Imperial College London

Department of Earth Science and Engineering

Acoustic full-waveform inversion in geophysical and medical imaging

Òscar Calderón Agudo

April 2018

Supervised by Joanna Morgan and Mike Warner

Submitted in part fulfilment of the requirements for the degree of

Doctor of Philosophy of Imperial College London

and the Diploma of Imperial College London

Declaration of originality

I herewith certify that all material in this dissertation which is not my own work has been properly acknowledged.

Òscar Calderón Agudo

Copyright declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

Imaging methods are employed in multiple scientific disciplines to obtain properties of targets that are otherwise invisible to the human eye, such as for imaging the earth's interior or the human body.

In geophysics, the full-waveform inversion (FWI) method is often used to recover high-quality high-fidelity models of the subsurface by recording seismic data in the field, and modelling wave propagation with a wave equation. But the visco-elastic nature of the subsurface is typically ignored, and the acoustic approximation of the wave equation is utilised to reduce the compute costs. This results in less well-resolved and inaccurate models, as accurate physics of wave propagation is not considered.

On the contrary, the application of FWI to medical imaging is still in its infancy, especially in 3D. Instead, ray-based tomographic methods are used, which do not account for the physics of finite-frequency wave propagation, leading to poorer recovered models of physical properties.

Here, I first propose a method to address elastic and viscous effects in FWI of seismic data, which is based on the use of matching filters and modelling elastic or viscous wave propagation prior to the inversion, while using the acoustic wave equation during the inversion. This represents a considerable time reduction over full elastic or viscous inversions. I show that the proposed method leads to more accurate and detailed P-wave velocity subsurface models, both on synthetic and field datasets.

Then, I investigate the feasibility of 3D FWI for breast cancer diagnosis and analyse, for the first time, its implementation to obtain 3D images of the human brain with ultrasound through the skull. The results demonstrate FWI is now ready to be applied to these datasets, and could result in more accurate images and faster and safer acquisitions than with current methods.

Acknowledgements

First of all I want to thank my supervisors, Joanna Morgan and Mike Warner, for giving me the opportunity to take part in this research project. Without their knowledge, intuition, guidance, advice, support and our long discussions all this would not have been possible.

I also want to thank past and present members of the FULLWAVE research group for their valuable ideas and various contributions to this work. Without any particular order, I want to thank Kenneth Irabor, Akela Silverton, Adrian Umpleby, Lluís Guasch, Melissa Gray, Tatiana Kalinicheva, Gang Yao, Jack Ashley, Tim Burgess, Nuno Vieira da Silva, Henry Debens, Jiashun Yao, Tim Lin, Tenice Nangoo, Michele Paulatto and Kajetan Chrapkiewicz.

In particular, I want to specially thank my colleagues Nuno Vieira da Silva and Lluís Guasch for supervising part of my work but, more importantly, for giving me excellent academic and personal advice throughout my PhD.

Also, I want to thank my new collaborator, Peter Huthwaite, and my colleagues Carlos Cueto Mondéjar, Tom Robins, Javier Cudeiro and Mengxing Tang: embarking into this new adventure would have been much more difficult and less promising without you.

This work has also been possible thanks to the funding provided by the Earth Science and Engineering Department at Imperial College London. Thus, I want to thank everyone that has contributed to make this possible.

Without a doubt, I also want to thank my friends in Vilafranca (David P., David L., Pere, Lidia, Marta, Francesc, Ricard and many more), the numerous people I have

met and friends I have made in London (particularly Kostis, Aimilios, Marina and Maria), as well as my best international university friend (Cédéric) for making this experience happier and more enjoyable. Thank you all!

Finally, I want to thank my parents, grandparents and sister: *Moltes gràcies pel vostre suport incondicional, per la vostra ajuda i per saber entendre què necessitava en tot moment. Us estimo!*.

My (future) parents-in-law: *Gràcies pels vostres consells, l'ajuda que ens heu donat a la Mireia i a mi i pels moments que hem passat junts durant aquests anys!*.

And of course, you, Mireia: *Tot això no hauria estat possible sense tu, ja ho saps. Simplement, gràcies, ets perfecte! T'estimo!*

Thank you so much to all of you!

Contents

List of tables **xiv**

List of figures **xxxi**

List of algorithms **xxxii**

List of acronyms **xxxiii**

1. Introduction **1**

2. Methods **5**

2.1. The wave equation 6

2.1.1. Anisotropic elastic media 6

2.1.2. Isotropic elastic media 10

2.1.3. Isotropic acoustic media 13

2.1.4. Anisotropic TI acoustic media 14

2.1.5. Visco-elastic and visco-acoustic media 15

2.2. The forward problem 18

2.2.1. Finite differences (FD) 19

2.2.2. Grid dispersion and numerical stability 21

2.2.3. Boundary conditions, modelling domain and source injection . 22

2.3. The inverse problem: full-waveform inversion (FWI) 24

2.3.1. Evolution and key features of FWI 25

2.3.2. Theoretical framework 30

2.4. Travel-time tomography 35

2.5. Acoustic FWI in a visco-elastic earth . . . 37
2.5.1. Wiener filters . . . 37
2.5.2. Workflow and algorithm . . . 42
2.6. Modelling and inversion in medical imaging . . . 54
2.7. Original work . . . 57

3. Acoustic FWI in an elastic earth 59
3.1. Introduction . . . 60
3.2. Application to 2D horizontally layered models . . . 65
3.3. Application to a marine synthetic dataset . . . 76
3.4. Application to a land synthetic dataset . . . 89
3.5. Application to a 3D marine field dataset . . . 97
3.6. Summary . . . 111

4. Acoustic FWI in a visco-acoustic earth 115
4.1. Introduction . . . 116
4.2. Application to a 2D horizontally layered model . . . 119
4.3. Application to a marine synthetic dataset . . . 130
4.3.1. Impact of noise . . . 136
4.4. Application to a marine field dataset . . . 139
4.4.1. Estimating Q . . . 140
4.4.2. Results . . . 149
4.5. Summary . . . 158

5. 3D imaging of the breast with ultrasound 162
5.1. Introduction . . . 163
5.2. 3D FWI of a numerical model of the breast . . . 166
5.3. Applicability to a 2D *in vivo* dataset . . . 177
5.4. Summary . . . 187

6. Towards 3D imaging of the brain with ultrasound 191
6.1. Introduction . . . 192

6.2. 3D brain phantom 195
6.2.1. 2D inversion 197
6.2.2. 3D inversion 209
6.3. Estimation of a skull model 214
6.3.1. Method 1 214
6.3.2. Method 2 218
6.4. Summary 222

7. Conclusions and next steps **226**
7.1. Addressing elastic and viscous effects in acoustic FWI of seismic data 226
7.2. 3D FWI of ultrasound data: breast and brain imaging 229

Appendices **232**

A. Relation between absorption coefficients in medical ultrasound and geophysics **233**

B. Estimating a skull starting model: method 1 **236**

C. Permissions **244**

Bibliography **270**

List of Tables

2.1. Characteristic scales between geophysical and medical imaging problems based on the examples presented in the following chapters. 56

List of Figures

2.1. Models of a) vertical transversely isotropic (VTI) and b) horizontal transversely isotropic (HTI) anisotropy caused by thin horizontal and vertical layering, respectively. Modified after Tsvankin (2001). 9

2.2. The effect of a) a compressional force P (hydrostatic pressure) acting on an element of volume V changes the volume of an elastic element, whereas b) a shear force F acting on an area ΔA changes the shape of an elastic element. Modified after Alsadi (2016). 12

2.3. Variation of the dissipation factor, Q^{-1}, with the logarithm of the angular frequency for a) a Zener model and b) a generalised Zener model with five dissipation mechanisms. Modified after Carcione (2001). 17

2.4. Comparison between the a) true, b) starting, c) inverted with FWI and d) inverted with AWI P-wave velocity models for the Marmousi2 dataset. Modified after Warner and Guasch (2014). 27

2.5. Models of a) velocity and b) acoustic impedance obtained after 3D travel-time tomography and 3D FWI at $\sim 40\,Hz$, respectively, for a field dataset in the Western Black Sea. The model in a) is used as a starting model for FWI. Modified after Routh et al. (2017). 28

2.6. Example of data matching of two very different time series using Wiener filters. Panel a) shows the input signal $\boldsymbol{y}$ and the target signal $\boldsymbol{x}$, both of length $N = 100$, whereas panel b) shows the Wiener filters computed with the QR and Levinson's methods with a length of the filters of $P = 2N$. The resulting estimated signals for both methods are overlaid with the target signal in panel c), and the computation times for different lengths of the filters are shown in d). 43

2.7. Data matching example for the Marmousi2 model in a) for P-wave velocity and b) for S-wave velocity models, which are used to generate c) acoustic modelled data and d) elastic modelled data for a shot at the centre of the model, and their difference is shown in e). Wiener filters of a length of $P = 2N$ are then computed trace-by-trace to match the elastic to the acoustic modelled data, and are then convolved with the elastic data to mitigate elastic effects, which leads to the gather in f). The difference of the acoustic data and the data in f) is shown in g). . 44

2.8. Workflow used to mitigate elastic effects based on Wiener filters. The dotted area represents the conventional acoustic FWI workflow. 46

2.9. Workflow used to mitigate visco-acoustic effects based on Wiener filters. The dotted area represents the conventional acoustic FWI workflow. 48

3.1. Snapshots of the wavefield in a three-layered model generated from a source situated at the centre of the model, within the water layer, both in an acoustic (left) and an elastic (right) media at different times (top to bottom). 62

3.2. Vertical profiles for the true P-wave and S-wave velocities for the horizontally layered 2D models A (left) and B (right). 67

3.3. Vertical profiles of the recovered P-wave velocity models for model A after a) acoustic FWI of the acoustic and elastic datasets and b) the acoustic and the matched datasets after 1 and 2 iterations of the data-matching workflow. The true, starting and recovered P-wave velocity model after acoustic FWI of acoustic data are overlaid. 69

3.4. In the left column, representative shot gathers for a) the observed elastic data and c), e), g) and i) the matched data obtained at the centre of the first elastic model for model A, the latter obtained with the following filtering parameters: c) trace-by-trace and 320 ms long, e) multiple traces and 320 ms long, g) multiple traces and 35 ms long, and i) multiple traces and 1.58 s long. The difference of the gathers on the left with the corresponding shot gather for the true acoustic data is shown on the right column. 71

3.5. Vertical profiles of the recovered P-wave velocity models for model B, in which all inversions are acoustic. The acoustic and elastic datasets are inverted in a) and c), whereas the corresponding matched datasets are inverted in b) and d) (for a single iteration of either workflow 1 or 2). The inversions in a) and b) are performed with the same starting model, whereas those in c) and d) use another starting model less accurate at the sea-bottom. The conventional FWI functional is used throughout, but the amplitude-balanced functional is also used in c). The true and starting models and the acoustic FWI of acoustic data are overlaid on all plots. 73

3.6. Vertical slices of the true marine Marmousi2 a) P-wave and b) S-wave velocity models. 77

3.7. Source signature in the time domain (left) and its amplitude spectrum (right) used to generate data for the Marmousi2 model. 77

3.8. Representative shot gathers at the centre of the Marmousi2 model for the a) acoustic, b) elastic and c) the matched data using a smooth v_p/v_s. Shot gathers in d) and e) show the difference of the panels above with the acoustic data in a). 78

3.9. Vertical slices of the Marmousi2 a) true, b) starting and c) - f) recovered P-wave velocity models. All inversions are acoustic and the data being inverted is c) the acoustic, d) the elastic, e) the matched data with smooth v_p/v_s estimate and f) matched data with average v_p/v_s estimate. The average model RMS difference with respect to the true model is computed for all models within the dotted box, and it is displayed at the bottom of each model. 80

3.10. Vertical slices of a) the true v_p/v_s, b) a smooth v_p/v_s obtained by mildly smoothing the true v_p/v_s and c) an average 1D v_p/v_s estimate obtained by averaging the true v_p/v_s within two regions and mildly smoothing the result. 81

3.11. Graphs of the model RMS difference with respect to the true model as a function of a) depth and b) position for the models in Figure 3.9. . . 83

3.12. Representative shot gathers at the centre of the Marmousi2 model for a the elastic data contamined with white Gaussian noise and b) the matched data obtained from the noisy elastic data, its acoustic inversion and the smooth v_p/v_s in Figure 3.10b. The gathers in c) and d) are the differences of the gathers abover with respect to the true acoustic data. 85

3.13. Vertical slices of the recovered P-wave velocity models obtained after a) acoustic FWI of the noisy elastic data and b) acoustic FWI of the matched elastic data with noise. 87

3.14. Vertical slices of the a) recovered P-wave velocity model and b) S-wave velocity model after full elastic FWI of the elastic data using the original starting P-wave velocity model, a starting S-wave velocity model built from the later and the average v_p/v_s and inverting for v_p and v_s simultaneously. 88

3.15. Vertical slices of the true a) P-wave and b) S-wave velocity models used to generate seismic data for the overthrust model. 90

3.16. Representative shot gathers at the centre of the 2D SEG/EAGE overthrust model for the a) observed acoustic, b) observed elastic and c) matched observed elastic data (obtained by matching 9 simultaneous traces and with a length of the filters of 450 ms). Panels d) and e) show the differences of the panels above with the true acoustic data. . 91

3.17. Vertical slices for the land 2D SEG/EAGE overthrust P-wave velocity model. From top to bottom: a) true model, b) starting model and recovered models after acoustic FWI of c) the acoustic, d) the elastic and e) the matched observed data. The average model RMS difference with respect to the true model within the dotted are are computed, and the values are overlaid on the models. 93

3.18. Vertical profiles for the models in Figure 3.17 at a) 3 km and b) 7.2 km position. 95

3.19. Average model RMS differences for each recovered model and the starting model in Figure 3.17 as a function of a) depth and b) position within the dotted box in Figure 3.17. 96

3.20. Graph of the average model RMS differences within the dotted area in Figure 3.17 as a function of iteration for the acoustic inversion of acoustic, elastic and matched observed data. 97

3.21. Shot gathers for a) the observed data and b) the acoustic modelled data obtained with the contractor's starting model for the Buzzard field dataset. Panels c) and d) display the F-K filtered observed and acoustic modelled gathers, respectively. Panel d) illustrates the presence of the post-critical reflection at the top-chalk and its multiples when elastic effects are negligible. Modified from (Stronge, 2018). 99

3.22. Reflection coefficients (solid line) and phase profiles (dashed lines) at the top-chalk of the Buzzard dataset for a) an acoustic and b) an elastic model, obtained using Zoeppritz equations for isotropic interfaces. Modified from (Stronge, 2018). 100

3.23. Representative shot gathers for a) the observed data, b) modelled acoustic data, c) modelled elastic data and d) matched observed data for the Buzzard dataset. The starting P-wave velocity model and an ane estimate of the v_p/v_s model based on the regional geology are used to generate synthetic data. 103
3.24. Vertical cross-line profiles at $X = 10.95\,km$ for the a) P-wave starting model, b) the estimated v_p/v_s and the recovered P-wave velocity models obtained with anisotropic 3D acoustic FWI of c) the observed and d) the matched datasets. The discontinuous line indicates the position in which a velocity profile is extracted. The chalk unit is dramatically thinned in pannel c), but is preserved in d). 105
3.25. Vertical in-line profiles at $Y = 5.95\,km$ for the a) P-wave starting model, b) the estimated v_p/v_s and the recovered P-wave velocity models obtained with anisotropic 3D acoustic FWI of c) the observed and d) the matched datasets. The discontinuous line indicates the position in which a velocity profile is extracted. The chalk unit is dramatically thinned in pannel c), but is preserved in d). 106
3.26. Velocity profiles for the starting model (black line) and the recovered P-wave velocity models obtained after anisotropic 3D acoustic FWI of observed data (red line) and matched data (blue line) for the well-log positions indicated in Figures 3.24 and 3.25. 108
3.27. Depth slices for the starting P-wave velocity model (first row) and the recovered models obtained after anisotropic 3D acoustic FWI of observed data (second row) and matched data (third row) at $300\,m$ depth (first column) and at $1150\,m$ depth (second column). 109
3.28. Depth slices at $50\,m$ depth for the a) P-wave and b) S-wave velocity models used for acoustic and elastic modelling. S-wave velocity is incorrectly set to non-zero values within certain regions in the water layer. 110

4.1. Recorded signal at 1490 m from a source in a homogeneous and isotropic acoustic medium (black line) and homogeneous and isotropic visco-acoustic medium with $Q = 20$ (gray line). 116
4.2. Vertical profiles of the 2D horizontally layered true P-wave velocity and absorption models A and B used to generate acoustic and visco-acoustic data. The starting model and the average of the two absorption models below the subsurface are also overlaid. 121
4.3. Vertical profiles at the centre of the 2D horizontally layered models using the conventional FWI functional – panels a) and b)– and an amplitude-balanced functional – panels c) and d)– after acoustic FWI of acoustic and visco-acoustic data – panels a) and c)– and after acoustic FWI of matched data – panels b) and d) – for one iteration of the workflow – panel d)– and for one and two iterations of workflows 1 and 2 – panel b). The true, starting and recovered velocity model for acoustic data are overlaid. 122
4.4. Shot gathers recorded at the centre of model B for the a) visco-acoustic data and the matched data obtained by computing Wiener filters c) trace-by-trace with a length of 325 ms and simultaneously matching 13 traces with lengths of e) 325 ms, g) 50 ms and i) 1.5 s. The gathers on the right column show the differences of those on the left with the acoustic data, in which the average RMS difference for all the traces in each shot gather with respect to the acoustic data has been overlaid. 126
4.5. Extracted trace at 4.2 km position for a central shot gather for acoustic (black line) and visco-acoustic (red line) data, as well as their difference (green line). 127

4.6. Trace comparison at 2 km position for the datasets in Figure 4.4 (red lines) with the true acoustic data (black line). Thus, panel a) corresponds to visco-acoustic data and panels b) to e) correspond to matched data obtained with filters of 325 ms length for panels b) and c), of 50 ms for panel d) and of 1.5 s for panel e). Wiener filters are computed trace-by-trace for panel b), whereas they are computed by matching 13 traces simultaneously for panels c) to e). 129

4.7. Vertical slices of the marine Marmousi2 a) true P-wave velocity, b) starting P-wave velocity, c) true Q and d) 1D average Q estimate models and the recovered models obtained after e) acoustic FWI of acoustic data, f) acoustic FWI of visco-acoustic data, g) acoustic FWI of matched data and h) visco-acoustic FWI of visco-acoustic data. . . 132

4.8. Vertical slice of the difference between the starting and the recovered models after acoustic FWI of the matched visco-acoustic data in Figures 4.7b and 4.7g, respectively. 134

4.9. Representative shot gathers recorded at the centre of the Marmousi2 model for a) acoustic, b) visco-acoustic and c) matched data. Panels d) and e) show the differences of the panels above with the same shot gather for the acoustic data. 135

4.10. Representative shot gathers at the centre of the Marmousi2 model for the a) visco-acoustic data with noise and b) the corresponding matched data. Panels c) and d) show the differences of the panels above with the noise-free acoustic data. 137

4.11. Vertical slices of the recovered P-wave velocity models after acoustic FWI of a) the noisy visco-acoustic data and b) the matched data using the original starting model, c) a new starting model obtained by smoothing a), and recovered velocity models after acoustic FWI of d) the matched data using the starting model in c) and e) the observed data using the starting model in c). Panel f) shows the recovered velocity model obtained from visco-acoustic FWI of the noisy visco-acoustic data using the original starting model and the Q model in Figure 4.7d, which is kept fixed during the inversion. 138

4.12. Q analysis based on Gabor transforms of the data. Panel a) shows the recorded trace after spherical divergence correction, its Gabor transform is computed in panel b) and the latter is converted to a 1D spectrum in panel c), which is used to estimate Q by linear fit of the logarithm of c), shown in panel d). The reference Q value is 80. 143

4.13. Interval Q analysis for a reflection trace as a function of time. Panel a) presents the trace after spherical divergence correction, the Gabor transform is plotted in panel b), the computed Q values as a function of time are shown in c) and the relative error of the estimated values in c) as a function of time is shown in d). The reference Q value is 80. 144

4.14. Extracted time sections for a) near-offset data, b) near-offset data after removing the direct arrival, c) near-offset data after removing the direct arrival and spherical divergence and d) the estimated Q model obtained using the data in c). 146

4.15. Seismic Q analysis for trace 600 of the extracted near-offset section in Figure 4.14c using Gabor transforms. Panel a) shows the extracted trace, panels b) to f) display the Gabor transforms and the windows in which 1D spectrum are computed (white window) with curves of constant $\chi = \omega\tau$ overlaid (black crosses), panels g) to k) show the computed 1D spectra for the panels above and panels l) to p) display their logarithm, the linear fit (in red) for a threshold of $5\,dB$ and the estimated Q value (at the top). 147

4.16. Models of the estimated quality factor in depth a) before lateral smoothing and b) after lateral smoothing. 148
4.17. Vertical slices of a) the contractors' starting velocity model, b) the estimated Q model and the contractor's models for the anisotropic Thomson's parameters c) δ and d) ϵ. 150
4.18. Vertical slices of the a) starting model and the recovered P-wave velocity models after VTI anisotropic acoustic FWI of the b) field observed data and the c) matched dataset. 152
4.19. Zoomed-in sections of the recovered models in Figures 4.18b and 4.18c for the recovered models after VTI anisotropic acoustic FWI of a) observed and b) matched data. 153
4.20. Recorded field data are displayed in panels a) and b) for two shots along the same sailing line and separated 22.5 km, whereas the corresponding matched data for these shots are respectively shown in panels c) and d) after matching 15 traces simultaneously and using a length of the filters of 325 ms. 154
4.21. Flattening of CIGs for a) the observed data using the recovered velocity model in Figure 4.18b, and for b) the observed data after Q-compensation with the Q estimated model in Figure 4.17b and the velocity model in Figure 4.18c after band-pass filtering. 156
4.22. Stacked migrated sections obtained from the observed data a) before Q-compensation and using the recovered P-wave velocity model after anisotropic acoustic FWI of the observed data and b) after Q-compensation and band-pass filtering using the recovered model after anisotropic acoustic FWI of the matched data. 157
4.23. Stacked migrated sections in Figure 4.22 overlaid with the velocity models in Figures 4.18b and 4.18c, which have been used to migrate the corresponding data. 158

5.1. Figure a) is a picture of an ultrasound tomography scanner, in which ultrasound waves are emitted and received by an array of transducers distributed in a semi-spherical surface, as in b). Figure c) is a drawing of the positioning of the patient during data acquisition. Modified after Ruiter et al. (2018). 164

5.2. 3D speed-of-sound numerical model of the breast used to generate ultrasound data. 167

5.3. Pairs of coronal and saggital views, respectively, for the numerical model of speed-of-sound of the breast. Panels a) and b) correspond to the true model, c) and d) to a starting model of just water, e) and f) are the result of 3D FWI using the water starting model and a starting frequency of $700\,kHz$, g) and h) are obtained from smoothing the true model with a Gaussian function in 3D, i) and j) are obtained after 3D FWI using the starting model in panels g) and h) and a starting frequency of $700\,kHz$ and k) and l) are obtained after 3D FWI using the water starting model and a minimum frequency of $300\,kHz$. Density and absorption are kept fixed during modelling and inversion, equal to those of water. 169

5.4. 3D numerical models for the a) density and b) quality factor for the breast. 171

5.5. Shot gather recorded at a single receiver ring array in an a) acoustic medium with fixed density and quality factor equal to those of water and b) in a visco-acoustic medium obtained with the variable density and Q models in Figure 5.4. 172

5.6. Pairs of coronal and sagittal views for the recovered models of the breast after 3D acoustic FWI of a) the acoustic data with constant density and absortion equal to those of water and b) of the visco-acoustic data ignoring density and absorption during the inversion. The inversions are performed using a water starting model and inverting the data from $300\,kHz$ up to $1\,MHz$. 173

5.7. Shot gathers recorded at a single ring array of receivers for the a) visco-acoustic data obtained with the density and absorption models in Figure 5.4 and b) the previous visco-acoustic data with added random noise. 174
5.8. Pairs of coronal and sagittal views for the recovered models of the breast after 3D acoustic FWI of the visco-acoustic data a) without and b) with random noise added in the data. The inversions are performed ignoring density and absorption effects, using a water starting model and inverting the data from $300\,kHz$ up to $1\,MHz$. 175
5.9. Representative ultrasound shot gathers obtained from a source in a single ring array of transducers surrounding a a) 2D slice of the *in silico* breast model and b) the full 3D model. 176
5.10. Coronal views of a) the true model and the recovered models after b) 2D FWI and b) 3D FWI of the 3D data modelled for a single ring array of transducers around the full 3D model. 177
5.11. Representative shot gather for the *in vivo* data a) before and b) after band-pass filtering using a zero-phase Butterworth band-pass filter that rejects data below $200\,kHz$ and beyond $1.75\,MHz$. 178
5.12. Selected traces used for extracting a source wavelet from the *in vivo* data for a single shot gather after LMO correction using a) the nominal and b) the calibrated location of the transducers. 180
5.13. Stacked traces obtained after LMO correction of the direct arrivals of a) the band-pass filtered and muted *in vivo* data and b) the modelled data for a water shot using a band-pass filtered Ricker wavelet as a first guess. 181
5.14. Plot of the a) extracted source wavelet in time and b) its amplitude spectrum used for 2D FWI. 182

5.15. Recovered models of speed-of-sound obtained with a starting model of just water and a) the nominal and b) the calibrated position of the transducers. The inversions are performed from $400\,kHz$ to $1.4\,MHz$ in a total of 100 iterations and using 85 sources per iteration. Data from $0\,\mu s$ to $180\,\mu s$ are used, hence including coherent noise. 183

5.16. Recovered models of speed-of-sound for the *in vivo* data obtained after a) 2D FWI with a starting model of just water and a minimum frequency of $500\,kHz$, b) bent-ray tomography and 2D FWI using the bent-ray tomography model as a starting model and c) a minimum frequency of $800\,kHz$ and d) a minimum frequency of $500\,kHz$. Data before $30\,mus$ and after $180\,mus$ is excluded and 85 sources per iteration are used for the inversions, which are performed up to $1.75\,MHz$. 185

5.17. Zoomed in sections of the recovered models of speed-of-sound after a) bent-ray tomography and b) 2D FWI with a minimum frequency of $500\,kHz$ shown in Figures 5.16b and 5.16d. 187

6.1. *In silico* models of the speed-of-sound within the head used for a) 3D and b) 2D experiments. 196

6.2. Three-cycle tone-burst source wavelet a) in time and b) its amplitude spectrum used to model 2D *in silico* data for a transverse section of the 3D head model. 197

6.3. Representative shot gather obtained for the 2D model of the head in Figure 6.1b a) before and b) after adding random noise. 198

6.4. Transverse sections of the head for the a) true and starting models with b) the exact skull and c) a smooth skull. The bottom pannels show the recovered models of the head obtained with starting models of d) just water, e) water and the exact skull as in b) and f) water and the smooth skull as in c). 199

6.5. Recovered transverse sections of the head with FWI using a) the exact skull starting model in Figure 6.4b and b) the smooth skull model in Figure 6.4c for the noisy observed data. 202

6.6. Recovered transverse sections of the brain obtained with the exact skull for a) the full dataset and b) only the nearest offsets (below 3 cm source-receiver offset). 202

6.7. Transverse sections of the brain for the true a) acoustic speed, b) density and c) absorption models. The bottom panels show the recovered models of acoustic speed obtained with FWI ignoring density and absorption effects for data generated d) with variable density and no absorption, e) with fixed density and variable absorption and f) with variable density and absorption. An amplitude-balanced functional is used to minimise the effect of the latter, and the exact skull starting model is used throughout. 204

6.8. Cross-plot of experimental values for density and acoustic velocity for different body tissues, which have been obtained from Iacono et al. (2015). A polynomial function of fourth order is fitted to the data, which can potentially be used to derive density from acoustic velocity in FWI. 205

6.9. Representative shot gathers obtained after modelling data with a) variable density and no absorption, c) fixed density and variable absorption and e) variable density and absorption. The panels on the right display the difference of the panels on the left with the data generated with fixed water density and no absorption shown in Figure 6.3a. 206

6.10. Transverse sections of the brain for a) the modified true model containing a blood clot and a hemorrhage, b) the exact skull starting model, c) the smooth skull starting model and the recovered models after 2D FWI of the noise-free data in d), e) and f) using a water starting model, the exact skull starting model and the smooth skull starting model, and of the noisy data using g) the exact skull starting model and h) the smooth starting model. 208

6.11. Representative shot gather obtained for the modified head model to simulate a stroke patient a) without and b) with noise. Panel c) shows the difference of the gather in a) with that obtained for a healthy head model, as displayed in Figure 6.3a. 209
6.12. Representative shot gather generated by a) modelling data in 2D across a transverse section of the head and b) modelling data in 3D with the full model of the head. In both cases, the transducers are in a single ellipsoidal array in the same plane as the transverse section shown in Figure 6.4a, and the amplitudes in both gathers are normalised with the average RMS of each shot for comparison. 210
6.13. Recovered models obtained by 2D FWI of a) the 2D data and b) the 3D data. Both inversions have been performed using the exact skull starting model and inverting the data from $100\,kHz$ to $550\,kHz$. . . . 211
6.14. Sagittal views of a) the true and starting speed-of-sound models of the head with b) the exact skull and c) the skull after smoothing is applied. The bottom panels show the recovered models of acoustic speed obtained with 3D FWI using a d) homogeneous water starting model, e) the exact skull starting model and f) the smooth skull starting model. 212
6.15. Transverse views of a) the true and starting speed-of-sound models of the head with b) the exact skull and c) the skull after smoothing is applied. The bottom panels show the recovered models of acoustic speed obtained with 3D FWI using a d) homogeneous water starting model, e) the exact skull starting model and f) the smooth skull starting model. 213
6.16. Transverse sections of estimates of the skull (left column) and of the recovered models with 2D FWI of data generated with variable density and no absorption using the estimates on the left column as starting models. The estimates of the skull correspond to a) the exact skull and the estimated skulls obtained with method 1 and b) the first strategy, c) the second strategy (flooding with constant thickness), d) inwards smoothing of the model in c) and e) smoothing in all directions of the model in c). 217

6.17. Each column corresponds to a different sagittal section at $X = 17.0\,cm$, $X = 14.2\,cm$ and $X = 11.4\,cm$ of the true model of speed-of-sound (first row), the recovered model obtained by 3D FWI of near-offset data (second row), estimated model of speed-of-sound of the skull (third row) and 3D FWI of the full dataset using the estimated model of the skull. 219

6.18. Each column corresponds to a different transverse section at $Z = 8.8\,cm$, $Z = 8.2\,cm$ and $Z = 6.5\,cm$ of the true model of speed-of-sound (first row), the recovered model obtained by 3D FWI of near-offset data (second row), estimated model of speed-of-sound of the skull (third row) and 3D FWI of the full dataset using the estimated model of the skull. 221

B.1. Normal incicdence (zero-offset) reflection dataset obtained for the transverse section of the head model with variable density and no attenuation in Figure 6.7, in which travel-time picks for the onset of a) the reflection at the outer table and b) at the inner table are overlaid (white crosses). 237

B.2. Succession of steps to find the outer table of the skull from reflection data. From left to right and top to bottom in green, a) circles with radii depending on the travel-time picks of the onset of the reflection, b) tangent points between pairs of circles closer to the centre of the model, removal of outliers after stages c) 1, d) 2 and e) 3, resulting in the clean points in f), g) clean points after smoothing and h) interpolation and i) polygon that goes through all the points (green line). The true model is overlaid for comparison purposes. 239

B.3. Distance matrices obtained during removal of outliers in step 1 a) before and b) after 2D smoothing. Those points in which the average distance across a column or a raw have a difference with the smooth distance matrix than a threshold are removed, and a distance value of zero is given resulting in the distance matrix in c). 240

B.4. 1D plots of the spatial coordinates for the filtered tangent points after step 1 a) before and b) after 1D smoothing. Those points in which the difference between the smooth and raw coordinates is larger than a certain threshold are removed, and assigned a zero distance value. This results in the plot in panel c). 241

B.5. Plots for steps 3 and 7 of a) the circles corresponding to the inner table (green cicles) and b) the extracted polygon (green line) for the first strategy to pick the onset of the reflection at the inner table. The bottom panels show the corresponding plots obtained for the second strategy to pick the arrival times, obtained from the outer table plus a constant thickness of the skull of $4\,mm$. The true model of acoustic speed is overlaid. 242

B.6. Final step to obtain a starting model for the skull using method 1. The region between the outer and inner tables in a) is filled with zeros, ones are assigned everywhere else and smoothing is applied, resulting in b). This is then converted to typical velocities of the skull in c) by assigning velocities to those regions with values between 0 and 0.4. The top panels show the results obtained for the inner table computed with the first strategy, whereas the bottom panels correspond to those obtained with the second strategy (constant skull thickness). 243

C.1. Permission granted for Figures 2.1, 2.4 and 2.5. Signature has been removed for privacy. 245

C.2. Permission granted for Figure 2.2. 246

C.3. Permission granted for Figure 2.3. 247

List of Algorithms

2.1. Computation of matched observed data using Wiener filters that match input traces (e.g. elastic or visco-acoustic modelled data) to desired traces (e.g. acoustic modelled data). 52

List of acronyms

1D	one dimension
2D	two dimensions
3D	three dimensions
AIC	Akaike information criterion
AWI	adaptive waveform inversion
CDP	comon depth point
CFL	Courant-Friedrichs-Lewy
CIG	common-image gather
CSF	cerebro-spinal fluid
CT	computed tomography
CURE	computer ultrasound risk evaluation
FD	finite-differences
FFT	fast Fourier transform
FWI	full-waveform inversion
HPC	high performance computing
HTI	horizontal transversely isotropic
LMO	linear move-out
MDS	multi-dimensional scaling
MPI	message passing interface
MRI	magnetic resonance imaging
PDE	partial differential equation
PET	positron emission tomography
PML	perfectly matched layer
RMS	root mean square
TI	transversely isotropic
TOF	time-of-flight
TT	travel-time tomography
TTI	tilted transversely isotropic
VTI	vertical transversely isotropic

1. Introduction

Imaging techniques play an essential role in many scientific fields, as they allow us to characterise and obtain a complete picture of an object otherwise invisible to the human eye. In particular, tomographic algorithms aim at reconstructing slices of quantitative models of physical properties by using observations of the wavefield that has travelled through an object, which is recorded at receivers around – or on one side only – of the object. In this manuscript, I focus on an iterative tomographic algorithm that uses the complete information of amplitude and phase of the signal as a function of time at each receiver, also known as full-waveform inversion (FWI). This technique requires the solution of an inverse problem, that is the inference of physical properties of an object from the measured data, which also demands accurate forward modelling of wave propagation. Due to its elevated computational cost and its complexity, forward modelling is typically performed using the acoustic wave equation. This assumption results in acceptable reconstructed models of physical properties under certain circumstances, but it can lead to inaccurate models in others.

FWI has been extensively applied in geophysics to obtain high-fidelity high-resolution models of the subsurface (Sirgue et al., 2010; Warner et al., 2013). In the ideal case, the full content of the recorded seismogram is used and a wave equation that accounts for the inherent visco-elastic nature of the earth is utilised to model wavefield propagation (Virieux and Operto, 2009). A model of the subsurface is then sought in an iterative fashion by minimising the misfit between the observed and modelled data. In principle, FWI has the potential to be used to obtain multiple parameter information of the subsurface, but these effects are typically ignored and the acoustic wave equation is used to reconstruct P-wave velocity models only. This is mainly due to three reasons: 1) the large computational cost of visco-elastic modelling, espe-

cially in three dimensions (3D) (Guasch et al., 2010; Hobro et al., 2014), 2) to avoid cross-talk between the different inverted parameters given the coupled effect of the different physical properties of the wavefield (Virieux and Operto, 2009) and 3) the complexity of modelling amplitude accurately without knowing the source wavelet exactly or having accurate information of density, S-wave velocity and attenuation (Tarantola, 1986).

Considering visco-elastic effects is also expected to lead to improved P-wave velocity models while only inverting for the latter parameter class (Warner et al., 2012; Kurzmann et al., 2013), as accounting for the right kinematics and dynamics of the wavefield as well as for phase conversions at interfaces should reduce the misfit between observed and modelled data. With this in mind, several authors have suggested strategies to account both for elastic and viscous effects in acoustic FWI. For instance, Shen (2010) proposed an amplitude-balanced data misfit – or functional – that focuses more on the phase rather than the amplitude of the signal by weighting the observed and modelled data by their root mean square (RMS) amplitude. This method is now standard practice in the oil industry as it mitigates some of the elastic and viscous effects during acoustic FWI without any added cost, but it is insufficient in the presence of strong elastic and viscous effects. Alternatively, other approaches have been suggested to address either elastic or viscous effects independently. In order to address elasticity to some extent in acoustic FWI and at low cost, Veitch et al. (2012) and Li et al. (2017) have recently demonstrated that the method proposed by Chapman et al. (2010), Hobro et al. (2014) and Chapman et al. (2014) to correct the amplitude of acoustic finite-difference simulations for elastic effects can improve the quality of the recovered P-wave velocity models obtained with acoustic FWI by including a correction term to the acoustic wave equation. Although this method accounts for amplitude effects, it neglects phase conversions and it can also degrade the result of acoustic simulations in the presence of strong velocity contrasts. On the other hand, Kurzmann et al. (2013) showed that, in the presence of viscous effects in the data, the quality of the recovered P-wave velocity models significantly improves if a fixed attenuation model is considered during modelling in FWI. This strategy considers attenuation to be a modelling parameter and not an inversion parameter, it

assumes density to be homogeneous and it is more expensive than conventional acoustic FWI as it requires modelling viscous effects at each iteration during the inversion. Additionally, Kurzmann et al. (2013) also demonstrate that neglecting viscous effects has a negative impact on the recovered P-wave velocity models for marine datasets with soft sediments.

Unlike in geophysics, the application of acoustic FWI in medical imaging is still in its infancy, especially on three-dimensional datasets generated with ultrasound. This is in part due to the computational cost of this method given that imaging speed is crucial for diagnosing diseases promptly, but new technological advances and increasingly faster computers means this technique can now be implemented in this field. The application of FWI to ultrasound data for medical imaging holds a great potential and could soon be adopted as a safe, portable and alternative imaging method. To date, most of the few applications of acoustic FWI to *in vivo* ultrasound datasets are focused on breast imaging, especially for breast cancer diagnosis (e.g. Pratt et al. (2007)). More recently, Bernard et al. (2017) have demonstrated on two-dimensional *in silico* (i.e. numerical) examples that acoustic FWI can be used to obtain quantitative images of bones, but this has not yet been applied in 3D or demonstrated on *in vivo* datasets.

In this thesis I first present a novel methodology to address elastic and viscous effects in acoustic FWI, which is based on the use of matching filters and does not require modelling elastic or viscous effects during the inversion, thus reducing the computation time as the acoustic wave equation is used during the inversion. I also demonstrate this method can be combined with that of Shen (2010), leading to better quality P-wave velocity models of the subsurface. At the same time, I analyse the impact of both elasticity and viscous effects on the recovered P-wave velocity models, and I show this method leads to improved P-wave velocity models of the subsurface compared to those obtained after ignoring the latter. Secondly, I demonstrate the potential of 3D acquisition and 3D FWI for imaging the breast to detect breast cancer with ultrasound, and analyse different aspects to assess its practical implementation. I then implement FWI on an ultrasound breast *in vivo* dataset having very little prior information on the acquisition system, and obtain a highly detailed recovered model

of P-wave velocity for the breast. Finally, I propose FWI of ultrasound data can be used as a non-invasive method to image the brain and I show, for the first time, its potential on 2D and 3D datasets, obtained using models of a human head.

In terms of organisation of this thesis, there are a total of seven chapters. In this first chapter I have briefly introduced the topics and motivation of this thesis, as well as the methodology used to tackle the stated problems and the results obtained. In the second chapter I discuss the historical background of FWI in more detail, I review the theoretical framework that leads to acoustic FWI, I present a new methodology to address elastic and viscous effects based on Wiener filters and I further explore the use of FWI for medical imaging. I then apply the proposed methodology to address elastic and visco-acoustic effects to geophysical problems in the third and fourth chapters, respectively, and I analyse its limitations and applicability for different scenarios. I then dedicate the fifth and sixth chapters to evaluate the applicability and discuss potential applications of acoustic FWI to medical imaging problems, in particular to image the breast and the brain, respectively. In these chapters I also discuss the difficulties encountered when performing FWI on these datasets and present solutions to most of these problems. Finally, I summarise the various conclusions of this thesis in the seventh chapter and I discuss the next steps of this research.

2. Methods

In the last decades, a thorough understanding of wave propagation has led to many far-reaching technological advances. The use and observation of waves have been crucial in disciplines such as astrophysics, quantum physics, medical imaging, telecommunications or seismology. As a result, there are numerous applications that are nowadays based on wave propagation. In particular, mechanical waves – i.e. oscillations of particles in a medium that results in a transfer of energy – are currently used to control the characteristics of musical instruments and its broadcasting, to image the earth's interior after recording a seismic signal at the surface or to monitor a baby in the womb and diagnose certain diseases in adults by using ultrasound waves.

Propagation of mechanical waves in a medium is described by the wave equation. In this chapter I begin by reviewing the formulation of the wave equation in different media, as these are the equations that are implemented in the forward modelling code used in the following chapters, both for imaging the earth and for imaging the human body. I then discuss two tomographic algorithms used to obtain models from observed data: travel-time tomography (TT) and FWI. I review the theoretical and historical background of FWI in more detail, as I use this technique throughout this research work, and I discuss what is current standard practice for imaging 3D models in geophysics (i.e. the acoustic approximation). Next, I present and discuss the methodology and algorithm I propose to use to mitigate elastic and viscous effects and obtain improved P-wave velocity models of the subsurface, which I then apply in the third and fourth chapters. Finally, I explore the similarities and differences of wave propagation in geophysics and medical imaging with ultrasound and I discuss how modelling and inversion differs from the former to the latter. The last section of this chapter is dedicated to describe the originality of my work, which is explained in

more detail at the beginning of the following chapters.

2.1. The wave equation

Propagation of mechanical waves in a material is described by the wave equation, which is mathematically formulated as a second-order linear partial differential equation (PDE) that involves time and space derivatives. Its solution on real scenarios requires the use of numerical methods, and it can be costly to compute. Thus, the wave equation is often approximated to reduce its associated computational cost. In geophysics, the acoustic or elastic approximations are normally used, with or without anisotropy and with or without considering viscous effects. In medical imaging, anisotropy is of minor importance and it is typically not considered. Here I review the basics of wave propagation relevant to this thesis, focusing on the formulation for wave propagation in the subsurface, and I discuss its numerical implementation. A more detailed discussion of the theory of wave propagation can be found in Carcione (2001) and Aki and Richards (2002). The similarities between the wave equation used in geophysics to model seismic wave propagation in the subsurface and the wave equation used to model wave propagation of ultrasound in the human body are discussed in section 2.6.

2.1.1. Anisotropic elastic media

Ignoring energy losses caused by absorption, the earth can be described as an anisotropic elastic medium (Chapman, 2004), that is a medium in which the mechanical properties differ in different directions and that reverts to its original state when the applied forces are removed (Aki and Richards, 2002). Formulation of the wave equation – or equation of motion – in anisotropic elastic media is obtained by the conservation of linear momentum – as a consequence of Newton's laws of motion –, and considering average properties of the materials instead of a detailed description of their atomic and crystalline structures. This leads to a partial differential equation (PDE) that is

the equation of motion (Auld, 1990; Aki and Richards, 2002):

$$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = \rho \frac{\partial^2 u_i}{\partial t^2} \tag{2.1}$$

in which I use Einstein's summation convention to denote summation over repeated indices, ρ is the mass density, $\partial/\partial x_i$ denotes the spatial partial derivative along the ith component, with $i = 1, 2, 3$ respectively representing the x, y and z Cartesian coordinates, t is time and u_i, f_i and σ_{ij} are the components of the particle displacement vector, the body force vector per unit volume and the stress tensor, respectively. The components of the stress tensor describe the internal forces that neighbouring particles in a material exert on each other across different planes when the material is deformed. Aki and Richards (2002) demonstrate that the conservation of angular momentum implies that the stress tensor is symmetric, i.e. $\sigma_{ij} = \sigma_{ji}$. To solve the wave equation, we need an additional relation that describes the characteristics of an anisotropic elastic medium. This requires introducing the concept of strain tensor, which describes the relative change in position of the end points of an element after deformation. In case these deformations are small, the latter is defined as follows:

$$\epsilon_{ij} \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{2.2}$$

We note the strain tensor is symmetric by definition, i.e. $\epsilon_{ij} = \epsilon_{ji}$. The definition of strain tensor is of importance as its relationship with the stress tensor, also known as the constitutive relation or constitutive law, characterises a medium. The simplest relationship between the stress and strain tensors for an anisotropic elastic medium is the so-called generalised Hooke's law (Aki and Richards, 2002), which states that the stress tensor is a linear combination of all the components of the strain tensor:

$$\sigma_{ij} = c_{ijkl}\epsilon_{kl} \tag{2.3}$$

where c_{ijkl} are the $3^4 = 81$ components of a fourth-order tensor of stiffness or elastic constants, $\boldsymbol{c}$, and i, j, k, l are the components in the Cartesian system of coordinates. There is experimental evidence that this constitutive relationship holds for small de-

formations, i.e. $\epsilon_{ij} \ll 1$, such as in most seismic applications, and thus this is a reasonable relationship when modelling seismic wave propagation (Chapman, 2004). For large deformations (e.g. close to an earthquake source, in the Earth's interior or in body tissues in some circumstances) this constitutive law is no longer valid and the finite-strain theory needs to be used (Hu, 2004). The linear constitutive law in equation (2.3) yields nine different equations for the stress components, such that each equation contains nine components of the strain tensor. Due to the symmetries of the strain and stress tensors, the components of the fourth-order tensor $\boldsymbol{c}$ reduce from 81 to 36 (Carcione, 2001; Aki and Richards, 2002; Chapman, 2004) and, considering heat exchange can be neglected during deformations (i.e. an adiabatic process), these further reduce to 21 components (Aki and Richards, 2002). Substitution of the constitutive law in equation (2.3) into equation (2.1) leads to the set of equations of motion for anisotropic elastic media.

These equations can be particularised to describe anisotropic media with certain symmetry properties. In particular, most anisotropic systems in the subsurface can be described as transversely isotropic (TI) media, or with hexagonal symmetry (Thomsen, 1986; Tsvankin, 2001; Nolet, 2008). These media have a single axis of symmetry and are isotropic perpendicular to the latter. As a result, the velocity of propagating waves depends on the angles relative to the axis of symmetry. For example, the intrinsic anisotropy of shales – which represent approximately 75 % of the clastic fill of sedimentary basins– can be well represented by a TI model formed by aligned clay particles that are typically horizontally layered (Tsvankin, 2001). Depending on the orientation of the symmetry axis, TI media are classified as vertical transversely isotropic (VTI) (vertical axis of symmetry), horizontal transversely isotropic (HTI) (horizontal axis of symmetry) or tilted transversely isotropic (TTI) (tilted axis of symmetry) media, as depicted in Figure 2.1. VTI anisotropic media have a vertical axis of symmetry and can be caused by thin horizontal layering (Figure 2.1a), whereas HTI systems have a horizontal axis of symmetry and can be caused by vertical layering (Figure 2.1b). If the axis of symmetry is tilted with respect to the vertical axis of symmetry, for instance if the layering is dipping, the system is known as TTI (HTI is a particular case of TTI).

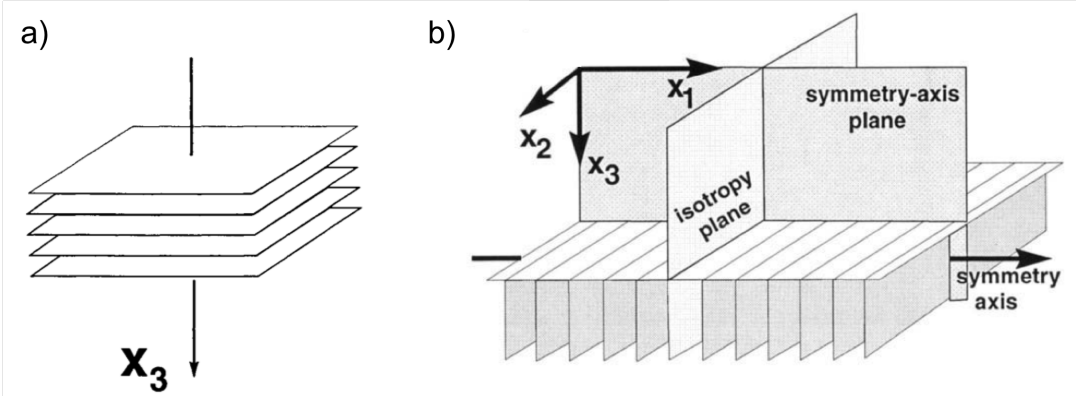


Figure 2.1.: Models of a) vertical transversely isotropic (VTI) and b) horizontal transversely isotropic (HTI) anisotropy caused by thin horizontal and vertical layering, respectively. Modified after Tsvankin (2001).

In VTI anisotropic media, the tensor of elastic constants that characterises the intrinsic symmetry reduces to (Carcione, 2001; Tsvankin, 2001):

$$\boldsymbol{c}^{(vti)} = \begin{pmatrix} c_{11} & c_{11}-2c_{66} & c_{13} & 0 & 0 & 0 \\ c_{11}-2c_{66} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{pmatrix} \tag{2.4}$$

with just five independent parameters, whereas the latter for HTI media takes the form (Tsvankin, 2001):

$$\boldsymbol{c}^{(hti)} = \begin{pmatrix} c_{11} & c_{13} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{33} & c_{33}-2c_{44} & 0 & 0 & 0 \\ c_{13} & c_{33}-2c_{44} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{55} \end{pmatrix} \tag{2.5}$$

also with five independent parameters. The tensor of elastic constants for TTI media is obtained by rotating tensor for VTI media in equation (2.4) according to

the tilt of the axis of symmetry. Substitution of these tensors into equation (2.3) and further into equation (2.1) leads to the wave equation for TI media as a function of the elastic coefficients. Thomsen (1986) introduced a parametrisation valid for VTI media with effective parameters that govern the seismic wavefield in these media in the case of weak anisotropy ($< 20\,\%$), which is often the case in the subsurface. With this parametrisation, the five independent elastic coefficients can be replaced by the vertical (along the axis of symmetry) P- and S-wave velocities (propagating in the direction of propagation and perpendicular to it), $v_{p,v}$ and $v_{s,v}$, and three dimensionless anisotropic parameters, ϵ, δ and γ, which take the form:

$$v_{p,v} \equiv \sqrt{\frac{c_{33}}{\rho}} \tag{2.6}$$

$$v_{s,v} \equiv \sqrt{\frac{c_{55}}{\rho}} \tag{2.7}$$

$$\epsilon \equiv \frac{c_{11} - c_{33}}{2c_{33}} \tag{2.8}$$

$$\delta \equiv \frac{(c_{13} + c_{55})^2 - (c_{33} - c_{55})^2}{2c_{33}(c_{33} - c_{55})} \tag{2.9}$$

$$\gamma \equiv \frac{c_{66} - c_{55}}{2c_{55}} \tag{2.10}$$

In this notation, the strength of anisotropy is described by ϵ, δ and γ, so that they are zero for isotropic media ($c_{11} = c_{33}$, $c_{55} = c_{66}$ and $c_{13} = c_{11} - 2c_{66}$), and the velocities of P-waves and S-waves with horizontal and vertical polarisation are determined by the vertical P- and S-wave velocities and the anisotropic parameters (Thomsen, 1986, equations 10a-10c).

2.1.2. Isotropic elastic media

For isotropic media, in which the mechanical properties do not depend on the direction of wave propagation, the number of independent elastic coefficients reduces to the

two Lamé moduli (Jeffreys and Jeffreys, 1972; Aki and Richards, 2002), λ and μ, which characterise the elastic properties of an isotropic medium. In this case, the components of $\boldsymbol{c}$ have the form:

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) \tag{2.11}$$

where δ is the Kronecker delta ($\delta_{ij} = 1$ if $i = j$ and zero otherwise), which can be derived from a TI medium provided $c_{11} = c_{33} = \lambda + 2\mu$, $c_{55} = c_{66} = \mu$ and $c_{13} = c_{12} = \lambda$ (Carcione, 2001). Substitution of equation (2.11) into the generalised Hooke's law in equation (2.3) yields the relation between stress and strain tensors in isotropic elastic media – also referred from here onwards as elastic media for simplicity:

$$\sigma_{ij} = \lambda\delta_{ij}\epsilon_{kk} + 2\mu\epsilon_{ij} \tag{2.12}$$

Considering the definition of the strain tensor for small deformations in equation (2.2), this yields:

$$\sigma_{ij} = \lambda\delta_{ij}\frac{\partial u_k}{\partial x_k} + \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{2.13}$$

The latter together with the equation of motion in (2.1) leads to a system of nine coupled equations for stress (six equations) and particle displacement (three equations). The numerical solution of this system is used to model wave propagation in isotropic elastic media (Guasch, 2012).

Isotropic elastic media can support both compressional and shear forces caused by a propagating wave, as illustrated in Figure 2.2.

The application of a compressional force reduces the volume of an elastic element, as in Figure 2.2a. After this force is removed, the element returns to its original position. Additionally, elastic elements also resist twisting motions caused by shear forces. As a result of shear forces, a cubic element changes its shape into a trapezoid, as in Figure 2.2b, without changing its volume, and it returns back to its original shape once shear forces are removed. In a real isotropic elastic medium, motion takes place when there is a stress difference between neighbouring particles (Alsadi,

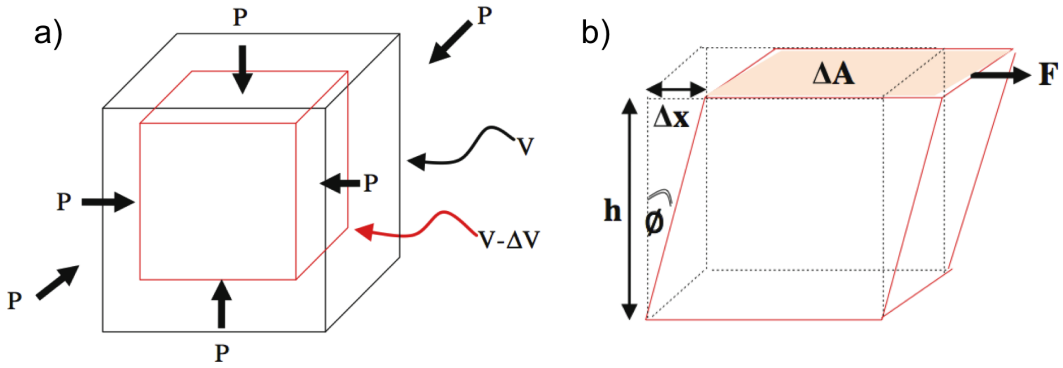


Figure 2.2.: The effect of a) a compressional force P (hydrostatic pressure) acting on an element of volume V changes the volume of an elastic element, whereas b) a shear force F acting on an area ΔA changes the shape of an elastic element. Modified after Alsadi (2016).

