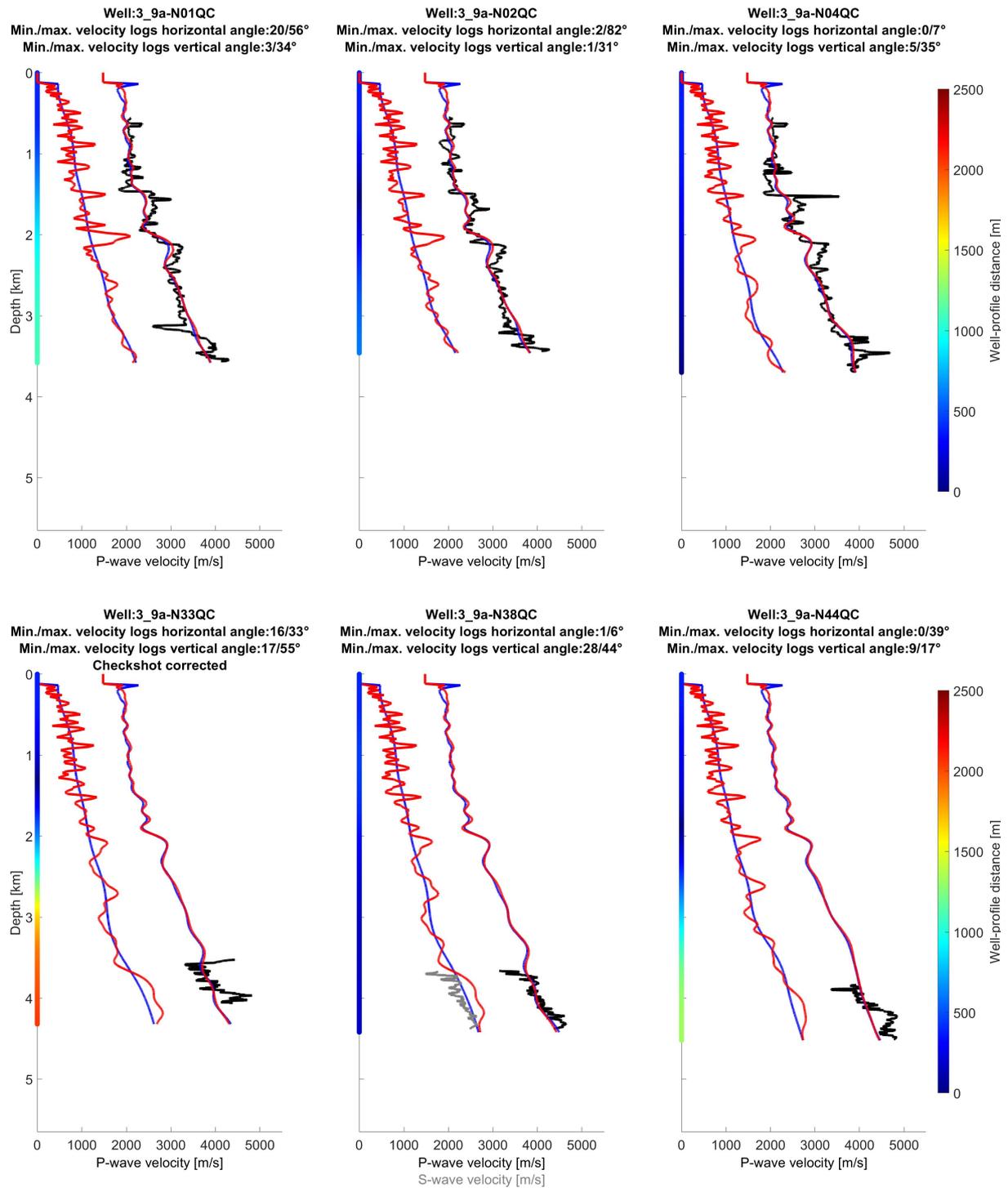


Figure K.6: Stage 2'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

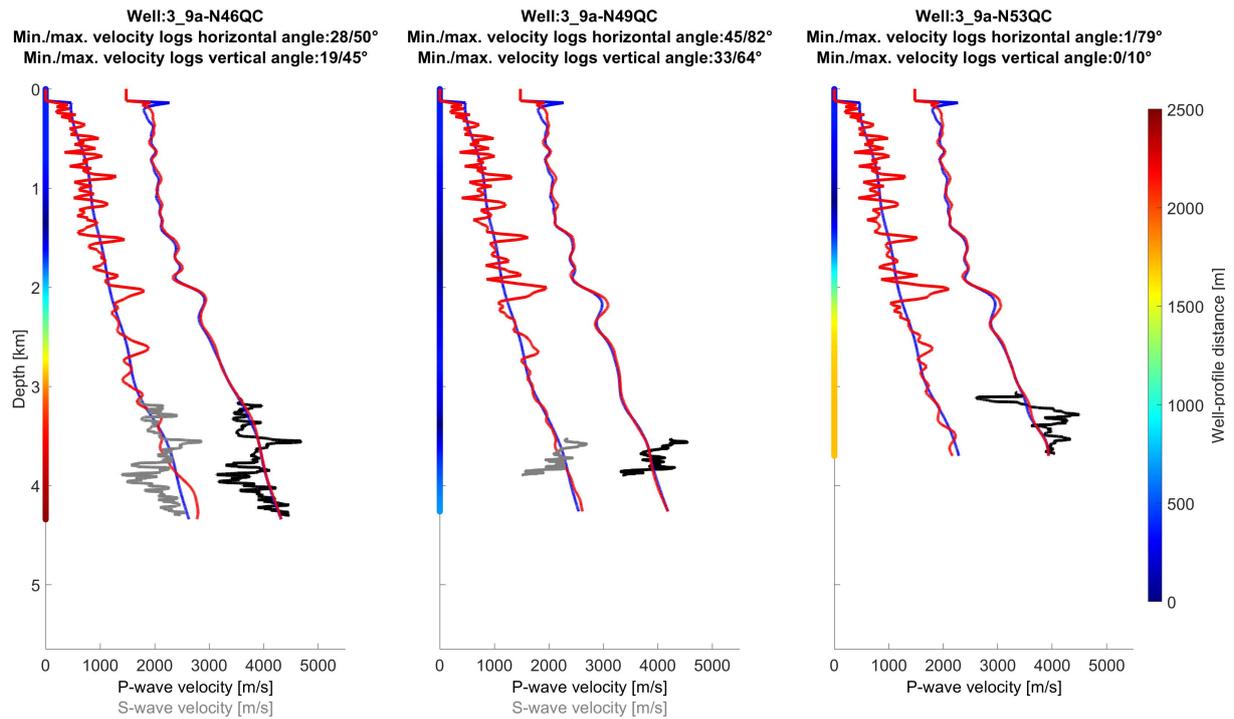


Figure K.7: Stage 2'. Initial (thin blue line) and final (thin red line) P- and S-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black and grey line respectively). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

K.2 Stage 1b' - first frequency range

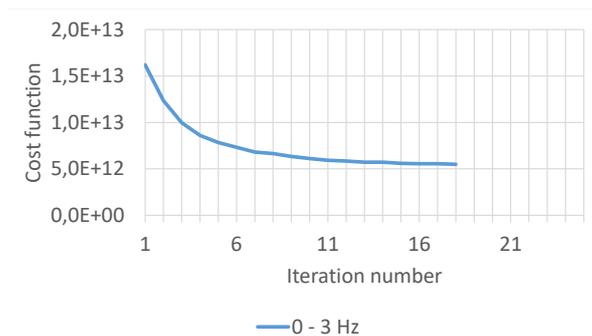


Figure K.8: Stage 1b'. Non-normalized cost function for the first frequency range.