2016). In these materials both compressional and shear deformations are possible, and these propagate at different velocities within the medium. These velocities can be derived for a source-free (i.e. $\boldsymbol{f} = \mathbf{0}$) homogeneous[1] isotropic elastic medium upon substitution of (2.13) into the equation of motion, equation (2.1), resulting into the following expression:

$$\lambda \frac{\partial u_k}{\partial x_i \partial x_k} + \mu \left(\frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial u_j}{\partial x_i \partial x_j} \right) = \rho \frac{\partial^2 u_i}{\partial t^2} \tag{2.14}$$

Express now this equation in vector notation considering $\nabla \cdot \boldsymbol{u} = \partial u_i / \partial x_i$ and the vector property $\nabla^2 \boldsymbol{u} = \nabla (\nabla \cdot \boldsymbol{u}) - \nabla \times (\nabla \times \boldsymbol{u})$, this results into the following expression (Wapenaar and Berkhout, 1989; Aki and Richards, 2002):

$$(\lambda + 2\mu) \nabla (\nabla \cdot \boldsymbol{u}) - \mu \nabla \times (\nabla \times \boldsymbol{u}) = \rho \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \tag{2.15}$$

where bold letters represent vectors. The previous equation can be separated into two simpler expressions that are related to the change in volume and shear motion. First, taking the divergence of equation (2.15) and considering the vector properties $\nabla^2 \Phi = \nabla \cdot (\nabla \Phi)$ and $\nabla \cdot (\nabla \times \boldsymbol{A}) = 0$, results into the following equation for the

[1] A homogeneous medium is one in which the properties are the same at every point, e.g. density and the Lamé moduli are the same within a layer of a certain material in the subsurface.

change in volume:

$$(\lambda + 2\mu)\,\nabla^2\theta = \rho\frac{\partial^2\theta}{\partial t^2} \tag{2.16}$$

where $\theta \equiv \nabla \cdot \boldsymbol{u}$. On the other hand, taking the curl of equation (2.15) and using the vector property $\nabla\,(\nabla\Phi) = 0$ yields:

$$\mu\nabla^2\Psi = \rho\frac{\partial^2\Psi}{\partial t^2} \tag{2.17}$$

in which $\Psi \equiv \nabla \times \boldsymbol{u}$. Equations (2.16) and (2.17) have the form of wave equations with velocities given by:

$$v_p = \sqrt{\frac{\lambda + 2\mu}{\rho}}, \qquad v_s = \sqrt{\frac{\mu}{\rho}} \tag{2.18}$$

The particle velocity v_p is curl-free and is associated to compressional waves or P-waves that change the volume of the material ($\nabla \cdot \boldsymbol{u} \neq 0$), which travel in the same direction as wave propagation. On the other hand, the velocity obtained from the second wave equation, v_s, is divergence-free and is related to shear waves or S-waves that propagate perpendicular to the direction of wave propagation (Wapenaar and Berkhout, 1989; Ikelle and Amundsen, 2005). By comparing the expressions for both velocities, P-waves are always faster than S-waves (the Lamé moduli are positive), hence the name of primary and secondary waves, respectively, as the former arrive earlier at receivers. These two types of waves are classified as body waves, as they travel within a material, and their effects can be seen on the recorded data in isotropic elastic media.

2.1.3. Isotropic acoustic media

Acoustic media can only support compressional forces and, hence, shear stresses are zero even in the presence of shear strains (Wapenaar and Berkhout, 1989; Chapman, 2004). By definition, ideal fluids are acoustic media as only their volume can be changed after the application of a compressional force, and shear stresses are always zero (Chapman, 2004). In isotropic acoustic media, the stress tensor is diagonal (the

off-diagonal terms represent shear stresses and, hence, need to be zero) and isotropic, with components $\sigma_{ii} = -p$, being p the hydrostatic pressure. As in these media there cannot be shear wave propagation, it follows from equation (2.18) that $\mu = 0$, and that $v_p = \sqrt{\lambda/\rho}$. The acoustic wave equation is derived from equation (2.1) considering the stress tensor is diagonal, yielding the following expression in vector form:

$$-\nabla p + \boldsymbol{f} = \rho \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \tag{2.19}$$

And the constitutive law that characterises acoustic media is derived by substituting $\mu = 0$ in equation (2.13), which results into the expression $\sigma_{ii} = \lambda \delta_{ii} \partial u_k / \partial x_k = -p$, or in vector form $\nabla \cdot \boldsymbol{u} = -p/\lambda$. The latter can be used to express the wave equation in terms of the hydrostatic pressure only after taking the divergence of equation (2.19) and substituting the constitutive relation, which yields:

$$\nabla^2 p - \frac{1}{v_p^2} \frac{\partial^2 p}{\partial t^2} = \nabla \cdot \boldsymbol{f} \tag{2.20}$$

2.1.4. Anisotropic TI acoustic media

Anisotropy is inherent in elastic media. However, an acoustic wave equation for TI media can be obtained if the velocity of S-waves in the equations of motion for these media is set to zero, which is called *pseudo-acoustic anisotropy* and is often shortened to *acoustic anisotropy*. The resulting equations describe the kinematics of P-wave propagation in TI media more accurately than that obtained with the isotropic acoustic wave equation (Alkhalifah, 2000) with an associated modelling cost inferior to that of solving the elastic anisotropic wave equation. With this aim, Alkhalifah (2000) and Zhang et al. (2005) derived acoustic wave equations for VTI and TTI media from dispersion relations and setting the shear wave velocity to zero. These equations are generally used to model P-wave velocity in TI anisotropic media, which use a notation based on Thomsen's parametrisation (equations (2.6) to (2.10)).

2.1.5. Visco-elastic and visco-acoustic media

So far I have considered wave propagation in ideal lossless elastic media. In reality, there is a loss of energy of propagating waves in the subsurface due to several dissipation mechanisms (Aki and Richards, 2002). Energy absorption leads to an amplitude decrease and phase dispersion of the wavefield due to visco-elasticity. In the context of linear visco-elasticity, the stress linearly depends on the strain rate, i.e. the strain at preceding times (Carcione, 2001) and, thus, it is said that the solid has memory. In mathematical terms, the constitutive law in these media is expressed as a convolution:

$$\sigma_{ij} = \Psi_{ijkl} * \frac{\partial \epsilon_{kl}}{\partial t} = \frac{\partial \Psi_{ijkl}}{\partial t} * \epsilon_{kl} \tag{2.21}$$

where Ψ_{ijkl} are the components of the relaxation tensor, which are assumed to be causal functions (i.e. the stress tensor depends on the variation of strain rate up to the current time, but it does not depend on future strain states) (Carcione, 2001; Butler, 2005), hence the second equality in the previous equation. Additionally, each component of the relaxation tensor is a decreasing function of time under the assumption that the stress is more affected by recent history variations of strain than previous variations. Carcione (2001) shows that, due to the symmetries of stress and strain tensors, the number of independent components of the relaxation tensor reduces to 36, and further reduce to 21 if Ψ is considered symmetric. In practice, the visco-elastic behaviour of the subsurface can be described by deriving relaxation functions using mechanical models that represent energy storage and dissipation. These models can contain springs with no weight, in which energy is stored, and dashpots, in which energy is dissipated. In this context, it is useful to introduce the dimensionless parameter Q, known as the quality factor, or its inverse, known as the dissipation factor, which describes energy dissipation in a material after the application of stress in cycles of frequency ω. A commonly used definition of the quality factor is:

$$Q(\omega) \equiv \frac{2E}{D} \tag{2.22}$$

where ω is the angular frequency, E is the time-averaged energy stored in the volume and D is the time-averaged dissipated energy in the cycle. For a plane wave in a visco-elastic solid in one dimension (1D), Carcione (2001) shows Q takes the following expression:

$$Q = \frac{\pi f}{\alpha v_{phase}} \tag{2.23}$$

where $f = \omega/2\pi$ is the frequency, v_{phase} is the phase velocity of the wave defined as $v_{phase} \equiv \omega/k$, being $k = \kappa - i\alpha$ the complex wavenumber and α the attenuation factor. Writing the displacement $u(x,t)$ of a plane wave in a 1D viscoelastic medium in terms of the complex wavenumber k and, in turn, in terms of the dissipation factor Q^{-1}, this yields:

$$u(x,t) = u_0 e^{i(\omega t - kx)} = u_0 e^{i(\omega t - \kappa x)} e^{-\frac{\pi f}{Q v_{phase}} x} \tag{2.24}$$

being u_0 the initial displacement. This equation describes the attenuation and dispersion introduced by Q in a 1D visco-elastic medium. In more general visco-elastic media, the dependence of Q with frequency is not always directly proportional, and depends on the constitutive relation of the mechanical model that describes that system. The simplest mechanical model is the Maxwell model, which consists of a spring and a dashpot in series (Maxwell, 1867). Carcione (2001) demonstrates that the quality factor linearly depends on the angular frequency in this model, $Q(\omega) = \omega\tau$, being τ the relaxation time – defined as the ratio between the viscosity of the dashpot and the elasticity constant of the spring. Another commonly used model to describe visco-elasticity is the Kelvin-Voigt model (Kelvin, 1875), which consists of a spring and a dashpot connected in parallel, which results in a quality factor inversely proportional to the frequency, $Q(\omega) = (\omega\tau)^{-1}$ (Carcione, 2001). A more general an accurate representation of the visco-elastic behaviour of rocks is given by the Zener model or standard linear solid (Zener, 1948), which consists of a spring in series with a Kelvin-Voigt model. In this case, the quality factor takes the form $Q = (1 + \omega^2\tau_\epsilon\tau_\sigma)/(\omega(\tau_\epsilon - \tau_\sigma))$, where τ_ϵ and τ_σ are the strain and stress relaxation times, and are related to the viscosity of the dashpot and the two elasticity constants

of the springs. Figure 2.3a depicts the dependence of Q^{-1} with frequency for a single standard linear solid, where the peak depends on the relaxation times and the viscosity.

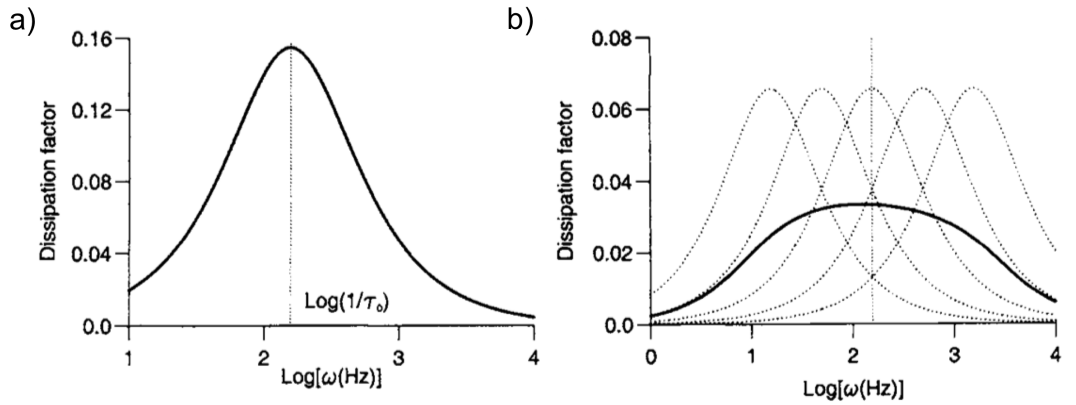


Figure 2.3.: Variation of the dissipation factor, Q^{-1}, with the logarithm of the angular frequency for a) a Zener model and b) a generalised Zener model with five dissipation mechanisms. Modified after Carcione (2001).

Nevertheless, none of the previous models accurately represents the typical measured values of Q for the subsurface. There is evidence that most rocks exhibit constant Q with frequency for a wide range of frequencies, including those typically used in exploration geophysics and seismology (McDonal et al., 1958; Kjartansson, 1979). This can be achieved by combining several standard linear solids in parallel (Liu et al., 1976; Carcione, 2001), resulting in a quality factor that is nearly-constant for a wide frequency band (Carcione, 2001, equations 2.177 and 2.178), as shown in Figure 2.3b for five relaxation mechanisms such that they give a constant value of $Q = 30$. This description is still used to model viscous effects in the time-domain, but requires the solution of a time-convolution equation that requires the previous history of stress or strain. To circumvent this problem, Carcione et al. (1988a) and Carcione et al. (1988b) introduced the concept of memory variables for visco-acoustic and visco-elastic media, which replace the convolution in time and lead to a coupled first-order system of equations of motion. Other models exist to obtain nearly constant Q with frequency, such as the Kjartansson's constant-Q model (Kjartansson, 1979). The latter uses relaxation functions of the form $t^{-2\gamma}$, where $\gamma << 1$ and t is time, it is mathematically simpler and it is often used in frequency domain applica-

tions as it requires the solution of fractional derivatives (Carcione, 2001), but it can also be solved in the time domain (Carcione, 2009; Zhu and Harris, 2013).

In certain situations, it is also relevant to consider Q can be anisotropic, and numerical methods exist in the literature to model wave propagation in these media (e.g. Zhu (2017)). However, in the remaining of this thesis I consider Q to be isotropic and I focus on wave propagation in visco-acoustic media, in which $\sigma_{ij} = -p\delta_{ij} = \Psi_{ijkl} * \frac{\partial \epsilon_{kl}}{\partial t}$. The wave equation for these media is obtained after taking the divergence of equation (2.1) and substituting the constituitive relation for visco-acoustic media in the latter, which can be solved with memory variables in the time domain (Carcione et al., 1988a). A more detailed derivation can be found in Carcione (2001) and references therein.

2.2. The forward problem

Prediction of data for a given model of the subsurface plays a key role in FWI. This is achieved by solving the forward problem that maps model parameters that characterise the subsurface, and which belong to the model space $\mathcal{M}$, into predicted measurements, which belong into the data space $\mathcal{D}$. This operation is mathematically represented as (Tarantola, 2005; Dmowska, 2010):

$$\mathcal{G}: \quad \mathcal{M} \rightarrow \mathcal{D} \tag{2.25}$$

and can also be written as:

$$\mathbf{G}(\mathbf{m}) = \mathbf{d} \tag{2.26}$$

where $\mathbf{m}$ is a vector of model parameters that describes the subsurface (for instance, P- and S-wave velocity and density), $\mathbf{d}$ is a vector that contains the predicted wavefield and $\mathbf{G}$ is a non-linear forward operator that describes how to calculate the data from the model parameters. In the acoustic case, the operator $\mathbf{G}$ is a computer code that implements the acoustic wave equation to find the pressure wavefield everywhere in the subsurface model given the models of P-wave velocity, density and the source

wavelet (Tarantola, 1984, equation 2).

There are several methods to solve the forward problem in equation (2.26) numerically given a wave equation, such as the finite-differences (FD) method (Kelly et al., 1976), finite-element methods (Marfurt, 1984) and pseudospectral methods (Danecek and Seriani, 2008), amongst others (Virieux and Operto, 2009). Here I briefly only discuss the FD method as it is the one used in the inversion and modelling code (namely *fullwave3D*) utilised throughout this thesis to solve the forward problem.

2.2.1. Finite differences (FD)

One of the most common approaches for solving the acoustic and elastic wave equations is the finite-differences scheme. The goal of FD is to solve PDEs in a discretised grid, which is achieved by substituting partial derivatives in a discretised model by a set of FD approximations derived from Taylor expansions of continuous functions around neighbouring points in the lattice. The first step of FD is to discretise the space and time domains in a lattice of points, and then find the solution to the PDE that approximates the solution at the points in the lattice. There are three main types of calculating partial derivatives with FD (Ames, 2014): forward, backward and central FD. Central FD are commonly used to solve the acoustic and elastic isotropic wave equations due to the higher accuracy of this method (Levander, 1988; Graves, 1996). The central FD approximations of the first and second space partial derivatives of a one-dimensional function $u(x)$ are given by (Courant et al., 1967; Smith, 1985):

$$\frac{\partial u(x)}{\partial x} \approx \frac{u(x+h) - u(x-h)}{2h} \tag{2.27}$$

$$\frac{\partial^2 u(x)}{\partial x^2} \approx \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} \tag{2.28}$$

with an error of order h^2, being h the separation between grid points, i.e. the grid spacing. Equivalent expressions can be obtained for the approximations of the partial time derivatives. A common choice to solve the acoustic and elastic wave equations for seismic applications is to approximate time and space derivatives so that they are

accurate to fourth-order in space and second order in time (Levander, 1988; Mufti, 1990; Graves, 1996). The fourth-order accurate approximation of the second-order space derivative for the one-dimensional function $u(x)$ is (Mufti, 1990):

$$\frac{\partial^2 u(x)}{\partial x^2} \approx \frac{-u(x-2h) + 16u(x-h) - 30u(x) + 16u(x+h) - u(x+2h)}{12h^2} \tag{2.29}$$

The latter, together with the second-order FD approximation of the second-order time derivative (from equation (2.28)), can be used to solve the acoustic wave equation in (2.19) for a body source $\boldsymbol{s} = \boldsymbol{f}/\lambda = s(t)\delta(x-x_s)\delta(y-y_s)\delta(z-z_s)$ that is zero everywhere except for the grid point where the source is located, at (x_s, y_s, z_s). Given that $\nabla p = -\lambda\nabla^2\boldsymbol{u}$ in an isotropic acoustic medium, being $\boldsymbol{u}$ the three-dimensional particle displacement, the acoustic wave equation is:

$$\nabla^2\boldsymbol{u} = v_p^{-2}\frac{\partial^2\boldsymbol{u}}{\partial t^2} + s(t)\delta(x-x_s)\delta(y-y_s)\delta(z-z_s) \tag{2.30}$$

And the FD solution fourth-order accurate in space and second-order accurate in time is (Mufti, 1990):

$$\begin{aligned} u^{n+1}(x,y,z) =& g(x,y,z)[u^n(x-2h,y,z) + u^n(x+2h,y,z) + u^n(x,y-2h,z) + \\ & u^n(x,y+2h,z) + u^n(x,y,z-2h) + u^n(x,y,z+2h) - \\ & 16(u^n(x-h,y,z) + u(x+h,y,z) + u^n(x,y-h,z) + \\ & u^n(x,y+h,z) + u^n(x,y,z-h) + u^n(x,y,z+h)) + \\ & 90u^n(x,y,z)] + 2u^n(x,y,z) - u^{n-1}(x,y,z) \end{aligned} \tag{2.31}$$

where the source term has been dropped for simplicty, $u(x,y,z)$ is the particle displacement for all points in the model, superindices indicate the time step, being n the current time t and $n+1$ or $n-1$ the future or past times, $t+\Delta t$ and $t-\Delta t$, respectively, and $g(x,y,z) = (v_p(x,y,z)\Delta t/h)^2/12$. Hence, the solution for the particle displacement at a future time step depends only on the current and previous

solutions for the displacement at the current and neighbouring grid points (this is a time-stepping solution, in which space and time are decoupled). In all the tests implemented in this thesis, finite-differences are used to solve any wave equation. In particular, the isotropic elastic wave equation is solved following the derivation of Graves (1996) in a staggered-grid in terms of velocity and stress, which is fourth-order accurate in space, second-order accurate in time and can be easily implemented on parallel computers (Graves, 1996; Guasch, 2012).

2.2.2. Grid dispersion and numerical stability

Due to the approximate nature of the solution of a PDE with finite-differences in a discretised model, the time sampling and grid spacing need to meet certain conditions in order to have numerical stability and little grid dispersion. These conditions depend on the accuracy of the FD operators used, and whether the problem is 1D, 2D or 3D. The first condition is related to the maximum grid spacing in all directions, $dx_{max} = dy_{max} = dz_{max}$, so that there is little dispersion of energy, which must obey:

$$dx_{max} = \frac{\lambda_{min}}{w} = \frac{v_{min}}{w f_{max}} \tag{2.32}$$

where v_{min} is the minimum velocity in the model – this would correspond to the minimum non-zero S-wave velocity in the presence of elastic effects–, f_{max} is the maximum frequency to be modelled or present in the source and w is a factor that depends on the accuracy of the FD operators. If the FD approximations are second-order accurate both in space and time, $w = 10$ (Mufti, 1990), whereas $w = 5$ for FD operators fourth-order accurate in space and second-order accurate in time (Alford et al., 1974; Levander, 1988; Graves, 1996). In other words, each wavelength must be crossed by $w = 5$ grid points in a fourth-order FD spatial approximation to avoid energy dispersion.

The second condition, also known as the Courant-Friedrichs-Lewy (CFL) condition (Courant et al., 1967, originally written in 1928), is related to the time sampling, and

it ensures there is numerical stability. This can be expressed as (Mufti, 1990):

$$dt_{max} = \frac{\mu dx_{max}}{v_{max}} \tag{2.33}$$

where v_{max} is the maximum velocity in the model and μ is a factor that depends on the dimensionality of the wave equation and the accuracy of the FD operators used (Lines et al., 1999). Using fourth-order accurate FD in space, $\mu = \sqrt{3/8}$ if the problem is in 2D (Lines et al., 1999), whereas $\mu = 1/2$ for 3D problems (Mufti, 1990; Graves, 1996; Lines et al., 1999).

The code I use throughout this thesis to model wave propagation, i.e. *fullwave3D*, has two different modelling kernels with different levels of accuracy, namely the low and the high kernels. The low kernel can model both 2D and 3D wave propagation, it can handle acoustic and elastic effects as well as VTI and TTI anisotropy and visco-acoustic effects, and is fourth-order accurate in space and second-order in time. Hence, the code is accurate at five gridpoints per wavelength and above (Warner et al., 2013) and less than 50 % of each cell must be crossed in a time step (i.e. $w = 5$ and $\mu = 0.5$). On the other hand, the high kernel can solve 3D acoustic and anisotropic VTI and TTI problems only, and is tenth-order accurate in space and can be either second-order or fourth-order accurate in time. The less accurate code in time from the high kernel is accurate at 3.6 wavelengths per grid point and above and less than 38 % of each cell must be crossed in a time step (i.e. $w = 3.6$ and $\mu = 0.38$), whereas the higher order type is accurate at 2.61 wavelengths per grid point and less than 66 % of each cell must be crossed in a time step (i.e. $w = 2.61$ and $\mu = 0.66$) (Silverton, 2015).

2.2.3. Boundary conditions, modelling domain and source injection

One of the problems of modelling wave propagation in a finite and discretised model is the presence of reflections generated at the boundaries of the model. A straightforward solution is to extend the model at the boundaries so that these are effectively further away, but this increases the computation time significantly (Cerjan et al., 1985). Alternatively, Cerjan et al. (1985) suggested a method in which the ampli-

tudes of the wavefield are gradually reduced within a few additional nodes along the boundaries of the model by multiplying the amplitudes of the wavefields by a negative exponential function. The exponential (or damping) function takes a value of 1 just at the boundary and it decreases within the padded area. Nevertheless, this method still requires padding the model with at least 20 cells per side, which increases the computational burden. Other methods to mitigate artificial reflections at boundaries exist, such as those based on the paraxial approximation of the wave equation (Clayton and Engquist, 1977), those based on projection operators and potentials (Lindman, 1975) or the well-known perfectly matched layer (PML) method (Berenger, 1994). In particular, *fullwave3D* mostly uses quadratic approximations of the wave equation at FD stencil locations just outside the domain, thus allowing full wave propagation up to the boundary (Debens, 2015). However, absorbing boundaries when solving the elastic wave equation are treated using the PML method (Guasch, 2012).

In most geophysical applications it is also crucial to include a free-surface boundary condition, i.e. an interface between a fluid or a solid and a vacuum. In marine settings, this is the condition that the wavefield needs to satisfy in the separation between air and water, whereas for land datasets this condition needs to be satisfied at the interface land-air. Theoretically, this is implemented by setting the traction component of the stress at the interface to zero (Chapman, 2004), which requires careful implementation in FD schemes to achieve numerical stability and good accuracy. Levander (1988) and Graves (1996) showed that a free-surface can be implemented in FD methods by setting the components of the stress $\sigma_{zz} = \sigma_{xz} = \sigma_{yz}$ to zero at the surface ($z = 0$) (z being the vertical axis, positive downwards), and assuming appropriate symmetry for the stress components around the free-surface (Graves, 1996, equation 15).

The solution of the forward problem with appropriate boundary conditions is solved throughout this manuscript in the time domain, as in the FD formulation of Graves (1996). Nevertheless, frequency-domain implementations also exist and are well-suited for problems that require the solution of time integrals, especially for modelling viscous effects (Stekl and Pratt, 1998). The implementation in both domains

is a capability of the *fullwave3D* code (Umpleby et al., 2010).

Another consideration in solving the forward problem is the injection of sources into the modelling. In the staggered-grid implementation (Graves, 1996), this can be achieved by either using the stress (equivalently, pressure in an acoustic medium) or the velocity components. In both cases, the source is represented as a discretised time series with a time sampling equal to that used for modelling (equation (2.33)), and with a maximum frequency in its spectrum that is lower than the frequency at which there is numerical dispersion (from equation (2.32)). The latter is to comply with the conditions of stability and dispersion of FD schemes. Further details of the injection of the source for FD modelling can be found in Guasch (2012).

2.3. The inverse problem: full-waveform inversion (FWI)

Full-waveform inversion is a model-building method that was originally formulated with the aim of imaging the subsurface from reflection seismic data by solving a non-linear inverse problem, in which a model of the subsurface that explains the observed data is sought. FWI has successfully been used in the last two decades to recover physical properties of the subsurface, locate and image natural resources and better understand the structure of the earth's crust, amongst other applications. To do so, FWI uses the full content of the recorded data as a function of time, as well as modelling of wave propagation through the subsurface with as much physics as possible. The recovered models have better resolution than those obtained with ray-based methods as FWI uses a more complete physics of finite-frequency wave propagation, but it also requires larger computational resources. Due to advances in computing and more efficient inversion algorithms proposed in the last decade, FWI is nowadays standard practice in the oil industry and for some academic datasets, also in 3D.

In this section I first review the historical development of FWI as well as some key features, I then derive and review the basic equations that leads to the recovery of high resolution models with FWI and I finally discuss what approximations are done

for its practical implementation.

2.3.1. Evolution and key features of FWI

Over three decades ago, Lailly (1983) and Tarantola (1984) formulated FWI as a local optimisation problem, in which the least-squares misfit between observed and modelled data, also known as the functional, is to be minimised. The non-linear inverse problem for seismic data is solved in FWI in an iterative fashion, and the full content of the observed data – unlike migration, in which only primary reflections are used – as well as full-wave modelling with accurate physics of wave propagation is performed at each iteration (Virieux and Operto, 2009). The aim is to obtain a model of the subsurface that predicts the observed data in the field. In the first iteration, FWI searches for an updated model of the subsurface in the vicinity of a starting model, which is generally smooth and simple. To achieve this goal, FWI first requires forward modelling from each source in the model to obtain a predicted (or modelled) dataset, construction of a data misfit (or data residual) by subtracting modelled and observed data trace-by-trace and sample-by-sample, back-propagation of the data residuals for each receiver and cross-correlation of the forward and backward propagated wavefields for each point in the model to create the gradient at the source, which leads to an updated model (Tarantola, 1984). The latter is then used as a starting model for the second iteration, and the process is repeated until a certain number of iterations or a certain data misfit criterion is reached.

Although FWI was initially proposed to be used with seismic reflection data only, many studies have demonstrated that long offsets or low frequencies are necessary to accurately reconstruct the intermediate and large wavelengths of the model (Mora, 1987; Pratt and Worthington, 1990; Pratt, 1990; Pratt et al., 1996), as well as wide-angle data (Virieux and Operto, 2009). Recently, several strategies have been proposed to perform reflection FWI that have shown some success (e.g. Hicks and Pratt (2001); Xu et al. (2013); Brossier et al. (2015); Yao and Wu (2017)), also to characterise the near-surface in elastic media (Provenzano et al., 2017), but this is not yet standard practice.

The presence of low frequencies also helps mitigate the non-linearity of the inverse problem. The latter arises from the fact that seismic data are oscillatory, hence the misfit function that is minimised in FWI is also oscillatory and its minimisation can lead to local minima that are not the global minimum of the inverse problem (Guasch and Warner, 2014; Warner and Guasch, 2014). Consequently, the updated model can be no better or even worse than the starting model (Warner et al., 2013). The solution of the inverse problem also relies on the Born approximation, that is the assumption that the updated model can be described as the starting model plus a small model perturbation and, hence, a linearisation of the non-linear misfit is possible. This requires that the starting model is close enough to the true model such that the modelled and observed data differ by less than half a cycle (Beydoun and Tarantola, 1988). Otherwise, the inversion can get stuck in a local minimum and lead to a spurious recovered model: this is known as cycle-skipping. Provifrf low-frequency data are available, recording of wide-angle long-offset data can also help mitigate the non-linearity and cycle skipping. Alternatively, several strategies have been proposed that help avoid cycle skipping, such as multi-scale strategies to progressively include high frequencies, layer-stripping methods to control the offsets at each step in the inversion (Virieux and Operto, 2009) or alternative data functionals (e.g. the adaptive waveform inversion (AWI) method of Guasch and Warner (2014) and Warner and Guasch (2014)).

Some of these strategies have produced acceptable recovered models from starting models that are far from the true one, as shown in Figure 2.4. In this test, the true model in Figure 2.4a is used to generate observed synthetic data using a source with dominant frequency at $10\,Hz$, which is then inverted using the simplistic 1D starting model in Figure 2.4b both with FWI and AWI, leading to the recovered models in Figures 2.4c and 2.4d, respectively. The conventional FWI functional leads to a poor-quality recovered model due to cycle-skipping effects, given that the starting model is not close enough to the true model, whereas AWI is able to overcome cycle-skipping and push the inversion towards the global minimum, leading to a high-quality high-fidelity recovered model.

Since its derivation, FWI has been implemented both in the time (Tarantola, 1984,

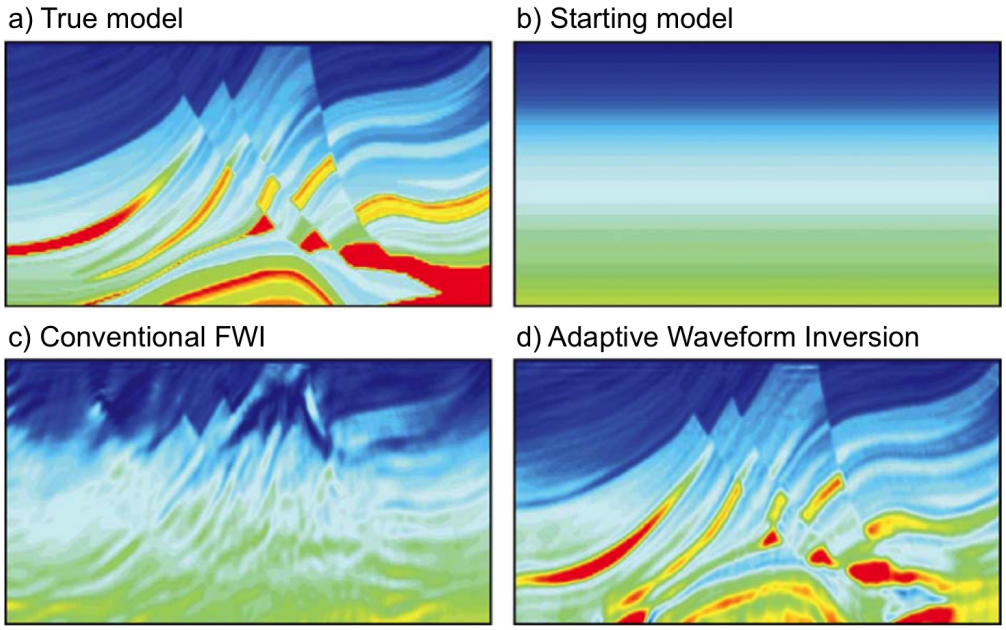


Figure 2.4.: Comparison between the a) true, b) starting, c) inverted with FWI and d) inverted with AWI P-wave velocity models for the Marmousi2 dataset. Modified after Warner and Guasch (2014).

1986; Mora, 1987) and in the frequency (Pratt and Worthington, 1990; Pratt, 1990; Pratt et al., 1996, 1998) domains. The choice of performing FWI in one domain or the other mainly depends on the ease of implementing multi-scale strategies that mitigate the non-linearity of the inverse problem, the capacity of time-windowing the data to exclude certain arrivals and the efficiency in modelling wave propagation in 3D (Virieux and Operto, 2009). Frequency-domain methods are well-suited for implementing multi-scale approaches by progressively inverting low to high-frequencies, and are also adequeate when inverting for attenuation due to the ease of modelling viscous effects in this domain. On the contrary, time-domain methods are better suited for time-windowing undesired arrivals, and are also more efficient in 3D (Virieux and Operto, 2009; Warner et al., 2013). Moreover, Bunks et al. (1995) showed a multi-scale approach can also be implemented in the time-domain by selecting subsets of data of increasing frequency by filtering the data with low-pass filters of different maximum frequency.

Due to the large computational cost of FWI either in the time or frequency domains,

most applications were initially restricted to 2D problems. Bunks et al. (1995) showed one of the first successful FWI applications to 2D synthetic data by using a time-domain multi-scale approach, whereas Shipp and Singh (2002) presented the results of 2D FWI on surface seismic data efficiently obtained in the time-domain. Some years later, Warner et al. (2008) showed the first applications of 3D FWI, whereas Sirgue et al. (2010) presented spectacular results obtained from a 3D field datset. These were possible due to more efficient modelling algorithms and advances in technology and parallel computing. Since then, 3D FWI has produced excellent recovered models (e.g. Warner et al. (2013) performed anisotropic 3D FWI on a field dataset acquired in the North Sea over a gas cloud). At the same time, Morgan et al. (2013) also demonstrated this technology was ready to be applied to more academic datasets, for instance to study the Earth's crust. The improved efficiency of inversion codes and technological advances have also led to other applications of FWI, such as to characterise the shallow subsurface by performing FWI of ultra-high-frequency ($0.2-4.0\,kHz$) reflection seismic data (Provenzano et al., 2017), which results in recovered models that agree with core-log data.

Moreover, Routh et al. (2017) recently published very high-quality FWI recovered models of a 3D field dataset, in which FWI was performed up to high frequencies ($\sim 40\,Hz$). Despite the large associated computational cost, the recovered model shows a high level of detail when compared to that obtained with ray-based methods (see Figure 2.5).

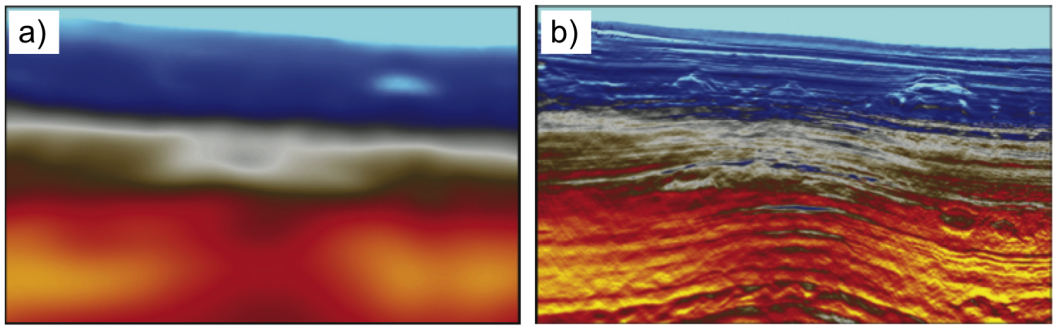


Figure 2.5.: Models of a) velocity and b) acoustic impedance obtained after 3D travel-time tomography and 3D FWI at $\sim 40\,Hz$, respectively, for a field dataset in the Western Black Sea. The model in a) is used as a starting model for FWI. Modified after Routh et al. (2017).

In fact, the main benefit of FWI with respect to ray-based methods is its enhanced resolution. Unlike ray-based methods, which have a resolution that is limited by the size of the first Fresnel zone (Williamson, 1991) – that is of the order of the square root of the product of the dominant wavelength and the length of the ray path or source-receiver offset– the resolution of FWI is half the dominant wavelength (Virieux and Operto, 2009). This is a consequence of honouring the finite-frequency of wave propagation in FWI and including as much physics as possible in the process.

Nevertheless, it is common practice to use the acoustic wave equation and only model the propagation of P-waves (Morgan et al., 2013), with or without anisotropy (Warner et al., 2013), and neglect elastic or viscous effects. In such case, P-wave velocity is the only parameter inverted for. Consideration of more realistic effects and inversion for S-wave velocity, anisotropic parameters and quality factor should, in principle, lead to improved quantitative models of the subsurface (Warner et al., 2012). However, this is not yet possible due to several challenges: 1) multi-parameter inversion increases the ill-posedness of the inverse problem as more degrees of freedom are introduced in the parametrisation, and the different sensitivity of the inversion to different parameter classes (Virieux and Operto, 2009), 2) coupling between different parameter classes as a function of aperture angle can lead to parameter cross-talk effects (Virieux and Operto, 2009), i.e. a parameter class cannot be independently inverted for without a leakage on the updated model for another parameter class, 3) the high computational cost of modelling elastic effects (Guasch, 2012; Hobro et al., 2014) and the increase in computation time when solving the visco-elastic wave equation in the time domain (Blanch et al., 1995) – as it requires several relaxation mechanisms – and 4) the difficulty in modelling the amplitudes correctly due to the usual poorly known subsurface density, attenuation and S-wave velocity as well as innacuracies in the source wavelet (Tarantola, 1986).

In the last two decades, various researchers have proposed different strategies to face these challenges. These include adding a correction term to the acoustic wave equation to account for some elastic effects (Chapman et al., 2010; Hobro et al., 2014; Chapman et al., 2014), inverting for the Lamé parameters rather than P-wave and

S-wave velocities (Shi et al., 2007), hierarchical strategies by inverting for a single parameter class first (e.g. P-wave velocity) while keeping the other fixed and then keeping the first fixed and inverting for another parameter class (e.g. S-wave velocity) (Tarantola, 1986; Sears et al., 2008), compensating the gradient for viscous effects (Xue et al., 2016), combining tomography for attenuation after having performed FWI for P-wave velocity only (Zhou et al., 2013) or a hybrid local-global approach to invert for anisotropic parameters (Debens et al., 2015a,b; Debens, 2015) or attenuation (da Silva et al., 2017). But the application of these strategies to field data in an efficient and automatic manner is still a matter of research.

2.3.2. Theoretical framework

Here I derive the basic equations used in FWI to solve the non-linear inverse problem locally in the time-domain, which are based on the original derivation of Lailly (1983) and Tarantola (1984). I use the notation and follow the derivation presented by Tarantola (2005), Virieux and Operto (2009) and Asnaashari et al. (2013), and references therein.

I first start by recalling that the forward problem can be expressed by $\mathbf{G}(\mathbf{m}) = \mathbf{d}$ (see equation (2.26)). FWI tries to solve the non-linear inverse problem, which is mathematically formulated as:

$$\mathbf{G}^{-1}(\mathbf{d}_{obs}) = \mathbf{m}' \tag{2.34}$$

where $\mathbf{m}'$ is the model we are trying to recover that explains the observed data vector $\mathbf{d}_{obs}$. Substitution of $\mathbf{m}'$ into the forward modelling expression should yield the dataset $\mathbf{d}$. However, the non-linear inverse problem has no analytical solution as $\mathbf{G}$ has no analytical inverse operator (Warner et al., 2013). Instead, FWI seeks an approximate and local solution to the linearised inverse problem, which is improved by iterating. First, the inverse problem is typically formulated as a least-squares optimisation problem of the form (the $\mathcal{L}_2$-norm is used for simplicity of the mathematics):

$$\phi_d(\boldsymbol{m}) = \sum_{N_s} \|\boldsymbol{W}_d \delta \boldsymbol{d}(\boldsymbol{m})\|_2^2 = \sum_{N_s} \delta \boldsymbol{d}(\boldsymbol{m})^T \boldsymbol{W}_d^T \boldsymbol{W}_d \delta \boldsymbol{d}(\boldsymbol{m}) \tag{2.35}$$

where $\phi_d(\boldsymbol{m})$ is a scalar quantity known as the data objective function or the data functional, which needs to be minimised, $\sum_{N_s}$ represents the sum over all shots, $\delta\boldsymbol{d} = \boldsymbol{d}_{obs} - \boldsymbol{d}_{pred}(\boldsymbol{m})$ is the misfit between the observed and predicted (or modelled) data for a given model $\boldsymbol{m}$ and $\boldsymbol{W}_d$ is a weighting operator that contains information on data uncertainties – if the traces are uncorrelated and assuming a Gaussian distribution, $\boldsymbol{W}_d$ is a diagonal matrix that contains the standard deviation of the data (Tarantola, 2005). The predicted data vector depends non-linearly on the N_m model parameters $\boldsymbol{m}$ according to the forward operator. The aim of FWI is to find a model $\boldsymbol{m}$ that minimises the misfit between observed and modelled data. For this purpose, the functional is linearised around a reference model, $\boldsymbol{m}_0$, also known as the starting model. This linearisation is known as the Born approximation. At the first iteration, $\boldsymbol{m}_0$ is the best guess of the true model. In the framework of the Born approximation, the updated or recovered model, $\boldsymbol{m}$, can be expressed as the reference model plus a small perturbation, $\Delta\boldsymbol{m}$, i.e. $\boldsymbol{m} = \boldsymbol{m}_0 + \Delta\boldsymbol{m}$. However, minimisation of the linearised data functional around the starting model leads to an ill-posed problem and, hence, it may not have a solution or it may have infinite solutions, i.e. a number of infinite models can provide a good data fit (Virieux and Operto, 2009). To reduce the ill-posedness of the inverse problem, prior model information and regularisation terms are added to the optimisation problem in equation (2.35), which leads to the more generic functional or misfit function (Asnaashari et al., 2013):

$$\phi(\boldsymbol{m}) = \phi_d(\boldsymbol{m}) + \lambda_1\phi_{m,1}(\boldsymbol{m}) + \lambda_2\phi_{m,2}(\boldsymbol{m}) \tag{2.36}$$

where λ_1 and λ_2 are regularisation parameters and $\phi_{m,1}(\boldsymbol{m})$ and $\phi_{m,2}(\boldsymbol{m})$ are penalty functions that penalise roughness in the model and variations from prior model information, respectively. The first penalty term, also known as the Tikhonov term, can be written as:

$$\phi_{m,1}(\boldsymbol{m}) = \|\boldsymbol{B}_x\boldsymbol{m}\|_2^2 + \|\boldsymbol{B}_z\boldsymbol{m}\|_2^2 = \frac{1}{2}\boldsymbol{m}^T\boldsymbol{D}\boldsymbol{m} \tag{2.37}$$

In this expression, $\boldsymbol{B}_x$ and $\boldsymbol{B}_z$ are first-order spatial derivative operators and $\boldsymbol{D}$ is the second-order Laplacian operator. Large variations of the model parameters in

any direction correspond to large first-order derivatives, which are penalised by these terms. The second penalty term can be computed as:

$$\phi_{m,2}(\boldsymbol{m}) = \|\boldsymbol{W}_m(\boldsymbol{m} - \boldsymbol{m}_p)\|_2^2 = \frac{1}{2}(\boldsymbol{m} - \boldsymbol{m}_p)^T \boldsymbol{W}_m^T \boldsymbol{W}_m(\boldsymbol{m} - \boldsymbol{m}_p) \tag{2.38}$$

where $\boldsymbol{W}_m$ is a weighting operator that contains information on the uncertainties of the prior model, $\boldsymbol{m}_p$. The latter is related to the previous iteration and can be set to be the prior model – for instance, the starting model – or be adapted during the inversion.

The solution of the inverse problem in equation (2.36) is sought in the vicinity of the starting model. Linearisation of the misfit functional with respect to the model parameters around the starting model at second-order yields:

$$\phi(\boldsymbol{m}) = \phi(\boldsymbol{m}_0) + \sum_{j=1}^{N_M} \frac{\partial \phi(\boldsymbol{m}_0)}{\partial m_j} \Delta m_j + \sum_{j=1}^{N_M} \sum_{k=1}^{N_M} \frac{\partial^2 \phi(\boldsymbol{m}_0)}{\partial m_j \partial m_k} \Delta m_j \Delta m_k + \mathcal{O}(\boldsymbol{m}^3) \tag{2.39}$$

Minimisation of the linearised misfit with respect to the model parameters takes place when the first derivative of the misfit function vanishes. Hence, taking derivatives in the previous equation and setting the right-hand-side to zero results in the following expression:

$$\Delta \boldsymbol{m} = -\left(\frac{\partial^2 \phi(\boldsymbol{m}_0)}{\partial \boldsymbol{m}^2}\right)^{-1} \frac{\partial \phi(\boldsymbol{m}_0)}{\partial \boldsymbol{m}} = -\boldsymbol{H}_0^{-1} \nabla_m \phi(\boldsymbol{m}_0) \tag{2.40}$$

where $\nabla_m \phi(\boldsymbol{m}_0)$ and $\boldsymbol{H}_0$ are the gradient and the Hessian of the misfit function evaluated at the starting model, respectively. Thus, the model update is searched in the direction of steepest descent (opposite to the gradient). After some algebraic calculations, the gradient (Asnaashari et al., 2013) and the Hessian can be respectively written as:

$$\nabla_m \phi(\boldsymbol{m}_0) = -\sum_{N_s} \boldsymbol{J}_0^T \boldsymbol{W}_d^T \boldsymbol{W}_d \delta \boldsymbol{d}_0 + \lambda_1 \boldsymbol{D} \boldsymbol{m}_0 + \lambda_2 \boldsymbol{W}_m^T \boldsymbol{W}_m(\boldsymbol{m}_0 - \boldsymbol{m}_p) \tag{2.41}$$

$$\boldsymbol{H}_0 = \sum_{N_s} \left(\boldsymbol{J}_0^T \boldsymbol{W}_d^T \boldsymbol{W}_d \boldsymbol{J}_0 - \frac{\partial \boldsymbol{J}_0^T}{\partial \boldsymbol{m}} \boldsymbol{W}_d^T \boldsymbol{W}_d \delta \boldsymbol{d}_0 \right) + \lambda_1 \boldsymbol{D} + \lambda_2 \boldsymbol{W}_m^T \boldsymbol{W}_m \tag{2.42}$$

where $J = \partial \boldsymbol{d}(\boldsymbol{m}) / \partial \boldsymbol{m}$ is the sensititvity matrix, formed by the Fréchet derivatives of the modelled data with respect to the model parameters. The method of solving the normal equations in equation (2.40) with the previous expressions for the gradient and the Hessian is known as the Newton method. The second term in the Hessian vanishes for linear problems, and it is usually neglected in FWI for solving non-linear inverse problems due to the large memory requirements for storing this term and the difficulty in computing the inverse of the Hessian matrix. Such type of approximation is known as the Gauss-Newton method (Virieux and Operto, 2009). Typically, only the diagonal terms of the approximate Hessian are computed. Neglecting regularisation terms and uncertainties, the diagonal terms of the approximate Hessian, $H_a = \boldsymbol{J}_0^T \boldsymbol{J}_0$, represent the zero-lag autocorrelation of the partial-derivative wavefields (Shin et al., 2001). An alternative method to compute the Hessian is the steepest-descend or the gradient method, which consists in approximating the inverse of the Hessian by a scalar α, also known as step-length. To improve the convergence of steepest-descend methods, a preconditioning term that consists of the diagonal elements of the approximate gradient is used (Shin et al., 2001; Virieux and Operto, 2009; Warner et al., 2013). Thus, each component of the model update in equation (2.40) is computed as follows (Warner et al., 2013):

$$\Delta m_i = -\frac{\alpha}{h_{ii}} \frac{\partial \phi(m_{0i})}{\partial m_i} \tag{2.43}$$

where h_{ii} is the approximate relative magnitude of the diagonal of the Hessian that corresponds to the model parameter m_i. The step-length can be estimated by a line-search method (Pratt et al., 1998; Virieux and Operto, 2009), which consists in perturbing the starting model a certain amount in the right direction and observe the effect this has on the residual wavefield.

On the other hand, the gradient can be efficiently computed with an adjoint-state

formulation without explicitly computing the Fréchet derivatives (Plessix, 2006). This requires forward modelling for each source in the starting model to obtain a predicted dataset, back-propagation of the residual wavefield obtained from the data residuals, and cross-correlation of the forward and backward propagated wavefields in time for each point in the model to form the gradient at each source, followed by stacking the resulting individual gradients for all sources (Tarantola, 1984; Virieux and Operto, 2009; Warner et al., 2013). This reduces the number of forward problems to solve for each source from N_M to just two (Virieux and Operto, 2009).

Nevertheless, the updated model in equations (2.40) or (2.43) after one iteration does not lead to the global minimum of the misifit function due to the non-linearity of the modelled data and the model parameters. Hence, the model is updated iteratively so that the starting model for each iteration is replaced by the result of the previous iteration, until convergence to the global minimum or a certain number of iterations are reached (Virieux and Operto, 2009; Warner et al., 2013).

Computation of the gradient via the adjoint-state method and estimation of the approximate Hessian in steepest-descend methods requires an efficient forward modelling algorithm. In practice, the acoustic wave equation is used to simplify both the cost of froward modelling and inversion algorithms, and neglect elastic or viscous effects. Although this reduces the computation time, it has also consequences on the recovered model updates. To mitigate these adverse effects, Shen (2010) proposed an alternative functional that enhances the importance of the phase of recorded data with respect to their amplitude by weighting both the observed and modelled data by their RMS energy trace-by-trace (Shen, 2010, equation 2). The solution of the inverse problem using this functional leads to similar expressions for the gradient and the Hessian to those obtained with conventional FWI, and can be derived as per above. Thus, the associated cost of solving this functional is similar to the the cost of conventional FWI, while still mitigating some of these adverse effects. The effectiveness of this method on mitigating elastic and visco-acoustic effects will be tested in the following two chapters, and will also be combined with the method I propose to address such effects. Other functionals can be suggested and implemented

to improve the quality of the recovered models, avoid cycle-skipping, use reflection data or perform multi-parameter inversion. Nevertheless, throughout this thesis I will use conventional FWI or the amplitude-balanced functional of Shen (2010).

2.4. Travel-time tomography

Travel-time tomography (TT) – also known as time-of-flight (TOF) within the medical imaging community– is an imaging technique that aims at providing reconstructed models by comparing measured and modelled travel-times of certain arrivals (Williamson, 1991), rather than the entire waveform as in FWI. This method is based on the ray approximation, which is valid at very high frequencies – or equivalently, for small wavelengths with respect to the size of the anomalies to be imaged. In the context of ray theory, the travel-time associated with a ray is given by the integrated slowness along the ray-path. This integral, known as the Fermat integral, can be expressed as follows (Bording et al., 1987; Williamson, 1991):

$$t_i = \int_{l_i} s(x,y,z)dl, \quad i = 1, ..., M \tag{2.44}$$

where t_i is the travel-time associated with the $i-th$ path from a source to a receiver and $s(x,y,z) = 1/v(x,y,z)$ is the slowness function (i.e. the inverse of the velocity) of a ray along the path l_i for a set of M ray-paths. Under the ray approximation, travel-times can be computed by solving the eikonal equation:

$$|\nabla t|^2 = \boldsymbol{s}^2 \tag{2.45}$$

with the additional condition that the traveltime at the source location is zero. The latter eikonal equation is only valid in the very high-frequency limit and for acoustic or elastic heterogeneous media (Cerveny, 2001), and can be solved with numerical methods (e.g. the finite-differences extrapolation method of Vidale (1990)). This is equivalent to solving the wave equation in FWI.

However, the difficulty in solving the integral equation (2.44) arises from the fact that the ray-path depends on the unknown slowness itself (Bording et al., 1987). Thus,

this equation is often linearised in TT around an initial slowness model (similar to the linearisation done in FWI). The problem is then to find the unknown slowness perturbations with respect to the starting model. After discretisation of the problem both in time and space, equation (2.44) is approximated by (Bording et al., 1987):

$$\Delta \boldsymbol{t} = \boldsymbol{D} \Delta \boldsymbol{s} \tag{2.46}$$

in which $\Delta \boldsymbol{t}$ is a column vector with the differences between the modelled and measured times (with dimension M), $\Delta \boldsymbol{s}$ is a vector of slowness differences between the initial model and the solution for each cell and $\boldsymbol{D}$ is a rectangular matrix whose element D_{ij} is the distance the $i-th$ ray travels through the $j-th$ cell. As each ray only covers a small portion of the discretised model, matrix $\boldsymbol{D}$ is typically sparse. A typical TT workflow (Bording et al., 1987) starts by obtaining arrival times from recorded data in the field (for instance, by picking first arrivals) and computing the difference of the latter with modelled travel-times obtained using the starting slowness model, for instance, by solving the eikonal equation (2.45) – or by ray-tracing methods from equation (2.44) (Langan et al., 1985; Stork, 1988; Ecoublet et al., 2002). Then, the distances that form matrix $\boldsymbol{D}$ are computed either assuming straight rays or using Snell's law for the starting model, and the system of equations in (2.46) is solved to find $\Delta \boldsymbol{s}$. Several methods to solve this system of equations exist: these include row-action methods that deal with individual rays (rows in $\boldsymbol{D}$) separately (Lines and Newrick, 2004) or the more efficient preconditioned conjugate gradients (Scales, 1987), which benefit from the sparsity of matrix $\boldsymbol{D}$. The process can then be repeated in an iterative fashion by setting the starting model for the next iteration to be the updated slowness model of the previous iteration until a certain stopping criterion is met.

Because TT is based on ray theory but emitted signals have a finite frequency, the resolution of the recovered models is limited by the size of the first Fresnel zone (Williamson, 1991): that is $0.5\sqrt{\pi L \lambda}$, with L being the distance that the wave has propagated and λ the dominant wavelength. Hence, it leads to recovered models with

worse resolution than those obtained with FWI – as the resolution of FWI is half the dominant wavelength (Virieux and Operto, 2009) – but the computation resources used in terms of memory and time are much lower. This makes TT a suitable imaging tool for obtaining approximate starting models for FWI.

In the remaining of this thesis I focus on wave-based methods, but I make use of starting models obtained from TT when inverting field data or *in vivo* data. This reduces the non-linearity of the inversion and produces high quality recovered models (e.g. see chapter 5).

2.5. Acoustic FWI in a visco-elastic earth

Despite the visco-elastic nature of the subsurface, viscous and elastic effects are often ignored in conventional FWI, and the acoustic wave equation is used to model wave propagation. Using approximate physics of wave propagation introduces inaccuracies in the recovered P-wave velocity models.

In this section I propose a method to mitigate viscous and elastic effects when performing acoustic FWI, which is based on Wiener (or matching) filters. Thus, I first introduce the concept and theory of Wiener filters and I explain their utility in matching a dataset to another. I then propose a workflow designed to address either viscous or elastic effects in acoustic FWI at a relatively low computational cost. The validity of this hypothesis and the potential of this method are demonstrated in chapters 3 and 4. In these chapters I also discuss the impact that the matching parameters have on the resulting dataset, as well as how to minimise their footprint.

2.5.1. Wiener filters

A Wiener filter is an optimal filter – in the least-squares sense– that transforms an input signal $\boldsymbol{y}$ into a target output signal $\boldsymbol{x}$ (Wiener, 1949). The filter output, denoted as $\boldsymbol{x'}$, is computed by linear time-invariant filtering of the observed input signal. Mathematically, this is expressed as the following convolution (Vaseghi, 2000):

$$\hat{x}(t) = \boldsymbol{y} * \boldsymbol{w} = \sum_{k=0}^{P-1} y(t-k)w(k), \quad t = 0, ..., N+P-2 \tag{2.47}$$

where $\hat{x}(t)$ is a time sample of the estimated output signal, $\boldsymbol{w}$ is the Wiener filter of length P and coefficients $w(k)$ and N is the length of the input and output signals before zero padding. In the previous equation, $\boldsymbol{y}$ is padded with $P-1$ zeros after the last sample to obtain a centred zero-padded vector $\boldsymbol{x}'$ that contains $(P-1)/2$ zeros before the first sample and also after the last sample.