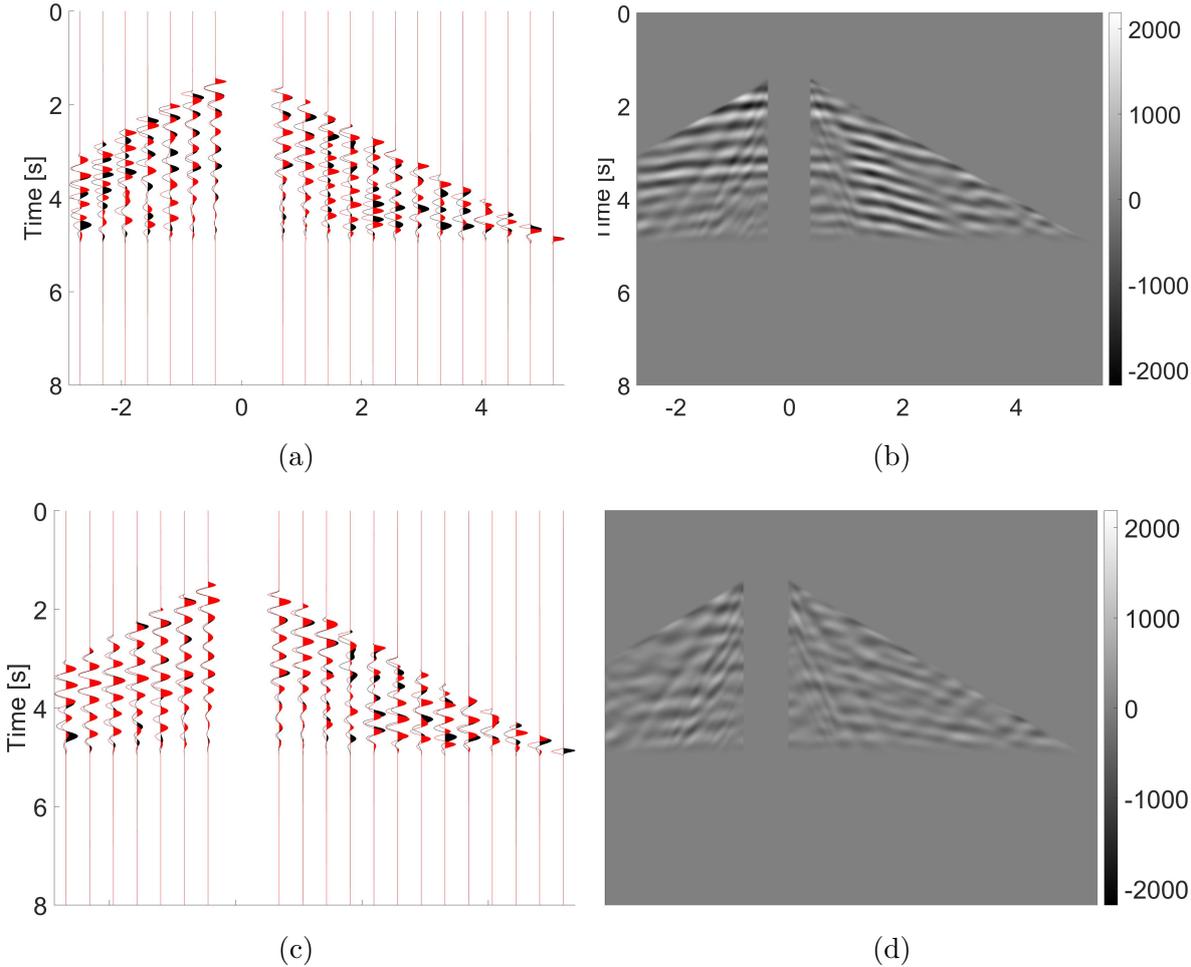


Figure K.9: Stage 1b' first frequency range. data inversion results. (a) Initial data fit; (b) Initial residual; (c) Final data fit; (d) Final residual. For a and c, the traces are normalized and the observed data are in black while the modeled data are in red. Western most pressure receiver gather.

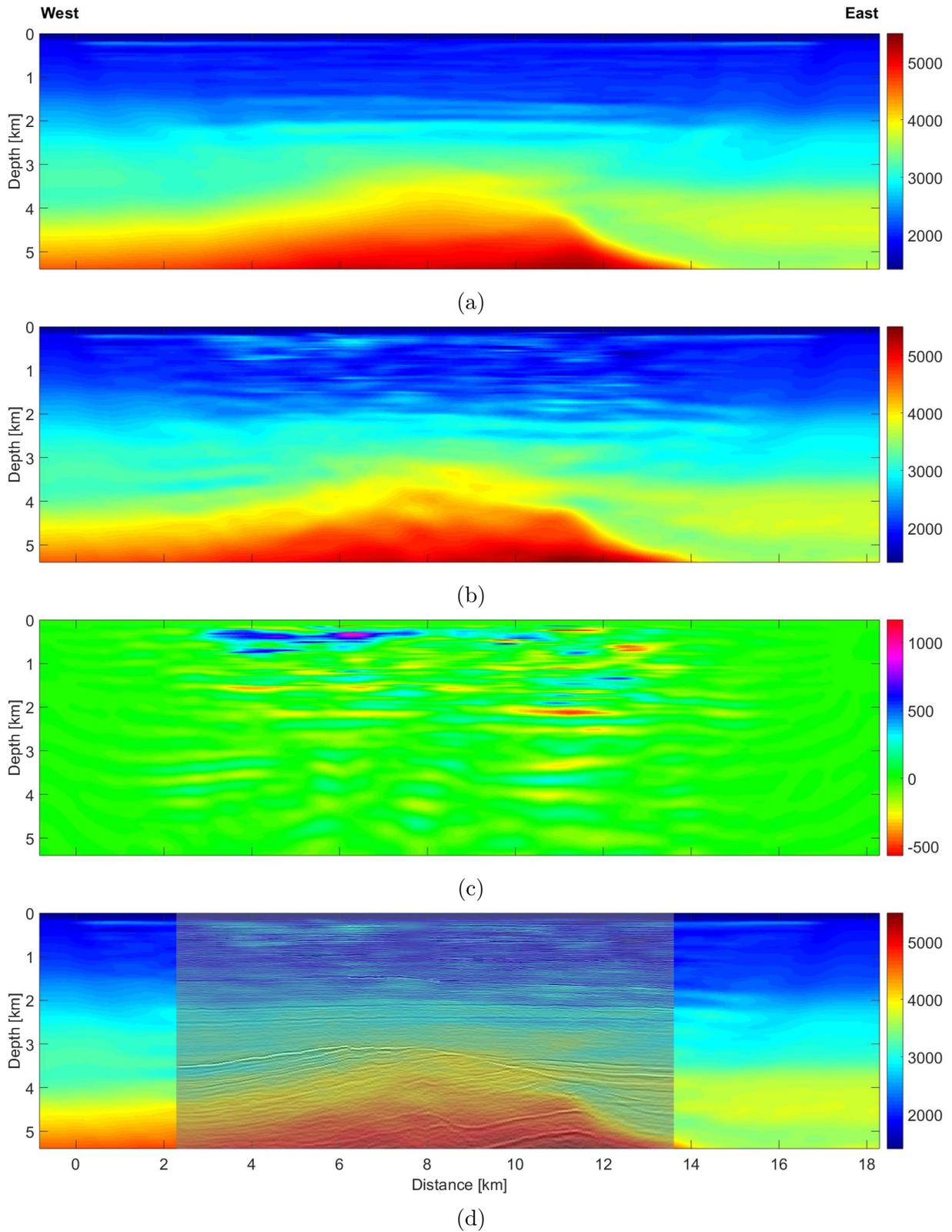


Figure K.10: Stage 1b' first frequency range. P-wave vertical velocity model inversion results. (a) Initial model; (b) Final model; (c) b-a; (d) Vintage migrated image overlaid above b.

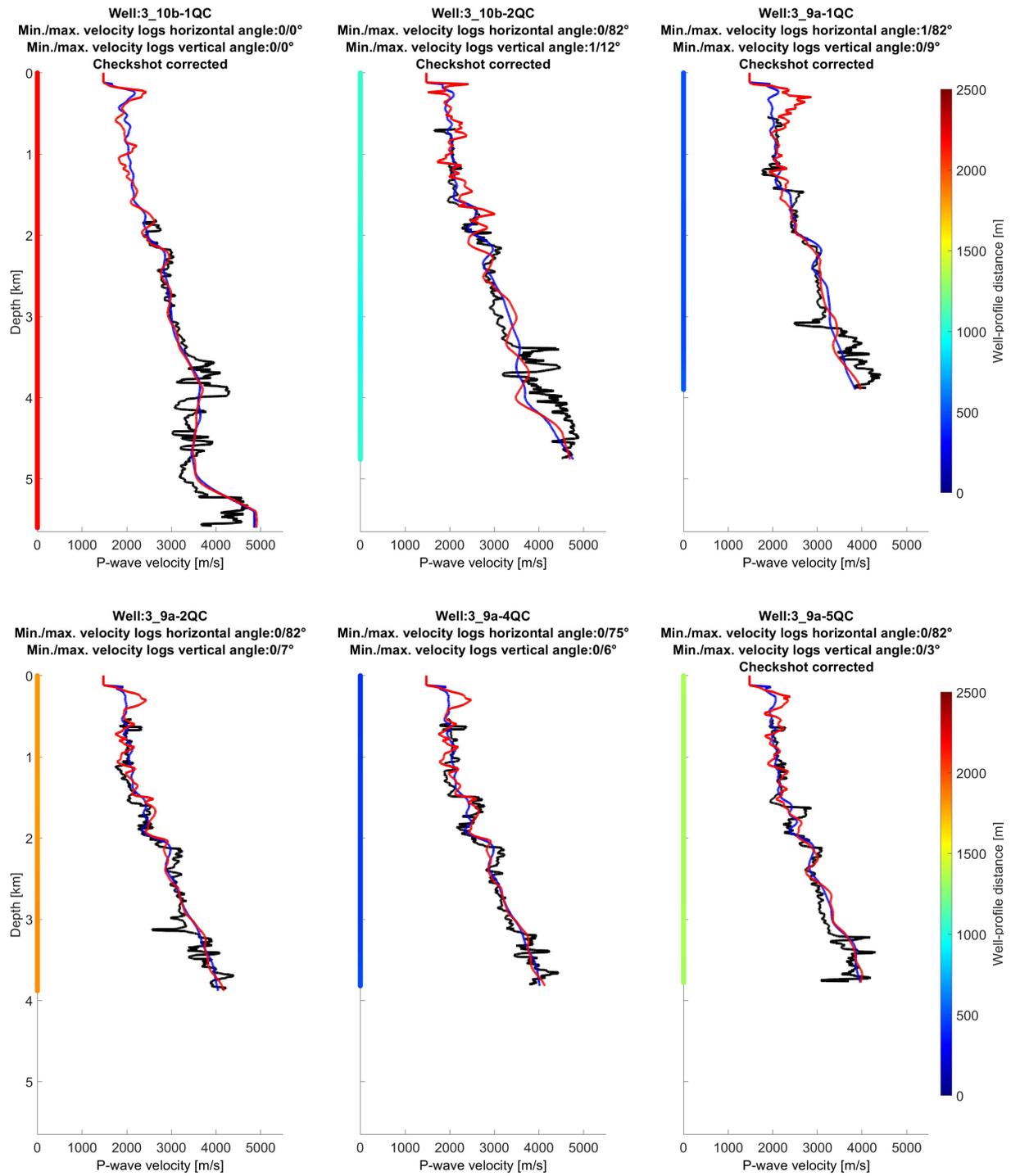


Figure K.11: Stage 1b' first frequency range. Initial (thin blue line) and final (thin red line) P-wave vertical velocity models compared with smoothed velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

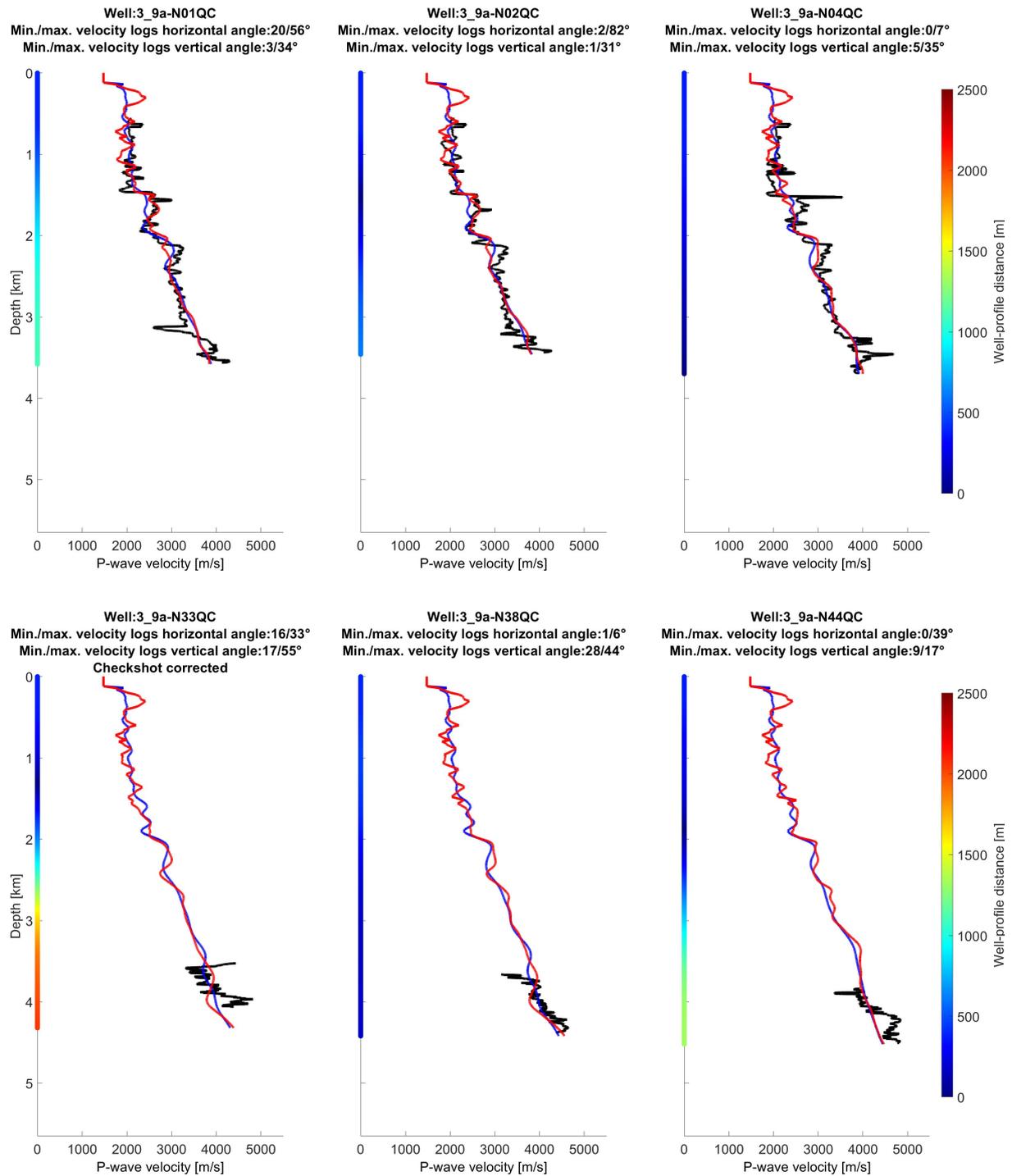


Figure K.12: Stage 1b' first frequency range. Initial (thin blue line) and final (thin red line) P-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations.

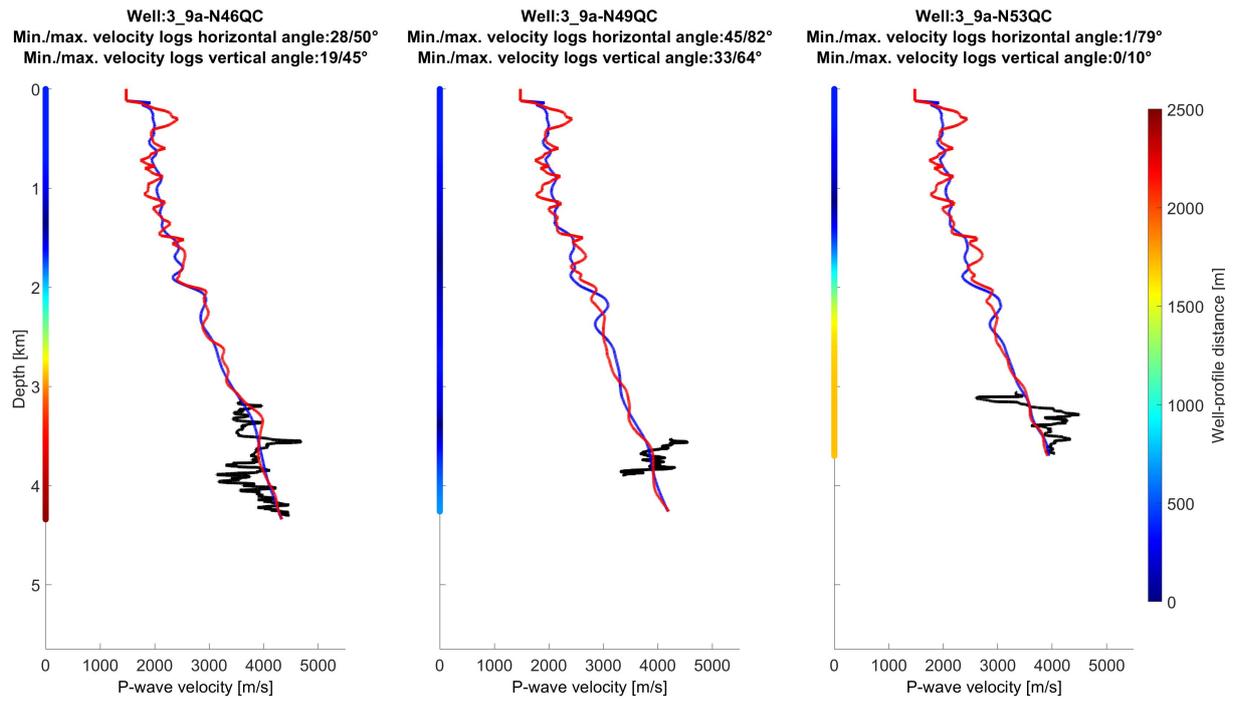


Figure K.13: Stage 1b' first frequency range. Initial (thin blue line) and final (thin red line) P-wave vertical velocity model compared with smoothed velocities from a selection of wells (thin black line). The smoothing is performed with a 40.5-meter-long vertical median filter. The multicolored vertical line provides the horizontal distance between the well and the profile. See figure 4.20 for well logs locations

Bibliography

- (1) Aghamiry H. S., Gholami A., and Operto S. (2019). Implementing bound constraints and total-variation regularization in extended full-waveform inversion with the alternating direction method of multiplier: application to large contrast media. *Geophysical Journal International* 218(2), 855–872.
- (2) Agudo Ò. C., Vieira da Silva N., Warner M., and Morgan J. (2018). Acoustic full-waveform inversion in an elastic world. *Geophysics* 83(3), R257–R271.
- (3) Aki K., and Richards P. G., *Quantitative Seismology*, Second Edition; University Science Books: 2002.
- (4) Alkhalifah T., and Tsvankin I. (1995). Velocity analysis for transversely isotropic media. *Geophysics* 60(5), 1550–1566.
- (5) Alves G. C. Elastic full waveform inversion of multicomponent data, Ph.D. Thesis, Stanford University, 2017.
- (6) Amestoy P., Brossier R., Buttari A., L’Excellent J.-Y., Mary T., Métivier L., Miniussi A., and Operto S. (2015). Fast 3D frequency-domain full-waveform inversion with a parallel block low-rank multifrontal direct solver: Application to OBC data from the North Sea. *Geophysics* 81(6), R363–R383.
- (7) Arecco M. A., Ruiz F., Pizarro G., Giménez M., Martínez P., and Ramos V. A. (2016). Gravimetric determination of the continental–oceanic boundary of the Argentine continental margin (from 36°S to 50°S). *Geophysical Journal International* 204(1), 366–385.
- (8) Arntsen B., and Carcione J. M. (2000). A new insight into the reciprocity principle. *Geophysics* 65(5), 1604–1612.
- (9) Arnulf A. F., Singh S. C., Harding A. J., Kent G. M., and Crawford W. (2011). Strong seismic heterogeneity in layer 2A near hydrothermal vents at the Mid-Atlantic Ridge. *Geophysical Research Letters* 38(13), L13320.
- (10) Arnulf A. F., Harding A. J., Singh S. C., Kent G. M., and Crawford W. (2012). Fine-scale velocity structure of upper oceanic crust from full waveform inversion of downward continued seismic reflection data at the Lucky Strike Volcano, Mid-Atlantic Ridge. *Geophysical Research Letters* 39(8), L08303.
- (11) Arnulf A. F., Harding A. J., Singh S. C., Kent G. M., and Crawford W. (2014). Nature of upper crust beneath the Lucky Strike volcano using elastic full waveform inversion of streamer data. *Geophysical Journal International* 196(3), 1471–1491.
- (12) Asnaashari A., Brossier R., Garambois S., Audebert F., Thore P., and Virieux J. (2013). Regularized seismic full waveform inversion with prior model information. *Geophysics* 78(2), R25–R36.