The output signal in equation (2.47), $\hat{\boldsymbol{x}}$, is the estimate of the desired output signal $\boldsymbol{x}$ that minimises the following least-squares error function:

$$\phi = \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|_2^2 = \frac{1}{2} \sum_{t=0}^{N+P-2} \left(x(t) - \sum_{k=0}^{P-1} y(t-k) w(k) \right)^2 \tag{2.48}$$

Minimisation of the previous least-squares error function with respect to a single filter coefficient requires setting the derivative $\partial\phi/\partial w(i)$ to zero, which yields the following equation for the $i-th$ component:

$$\sum_{t=0}^{N+P-2} \left[\left(x(t) - \sum_{k=0}^{P-1} y(t-k) w(k) \right) \left(\sum_{k=0}^{P-1} y(t-k) \delta_{ki} \right) \right] = 0 \tag{2.49}$$

in which $\delta_{ki} = 1$ if $k = i$ and $\delta_{ki} = 0$ otherwise. Minimisation with respect to all coefficients results in the following system of normal equations (Robinson and Durrani, 1986):

$$\sum_{t=0}^{N+P-2} x(t) y(t-i) = \sum_{k=0}^{P-1} w(k) \sum_{t=0}^{N+P-2} y(t-k) y(t-i), \quad i = 0, ..., P-1 \tag{2.50}$$

The latter can be re-written as:

$$c_{xy}(i) = \sum_{k=0}^{P-1} w(k) R_{yy}(i-k), \quad i = 0, ..., P-1 \tag{2.51}$$

which is known as the Wiener-Hopf equation with:

$$c_{xy}(i) = \sum_{t=0}^{N+P-2} x(t) y(t-i), \quad i = 0, ..., P-1 \tag{2.52}$$

and:

$$R_{yy}(i-k) = \sum_{t=0}^{N+P-2} y(t-k)y(t-i), \quad i = 0, ..., P-1 \tag{2.53}$$

where c_{xy} is the cross-correlation of the target signal and the input signal (left-hand-side in equation (2.50)) and $R_{yy}(i)$ is the auto-correlation of the input signal. Thus, the right-hand-side in equations (2.50) and (2.51) corresponds to the convolution of the Wiener filter with the auto-correlation of the input data. Equation (2.51) can easily be expressed in matrix form as:

$$\boldsymbol{c}_{xy} = \boldsymbol{R}_{yy}\boldsymbol{w} \tag{2.54}$$

in which $\boldsymbol{w}$ and $\boldsymbol{c}_{xy}$ are column vectors of dimension P with the filters' coefficients and the cross-correlations of the input and the target signals, respectively, and:

$$\boldsymbol{R}_{yy} = \begin{pmatrix} R_{yy}(0) & R_{yy}(1) & R_{yy}(2) & \dots & \dots & R_{yy}(P-1) \\ R_{yy}(1) & R_{yy}(0) & R_{yy}(1) & R_{yy}(2) & \dots & R_{yy}(P-2) \\ R_{yy}(2) & R_{yy}(1) & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & R_{yy}(-1) & R_{yy}(-2) \\ R_{yy}(P-2) & \dots & \ddots & R_{yy}(1) & R_{yy}(0) & R_{yy}(1) \\ R_{yy}(P-1) & \dots & \dots & R_{yy}(2) & R_{yy}(1) & R_{yy}(0) \end{pmatrix} \tag{2.55}$$

is a square matrix of dimensions $P \times P$ that contains the auto-correlations of the input data. In the previous expression I have taken into account that the autocorrelation of two stationary time series is an even function, i.e. $R_{yy}(i) = R_{yy}(-i)$, which follows from the fact that $c_{xy}(i) = c_{yx}(-i)$ (Robinson and Treitel, 1980, chapter 6). The autocorrelation matrix displays a particular symmetry: the values along the diagonals are constant, i.e. each element in the diagonal descending from left to right is constant. Due to this symmetry, the matrix is said to be a Toeplitz matrix (Robinson and Treitel, 1980; Claerbout, 1985; Robinson and Durrani, 1986; Golub and Van Loan, 2013), which is a particular type of persymmetric matrix, and the system of equations in equation (2.54) is a Toeplitz system.

The coefficients of the Wiener filters are found by solving equation (2.54). Thus, $\boldsymbol{w} = \boldsymbol{R}_{yy}^{-1}\boldsymbol{c}_{xy}$, being $\boldsymbol{R}_{yy}^{-1}$ the inverse of the auto-correlation matrix. Several strategies exist to efficiently invert this autocorrelation matrix by exploiting the symmetry properties of a Toeplitz matrix (Golub and Van Loan, 2013). These are mainly classified in three categories depending on the number of arithmetic operations: 1) classical methods, 2) fast methods and 3) super-fast methods.

The first category, classical methods, are robust but relatively expensive to implement as these require $\mathcal{O}(P^3)$ operations (Musicus, 1988). The most popular methods in this category include QR decomposition, Gaussian elimination and Cholesky decomposition, from which QR decomposition is generally preferred due to its mathematical simplicity and its robustness (Bojanczyk et al., 1995; Vaseghi, 2000). The latter consists in decomposing the auto-correlation matrix in equation (2.55) into a product of an orthogonal matrix and an upper-triangular matrix to compute the Wiener filters' coefficients by back-substitution (Francis, 1961, 1962; Kublanovskaya, 1962). The computation burden of this method mainly lies on the decomposition step of the Toeplitz matrix, which can be performed for instance by using the Gram-Scmidt process. Nevertheless, fast implementations of the QR algorithm that require $\mathcal{O}(P^2)$ operations also exist (e.g. Bojanczyk et al. (1995)), but these are not unconditionally stable (Brent, 1988).

On the other hand, fast methods require $\mathcal{O}(P^2)$ operations and, therefore, are more affordable. The most frequently used fast method is the Levinson's algorithm (Levinson, 1949) after modifications from the original derivation to improve its stability (Bunch, 1985). The latter consists in a recursive procedure to solve a $M+1$-dimensional Toeplitz system with $M = 0, 1, ..., P-1$ (Morettin, 1984; Bunch, 1985). This method starts by first solving for the first coefficient of the equation $\boldsymbol{R}_{yy}\boldsymbol{w} = \boldsymbol{c}_{xy}$ as follows (O'Dowd, 1991):

$$R_{yy}(0)w(0) = c_{xy}(0) \tag{2.56}$$

followed by solving:

$$\begin{pmatrix} R_{yy}(0) & R_{yy}(1) \\ R_{yy}(1) & R_{yy}(0) \end{pmatrix} \begin{pmatrix} w(0) \\ w(1) \end{pmatrix} = \begin{pmatrix} c_{xy}(0) \\ c_{xy}(1) \end{pmatrix} \tag{2.57}$$

and M linear systems of equations are progressively solved to find all coefficients in a similar way. For more details on the implementation of a stable Levinson's algorithm I refer to Press et al. (2007): I have implemented the algorithm described in this reference to compute Wiener filters using the Levinson's recursion method.

The third category of methods for solving Toeplitz systems, namely super-fast methods (Bunch, 1985; de Hoog, 1987; Turnes et al., 2014), require $\mathcal{O}(Plog^2P)$ operations, which is significantly less than those required by fast methods. This can be achieved by using fast Fourier transforms (FFTs) (Ammar and Gragg, 1988; Brent, 1988) and by transforming Toeplitz system into a special type of polynomial problems (Turnes et al., 2014). However, super-fast algorithms were typically considered unstable (Bunch, 1985). Thus, various techniques and formulations have been proposed during the last decades to stabilise these methods (Bunch, 1985; Brent, 1988; Lakshmikantham et al., 1994; Bojanczyk et al., 1995; Higham, 2002; Turnes et al., 2014).

To avoid stability problems, super-fast methods are not further explored in the remaining of this thesis. Instead, I have developed a parallel Fortran code that can be used in a high performance computing (HPC) facility, which uses the Levinson's recursion algorithm as derived by Press et al. (2007) to efficiently compute the coefficients of the Wiener filters when matching one seismic trace to another. For further details on this implementation I refer to the next section and chapters 3 and 4.

Figure 2.6 shows the potential of using Wiener filters to match two distinct stationary time-series of length N, shown in Figure 2.6a, which have different phase and amplitude content. The first signal, $\boldsymbol{x}$, is a periodical function with respect to the number of samples, whereas the second signal is also a periodic function but with smaller time period (i.e. higher frequencies) and with amplitudes that are exponen-

tially damped with increasing number of samples. Next, I compute Wiener filters that match signal $\boldsymbol{y}$ into signal $\boldsymbol{x}$ by solving equation (2.54) and using both the classical QR method and the Levinson's method to invert the auto-correlation matrix. The resulting filters, computed with a filter length of $P = 2N = 200$ samples for both methods, are nearly identical, and are overlaid in Figure 2.6b. These filters are then convolved with the input signal, $\boldsymbol{y}$, to produce the estimate $\hat{\boldsymbol{x}} = \boldsymbol{y} * \boldsymbol{w}$, which is an optimal estimate of the desired signal $\boldsymbol{x}$ in the least-squares sense. This is shown in Figure 2.6c, in which the desired signal and the estimated signals with both methods, $\hat{\boldsymbol{x}}_{QR}$ and $\hat{\boldsymbol{x}}_{Levinson}$ respectively for the QR and Levinson's methods, are plotted. In the absence of noise, the computed Wiener filters lead to very good approximates of the desired signal with both methods. However, there is a noticeable difference in terms of compute time, as depicted in Figure 2.6d for different lengths of the filters for both methods. This plot shows that the computation time of the Wiener filters increases with the length of the filter, and it increases more rapidly for the QR method given that the number of required operations is $\mathcal{O}(P^3)$ rather than $\mathcal{O}(P^2)$. Thus, the Levinson's method is used from here onwards.

The test in Figure 2.6 also demonstrates the potential of Wiener filters. In fact, Wiener filters have been used in geophysics for multiple purposes: these include noise supression (del Negro, 1996; Karsl and Bayrak, 2004), wavelet shaping (Robinson and Treitel, 1980; Yilmaz, 1987), seismic deconvolution (Robinson and Treitel, 1980; Yilmaz, 1987) and also to avoid cycle skipping in FWI (Warner and Guasch, 2014; Guasch and Warner, 2014). Here I suggest Wiener filters can be used to mitigate elastic and viscous effects in acoustic FWI. This hypothesis is further discussed now.

2.5.2. Workflow and algorithm

The data matching application between two time traces in Figure 2.6, based on Wiener filters, can be extended to match two entire datasets, gather-by-gather and trace-by-trace. This is demonstrated in Figure 2.7, in which acoustic and elastic data have been modelled for a shot at the centre of the well-known Marmousi2 model (Bourgeois et al., 1990; Martin et al., 2006). The corresponding P-wave and S-wave velocity models are depicted in Figures 2.7a and 2.7b, respectively, whereas the acoustic and

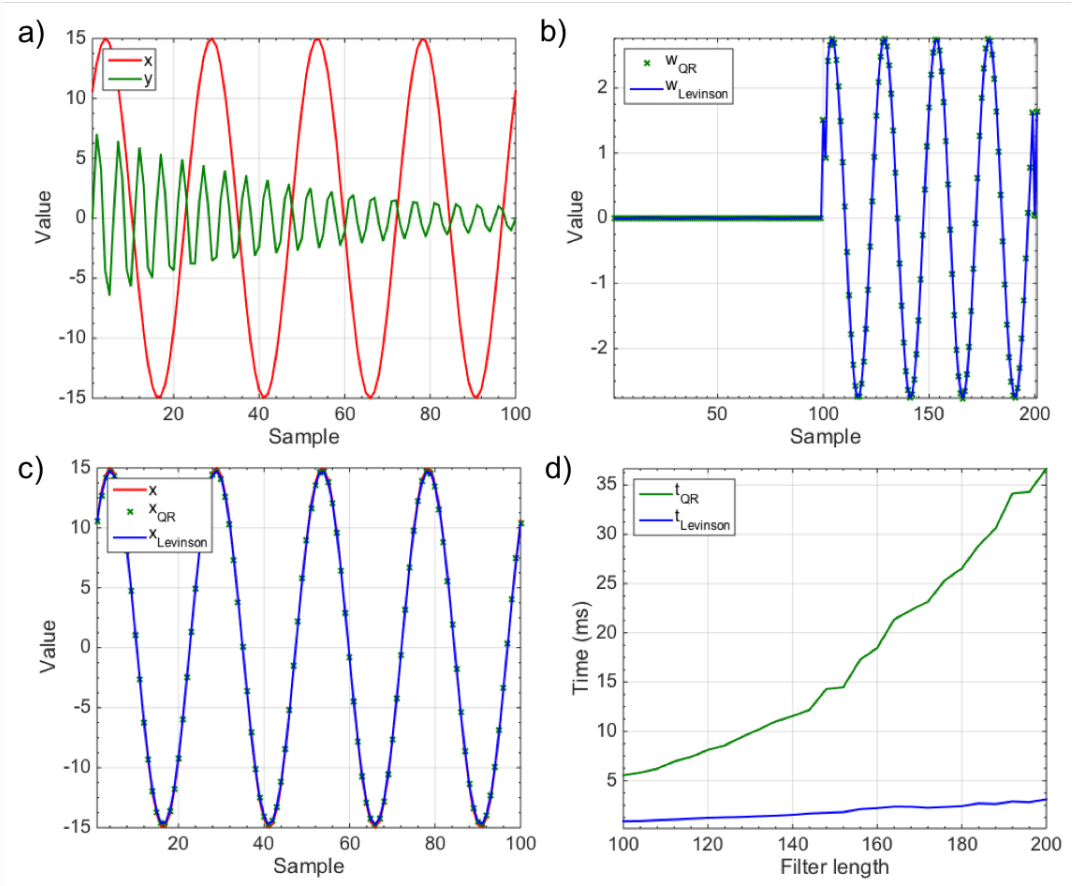


Figure 2.6.: Example of data matching of two very different time series using Wiener filters. Panel a) shows the input signal $\boldsymbol{y}$ and the target signal $\boldsymbol{x}$, both of length $N = 100$, whereas panel b) shows the Wiener filters computed with the QR and Levinson's methods with a length of the filters of $P = 2N$. The resulting estimated signals for both methods are overlaid with the target signal in panel c), and the computation times for different lengths of the filters are shown in d).

elastic modelled data at receivers spread across the model are displayed in Figures 2.7c and 2.7d. Figure 2.7e shows the trace-by-trace differences between the elastic and the acoustic traces.

Analogously to the test in Figure 2.6, I compute Wiener filters that match each trace in the elastic dataset (the input data) to each trace in the acoustic dataset (the desired dataset). I then convolve each trace in the elastic dataset with the corresponding computed Wiener filter, which leads to the filtered (or matched) dataset in Figure 2.7f: this is an estimate of the desired acoustic data. As shown in Figure 2.7g, the differences between the filtered dataset and the acoustic dataset are significantly

smaller than those between the elastic and the acoustic dataset in Figure 2.7e.

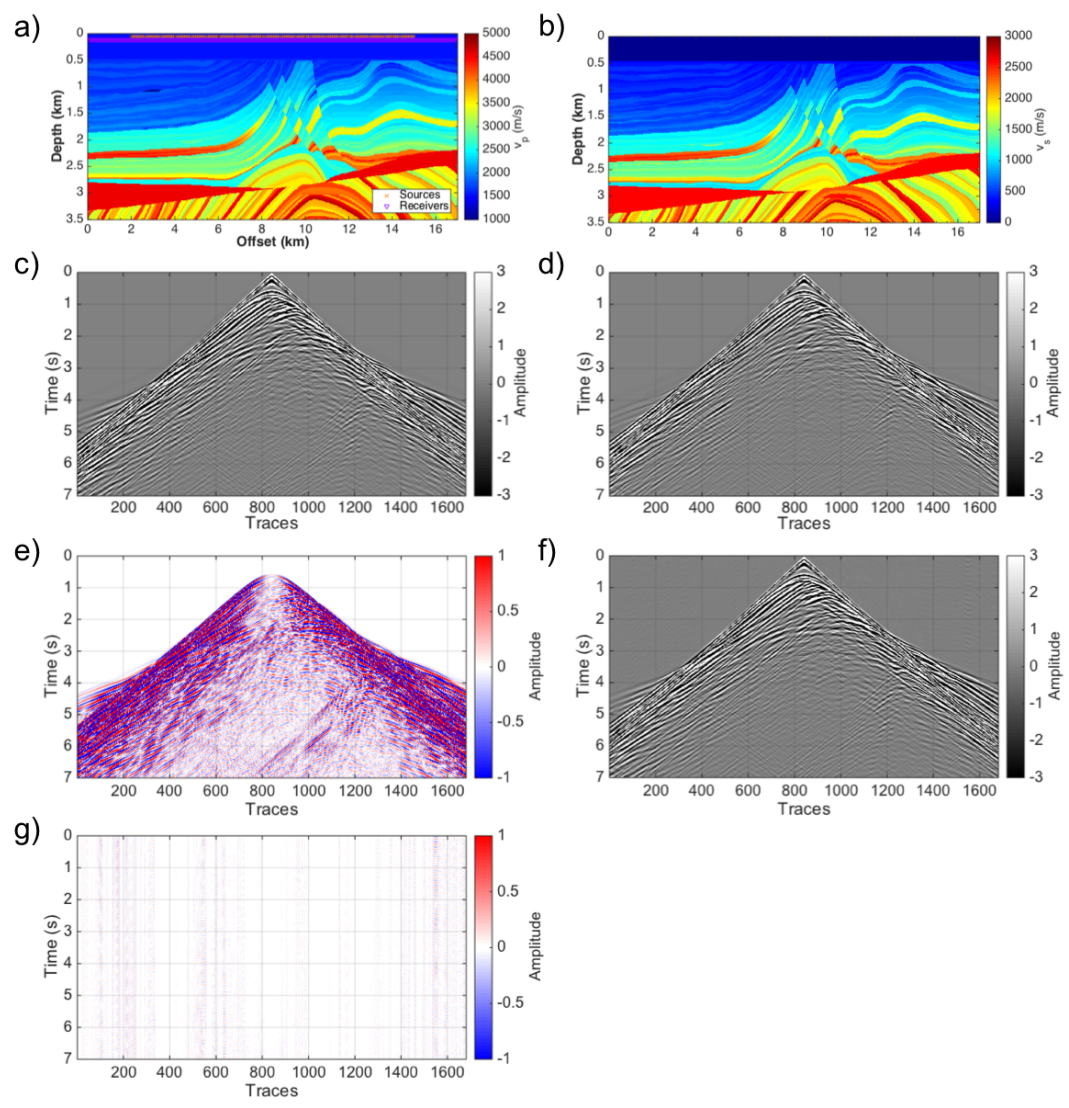


Figure 2.7.: Data matching example for the Marmousi2 model in a) for P-wave velocity and b) for S-wave velocity models, which are used to generate c) acoustic modelled data and d) elastic modelled data for a shot at the centre of the model, and their difference is shown in e). Wiener filters of a length of $P = 2N$ are then computed trace-by-trace to match the elastic to the acoustic modelled data, and are then convolved with the elastic data to mitigate elastic effects, which leads to the gather in f). The difference of the acoustic data and the data in f) is shown in g).

The previous example demonstrates that it is possible to match an elastic to an acoustic dataset obtained from a realistic subsurface model in the absence of noise and if we know what are the right coefficients of the Wiener filters. In practice, recorded field seismic data contains elastic and viscous effects (anisotropy is neglected here for

simplicity), and no prior information on accurate P-wave and S-wave velocity models or a Q model are typically available. Nevertheless, if one makes an assumption on these models it is in principle possible to model acoustic and also elastic or visco-acoustic data, compute Wiener filters that match the elastic or the visco-acoustic data to the acoustic data and filter the observed data – recorded in the field – with the computed Wiener filters to produce a filtered (or matched) dataset, in which elastic or viscous effects are removed. Similarly to the example in Figure 2.7, this should produce the best results if the observed data only contains elastic or viscous effects and no noise (in that case the observed data would be equivalent to the elastic data in the example in Figure 2.7) and also when the estimates of the P-wave and S-wave velocity models or the P-wave velocity and the Q models are close to the true models.

In case the data are contaminated with noise or/and the estimated velocity or attenuation models are not as accurate, mitigation of elastic or viscous effects might not be as effective, but it could potentially still mitigate significant elastic or viscous effects. The matched data rather than the original observed data could then be used as an input for an acoustic inversion, potentially leading to more accurate recovered P-wave velocity models as the physics used in the modelling would yield a dataset closer to the matched dataset than the original elastic or visco-acoustic observed dataset.

In order to implement this in practice I propose two workflows: the first one is designed to mitigate elastic effects (Figure 2.8), and the second is designed to mitigate visco-acoustic effects (Figure 2.9).

The aim of these two workflows is only to improve the quality of recovered P-wave velocity models from acoustic FWI of either realistic synthetic seismic data that is contaminated with elastic and viscous effects or field seismic data, and it is not to obtain models of S-wave velocity or Q models. Estimates of the latter are necessary for the application of the workflow and need to be obtained via alternative methods. This is discussed further in chapters 3 and 4. One important conclusion from several tests discussed in these chapters is that these estimates do not need to be very accurate.

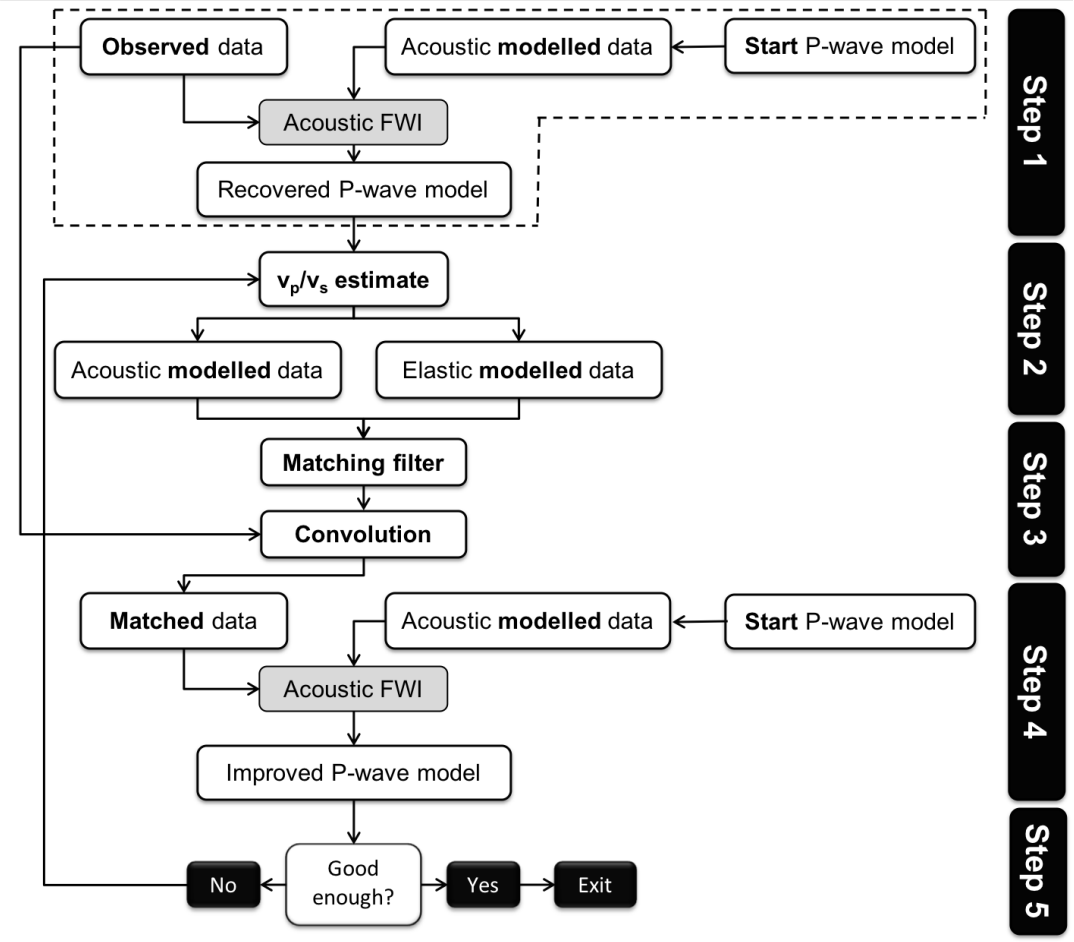


Figure 2.8.: Workflow used to mitigate elastic effects based on Wiener filters. The dotted area represents the conventional acoustic FWI workflow.

In essence, the workflows in Figures 2.8 and 2.9 are very similar. Both can be divided in 5 different steps. Step 1 consists in performing conventional acoustic FWI of the observed data by minimising the misfit between the latter and modelled data, as discussed in section 2.3.2. A starting P-wave velocity model is required at this stage, which is used to model acoustic data during the first iteration of acoustic FWI. Here there is no restriction in the type of functional used in the inversion: the aim is to obtain a first estimate of the P-wave velocity model as accurate as possible in the presence of elastic and viscous effects in the data. Thus, I recommend either the use of the classical least-squares data misfit or the use of the amplitude-balanced functional of Shen (2010), which can help mitigate some elastic and viscous effects.

The second step of the workflow, Step 2, consists in modelling data using the

recovered model in Step 1 for all shots in the model. This differs in both schemes. In case the objective is to mitigate elastic effects (Figure 2.8), acoustic and elastic data are modelled by using the recovered P-wave velocity model obtained after Step 1 and an estimate of v_p/v_s. On the other hand, if the aim is to mitigate visco-acoustic effects then acoustic and visco-acoustic data are modelled for all shots by using the recovered P-wave velocity model and an estimate of Q.

Once Step 2 is complete, the remaining steps are the same in both workflows. Next, Wiener filters that match the modelled elastic or visco-acoustic data – depending on the application– to the modelled acoustic data for all shots are computed, trace-by-trace. The computed filters are then convolved with the observed data (Step 3), which lead to a matched dataset in which elastic or viscous effects are mitigated, provided the estimates for P-wave and S-wave velocities and the Q model are accurate enough.

The matched dataset is then used as an input for acoustic FWI in Step 4. The latter is expected to lead to an improved recovered P-wave velocity model of the subsurface as the input data for the inversion contains less elastic or viscous effects and, thus, acoustic modelling during acoustic FWI is more accurate.

Finally, the workflow can be repeated a number of times in Step 5 to further improve the result by using the recovered P-wave velocity model in Step 4 as an input for acoustic and elastic or visco-acoustic modelling in Step 2. Nevertheless, each additional iteration of the workflow increases the total computational cost, especially due to the inversion in Step 4. This will be further discussed in chapters 3 and 4, in which I demonstrate that the first iteration of the workflow has the largest impact on the result and that one or two iterations provide good quality results without compromising on the computational time.

Variations of the previous workflows are indeed possible, and these might be needed under certain circumstances – for instance, in the presence of strong elastic or viscous effects or when the data is contaminated with a high level of noise. These aspects will be covered in more detail in chapters 3 and 4.

But, independently of the modifications introduced, the implementation of Step 3 is key for the good performance of the workflow. This requires an efficient and

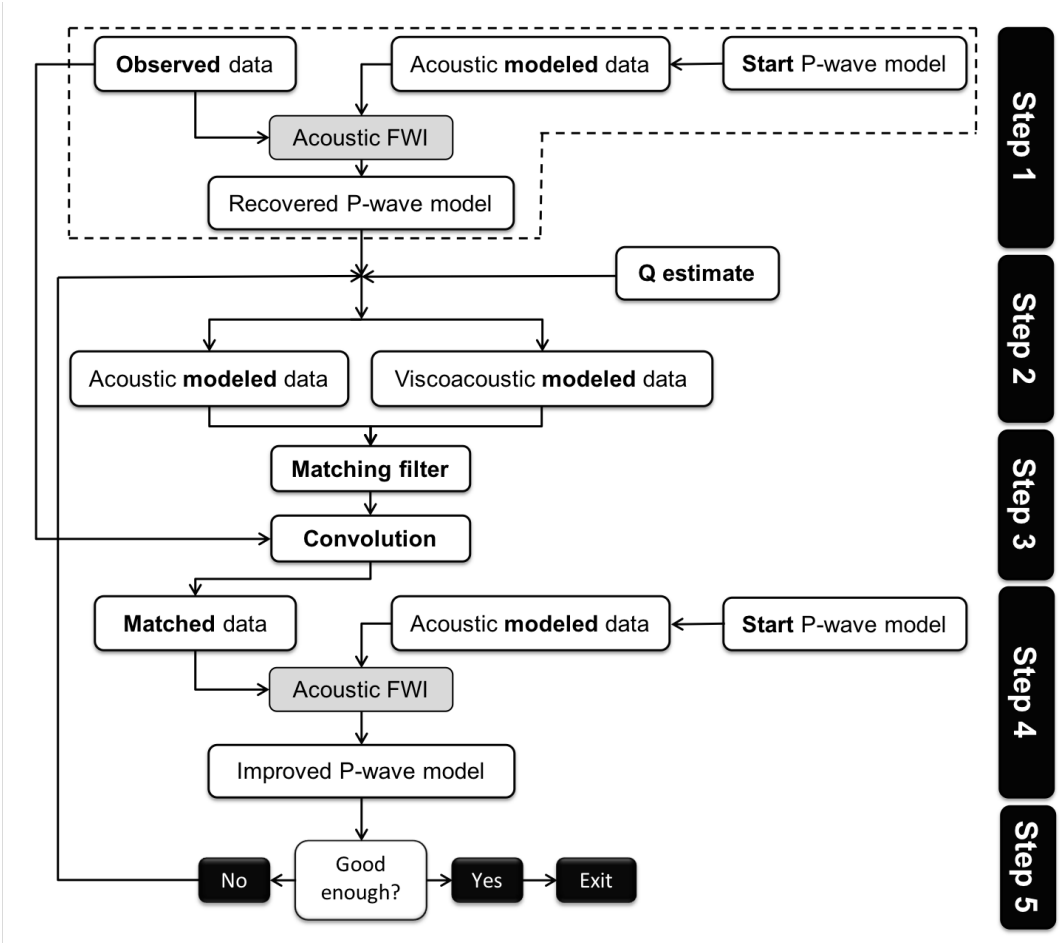


Figure 2.9.: Workflow used to mitigate visco-acoustic effects based on Wiener filters. The dotted area represents the conventional acoustic FWI workflow.

robust code that computes Wiener filters that match one dataset to another, such that the resulting matched observed dataset obtained after the convolution with the Wiener filters contains as little artefacts from filtering as possible. There are various potential sources of error when matching two different datasets. First, the Levinson's algorithm used to calculate Wiener filters might not be stable if the diagonal elements of the auto-correlation matrix are too small. Second, computing a single filter that matches a whole seismic trace to another one that may contain different events might introduce artefacts in the data if the spectra of the two traces are very different. This can be observed in the example in Figures 2.7f and 2.7g, in which each trace from the elastic data is matched to the corresponding trace in the acoustic dataset – note spectra might be different because there are S-wave related events in the elastic data

that are not present in the acoustic data and also because amplitudes are different. In fact, the estimation of a long single filter per trace also has consequences in terms of compute time, as the longer the filter the more arithmetic operations are required (recall this is proportional to P^2 for the fast Levinson's method). A third potential artefact might appear due to the fact that data matching is performed between two signals of finite length, and the latter can introduce some unwanted noise around the edges of the trace being matched. Finally, another source of error might be not accounting for the lateral continuity of the data, as seismic data presents continuity both in time and along the events. In order to tackle all these potential sources of error I introduce the following four different mechanisms:

1. Zero-padding and tapering of the data
2. Pre-whitening of the auto-correlation matrix
3. Lateral smoothing by matching multiple traces simultaneously
4. Computing short filters in time

The first strategy consists in extending each trace with zeros on each side prior to computing Wiener filters, and multiplying the padded trace with a discrete window function that smoothly increases its values from 0 on the extremes of the trace to 1 at the centre. In particular, I use a Hann (Oppenheim and Schafer, 1989; Press et al., 2007) or a Blackman window functions (Oppenheim and Schafer, 1989). This method increases the computation time slightly, as the trace is extended by a small percentage of the total number of samples, but it ensures that any edge effects are reduced and that they happen within the padded region.

The second mechanism is used to stabilise the Levinson's algorithm, and consists in adding a constant value to the diagonal elements of the auto-correlation matrix such that $R_{yy}(0)$ transforms to $(1+\epsilon)R_{yy}(0)$, being ϵ a value between 0 and 1 (Gubbins, 2004). By inspection of equation (2.56), it is clear that adding a small value to $R_{yy}(0)$ can transform a division by zero into a division by a small number, thus making the Levinson's algorithm stable. The main effect of adding this small term is to flatten the spectrum of the filters (Gubbins, 2004), thus avoiding notches. However, if the

diagonal terms are too small, adding a percentage of itself might not be enough to stabilise the problem. Thus, I compute the maximum absolute value in the auto-correlation matrix and I add a small percentage of this value to the diagonal terms.

The third mechanism consists in imposing some smoothness across neighbouring traces to have better continuity of the events and less artefacts of trace-by-trace matching. This is achieved by reformulating the least-squares problem in equation (2.48), so that the aim is to compute a single Wiener filter that matches several consecutive traces of the input data to several traces of the desired data. In order to achieve this, I propose equation (2.48) can be reformulated as follows:

$$\tilde{\phi} = \left\| \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_r \end{pmatrix} - \begin{pmatrix} \boldsymbol{y}_1 * \boldsymbol{w} \\ \boldsymbol{y}_2 * \boldsymbol{w} \\ \vdots \\ \boldsymbol{y}_r * \boldsymbol{w} \end{pmatrix} \right\|_2^2 = \frac{1}{2} \sum_{j=1}^{r} \left[\sum_{t=0}^{N+P-2} \left(x_j(t) - \sum_{k=0}^{P-1} y_j(t-k) w(k) \right)^2 \right] \tag{2.58}$$

where r is the number of traces to be matched simultaneously and the misfit contains the difference between two long vectors that respectively contain the desired traces and the convolution of the filter with the input traces, concatenated in single vectors of dimension $(N + P - 1) \times r$. Similarly to the derivation for a single trace, the coefficients of the Wiener filters that match multiple traces can be found by minimising the misfit $\tilde{\phi}$ with respect to the filter coefficients and setting that to zero, i.e. $\partial\tilde{\phi}/\partial w(i) = 0$. This leads to the following set of normal equations:

$$\sum_{j=1}^{r} \sum_{t=0}^{N+P-2} x_j(t) y_j(t-i) = \sum_{k=0}^{P-1} w(k) \sum_{j=1}^{r} \sum_{t=0}^{N+P-2} y_j(t-k) y_j(t-i), \quad i = 0, ..., P-1 \tag{2.59}$$

Which can be expressed as:

$$\tilde{\boldsymbol{c}}_{xy} = \tilde{\boldsymbol{R}}_{yy} \boldsymbol{w} \tag{2.60}$$

with:

$$\tilde{c}_{xy}(i) = \sum_{j=1}^{r} \sum_{t=0}^{N+P-2} x_j(t) y_j(t-i), \quad i = 0, ..., P-1 \tag{2.61}$$

and:

$$\tilde{R}_{yy}(i-k) = \sum_{j=1}^{r} \sum_{t=0}^{N+P-2} y_j(t-k) y_j(t-i), \quad i = 0, ..., P-1 \tag{2.62}$$

Therefore, the latter results in a Toeplitz system similar to that in equation (2.54), but now $\tilde{\boldsymbol{c}}_{xy}$ and $\tilde{\boldsymbol{R}}_{yy}$ are the sum of the cross-correlation between the input and desired signals and the auto-correlation of the input signals for multiple traces, respectively. For the particular case of just one trace, the previous equations reduce to those derived earlier for a single trace. The estimation of Wiener filters for multiple traces simultaneously has a small impact in the computation time with respect to a trace-by-trace application, as the calculation of cross-correlation and auto-correlation coefficients is the sum of more elements. However, this does not significantly affect the compute time as the length of the filter and the order of the cross-correlation vector and auto-correlation matrix are unchanged.

Finally, in the fourth mechanism I introduce some degree of smoothness in time to avoid artefacts caused by matching entire traces that have different spectra. To do so, I obtain matched data for small portions of a whole trace, defined by a certain length (i.e. number of samples), so that the matched portions of the data overlap in time. I then use a Blackman window weighting function to compute the final matched data such that the sample at the centre of the current matched data has a larger weight than the sample of the next overlapping matched data – and with the constrain that the sum of the weights must be equal or smaller than 1. This has a smoothing effect along the trace and also leads to a reduction in the compute time as this mainly depends on the squared of the length of the filters, which is reduced.

To further reduce the compute time, I start matching the data for each trace when, starting from time zero, the value of the forward acoustic modelled data is larger than a certain value. This simplifies the computations as filters are not calculated

before the first arrivals and, hence, less matching filters need to be computed. This measure also reduces stability issues when the length of the filters is short and the data to be matched contains only zeros (or very small values) before the first arrivals, in which case the elements of the auto-correlation matrix are very small and, thus, the Levinson's method may be unstable.

Both the impact of the length of the filters and the number of neighbouring traces matched simultaneously to impose smoothness have the biggest impact on the quality of the resulting matched datased. Thus, these aspects are further discussed in chapters 3 and 4.

I now briefly discuss how the data matching code is implemented in practice. Algorithm 2.1 schematically describes the simplified structure of the code I have developed in Fortran to implement the workflows in Figures 2.8 and 2.9. The output of this algorithm is the matched observed dataset, which is obtained from the input modelled data, the desired modelled data and the observed data. The implementation of the four mechanisms described above to overcome some limitations of trace-by-trace filtering is not discussed here, as it does not change the structure of the code significantly.

Algorithm 2.1 Computation of matched observed data using Wiener filters that match input traces (e.g. elastic or visco-acoustic modelled data) to desired traces (e.g. acoustic modelled data).

Input: $\boldsymbol{y}$ (input data), $\boldsymbol{x}$ (desired data), $\boldsymbol{d}_{obs}$ (observed data)
Output: $\boldsymbol{d}_{matched}$ (matched observed data)

Open and read: $\boldsymbol{y}$, $\boldsymbol{x}$, $\boldsymbol{d}_{obs}$
Read: filtering parameters
Order data
Initialise variables
$i = 1$
while $i \leq N_{traces}$ **do**
    call Wiener_filtering($\boldsymbol{y}_i$, $\boldsymbol{x}_i$, filtering parameters, $\boldsymbol{w}_i$)
    call convolution($\boldsymbol{d}_{obs,i}$, $\boldsymbol{w}_i$, $\boldsymbol{d}_{matched,i}$)
    Write/append: $\boldsymbol{d}_{matched,i}$
    i=i+1
end while
Close data files

Thus, the code starts by reading the entire datasets as well as the filtering parame-

ters: these include the length of the padded region as a percentage of the length of the data to be filtered, the length of the filters, the value of the pre-whitening constant ϵ and the number of neighbouring traces to be used to compute Wiener filters.

In the next part of the code, the different datasets are ordered with respect to shot identification number and channel number. This is important as the three datasets that have been read at the previous step do not need to be in the same order initially (for instance, if the modelled data has been generated with a parallel code and each shot is written as soon as the data is computed), and it is crucial to select the same traces when filtering. The different variables used for the code are then initialised (including the index i, which indicate trace number) and the main part of the code begins. In this part of the code Wiener filters and the corresponding matched data are computed for all traces in a loop that starts at the first trace present in the data and ends at the last trace. In more detail, the cross-correlation vector and auto-correlation matrix are first computed for a whole trace, and Wiener filters are then computed using the Levinson's method – when particularising it to part of the trace the algorithm is the same, but the length of the filters and the part of the data to be used is smaller. In this step, the data is also conveniently padded with zeros, and pre-whitening is added to the diagonal elements of the auto-correlation matrix to avoid stability issues.

The resulting Wiener filters are then convolved with the corresponding trace (or part of the trace) of the observed data, which leads to a matched trace – if using short overlapping filters in time, the matched trace is obtained by weighting the different portions of matched data along the length of the trace with Blackman window functions. This is then written to disk into a file that has the same structure as the original data files so that it can then be used by the FWI code as an input observed data. Data is appended to this file until the last trace has been matched, and all data files are properly closed.

For a faster and more efficient computation, I have implemented this algorithm in parallel using message passing interface (MPI) library routines in Fortran. This allows the implementation of the code across multiple nodes (and the use of all cores

in each node), thus speeding up the workflow significantly. This does not change the main structure of the code described in algorithm 2.1, but it does change how the data is handled. In my implementation, I use a master node that reads all the input data and parameters and then distributes part of the data and key parameters to the different slaves. Thus, the slaves do not store all the data but only receive part of the data, filter the data and then send it back to the master node, which then appends the received data to the output data file. Parallelisation of the data matching code results in a considerable reduction in the compute time, such that running it for a whole dataset typically represents about a 5 % of the total time it takes to perform acoustic FWI for the same data provided the number of computer resources is the same. This is the final version of the code that I use in chapters 3 and 4 to mitigate elastic and visco-acoustic effects, respectively.

2.6. Modelling and inversion in medical imaging

In the second part of this thesis I demonstrate the potential of FWI in improving the quality of sound-of-speed models of biological tissues, in particular for breast and brain imaging. This work is based on the use of ultrasound waves, i.e. sound waves with frequencies that range from $20\,kHz$ – beyond the human audible limit – to several gigahertz, and which can be generated in the laboratory using piezoelectric materials (Manbachi and Cobbold, 2011). Since its first practical application to detect submarines during World War I (this was proposed and tested by Paul Langevin around 1917) (Manbachi and Cobbold, 2011), ultrasound waves have been used with great success for imaging in non-destructive testing and in medicine (also for therapeutic purposes). In particular, several medical imaging modalities based on ultrasound exist nowadays: the most popular include B-mode imaging (Chan and Perlas, 2011), colour Doppler imaging (Evans et al., 2011), elastography (Gennisson et al., 2013) and opto-acoustic methods (Mertz, 2015).

All these techniques fundamentally rely on the wave nature of ultrasound. Ultrasound propagation within biological tissues can be described in most cases by the acoustic or elastic wave equation. Thus, the theory of linear wave propagation for

seismic waves described in this chapter can also be used to model the propagation of ultrasound waves. In this context, linear propagation means that the ultrasound signal is emitted at one fundamental frequency, and no other new frequencies can be created, i.e. non-existing frequencies in the original signal cannot be generated (Thomas and Rubin, 1998). Nevertheless, ultrasound propagation is only a linear process when amplitudes and frequencies are low, and non-linear effects need to be considered at high frequencies and large amplitudes (Zhao et al., 2014).

In the non-linear regime, the phase of ultrasound waves is distorted as it propagates, resulting in the generation of harmonics of the fundamental frequency transmitted by the ultrasound transducer (Pinton et al., 2009; Othman et al., 2014) – unlike in seismic imaging using pressure source, in which only the linear regime corresponding to the fundamental component is relevant as amplitudes and frequencies are not very large. The wave equation that models non-linear wave propagation mainly contains the linear wave equation and additional terms that account for the non-linearity, the thermo-viscous diffusivity, variations in density and several relaxation mechanisms (Pinton et al., 2009; Zhao et al., 2014). Non-linear ultrasound imaging techniques have seen a significant development in the last decade to improve the resolution and penetration of ultrasound scanners, both using harmonics generated by reflections of the ultrasound waves at tissues or at contrast agents (typically micro-bubbles) (Thomas and Rubin, 1998; Tranquart et al., 1999). However, the energy of the harmonic components received is always lower than that of the fundamental component, and wide-band transducers able to transmit and receive ultrasound signals at high frequencies are required.

On the other hand, ultrasound tomographic techniques that neglect non-linear effects and use transmitted and reflected energy have been used since the 1970's, in particular for breast cancer detection (Huthwaite et al., 2010). In these applications, the amplitudes and frequencies of the emitted ultrasound waves are low, and non-linearity introduced by heterogeneities in body tissue properties are often neglected. The use of transmitted energy rather than the usual reflected and back-scattered (B-mode) energy has the potential to lead to more accurate and quantitative models

of acoustic velocities (Pratt et al., 2007) – the latter also referred here and in the medical community as speed-of-sound.

Similarly to geophysics, the simplest of these methods is travel-time tomography (TT) (see section 2.4), also known as time-of-flight (TOF) in medical imaging, in which arrival times of first arrivals are extracted from transmitted arrivals, and ray-based iterative methods are used to obtain a reconstruction of velocity models (Pratt et al., 2007). Ray-based methods provide an approximate estimate of the speed-of-sound models but are unable to resolve features that are separated more than the size of the first Fresnel zone: that is typically $10 - 15\,mm$ (Huthwaite et al., 2010). Despite this limitation, TT can be used to rapidly obtain smooth models to be used as starting models by more sophisticated algorithms, or to reconstruct the approximate geometry of limbs (Fincke et al., 2016). However. it is not sufficient when imaging small soft tissue structures such as those in the breast (Fincke et al., 2016). To overcome these limitations, diffraction tomography methods were suggested, but the latter are unable to obtain large-scale variations (Huthwaite et al., 2010) and they rely on the assumption of homogeneity in the background (Pratt et al., 2007).

Alternatively, Pratt et al. (2007) applied FWI to image a breast phantom and a 2D dataset from an *in vivo* breast experiment without any assumption on the homogeneity of the background, obtaining a resolution of up to half the dominant wavelength. The applicability of FWI as developed in geophysics to solve medical imaging problems is possible as both problem types show similar scales characteristics (Pratt et al., 2007). Table 2.1 compares characteristic scales for the range of values used in chapters 3 to 6.

Table 2.1.: Characteristic scales between geophysical and medical imaging problems based on the examples presented in the following chapters.

	Geophysics	**Medical ultrasound**
Velocities	$1450 - 4500\,m/s$	$1450 - 3000\,m/s$
Useful frequencies for FWI	$2 - 30\,Hz$	$0.2 - 2\,MHz$
Wavelengths	$50 - 2000\,m$	$0.7 - 15\,mm$
Grid spacing	$5 - 50\,m$	$0.5 - 5.0\,mm$
Time sampling	$0.5 - 8\,ms$	$0.5 - 8\,\mu s$

Therefore, based on table 2.1, the cost of modelling sound wave propagation in

a 2D model of dimension $1\,km \times 1\,km$ in geophysics with a grid spacing of $10\,m$ is equivalent to modelling ultrasound wave propagation in a model of size $10\,cm \times 10\,cm$ with a grid spacing of $1\,mm$, as the total number of cells in the model is 101×101. Recent applications of FWI to medical imaging problems have shown some success in obtaining acoustic velocity models of the breast with source encoding techniques (Matthews et al., 2017), obtaining acoustic velocity and attenuation models for synthetic breast models in a time-domain formulation (Pérez-Liva et al., 2017) and even on imaging bones on synthetic data (Bernard et al., 2017).

In parallel to these recent studies, I have applied conventional acoustic FWI to demonstrate the potential of 3D FWI to image the breast and the brain, leading to high resolution models of speed-of-sound with resolution of under $1\,mm$ for soft tissue at the highest frequencies. To my knowledge, this is one of the very first applications of 3D FWI for breast imaging and the very first application in brain imaging, both in 2D and in 3D. For the implementation of FWI, I neglect non-linear ultrasound propagation effects within tissue. This can be justified as I use data acquired or generated with de-focused ultrasound transducers – in practice I consider point sources, hence low amplitude – operating at relatively low frequencies – with transmit peak frequencies always below $1\,MHz$, and a maximum recorded frequency below the second harmonic frequency of $2\,MHz$. Thus, I use the expressions for the model update obtained with the Born approximation as derived in this chapter. In addition, I explore the importance of low frequencies and the impact of ignoring attenuation and density during acoustic FWI for breast imaging in chapter 5. In chapter 6, I discuss in detail how FWI can be applied to image the brain through the skull using ultrasound, which is considered very difficult (Ylitalo et al., 1990; Smith et al., 2009) and has never been demonstrated before.

2.7. Original work

Developing a practical an efficient 3D FWI code that is able to handle (anisotropic) acoustic, elastic and visco-acoustic modelling is not a simple task. In this thesis I

use the *fullwave3D* code to perform all the inversions. The latter has been mainly developed by Adrian Umpleby (Earth Science and Engineering Department, Imperial College London) as part of the FULLWAVE consortium, with contributions from others. In particular for this thesis, I use the elastic modelling code developed by Lluis Guasch (Earth Science and Engineering Department, Imperial College London) during his PhD thesis and the forward modelling visco-acoustic kernel developed by Nuno Vieira da Silva (Earth Science and Engineering Department, Imperial College London), both implemented in the time domain into *fullwave3D* by Adrian.

The original work for the third and fourth chapters of this thesis is based on the implementation of the data matching algorithm described in this chapter to remove elastic and viscous effects from seismic data. To do so, I have developed a parallel Fortran code, which uses modelled data generated from *fullwave3D* to obtain a matched dataset with less elastic and viscous effects, which is then inverted using *fullwave3D*. This code has seen numerous changes in the last three years since I first developed it: now it is designed so that it uses the same type of input and output files as *fullwave3D* and, hence, it can work with field data in 2D and in 3D. I have also developed and implemented shorter MATLAB codes to estimate Q from seismic data, to calibrate the location of transducers for in vivo imaging of the breast or to find a good starting model to image the brain, amongst others. To test my ideas on more realistic examples, I have also had the contribution of other collaborators such as Tatiana Kalinicheva (Earth Science and Engineering Department, Imperial College London), who pre-processed and performed FWI on the field data presented in the fourth chapter, and Peter Huthwaite (Mechanical Engineering Department, Imperial College London), who provided me with a starting model for the *in vivo* breast dataset. The original work in the fifth and sixth chapters is based on the hypothesis that FWI can be applied to image not only the breast, but also the human brain.

Other contributions and publications arising from my research are further discussed at the beginning of each chapter.

3. Acoustic FWI in an elastic earth

Most of the current applications of FWI are based on the acoustic wave equation and the recovery of P-wave velocity models of the subsurface. Thus, the elastic nature of the earth is frequently neglected mainly to reduce the compute cost and run times of forward modelling elastic wave propagation, and secondarily to reduce the effect of cross-talk between the inverted parameters. In this chapter I implement the methodology proposed in chapter 2, which is based on Wiener filters, to mitigate elastic effects in acoustic FWI, which leads to better quality recovered P-wave velocity models at a much lower cost than elastic FWI.

I have presented some of the results in this chapter at the SEG International Exposition and 86th Annual Meeting in Dallas:

> Agudo, O. C., da Silva, N. V., Warner, M., and Morgan, J. (2016). Acoustic full-waveform inversion in an elastic world. In *SEG Technical Program Expanded Abstracts 2016*, volume 24, pages 1058–1062. Society of Exploration Geophysicists.

Additionally, the methodology and results included in this chapter have also been published in the form of a scientific paper:

> Agudo, O. C., da Silva, N. V., Warner, M., and Morgan, J. (2018b). Acoustic full-waveform inversion in an elastic world. *Geophysics*, 83(3):1–15.

The current chapter contains those results I have already published and further tests that complement those above. Moreover, I include a section on the preliminary application to a 3D field dataset that has not yet been published elsewhere.

3.1. Introduction

In full-waveform inversion, the non-linear misfit between observed and modelled data is minimised iteratively and, hence, wave propagation in the subsurface needs to be computed at each iteration. Ideally, the full content of the recorded waveforms is utilised in FWI, and modelling is performed as accurately as possible by accounting for the full physics of wave propagation, thus including all types of generated waves (Virieux and Operto, 2009).

Correctly addressing the kinematics and dynamics of the wavefield by accounting for the anisotropic and visco-elastic subsurface behaviour should in principle lead to high quality P-wave recovered models of the subsurface and the recovery of other parameter classes, such as S-wave velocity and attenuation (Warner et al., 2012). However, using an accurate physics of wave propagation in forward modelling largely increases the computational cost, especially due to modelling elastic waves. Moreover, inverting for multiple parameter classes also increases the ill-posedness of the inverse problem as there are more degrees of freedom in the parametrisation, and the inversion can also be sensitive to various parameter classes in different ways (Virieux and Operto, 2009).

In particular, elastic FWI is not widely used due to the high computational burden of elastic modelling (Guasch et al., 2010; Hobro et al., 2014), the non-uniqueness and coupling effects of the different physical properties of the wavefield (Virieux and Operto, 2009) and the difficulty in modelling amplitudes correctly without knowing the subsurface density, attenuation, source wavelet and S-wave velocity accurately prior to the inversion (Tarantola, 1986). Of the latter, the main reason elastic FWI is not yet possible is the high computational cost of elastic modelling, especially when inverting 3D datasets efficiently. This is due to the number of model parameters required to specify an elastic model (i.e. 21 parameters and the density when accounting for anisotropy), the large number of field variables required in elastic modelling (i.e. particle velocity and stress components), the cost of solving the elastic wave equation and the fine grid spacing required to avoid dispersion of S-waves with much lower values than P-wave velocities, especially close to the surface or sea-bottom (Hobro

et al., 2014). The last factor has the largest weight on the total computation cost of elastic modelling and, hence, elastic inversion problems.

Several strategies have been suggested to efficiently perform elastic FWI and address some of the issues preventing elastic FWI (Raknes and Arntsen, 2017). These include hierarchical strategies to invert different parameter classes successively (Sears et al., 2008; Brossier et al., 2009), the use of graphic processing units to efficiently perform elastic forward modelling (Venstad, 2016), wavefield reconstruction methods to reduce the amount of data stored (Nguyen and McMechan, 2015; Raknes and Weibull, 2016) and elastic mode decomposition to compute the gradients and reduce cross-talk effects (Wang and Cheng, 2017), amongst others.

Nevertheless, these strategies are not yet widely implemented and elastic FWI is rarely performed in order to reduce compute costs. Instead, the anisotropic acoustic wave equation is utilised for forward modelling during acoustic FWI, and only P-wave velocity models of the subsurface are recovered. Anisotropic acoustic FWI allows efficient inversion of large 3D datasets (Warner et al., 2013), and it provides good quality results as it accounts for the correct kinematics of the wavefield.

However, ignoring elastic effects has a negative impact on the quality of the recovered velocity models, especially as phase conversions at interfaces (for instance, P- to S-wave conversions in an isotropic subsurface model) are not properly accounted for. The latter also has an impact on the amplitude of all arrivals, including P-wave arrivals (Hobro et al., 2014). This is illustrated in Figure 3.1, in which the wavefield in a three-layered model is shown at three different times in an acoustic (left column) and an elastic (right) media.

In this example, the wavefield is generated by a point source located at the center of the model within the (top) water layer, in which S-wave velocity is set to zero for both media. After $1\,s$, the wavefield propagates across the water layer, is recorded at the receivers – this is the direct wavefield–, and it reaches the fluid-solid interface. At this point, part of the wavefield is reflected upwards towards the receivers, and the rest is transmitted downwards to the second layer – hence, the energy of the transmitted wave is not the same as the incident wave, as it splits in a reflected and

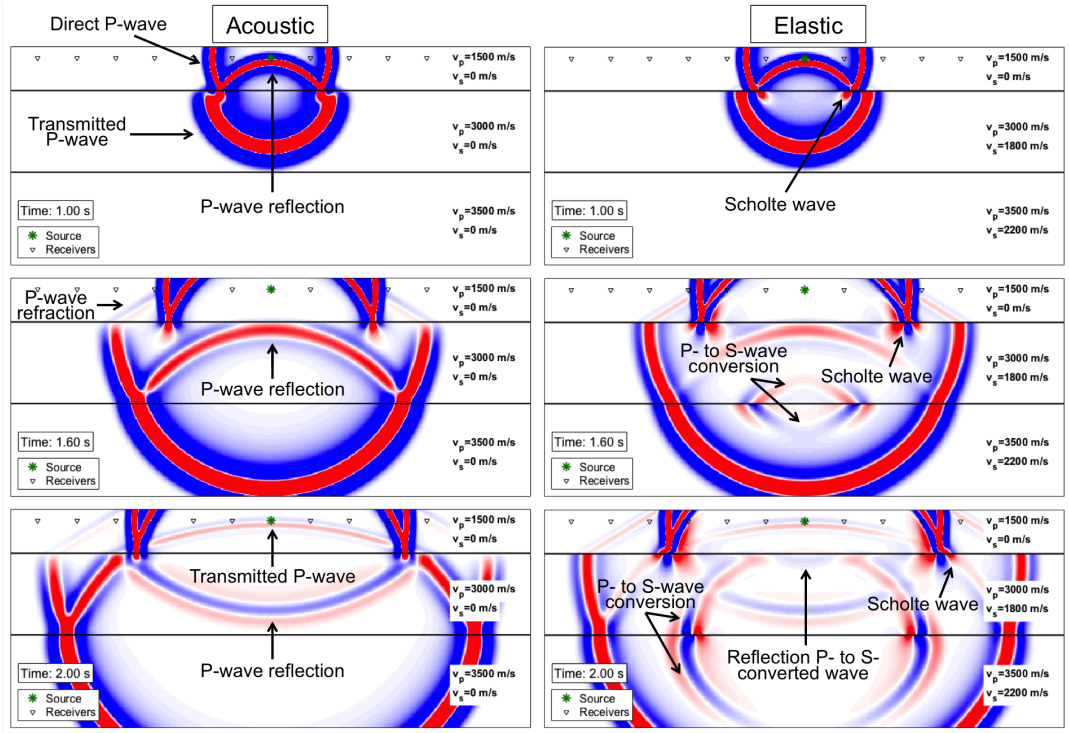


Figure 3.1.: Snapshots of the wavefield in a three-layered model generated from a source situated at the centre of the model, within the water layer, both in an acoustic (left) and an elastic (right) media at different times (top to bottom).

a transmitted waves.