- (13) Audhkhasi P., and Singh S. C. (2019). Seismic structure of the upper crust from 0–75 Ma in the equatorial Atlantic ocean on the African plate using ultralong offset seismic data. *Geochemistry, Geophysics, Geosystems* 20(12), 6140–6162.
- (14) Auer L., Nuber A. M., Greenhalgh S. A., Maurer H., and Marelli S. (2013). A critical appraisal of asymptotic 3D-to-2D data transformation in full-waveform seismic crosshole tomography. *Geophysics* 78(6), R235–R247.
- (15) Bale R., Marchand T., Wilkinson K., Wikel K., and Kendall R. (2013). The signature of shear-wave splitting: Theory and observations on heavy oil data. *The Leading Edge* 32(1), 14–24.
- (16) Barnes C., and Charara M. (2009). Full-waveform inversion results when using acoustic approximation instead of elastic medium. *Geophysics* 74(6), WCC91–WCC103.
- (17) Belahi T., Fuji N., and Singh S. C. (2015). Elastic Versus Viscoelastic Full Waveform Inversion of Near-offset and Wide-angle Data in the Presence of Attenuation. *Conference Proceedings: 77th EAGE Conference & Exhibition 2015*.
- (18) Berryman J. G., Berge P. A., and Bonner B. P. (2002). Estimating rock porosity and fluid saturation using only seismic velocities. *Geophysics* 67(2), 391–404.
- (19) Bergslid T. S., Raknes E. B., and Arntsen B. (2015). The influence of anisotropy on elastic full-waveform inversion. *SEG Technical Program Expanded Abstracts*, 1425–1429.
- (20) Berryhill J. R. (1984). Wave-equation datuming before stack. *Geophysics* 49(11), 2064–2066.
- (21) Berenger J. P. (1994). A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics* 114(2), 185–200.
- (22) Betti E. (1872). Teoria della elasticita. *Il nuovo Cimento* 7-8, 5–21.
- (23) Beydoun W. B., and Tarantola A. (1988). First Born and Rytov approximation: Modeling and inversion conditions in a canonical example. *The Journal of the Acoustical Society of America* 83(3), 1045–1055.
- (24) Bickel S. M. (1990). Velocity-depth ambiguity of reflection traveltimes. *Geophysics* 55(3), 266–276.
- (25) Bleistein N. (1986). Two-and-one-half dimensional in-plane wave propagation. *Geophysical Prospecting* 34(5), 686–703.
- (26) Boiero D., Strobba C., Velasco L., and Vermeer P. (2013). Guided Waves - Inversion and Attenuation. *Conference Proceedings: 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*.
- (27) Boiero D., Wiarda E., and Vermeer P. (2013). Surface- and guided-wave inversion for near-surface modeling in land and shallow marine seismic data. *The Leading Edge* 32(6), 638–646.
- (28) Boiero D., Wiarda E., West L. R., and Vermeer P. (2014). Near-surface Modelling in Shallow Marine Environments Using Surface and Guided Waves. *Conference Proceedings: 76th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*.

- (29) Borisov D., and Singh S. C. (2015). Three-dimensional elastic full waveform inversion in a marine environment using multicomponent ocean-bottom cables: a synthetic study. *Geophysical Journal International*, 1215–1234.
- (30) Borisov D., Gao F., Williamson P., Simons F. J., and Tromp J. (2019). Robust surface-wave full-waveform inversion. *SEG Technical Program Expanded Abstracts*, 5005–5009.
- (31) Brossier R., Operto S., and Virieux J. (2009). Seismic imaging of complex onshore structures by two-dimensional elastic frequency-domain full-waveform inversion. *Geophysics* 74(6), WCC105–WCC118.
- (32) Brossier R., Operto S., and Virieux J. (2009). Two-dimensional seismic imaging of the Valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion. *SEG Technical Program Expanded Abstracts*, 2293–2297.
- (33) Brossier R., Operto S., and Virieux J. (2010). Which data residual norm for robust elastic frequency-domain full waveform inversion? *Geophysics* 75(3), R37–R46.
- (34) Broyden C. G. (1970). The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA Journal of Applied Mathematics* 6(1), 76–90.
- (35) Bunks C., Saleck F. M., Zaleski S., and Chavent G. (1995). Multiscale seismic waveform inversion. *Geophysics* 60(5), 1457–1473.
- (36) Byrd R. H., Lu P., Nocedal J., and Zhu C. (1995). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing* 16(5), 1190–1208.
- (37) Cary P., Li X., Popov G., and Zhang C. (2010). Shear-wave splitting in compliant rocks. *The Leading Edge* 29(10), 1278–1285.
- (38) Carcione J. M., *Wave Fields in Real Media: Wave Propagation in Anisotropic, Anelastic, Porous and Electromagnetic Media*, Third Edition; Handbook of geophysical exploration: Seismic exploration, Vol. 31; Elsevier: 2014.
- (39) Castagna J. P., Batzle M. L., and Eastwood R. L. (1985). Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. *Geophysics* 50(4), 571–581.
- (40) Chapman C. H., Hobro J. W. D., and Robertsson J. O. A. (2014). Correcting an acoustic wavfield for elastic effects. *Geophysical Journal International* 197(2), 1196–1214.
- (41) Chelminsky S., Watson L. M., and Ronen S. (2019). Research Note: Low-frequency pneumatic seismic sources. *Geophysical Prospecting* 67(6), 1547–1556.
- (42) Choi Y., and Shin C. (2007). Frequency domain elastic full waveform inversion using the new pseudo-Hessian matrix: elastic Marmousi-2 synthetic test. *SEG Technical Program Expanded Abstracts*, 1908–1912.
- (43) Cho H., Min D.-J., Cha Y. H., Shin C., and Ko S. (2008). A comparative study of cascaded frequency-selection strategies for 2D frequency-domain acoustic waveform inversion. *SEG Technical Program Expanded Abstracts*, 1930–1934.
- (44) Cholach P. Y., and Schmitt D. R. (1990). Seismic anisotropy of shales. *Recorder* 28(7).

- (45) Courant R., Friedrichs K., and Lewy H. (1967). On the partial difference equations of mathematical physics. *IBM journal of Research and Development* 11(2), 215–234.
- (46) Crampin S., Evans R., Üçer B., Doyle M., Davis J. P., Yegorkina G. V., and Miller A. (1990). Observations of dilatancy-induced polarization anomalies and earthquake prediction. *Nature* 286(5776), 874–877.
- (47) Crase E., Pica A., Noble M., McDonald J., and Tarantola A. (1990). Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics* 55(5), 527–538.
- (48) Dablain M. A. (1986). The application of high-order differencing to the scalar wave equation. *Geophysics* 51(1), 54–66.
- (49) de Hoop A. T. (1966). An elastodynamic reciprocity theorem for linear, viscoelastic media. *Applied Scientific Research* 16, 39–45.
- (50) Díaz E., and Guitton A. (2011). Fast full waveform inversion with random shot decimation. *SEG Technical Program Expanded Abstracts* 16, 2804–2808.
- (51) Dillen M. W. P. Time-lapse seismic monitoring of subsurface stress dynamics, Ph.D. Thesis, Delft University of Technology, 2000.
- (52) Eisner E., and Clayton R. W. (2001). A reciprocity method for multiple-source simulations. *Bulletin of Seismological Society of America* 91(3), 553–560.
- (53) Ernst F. E., Herman G. C., and Ditzel A. (2002). Removal of scattered guided waves from seismic data. *Geophysics* 67(4), 1240–1248.
- (54) Ernst F. E., and Herman G. C. (1998). Removal of guided waves from seismic data in laterally varying media. *Wave Motion* 28(2), 173–189.
- (55) Fenati D., and Rocca F. (1984). Seismic reciprocity field tests from Italian Peninsula. *Geophysics* 49(10), 1690–1700.
- (56) Ferreira A. M. G., Faccenda M., Sturgeon W., Chang S.-J., and Schardong L. (2019). Ubiquitous lower-mantle anisotropy beneath subduction zones. *Nature Geoscience* 12(4), 301–306.
- (57) Fichtner A., *Full Seismic Waveform Modelling and Inversion*; Springer-Verlag Berlin Heidelberg: 2011.
- (58) Fletcher R., and Reeves C. M. (1964). Function minimization by conjugate gradients. *The Computer Journal* 7(2), 149–154.
- (59) Fletcher R. (1970). A new approach to variable metric algorithms. *The Computer Journal* 13(3), 317–322.
- (60) Fliedner M. M., and White R. S. (2001). Seismic structure of basalt flows from surface seismic data, borehole measurements, and synthetic seismogram modeling. *Geophysics* 66(6), 1925–1936.
- (61) Fokkema J. T., and van den Berg P. M., *Seismic applications of acoustic reciprocity*; Elsevier: 1993.
- (62) Gaiser J., Moldoveanu N., Macbeth C., Michelena R., and Spitz S. (2001). Multi-component technology: the players, problems, applications, and trends: Summary of the workshop sessions. *The Leading Edge* 20(9), 974–977.

- (63) Gangi A. F. (2000). Constitutive equations and reciprocity. *Geophysical Journal International* 143(2), 311–318.
- (64) Gangi A. F. (1970). A Derivation of the Seismic Representation Theorem Using Seismic Reciprocity. *Journal of Geophysical Research* 75(11), 2088–2095.
- (65) Gardner G. H. F., Gardner L. W., and Gregory A. R. (1974). Formation Velocity and Density-The Diagnostic Basics for Stratigraphic Traps. *Geophysics* 39(6), 770–780.
- (66) Gholami Y., Brossier R., Operto S., and Ribodetti A. (2013). Which parameterization for acoustic vertical transverse isotropic full waveform inversion? - Part 1: sensitivity and trade-off analysis. *Geophysics* 78(2), R81–R105.
- (67) Goldfarb D. (1970). A Family of Variable Metric Updates Derived by Variational Means. *Mathematics of Computation* 24(109), 23–26.
- (68) Granli J. R., Arntsen B., Sollid A., and Hilde E. (1999). Imaging through gas-filled sediments using marine shear-wave data. *Geophysics* 64(3), 668–677.
- (69) Grimbergen J. L. T., W. (1996). Reciprocity theorems for one-way wavefields. *Geophysical Journal International* 127(1), 169–177.
- (70) Gumble J. E., and Gaiser J. E. (2006). Characterization of layered anisotropic media from prestack PS-wave-reflection data. *Geophysics* 71(5), D171–D182.
- (71) Guo Y., Huang J., Cui C., Li Z., Fu L., and Li Q. (2019). Multi-source multi-scale source-independent full waveform inversion. *Journal of Geophysics and Engineering* 16(3), 479–492.
- (72) Guo P., Visser G., and Saygin E. (2020). Bayesian trans-dimensional full waveform inversion: synthetic and field data application. *Geophysical Journal International* 222(1), 610–627.
- (73) Hale D. (2009). Structure-oriented smoothing and semblance. *CWP* 635.
- (74) Hamilton E. L. (1976). Shear-wave velocity versus depth in marine sediments: a review. *Geophysics* 41(5), 985–996.
- (75) Hamilton E. L. (1978). Sound velocity-density relations in sea-floor sediments and rocks. *The Journal of the Acoustical Society of America* 63(2), 366–377.
- (76) Han F.X., Sun J.-G., and Wang K. (2012). The influence of sea water velocity variation on seismic traveltimes, ray paths, and amplitude. *Applied Geophysics* 9(3), 319–325.
- (77) Harker S. D., Richardson G., Sides L. E., and Cooper R. (2003). Alwyn North Triassic main gas-condensate; drilling deeper promotes production. *Petroleum Geoscience* 9(2), 133–143.
- (78) He W., and Plessix R.-E. (2016). Analysis of different parameterisations of waveform inversion of compressional body waves in an elastic transverse isotropic Earth with a vertical axis of symmetry. *Geophysical Prospecting* 65(4), 1004–1024.
- (79) He W., Plessix R.-E., and Singh S. (2018). Parametrization study of the land multiparameter VTI elastic waveform inversion. *Geophysical Journal International* 213(3), 1660–1672.
- (80) He W. Elastic full waveform inversion in vertically transversely isotropic media, Ph.D. Thesis, Institut de Physique du Globe de Paris, 2017.

- (81) Hestenes M. R., and Stiefel E. (1952). Methods of Conjugate Gradients for Solving Linear Systems. *Journal of Research of the National Bureau of Standards* 49(6), 409–436.
- (82) Hicks E., Hoeber H., Houbiers M., Pannetier Lescoffit S., Ratcliffe A., and Vinje V. (2016). Time-lapse full-waveform inversion as a reservoir-monitoring tool — A North Sea case study. *The Leading Edge* 35(10), 850–858.
- (83) Hicks G. J. (2002). Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions. *Geophysics* 67(1), 156–166.
- (84) Huot G., and Singh S. C (2018). Seismic Evidence for Fluid/Gas Beneath the Mentawai Fore-Arc Basin, Central Sumatra. *Geophysical Journal International* 123(2), 957–76.
- (85) Ishiyama W. B., and Mainprice D. (1998). An olivine fabric database: an overview of upper mantle fabrics and seismic anisotropy. *Tectonophysics* 293, 145–157.
- (86) Ishiyama T. (2013). Dispersive multi-modal mud-roll elimination using feedback-loop approach. *SEG Technical Program Expanded Abstracts*, 4417–4421.
- (87) Johnson B. D., and van Klinken G. (1979). Some equivalent bodies and ambiguity in magnetic and gravity interpretation. *Exploration Geophysics* 10(1), 109–110.
- (88) Kamath N., Tsvankin I., and Díaz E. (2017). Elastic full-waveform inversion for VTI media: A synthetic parameterization study. *Geophysics* 82(5), C163–C174.
- (89) Karakostas F., Lognonné P., Larmat C., and Schmerr N. (2019). A Martian Impact Full Rayleigh Waveform Inversion Technique for 1D Identification of Crustal Structure. *50th Lunar and Planetary Science Conference 2019*.
- (90) Katou M., Abe S., Saito H., and Sato H. (2017). Reciprocal data acquisition and subsequent waveform matching for integrated onshore-offshore seismic profiling. *Geophysical Journal International* 212(1), 509–521.
- (91) Kayum S. N. (2020). GeoDRIVE – a high performance computing flexible platform for seismic applications. *First Break* 38(2), 97–100.
- (92) Kelly K. R., Ward R. W., Treitel S., and Alford R. M. (1976). Synthetic seismograms: a finite-difference approach. *Geophysics* 41(1), 2–27.
- (93) Kendall J.-M., Fisher Q. J., Crump S. C., Maddock J., Carter A., Hall S. A., Wookey J., Valcke S. L. A., Casey M., Lloyd G., and Ismail W. B. (2007). Seismic anisotropy as an indicator of reservoir quality in siliciclastic rocks. *Geological Society* 292, 123–136.
- (94) Kennett B. L. N., Sambridge M. S., and Williamson P. R. (1988). Subspace methods for large inverse problems with multiple parameter classes. *Geophysical Journal International* 94(2), 237–247.
- (95) Kennett B. L. N., *Seismic wave propagation in stratified media*; Cambridge University Press: 1983.
- (96) Kim Y., Cho H., Min D.-J., and Shin C. (2011). Comparison of Frequency-Selection Strategies for 2D Frequency-Domain Acoustic Waveform Inversion. *Pure and Applied Geophysics* 168(10), 1715–1727.
- (97) Klein G., Bohlen T., Theilen F., Kugler S., and Forbriger T. (2005). Acquisition and Inversion of Dispersive Seismic Waves in Shallow Marine Environments. *Marine Geophysical Research* 26(2-4), 287–315.

- (98) Kmenbach M. (1994). Multicomponent source equalization. *SEG Technical Program Expanded Abstracts*, 1449–1452.
- (99) Knopoff L., and Gangi A. F. (1959). Seismic reciprocity. *Geophysics* 24(4), 681–691.
- (100) Knopoff L. (1999). BENO GUTENBERG 1889-1960 - A Biographical Memoir by LEON KNOPOFF.
- (101) Korenaga J., Holbrook W. S., Kent G. M., Kelemen P. B., Detrick R. S., Larsen H.-C., Hopper J. R., and Dahl-Jensen T. (2000). Crustal structure of the Southeast Greenland margin from joint refraction and reflection seismic tomography. *Journal of Geophysical Research* 105(B9), 21, 591–21, 614.
- (102) Kotelnikov V. A. (1933). On the carrying capacity of the ether and wire in telecommunications. *Material for the First All-Union Conference on Questions of Communication*.
- (103) Kugler S., Bohlen T., Bussat A., and Klein G. (2005). Variability of Scholte-wave Dispersion in Shallow-water Marine Sediments. *Journal of Environmental and Engineering Geophysics* 10(2), 203–218.
- (104) Kuzuoglu M., and Mittra R. (1996). Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers. *IEEE Microwave and Guided Wave Letters* 6(12), 447–449.
- (105) Landrø M., and Amundsen L. (2010). Marine Seismic Sources Part I. *GEOExPro* 7(1), 203–218.
- (106) Landrø M., Amundsen L., and Barker D. (2011). High-frequency signals from air-gun arrays. *Geophysics* 76(4), Q17–Q25.
- (107) Landrø M., and Amundsen L. (2014). Is it optimal to tow air guns shallow to enhance low frequencies? *Geophysics* 79(3), A13–A18.
- (108) Le Meur D., Benjamin L., Twigger L., Garceran K., Delmas L., and Poulain G. (2010). Adaptive attenuation of surface-wave noise. *First Break* 28(9), 83–88.
- (109) Levin V., and Park J. (1998). P-SH Conversions in Layered Media with Hexagonally Symmetric Anisotropy: A CookBook. *Pure and Applied Geophysics* 152, 669–697.
- (110) Levander A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics* 53(11), 1425–1436.
- (111) Li X.-Y., and Zhang Y.-G. (2011). Seismic reservoir characterization: How can multicomponent data help? *Journal of Geophysics and Engineering* 8(2), 123–141.
- (112) Li Y. E., and Demanet L. (2016). Full-waveform inversion With extrapolated low-frequency data. *Geophysics* 81(6), R339–R348.
- (113) Louboutin M., Guasch L., and Herrmann F. J. (2017). Data Normalization Strategies for Full-waveform Inversion. *79th EAGE Conference & Exhibition*.
- (114) Madariaga P. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America* 66(3), 639–666.
- (115) Martin G. S., Wiley R., and Marfurt K. J. (2006). Marmousi2: An elastic upgrade for Marmousi. *The Leading Edge* 25(2), 113–124.