In the elastic case, P-waves are generated at the source and no S-waves are created or propagate through the water layer, given that there is zero shear wave velocity. Nevertheless, surface waves are generated at the fluid-solid interface, which propagate within this region with the S-wave velocity of the second layer and have amplitudes that exponentially decrease away from the interface. These are known as Scholte waves, and are difficult to observe in a single snapshot but can be seen if data are recorded along the interface or by inspection of the time-evolution of the wavefield for multiple snapshots.

At 1.6 s (second row in Figure 3.1), a critical P-wave refraction can be easily observed within the first layer in both media, which propagates with the velocity of the second layer and reaches those receivers at an offset greater than the critical distance earlier than the direct arrival does. On the other hand, the transmitted wavefield to the second layer continues propagating and hits the second interface. At this point,

part of the P-wave energy is transmitted to the bottom layer and the rest is reflected upwards as a reflected P-wave in both media. Additionally, there is a P- to S-wave phase conversion at this interface in the elastic case, which propagates within the second layer with the S-wave velocity of the second layer, and does not exist in the acoustic medium. This splitting of P-waves also has an impact on the variation of amplitude with offset of different arrivals, including P-waves (as stated by Hobro et al. (2014)): for example, the amplitude versus offset of the reflected P-wave at the second interface travelling upwards is much smaller in the elastic case, whereas there is no visible difference in the amplitudes of the water wave (the direct arrival) and the transmitted P-wave to the second layer.

Later in time, the different waves continue propagating and they are transmitted and reflected at the different interfaces. In particular, the wavefield is overall more complicated in the elastic case as the phase-converted waves interfere with the transmitted and reflected P-waves, thus changing their amplitude behaviour too. Moreover, Scholte waves at the surface are more visible at $2\,s$ as they are ahead of the direct P-waves arrivals and have covered a longer distance with respect to the direct P-wave: this is because the S-wave velocity of the second layer is slightly above the water velocity.

Thus, the recorded data at the receivers in both media differ, which can compromise the quality of the recovered P-wave velocity models obtained with acoustic FWI as the physics of elastic wave propagation is not considered.

To mitigate the impact of elastic effects in acoustic FWI at a reduced cost, some strategies have been proposed. For instance, Chapman et al. (2010) proposes a method to correct acoustic modelling for some elastic effects by adding a correction term to the acoustic wave equation. This method can be used to correct the amplitudes of reflected P-waves at a cost higher than acoustic simulations, but smaller than the cost of elastic modelling. Although the latter can be used in FWI to partly simulate elastic wave propagation, it does not account for phase conversions and it can degrade the quality of acoustic simulations when velocity contrasts are large (Hobro et al., 2014).

On the other hand, a widely used method to mitigate elastic (and also viscous) effects in acoustic FWI is the amplitude-balancing method of Shen (2010), which consists of using an objective function that uses more of the phase information rather than the amplitude of the arrivals. This is achieved by weighting the amplitude of the observed and modelled data in the objective function by their RMS energy trace-by-trace, such that equation (2.35) becomes:

$$\phi_{d,norm}(\boldsymbol{m}) = \sum_{N_s} \left\| \boldsymbol{W}_d \left(\frac{\mathbf{d}_{obs}}{\sqrt{\mathbf{d}_{obs}^T \mathbf{d}_{obs}}} - \frac{\boldsymbol{d}_{pred}(\boldsymbol{m})}{\sqrt{\boldsymbol{d}_{pred}^T(\boldsymbol{m}) \boldsymbol{d}_{pred}(\boldsymbol{m})}} \right) \right\|_2^2 \tag{3.1}$$

This objective function is widely used in the industry as it mitigates some elastic effects while having a minor impact on the total computational cost. However, the latter is not sufficient in the presence of strong elastic effects.

Here, I implement the alternative method suggested in section 2.5, which is based on Wiener filters, to mitigate a wider range of elastic effects while still performing acoustic FWI, even in the presence of noise or strong elastic effects. The suggested method, which is applied here to pressure data but has the potential to be applied to multi-component data, requires an estimate of v_p/v_s and it is not designed to obtain S-wave velocity models of the subsurface, but to improve the quality of the recovered P-wave velocity models with acoustic FWI in the presence of elasticity. Moreover, this method allows the use of any objective functional during acoustic FWI, such as the conventional functional in equation (2.35) or the amplitude-balanced functional in equation (3.1). In particular, the combination of the current method with the amplitude-balanced functional of Shen (2010) provides the best results.

Within this chapter, I evaluate the impact of elastic effects on the recovered P-wave velocity models of acoustic FWI of synthetic data, and I implement the data-matching workflows suggested in the previous chapter to mitigate elastic effects on the latter. Finally, I demonstrate on a marine 3D field dataset that elastic effects need to be addressed for certain marine environments, as ignoring these may result in unrealistic recovered models. I then present the results and discuss the benefits of using the data-matching method to mitigate elastic effects for this particular dataset.

In more detail, I validate the current method on 2D horizontally layered media in the next section (section 3.2), I study the impact of the filtering parameters on the resulting matched data and I investigate the limitations of this method and how this can be modified to address strong elastic effects. I then use this knowledge in section 3.3 to study the impact of elasticity and mitigate its adverse effect on the recovered velocity models in a more realistic 2D marine synthetic dataset, also in the presence of strong noise, and I compare the resulting P-wave velocity models to those obtained with a fully elastic inversion. In section 3.4 I repeat the experiments on a realistic 2D land synthetic dataset, in which elastic effects are more important, and I show that the data-matching method leads to higher quality recovered P-wave velocity models. Next, I implement this strategy on a 3D field dataset and I show the dramatic effect that neglecting elastic effects has on the recovered velocity models, even for a marine dataset, whereas applying the data-matching method provides much more plausible recovered velocity models. Finally, I discuss the results obtained in this chapter and summarise the main conclusions in section 3.6.

For all inversions performed in this and the following chapters, density is computed from the true P-wave velocity model and using Gardner's law (Gardner, 1974) during the modelling, whereas it is computed from the recovered models at each iteration and Gardner's law during the inversions.

3.2. Application to 2D horizontally layered models

Initially, I test the impact of acoustically inverting synthetic seismic data obtained from 2D horizontally layered susburface models. This is similar to the study of Barnes and Charara (2009) for 1D models with different v_p/v_s profiles, but here I perform 2D inversions with strong horizontal smoothing and a much higher number of inverted parameters. Next, I apply the data-matching workflow to mitigate elastic effects by using the recovered P-wave velocity models and the true v_p/v_s, and I assess the impact of the length of the filters on the quality of the matched data and the noise-reduction introduced by matching several traces simultaneously when compared to matching them trace-by-trace. I then demonstrate in this simple model that both FWI (using

the conventional functional or the amplitude balanced functional of Shen (2010)) and the current method fail in recovering an accurate P-wave velocity model when elastic effects are strong or when the starting model is far from the true model, and I demonstrate that a modified data-matching workflow, less computationally demanding than that suggested in Figure 2.8, provides high-quality recovered P-wave velocity models in these cases.

The two reference sets of models used to generate acoustic and elastic synthetic single-component pressure data, namely true models A and B, are shown in Figure 3.2. These are respectively designed to simulate two different type of geological settings in which P- and S-wave velocities within layers are constant and there are large associated velocity jumps with respect to having smoothly varying gradients in velocity, respectively. Thus, model A is consists of two layers with constant P- and S-wave velocities and large impedance contrasts for P- and S-waves at the sea-bottom ($450\,m$) and at $640\,m$ depth. On the other hand, model B consists of four layers: two of them with gradients in P-wave velocity and the other two with constant P-wave velocities. The water layer for model B is shallower ($340\,m$ depth) than for model A, impedance contrasts are smaller and maximum P- and S-wave velocities are higher. For each model set, I generate elastic data with two different S-wave velocity profiles that lead to different degree of elastic effects. Thus, the first S-wave velocity profile for both datasets (with associated S-wave velocity v_{s1}, represented by a black line with filled triangles in Figure 3.2) leads to weaker elastic effects than those obtained with the second S-wave velocity model (with associated S-wave velocity v_{s2}, represented by a grey line with filled circles in Figure 3.2) as impedance contrasts for S-waves at interfaces are smaller.

Acoustic and elastic data are generated for both sets of models with independent finite-difference modelling codes and the same source signature and modelling parameters. First, two-dimensional horizontally layered models are generated by laterally extending the reference models in Figure 3.2, so that these are $15\,km$ long by $900\,m$ deep, and I use a grid spacing of $3\,m$ that ensures there is no numerical dispersion for both P- and S-waves (according to equation (2.32) for a 2D problem). Synthetic

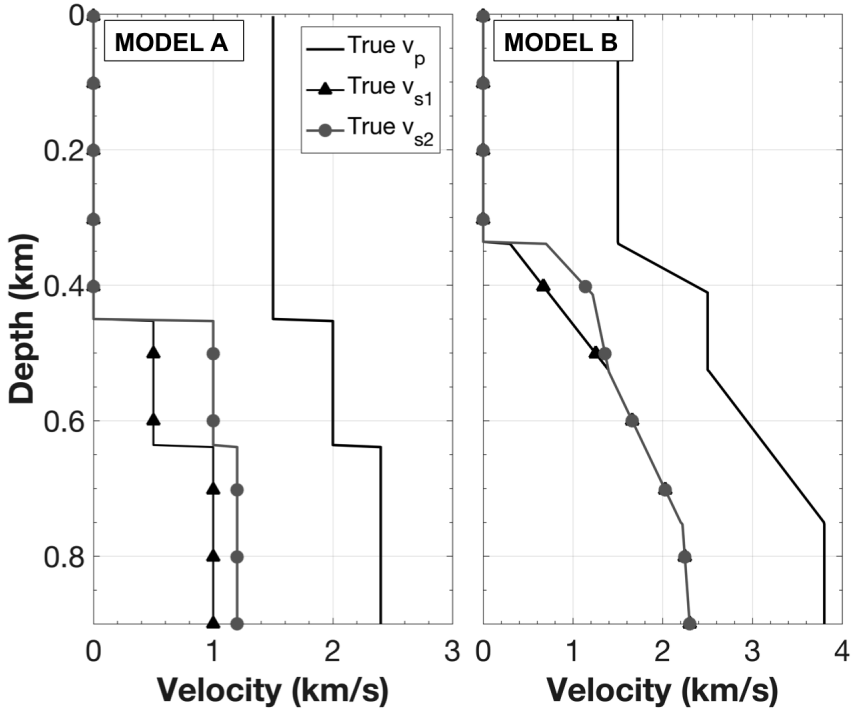


Figure 3.2.: Vertical profiles for the true P-wave and S-wave velocities for the horizontally layered 2D models A (left) and B (right).

data are generated by deploying 41 sources at $6\,m$ depth with a constant shot spacing of $300\,m$ that emit a pressure signal corresponding to a Ricker wavelet with $10\,Hz$ centre frequency. The seismic signal is then recorded by 967 hydrophones at $16\,m$ depth, with a receiver spacing of $15\,m$. The maximum offset of the recorded signal is $7.25\,km$, which is enough to record turning rays that have travelled through all the layers in the models. To simulate realistic marine seismic data, absorbing boundaries are used on the sides of the models and at the bottom, whereas I use a top free-surface to model the air-water interface. Hence, source and receiver ghosts and surface-related multiples are present in the data.

Once synthetic acoustic and elastic data are generated I perform acoustic FWI of the synthetic data for models A and B by using the same inversion parameters. The inversion is performed using the conventional objective functional, and a multi-scale approach such that data are inverted from $3\,Hz$ to $15\,Hz$ in a total of 60 iterations. Figure 3.3 shows the true (black line), starting (dashed black line) and recovered

velocity profiles at the centre of model A. In particular, Figure 3.3a also shows the recovered models after acoustic FWI of the acoustic (green line with filled squares) and elastic datasets (blue line with filled triangles and red line with filled circles for the first and second elastic models, respectively).

As expected, acoustic FWI of acoustic data leads to the most accurate result when compared to the true profile, which is only limited by the finite range of inverted frequencies. The latter is more accurate at the sea-bottom, as the starting model is the same as the true model, and is less accurate at the second interface due to the limited bandwidth and also because the starting model is further from the true model.

On the other hand, the recovered model after acoustic FWI of the first elastic dataset leads to a slightly more accurate recovered model than that obtained for the second elastic dataset. Both acoustic inversions of elastic data result in recovered models that have a ripple at the sea-bottom: this is caused by ignoring elastic effects during the inversion given the strong impedance contrast for S-waves at this interface, which has an effect on the accuracy of the recovered model within the first layer.

In order to mitigate elastic effects and improve the quality of the recovered P-wave velocity models, I then apply the data-matching workflow using the recovered P-wave velocity models and the true v_p/v_s obtained from Figure 3.2, and the results are shown in Figure 3.3b. Thus, acoustic and elastic data are generated for all sources using the latter P- and S-wave velocity models in Step 2 of the workflow in Figure 2.8, and Wiener filters that match the elastic to the acoustic synthetic data are computed and convolved with the observed elastic data in Step 3 of the workflow to produce a matched observed dataset, in which elastic effects are mitigated. A second acoustic FWI step (Step 4 of the workflow) results in the recovered models shown in blue and red in Figure 3.3b for both elastic datasets: this is one iteration of the data-matching workflow. The recovered models after applying the current workflow are more accurate than those obtained with conventional acoustic FWI (see Figure 3.3a) and the inaccuracy at the sea-bottom is considerably reduced, especially for the first elastic model as elastic effects are weaker. To further reduce the impact of elastic effects and improve the quality of the recovered model on the second elastic dataset,

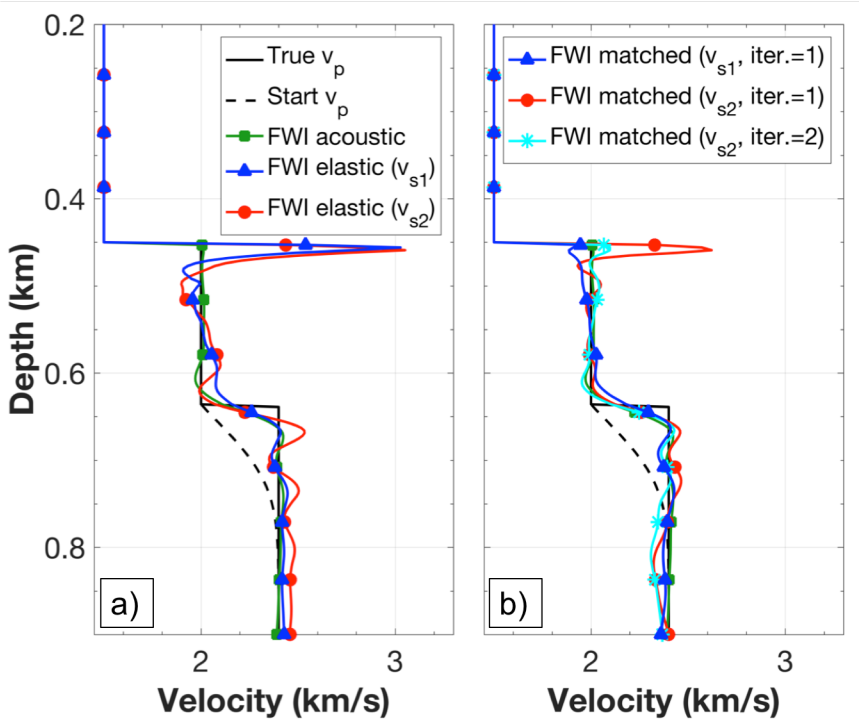


Figure 3.3.: Vertical profiles of the recovered P-wave velocity models for model A after a) acoustic FWI of the acoustic and elastic datasets and b) the acoustic and the matched datasets after 1 and 2 iterations of the data-matching workflow. The true, starting and recovered P-wave velocity model after acoustic FWI of acoustic data are overlaid.

I apply again the data-matching workflow. At this stage, I use the true v_p/v_s again and the recovered model after the first iteration of the workflow. This results in an improved velocity model, which is shown in cyan (with stars) in Figure 3.3b: this is the second iteration of the workflow. Thus, a second iteration of the workflow can be used to further mitigate elastic effects in the presence of strong elasticity. However, this increases the computation time, mainly due to the cost of an additional acoustic inversion, but not as much as a full elastic inversion.

In the previous examples, I have used the same filtering parameters in all data-matching tests. The choice of these parameters has an impact on the quality of the matched data and, hence, the quality of the recovered P-wave velocity models obtained with this method. Therefore, I now study how the matching parameters influence the quality of the matched data by computing Wiener filters with different

lengths and matching data trace-by-trace or matching several traces simultaneously. As changes in these parameters lead to similar conclusions on several datasets, I just discuss the impact of the latter for the first elastic model for model set A (obtained from the P-wave velocity profile for v_{s1} in Figure 3.3).

Thus, Figure 3.4 shows a shot gather at the centre of the model for the observed elastic data (Figure 3.4a) and the matched data obtained with different filtering parameters (remaining panels on the left column in Figure 3.4), whereas the difference of the panels on the right column with the true acoustic data – i.e. generated with zero S-wave velocity – are shown on the right column in Figure 3.4. With respect to the number of traces matched simultaneously, the result in Figure 3.4c is obtained matching data trace-by-trace, whereas the matched data in Figure 3.4e, 3.4g and 3.4i are obtained by matching 11 neighbouring traces simultaneously: that is the separation between the first and the last traces to be matched is $150\,m$ – note receiver spacing is $15\,m$. On the other hand, the length of the Wiener filters is of $320\,ms$ in Figures 3.4c 3.4e, of $35\,ms$ in Figure 3.4g and of $1.58\,s$ in Figure 3.4i.

By comparing the difference plots in Figure 3.4 several conclusions can be drawn. First, differences with the true acoustic data are overall smaller for the matched data with respect to those for the observed elastic data, except for localised areas with larger errors on the matched data. The latter are due to the fact that acoustic and elastic data modelled in Step 2 of the workflow are not generated from the true P-wave model, but from the less accurate recovered P-wave velocity model after acoustic FWI of the observed data. Nevertheless, all tests in this thesis demonstrate that these small errors do not hinder the improvement in the quality of the recovered P-wave velocity models.

A second conclusion obtained from Figure 3.4 is that trace-by-trace filtering leads to matched data that is more noisy, and that simultaneously matching several neighbouring traces introduces horizontal smoothing in the data so that the different events are more continuous. However, matching many traces simultaneously may have a negative impact on the result if various events (for instance due to P-waves and S-waves) are involved in this computation, which may contaminate the results.

Finally, the third conclusion is that the length of the filters can significantly affect

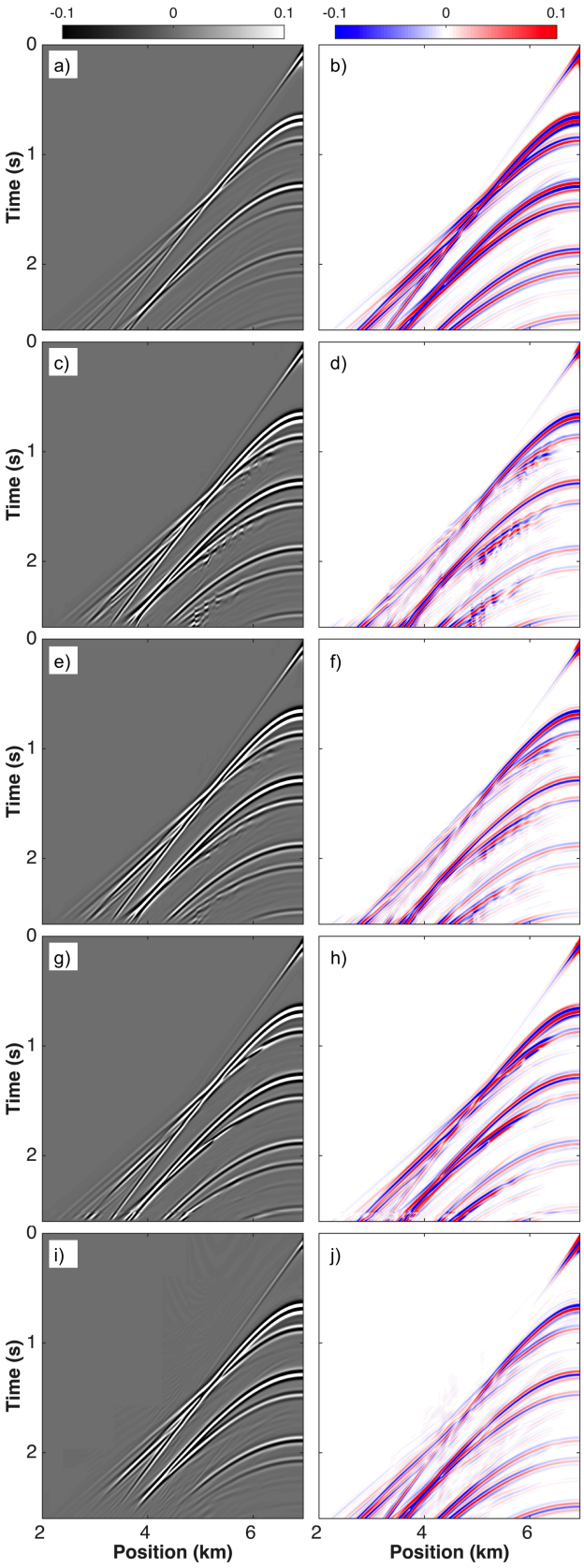


Figure 3.4: In the left column, representative shot gathers for a) the observed elastic data and c), e), g) and i) the matched data obtained at the centre of the first elastic model for model A, the latter obtained with the following filtering parameters: c) trace-by-trace and 320 ms long, e) multiple traces and 320 ms long, g) multiple traces and 35 ms long, and i) multiple traces and 1.58 s long. The difference of the gathers on the left with the corresponding shot gather for the true acoustic data is shown on the right column.

the quality of the matched data and the compute time. Thus, short filters are fast to compute – as the data-matching algorithm depends on the number of samples to be filtered– but produce errors at short offsets and in regions where multiple events cross each other, whereas longer filters give better results at short offsets and at crossing points, but lead to worse results at large offsets. Figure 3.4e demonstrates that an intermediate filter length provides a good quality matched data without the large computational burden associated with long filters.

After all the tests in this chapter, I recommend to use filters with a length between $300\,ms$ and $400\,ms$, while matching several traces simultaneously so that the spacing between the first and the last matched traces is between $100\,m$ and $850\,m$.

Next, I perform acoustic FWI of the acoustic and elastic data obtained for model B and I study the limitations of the data-matching workflow and conventional acoustic FWI when elastic effects are strong and the starting model is not very accurate. The P-wave velocity profiles at the centre of the models for all these tests are displayed in Figure 3.5 for two different starting models: the first is more accurate within the first layer, where elastic effects are more important due to the impedance contrasts (Figures 3.5a and 3.5b), whereas the second is less accurate within the first layer and more accurate towards the bottom of the model (Figures 3.5c and 3.5d).

For both starting models, acoustic FWI of the acoustic data (green line with filled squares overlaid in all plots) results in a recovered model very close to the true one, even at interfaces as the acoustic impedance changes are less pronounced than for model A – i.e. there are no sudden jumps in P-wave velocity, but gradual changes associated with the gradients in velocity. The recovered profiles after acoustic FWI of the two elastic datasets are displayed in Figures 3.5a and 3.5c for the two starting models, respectively, in blue (with filled triangles) and red lines (the latter with filled circles). Acoustic FWI of the elastic datasets using the first starting model (see Figure 3.5a) results in good quality recovered models down to $750\,m$ depth, except for inaccuracies at the sea-bottom and at interfaces caused by the ignored elastic impedance contrasts. Below $750\,m$, the combination of elastic effects and a poor starting model affect the goodness of the recovered models, which are slower than the

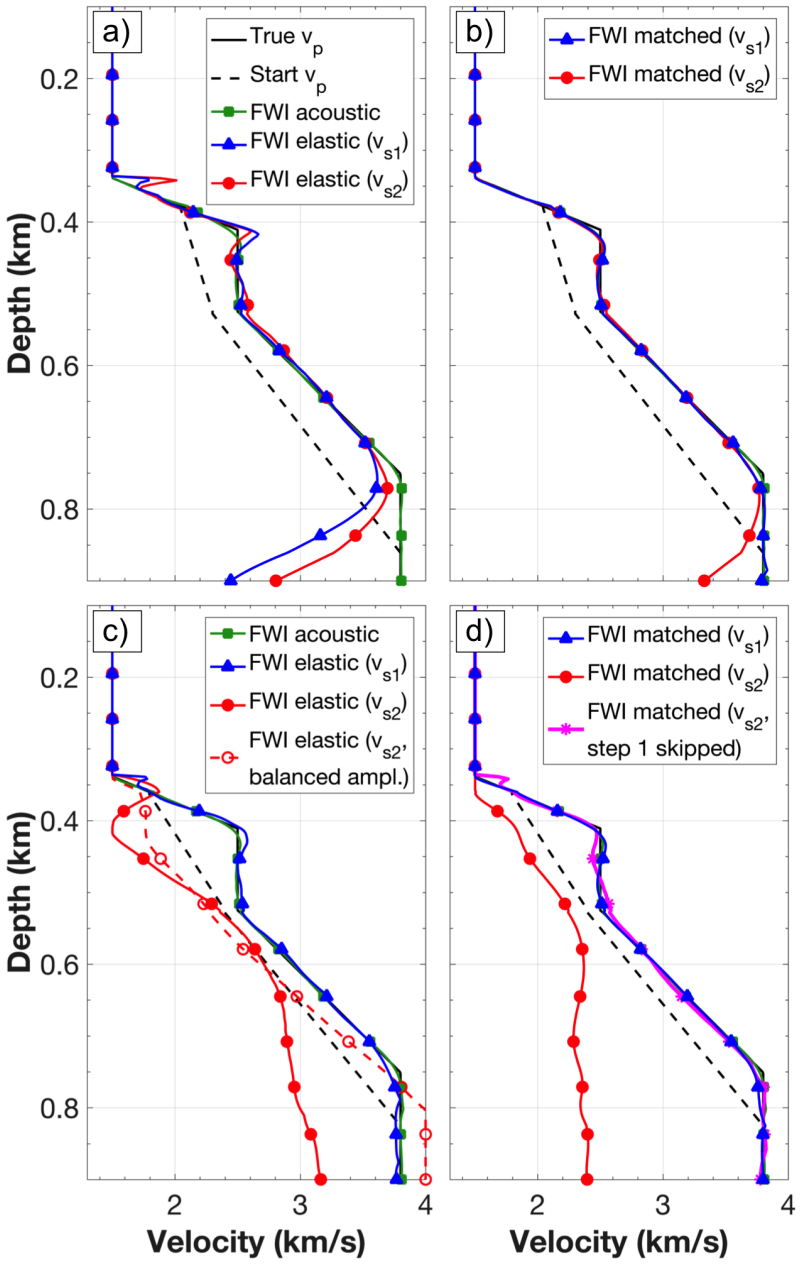


Figure 3.5.: Vertical profiles of the recovered P-wave velocity models for model B, in which all inversions are acoustic. The acoustic and elastic datasets are inverted in a) and c), whereas the corresponding matched datasets are inverted in b) and d) (for a single iteration of either workflow 1 or 2). The inversions in a) and b) are performed with the same starting model, whereas those in c) and d) use another starting model less accurate at the sea-bottom. The conventional FWI functional is used throughout, but the amplitude-balanced functional is also used in c). The true and starting models and the acoustic FWI of acoustic data are overlaid on all plots.

true model and this gets worse with depth. Figure 3.5b demonstrates that a single iteration of the data-matching workflow using the latter recovered P-wave velocity models and the true v_p/v_s provides improved P-wave velocity models, especially for the first elastic dataset in which elastic effects are weaker. Analogously to the tests for model A, the recovered profile for the second elastic dataset could be further improved with a second iteration of the data matching workflow if needed.

On the other hand, acoustic inversion of both datasets with the conventional objective functional and the second starting model (see Figure 3.5c) leads to very different results. Whereas acoustic FWI of the first elastic dataset is able to recover an accurate model (blue line with triangles in Figure 3.5c), it fails in recovering an accurate model for the second elastic dataset (red line with filled circles in Figure 3.5c). Thus, when elastic effects are strong and the starting model is not accurate in the shallow subsurface, conventional acoustic FWI of the elastic data leads to incorrect recovered models, probably due to cycle-skipping effects. Application of the data-matching workflow using this recovered model and the true v_p/v_s does not improve the result, as shown in Figure 3.5d (red line with filled circles). Hence, this is a limitation of the workflow: the workflow in Figure 2.8 fails when elastic effects are strong and starting model close to the sea-bottom is poor as the P-wave velocity model used to generate acoustic and elastic data is far from the true model.

Alternatively, I use the objective functional proposed by Shen (2010) to acoustically invert the second elastic dataset, which is commonly used to mitigate elastic effects. This results in the recovered model shown in Figure 3.5c (dashed red line with empty circles), which is better than that obtained with the conventional functional, but it is still far from the true model. Therefore, the amplitude-balanced method is also insufficient in the presence of strong elastic effects and poor starting models in the shallow subsurface.

In such cases I propose to use a variation of the workflow proposed in Figure 2.8. This new workflow, referred from now onwards as workflow 2 – in contrast to the original workflow, or workflow 1 – is based on the observation that acoustically inverting an elastic dataset that contains strong elasticity and in which the starting model is not very accurate leads to a recovered P-wave velocity model that may be

worse than the starting model. Thus, what I suggest is to skip Step 1 in the workflow in Figure 2.8 – hence, avoiding the acoustic inversion of the observed data– and proceed with the remaining steps: that is modelling acoustic and elastic data from the starting P-wave velocity model directly and the estimate of v_p/v_s, computing matching filters that match the elastic to the acoustic data, convolving them with the observed elastic data and performing acoustic FWI of the matched data for the first time.

The application of workflow 2 to the second elastic dataset for the second starting model is shown in Figure 3.5d (magenta line with stars). The latter leads to a very accurate recovered P-wave velocity model in the presence of large elasticity and poor quality starting model at the sea-bottom at a cost which is slightly higher than a single conventional acoustic inversion and is nearly half the cost of the original workflow, given that the first acoustic inversion step is not required. Moreover, the modified workflow provides a more accurate result than that obtained with the amplitude-balanced functional. Nevertheless, after various tests on several 2D datasets I recommend the use of the original workflow when possible, and utilisation of the modified workflow only when elasticity effects are large and acoustic FWI of elastic data leads to results that are not geologically plausible.

Similar to the conclusions of Barnes and Charara (2009) for 1D models, the tests in this section for 2D horizontally layered models suggest that recovered P-wave velocity models obtained with acoustic FWI in marine environments are not reliable when S-wave impedances are large, in which elastic effects need to be considered. Moreover, I have demonstrated that the combination of large elastic effects and poor starting P-wave velocity models close to the sea-bottom leads to inaccurate recovered P-wave velocity models when using conventional acoustic FWI or an amplitude-balanced functional. For all these situations I have shown that the proposed data-matching workflow – either workflows 1 or 2– can be used to address elastic effects and obtain high quality recovered models at a lower cost than elastic FWI, which is normally between one and three times the cost of conventional acoustic FWI.

3.3. Application to a marine synthetic dataset

I now address elastic effects in a more realistic marine synthetic dataset obtained from the Marmousi2 model (Martin et al., 2006), which is an extension of the well-known Marmousi P-wave velocity model in depth and width that is fully elastic. The latter captures the geology of the North Quenguela Trough (Quanza Bazin, Angola), which is mainly composed of shale units, some sand layers and a complex faulted anticline at the centre of the model that contains marl. In the deeper sections, an unconformity and a small salt layer separate the central anticline with the anticline below, which is mainly formed by shale and sand layers. A density model for generating observed data is obtained by using Gardner's law and the true P-wave velocity model.

Tests in this section are designed to study the impact of elastic effects in acoustic FWI of a typical marine dataset, investigate how can these be addressed with the data-matching workflow and what is the consequence of using a rough estimate of the v_p/v_s in the workflow as well as assessing the results when the data is contaminated with noise. For this purpose, acoustic and elastic synthetic pressure data are first generated using the original true models and limiting the water depth to be $200\,m$ (instead of the original $450\,m$), as shown in Figure 3.6, in order to reduce the computation time and reinforce the presence of elastic effects.

Observed data are generated with finite-differences modelling codes that are able to implement the acoustic and elastic wave equations. I use a grid spacing of $2.5\,m$, which ensures there is little dispersion both for P-waves and S-waves as the minimum S-wave velocity is $270\,m/s$ (as per equation (2.32)), which results in a relatively large 2D model with 1301×6801 cells. A total of 111 sources are deployed at $6\,m$ depth with a shot spacing of $150\,m$, which emit a pressure signal with a useful bandwidth from $2\,Hz$ to $20\,Hz$, as depicted in Figure 3.7. A total of $6.7\,s$ of data are modelled using a time sampling of $0.25\,ms$ to ensure numerical stability given the maximum P-wave velocity and the grid spacing (as per equation (2.33)). The data is then recorded by 3361 receivers deployed at $10\,m$ depth, with a constant receiver spacing of $5\,m$. Absorbing boundaries are used on the sides and the bottom of the model, whereas a free surface is used at the top in order to simulate marine data: thus, source and

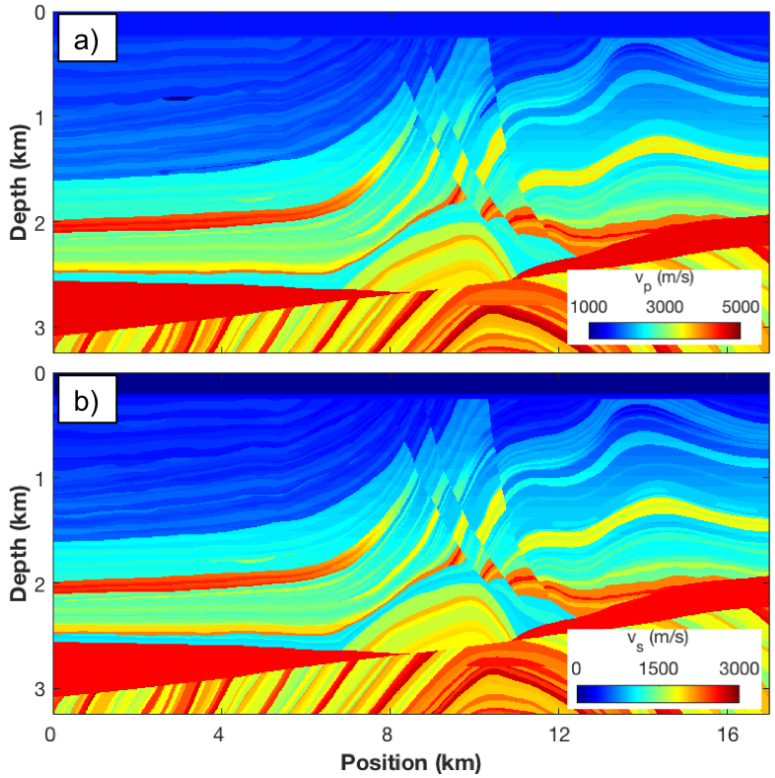


Figure 3.6.: Vertical slices of the true marine Marmousi2 a) P-wave and b) S-wave velocity models.

receiver ghosts and surface multiples are present in the data.

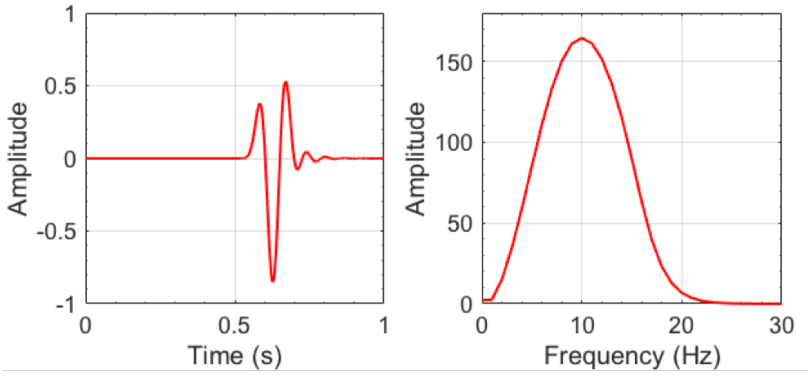


Figure 3.7.: Source signature in the time domain (left) and its amplitude spectrum (right) used to generate data for the Marmousi2 model.

Figures 3.8a and 3.8b show shot gathers for the acoustic and elastic observed data, obtained at the centre of the model, whereas Figure 3.8d illustrates the difference between the elastic and the acoustic data. The differences between the two datasets are small, as expected in a marine environment, but they are visible especially in

terms of amplitude of P-wave arrivals even with the same color-scale used for the raw data. Additionally, P- to S-wave converted waves are also present, but these are difficult to observe as the dataset contains many events given that the model contains many features and, therefore, waves travelling through this model are refracted and reflected at many interfaces.

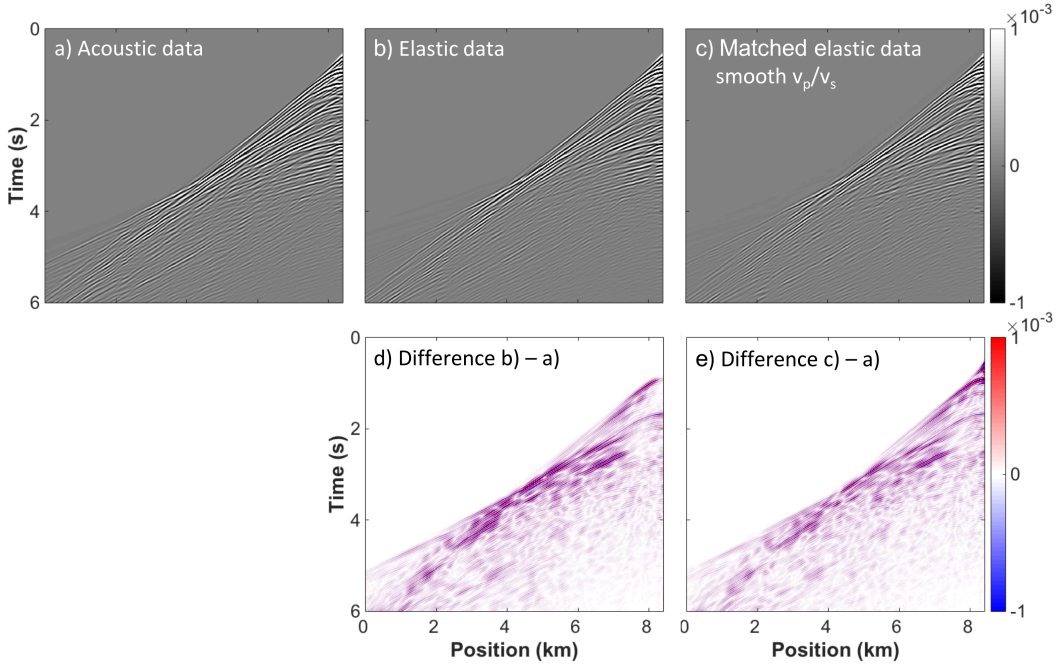


Figure 3.8.: Representative shot gathers at the centre of the Marmousi2 model for the a) acoustic, b) elastic and c) the matched data using a smooth v_p/v_s. Shot gathers in d) and e) show the difference of the panels above with the acoustic data in a).

Next, I invert the noise-free acoustic and elastic observed data independently using acoustic FWI to assess the impact of elasticity on the inverted results for a typical marine dataset, and I apply the data-matching workflow to mitigate these effects. The results are shown in Figure 3.9. Figure 3.9a shows the true P-wave velocity model again for comparison, in which I have overlaid a dotted area where the average RMS difference for all cells within this region with respect to the true model is computed. Thus, the closer this value is to zero the more accurate the recovered model is (hence, the average model RMS for the true model is $0\,m/s$). Average model RMS values are computed for all models in Figure 3.9 to better assess the quality of the results.

First, the observed acoustic and elastic datasets are inverted using the model in

Figure 3.9b as a starting model, which I obtained by smoothing the true P-wave velocity model using a Gaussian function of $7.5\,m$ correlation length. I applied the latter smoothing function several times depending on the depth level, so that the top layer of the starting model is better resolved than deeper layers. This is usually the case, as travel-time tomography (TT) is used to generate starting models and ray coverage and penetration is better close to the surface. All inversions in this section are carried out using the conventional FWI functional and using a multi-scale approach by inverting the data in the time-domain from $3\,Hz$ to $18\,Hz$ in a total of 145 iterations.

Figures 3.9c and 3.9d illustrate the recovered models after acoustic FWI of acoustic and elastic data, respectively. As expected, the differences between the two inverted model are small, but visible. For instance, I note the loss in resolution of the individual layers in the model in Figure 3.9d when compared to that in Figure 3.9c (i.e. the thickening of the nearly horizontal layer with high velocity at $2\,km$ depth from $0\,km$ to $6\,km$ distance, indicated by a purple arrow), the deterioration in the accuracy of the velocity (e.g. slower velocities than those in the true model where the black arrow is pointing) and the noisier behaviour of the model below $2.2\,km$ depth. These differences translate into an increase of the average model RMS difference within the dotted area in the elastic case (of $207.45\,m/s$) when compared to the acoustic case (of $167.79\,m/s$). The data-matching workflow should partially compensate for these effects and produce higher quality recovered P-wave velocity models.

Therefore, the next step is to implement the workflow in Figure 2.8 to address elastic effects. As elastic effects are not very strong on this dataset, I implement workflow 1 by using the recovered model in Figure 3.9d and a guess of v_p/v_s. Initially, I use an estimate of the v_p/v_s model obtained by smoothing the true v_p/v_s model in Figure 3.10a, yielding the model in Figure 3.10b – I study the impact of having a very approximate v_p/v_s estimate in the next sub-section.

I proceed with the data-matching workflow by modelling acoustic and elastic data and computing Wiener filters that match 35 traces simultaneously, so that the spacing between the first and the last trace being matched is of $170\,m$, and I convolve the computed filters with the observed data. I then run acoustic FWI of the matched

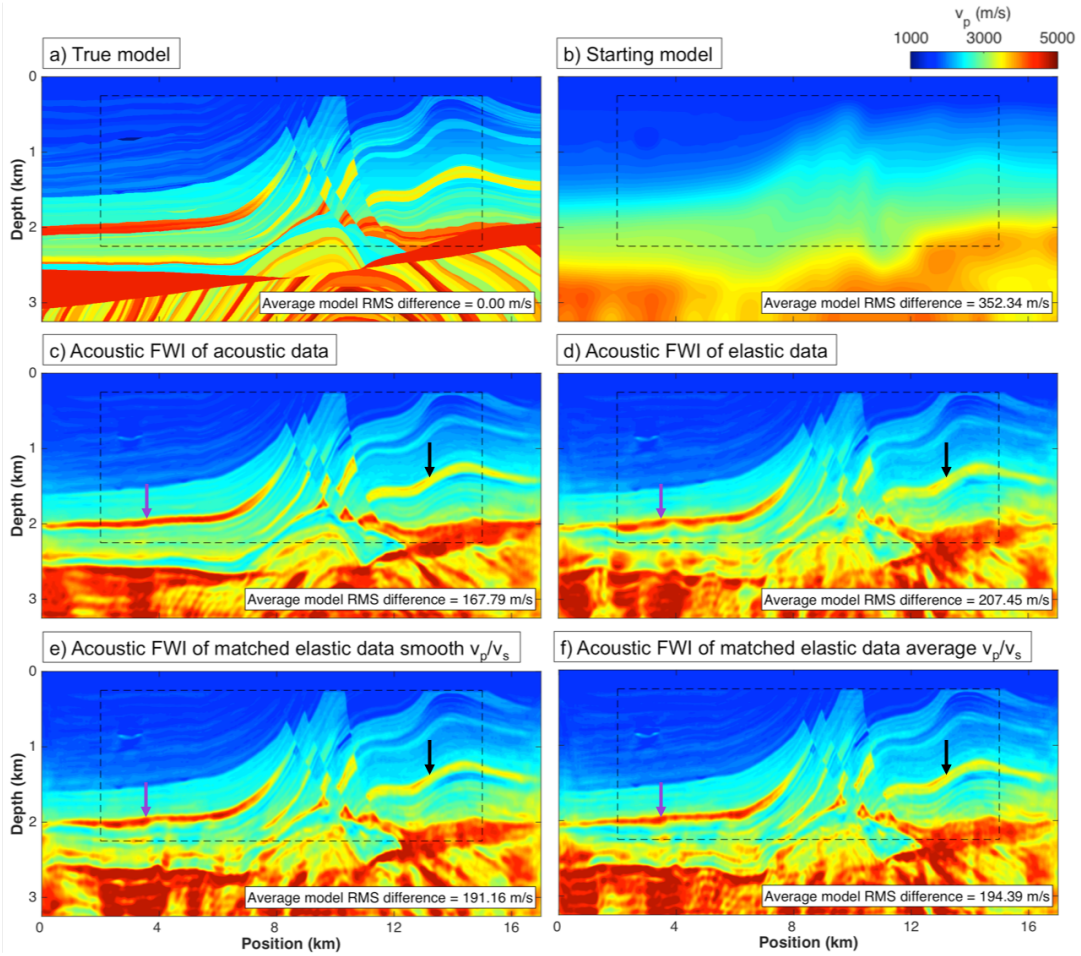


Figure 3.9.: Vertical slices of the Marmousi2 a) true, b) starting and c) - f) recovered P-wave velocity models. All inversions are acoustic and the data being inverted is c) the acoustic, d) the elastic, e) the matched data with smooth v_p/v_s estimate and f) matched data with average v_p/v_s estimate. The average model RMS difference with respect to the true model is computed for all models within the dotted box, and it is displayed at the bottom of each model.

data in Step 4 of the workflow and I perform a second iteration of the data-matching workflow in Step 5 to further improve the quality of the recovered model: the total compute cost is therefore between three to four times the cost of conventional acoustic FWI. Figure 3.8c shows the resulting shot gather for the matched data at the second iteration of the workflow, whereas its difference is plotted in Figure 3.8e. When compared to the elastic data in Figure 3.8b and its difference with the acoustic data in Figure 3.8d, the matched data is closer to the acoustic data and, hence, the overall differences are reduced (except at very short offsets, which is due to the length of

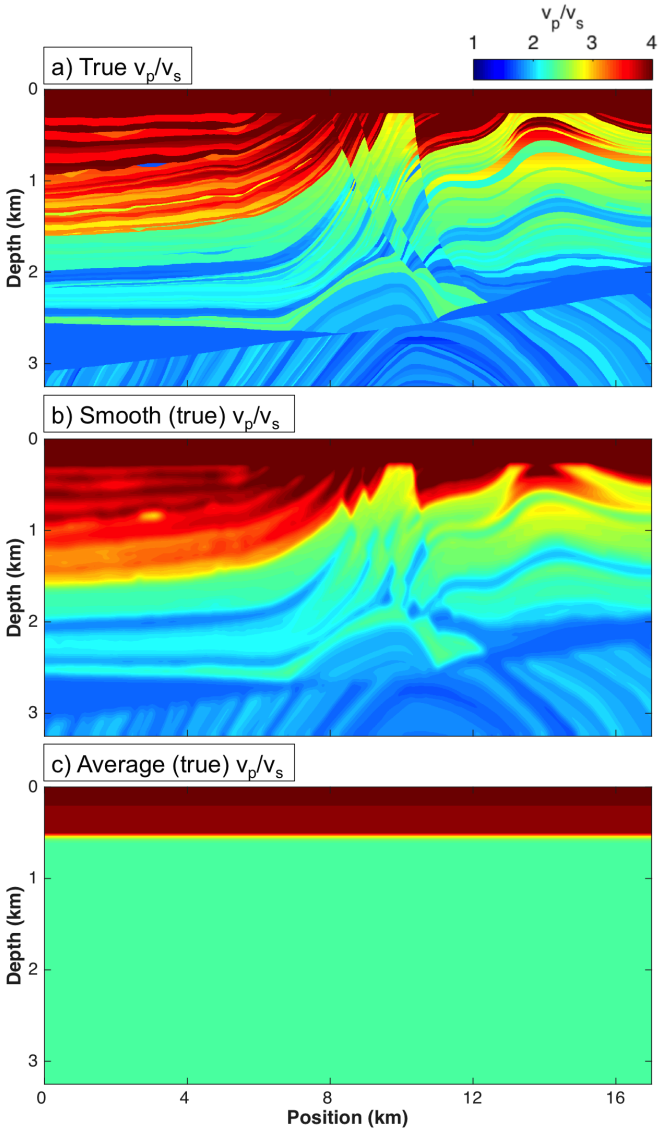


Figure 3.10: Vertical slices of a) the true v_p/v_s, b) a smooth v_p/v_s obtained by mildly smoothing the true v_p/v_s and c) an average 1D v_p/v_s estimate obtained by averaging the true v_p/v_s within two regions and mildly smoothing the result.

the filters and has little impact on the inversions, which are dominated mostly by refractions and less by reflections). For this reason, acoustic inversion of the matched dataset results in an improved P-wave velocity model, as shown in Figure 3.9e. The latter is better resolved than the model obtained after acoustic FWI of elastic data in Figure 3.9d, so that interfaces are sharper and layers are thinner (e.g. high-velocity layer indicated by the purple arrow), and it is also overall more accurate (e.g. there is an increase in the velocities in the layer indicated with a black arrow, which is closer to the values in the true model).

In summary, the recovered velocity model using the data-matching workflow is better than the model obtained after acoustic FWI of elastic data and not as accurate as

that obtained with acoustic FWI of the acoustic data – this is expected as the latter is the reference result (no elastic effects). This can also be seen by the reduction on the average model RMS value within the dotted area, which decreases from $207.45\,m/s$ to $191.16\,m/s$. More iterations of the workflow could further improve the result but this would increase the computation cost as an acoustic inversion is involved at each step. Moreover, this and other tests indicate that the first iteration of the workflow is the most important – it is in the first iteration that elastic effects are mitigated for the first time.

I further assess the quality of these results by computing the model RMS difference as a function of depth – i.e. I average the model RMS difference for all positions for a certain depth level and I plot the values for all depths – and position within the dotted box for all models. These plots, shown in Figure 3.11, give more detailed information on the differences with respect to the true model on average for each depth level (Figure 3.11a) and each position (Figure 3.11b) than the average model RMS difference on the plots in Figure 3.9.

From Figure 3.11 I observe the following: 1) in a marine dataset in which elastic effects are small, acoustic FWI of either acoustic or elastic data leads to more accurate results than the starting model, 2) acoustic FWI of acoustic data (red line with filled circles) is more accurate than acoustic FWI of elastic data (black line) for all depths and positions except for localised anomalies, 3) acoustic FWI of the matched data using the smooth v_p/v_s model leads to a model closer to the true one than that after acoustic FWI of elastic data, especially below $1375\,m$ depth and for positions greater than $8\,km$. More importantly, I conclude that acoustic FWI of elastic data for the marine Marmousi2 model, which represents a marine dataset with weak elastic effects, provides good quality results down to $1.6\,km$ depth – this can also be observed in Figures 3.9c and 3.9d–, and it degrades below this level. The data-matching workflow is capable of mitigating such small elastic effects and lead to a better resolved and more accurate model below this depth. Therefore, I suggest weak elastic effects in marine datasets can be addressed with the data-matching workflow, and the data-matching method may provide more impressive results on

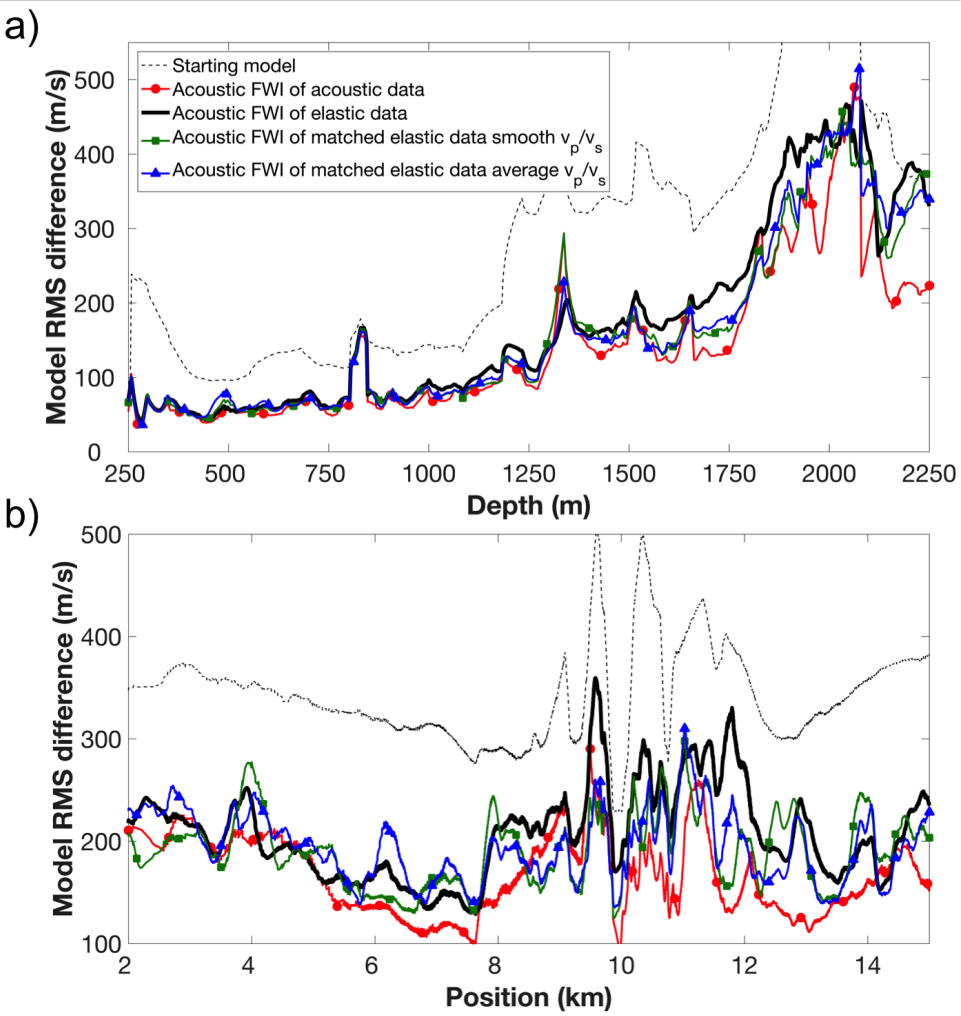


Figure 3.11.: Graphs of the model RMS difference with respect to the true model as a function of a) depth and b) position for the models in Figure 3.9.

marine datasets with stronger elastic effects (as in the 3D field dataset) or on land datasets, where elastic effects are inherently stronger. The latter is discussed in the next section, but first I assess the consequences of using a very approximate v_p/v_s model within the data-matching workflow, the impact of noise on the results and how the best results in Figure 3.9 compare to a full elastic inversion.

3.3.0.1. Impact of the v_p/v_s estimate

In the tests performed so far, I have used the true v_p/v_s model in order to mitigate elastic effects using the data-matching workflow. However, the latter is typically unknown on a field dataset, especially if well-logs are not available. Thus, I now

study how accurate this estimate needs to be to lead to an improvement using the current method for the Marmousi2 model.

For this purpose, I design a very smooth and approximate 1D v_p/v_s model by averaging the true ratio in just two regions, and applying mild Gaussian smoothing to avoid artificial diffractions at interfaces. The latter, which I refer to as the average true v_p/v_s, is shown in Figure 3.10c: this v_p/v_s estimate is more accurate near the sea-bottom to capture the elastic effects generated in this area, and much less so below $600\,m$.