- (116) Martin R. D., Komatitsch D., and Ezziani A. (2008). An unsplit convolutional perfectly matched layer improved at grazing incidence for seismic wave propagation in poroelastic media. *Geophysics* 73(4), T51–T61.
- (117) Marelli S., Maurer H., and Manukyan E. (2012). Validity of the acoustic approximation in full-waveform seismic crosshole tomography. *Geophysics* 77(5), R129–R139.
- (118) Marjanović M., Plessix R.-E., Stopin A., and Singh S. C. (2019). Elastic versus acoustic 3-D Full Waveform Inversion at the East Pacific Rise 9°50N. *Geophysical Journal International* 216(3), 1497–1506.
- (119) Massoni I., Boelle J.-L., Brossier R., and Virieux J. (2016). Layer stripping FWI for surface waves. *SEG Technical Program Expanded Abstracts*, 1369–1373.
- (120) Mattesini M., Belonoshko A. B., Buforn E., Ramírez M., Simak S. I., Udías A., Mao H.-K., and Ahuja R. (2010). Hemispherical anisotropic patterns of the Earth’s inner core. *SEG Technical Program Expanded Abstracts* 107(21), 9507–9512.
- (121) Mazzotti A., Bienati N., Stucchi E., Tognarelli A., Aleardi M., and Sajeve A. (2016). Two-grid genetic algorithm full-waveform inversion. *The Leading Edge* 35(12), 1068–1075.
- (122) Métivier L., Bretaudeau F., Brossier R., Operto S., and Virieux J. (2014). Full waveform inversion and the truncated Newton method: quantitative imaging of complex subsurface structures. *Geophysical Prospecting* 62(6), 1353–1375.
- (123) Métivier L., Brossier R., Mérigot Q., Oudet E., and Virieux J. (2016). Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. *Geophysical Journal International* 205(1), 345–377.
- (124) Miller S. L. M., and Stewart R.R. (1990). Effects of lithology, porosity and shaliness on P- and S-wave velocities from sonic logs. *Canadian Journal of Exploration Geophysics* 26(1&2), 94–103.
- (125) Mittet R., and Hokstad K. (1995). Transforming walk-away VSP data into reverse VSP data. *Geophysics* 60(4), 968–977.
- (126) Mitchell N. C. (1998). Sediment accumulation rates from Deep Tow profiler records and DSDP Leg 70 cores over the Galapagos spreading centre. *Geological Evolution of Ocean Basins: Results from the Ocean Drilling Program*, Cramp, A., Macleod, C. J., Lee, S. V. and Jones, E. J. W., Geological Society of London, 199–209.
- (127) Moczo P., Kristek J., Vavryčuk V., Archuleta R. J., and Halada L. (2002). 3D heterogeneous staggered-grid finite-difference modeling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities. *Bulletin of the Seismological Society of America* 92(8), 3042–3066.
- (128) Moghaddam P. P., Libak A., Keers H., and Mjelde R. (2012). Efficient and accurate modeling of ocean bottom seismometer data using reciprocity. *Geophysics* 77(6), T211–T220.
- (129) Moghaddam P. P., Keers H., Herrmann F. J., and Mulder W. A. (2013). A new optimization approach for source-encoding full-waveform inversion. *Geophysics* 78(3), R125–R132.

- (130) Mondol N. H., Avseth P., Fawad M., and Smith T. (2010). Vs Prediction in Unconsolidated Sands - Physical and Geological Controls on Shear Wave Velocities. *72nd EAGE Conference & Exhibition incorporating SPE EUROPEC 2010*.
- (131) Morales J. L., and Nocedal J. (2011). Remark on “algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization”. *ACM Transactions on Mathematical Software* 38(1).
- (132) Mora P., and Wu Z. (2018). Elastic versus acoustic inversion for marine surveys. *Geophysical Journal International* 214(1), 596–622.
- (133) Mulder W. A. (2005). Rigorous redatuming. *Geophysical Journal International* 161(2), 401–415.
- (134) Naeini E., and Hale D. (2015). Image- and horizon-guided interpolation. *Geophysics* 80(3), V47–V56.
- (135) Nicolas B., Mars J., Lacoume J.-L., and Fattaccioli D. (2011). Are ultra low frequency waves suitable for detection? *Conference paper: OCEANS '02 MTS/IEEE* 2, 1109–1113.
- (136) Nocedal J., and Wright S. J., *Numerical optimization*, Second Edition; Springer: 2006.
- (137) Nowack R., and Chen W.-P. (1999). Source-Receiver Reciprocity and Empirical Green’s Functions from Chemical Blasts. *Bulletin of the Seismological Society of America* 89(2), 538–543.
- (138) Nyquist H. (1928). Certain topics in telegraph transmission theory. *Trans. AIEE* 47(2), 614–644.
- (139) Ogarague D. O. (2008). An F-K Procedure of Eliminating Guided Wave Energy from Shallow Marine Seismic Data. *The Pacific Journal of Science and Technology* 9(2), 502–508.
- (140) Oh J.-U., Kalita M., and Alkhalifah T. (2018). 3D elastic full-waveform inversion using P-wave excitation amplitude: Application to ocean bottom cable field data. *Geophysics* 83(2), R129–R140.
- (141) Operto S., Virieux J., Dessa J.-X., and Pascal G. (2006). Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography: Application to the eastern Nankai trough. *Journal of Geophysical Research* 11, B09306.
- (142) Operto S., Miniussi A., Brossier R., Combe L., Métivier L., Monteiller V., Ribodetti A., and Virieux J. (2015). Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation. *Geophysical Journal International* 202(2), 1362–1391.
- (143) Pica A., Diet J. P., and Tarantola A. (1990). Nonlinear inversion of seismic reflection data in a laterally invariant medium. *Geophysics* 55(3), 284–292.
- (144) 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.

- (145) Polak E., and Ribière G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue française d'informatique et de recherche opérationnelle. Série rouge* 3(16), 35–43.
- (146) Prioux V., Brossier R., Gholami Y., Operto S., Virieux J., Barkved O. I., and Kommedal J. H. (2011). On the footprint of anisotropy on isotropic full waveform inversion: The Valhall Case Study. *Geophysical Journal International* 187(3), 1495–1515.
- (147) Prioux V., Brossier R., Operto S., Virieux J., Barkved O. I., and Kommedal J. H. (2012). Two-dimensional anisotropic visco-elastic full waveform inversion of wide-aperture 4C OBC data from the Valhall field. *74th EAGE Conference & Exhibition incorporating SPE EUROPEC 2012*.
- (148) Raknes E. B., Arntsen B., and Weibull W. (2015). Three-dimensional elastic full waveform inversion using seismic data from the Sleipner area. *Geophysical Journal International* 202(3), 1877–1894.
- (149) Ratcliffe A., Win C., Vinje V., Conroy G., Warner M., Umpleby A., Stekl I., Nangoo T., and Bertrand A. (2011). Full Waveform Inversion: a North Sea OBC case study. *SEG Technical Program Expanded Abstracts*, 2384–2388.
- (150) Rattey R. P., and Hayward A. B. (1993). Sequence stratigraphy of a failed rift system: the Middle Jurassic to Early Cretaceous basin evolution of the Central and Northern North Sea. *Geological Society, London, Petroleum Geology Conference series* 4, 215–249.
- (151) Rayleigh Lord (Strutt J. W.), *Some general theorems related to vibrations*, 1873; Vol. s1-4(1), pp 357–368.
- (152) Rayleigh Lord (Strutt J. W.), *The theory of sound*; Macmillan and Company: 1878.
- (153) Razavy M., and Lenoach B. (1986). Reciprocity principle and the approximate solution of the wave equation. *Bulletin of the Seismological Society of America* 76(6), 1776–1789.
- (154) Reilly J. M. (1993). Integration of well and seismic data for 3D velocity model building. *First Break* 11(6), 247–260.
- (155) Ren Z., Li Z., and Gu B. (2019). Elastic reflection waveform inversion based on the decomposition of sensitivity kernels. *Geophysics* 84(2), R235–R250.
- (156) Rickers F., Fichtner A., and Trampert J. (2013). The Iceland-Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: Evidence from full-waveform inversion. *Earth and Planetary Science Letters* 367(1), 39–51.
- (157) Roberts M. A. Full waveform inversion of walk-away VSP data, Ph.D. Thesis, Institut de Physique du Globe de Paris, 2005.
- (158) Robertsson J. O. A. (1996). A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography. *Geophysics* 61(6), 1921–1934.
- (159) Roden J. A., and Gedney S. D. (2000). Convolution PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media. *Microwave and Optical Technology Letters* 27(5), 334–339.