After using the average v_p/v_s model and the recovered P-wave velocity model after acoustic FWI of elastic data in the implementation of the data-matching wokflow, the obtained recovered model is shown in Figure 3.9f. Similarly to the model obtained with the smooth v_p/v_s presented in Figure 3.9e (the same inversion and filtering parameters are used), the P-wave velocity model in Figure 3.9f is more accurate and has more detail than that after acoustic FWI of elastic data (Figure 3.9d), and is very similar to the model obtained with the smooth true v_p/v_s. This can be seen by visual inspection of Figures 3.9f and 3.9e, by comparison of their average model RMS difference within the dotted box ($194.39\,m/s$ with respect to $191.16\,m/s$) and by their similar model RMS variations with respect to depth and position (Figure 3.11).

Consequently, using an average 1D v_p/v_s model is enough to mitigate most of the elastic effects for this model as the most important elastic effects are generated at the sea-bottom, where the impedance change for S-waves is larger, and this is in average captured by the average v_p/v_s model. Moreover, this results suggest that, even though v_p/v_s and v_s are not well known, the performance of acoustic FWI of elastic data for marine datasets can be improved by using the data-matching workflow and a rough estimate of the v_p/v_s that captures the approximate value at the sea-bottom – or in those regions where strong impedance contrasts for S-waves are possible, if applicable.

3.3.0.2. Impact of noise

I now test the robustness of the data-matching workflow with respect to added random noise on the observed elastic data. First, white Gaussian noise filtered with the spectrum of the source – so that it has the same bandwidth as the data– is added to the elastic data, and a representative shot gather is shown in Figure 3.12a, whereas its difference with the true acoustic data is presented in Figure 3.12c.

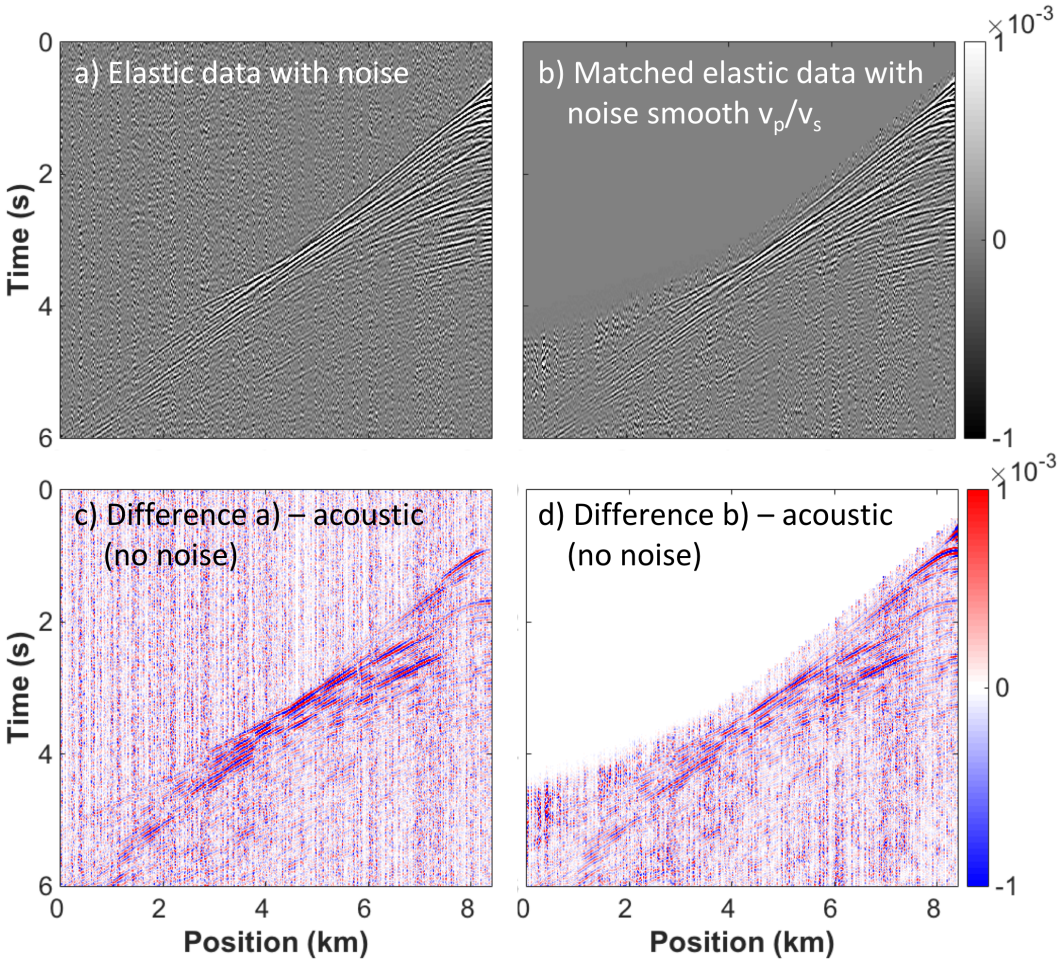


Figure 3.12.: Representative shot gathers at the centre of the Marmousi2 model for a the elastic data contamined with white Gaussian noise and b) the matched data obtained from the noisy elastic data, its acoustic inversion and the smooth v_p/v_s in Figure 3.10b. The gathers in c) and d) are the differences of the gathers abover with respect to the true acoustic data.

The noisy elastic data is then acoustically inverted with the same inversion parameters as for the previous tests, which includes mild smoothing to reduce the ill-posedness of the inverse problem, and the recovered P-wave velocity model is shown

in Figure 3.13a. This model is overall noisier than that in Figure 3.9d for the noise-free data as reducing the misfit between observed and acoustically modelled data during acoustic FWI is now more difficult due to the added noise and also the elastic effects. As a result, there is a loss in accuracy and a larger average model RMS difference of $217.18\,m/s$ with respect to the noise-free scenario (in which the latter was $207.45\,m/s$).

I then use the recovered P-wave velocity model in Figure 3.13a as an input to Step 2 of the data-matching workflow, and I proceed with Steps 2 to 5 as in the previous tests with the same filtering and inversion parameters, so that two data-matching iterations are applied in total. The smooth v_p/v_s model is used and Figure 3.13b shows a representative shot gather for the matched data, whereas their difference with the noise-free acoustic data is plotted in Figure 3.12d. In the latter, I apply an automatic mute during the matching to reduce the compute time, as explained in chapter 2. Thus, matched data is only computed for each trace when a certain amplitude threshold is reached in the forward acoustic modelled data to reduce the number of Wiener filters to be computed.

Similarly to the results in Figure 3.8, the differences with respect to the (noise-free) acoustic data are reduced for the matched data within the region where Wiener filters are computed. A final acoustic inversion of the matched data results in the P-wave velocity model in Figure 3.13b. This model is not as well resolved as that obtained for the noise-free data (Figure 3.9e), but there is an improvement with respect to the acoustic FWI of the noisy elastic data in Figure 3.13a – for instance, there is a thinning of the layers indicated with black and purple arrows, and an overall reduction of the noise in the model due to the smoothing introduced by matching several traces simultaneously. This improvement is however smaller in comparison to the noise-free case: the RMS reduction from Figure 3.9d to 3.9f was of 8 %, whereas it is of 1.4 % from Figures 3.13a to 3.13b.

Thus, the presence of strong random noise in the observed data limits the performance of the data-matching workflow as the matched data also contains part of this noise. Nevertheless, the recovered models obtained with the proposed method

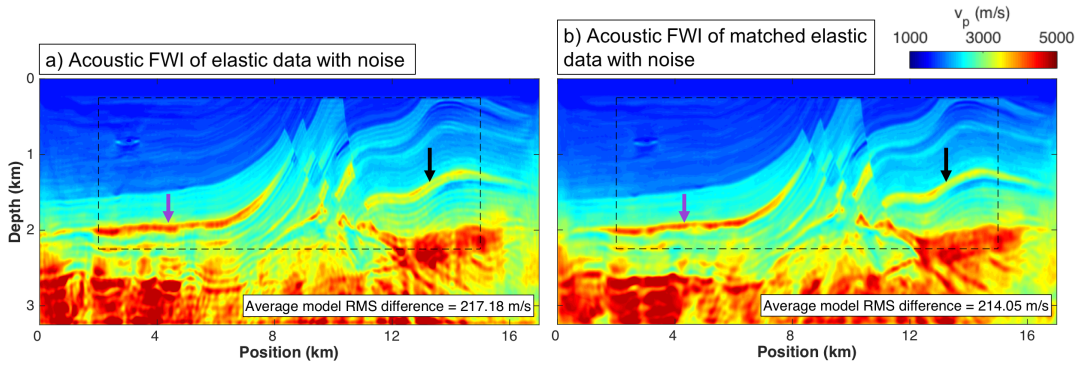


Figure 3.13.: Vertical slices of the recovered P-wave velocity models obtained after a) acoustic FWI of the noisy elastic data and b) acoustic FWI of the matched elastic data with noise.

are still more accurate and better resolved than those obtained with conventional acoustic FWI of the observed data.

3.3.0.3. Comparison with full elastic FWI

Finally, I compare the results obtained in the previous tests with respect to a full elastic 2D inversion of the noise-free elastic data. I perform elastic FWI using a forward modelling code that uses the isotropic heterogeneous elastic wave equation (Guasch, 2012) and the same inversion parameters and acquisition geometry as in the previous tests, but I invert for the P-wave and S-wave slownesses simultaneously – other multi-parameter inversion strategies exist and may yield better results, but this is not further explored here. The starting model for v_p in Figure 3.9b is again used for P-waves, whereas a starting model for v_s is built by using the latter and the average v_p/v_s model in Figure 3.10c, assuming that we have the same information on the v_p/v_s as the data-matching test in Figure 3.9f. Figures 3.14a and 3.14b respectively show the recovered P-wave and S-wave velocity models.

The recovered P-wave velocity model is not as accurate as the model obtained after acoustic FWI of acoustic data due to the cross-talk between P- and S-wave model parameters during the inversion, it is more accurate than the result of acoustic FWI of elastic data in Figure 3.9d and it is slightly better in terms of accuracy and resolution to that obtain with the data-matching workflow in Figure 3.9f. However, it is better resolved and less noisy below $2.4\,km$ depth as S-waves and amplitudes of P-waves

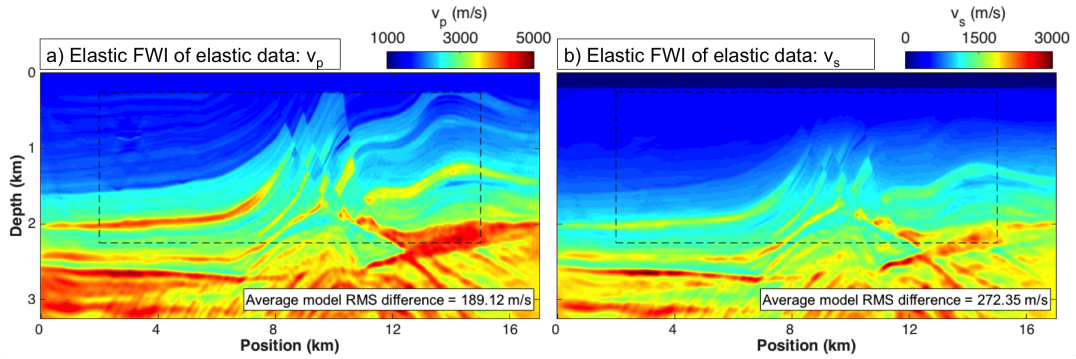


Figure 3.14.: Vertical slices of the a) recovered P-wave velocity model and b) S-wave velocity model after full elastic FWI of the elastic data using the original starting P-wave velocity model, a starting S-wave velocity model built from the later and the average v_p/v_s and inverting for v_p and v_s simultaneously.

are better handled. On the other hand, elastic FWI also provides a recovered S-wave velocity model. When compared to the true v_s model in Figure 3.6b, the former is poorly resolved, especially within the first $400\,m$ as the starting model for S-waves is obtained from the average v_p/v_s and does not contain any structure within this zone. However, the average recovered S-wave velocity is approximately correct in this region, which leads a good recovery of the P-wave velocity model for the shallow subsurface.

In terms of compute cost, elastic FWI is at least an order of magnitude more expensive than acoustic FWI. Thus, I use 56 nodes with 2 tasks per node for the acoustic inversion and 112 nodes and a single task per node for elastic FWI, as well as double the number of threads. This difference in the number of resources yields a similar elapsed time per iteration, which is of $51\,h$ for both acoustic and elastic FWI for this dataset. This difference should in fact be larger as the grid spacing used for the previous acoustic inversions is the grid spacing to avoid there is no dispersion of S-waves, and no S-waves are modelled during acoustic inversions. However, the same spacing was used throughout to avoid the hassle of resampling elastic modelled data to match it to the acoustic data during the data-matching workflow. This is now easily implemented and, therefore, I recommend to use a coarser grid spacing and time sampling for acoustic inversions and acoustic modelling with respect to the finer grid spacing and time sampling required to avoid energy dispersion and numerical

instabilities during elastic modelling – and, hence, during elastic inversions. Thus, when matching elastic to acoustic data the former needs to be resampled to the same time sampling as the acoustic data. I follow this strategy on the 3D marine field dataset in section 3.5.

In conclusion, the data-matching workflow with a very approximate average v_p/v_s estimate provides significant improvements with respect to conventional acoustic FWI of elastic data and slightly worse results than a full elastic inversion, but also at a much lower cost. For this reason, I recommend to initially use the data-matching workflow for marine datasets, especially in 3D, and then move to a full elastic inversion to obtain more accurate velocity models or if S-wave velocity models are required.

3.4. Application to a land synthetic dataset

Next, I implement the data-matching strategy to mitigate elastic effects on a realistic synthetic land dataset. Data are generated from the marine P-wave 2D SEG/EAGE overthrust velocity model (Aminzadeh et al., 1997), shown in Figure 3.15a, which I have converted to a land dataset by removing the water layer. The S-wave velocity model is built by using the P-wave velocity model in Figure 3.15a and a constant velocity ratio $v_p/v_s \approx 1.71$ – obtained from a constant Poisson's ratio of $\nu = 0.24$ such that $(v_p/v_s)^2 = (2(1-\nu))/(1-2\nu)$ (Brossier et al., 2009) – yielding the S-wave velocity model in Figure 3.15b.

Single-component acoustic and elastic pressure data are generated using an FD modelling code that implements the acoustic and elastic wave equations and the models in Figure 3.15, which are discretised so that the grid spacing is $10\,m$. The latter ensures there is little dispersion of either P- or S-waves below $17.5\,Hz$ (as per equation (2.32)) – note the minimum non-zero velocity is larger than in the previous marine dataset. Thus, the size of the velocity models is 239×1001 (depth $\times$ width cells). A total of 95 sources are deployed at $5\,m$ depth with a shot spacing of $100\,m$, which emit the pressure signal in Figure 3.7 band-pass filtered so that the useful bandwidth spans from $2\,Hz$ to $17\,Hz$ and with a time sampling of $0.5\,ms$. Sources

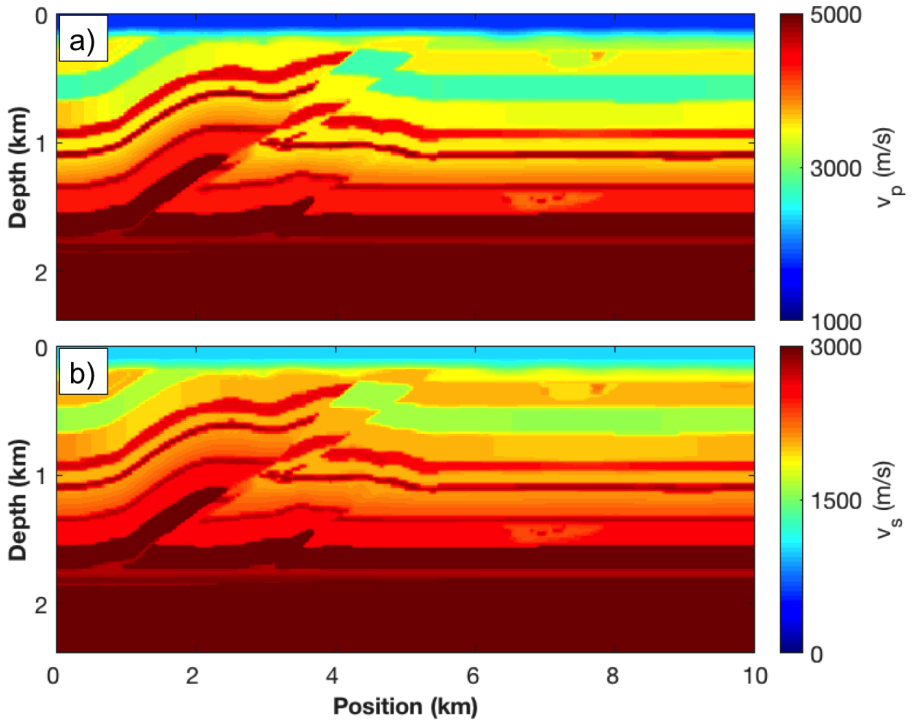


Figure 3.15.: Vertical slices of the true a) P-wave and b) S-wave velocity models used to generate seismic data for the overthrust model.

are fired one at a time and, unlike in the marine case, these are located in a layer with non-zero S-wave velocity. Therefore, elastic waves are directly generated near the source (S-waves are generated via conversion of P- to S-waves in a typical marine seismic experiment). The signal is then recorded by 961 receivers at 5 m depth, with a receiver spacing of 10 m. Absorbing boundaries are used on all sides of the model to simulate land data properly and avoid artificial reflections at the edges.

The modelled acoustic and elastic observed data for a representative shot gather at the centre of the model are displayed in Figures 3.16a and 3.16b, whereas the difference of the elastic and the acoustic gathers is plotted in Figure 3.16d. Two main differences can be observed by comparing these shot gathers: first, there is a visible amplitude difference of all P-wave arrivals, including the direct arrival, and second, there are events in the elastic data that are not present in the acoustic data. The former takes place as the wavefield splits into P- and S-waves right at the start, next to the source, and amplitude of all arrivals in both media are different due to the different energy splitting at interfaces for P- and S-waves. The second observation relates to

events generated by propagation of S-waves, which are present in the elastic case and do not exist in the acoustic case. Thus, the elastic data contains additional refractions and reflections due to S-waves (for instance, as the refractions with different slope or the reflections indicated by red arrows in Figure 3.16b).

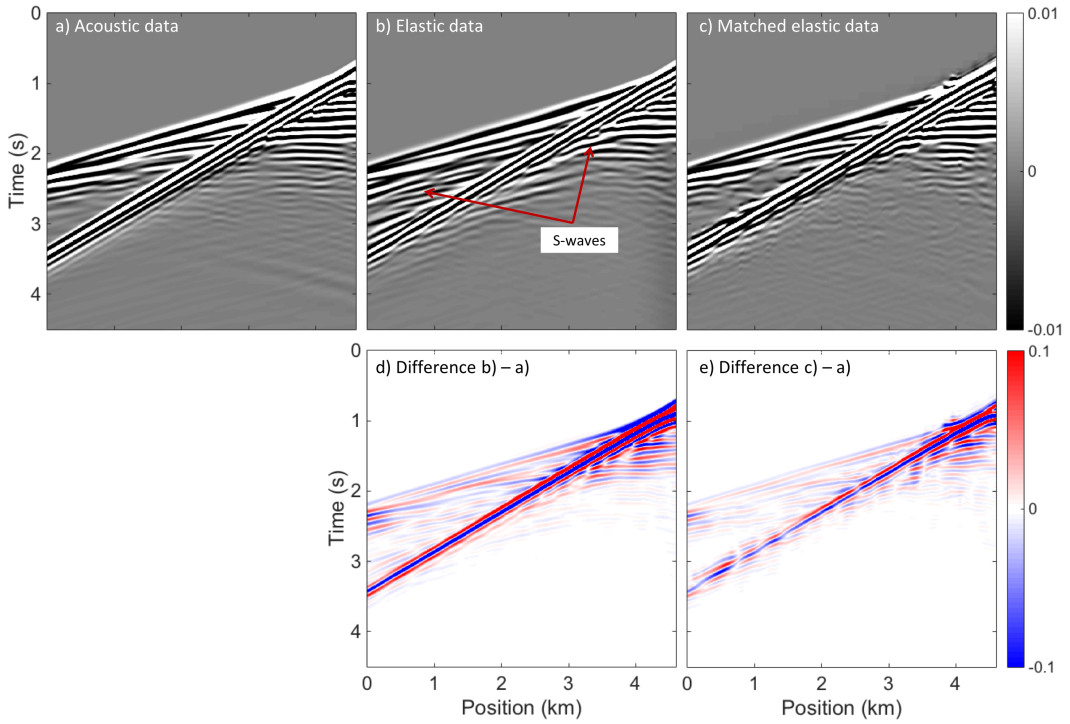


Figure 3.16.: Representative shot gathers at the centre of the 2D SEG/EAGE overthrust model for the a) observed acoustic, b) observed elastic and c) matched observed elastic data (obtained by matching 9 simultaneous traces and with a length of the filters of 450 ms). Panels d) and e) show the differences of the panels above with the true acoustic data.

For these reasons, it is expected that acoustic inversion of the acoustic and the elastic data perform differently. To assess the impact of the elastic effects in acoustic FWI of this land dataset, I now perform acoustic FWI of the acoustic and elastic data independently. I use the amplitude-balanced functional of Shen (2010) in all inversions for this dataset to further mitigate elastic effects, given than these are stronger on land data, and I invert the data using a multi-scale approach by inverting the data from 3.5 Hz to 16.7 Hz in a total of 144 iterations. The results of the inversions are presented in Figure 3.17, which also contains the improvement achieved by applying the data-matching workflow.

Figure 3.17a displays the true P-wave velocity model again as a reference, and Figure 3.17b is the starting model used in the inversions, which is obtained by smoothing the true model 180 times with a Gaussian function of 30 m correlation length. As in the previous section, the quality of the results is further assessed by computing the average model RMS difference within the dotted area with respect to the true model.

Acoustic FWI is first performed on the acoustic data, and the recovered velocity model is shown in Figure 3.17c. The latter is much better resolved than the starting model as the different layers in the overthrust can be distinguished, and the decrease in the average model RMS difference from 366.44 m/s to 217.97 m/s indicates that velocities are also more accurate. Additionally, the faulting system at 3.5 km position is well recovered, as the layers are well separated at the fault zone (indicated by a black ellipse), and the high-velocity layer at 300 m depth (indicated by a purple arrow) is well recovered. The quality of the result is mainly limited by the finite frequency band of the data and the inversion.

On the other hand, neglecting elastic effects and performing acoustic FWI of the elastic data results in a poor quality recovered P-wave velocity model, as shown in Figure 3.17d. The different layers are less well defined in the latter than in the result in Figure 3.17c, and the fault zone (indicated by a black circle) is more difficult to be identified. Furthermore, the low-velocity channel at 800 m (indicated by a black arrow) is incorrectly pushed down due to the slow S-wave velocities at the top of the model. All these effects result in an increase of the average model RMS difference with respect to that in Figure 3.17c to 339.58 m/s.

Hence, despite the use of an amplitude-balanced functional elasticity has a relevant and clearly visible adverse impact on the recovered P-wave velocity models of the overthrust land dataset and further needs to be done to improve the quality of the recovered P-wave velocity model. Thus, I combine the result of amplitude-balanced acoustic FWI with the data-matching worklfow. First, I model acoustic and elastic data (Step 2) by using the P-wave velocity model in Figure 3.17d and an S-wave velocity model built from the latter and the true v_p/v_s. I then compute Wiener filters that match the elastic to the acoustic data (nine traces are matched simultaneously

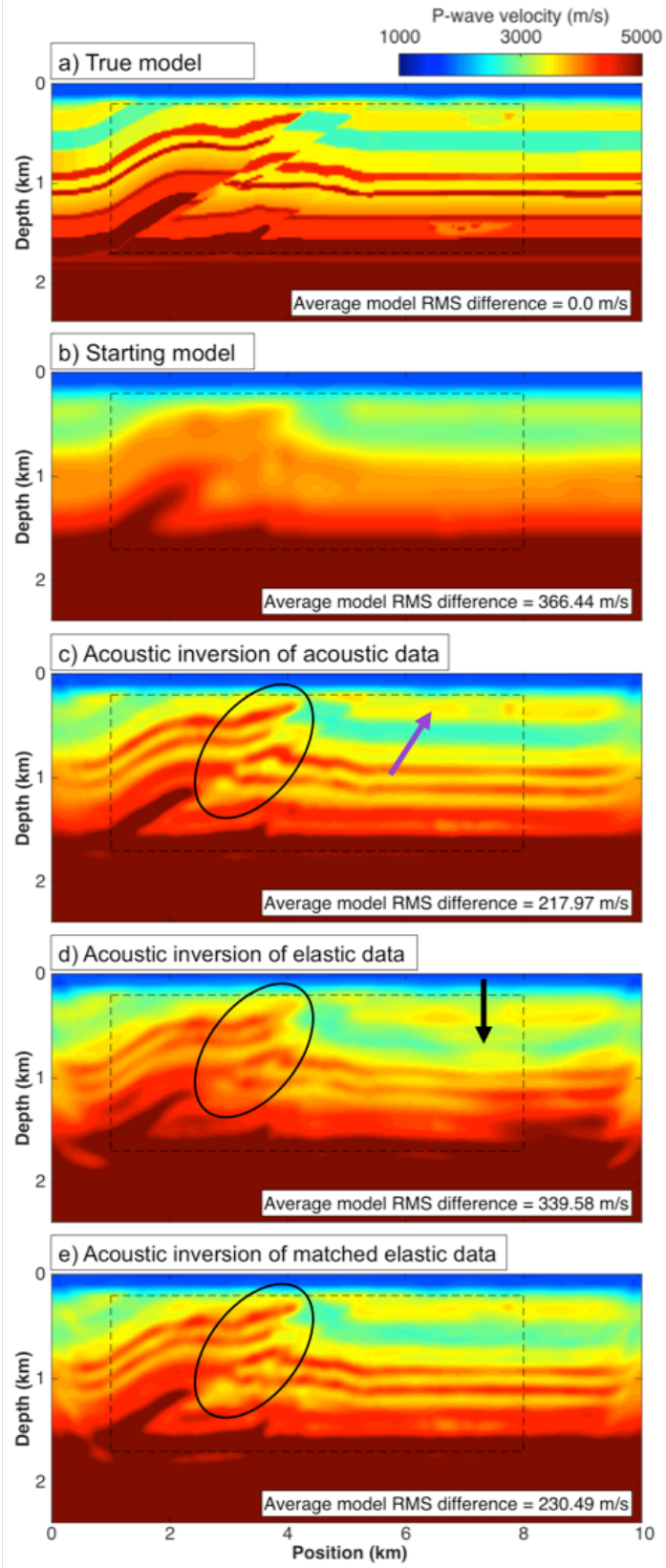


Figure 3.17: Vertical slices for the land 2D SEG/EAGE overthrust P-wave velocity model. From top to bottom: a) true model, b) starting model and recovered models after acoustic FWI of c) the acoustic, d) the elastic and e) the matched observed data. The average model RMS difference with respect to the true model within the dotted are are computed, and the values are overlaid on the models.

and the length of the filters is set to 900 samples, i.e. 450 ms) and I convolve them with the observed elastic data (Step 3). Acoustic FWI is then performed on the

resulting dataset (Step 4) and the data-matching workflow is repeated once more (Step 5) to further improve the result. Figure 3.16c shows the resulting matched data, and Figure 3.16e displays its difference with the acoustic data. Visibly, the overall differences with the acoustic data are reduced when compared to the elastic data in Figure 3.16d for all arrivals, including the direct P-wave arrival and the refractions, and the additional S-wave events indicated by a red arrow in Figure 3.16b have been mitigated in Figure 3.16c.

Acoustic FWI of the matched data yields the P-wave velocity model in Figure 3.17e. In comparison to the model obtained after acoustic FWI of elastic data, the data-matching workflow leads to a more detailed and more accurate result, in which layers are again clearly defined, the fault zone (black ellipse) is more clearly differentiated and the low-velocity channel at $800\,m$ depth is flat and in the correct position. All these improvements result in a reduction of the average model RMS down to $230.49\,m/s$, close to that obtained for the acoustic inversion of acoustic data.

To further assess the benefits of the data-matching workflow on land data I perform three quality checks. First, I extract and compare two well-logs at $3\,km$ and at $7.2\,km$ position for all the models in Figure 3.17, and I display the profiles in Figures 3.18a and 3.18b, respectively. These correspond to the overthrust region (Figure 3.18a) and away from the overthrust (3.18b).

Both profiles in Figure 3.18 show that acoustic FWI of acoustic data (blue line with stars) provides and improvement with respect to the starting model and that acoustic FWI elastic data (red line with filled circles) results in less accurate recovered profiles than the former, especially in the regions indicated with black arrows (both solid and dashed arrows). The difference in accuracy between the latter two inversions with respect to the true model reaches its maximum in the profile at $7.2\,km$ position between $0.5 - 0.65\,km$ depth (black continuous arrow in Figure 3.18b), which corresponds to the push-down observed in the low-velocity channel in Figure 3.17d. On the contrary, the profile corresponding to the acoustic FWI of matched data (green line with filled squares) is closer to the results of acoustic FWI of acoustic data, and hence to the true model. These improvements can be observed throughout the profiles, especially

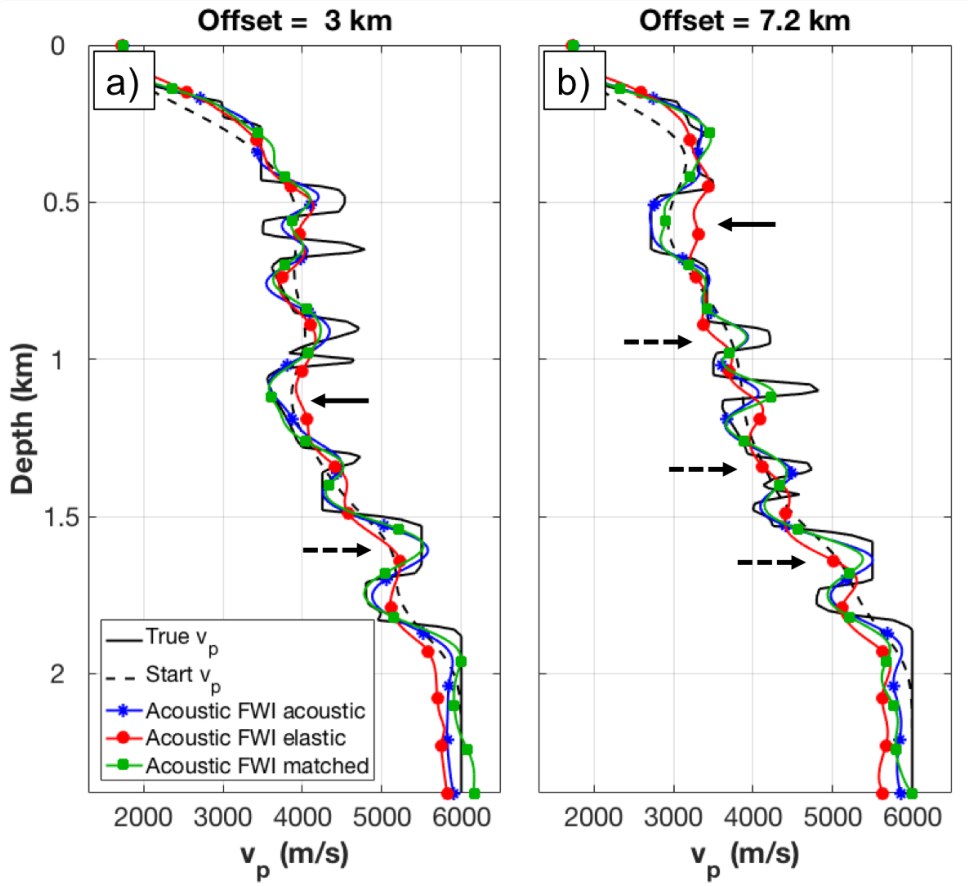


Figure 3.18.: Vertical profiles for the models in Figure 3.17 at a) 3 km and b) 7.2 km position.

where indicated by the arrows: this includes the correction in terms of velocity of the low-velocity channel in Figure 3.18b (solid black arrow).

The second test to evaluate the quality of the results is to compute the average model RMS difference as a function of depth and position within the dotted box in Figure 3.17 for the recovered and the starting models. This is captured in Figure 3.19, which shows the benefit of using the data-matching workflow as the values associated to the latter (gray line with filled squares) in terms of position and in depth are smaller than those obtained for acoustic FWI of the elastic data (thick black line with filled circles), and they are close – but not smaller, as expected– to the values for acoustic data (thin black line).

Finally, I compute the average model RMS difference within the dotted box per

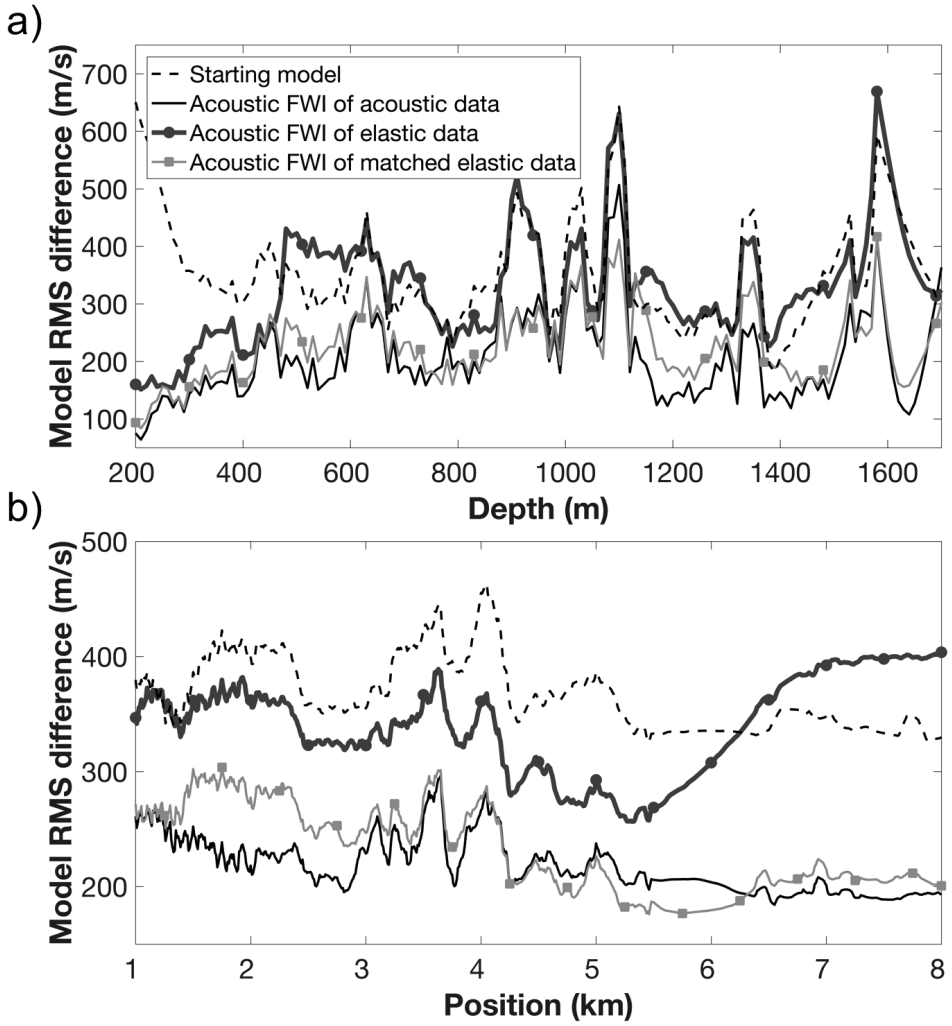


Figure 3.19.: Average model RMS differences for each recovered model and the starting model in Figure 3.17 as a function of a) depth and b) position within the dotted box in Figure 3.17.

iteration during the inversions, and the result is presented in Figure 3.20.

This graph clearly shows that the accuracy of the recovered model obtained with the data-matching worklfow improves as the inversion algorithm iterates, and that model differences are significantly smaller than for the acoustic inversion of elastic data for all iterations. Moreover, the decrease in the model differences with respect to the true model per iteration for the elastic data is less pronounced than the decrease observed for the matched observed data – hence, the acoustic inversion of the matched data converges faster to the true model. Figure 3.20 also shows that the recovered P-wave velocity model with the proposed strategy is never as good as that obtained

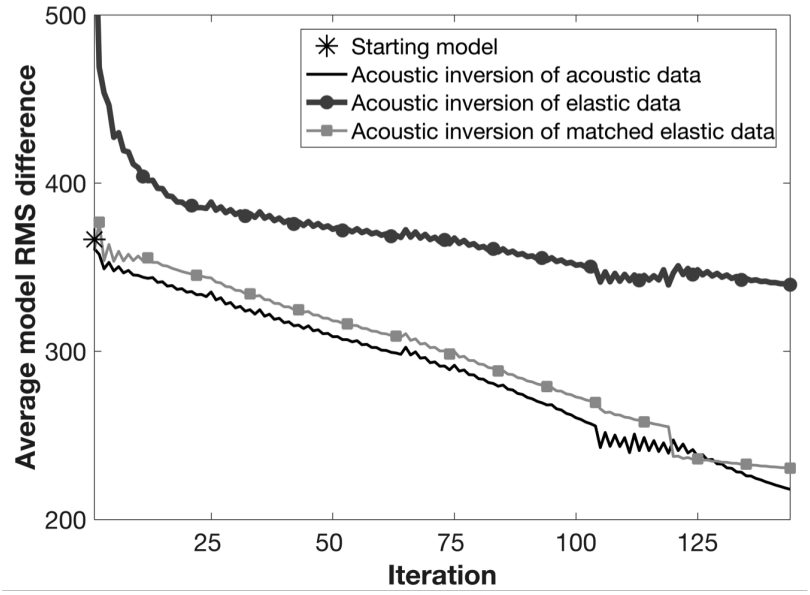


Figure 3.20.: Graph of the average model RMS differences within the dotted area in Figure 3.17 as a function of iteration for the acoustic inversion of acoustic, elastic and matched observed data.

when the dataset is acoustic, as the latter is the ideal case.

Consequently, the previous tests carried out in this section suggest that the use of an amplitude-balanced functional is not enough to mitigate elastic effects in acoustic FWI of land data, and that the proposed data-matching workflow can be used to address elasticity and provide recovered P-wave velocity models that have significantly better resolution and accuracy for such datasets.

3.5. Application to a 3D marine field dataset

As shown in the previous sections, elastic effects in marine settings are generally much weaker than in land datasets. However, this is not always the case, and elastic effects in certain marine environments can significantly deteriorate the quality of the recovered P-wave velocity models obtained with acoustic FWI of hydrophone-recorded data. It is on these datasets that the data-matching approach can be most beneficial.

This is illustrated here for a marine 3D field dataset acquired in the North Sea, over the Buzzard field, where a dense ocean-bottom-node 3D survey was carried out in 2014. From top to bottom, the Buzzard field is formed by a 1 km-thick unit of

young clastic rocks (mainly from late-Paleocene/early-Eocene), followed by a 500 m-thick chalk formation, then interbedded sandstones and mudstones and finally the reservoir, which is composed of deep marine turbidities (Doré and Robbins, 2005; Ray et al., 2010; Stronge, 2018). As a result of the large P- and S-wave impedance contrasts at the top-chalk interface and, more precisely, due to ignoring elastic effects caused by the shallow chalk unit, acoustic FWI fails and results in an unrealistic recovered P-wave velocity model of the subsurface.

Here I show the initial tests of using the data-matching approach to recover a P-wave velocity model that is more accurate at the top-chalk interface and within the chalk unit. Nevertheless, the results presented here are still preliminary, and further research is needed.

First of all, the importance of considering elastic effects for this dataset is reviewed. Figures 3.21a and 3.21b show the same shot gather for the field data and acoustic modelled data obtained using the starting model and neglecting elastic effects. Unlike for typical marine datasets, the shot gather for field data seems simpler than the acoustically modelled shot gather, and the latter contains additional reflections that interfere with the refracted events, and which are not present in the field data.

These is due to the post-critical P-wave reflection at the top-chalk interface, and its associated multiples. The latter are absent in the field data, even after removing direct arrivals applying an F-K filter (Figure 3.21c), but they are present in the acoustic modelled data, as it can be seen in the F-K filtered shot gather in Figure 3.21d. Moreover, the amplitude of the P-wave post-critical reflection for field data in Figure 3.21c is more attenuated with offset than that in the acoustic modelled data in Figure 3.21d. Thus, the hypothesis is that these differences are caused by elasticity.

To validate whether the attenuation of the post-critical P-wave reflection at the top-chalk interface and its associated multiples is due to elastic effects, the reflection coefficient and the phase of the P-wave reflections at the top-chalk interface with offset are computed in an acoustic and an elastic model that simulate the shallow structure of the Buzzard field (Stronge, 2018). The graphs in Figure 3.22 illustrate that the reflection coefficient of the reflection at the top-chalk interface slowly increases

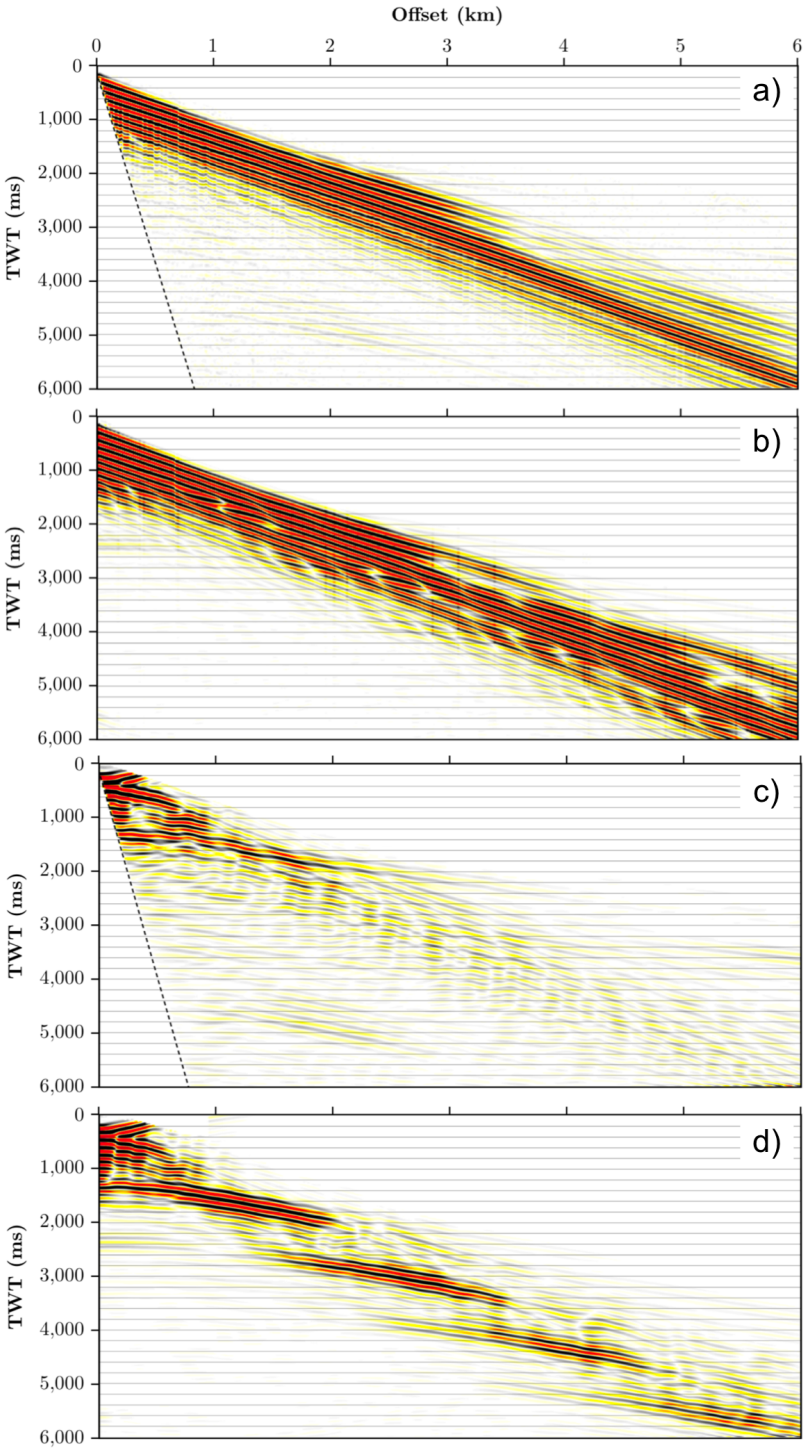


Figure 3.21.: Shot gathers for a) the observed data and b) the acoustic modelled data obtained with the contractor's starting model for the Buzzard field dataset. Panels c) and d) display the F-K filtered observed and acoustic modelled gathers, respectively. Panel d) illustrates the presence of the post-critical reflection at the top-chalk and its multiples when elastic effects are negligible. Modified from (Stronge, 2018).

from zero up to 1150 m offset in both scenarios, which corresponds to the critical distance, after which the reflection becomes post-critical. At this point, the reflection coefficients stays constant at a value of one in the acoustic case (Figure 3.22a), but decreases for the elastic case (Figure 3.22b) to a value of zero, and then steadily increases with offset until it reaches a coefficient similar to that at near-offsets.

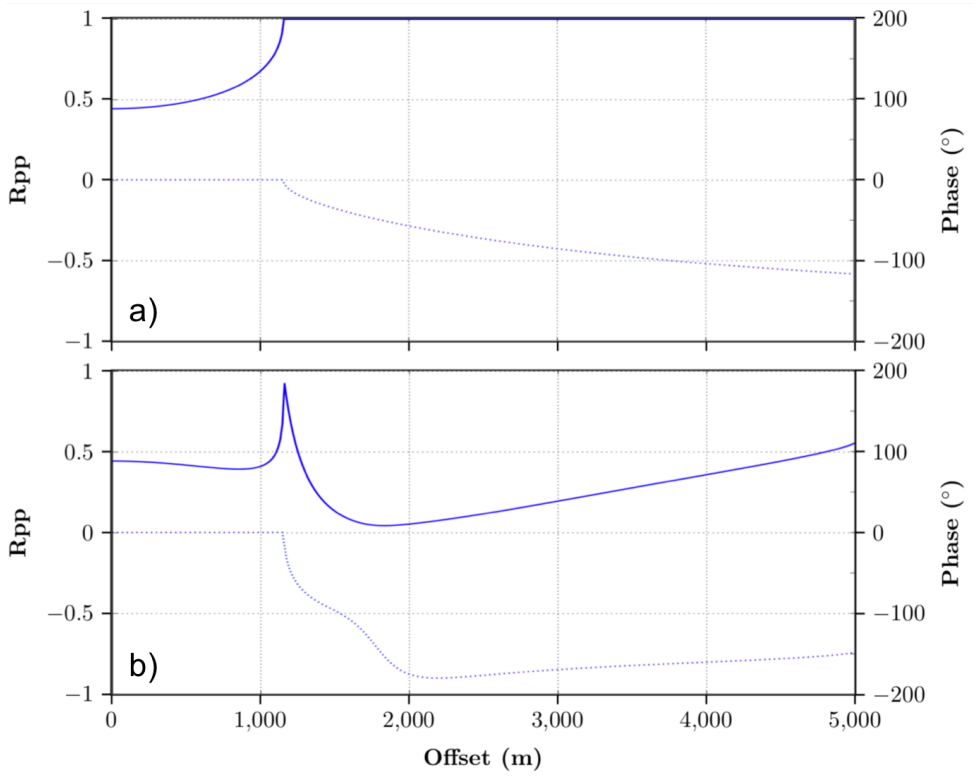


Figure 3.22.: Reflection coefficients (solid line) and phase profiles (dashed lines) at the top-chalk of the Buzzard dataset for a) an acoustic and b) an elastic model, obtained using Zoeppritz equations for isotropic interfaces. Modified from (Stronge, 2018).

As a consequence, post-critical reflections are strong in the acoustic data, but are much weaker for elastic data. This explains the amplitude behaviour with offset for the post-critical reflection and its multiples in the infinite frequency limit, given that the Zoepritzz equations use ray theory to model wave propagation. To demonstrate that the latter is also valid when considering the finite-frequency nature of seismic waves, acoustic and elastic data are modelled using independent modelling codes that solve the acoustic and elastic wave equations for the entire model.

These modelled data are also used to mitigate elastic effects by applying the data-matching workflow. Given that elastic effects are strong due to the shallow high impedance contrast at the top-chalk interface, acoustic FWI fails (Stronge, 2018), and the data-matching workflow is applied here directly by using the starting P-wave velocity model and an estimate of the v_p/v_s model.

For the current tests, I design a v_p/v_s model based on the regional geology and not on measurements within the area, as these were not available. A more accurate v_p/v_s model could be re-designed in future tests by incorporating well-log information. The resulting v_p/v_s model has a value of 1.8 within the chalk (Hamilton, 1979; Bhakta and Landro, 2013), of 3 for the unit of clastic rocks above the chalk (Hamilton, 1979) and of 2 everywhere else. S-wave velocity is set to zero within the water layer and, to reduce the computational cost, v_p/v_s is set to 2.5 for cells in the model with P-wave velocity between the water velocity and $1600\,m/s$, which mainly correspond to cells close to the sea-bottom. Higher values of v_p/v_s should be expected close to the sea-floor, reaching values as high as 13 for soft sediments at the sea-floor (Hamilton, 1979). However, large values of v_p/v_s represent a high computational cost, as the grid spacing needs to be refined to account for slower S-wave velocities and have little energy dispersion while modelling. As a consequence, elastic effects close to the sea-bottom might not be properly considered.

The resulting 3D P- and S-wave velocity models for acoustic and elastic modelling have dimensions of $91 \times 364 \times 378$ and $181 \times 727 \times 755$ in cells, respectively, with an associated grid spacing of $50\,m$ and $25\,m$ to avoid having dispersion of P- and S-waves. Density is computed from the P-wave velocity model using Gardner's law. Data are modelled for 973 pressure sources and recorded with a total of 42930 receivers, recording $6\,s$ of data with a time sampling of $4\,ms$ and of $2\,ms$ for acoustic and elastic modelling, respectively, in order to have numerical stability. Once elastic modelling is complete, the resulting data are resampled to $4\,ms$ using a sinc function, which takes into account the oscillatory nature of the seismic signal.

The data-matching workflow is then applied by matching the elastic to the acoustic

data with filters of 75 ms length, and matching 3 traces simultaneously – so that the spacing between the first and the last matched traces is 150 m. The computed filters are then convolved with the observed data to produce a matched dataset, in which elastic effects are mitigated.

A comparison for a shot gather for the field data, the acoustic and the elastic modelled data and the matched data is shown in Figure 3.23.

The presence of strong post-critical P-wave reflections in the acoustic data in Figure 3.23b, which are strongly attenuated in the elastic data in Figure 3.23c, demonstrates that this is an effect that is also observed when considering finite-frequency wave propagation. As a result, the elastic data is closer to the observed data, both in terms of amplitude but also because elastic data does not contain strong post-critical P-wave reflections. Nevertheless, slow surface waves are present in the elastic data, which have been previously muted on the observed dataset. These do not have a significant impact when applying the data-matching workflow, as the matched data is set to zero within regions in which the observed data is also zero. Finally, the gather for the matched data in Figure 3.23d illustrates the effect of removing elasticity from field data, so that post-critical P-wave reflections are incorporated and the amplitudes of the different arrivals are modified.

Next, anisotropic 3D FWI – using the contractor's anisotropic parameters (Stronge, 2018)– is performed both on the observed and on the matched datasets. The inversions are performed using a multi-scale approach by inverting data from 3 Hz up to 10.4 Hz in a total of 140 iterations, using only 98 sources per iteration and an amplitude-balanced functional to further mitigate elastic effects. This results in a compute time of 6 days and 10 hours using 11 nodes and 3 tasks per node.

I note that elastic FWI would be significantly more expensive for this dataset, given that 3D elastic FWI when compared to 3D acoustic FWI represents an increase in computational time of a factor of two when using the same mesh (Guasch, 2012). But, in order to account for slow S-wave velocities, the mesh used for elastic FWI should be reduced to at least half the grid size for acoustic inversions. Thus, 3D elastic FWI when compared to 3D acoustic FWI with a grid spacing of half the spacing used for the acoustic inversion would result in a runtime that is at least

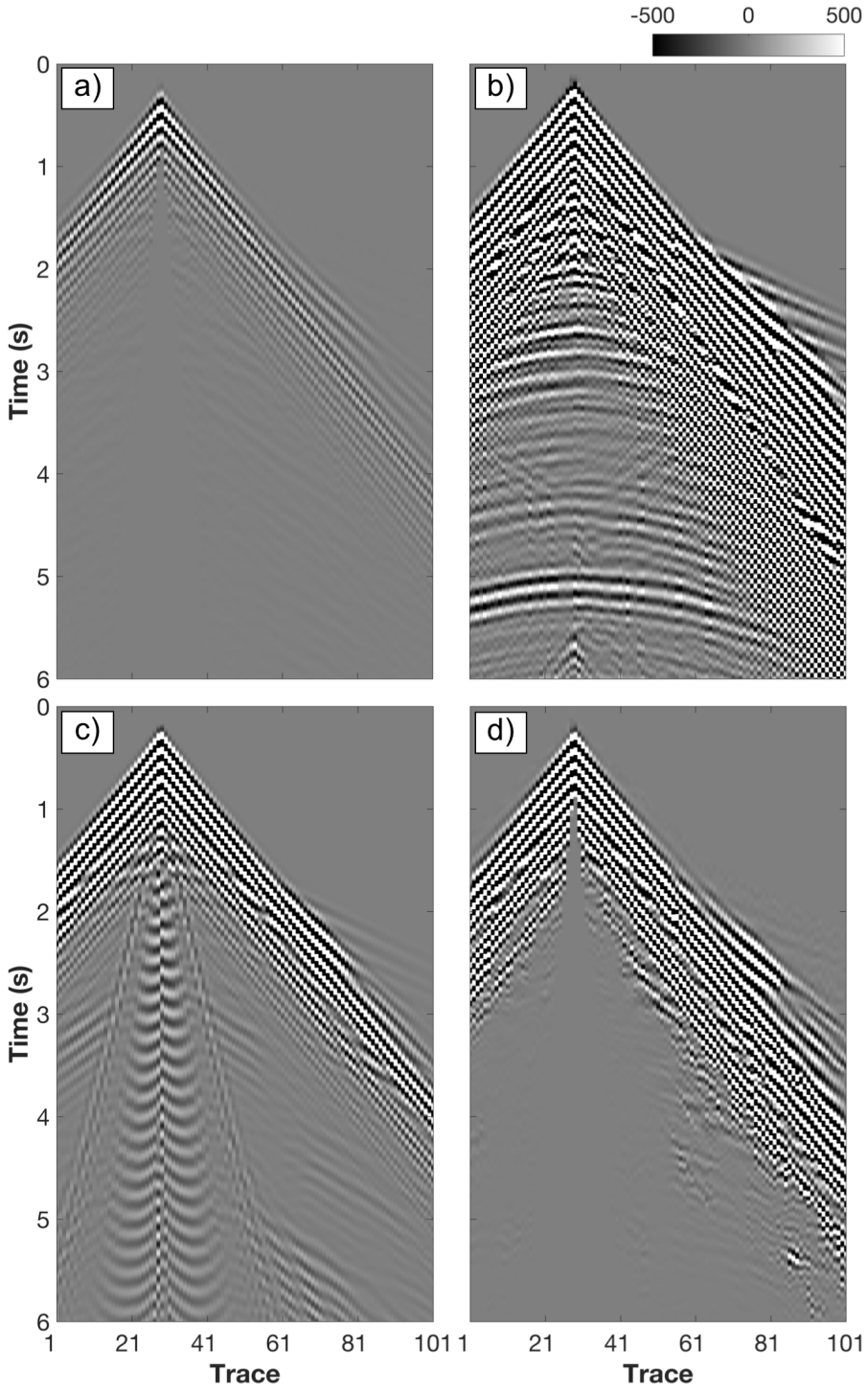


Figure 3.23.: Representative shot gathers for a) the observed data, b) modelled acoustic data, c) modelled elastic data and d) matched observed data for the Buzzard dataset. The starting P-wave velocity model and an ane estimate of the v_p/v_s model based on the regional geology are used to generate synthetic data.

$2^5 = 32$ times longer[1], and a difference in memory of $9/2 \times 2^4 = 72$ times the storage required for an acoustic inversion[2]. Therefore, elastic FWI for this dataset would most likely be unaffordable nowadays. On the contrary, the data-matching workflow allows mitigating elastic effects at a much affordable cost, which is similar to that of a single acoustic inversion here.

Figures 3.24 and 3.25 respectively show the results obtained for a cross-line and an in-line profiles. Each of these figures displays the starting P-wave velocity model, the estimated v_p/v_s models and the respective recovered P-wave velocity models obtained from 3D anisotropic acoustic FWI of the observed and matched datasets.

The first and most important observation from these figures is that acoustic FWI of the observed data results in a model in which the high-velocity chalk layer is dramatically thinned, and velocities are pushed down as if the chalk layer did not exist. This is due to the large impedance contrast at the top-chalk and not correctly accounting for elastic effects. More precisely, the modelled data during acoustic FWI contains strong P-wave post-critical reflections, but these are not present in the observed data. Thus, in order to reduce the data misfit, the chalk layer is removed in acoustic FWI, resulting in modelled data that is more similar to the observed data, as it does not contain P-wave post-critical reflections. However, this leads to an unreliable recovered model that does not contain a continuous chalk unit. Therefore, just using an amplitude-balanced functional is not sufficient to mitigate elastic effects for this dataset.

On the other hand, the recovered velocity model obtained from acoustic inversion of matched data preserves the chalk layer given that post-critical P-wave reflections at the top-chalk are present in the matched data as well as in the modelled data. As

[1]By reducing the grid spacing to half the value used for acoustic inversion, the number of cells increases by a factor of 2 in each of the three space dimensions, by a factor of 2 in time (to avoid numerical dispersion) and by another factor of 2 as elastic inversion takes roughly twice the time as an acoustic inversion when computed using the same mesh (Guasch, 2012). Overall, the runtime of elastic 3D FWI is 2^5 times larger than that of acoustic 3D FWI when the grid spacing used for the latter is double the size used in the former.

[2]Reducing the grid size to half the original size means an increase in storage of the wavefield by a factor of 2 for each spacial dimension, a factor of 2 in time and a factor of 9/2 given that 9 variables need to be stored in the elastic case (first order PDE in which velocity, stress and strain need to be stored at the current time) (Guasch, 2012), whereas only pressure at the current and previous times need to be stored in the acoustic case (second order PDE in equation (2.20)).

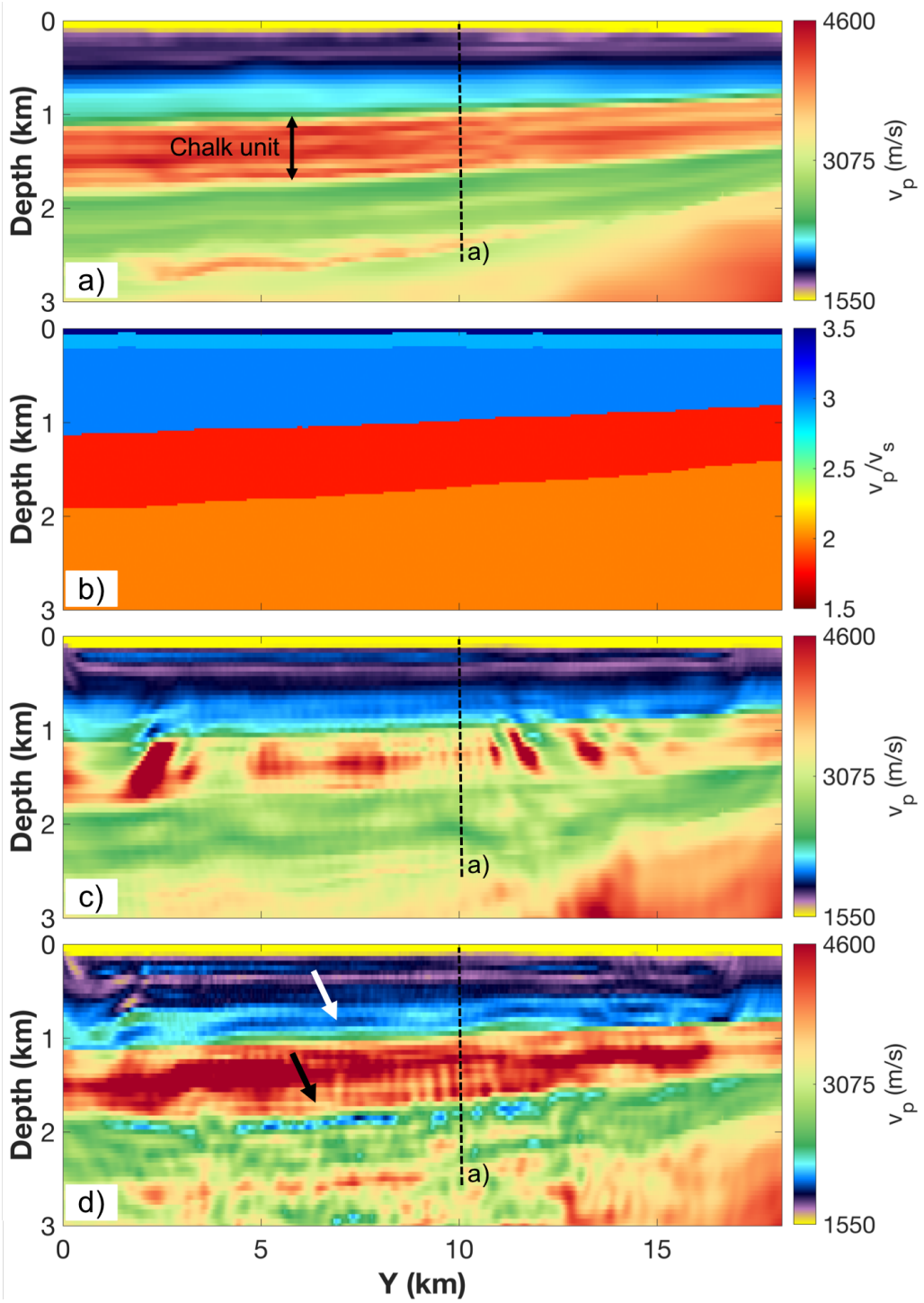


Figure 3.24.: Vertical cross-line profiles at $X = 10.95\,km$ for the a) P-wave starting model, b) the estimated v_p/v_s and the recovered P-wave velocity models obtained with anisotropic 3D acoustic FWI of c) the observed and d) the matched datasets. The discontinuous line indicates the position in which a velocity profile is extracted. The chalk unit is dramatically thinned in pannel c), but is preserved in d).

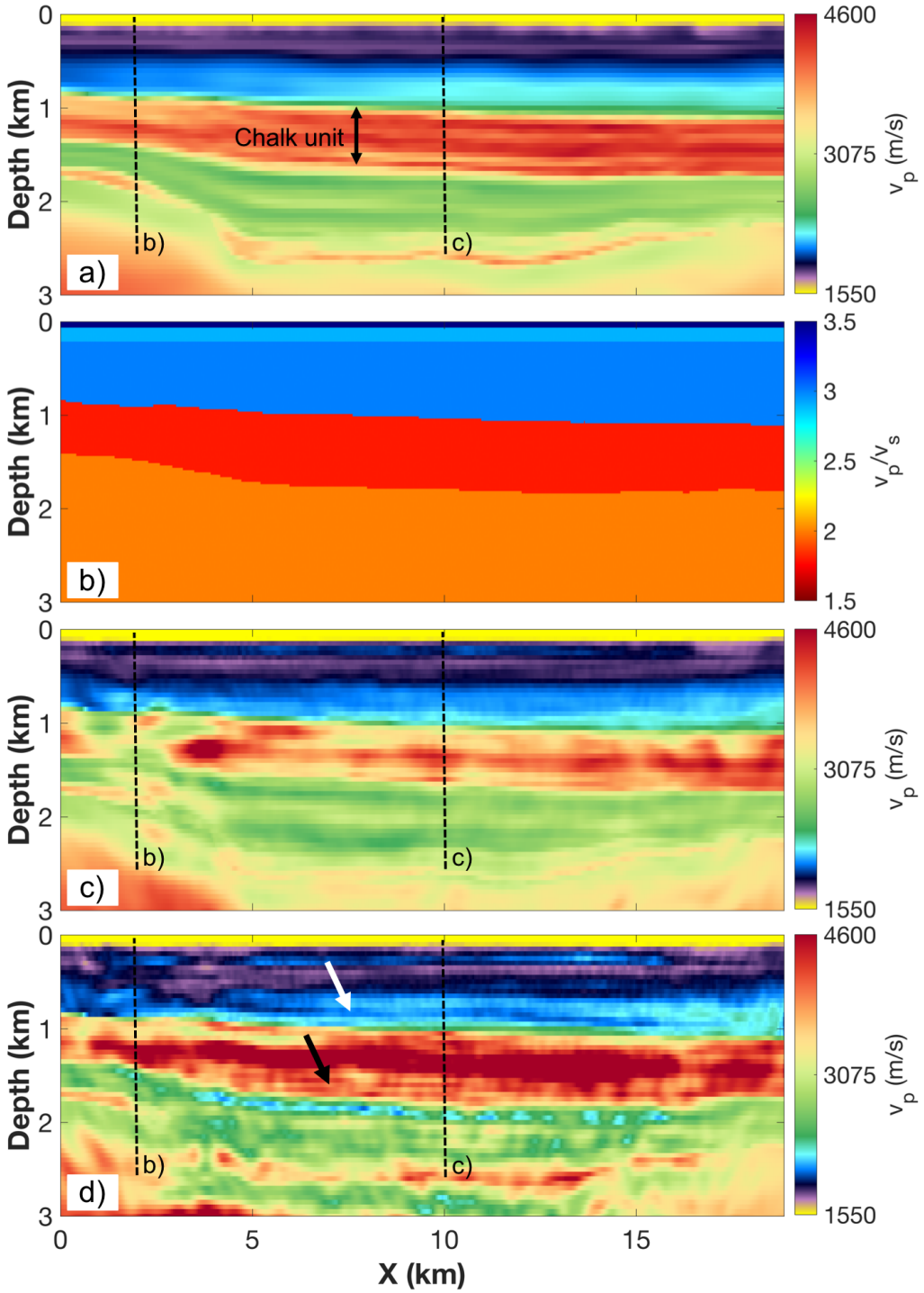


Figure 3.25.: Vertical in-line profiles at $Y = 5.95\,km$ for the a) P-wave starting model, b) the estimated v_p/v_s and the recovered P-wave velocity models obtained with anisotropic 3D acoustic FWI of c) the observed and d) the matched datasets. The discontinuous line indicates the position in which a velocity profile is extracted. The chalk unit is dramatically thinned in pannel c), but is preserved in d).

a result, the P-wave velocity within the chalk layer is updated so that finer structure can be observed for the cross-line and in-line profiles when compared to the starting model (indicated by black arrows in Figures 3.24d and 3.25d). P-wave velocities are also increased in some regions up to values of $5300\,m/s$ in average (the maximum allowed velocity during the inversions is $5900\,m/s$). Additionally, some improvements can also be seen right above the top-chalk in the recovered model obtained for the matched data, in which a continuous layer with velocities between $1700\,m/s$ and $1800\,m/s$ is recovered (white arrow).

However, the recovered model along the cross-line profile in Figure 3.24d is more noisy, suggesting stronger smoothing is required along this direction. The latter was expected, given that data-matching has been performed by matching multiple traces simultaneously along the in-line direction and, therefore, no smoothing along the cross-line direction has been applied when matching the data.

To further assess the differences between the recovered models obtained with the observed and the matched datasets, three velocity profiles are extracted from the starting and recovered models in Figures 3.24 and 3.25, as denoted on these models by the discontinuous vertical lines. Figure 3.26 displays the velocity profiles extracted from these models.

The velocity profiles in Figure 3.26 further illustrate how the chalk layer is thinned in acoustic inversion of observed data (red line), whereas it is preserved and updated after acoustic inversion of the matched data, resulting in higher speeds and a more clear differentiation of layers within the chalk unit.

Nevertheless, the accuracy of the updates has not yet been compared to well-logs, or assessed by investigating the flatness of migrated gathers and the match of the recovered model with the main events in the migrated stacked section. Hence, tests to assess the quality of the recovered models should be performed in the near-future.

For instance, the quality of the recovered model obtained for the matched data for the first $400\,m$ below the sea-floor is less reliable than that obtained for the observed data. This can be observed in the depth slices at $300\,m$ depth in Figures 3.27a-c,

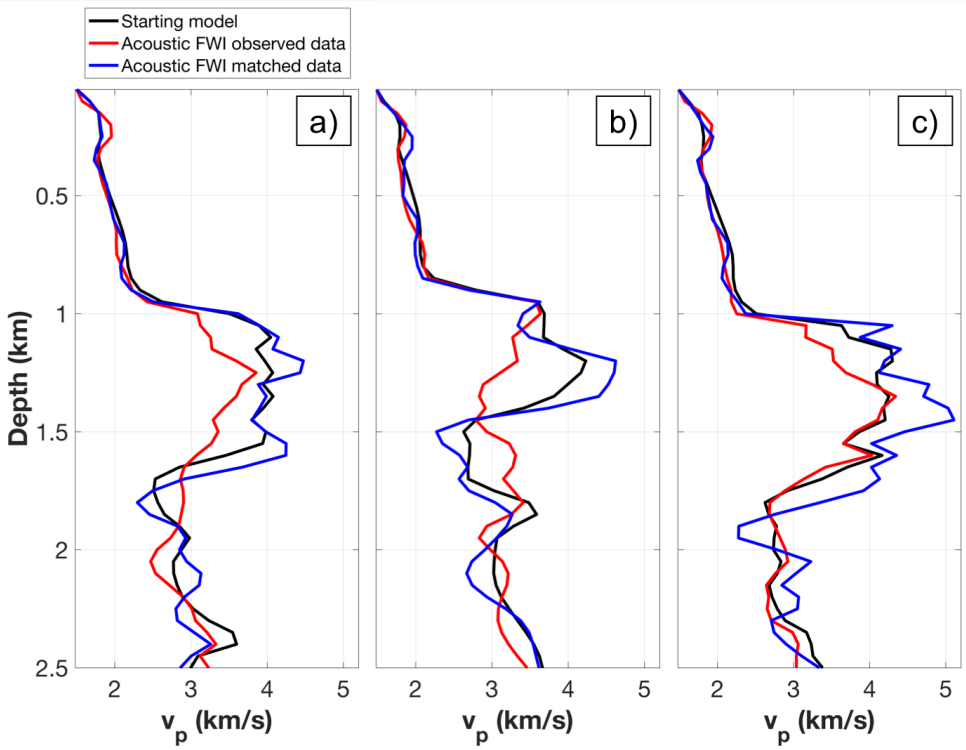


Figure 3.26.: Velocity profiles for the starting model (black line) and the recovered P-wave velocity models obtained after anisotropic 3D acoustic FWI of observed data (red line) and matched data (blue line) for the well-log positions indicated in Figures 3.24 and 3.25.

and also in the lack of continuity of layers within the first 400 m in Figures 3.24d and 3.25d.

Whereas a structure that may be a channel can be observed in the depth slice obtained for the observed data in Figure 3.27b, this is hardly visible in the depth slice for the matched data in Figure 3.27c. Possibly, this is due to not properly accounting for the slow S-wave velocities at the sea-floor in elastic modelling, which negatively affects the quality of the recovered model within the first 400 m. Moreover, the S-wave velocity model is not properly designed close to the sea-floor and within the water layer, so that non-zero S-wave velocities are incorrectly assigned to certain regions of the model within the water layer. This is shown in Figure 3.28 for a depth slice at 50 m depth.

Most likely, this was caused when mildly smoothing the final S-wave velocity model, and incorrectly setting the S-wave velocity within the water layer to zero after smoothing. For this reason, future tests should also take this into account.

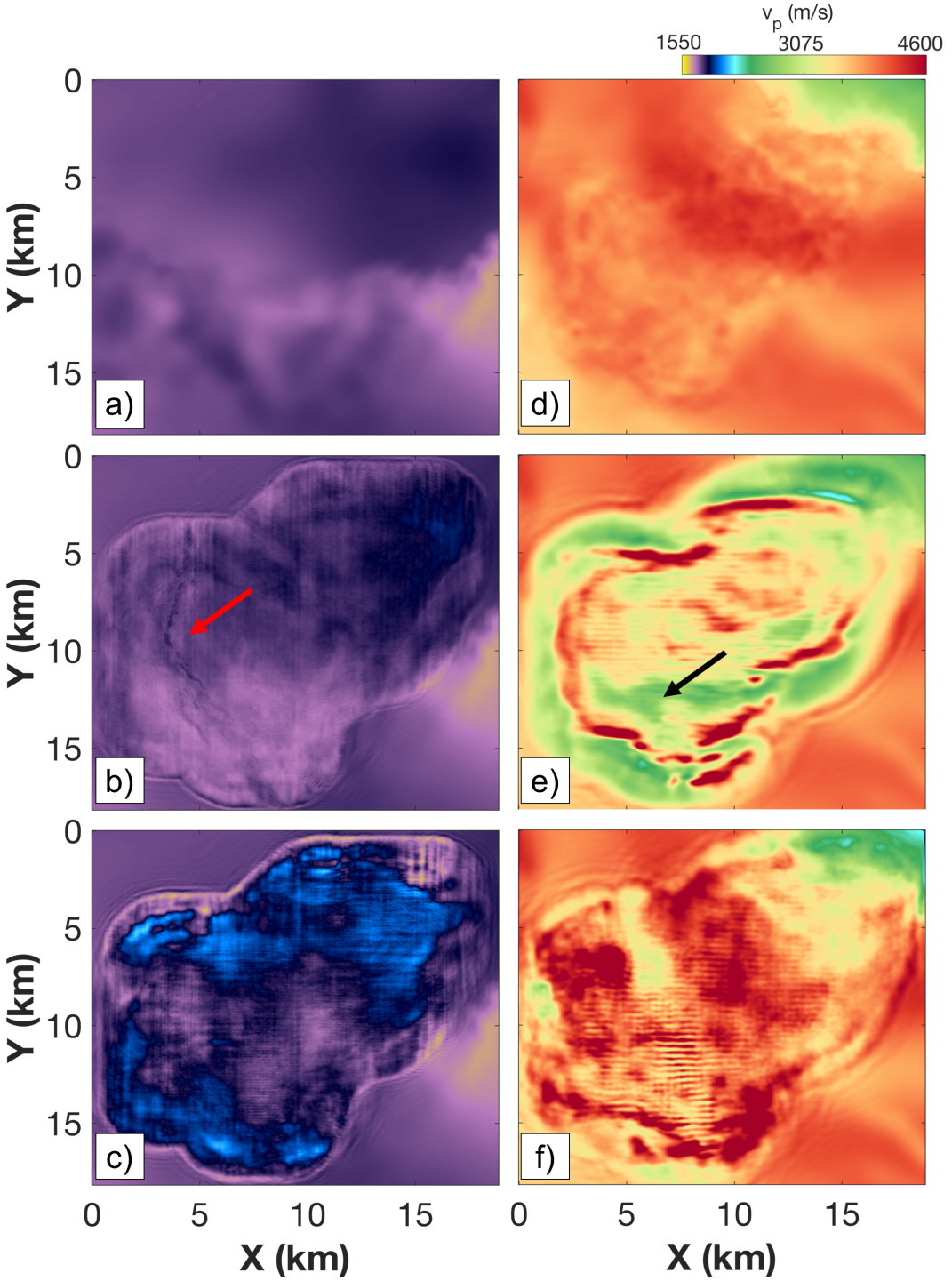


Figure 3.27.: Depth slices for the starting P-wave velocity model (first row) and the recovered models obtained after anisotropic 3D acoustic FWI of observed data (second row) and matched data (third row) at 300 m depth (first column) and at 1150 m depth (second column).

On the other hand, the depth slices at 1150 m in Figures 3.27d-f illustrate again that velocities are incorrectly pushed down within the chalk layer when using the ob-

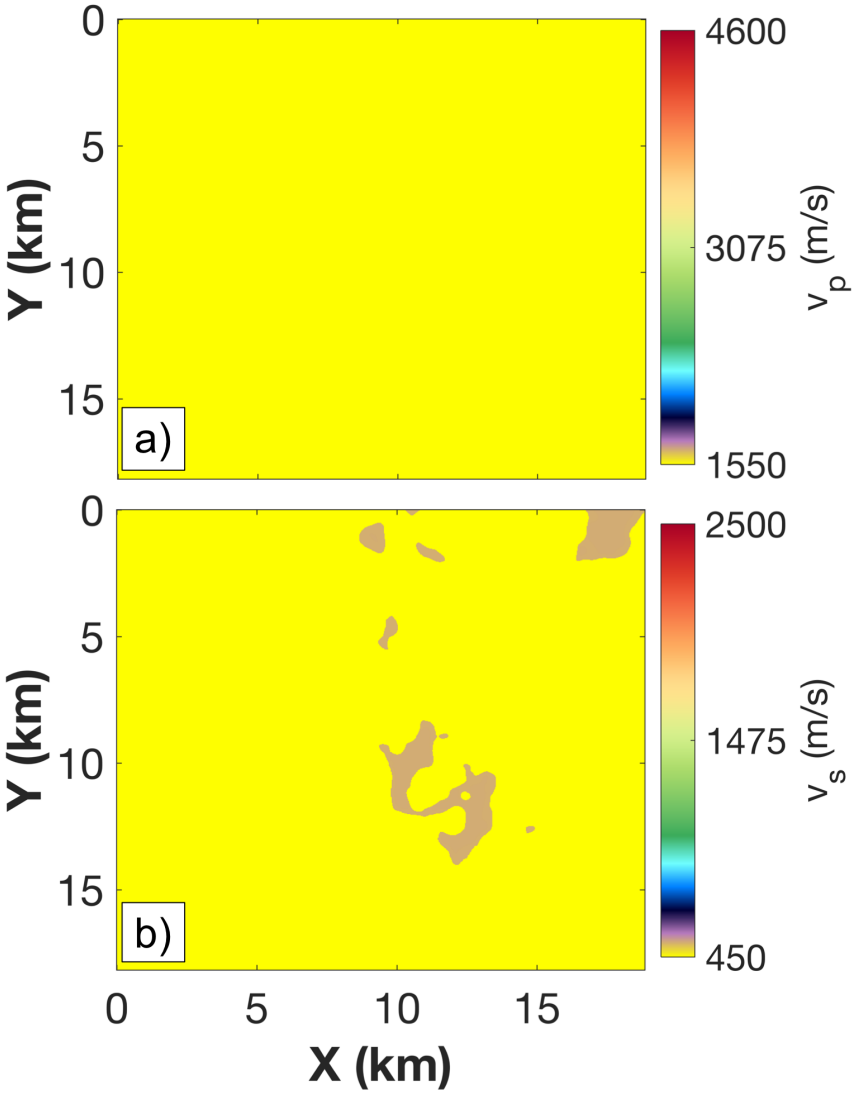


Figure 3.28.: Depth slices at 50 m depth for the a) P-wave and b) S-wave velocity models used for acoustic and elastic modelling. S-wave velocity is incorrectly set to non-zero values within certain regions in the water layer.

served data (Figure 3.27e), but these are correctly preserved and also updated when using the matched data (Figure 3.27f).

Overall, the previous preliminary tests on the 3D Buzzard field dataset illustrate that elastic effects can also be relevant in marine datasets in which layers of large impedance contrasts are present at the shallow subsurface, and that the data-matching method can be used to improve the quality of the recovered models in these situations with a cost that is much lower than a full elastic inversion. The results

also suggest that more accurate information of the v_p/v_s model at the sea-floor is required, as well as a better design of the v_s model within the water layer, and future tests should therefore benefit from incorporating well-log information of the shallow subsurface.

3.6. Summary

In this chapter I have shown that ignoring elasticity in conventional acoustic FWI has an adverse effect on the quality of the recovered P-wave velocity models of the subsurface, leading to a visible loss in resolution and accuracy when compared to the results of acoustic FWI of acoustic data, the latter being the ideal case. In practice, the subsurface exhibits an elastic behaviour and techniques designed to address elastic effects are necessary to obtain high quality velocity models of the subsurface. In general, elastic effects are less significant on marine datasets and more relevant for land datasets, in which mechanical waves are generated within a region of non-zero S-wave velocity. However, elasticity also deteriorates the quality of the recovered P-wave velocity models on typical marine datasets obtained with acoustic FWI, and can lead to geologically non-plausible results under certain circumstances – for instance, when a chalk layer with low S-wave velocities is close to the sea-bottom and its elastic component is not considered. Standard practice methods to mitigate these effects, such as the amplitude-balanced objective functional, are not sufficient in these situations.

In order to address these adverse effects and obtain higher quality recovered P-wave velocity models while performing acoustic FWI, I have proposed a method that is based on Wiener filters, and which results in a matched observed dataset that contains less elastic effects and, hence, is more suitable for acoustic FWI. I have demonstrated that this methodology provides significant improvements in the recovered P-wave velocity model with respect to ignoring elasticity at a cost that is lower than a full elastic inversion –typically between two to three times the cost of acoustic FWI– and, thus, it is well-suited to address elasticity on large 3D datasets, as demonstrated on a field dataset.

Nevertheless, several aspects need to be addressed for a successful implementation of the data-matching workflow. First, this workflow involves computing Wiener filters that match modelled elastic to modelled acoustic data, which can have very different spectra. This is because elastic data contains S-wave related events which are not present in the acoustic data, and the amplitudes of P-wave arrivals are different due to energy splitting differently at interfaces. This may result in localised artefacts – generally seen as ripples– in the matched data. In order to minimise their impact on the data, I have proposed to use overlapping short filters in time and compute a single Wiener filter that matches multiple traces simultaneously. The former introduces smoothing along a trace, reduces the computation cost and ensures small portions of the data are matched, which are more likely to have similar spectra. The latter introduces a certain degree of lateral smoothing across neighbouring traces. I have demonstrated on synthetic data that these two measures improve the quality of the matched data by mitigating the artefacts introduced by matching two very different traces. After several tests on synthetic and field data I recommend to use filters with a length of between $300\,ms$ and $400\,ms$, and compute single Wiener filters that match a set of neighbouring traces such that the spacing between the first and last matched traces is between $100\,m$ and $800\,m$.

A second aspect that needs to be considered when using the suggested data-matching workflow is that this requires an estimate of the velocity ratio v_p/v_s, and this is typically not known with good accuracy. However, the tests I have performed on marine synthetic and field data indicate that the estimate of v_p/v_s does not need to be known very accurately as long as it captures the most important elastic effects, typically due to the large impedance contrasts for S-waves at the sea-bottom or the impedance contrasts at other interfaces (e.g. sediments and chalk). The latter can be obtained from additional observations, such as shallow well-logs or knowledge of the regional geology. However, further research should focus on methods of estimating v_p/v_s from the data. These may include global inversion methods with only a few parameters – given that only a smooth and approximate estimate of the velocity ratio is necessary– prior to the application of the data-matching workflow.

Finally, a third aspect to consider is the limitations of the data-matching workflow,

and how can these be minimised. I have shown that the data-matching workflow – and also acoustic FWI with an amplitude-balanced functional– fails when addressing strong elastic effects if the starting P-wave velocity model for the inversion is poor, as acoustic FWI of the observed data leads to a result that is cycle-skipped. This introduces a limitation on the data-matching workflow as the latter uses the recovered P-wave velocity model to mitigate elastic effects. However, I have suggested an alternative workflow that overcomes this problem by utilising the starting P-wave velocity model directly to model acoustic and elastic data, and computation of matched observed data prior to any inversion – hence skipping Step 1 of the workflow. Unlike the original workflow, the latter leads to a very accurate recovered P-wave velocity model at a cost slightly larger than conventional acoustic FWI in those cases where elastic effects are strong and starting models are poor.

Although the data-matching workflow has been used in this chapter on pressure data, application to multi-component data is in principle possible by modelling the components of particle velocity in Step 2 of the workflow. This will be further explored in the future.

One of the benefits of the data-matching method when addressing elastic effects in large 2D datasets or in 3D datasets is that all inversions are acoustic. Therefore, the grid spacing and time sampling do not need to be as small as for elastic inversions, in which elastic waves need to be modelled. This is mainly because S-waves generally have much slower velocities than P-waves, and time sampling and grid spacing need to be reduced to avoid energy dispersion and instability issues during modelling for each iteration in the inversion. Instead, elastic modelling is only required once in the data-matching workflow, prior to an acoustic inversion. For this reason, acoustic modelling can be performed on a coarse grid, whereas elastic modelling can be performed using a finer grid followed by resampling of the data to match the time sampling of the acoustic and observed data. I have used this strategy to mitigate elasticity on the 3D field dataset, and I recommend it to further reduce the costs of the algorithm.

In fact, the application to the 3D marine field dataset demonstrates that accounting for elasticity can be of great importance for marine datasets with a geological

structure that contains large impedance contrasts at interfaces and that is in the shallow subsurface, as for example a clastic-chalk interface. I have shown that in the Buzzard field dataset example, acoustic FWI of the observed data fails within the chalk layer, incorrectly pushing its P-wave velocity down and leading to a dramatic thinning of the chalk unit. The application of the data-matching workflow shows promising results in mitigating elastic effects, such that the chalk layer is preserved during acoustic inversion of the matched data. Nevertheless, these results are still preliminary, and further tests need to be performed using a more accurate estimate of v_p/v_s, possibly obtained from nearby well-logs.

In fact, elastic effects are very strong for this dataset. Thus, what I propose is to use a new strategy based on the results in this chapter, which would consist of converting predicted acoustic data to predicted elastic data at each iteration, and performing elastic inversion in a coarse grid. In order to avoid performing elastic modelling at each iteration, which would considerably increase the computation cost, I suggest using trained neural networks to obtain elastic data from modelled acoustic data. This, however, requires further research.

Alternative applications of machine learning methods would be automatic computation of Wiener filters provided an acoustic and an elastic datasets, or automatic generation of the elastic wavefield from an acoustic wavefield at each iteration of the inversion, thus enabling matching of the wavefields at each inversion step. Although the latter may result in better results, it could potentially increase the total computational burden as computation of Wiener filters at each inversion step would represent a large associated cost.

Other applications of the data-matching workflow also need to be explored, for instance to mitigate other unwanted amplitude effects not typically accounted for in acoustic FWI. This includes accounting for attenuation, which is generally disregarded in acoustic FWI. I explore this possibility in the next chapter by slightly modifying the workflow used to mitigate elastic effects to account for attenuation, which requires an approximate estimate of Q. I explain how the latter can be achieved and I successfully apply the method to synthetic data and a marine field dataset.

4. Acoustic FWI in a visco-acoustic earth

Similarly to elasticity, viscous effects are typically neglected in conventional acoustic full-waveform inversion (FWI). This has negative consequences on the quality of the recovered P-wave velocity models of the subsurface due to the viscous nature of the earth. Here I extend the methodology used in the previous chapter to address elastic effects in order to account for viscous effects, in particular visco-acoustic effects.

The research work described in this chapter has been recently presented at the 79th EAGE Conference and Exhibition in Paris:

> Agudo, O. C., Vieira da Silva, N., Warner, M., and Morgan, J. (2017). Addressing viscous effects in acoustic full-waveform inversion. In *79th EAGE Conference and Exhibition*.

Moreover, I have also submitted a related manuscript to a scientific journal, and this is currently under review:

> Agudo, O. C., da Silva, N. V., Warner, M., Kalinicheva, T., and Morgan, J. (2018a). Adressing viscous effects in acoustic full-waveform inversion. *Geophysics*, (Under review).

In this chapter I describe the main results contained in the publications above, as well as additional tests that complement the latter.

I want to thank and acknowledge Woodside Energy for providing our research group with the field dataset presented in this section, as well as my colleague Tatiana

Kalinicheva for processing the field data and preparing all the necessary files for running a successful inversion.

4.1. Introduction

Seismic wave propagation in the subsurface is not purely elastic, as part of the energy of the wavefield is dissipated into heat. Energy absorption due to the viscosity of the rocks introduces an amplitude attenuation and phase dispersion of the recorded waveforms with respect to those obtained in viscous-free media (Causse et al., 1999), as shown in Figure 4.1. This figure compares the recorded pressure signal at $1490\,m$ from a source in a homogeneous and isotropic medium with constant velocity ($v_p = 2000\,m/s$) when there is no viscosity (i.e. an acoustic medium, black line) and when the medium is visco-acoustic (with associated constant quality factor of $Q = 20$, gray line). With respect to the acoustic medium, the amplitude recorded in the visco-acoustic medium has lower amplitudes and is delayed in time (as observed by the difference in time of peaks and troughs in both signals).

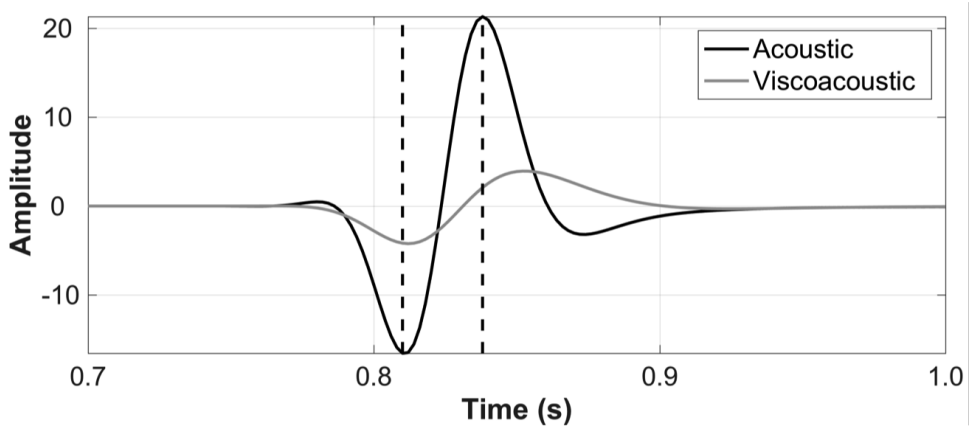


Figure 4.1.: Recorded signal at $1490\,m$ from a source in a homogeneous and isotropic acoustic medium (black line) and homogeneous and isotropic visco-acoustic medium with $Q = 20$ (gray line).

Amplitude attenuation and phase dispersion due to viscosity are not included when performing conventional acoustic FWI, and this results in an increase of the data misfit between observed and modelled acoustic data at each iteration. This hampers the convergence of the inversion towards a global minimum, and adversely affects

the quality of recovered P-wave velocity models. In fact, viscous effects need to be considered in the inversions to obtain fully quantitative subsurface models (Causse et al., 1999; Warner et al., 2012). However, viscous effects are typically neglected in conventional 3D time-domain applications due to the large computational burden of modelling viscous effects in the subsurface: this requires solving the visco-acoustic or visco-elastic wave equations with several relaxation mechanisms to have constant attenuation with frequency, as most rocks and sediments show this behaviour (Blanch et al., 1995).

The use of an accurate P-wave velocity model and a smooth estimate of the attenuation model – described by the quality factor, Q – can improve the quality of seismic migrated images, for instance via Q-compensated migration (Dai and West, 1994; Mittet et al., 1995; Causse and Ursin, 2000; Zhu et al., 2014). For this reason, several researchers have made important steps in the last decade towards obtaining accurate v_p and Q models from surface seismic data, despite the ill-posedeness of the inverse problem in recovering multiple parameter classes independently.

While some strategies for visco-acoustic FWI have been recently implemented in the time-domain (Cheng et al., 2015; Plessix et al., 2016; Bai et al., 2017), most of the implementations have been done in the frequency domain (Pratt and Worthington, 1990; Pratt, 1990; Song et al., 1995; Pratt, 1999; Plessix et al., 2012; Operto et al., 2015), in which attenuation is included naturally in the formulation by using complex velocities. However, these applications have only been partially successful, especially on large datasets. Other approaches include recovering attenuation by compensating the gradient for viscous effects (Xue et al., 2016, 2017), using travel-time tomography (TT) for Q after having performed TT and FWI for v_p only (Zhou et al., 2013) or using a semi-global inversion algorithm for Q and v_p (da Silva et al., 2017).

Another successful approach was recently implemented by Kurzmann et al. (2013), who performed visco-acoustic FWI by incorporating attenuation in the modelling as a fixed inversion parameter, and showed this improves the quality of recovered P-wave velocity models. However, this can result in a significant increase in the computation time with respect to acoustic FWI, as visco-acoustic modelling is required at each inversion step. To reduce the computation time, it is therefore standard practice to

use the amplitude-balanced functional of Shen (2010) (see equation (3.1)) in acoustic FWI.

In this chapter I apply the methodology described in chapter 2 (see Figure 2.9), based on Wiener filters, in order to mitigate viscous effects in acoustic FWI in a cost-effective manner, and I particularise it to visco-acoustic effects. Similarly to the methods of Shen (2010) and Kurzmann et al. (2013), the aim of the proposed method is to provide improved P-wave velocity models of the subsurface and not to obtain attenuation models. Instead, an estimate of Q_p for P-waves – referred from here onwards as Q for simplicity– is computed from near-offset data using Gabor transforms prior to the application of the workflow, as detailed in section 4.4. I also demonstrate on synthetic data that the proposed data-matching workflow can be combined with the amplitude-balanced method of Shen (2010), and this leads to the best results.

With respect to the application to mitigate elastic effects explained in the previous chapter, there are three main differences: 1) an estimate of Q is necessary and can be obtained from the data independently and prior to any inversion, 2) it does not require estimating v_p/v_s from the recovered P-wave velocity model obtained in Step 1 and 3) visco-acoustic modelling is performed in Step 2 of the workflow rather than elastic modelling and, therefore, the same grid spacing and time-sampling as for acoustic modelling can be used.

For its implementation, I use the workflow in Figure 2.9, which I will refer as workflow 1 in the remaining of this chapter. Additionally, I suggest two alternative workflows in the following sections to improve the performance of this method in the presence of strong attenuation effects (workflow 2) and strong random noise (workflow 3). These variations consist in skipping Step 1 in workflow 1 and modelling acoustic and visco-acoustic data directly from the starting model (workflow 2), and using a better starting model for the acoustic inversion in Step 4, obtained by severely smoothing the recovered P-wave velocity model in Step 1 (workflow 3). These approaches will be further detailed when appropriate.

In terms of structure of this chapter, I first assess the impact of neglecting viscous effects on 2D horizontally layered models, I implement the proposed methodology and combine it with that of Shen (2010) to demonstrate the potential in improving the quality of recovered P-wave velocity models, also in the presence of strong attenuation. Similarly to the previous chapter, I investigate the impact of the filtering parameters on the quality of the matched observed data and I suggest a range of parameters suitable for most applications. Next, I repeat the tests on a more realistic 2D marine synthetic dataset, I compare the results to those obtained with visco-acoustic FWI with fixed attenuation in the modelling (as proposed by Kurzmann et al. (2013)) and I assess the impact that strong added noise on the observed data has on the recovered velocity models obtained with the data-matching workflow. For all these tests I use a rough average of the Q model, and I demonstrate this is sufficient to mitigate most viscous effects. I then implement the two alternative workflows to mitigate strong visco-acoustic effects and address a high level of random noise in the observed data. Finally, I apply this method to a field dataset by estimating Q from the data using Gabor transforms, and I show the improvements in terms of continuity of the events, level of detail, accuracy, flattening of the gathers and higher quality of seismic migrated sections.

4.2. Application to a 2D horizontally layered model

The first application of the data-matching workflow is on 2D horizontally layered models, which are used to study the optimal choice of filtering parameters and the performance of the data-matching workflow on mitigating viscous effects when combined with an amplitude-balanced objective functional. The true P-wave velocity model is shown in Figure 4.2 (solid black line), which consists of a two-layer model with constant velocities within each layer and a top water layer of $600\,m$ depth. Figure 4.2 also contains the starting model (dashed black line) used for the inversions and two absorption models, defined by two vertical variations of the quality factor, Q_1 and Q_2. These models are designed to simulate two different geological scenarios: one in which there is strong intrinsic attenuation at the sea-bottom and in which Q

is constant where P-wave velocity is constant, assuming compaction has no effect on Q within the same type of rocks, and the other in which Q increases smoothly with depth due to compaction even when P-wave velocity is constant – here it is assumed that porosity is reduced due to compaction leading to less energy absorption.

Thus, the first model, referred to as model A (solid red line), consists of two layers of constant and low Q values, whereas the quality factor changes gradually in the second model, referred to as model B (solid blue line). Thus, viscous effects are larger in model A, as these are inversely proportional to the quality factor, and this model is used to test the performance of the data-matching workflow when viscous effects are strong. In both models, the quality factor is very large within the water layer to simulate a realistic marine dataset – absorption is negligible in water. Additionally, average models computed for the two Q models below the sea-bottom are overlaid (dashed red and blue lines), which have the same Q value within the water layer as that for the true model.

Acoustic and visco-acoustic data are first generated with two independent FD modelling codes with the same level of accuracy (second order in time and sixth order in space). Two-dimensional laterally invariant models obtained by laterally extending the vertical profiles in Figure 4.2 are used, so that the models are $15\,km$ long (and $1.2\,km$ deep), and the grid spacing is set to $10\,m$ to ensure there is little dispersion of P-waves below $30\,Hz$ (as per equation (2.32)). A total of 41 sources are deployed at $6\,m$ depth, which emit a pressure signal corresponding to a Ricker wavelet with peak frequency at $15\,Hz$. Three seconds of data are recorded by 967 receivers at $16\,m$ depth, with a time sampling of $1\,ms$ to ensure numerical stability (as per equation (2.33)). A top free surface and absorbing boundaries on the sides and at the bottom of the model are used to simulate marine data, which includes surface-related multiples and seismic ghosts.

Next, the generated acoustic and visco-acoustic observed data for models A and B are acoustically inverted using a multi-scale approach by inverting the data from $3\,Hz$ to $15\,Hz$ in 11 iteration blocks and a total of 60 iterations. The same inversion parameters are used for all inversions, and the results at the centre of the 2D models

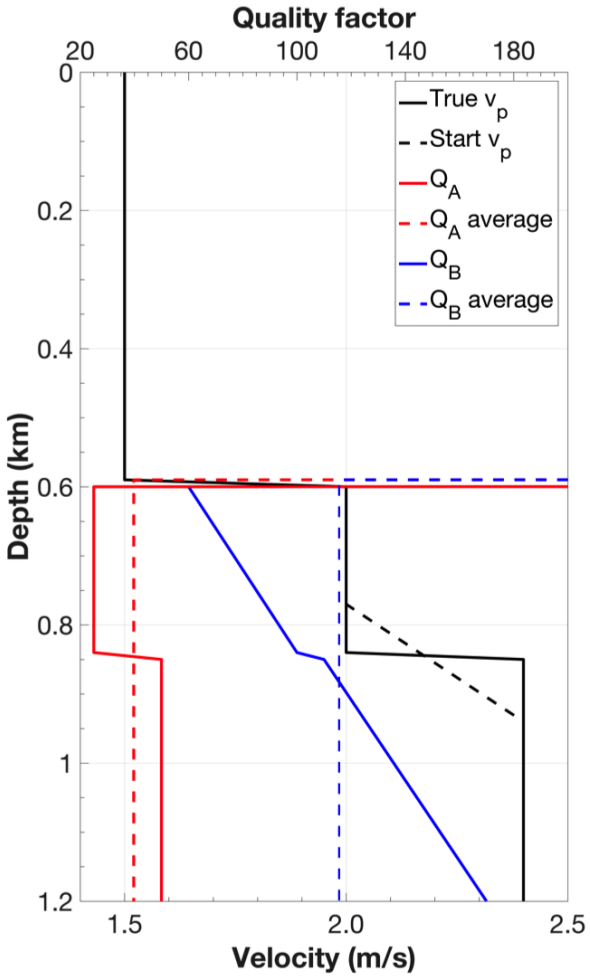


Figure 4.2: Vertical profiles of the 2D horizontally layered true P-wave velocity and absorption models A and B used to generate acoustic and visco-acoustic data. The starting model and the average of the two absorption models below the subsurface are also overlaid.

are shown in Figure 4.3a when using the conventional objective functional, and in Figure 4.3c when using the amplitude-balanced functional of Shen (2010).

The first observation from these plots is that acoustic FWI of acoustic data (green line) leads to a very accurate result: this is the reference result as visco-acoustic effects are absent both in the modelling and in the inversion. Secondly, acoustic FWI of visco-acoustic data leads to less accurate and slower recovered P-wave velocity models than those obtained with acoustic FWI of acoustic data, at the sea-bottom but also especially with depth, as the time-delay and amplitude attenuation caused by viscous effects are neglected during the inversions. Given that model A has smaller quality factor values, the impact of viscous effects on the recovered P-wave velocity models is larger in model A (red line) than in model B (blue line) in Figures 4.3a and 4.3c, resulting in velocities up to $400\,m/s$ slower than the reference values for model A when using the conventional functional. A third and important observation

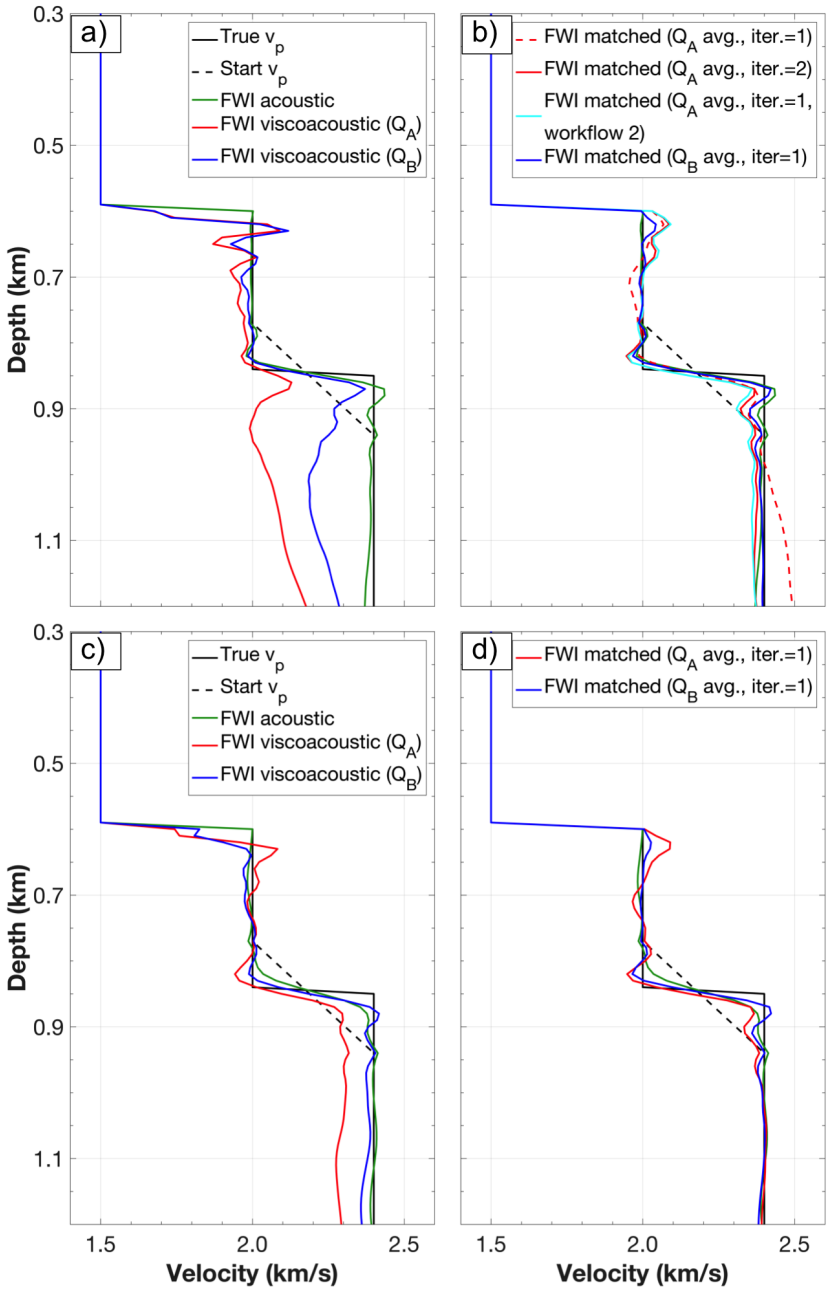


Figure 4.3.: Vertical profiles at the centre of the 2D horizontally layered models using the conventional FWI functional – panels a) and b)– and an amplitude-balanced functional – panels c) and d)– after acoustic FWI of acoustic and visco-acoustic data – panels a) and c)– and after acoustic FWI of matched data – panels b) and d) – for one iteration of the workflow – panel d)– and for one and two iterations of workflows 1 and 2 – panel b). The true, starting and recovered velocity model for acoustic data are overlaid.

is that viscous effects are significantly mitigated when using an amplitude-balanced functional (Figure 4.3c), yielding more accurate results than those obtained in Figure 4.3a. This improvement is sufficient to mitigate viscous effects for model B and some viscous effects for model A, although there is still a visible imprint of the latter for model A at the sea-bottom and deep in the model due to the presence of strong absorption. As a minor observation, the result of acoustic FWI of acoustic data is not as good as that in Figure 4.3a as the amplitude-balanced functional disregards part of the amplitude information, which is well-known in this case as this is the ideal situation.

In order to address viscous effects when using both objective functionals, I apply the data-matching workflow in Figure 2.9 using the recovered P-wave velocity models after acoustic FWI of visco-acoustic data in Figures 4.3a and 4.3c and estimates of the Q models obtained by averaging the true models below the sea-bottom, which are shown in Figure 4.2 (dashed red and blue lines): these are very approximate models. Modelled acoustic and visco-acoustic data are generated in Step 2 of the workflow, and matching filters that match the visco-acoustic to the acoustic data are computed and applied to the observed visco-acoustic data in Step 3. The length of the matching filters is set to $325\,ms$ and 13 traces are matched simultaneously, i.e. the spacing between the first and the last neighbouring traces matched simultaneously is $180\,m$. Acoustic FWI of the matched data (Step 4 of the data-matching workflow) results in the recovered models shown in red (dashed) and blue (solid) in Figure 4.3b for models A and B using the conventional FWI functional, and in red and blue (both solid lines) for models A and B in Figure 4.3d using an amplitude-balanced functional. Figure 4.3b also shows the result of applying a second iteration of the data-matching workflow (Step 5) for model A (solid red line) to further mitigate viscous effects. I make the following observations from Figures 4.3b and 4.3d:

1. An average Q estimate is sufficient to mitigate most viscous effects yielding more accurate and faster recovered velocity models, as long as the quality factor is of the same order of magnitude as the reference values.

2. The data-matching workflow with an average Q estimate and the conventional FWI functional performs well in the presence of strong viscous effects (Figure 4.3b) and improves the accuracy of the recovered model obtained after acoustic FWI of the visco-acoustic data, but the effects of absorption are still visible within the second layer (dashed red line).

3. A second iteration of the data-matching workflow for model A and the conventional FWI functional (Figure 4.3b, solid red line) further improves the quality of the recovered model in the presence of strong viscous effects, but increases the cost of the workflow mainly due to an additional acoustic inversion step.

4. Using an amplitude-balanced functional in combination with the data-matching workflow (Figure 4.3d) provides the best results for a single iteration of the data-matching workflow and yields recovered models with minimal imprint of viscous effects, even when these are large (model A).

To further reduce the computation cost when using the conventional FWI functional in the presence of large viscous effects in the data, I suggest an alternative workflow, namely workflow 2. The latter consists in eliminating Step 1 from the workflow in Figure 2.9 and modelling acoustic and visco-acoustic data directly from the starting model and the estimate of Q. Workflow 2 then proceeds with the remaining steps of the workflow (3 to 5) resulting in an improved P-wave velocity model at a cost similar to a single acoustic inversion: this is analogous to the second workflow suggested to mitigate strong elastic effects in chapter 3. When applied to the visco-acoustic data for model A using the conventional FWI functional, this results in the recovered model shown in cyan (solid line) in Figure 4.3b, which is faster and more accurate than the recovered model after acoustic FWI of visco-acoustic data in Figure 4.3a (solid red line). It is also more accurate than the model in Figure 4.3c obtained with an amplitude-balanced functional at roughly the same cost, and is very similar to that obtained after two iterations of the data-matching workflow (Figure 4.3b, solid red line). Thus, workflow 2 is recommended in the presence of large viscous effects.

The simple synthetic tests in Figure 4.3 demonstrate that neglecting visco-acoustic effects has a negative impact on the quality of the recovered P-wave velocity mod-

els for marine datasets, which can be partly addressed using the amplitude-balanced functional proposed by Shen (2010). For this reason, I use an amplitude-balanced functional in the remaining tests within this chapter. Nevertheless, residual effects are still visible when using the latter, especially if viscous effects are large, but can be further mitigated when combining this functional and the data-matching workflow.

Analogous to the application to mitigate elastic effects in chapter 3, the choice of appropriate filtering parameters is also important to obtain good quality matched observed data, such that events are continuous laterally and vertically and artefacts caused by matching data with different spectra are minimised. For this reason, I assess the quality of the matched data obtained for model B – the conclusions for model A are very similar as the P-wave velocity model is the same and only the absorption model changes– for different filtering parameters, i.e. for filters with different lengths and matching neighbouring traces independently or simultaneously.

The quality of the matched data is assessed by comparing a shot gather with the same gather for acoustic data. The data-matching method should result in gathers with smaller differences when compared to the acoustic data than the visco-acoustic data. This is illustrated in Figure 4.4, which shows the data for a shot gather at the centre of the model for visco-acoustic (Figure 4.4a) and matched data (Figures 4.4c, 4.4e, 4.4g and 4.4i) obtained with different parameters (left panels), as well as their difference with respect to the true acoustic data (right panels) and the average RMS differences for each shot with respect to the acoustic data are overlaid. Matched data are obtained computing Wiener filters of $325\,ms$ length for panels 4.4c and 4.4e, of $50\,ms$ length for panel 4.4g and of $1.5\,s$ for panel 4.4i, whereas the matching is done trace-by-trace for panel 4.4c and matching 13 traces simultaneously for panels 4.4e, 4.4g and 4.4i so that the spacing between the first and last matched traces is $180\,m$.