- (160) Saengar E. H., and Bohlen T. (2004). Finite-difference modeling of viscoelastic and anisotropic wave propagation using the rotated staggered grid. *Geophysics* 69(2), 583–591.
- (161) Sajeve A., Aleardi M., and Mazzotti A. (2017). Genetic algorithm full-waveform inversion: uncertainty estimation and validation of the results. *Bollettino di Geofisica Teorica ed Applicata* 58(4), 395–414.
- (162) Sayers C. M., *Geophysics under stress: Geomechanical applications of seismic and borehole acoustic waves*; Society of Exploration Geophysicists, the European Association of Geoscientists, and Engineers: 2010.
- (163) Schön J., *Physical Properties of Rocks*, Second Edition; Elsevier: 2015; Vol. 65.
- (164) Sears T. J., Singh S. C., and Barton P. J. (2008). Elastic full waveform inversion of multi-component OBC seismic data. *Geophysical Prospecting* 56(6), 843–862.
- (165) 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.
- (166) Sears T. J. Elastic full waveform inversion of multi-component ocean-bottom cable seismic data, Ph.D. Thesis, University of Cambridge, 2007.
- (167) Seifoullaev R. K., Stoffa P. L., and Sen M. K. (2005). Use of reciprocity in a double plane wave Kirchoff depth migration. *SEG Technical Program Expanded Abstracts*, 1636–1640.
- (168) Shannon C. E. (1949). Communication in the presence of noise. *Proceedings of the Institute of Radio Engineers* 37(1), 10–21.
- (169) Shanno D. F. (1970). Conditioning of Quasi-Newton Methods for Function Minimization. *Mathematics of Computation* 24(111), 647–656.
- (170) Sheen D.-H., Tuncay K., Baag C.-E., and Ortoleva P. J. (2004). Efficient finite difference calculation of partial derivative seismic wavefield using reciprocity and convolution. *SEG Technical Program Expanded Abstracts*, 1850–1853.
- (171) Sheen D.-H., Tuncay K., Baag C.-E., and Ortoleva P. J. (2006). Time domain Gauss-Newton seismic waveform inversion in elastic media. *Geophysical Journal International* 167(3), 1373–1384.
- (172) Shipp R. M., and Singh S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International* 151(2), 325–344.
- (173) Shi Y., Zhao W., and Cao H. (2007). Nonlinear process control of wave-equation inversion and its application in the detection of gas. *Geophysics* 72(1), R9–R18.
- (174) Simmons J. L. Jr. (2009). Converted-wave splitting estimation and compensation. *Geophysics* 74(1), D37–D48.
- (175) Sirgue L., and Pratt R. G. (2004). Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics* 69(1), 231–248.
- (176) Sirtori A., Mantovani M., Epifani A., and Miotti F. (2019). Integrated Simultaneous Joint and Full-Waveform Inversion Workflow for Multiphysics Near-Surface Modeling. *81st EAGE Conference & Exhibition 2019*.

- (177) Skeels D. C. (1947). AMBIGUITY IN GRAVITY INTERPRETATION. *Geophysics* 12(1), 43–56.
- (178) Solano C. A. P., Stopin A., and Plessix R.-E. (2013). Synthetic Study of Elastic Effects on Acoustic Full Waveform Inversion. *75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*.
- (179) Song Z.-M., and Williamson P. R. (1995). Frequency-domain acoustic-wave modeling and inversion of crosshole data: Part 1 - 2.5-D modeling method. *Geophysics* 60(3), 784–795.
- (180) Song Z.-M., Williamson P. R., and Pratt R. G. (1995). Frequency-domain acoustic-wave modeling and inversion of crosshole data: Part II-Inversion method, synthetic experiments and real-data results. *Geophysics* 60(3), 796–809.
- (181) Song X. (1997). Anisotropy of the Earth's inner core. *Reviews of Geophysics* 35(3), 297–313.
- (182) Stewart R. R., Gaiser J. E., Brown R. J., and Lawton D. C. (2003). Converted-wave seismic exploration: Applications. *Geophysics* 68(1), 40–57.
- (183) Stolt R. H., and Weglein A. B., *Seismic Imaging and Inversion: Application of Linear Inverse Theory*; Cambridge University Press: 2012.
- (184) Strobbia C. L., Vermeer P., Laake A., Glushchenko A., and Re S. (2010). Surface waves: processing, inversion and removal. *First Break* 28(8), 85–91.
- (185) Szymańska-Małysa Ż., and Dubiel P. (2019). 3C seismic data processing and interpretation: a case study from Carpathian Foredeep basin. *Acta Geophysica* 67, 2031–2047.
- (186) Tan T. H. (1987). Reciprocity theorem applied to the geophone-ground coupling problem. *Geophysics* 52(12), 1715–1717.
- (187) Tarantola A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics* 49(8), 1259–1266.
- (188) Tarantola A. (1986). A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics* 51(10), 1979–2018.
- (189) Tarantola A. (1988). Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure and Applied Geophysics* 128(1-2), 365–399.
- (190) Thomas C., and Kendall J.-M. (2002). The lowermost mantle beneath northern Asia—II. Evidence for lower-mantle anisotropy. *Geophysical Journal International* 151(1), 296–308.
- (191) Thomsen L., Tsvankin I., and Mueller M. C. (1999). Coarse-layer stripping of vertically variable azimuthal anisotropy from shear-wave data. *Geophysics* 64(4), 1126–1138.
- (192) Thomsen L. (1986). Weak elastic anisotropy. *Geophysics* 51(10), 1954–1966.
- (193) Tomar G., Shapiro N. M., Mordret A., Singh S., and Montagner J. P. (2015). Noise-based Seismic Tomography at the Valhall Oil Field with Using Scholte and Love Waves. *Conference Proceedings: 77th EAGE Conference & Exhibition 2015*.

- (194) Tromp J., Tape C., and Liu Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International* 160(1), 195–216.
- (195) Tromp J. (2020). Seismic wavefield imaging of Earth’s interior across scales. *Nature Reviews Earth & Environment* 1, 40–53.
- (196) Tromp J. (1993). Support for anisotropy of the Earth’s inner core from free oscillations. *Nature* 366, 678–681.
- (197) Tsuji T., and Iturrino G. J. (2008). Velocity-porosity relationships in oceanic basalt from eastern flank of the Juan de Fuca Ridge: The effect of crack closure on seismic velocity. *Exploration Geophysics* 39(1), 41–51.
- (198) Ung K., Bona A., and Madadi M. (2015). Reciprocity principle in finite difference modelling of waves in elastic media. *ASEG Extended Abstracts*.
- (199) van Vossen R., Robertsson J. O. A., and Chapman C. H. (2002). Finite-difference modeling of wave propagation in fluid-solid configuration. *Geophysics* 67(2), 618–624.
- (200) van Vossen R., Trampert J., Curtis A., and Laake A. (2005). Source and or receiver amplitude equalization using reciprocity - Application to land seismic data. *SEG Technical Program Expanded Abstracts*, 2177–2180.
- (201) van Manen D.-J., Curtis A., and Robertson J. O. (2006). Interferometric modeling of wave propagation in inhomogeneous elastic media using time reversal and reciprocity. *Geophysics* 71(4), SI47–SI60.
- (202) van Vossen R., Curtis A., Laake A., and Trampert J. (2006). Surface-consistent deconvolution using reciprocity and waveform inversion. *Geophysics* 71(2), V19–V30.
- (203) van Vossen R., Curtis A., and Trampert J. (2006). Surface-consistent amplitude corrections for single or multicomponent sources and receivers using reciprocity and waveform inversion. *Geophysical Journal International* 165(1), 311–322.
- (204) van Borselen R. G., Fokkema J., and van den Berg P. (2013). Wavefield decomposition based on acoustic reciprocity: Theory and applications to marine acquisition. *Geophysics* 78(2), WA41–WA54.
- (205) Vassiliou A., Stromberg R., and Woog L. (2017). Elastic waveform inversion for rock property estimation and for imaging: A Midland Basin case study. *SEG Technical Program Expanded Abstracts*, 564–568.
- (206) Vermeer G. J. O. (1991). Symmetric sampling. *The Leading Edge* 10(11), 21–27.
- (207) Vermeer G. J. O. (1998). 3-D symmetric sampling. *Geophysics* 63(5), 1629–1647.
- (208) Vidal C. A., van der Neut J., and Wapenaar K. (2019). Reciprocity-based passive monitoring with individual sources. *Geophysics* 84(2), KS29–KS38.
- (209) Vigh D., Starr E. W., and Elapavuluri P. (2009). Acoustic Waveform Inversion vs. Elastic data. *SEG Technical Program Expanded Abstracts*.
- (210) Vigh D., Jiao K., Watts D., and Sun D. (2014). Elastic full-waveform inversion application using multicomponent measurements of seismic data collection. *Geophysics* 79(2), R62–R76.
- (211) Vinh P. C. (2013). Scholte-wave velocity formulae. *Wave Motion* 50(2), 180–190.

- (212) Virieux J., and Operto S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74(6), WCC127–WCC152.
- (213) Virieux J., Etienne V., Cruz-Atienza V., Brossier R., Chaljub E., Coutant O., Garambois S., Mercerat D., Prioux V., Operto S., Ribodetti A., and Tago J. In 2012; Chapter 13, pp 253–304.
- (214) Virieux J. (1984). SH wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics* 49(11), 1933–1942.
- (215) Virieux J. (1986). P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics* 51(4), 889–901.
- (216) Visser G., Guo P., and Saygin E. (2019). Bayesian transdimensional seismic full-waveform inversion with a dipping layer parameterization. *Geophysics* 84(6), R845–R858.
- (217) Wang N., Montagner J.-P., Fichtner A., and Capdeville Y. (2013). Intrinsic versus extrinsic seismic anisotropy: The radial anisotropy in reference Earth models. *Geophysical Research Letters* 40(16), 4284–4288.
- (218) Wang H., Singh S. C., Audebert F., and Calandra H. (2015). Inversion of seismic refraction and reflection data for building long-wavelength velocity models. *Geophysics* 80(2), R81–R93.
- (219) Wang Z., Hu H., and Yang Y. (2015). Reciprocity relations for the elastodynamic fields generated by multipole sources in a fluid-solid configuration. *Geophysical Journal International* 203(2), 883–892.
- (220) Wang Y., Miller R. D., Peterie S. L., Sloan S. D., Moran M. L., Cudney H. H., Smith J. A., Borisov D., Modrak R., and Tromp J. (2019). Case History: Tunnel detection at Yuma Proving Ground, Arizona, USA — Part 1: 2D full-waveform inversion experiment. *Geophysics* 84(1), B95–B105.
- (221) Wang Z., Singh S. C., and Noble M. (2020). True-amplitude versus trace-normalized full waveform inversion. *Geophysical Journal International* 220(2), 1421–1435.
- (222) Wang Y., and Schuster G. T. (1996). Finite-difference variable grid scheme for acoustic and elastic wave equation modeling. *SEG Technical Program Expanded Abstracts*, 674–677.
- (223) Wang Z. (2002). Seismic anisotropy in sedimentary rocks, part 2: Laboratory data. *Geophysics* 67(5), 1423–1440.
- (224) Wapenaar K., Slob E., and Fokkema J. T. (2004). Reciprocity and power balance for piecewise continuous media with imperfect interfaces. *Journal of Geophysical Research: Atmospheres* 109, B10301.
- (225) Wapenaar K., Slob E., Snieder R., and Curtis A. (2010). Tutorial on seismic interferometry: Part 2 - Underlying theory and new advances. *Geophysics* 75(5), A211–A227.
- (226) Wapenaar C. P. A., Verschuur D. J., and Herrmann P. (1992). Amplitude pre-processing of single and multicomponent seismic data. *Geophysics* 57(9), 1178–1188.
- (227) Wapenaar K. (1998). Reciprocity properties of one-way propagators. *Geophysics* 63(5), 1795–1798.

- (228) Whale R., Bale R., Poplavskii K., Douglas K., Li X., and Slind C. (2009). Estimating and compensating for anisotropy observed in PS data for a heavy oil reservoir. *SEG Technical Program Expanded Abstracts*, 1212–1216.
- (229) Whittaker E. T. (1915). On the Functions Which are Represented by the Expansions of the Interpolation Theory. *Proc. Royal Soc.* 35, 181–194.
- (230) White J. E. (1960). Use of reciprocity theorem for computation of low-frequency radiation patterns. *Geophysics* 25(3), 613–624.
- (231) Wikel K., Kendall R., Bale R., Grossman J., and DeMeersman K. (2012). 4D-3C geomechanical study of in-situ bitumen recovery utilizing toe to heel air injection (THAI[®]), Alberta, Canada. *First Break* 30(2), 43–53.
- (232) Wu R.-S., Luo J., and Wu B. (2014). Seismic envelope inversion and modulation signal model. *Geophysics* 79(3), WA13–WA24.
- (233) Yuan S., Fuji N., Singh S., and Borisov D. (2017). Localized time-lapse elastic waveform inversion using wavefield injection and extrapolation: 2-D parametric studies. *Geophysical Journal International* 209(3), 1699–1717.
- (234) Zhang Q., Zhou H., Li Q., Chen H., and Wang J. (2017). Robust source-independent elastic full-waveform inversion in the time domain. *Geophysics* 81(2), R29–R44.
- (235) Zhao Z., Sen M. K., and Stoffa P. L. (2017). Reciprocity and double plane-wave migration. *Geophysics* 82(6), S453–S466.
- (236) Zhang Z., and Alkhalifah T. (2019). Local-crosscorrelation elastic full-waveform inversion. *Geophysics* 84(6), R897–R908.
- (237) Zhao Z., and Sen M. K. (2019). A gradient based MCMC method for FWI and uncertainty analysis. *SEG Technical Program Expanded Abstracts*, 1465–1469.
- (238) Zhu C., Byrd R. H., Lu P., and Nocedal J. (1997). Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)* 23(4), 550–560.

Inversion de forme d'onde élastique en milieux anisotropes

Méthodologie et Application à des données de nœuds de fond de mer de mer du Nord

par Aline ROBIN

- Résumé

L'inversion de forme d'onde (FWI) est apparue comme une formidable classe d'algorithmes pour obtenir des modèles quantitatifs et haute résolution des propriétés mécaniques du sous-sol. Cependant, la FWI souffre de coûts de calculs élevés et des limitations inhérentes à l'algorithme d'inversion utilisé. En ce qui concerne l'imagerie de la Terre en environnement marin à l'échelle locale, ces limitations ont conduit à la restriction initiale de la FWI à une représentation fluide de la Terre, où seule la vitesse des ondes P est mise à jour. L'augmentation des ressources de calcul a récemment permis d'aboutir à une stratégie d'inversion 2D tant des vitesses des ondes P que celles des ondes S, pour les environnements présentant de faibles vitesses S en fond de mer. La stratégie a été élaborée en approximant le sous-sol par un milieu isotrope et en utilisant des données de câbles de fond de mer (OBC). Cette thèse se concentre sur des dispositifs d'enregistrement plus contemporains et assouplit partiellement la limitation isotrope, afin de répondre aux besoins des applications modernes. Les OBC sont en effet remplacés par des sismomètres/nœuds de fond de mer (OBS/OBN). Sans l'utilisation du principe de réciprocité, où les rôles des sources et des récepteurs sont interchangés, le coût de calcul accru de ces types d'acquisition serait prohibitif. Je fournis ici un guide complet sur la réciprocité pour la FWI 2D/3D en milieux élastiques anisotropes. Par ailleurs, la stratégie d'inversion étudiée a été développée dans des environnements isotropes, alors que la Terre est anisotrope. Après les adaptations nécessaires, j'évalue la performance de la stratégie pour l'inversion des vitesses verticales, lorsque les autres paramètres d'anisotropie sont légèrement erronés et ne sont pas mis à jour. Je me concentre sur un type d'anisotropie spécifique nommée « isotropie transverse verticale » (VTI). Je propose également une variante pour des applications réelles, qui, par exemple, ne prendraient pas en compte l'atténuation de la Terre. Je confirme finalement les conclusions synthétiques et théoriques avec un jeu de données réelles 2D OBN, acquis en mer du Nord.

- Mots clés

Imagerie sismique ; inversion de forme d'onde (FWI) ; environnement marin à faibles vitesses S ; élasticité ; isotropie transverse verticale (VTI) ; sismomètre/nœud de fond de mer (OBS/OBN) ; réciprocité ; paramétrisation ; coefficient de réflexion ; diagramme de réflexion.



Elastic full waveform inversion in anisotropic media

Methodology and Application to ocean bottom node data from the North Sea par
Aline ROBIN

- Abstract

Full Waveform Inversion (FWI) has emerged as a powerful class of algorithms to retrieve quantitative and high resolution models of mechanical properties of the subsurface. However, FWI suffers from high computing costs and from the intrinsic limitations of the embedded inversion algorithm. Focusing on Earth local imaging in marine environments, these limitations led to the initial restriction of FWI to a fluid characterization of the Earth, where only the P-wave velocity was updated. Benefiting from increased computing resources, recent efforts produced a successful and much needed 2D inversion strategy for both P- and S-wave velocities in soft seabed environments. The strategy was developed assuming an isotropic subsurface and using ocean-bottom cable data (OBC). This thesis focuses on more contemporary recording devices, and partially relaxes the isotropic limitation. Over the years, ocean-bottom cables are indeed replaced by ocean-bottom seismometers/nodes (OBS/OBN). Without the use of the reciprocity principle, where the roles of the airgun sources and the OBS/OBN receivers are interchanged, the increased computing cost would be prohibitive. I provide here an extensive guide of such a use of reciprocity in elastic anisotropic 2D/3D FWI. Furthermore, the available soft seabed strategy was developed in isotropic environments, while the real Earth is anisotropic. After the necessary adaptations, I assess the performance of the strategy for vertical velocity retrieval when slightly inaccurate remaining anisotropy parameters are used, those parameters not being updated. I focus on a specific kind of anisotropy called “vertical transverse isotropy”. I also propose a variant strategy for real applications that would not take attenuation into account for instance. Finally, I support my synthetic and theoretical findings with a real 2D OBN dataset from the North Sea.

- Keywords

Seismic imaging; full waveform inversion (FWI); soft seabed environments; elasticity; vertical transverse isotropy (VTI); ocean bottom seismometer/node (OBS/OBN); reciprocity; parameterization; reflection coefficient; reflection pattern.