Visco-acoustic data presents the largest differences with respect to the acoustic data for those events generated within the subsurface except for localised regions, and smaller differences associated with the direct arrival as the latter propagates through the water only, which has very large Q values. Large differences in Figure 4.4b associated with reflections and deep refractions are mainly due to the time-delay

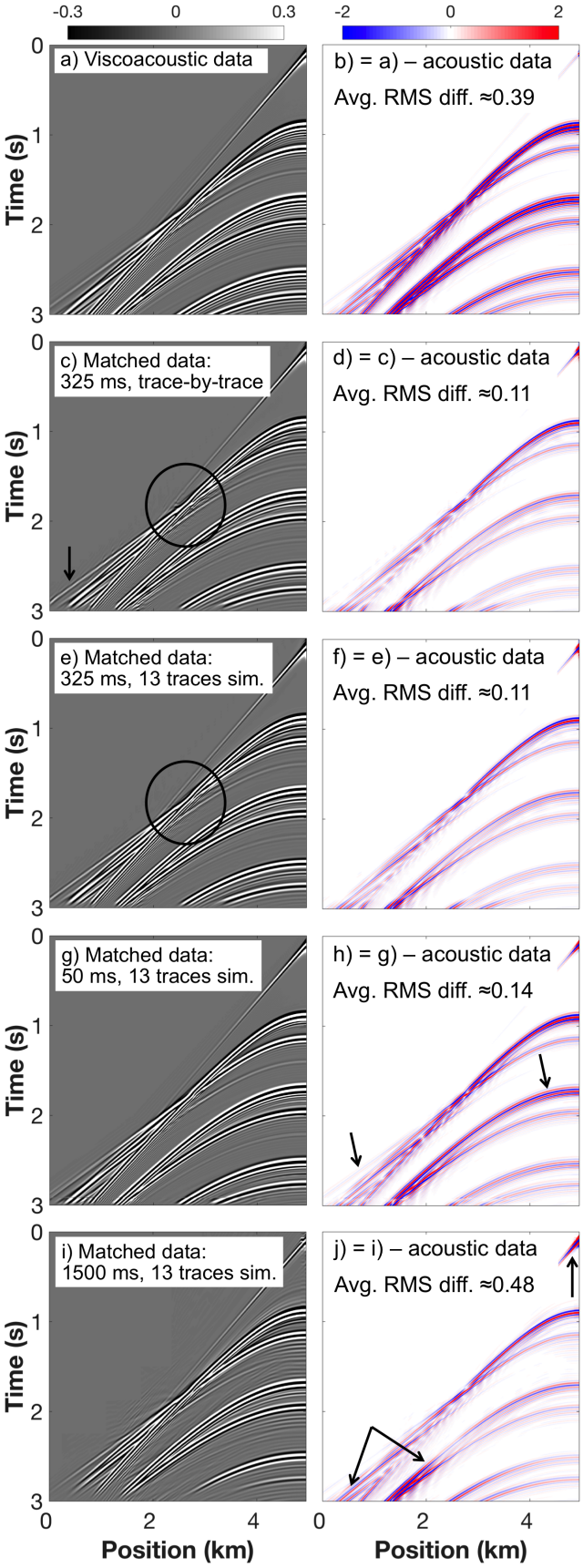


Figure 4.4: Shot gathers recorded at the centre of model B for the a) visco-acoustic data and the matched data obtained by computing Wiener filters c) trace-by-trace with a length of 325 ms and simultaneously matching 13 traces with lengths of e) 325 ms, g) 50 ms and i) 1.5 s. The gathers on the right column show the differences of those on the left with the acoustic data, in which the average RMS difference for all the traces in each shot gather with respect to the acoustic data has been overlaid.

introduced by viscous effects, and secondly due to amplitude attenuation. This can be appreciated in Figure 4.5, which illustrates the difference (green line) for a single trace at 4.2 km position for acoustic (black line) and visco-acoustic data (red line). Despite amplitude differences being small between the same trace for acoustic and visco-acoustic data, the small but visible time-delay introduced by viscous effects yields a large difference when this is computed sample-by-sample.

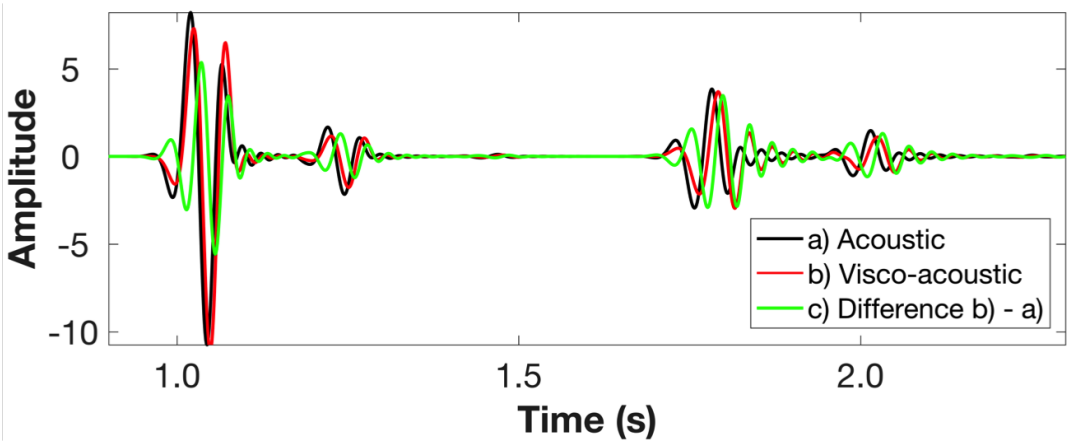


Figure 4.5.: Extracted trace at 4.2 km position for a central shot gather for acoustic (black line) and visco-acoustic (red line) data, as well as their difference (green line).

On the other hand, using data-matching workflow with an average Q estimate results in matched data that have smaller differences with respect to the acoustic data, mainly due to a phase correction and in part due to an amplitude adjustment when using the current method – the latter can be for example observed by comparing the amplitude increase for the refraction at the second layer in Figure 4.4c with respect to Figure 4.4a, indicated by an arrow. However, trace-by-trace matching introduces some distortion of the events where multiple waves interfere, such as the region indicated by a circle in Figure 4.4c, which are mitigated when matching multiple traces simultaneously and leaving the length of the filters unchanged, as shown in Figure 4.4e. The difference plot in Figure 4.4f is in general less noisy than that in Figure 4.4d due to the mild smoothing introduced by this strategy. Thus, after the tests in this section and those in the remaining sections of this chapter I recommend to compute Wiener filters that match several neighbouring traces simultaneously such that

the spacing between the first and the last trace being matched is comprised between $150\,m$ and $250\,m$: this generally results in good quality data with small artefacts due to filtering, good continuity of the events while still addressing viscous effects.

The length of the filters also has a visible impact on the quality of the data – as indicated by the average RMS values indicated on Figure 4.4 – and hence the subsequent acoustic inversions. Filters that are short result in noisy data and large data differences both for reflections and refractions, as shown in Figure 4.4h (black arrows) obtained with filters of $50\,ms$ length, whereas long filters result in smoother and more continuous data but larger errors close to the source and at late times, as shown by the black arrows in Figure 4.4j. Additionally, the longer the filter the more computationally demanding it is to compute its coefficients. To further compare the performance of the filtering parameters tested in Figure 4.4, a trace is extracted at $2\,km$ position – i.e. at a relatively long offset – and compared with the corresponding trace for the acoustic data (black line) in Figure 4.6 for all datasets in Figure 4.4 (red lines in Figure 4.6).

Figure 4.6a shows the expected amplitude attenuation and time delay introduced by viscous effects for this trace, which are mitigated when applying the data-matching workflow as shown in Figures 4.6b to 4.6e. Similarly to Figure 4.4, Figure 4.6b is obtained by matching data trace-by-trace, whereas the remaining panels 4.6c to 4.6e are obtained by matching 13 traces simultaneously. In terms of length of the filters, this is set to $325\,ms$ in Figures 4.6b and 4.6c, to $50\,ms$ in Figure 4.6d and to $1.5\,s$ in panel 4.6e. Independently of the filtering parameters used, the time delay caused by viscous effects is appropriately compensated in all traces for the matched data (Figures 4.6b to 4.6e), as the positions of peaks and troughs are coincident with those of the acoustic trace. In terms of amplitude corrections, each choice of the filtering parameters results in a different correction except for traces in Figures 4.6b and 4.6c, which result in very similar results. The main difference of matching data on a trace-by-trace basis or doing it for multiple traces simultaneously is not visible here, but where multiple events cross each other. On the other hand, I observe that filters of $50\,ms$ length result in poor amplitude correction, especially for early arrivals, and

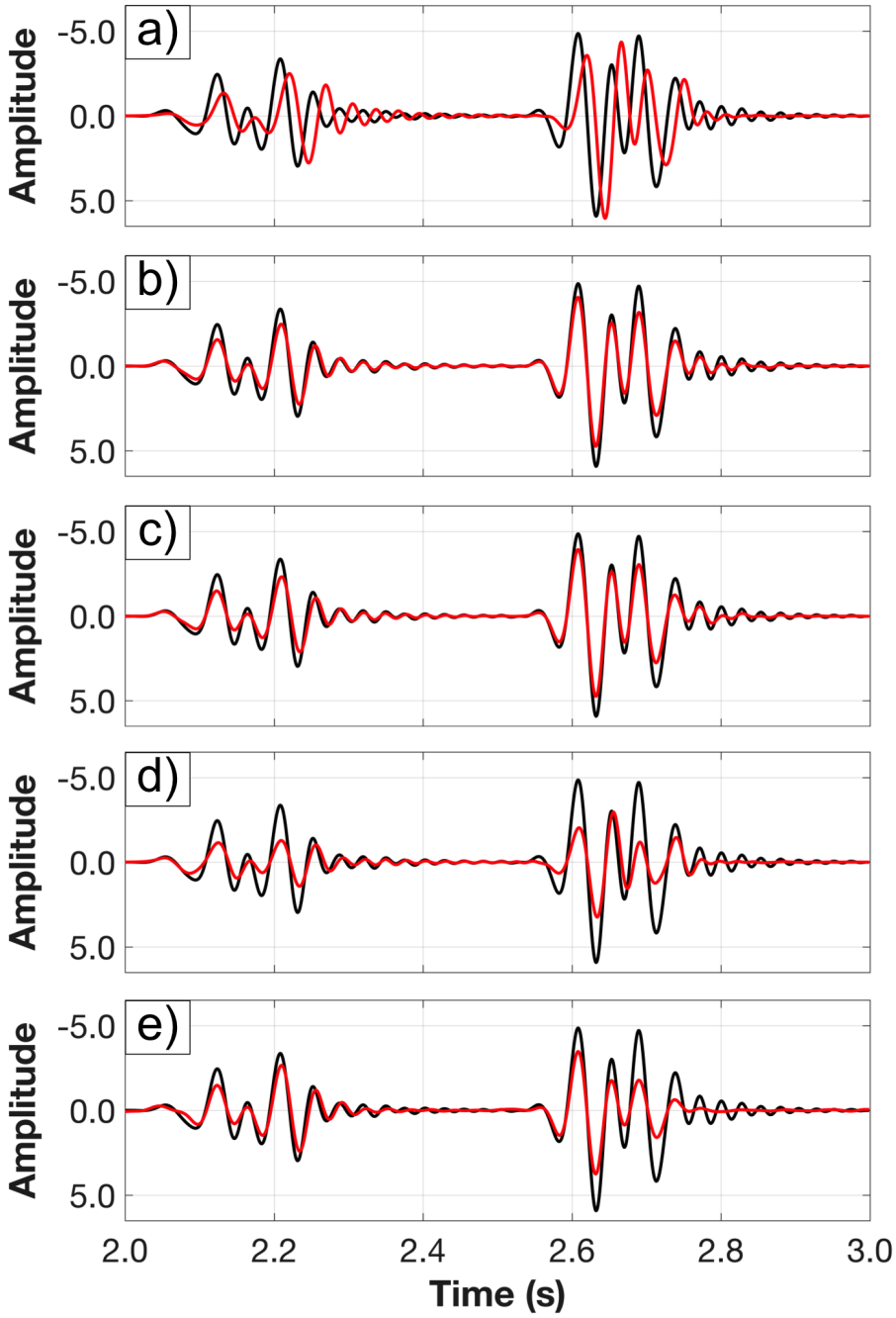


Figure 4.6.: Trace comparison at $2\,km$ position for the datasets in Figure 4.4 (red lines) with the true acoustic data (black line). Thus, panel a) corresponds to visco-acoustic data and panels b) to e) correspond to matched data obtained with filters of $325\,ms$ length for panels b) and c), of $50\,ms$ for panel d) and of $1.5\,s$ for panel e). Wiener filters are computed trace-by-trace for panel b), whereas they are computed by matching 13 traces simultaneously for panels c) to e).

that longer filters (of 1.5 s) are more expensive to compute and yield better amplitude behaviour for early arrivals, which deteriorates at later times. Therefore, filters with an intermediate length – typically between 0.3 s to 0.4 s– result in matched data that is optimally matched both for early and late arrivals at a relatively low cost. Moreover, using overlapping filters that are short compared to the length of the full trace helps mitigating the non-stationary behaviour of absorption with time, i.e. along the trace.

The study of the performance of the data-matching workflow and the quality of the matched data done in this section are now utilised to mitigate visco-acoustic effects in a more realistic marine synthetic dataset and in a field dataset.

4.3. Application to a marine synthetic dataset

The impact of visco-acoustic effects and the performance of the data-matching workflow to provide high quality recovered P-wave velocity models, even in the presence of noise, are now investigated on the more realistic marine Marmousi2 dataset (Martin et al., 2006), which presents both lateral and vertical velocity variations. Additionally, I compare the recovered models to those obtained with visco-acoustic FWI with a fixed absorption model, as suggested by Kurzmann et al. (2013).

First, I generate acoustic and visco-acoustic data using the true P-wave velocity model in Figure 4.7a, which contains a 450 m-deep water layer, and the reference Q model in Figure 4.7c. The latter is a smooth absorption model that is representative of the subsurface, with smooth lateral and vertical variations and values that span from 20 at the sea-bottom – typically being more attenuative due to porosity, for instance – to 100 at the deepest part of the model. The models are discretised with a grid spacing of 10 m to avoid dispersion of P-waves below 30 Hz (as per equation (2.32)), and 6 s of pressure data sampled every 1 ms to ensure numerical stability (as per equation (2.33)) are generated by 99 sources deployed at 6 m depth uniformly distributed across the model with a shot spacing of 160 m, which are sequentially set off. The source wavelet is a Ricker wavelet with peak frequency at 11 Hz, which

results in useful data within the bandwidth that spans from $2\,Hz$ to $20\,Hz$. Data are recorded by 791 receivers uniformly distributed across the model every $20\,m$ and deployed at $10\,m$ depth. A top free surface and absorbing boundaries on the sides and the bottom of the model are used to model marine data and include surface multiples and seismic ghosts.

Following data generation, I perform acoustic FWI of the acoustic and visco-acoustic observed data independently using the starting model in Figure 4.7b, which is obtained by strongly smoothing the true velocity model in Figure 4.7a so that it resembles a model obtained with travel-time tomography. More specifically, the latter is built by separating the true model in three parts as a function of depth and applying a Gaussian smoothing filter with a correlation length of 2 cells several times in each region, being the deeper section the region of the model where the smoothing filter is applied more times – this simulates the usually coarser resolution of models obtained with travel-time tomography in deeper regions. To assess the closeness of the starting model with the true model within the overlaid dotted box, I compute the absolute value of the relative error for each cell within this box and I calculate the average: this indicates the accuracy with respect to the true model, and is shown as a percentage on the right corner of this model and also for the recovered models. For the inversions, I use a multi-scale approach by inverting the data from $3\,Hz$ to $18\,Hz$ in 130 iterations in the time domain and using the amplitude-balanced functional of Shen (2010) to further mitigate viscous effects. The recovered models are displayed in Figures 4.7e and 4.7f for the acoustic and visco-acoustic data, respectively.

The recovered model in Figure 4.7e is the reference result, as the same constitutive laws have been used to generate the data and invert it. The latter recovered model presents most of the interfaces in the true model with accurate velocities, so that the relative average error is reduced from $8.42\,\%$ to $3.11\,\%$, and is mainly limited in resolution by the range of inverted frequencies (as there is no added noise) and the acquisition geometry. On the other hand, the recovered model after acoustic FWI of visco-acoustic data in Figure 4.7f is overall slower – for instance, there is a reduction in velocity at the horizontal high velocity layer at $2.2\,km$ depth that spans across

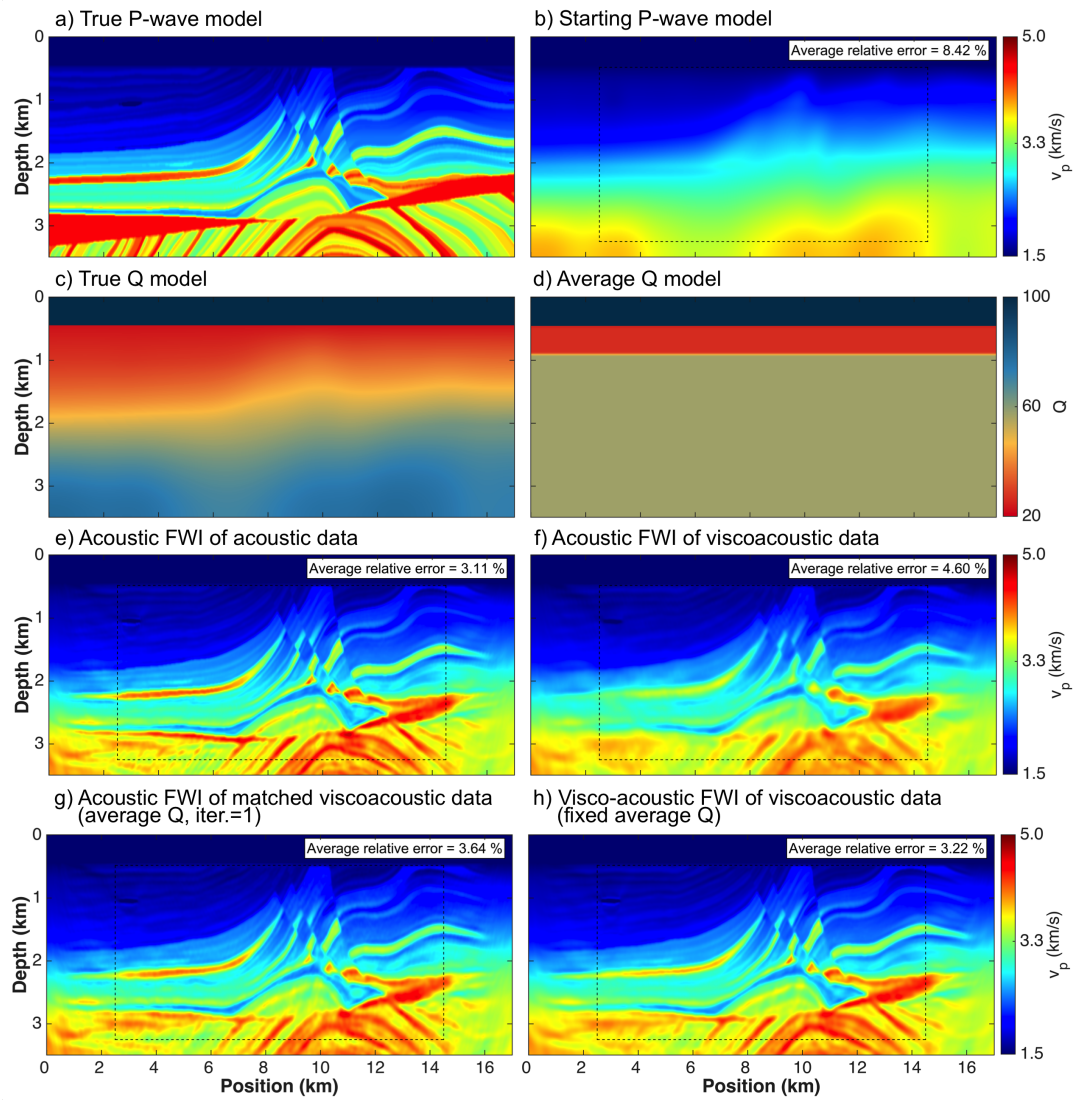


Figure 4.7.: Vertical slices of the marine Marmousi2 a) true P-wave velocity, b) starting P-wave velocity, c) true Q and d) 1D average Q estimate models and the recovered models obtained after e) acoustic FWI of acoustic data, f) acoustic FWI of visco-acoustic data, g) acoustic FWI of matched data and h) visco-acoustic FWI of visco-acoustic data.

the left-hand-side of the model – and less detailed – for instance, see below 2.9 km depth. This is mainly due to the time delay caused by attenuative media not being considered, and also due to the amplitude effects being ignored, which causes an increase on the average relative error with respect to Figure 4.7e to 4.60 %. Thus, an amplitude-balanced functional is not enough to recover a high-fidelity velocity model for this marine dataset in an attenuative medium, and the application of the data-matching workflow to further improve the results is justified.

For this reason, I implement the data-matching workflow in Figure 2.9 to address visco-acoustic effects in acoustic FWI. First, I model acoustic and visco-acoustic data using the recovered model in Figure 4.7f and a 1D Q estimate (Step 2) built by averaging the true model in Figure 4.7d in two different regions: the first one comprises the area between the sea-bottom and the 950 m depth level, and the second one comprises the remaining bottom part of the model. The average is mildly smooth at the transition between these layers to avoid artificial diffractions at the interface. I then compute matching filters in Step 3 of the workflow that match the visco-acoustic to the acoustic modelled data, and apply these to the observed visco-acoustic data, which results in a matched dataset. In this step I compute filters that match 11 traces simultaneously – i.e. with a spacing of 200 m between the neighbouring traces simultaneously matched– and I set the length of the filters to 325 ms. An additional acoustic inversion (Step 4) results in the recovered model in Figure 4.7g.

Despite the use of a very approximate Q estimate within the data-matching workflow, the recovered velocity model in Figure 4.7g is more accurate, sharper and overall faster than the model in Figure 4.7f, leading to a reduction of the average relative error down to 3.64 %. This could be further improved by an additional iteration of the data-matching workflow, but it would also increase the computational cost and the benefit would be smaller than that of the first iteration.

To further study the improvement after the application of the data-matching workflow, Figure 4.8 shows the difference between the starting and recovered model after acoustic FWI of the matched visco-acoustic data, i.e. the difference between Figures 4.7b and 4.7g. The latter shows the modifications of the model introduced during the inversion that lead to the improved P-wave velocity model in Figure 4.7g.

As an additional comparison, I finally perform visco-acoustic FWI of the visco-acoustic data by keeping Q as a fixed parameter during the inversion, equal to the model in Figure 4.7d, as proposed by Kurzmann et al. (2013). This yields the reference recovered model in Figure 4.7h, as it has been obtained using a modelling code that uses the same constitutive law (the visco-acoustic wave equation) for generating the observed dataset and inverting it. The quality of the latter is slightly better than that

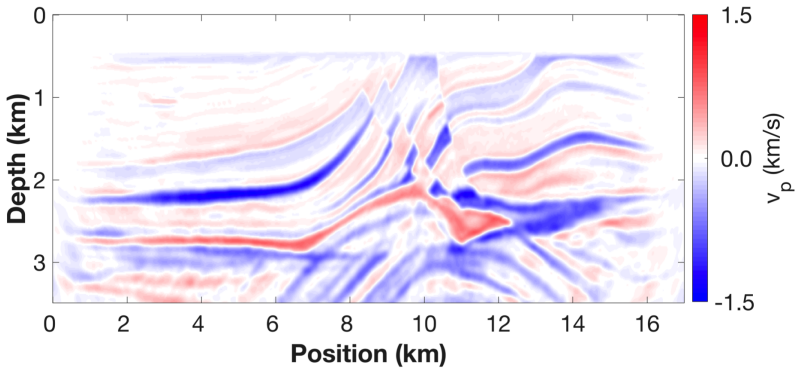


Figure 4.8.: Vertical slice of the difference between the starting and the recovered models after acoustic FWI of the matched visco-acoustic data in Figures 4.7b and 4.7g, respectively.

in Figure 4.7g, in which different wave equations were used for the modelling and the inversion of the data. Nevertheless, the nearly-horizontal high-velocity layer at 2.2 km depth is better recovered with the data-matching method. The main advantages of using the data-matching workflow with respect to a visco-acoustic inversion in the time domain are:

1. All inversions are acoustic, resulting in a reduction of the computation time. On the contrary, visco-acoustic FWI uses the visco-acoustic wave equation at each iteration with Q being nearly constant for the range of frequencies used in seismic exploration, and this requires including several relaxation mechanisms to model viscous effects in the time domain. The latter results in an increase in the compute time, especially for 3D datasets (acoustic FWI for this dataset takes two hours on 50 nodes and 2 tasks per node, whereas visco-acoustic FWI takes 3.8 h using the same resources, and this difference will increase for 3D datasets).

2. It results in a matched dataset that contains less viscous effects than the original observed data.

3. It can be potentially combined with the method presented in chapter 3 to mitigate visco-elastic effects all at once, although this has not been further explored.

The improvement introduced by the data-matching workflow is due to the changes

introduced in the observed data used for the inversions. For this reason, it is important to check the changes introduced by the data-matching workflow in the observed data, as shown in Figure 4.9. This figure displays a representative shot gather for the acoustic (Figure 4.9a), visco-acoustic (Figure 4.9b) and matched data (Figure 4.9c) that illustrates the differences between these datasets. Their difference with respect to the same shot for the acoustic data is presented in Figures 4.9d and 4.9e.

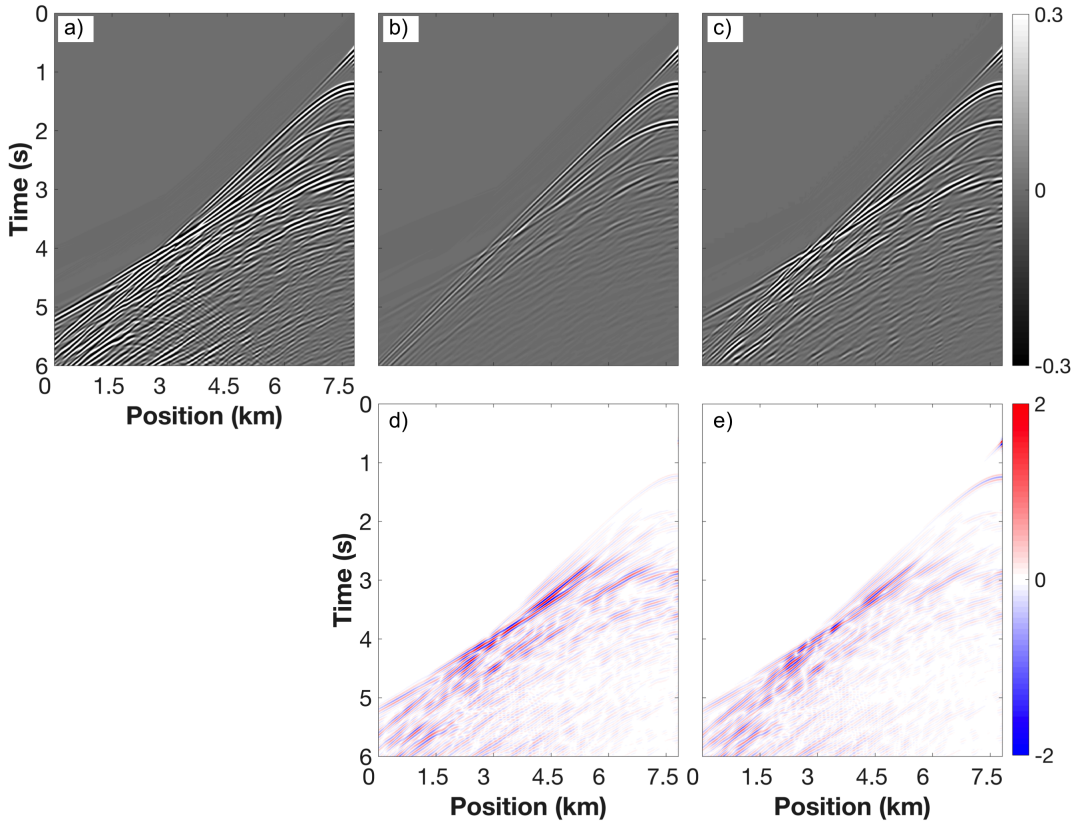


Figure 4.9.: Representative shot gathers recorded at the centre of the Marmousi2 model for a) acoustic, b) visco-acoustic and c) matched data. Panels d) and e) show the differences of the panels above with the same shot gather for the acoustic data.

As expected, the visco-acoustic data has smaller amplitudes than the acoustic data. However, phase changes are now more difficult to observe as there are many more events than for the previous 2D horizontally layered models. These effects are partially corrected when applying the data-matching workflow with a rough estimate of Q, and the data in Figure 4.9c has larger amplitudes than the visco-acoustic data in Figure 4.9b, both for reflections and refractions. As a result, the differences displayed

in Figure 4.9e are smaller than those in Figure 4.9d, which results in the high quality recovered velocity model in Figure 4.7g.

4.3.1. Impact of noise

Before the application to a field dataset, I test the robustness of the data-matching workflow in the presence of strong random noise in the observed data. This will lead to a new alternative workflow, namely workflow 3, that is recommended in the presence of strong random noise.

First, I add random noise to the original observed visco-acoustic data by computing the mean of the absolute value of the difference between the maximum and minimum values of each trace, and adding a percentage of the latter to each sample in the dataset that is comprised between 0 % and 50 %. The added noise is additionally filtered with the spectrum of the source so that the bandwidth of the noise is within the bandwidth of the data. Figure 4.10a shows a representative shot gather of the noisy visco-acoustic data, and its difference with the noise-free acoustic data is displayed in Figure 4.10c.

Acoustic inversion of the latter noisy visco-acoustic data with the same inversion parameters as in the previous tests and an amplitude-balanced FWI objective functional leads to the recovered model in Figure 4.11a. This model is more noisy and less accurate than the recovered model for the noise-free data in Figure 4.7f (the average relative error has increased to 5.15 %). The latter together with the average Q model in Figure 4.7d is used in Step 2 of the data-matching workflow to mitigate viscous effects, which results in a matched dataset – the same filtering parameters as before are used. A representative shot gather for the latter is shown in Figure 4.9b. Similarly as for the noise-free data, the amplitude of the refractions and reflections is overall increased, leading to smaller differences with respect to the acoustic data, but now the amplitude of noise is also boosted. Note that the matched data is muted: this is done automatically to reduce the cost of computing matching filters, which are computed only where the acoustic modelled data is larger than a certain small threshold value. In order to determine whether the boosting of the noise has a larger

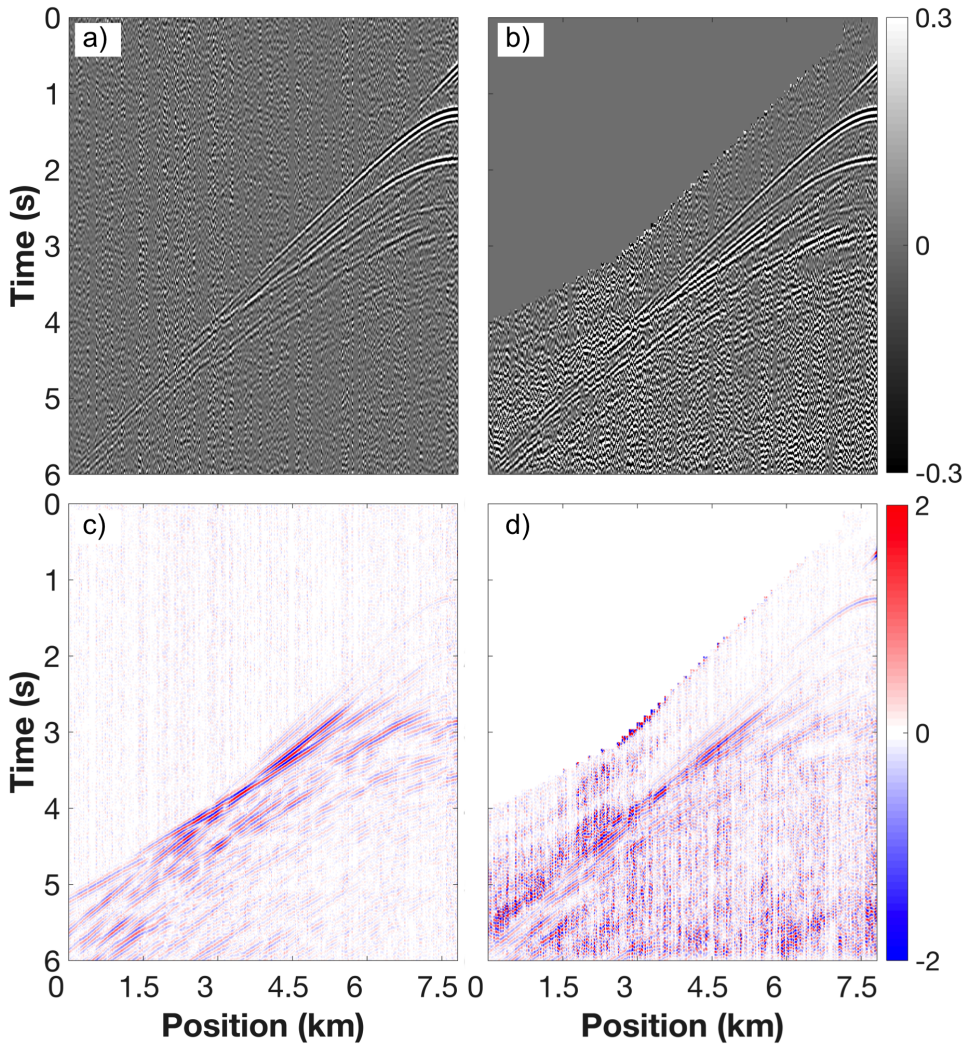


Figure 4.10.: Representative shot gathers at the centre of the Marmousi2 model for the a) visco-acoustic data with noise and b) the corresponding matched data. Panels c) and d) show the differences of the panels above with the noise-free acoustic data.

effect than the correction in amplitude and phase of the different events, I then proceed with Step 4 of the workflow and perform acoustic FWI of the matched data, resulting in the recovered model in Figure 4.11b. Although the latter is overall faster and more accurate in some regions – for instance, across the horizontal high-speed layer at 2.2 km depth–, noise has introduced many localised artefacts in the model that negatively affect the accuracy of the model, such that there is no reduction of the average relative error within the dotted area but an increase to 5.56 %.

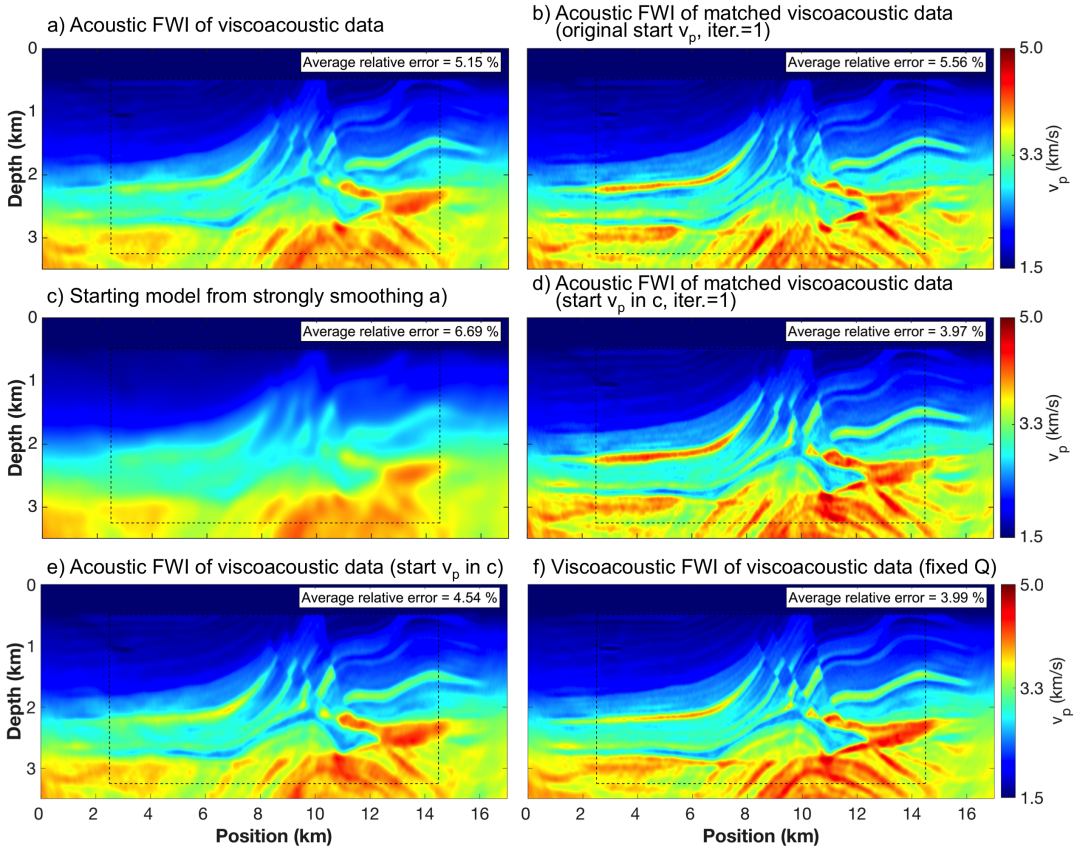


Figure 4.11.: Vertical slices of the recovered P-wave velocity models after acoustic FWI of a) the noisy visco-acoustic data and b) the matched data using the original starting model, c) a new starting model obtained by smoothing a), and recovered velocity models after acoustic FWI of d) the matched data using the starting model in c) and e) the observed data using the starting model in c). Panel f) shows the recovered velocity model obtained from visco-acoustic FWI of the noisy visco-acoustic data using the original starting model and the Q model in Figure 4.7d, which is kept fixed during the inversion.

To mitigate the impact of noise I suggest a modified data-matching workflow, referred to as workflow 3, that consists in performing Step 3 of the workflow as in the original workflow – i.e. computing the matched data from the recovered velocity model after acoustic FWI of the observed data and the estimate of Q– and using a different starting model for the acoustic inversion in Step 4, which is obtained by heavily smoothing the recovered model in Figure 4.11a. This speeds up the convergence of the inversion using the data-matching workflow, as demonstrated in Figures 4.11c and 4.11d, which respectively show the starting model obtained by smoothing the recovered model in Figure 4.11a and the recovered model after acoustic FWI of

the matched data using the recovered model in Figure 4.11c as a starting model and just one iteration of the data-matching workflow. The latter is faster, more accurate and sharper than the recovered models in Figures 4.11a and 4.11b, with an average relative error that is significantly smaller, of 3.97 %. In fact, the quality of the recovered model in Figure 4.11d is comparable to the result for the noise-free data in Figure 4.7g. In order to determine whether this is only due to the starting model used and not the combination of the latter and the use of the matched data, I perform acoustic FWI of the noisy visco-acoustic data but using the starting model in Figure 4.11c, and the result is shown in Figure 4.11e. The recovered model in this case is slightly better than that in Figure 4.11a as the starting model is more accurate, but the changes are small and the overall velocity is still too slow because the observed data contains viscous effects. Hence, using the matched data in combination with the smooth starting model leads to a very accurate recovered model in the presence of strong noise. As a final comparison, the result of visco-acoustic FWI of the noisy visco-acoustic data with a fixed Q model during the inversion, equal to that in Figure 4.7d, is displayed in Figure 4.11f. This is comparable to the model in Figure 4.11d, but it has been obtained using the same constitutive law during the modelling and the inversion and requires longer computation time as the inversion has been implemented in the time-domain, where several relaxation mechanisms are necessary to model a subsurface with constant Q with frequency.

Thus, I recommend the use of workflow 3 to mitigate viscous effects when the observed data is very noisy.

4.4. Application to a marine field dataset

I now implement the data-matching workflow designed to mitigate visco-acoustic effects to a marine field dataset acquired in the Carnarvon basin, on the North-West Australian Shelf (Kalinicheva et al., 2017). The acquisition geometry for this survey was designed and optimised for a successful implementation of FWI. To achieve this, a low frequency source with central frequency at $8\,Hz$ was deployed at $10\,m$ depth and set off every $50\,m$, generating useful signal below $2.5\,Hz$. Additionally, a

$10000 - m$ long streamer cable towed at $25\,m$ depth with a receiver spacing of $12.5\,m$ was used to record the data, which enabled the recording of sufficient refracted energy. The survey was carried out in a marine environment, with water depth ranging from $600\,m$ to $1600\,m$ across the survery. For the present study, I use data acquired across a single sailing line that contains strong water multiples and a good signal-to-noise at low frequencies. Some minimal pre-processing was applied to the data for preparation for the inversions: this includes band-pass filtering the data from $1.5\,Hz$ to $24.5\,Hz$ to suppress swell noise and extracting a source wavelet by using matching filters (Kalinicheva et al., 2017).

The minimally pre-processed data was used as an input for an acoustic inversion. In parallel, I estimated a Q subsurface model from near-offset data, which was then used to mitigate visco-acoustic effects and improve the quality of the recovered model. Thus, I first introduce the method utilised to estimate Q from the data and how has this been applied to this particular dataset. I note that estimating Q from the data is not an easy problem and several methods have been suggested during the last decades that have shown more or less success, but only a smooth and very approximate absorption model is required for the data-matching workflow to work, as demonstrated in the previous sections.

4.4.1. Estimating Q

Viscous effects degrade the quality of recorded seismic data and, thus, several strategies for filtering absorption effects have been proposed during the last decades. However, most of them require an estimation of the quality factor Q (Lupinacci and Oliveira, 2015), as absorption is inversely proportional to the latter, and several methods have been suggested to estimate Q from reflection data: this is referred to as Q analysis. Some of the most utilised methods to estimate Q include the spectral ratio method (Hauge, 1981; Raikes and White, 1984; Sams and Goldberg, 1990; Pinson et al., 2008), the frequency shift method (Quan and Harris, 1997; Li et al., 2015) and spectral decomposition by using Gabor transforms (Wang, 2004, 2014), amongst others. Here I use the method of Wang (2004) and Wang (2014) to estimate Q, which

is based on Gabor transforms of the data. This method allows robust and automatic estimation of Q trace-by-trace from reflection field data, and performs better than other methods (Lupinacci and Oliveira, 2015).

Gabor transforms are windowed Fourier transforms of the data, in which the window is a Gaussian function, and can be used to determine the frequency content of a trace as a function of time – unlike Fourier transforms, which consider an infinite time interval. Thus, the Gabor transform of a trace $u(t)$ leads to an amplitude spectrum $U(\tau, \omega)$ that is defined by (Lupinacci and Oliveira, 2015):

$$U(\tau, \omega) = \int_{-\infty}^{\infty} u(t) e^{-\frac{(t-\tau)^2}{2\sigma^2}} e^{-i\omega t} dt \tag{4.1}$$

where σ controls the width of the Gaussian function. If we consider reflection data, the amplitude of a plane wave, $A(\tau, \omega) = |U(\tau, \omega)|$, decays exponentially due to absorption as follows:

$$A(\tau, \omega) = A_0 e^{-\frac{\omega\tau}{2Q}} \tag{4.2}$$

where A_0 is the amplitude at $\tau = 0\,s$. The latter can be linearised by taking logarithms in the previous expression, resulting in the following equation:

$$y(\chi) = -\frac{1}{Q}(\chi - \chi_a) \tag{4.3}$$

being $\chi \equiv \omega\tau$, $y(\chi) \equiv \ln\left[A^2(\chi)/A^2(\chi_a)\right]$ and χ_a the coordinate where the power $A^2(\chi)$ is maximum (Wang, 2004). The data can thus be fitted with equation (4.3) to determine Q^{-1} by setting up a linear least-squares problem, such that the inverse of the quality factor can be determined as (Wang, 2004):

$$Q^{-1} = -\frac{\boldsymbol{x}^T \boldsymbol{y}}{\boldsymbol{x}^T \boldsymbol{x}} \tag{4.4}$$

where $\boldsymbol{x}$ is the discretised χ variable and $\boldsymbol{y}$ is a vector containing the data $y(\chi)$. The data fit is performed from χ_a to a certain small value of χ, i.e. χ_b, depending on the signal-to-noise of the data in order to cut small values of $y(\chi)$ given the noise

level of the data and the band-limited nature of the recorded seismic signals.

Figure 4.12 illustrates an example on how to estimate Q from a reflection seismic trace obtained in a homogeneous medium with constant quality factor of $Q = 80$. For this example, a source emits a Ricker wavelet that propagates through the medium, and data are recorded at different times along the vertical direction and stacked in order to simulate reflection data, which results in the waveform in Figure 4.12a after correction for spherical divergence – this is a step necessary for estimating Q with this method. Next, the Gabor transform of the data as in equation (4.1) is computed as a function of time and frequency, and the amplitude spectrum is shown in Figure 4.12b: this reveals the characteristic seismic wavelets of the recorded signals as localised energy packages. The 2D spectrum in Figure 4.12b is then transformed into a 1D spectrum by stacking all samples along curves with the same value of the product of time and frequency, i.e. $\chi = \omega\tau$, as indicated on top of Figure 4.12b for some of these values. This results in the 1D spectrum in Figure 4.12c, which decreases monotonically from the maximum value χ_a. Finally, Q^{-1} can be estimated within the support region defined by $[\chi_a, \chi_b]$ by taking logarithms of the curve in Figure 4.12c and fitting the data with equation (4.3) inside this region. In this example, I choose the minimum threshold to be $10\,dB$, but this depends mainly on the signal-to-noise of the data and the choice of the width of the Gaussian function used for the Gabor transform. With this choice, the estimated Q value is of $Q_{est} = 78.58$ with a square of the correlation coefficient of 0.99.

However, the latter only considers an average Q for the whole trace, i.e. a single subsurface layer with a constant Q value along the vertical direction. More realistically, different layers might have different Q values. Therefore, it is more adequate to perform a time-interval Q analysis by estimating Q at different intervals. This is similar to velocity analysis in seismic processing, so that average Q values from the start of the trace to different time levels are converted to a series of layered Q values. The equations on how to compute interval Q values from average ones can be found in Wang (2004) and Wang (2009). Figure 4.13 shows the application to the same example as in Figure 4.12, but Q is now computed within 5 layers with the same

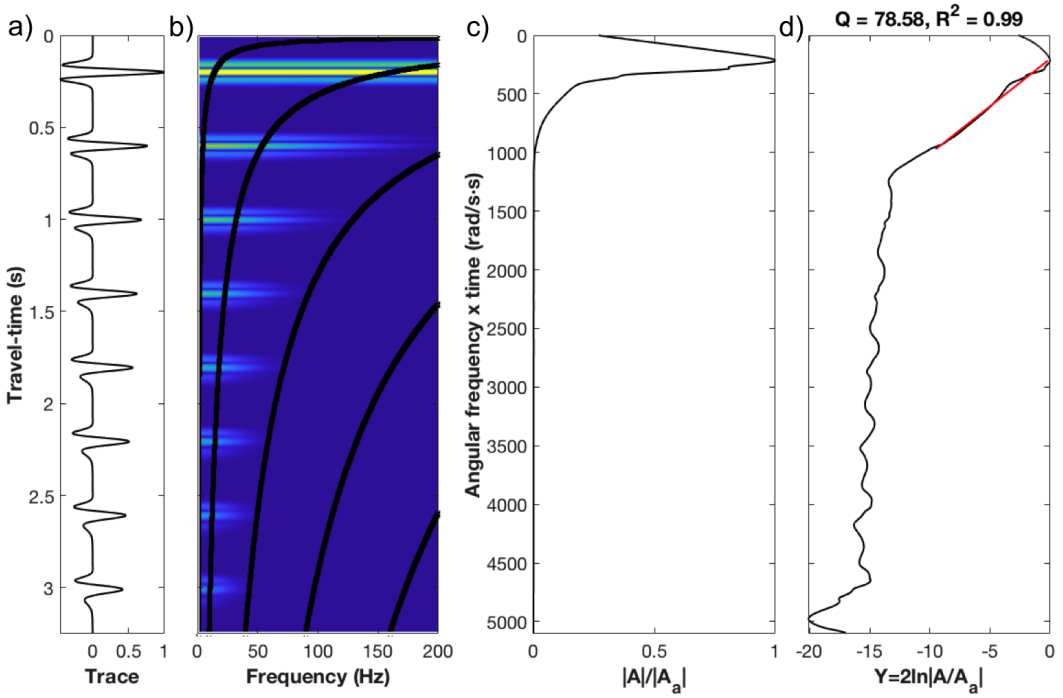


Figure 4.12.: Q analysis based on Gabor transforms of the data. Panel a) shows the recorded trace after spherical divergence correction, its Gabor transform is computed in panel b) and the latter is converted to a 1D spectrum in panel c), which is used to estimate Q by linear fit of the logarithm of c), shown in panel d). The reference Q value is 80.

threshold of $10\,dB$. Again, Figures 4.13a and 4.13b display the trace and its Gabor transform, whereas Figure 4.13c shows the interval Q values obtained as a function of time, and the relative error for each time sample with respect to the known value of $Q = 80$ is plotted in Figure 4.13d.

This results in a variation of Q with time, and hence in depth, that is close to the true value with relative errors smaller than $\pm 20\,\%$. More importantly, these are of the same order of magnitude as the true value – this is important as attenuation damps the amplitudes of a plane wave exponentially.

I implement this methodology in MATLAB and I estimate Q values in time for the observed data doing several simplifications, as only an approximate Q model is required for the data-matching workflow to be successful. One of the requirements of the selected Q estimation method is that the data needs to be a reflection dataset. Instead, I extract the first trace for all shots along the seismic line – the separation between each shot and the first receiver is $100\,m$– and I use this to estimate Q, as a

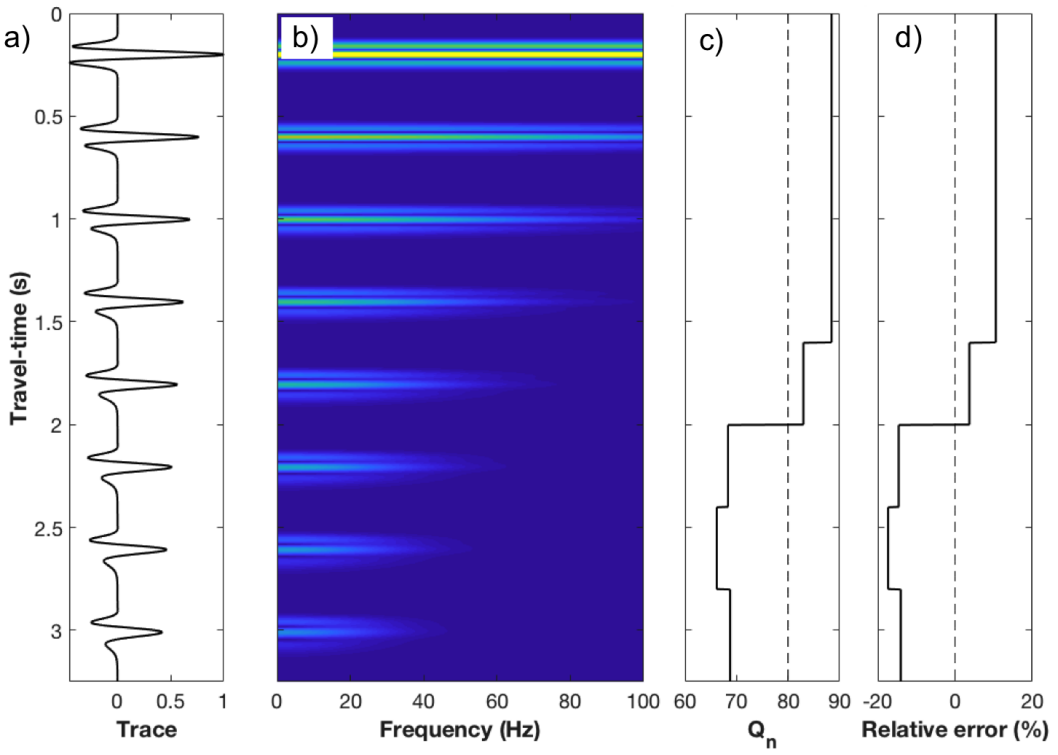


Figure 4.13.: Interval Q analysis for a reflection trace as a function of time. Panel a) presents the trace after spherical divergence correction, the Gabor transform is plotted in panel b), the computed Q values as a function of time are shown in c) and the relative error of the estimated values in c) as a function of time is shown in d). The reference Q value is 80.

reflection section is not available. The latter is a near-offset section and is displayed in Figure 4.14a. However, this is not suitable for Q estimation as it also contains the direct arrival through the water. Thus, I first remove the direct arrival by estimating the depth of the water-bottom below each extracted receiver given the starting P-wave velocity model, a fixed water velocity and the assumption that the extracted section is a reflection section: hence, $t_{sb} = 2z_{sb}/v_{water}$ with z_{sb} the depth of the water bottom according to the contractor's starting model. This results in the near-offset section in Figure 4.14b: the direct arrival is efficiently removed, hence the approximation of a reflection section seems suitable.

The next step is to correct the data for spherical divergence effects. I perform the latter by using an RMS velocity model built from the starting model and multiplying each time sample by the time of the sample and the corresponding RMS velocity sample (Newman, 1973), which results in the section displayed in Figure 4.14c. The

latter is the dataset used to estimate Q using Gabor transforms.

Q-analysis is done trace-by-trace for the entire seismic line but, first, I study the choice of the Q analysis' parameters on a a few single traces. Figure 4.15 shows this study for extracted trace number 600, plotted in Figure 4.15a as a function of time. The panels in Figures 4.15b to 4.15f display the Gabor transforms of the seismic trace in which I overlay the time and frequency windows (in white) for which 1D amplitude spectra are computed, as well as curves of constant product of time and frequency (black crosses). Panels in Figures 4.15g to 4.15k show the computed 1D amplitude spectra within the region specified in the panels above, and the logarithm of the latter plots as well as the linear fit to determine the quality factor for each region are displayed in panels in Figures 4.15l to 4.15p. The quality factor for each of this plots is displayed at the top of each of the previous panels.

This example for trace 600 shows that linear fits of the data provide relatively sensible values of Q between the range from 25 to 200 for a threshold of $5\,dB$ (and the same size of the Gaussian window function for the Gabor transform). After experimenting on other traces, such a threshold value provides reasonable results across the dataset. Thus, I use this value and I estimate Q trace-by-trace in 6 time intervals for each trace, defined from the time at the sea-bottom as the vector $[t_{sb} + 1.2, t_{sb} + 1.6, t_{sb} + 2.0, t_{sb} + 2.5, t_{sb} + 4.2, t(end)]$, where all the units are in seconds and $t(end)$ is the total time for the recorded data.

Estimation of Q using Gabor transforms for multiple layers can be somehow expensive if the dataset is very large, and parallelisation of the code can be used to drastically reduce the computation time. However, the code I used for Q estimation here was run on a single computer node, given the near-offset dataset was not too large. In total, Q estimation took between three to four hours for the entire dataset and resulted in the model shown in Figure 4.14d. The estimated Q values increase with time (and hence with depth), and range from 20 near the sea-bottom to $400-500$ at $8\,s$. This is consistent with the geology of the area, as there is no gas-cloud or fine layering and viscous effects are expected to be larger at shallow regions as sedimentary rocks are less compacted, resulting in a larger loss of energy.

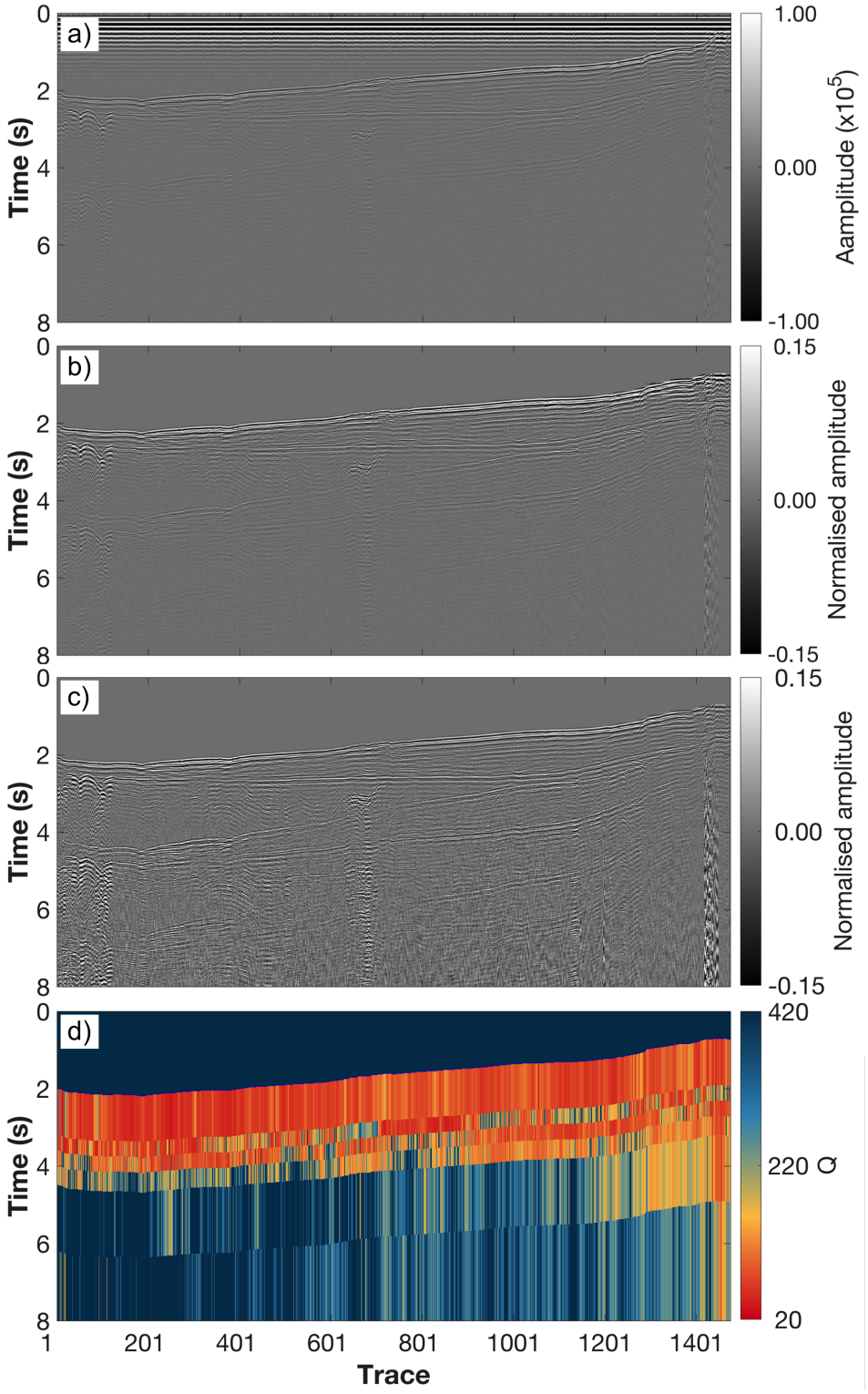


Figure 4.14.: Extracted time sections for a) near-offset data, b) near-offset data after removing the direct arrival, c) near-offset data after removing the direct arrival and spherical divergence and d) the estimated Q model obtained using the data in c).

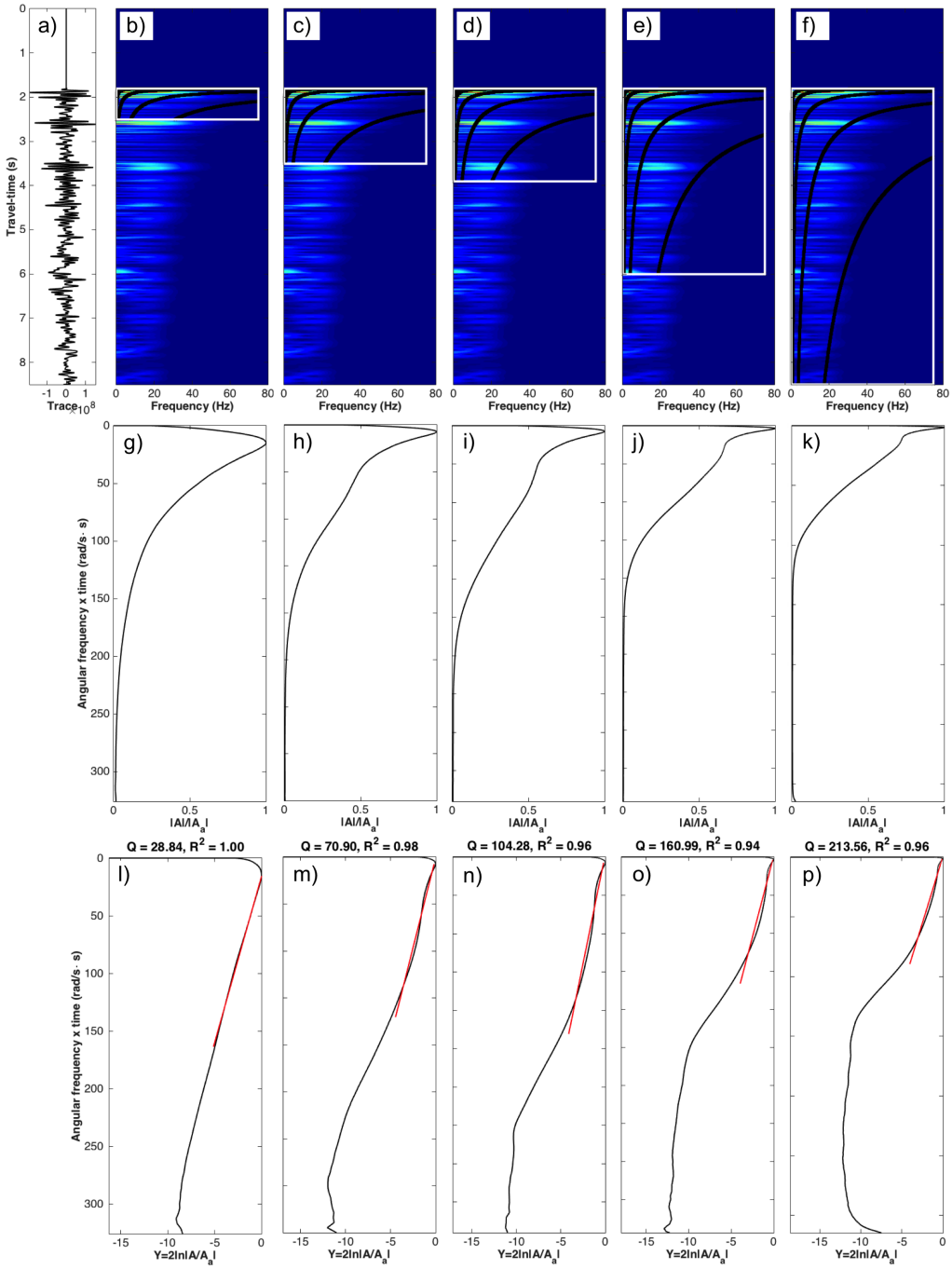


Figure 4.15.: Seismic Q analysis for trace 600 of the extracted near-offset section in Figure 4.14c using Gabor transforms. Panel a) shows the extracted trace, panels b) to f) display the Gabor transforms and the windows in which 1D spectrum are computed (white window) with curves of constant $\chi = \omega\tau$ overlaid (black crosses), panels g) to k) show the computed 1D spectra for the panels above and panels l) to p) display their logarithm, the linear fit (in red) for a threshold of $5\,dB$ and the estimated Q value (at the top).

The latter estimated Q model needs to be converted to depth in order to be used to model visco-acoustic data with existing modelling codes. Thus, I convert the model in Figure 4.14d to depth by using the starting velocity model and a standard time-to-depth conversion workflow that uses the starting velocity of interval velocities in depth, it converts it to average velocities in time and uses that to convert the Q model in time to depth. The resulting estimated model in depth is shown in Figure 4.16a, whereas Figure 4.16b displays the latter after severe lateral smoothing performed differently within three depth levels in order to impose lateral continuity of the Q model.

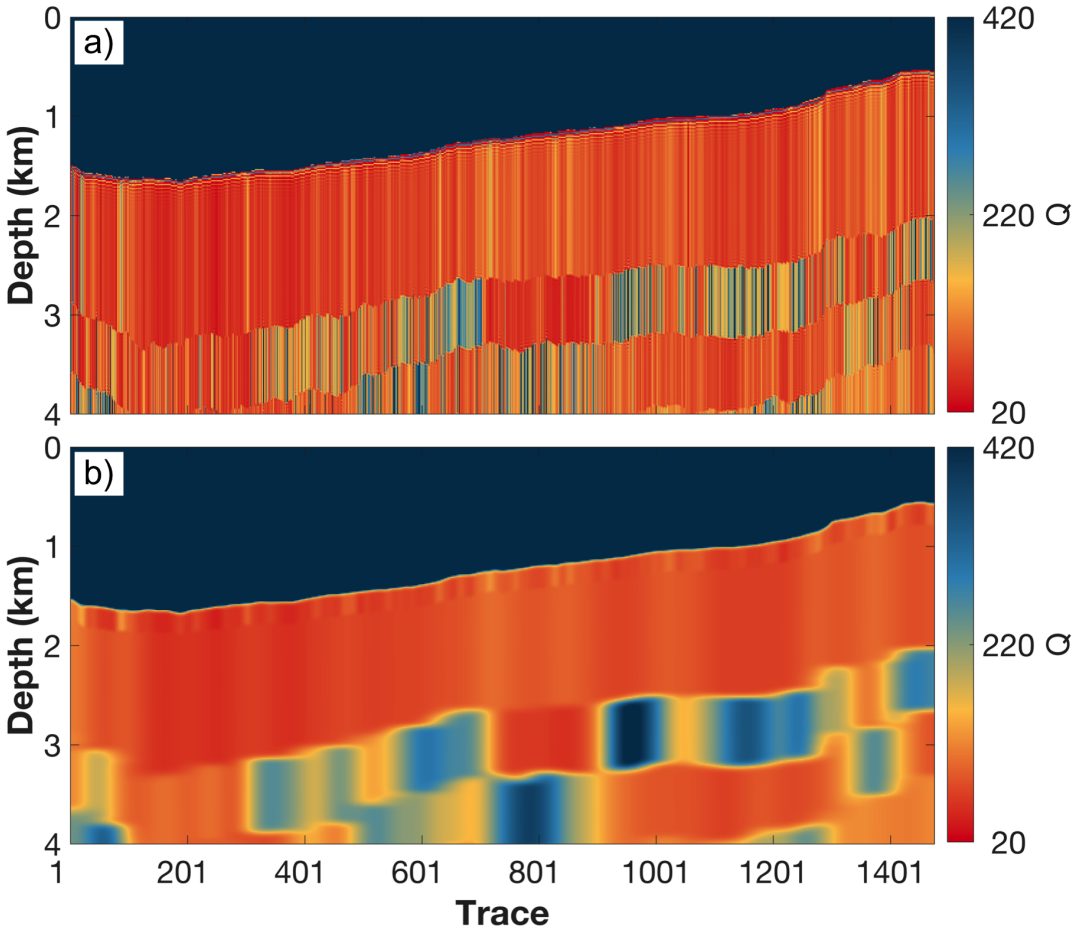


Figure 4.16.: Models of the estimated quality factor in depth a) before lateral smoothing and b) after lateral smoothing.

Finally, the Q model is extended to the length of the model by using the location of sources and the first receiver for each shot, from which the data has been extracted, interpolating the values for Q for those cells in the model that fall between two of

the extracted traces and repeating either the first or last extracted trace for those columns in the model that fall either before or after the position of the first or last extracted traces. In this step, I also take into account the depth of the sea-bottom for each column of the model, and that Q must be very large within the water layer.

At this stage, all files are ready for the application of the data-matching workflow and the inversion of the data. The starting and estimated models used for this purpose are displayed in Figure 4.17, which shows the starting contractor's velocity model (Figure 4.17a) – this is based on pre-stack time migration stacking velocities (Kalinicheva et al., 2017)–, the final version of the estimated Q model (Figure 4.17b) and the contractor's models for the Thomsons' parameters for anisotropy, δ and ϵ (Figures 4.17c and 4.17d).

In summary, the estimated Q model obtained in this section is approximated for the following reasons:

1. The dataset used to extract a Q model is not a pure reflection dataset, but a near-offset reflection dataset.

2. Surface-related multiples are kept in the data in order to avoid changing the spectra of the data incorrectly before the first multiple, which results in over-estimating Q values after the first multiple – fortunately, this is well below the model in Figure 4.17a, starting slightly below $4\,s$ in Figure 4.14c.

3. A limited number of time intervals is used to estimate Q trace-by-trace.

4. Spherical divergence is corrected using an average RMS velocity model.

Nevertheless, in the next section I demonstrate the latter is sufficient to improve the quality of the recovered velocity models using the data-matching workflow.

4.4.2. Results

Simultaneously to the previous Q analysis, VTI anisotropic acoustic FWI is performed on the minimally pre-processed observed data. The models used for the anisotropic parameters are those in Figures 4.14c and 4.14d, and these are fixed during the

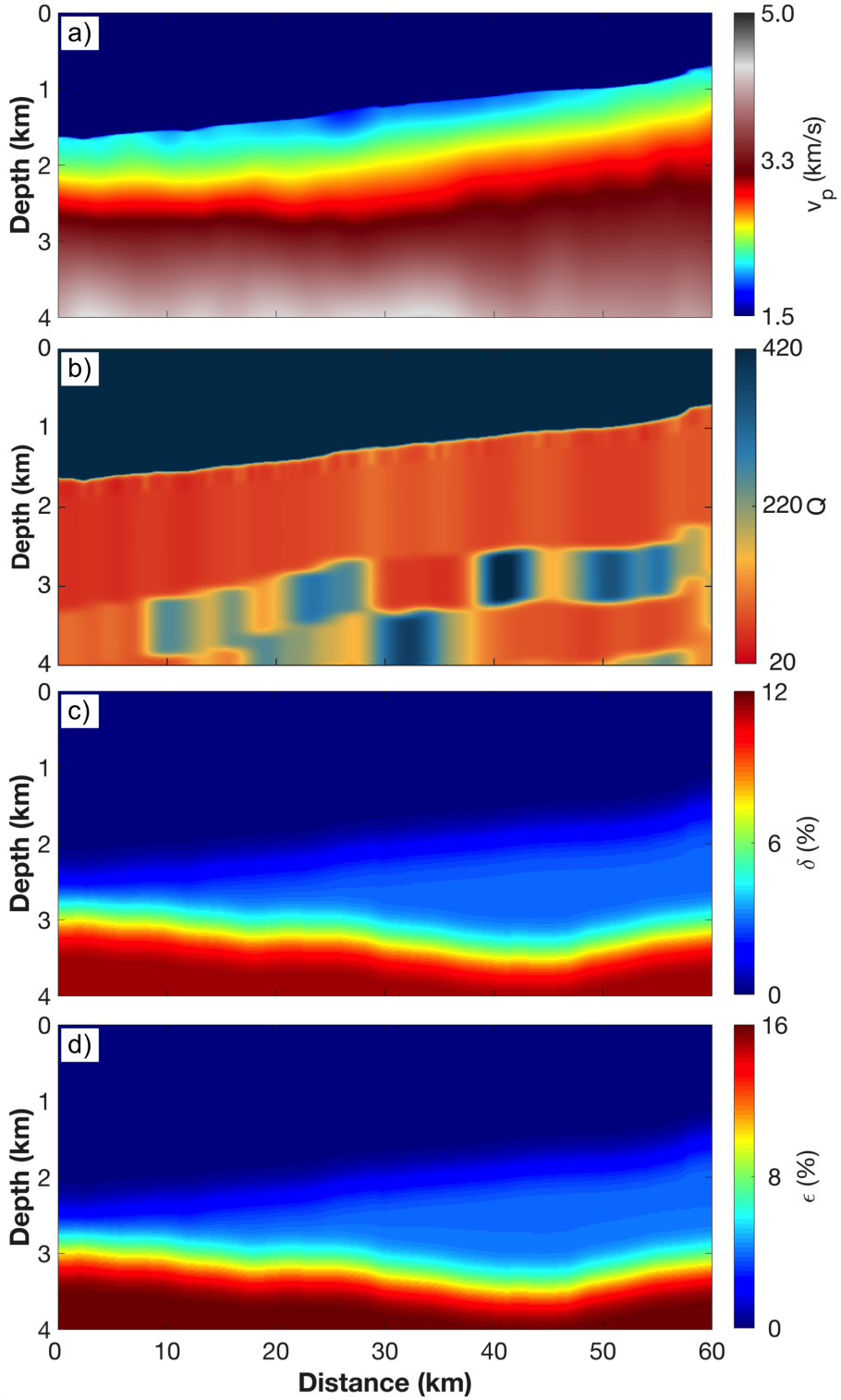


Figure 4.17.: Vertical slices of a) the contractors' starting velocity model, b) the estimated Q model and the contractor's models for the anisotropic Thomson's parameters c) δ and d) ϵ.

inversion. The data is inverted in a multi-scale fashion by inverting from 2.5 Hz to 24 Hz in blocks of 4 iterations per frequency band and using an amplitude-balanced functional. Anisotropic acoustic FWI of the observed data using the starting model in Figures 4.17a and 4.18a results in the recovered velocity model in Figure 4.18b. The latter has a higher level of detail when compared to the starting model and contains various features that are not present in the former, such as the presence of low velocity channels in the shallow subsurface, a major unconformity that crosses the model from left to right at depths between 2 km and 2.5 km depth (indicated by a black arrow) and reflectors below 2.5 km depth, which are interpreted as basaltic intrusions (Kalinicheva et al., 2017).

In order to improve the quality of the recovered velocity model by mitigating visco-acoustic effects, a strategy for implementing the data-matching workflow needs to be chosen. As explained in previous sections, this depends on the estimated Q model and the level of noise in the observed data. Thus, I use workflow 2 to account for possible large viscous effects as the Q estimate contains relatively low Q values in the shallow subsurface, which leads to strong absorption effects, and the signal-to-noise ratio of the data is high. In other words, I mitigate visco-acoustic effects directly from the starting model and the estimate of Q by first modelling acoustic and visco-acoustic data using the latter, computing Wiener filters that match the visco-acoustic to the acoustic data and applying them to the observed data, which leads to a matched dataset. For the present application, I match 15 traces simultaneously, i.e. a spacing of 175 m between the first and last traces matched simultaneously, and a length of the filters of 325 ms. The matched data is then acoustically inverted to obtain a recovered P-wave velocity model. Overall, the elapsed compute time is between one and two times the cost of a conventional acoustic inversion. In this workflow, modelling of acoustic and visco-acoustic data is performed without considering anisotropy given that the model of Q is isotropic, whereas the inversion is anisotropic acoustic with fixed anisotropic parameters – equals to those provided by the contractor.

The resulting recovered velocity model in Figure 4.18c shows the benefits of the data-matching workflow. Although differences in the shallow subsurface with respect

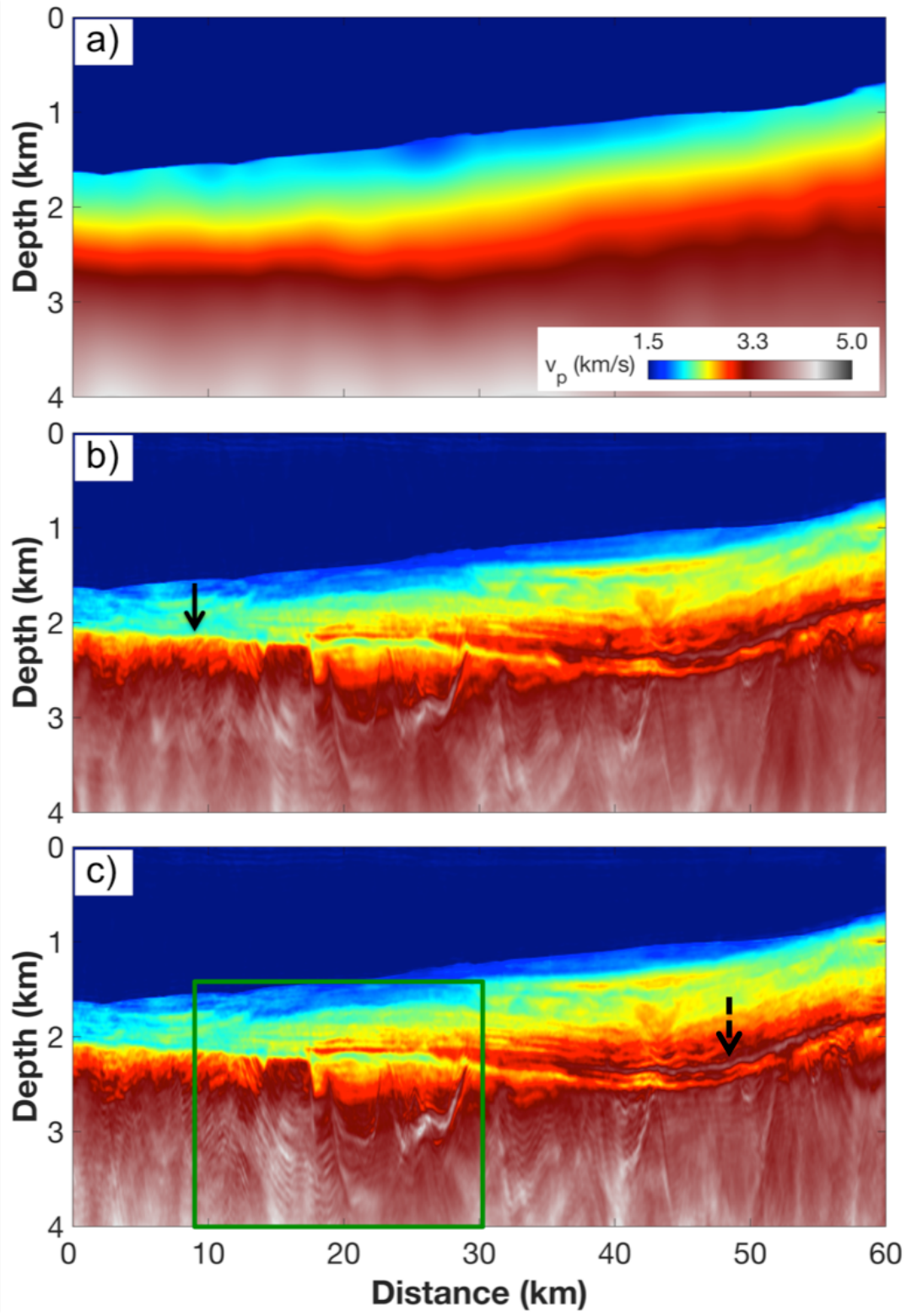


Figure 4.18.: Vertical slices of the a) starting model and the recovered P-wave velocity models after VTI anisotropic acoustic FWI of the b) field observed data and the c) matched dataset.

to the model in Figure 4.18b are small, the velocity model obtained from the matched data in Figure 4.18c is in average faster below the unconformity, it is overall more detailed – for instance, there is a thinning of the high velocity layer indicated by a dashed black arrow – and structures below the unconformity are now clearly seen and can be followed across the model. In order to better observe the differences below the unconformity, Figures 4.19a and 4.19b display a zoomed-in section of the models in Figures 4.18b and 4.18c, respectively, corresponding to the area inside the green box in Figure 4.18c. In this comparison it can be observed that layers below the unconformity are better identified and are more continuous in Figure 4.19b when compared to the model in Figure 4.19a, and vertical features that may indicate faulting – this is indicated by blue arrows – can also be detected. Thus, we expect the latter should lead to a more accurate and better resolved migrated section.

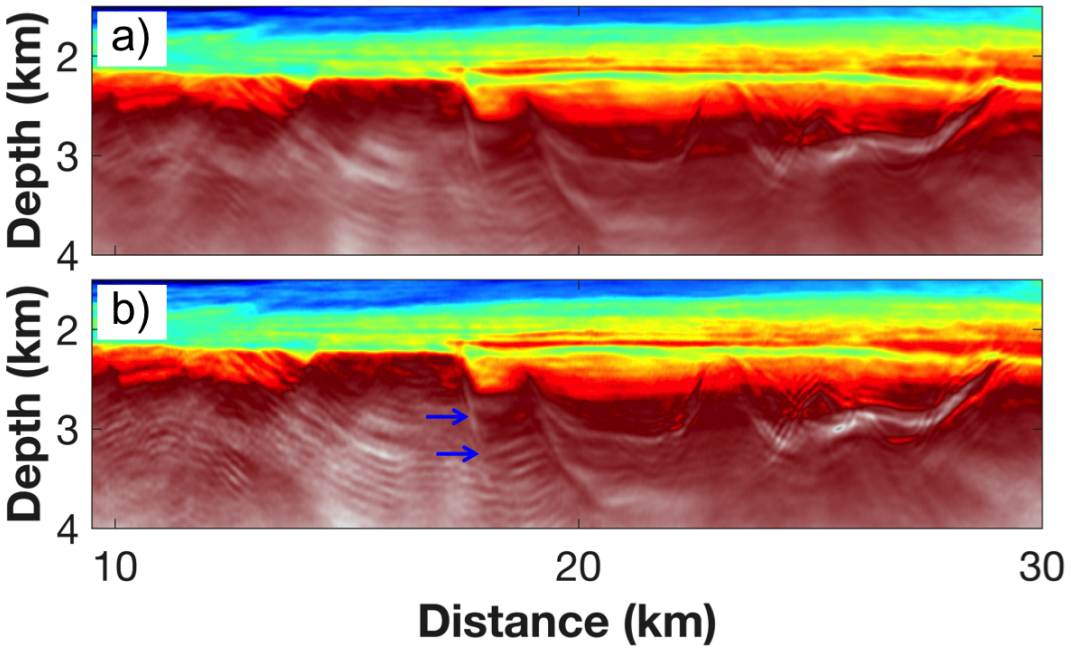


Figure 4.19.: Zoomed-in sections of the recovered models in Figures 4.18b and 4.18c for the recovered models after VTI anisotropic acoustic FWI of a) observed and b) matched data.

In terms of the changes with respect to the observed data, these are compared in Figure 4.20, which compares the observed data after minimal pre-processing for two different shots with the respective matched data. The two shot gathers shown in Figures 4.20a and Figure 4.20b have been recorded along the same seismic line,

and with a shot separation of 22.5 km. The level of noise in the data is low, and only some localised random noise can be observed for the first 0.5 km of the second shot gather: this justifies the use of the second data-matching workflow instead of workflow 3. After the application of the data-matching workflow, the corresponding matched data are displayed in Figures 4.20c and 4.20d for the first and second shot gathers, respectively.

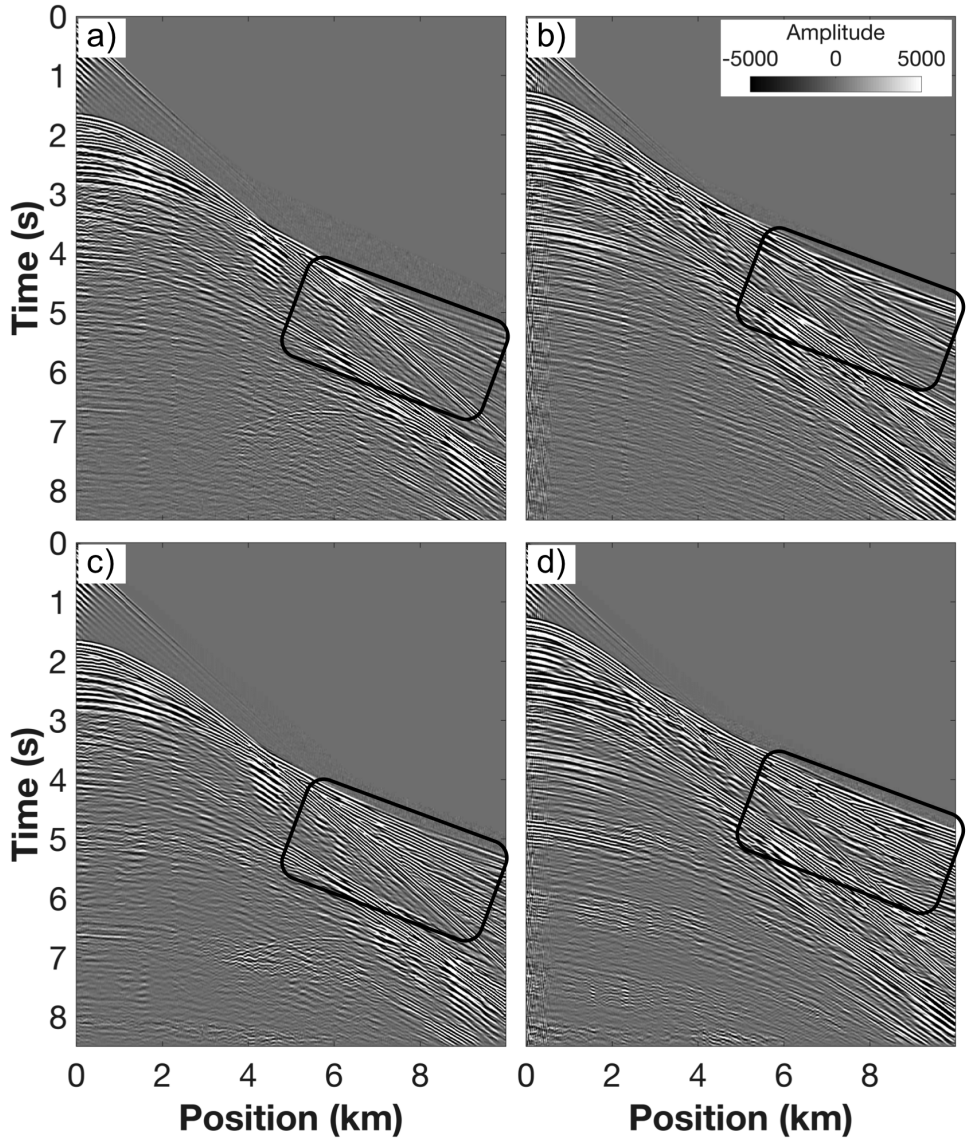


Figure 4.20.: Recorded field data are displayed in panels a) and b) for two shots along the same sailing line and separated 22.5 km, whereas the corresponding matched data for these shots are respectively shown in panels c) and d) after matching 15 traces simultaneously and using a length of the filters of 325 ms.

As expected, the amplitude of the events increases with respect to the observed data after mitigating visco-acoustic events, for instance for the refractions inside the black boxes in the different panels in Figure 4.20. Additionally, some of the noise present in the shot gather in Figure 4.20b is mitigated in the data in Figure 4.20d, given that matching several traces simultaneously introduces some smoothness in the matched data. Similar to the tests for the Marmousi2 model, phase changes are difficult to observed as there are many events in the data and changes are of the order of milliseconds. I also want to highlight that the choice of the matching parameters is such that little artefacts due to filtering can be observed on the matched data, yielding the high quality recovered models in Figures 4.18c and 4.19b.

In order to assess the quality of the results and analyse whether the differences in the recovered P-wave velocity model after the application of the data-matching workflow are real improvements, I perform 2D anisotropic Kirchoff post-stack depth migration of the data: the flatness of common-image gathers (CIGs) and the fit of the recovered velocity model with the stacked migrated sections are used to assess these improvements. First, I migrate the observed data using the recovered model obtained after anisotropic acoustic FWI of the observed data in Figure 4.18b, which results in the CIGs panels shown in Figure 4.21a. I then migrate the observed data after applying Q-compensation and band-pass filtering the data – this is done with the same filter used to band-pass filter the observed data during pre-processing to have the same bandwidth of the data – and using the recovered velocity model obtained after anisotropic acoustic FWI of the matched data, resulting in the CIGs shown in Figure 4.21b.

The latter panel of CIGs shows that these are slightly flatter after Q-compensation and using the recovered P-wave velocity model after the data-matching workflow, as indicated in different locations by blue arrows. Additionally, the events are thinner, amplitude effects are compensated for, especially with depth, and the position of some of the events changes slightly such that they are shallower than those in Figure 4.21a after accounting for absorption effects. Despite these improvements, I note that a more sophisticated migration algorithm that is able to account for visco-acoustic

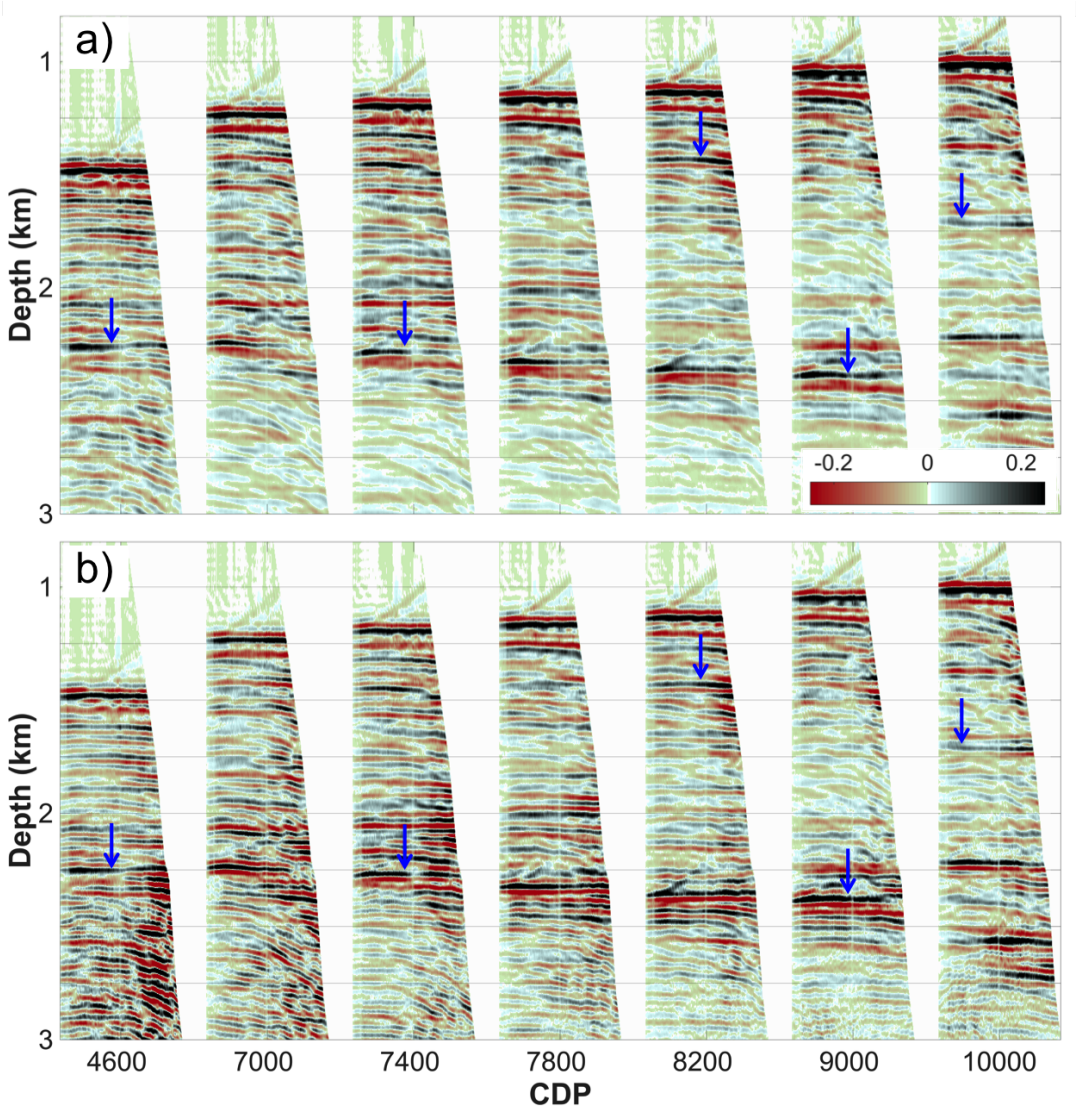


Figure 4.21.: Flattening of CIGs for a) the observed data using the recovered velocity model in Figure 4.18b, and for b) the observed data after Q-compensation with the Q estimated model in Figure 4.17b and the velocity model in Figure 4.18c after band-pass filtering.

effects during migration should be used, but I do not have access to such software.

Finally, Figures 4.22a and 4.22b show the stacked sections for the CIGs in Figure 4.21.

The use of the Q estimated model and the recovered velocity model result in a stacked migrated section in which the layers are better resolved, hence thinner, and more continuous, especially below the unconformity. In fact, layers below the un-

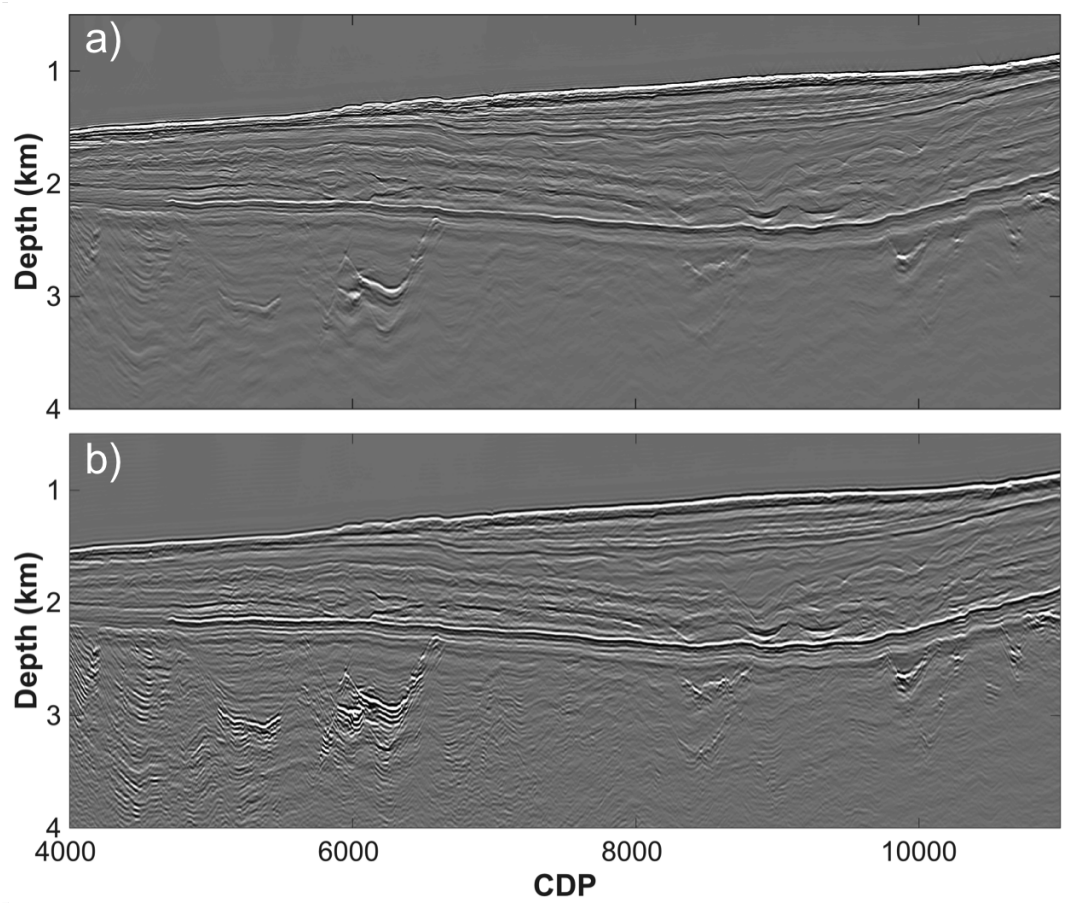


Figure 4.22.: Stacked migrated sections obtained from the observed data a) before Q-compensation and using the recovered P-wave velocity model after anisotropic acoustic FWI of the observed data and b) after Q-compensation and band-pass filtering using the recovered model after anisotropic acoustic FWI of the matched data.

conformity are unravelled, which spread from comon depth points (CDPs) 4000 to 6500. This is consistent with the improvements observed below the unconformity in Figures 4.18c and 4.19b, as demonstrated in Figure 4.23, in which the corresponding recovered velocity models are overlaid on the stacked migrated sections.

The layers observed in the recovered velocity model below the unconformity after applying the data-matching workflow do correspond to layers in the stacked migrated section in Figure 4.23b, whereas these are less visible on the section in Figure 4.23a.

I conclude that the data-matching workflow leads to improvements in the quality of the recovered P-wave velocity models obtained with acoustic FWI in the time-domain,

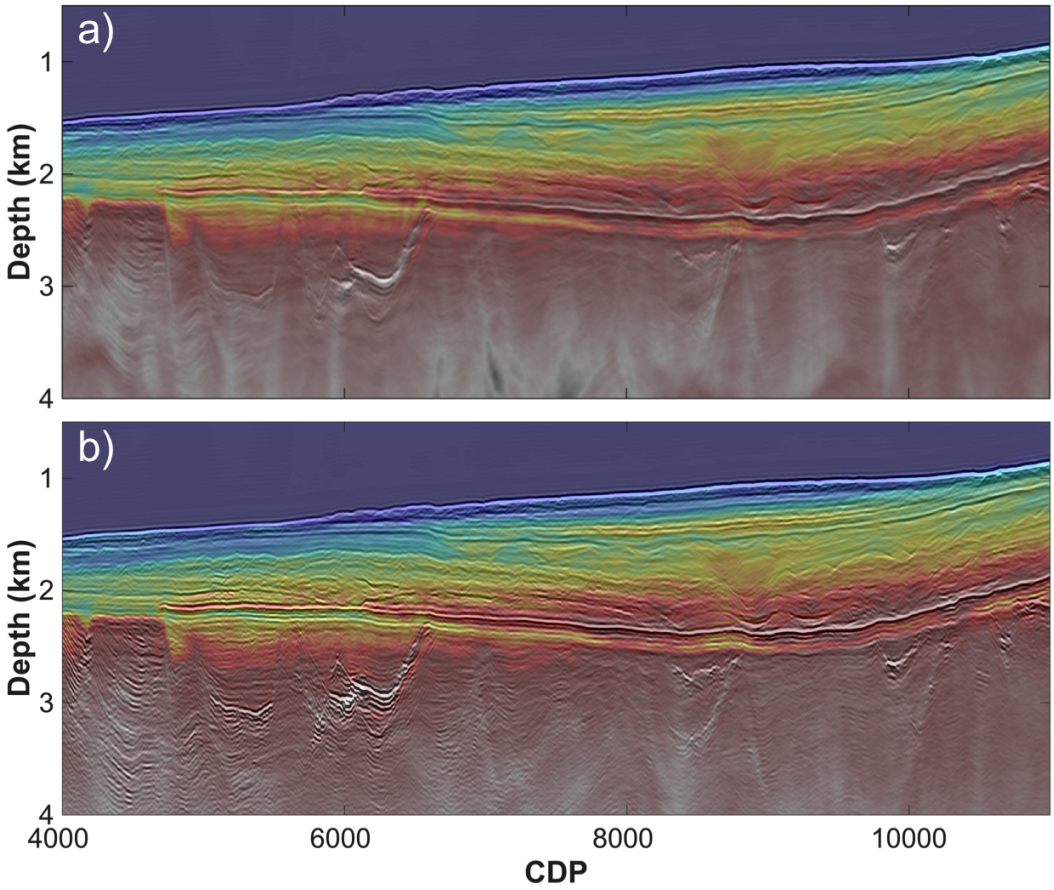


Figure 4.23.: Stacked migrated sections in Figure 4.22 overlaid with the velocity models in Figures 4.18b and 4.18c, which have been used to migrate the corresponding data.

also of field data, at a relatively low cost. Hence, I suggest the proposed method in combination with an amplitude-balanced functional as an efficient strategy to mitigate visco-acoustic effects in acoustic FWI.

4.5. Summary

The viscous nature of the subsurface modifies the amplitude and phase of the recorded data compared to that recorded in a viscous-free subsurface. In this chapter I have shown that not considering visco-acoustic effects has adverse consequences on the quality of the recovered P-wave velocity models obtained with acoustic FWI, yielding poorly resolved and inaccurate P-wave velocity models, even when an amplitude-balanced objective functional designed to mitigate these effects is utilised. With the

aim of improving the quality of the recovered velocity models, I have adapted the methodology implemented in chapter 3, first designed to mitigate elastic effects and based on Wiener filters, in order to address visco-acoustic effects. The application of the proposed data-matching method in this chapter, both on the synthetic and field data, demonstrates the benefits in terms of accuracy and resolution of the recovered velocity models with respect to those obtained with conventional acoustic FWI. This has been verified against well-log profiles and comparisons with the true model for synthetic data, and the flatness of the CIGs and correlation of the migrated sections and the velocity models for field data.

As the inversions using the proposed method are always acoustic, the data-matching workflow represents an improvement in compute cost when compared to visco-acoustic FWI in the time domain, in which several relaxation mechanisms need to be modelled to account for a constant quality factor with frequency.

To overcome the limitations of the original workflow in the presence of strong viscous effects or noisy data, I have suggested and implemented two alternative workflows, namely workflows 2 and 3, that consist on mitigating visco-acoustic effects using the starting model directly and using a starting model for the acoustic inversion of matched data obtained by smoothing the recovered model of the acoustic inversion of acoustic data, respectively. Workflow 2 successfully mitigates strong viscous effects and results in a computation time slightly over that of a single conventional acoustic inversion, as shown on a synthetic and a field dataset, whereas workflow 3 has a cost equal to the original workflow and its application has only been shown on a realistic synthetic dataset.

The tests performed on synthetic data in this chapter have also been useful to conclude that a smooth and approximate background Q estimate is sufficient to mitigate the most relevant visco-acoustic effects. Due to this result, I have performed an approximate Q analysis of near-offset field data that resulted in a smooth Q estimate in depth. The assessments on the quality of the recovered models using the latter and

the data-matching workflow also indicates that using such an approximate Q model is enough to improve the quality of the recovered P-wave velocity models. Nevertheless, more accurate methods to estimate Q, for instance by hybrid local-global methods such as that of da Silva et al. (2017), could further improve the results.

The conclusion with respect to the filtering parameters when computing matching filters from the tests in this chapter are similar to those in chapter 3: I recommend a length of the filters between $300\,ms$ and $400\,ms$ and matching several traces simultaneously such that the spacing between first and last matched neighbouring traces is at least $100\,m$, and up to $250\,m$. I want to highlight that the latter was larger for elastic effects, of $800\,m$, but this depends on the dataset: for instance, elastic data with slower velocities will contain events that have a larger dip (i.e. that are more vertical), and hence more lateral variation. Thus, I recommend to test the filtering parameters on a single shot gather first, and assess the quality of the results. For instance, the presence of ripples on the matched data that are not present in the observed data may indicate this is caused by the choice of matching parameters. For future implementations, I suggest using an automatic method for selecting the filtering parameters. For instance, this could be based on testing different parameters randomly within a certain range for a set of random traces, selecting those parameters that lead to the smallest RMS difference with respect to the modelled acoustic traces and using the selected parameters to compute matching filters for the entire dataset. However, this has not yet been tested.

Another important remark about the data-matching method is that this can address visco-acoustic effects for complex geologies, as wave-field modelling is performed, and accounts for multiples and seismic ghosts. On the other hand, inverse Q-filtering methods that remove viscous effects from the data are not suited for acoustic FWI, as these assume a 1D-layered earth (Guo and McMechan, 2015). What is more, the data-matching workflow implemented here can be combined with the data-matching workflow used in chapter 3 to mitigate visco-elastic effects. This would consist of the same steps as those in the workflows in Figures 2.8 and 2.9, but acoustic and

visco-elastic data would need to be modelled in Step 2. Thus, it would require an estimate of v_p/v_s and also estimates for Q_p and Q_s, the quality factors associated with P- and S-waves. Given that only smooth and approximate estimates are required for the data-matching workflow, these estimated models could be obtained by using a hybrid local-global inversion approach as that suggested by da Silva et al. (2017) by parametrising the subsurface with only a few parameters. Acoustic inversion of the resulting matched data would, in principle, lead to an improved velocity P-wave velocity model without the need of performing visco-elastic FWI and the consequent cost of modelling elastic waves and viscous effects in the time domain during the inversion.

Finally, application of the data-matching algorithm to account for 3D visco-acoustic (and also visco-elastic) effects is straightforward. Nevertheless, this requires estimating a 3D Q model, for instance, by using Gabor transforms. Given that this can be computationally demanding if multiple time layers are used, I recommend using a parallel code to estimate Q in which the computation for each trace is distributed across multiple nodes and tasks. Once a model of Q is obtained, the remaining steps of the data-matching workflow are unchanged and only the acquisition geometry needs to be considered when matching neighbouring traces simultaneously. Further research will focus on implementing the current code to 3D field datasets.

5. 3D imaging of the breast with ultrasound

Medical imaging methods aim at obtaining visual representations of physical properties of the body that are otherwise invisible to the human eye. These are vital, as they provide the means for clinicians to diagnose and treat diseases.

In this chapter I explore the potential of FWI of ultrasound data in imaging breast cancer, both on an *in silico*[1] and an *in vivo*[2] breast datasets.

I have presented the research work carried out in this chapter at the international MUST conference in Speyer (Germany):

> Agudo, O. C., Guasch, L., Huthwaite, P., and Warner, M. (2018c). 3D imaging of the breast using full-waveform inversion. In *Proceedings of the International Workshop on Medical Ultrasound Tomography*, pages 99–110.

This chapter contains the results published at this conference, as well as additional intermediate results to provide a better overall understanding. I would like to acknowledge the contributions of two researchers to this work: first, I want to thank Neb Duric and the Karmanos Cancer Institute (Detroit, USA) for providing me with the *in vivo* data and the permission to publish the results and, secondly, Peter Huthwaite (Mechanical Engineering Department, Imperial College London) for giving me access to the synthetic breast phantom and providing me with a starting model and source calibration for the *in vivo* data.

[1] The term *in silico* means performed in a computer, and hence an *in silico* model or dataset is equivalent to a synthetic model or dataset in geophysics.

[2] An *in vivo* experiment is that in which the latter is performed on live cells or organisms in the laboratory. This is analogous to a field dataset in geophyiscs.

5.1. Introduction

Breast cancer is the second most common type of cancer worldwide and the most common amongst women (Ferlay et al., 2015). In Europe, breast cancer is the leading cause of cancer death amongst women, and the third overall, which caused 113,000 deaths in 2012 (Ferlay et al., 2013, 2015). Advances in medical imaging techniques are crucial to help detect breast cancer at an early stage, thus accelerating diagnosis and treatment. In the last decades, mammographies have been frequently used for standard routine screenings, principally due to their low cost. Mammographies consist on the projection of low-energy X-rays through the breast, and the formation of an image based on the different absorption of X-rays for normal and cancerous tissues due to their different densities. The main benefit of mammography screening is the reduction in deaths due to breast cancer (Løberg et al., 2015), but there are also several risks associated with the latter:

1. It uses ionising radiation, which can induce cancer (400 to 800 breast cancers were detected in England in 2017 for every cancer induced, according to Warren et al. (2017)).

2. It can lead to overdiagnosis (Løberg et al., 2015), i.e. detection and treatment of tumours that are not life-threatening due to lack of adequate markers (based on recent estimates from the United States, 15 out of 1,000 women being invited to biennial mammography scans are overdiagnosed according to Løberg et al. (2015)).

3. It can lead to false-positive diagnostics when trying to detect breast cancer early (Heywang-Köbrunner et al., 2011; Løberg et al., 2015) (the risk of experiencing a false-positive diagnostic in Europe for biennial screening is approximately 20 %, according to Hofvind et al. (2012)).

4. It has a low sensitivity for dense breasts, leading to false-negative diagnostics (Given-Wilson et al., 1997; Murphy et al., 2007)

An alternative method for breast cancer screening is ultrasound tomography, which utilises ultrasound to provide images of physical properties of the breast, such as the

propagation speed-of-sound[3] in tissue, absorption of sound waves and tissue density. The use of ultrasound can overcome some of the disadvantages of mammographies. For instance, ultrasound waves are non-ionising radiation and, hence, there is no risk of induced cancer. Moreover, ultrasound tomography results in quantitative models of physical properties of the tissues, which can potentially be used to separate different regions and reduce the number of overdiagnosed patients and false-positive and false-negative diagnostics. Additional benefits of ultrasound tomography are the portability of the imaging system and the fact that the acquisition is painless – unlike the possible discomfort associated with a mammography. In such systems, as that shown in Figure 5.1a, the patient lies prone on a bed with a breast suspended in a warm water bath (Figure 5.1c), in which ultrasound waves are emitted and received by ultrasound transducers[4] (Figure 5.1b). The data is then analysed and a model of physical properties of the breast is obtained.

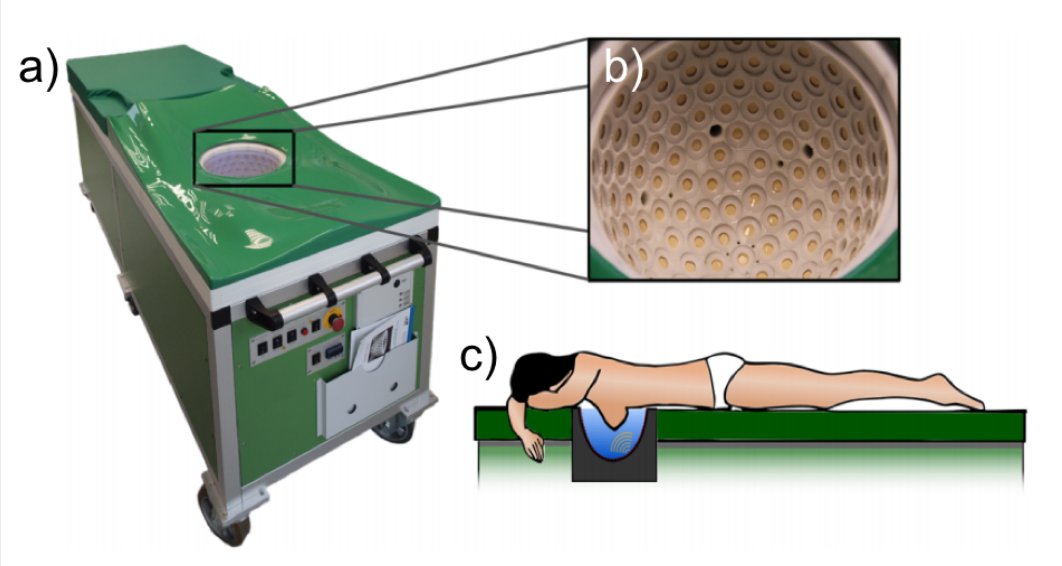


Figure 5.1.: Figure a) is a picture of an ultrasound tomography scanner, in which ultrasound waves are emitted and received by an array of transducers distributed in a semi-spherical surface, as in b). Figure c) is a drawing of the positioning of the patient during data acquisition. Modified after Ruiter et al. (2018).

In the example in Figure 5.1b, transducers are located around the breast in a

[3]Speed-of-sound refers to the velocity of acoustic waves in a medium, and is equivalent to P-wave velocity in geophysics, v_p.
[4]Transducers are sensors that can act both as sources and receivers.

semi-spherical geometry. Other ultrasound tomography scanners use circular rings of transducers, which can be fixed or moved to acquire data at different planes. Tomographic reconstruction algorithms are then used to reconstruct a model of the breast. Typically, ray-based algorithms are used to obtain a model of speed of sound, which can give vital information of the presence of a tumour given that the latter are generally denser and sound propagates faster than in the surrounding tissue.

However, the spatial resolution of ray-based methods such as time-of-flight (TOF)[5] (see section 2.4) is limited by the size of the first Fresnel zone (Williamson, 1991). In practice, this results in a spatial resolution of $10\,mm$, which means that small cancer masses, typically of the size of $5\,mm$, are difficult to detect using ray-based algorithms (Huthwaite et al., 2010; Huthwaite, 2012). Thus, some researchers have proposed the use of diffraction tomography algorithms (Simonetti and Huang, 2008) to recover material properties by analysing the distortion in space of the wavefield scattered at the target under the Born approximation for the incident wavefield, such that the total wavefield can be computed as an incident plane wave and a scattered wavefield (Simonetti and Huang, 2008; Müller et al., 2015). These methods can resolve smaller features, but fail in reconstructing larger objects, especially when the contrast in speed of sound between objects is large (Huthwaite et al., 2010). For this reason, methods that combine TOF tomography and diffraction tomography have been suggested as an alternative (e.g. the Hybrid Algorithm for Robust Breast Ultrasound Tomography or HARBUT, proposed by Huthwaite (2012)).

On the other hand, full-waveform inversion (FWI) algorithms honour the physics of finite-frequency wave propagation to model wave propagation through an object at sub-wavelength scales and, therefore, are able to model all type of waves. Thus, FWI methods are, in principle, able to overcome the limitations of other methods in terms of spatial resolution, but at a much larger computational cost. Nevertheless, 3D FWI is nowadays standard practice in the oil industry thanks to recent advances in computation and the development of more robust FWI algorithms. Although FWI algorithms have been applied in the past to recover 2D models from *in vivo* data of

[5]This is equivalent to travel-time tomography (TT) in geophysics.

the breast (e.g. Pratt et al. (2007)), its implementation to 3D ultrasound datasets of both *in vivo* and *in silico* data have not yet been demonstrated, partly due to the lack of adequate data-acquisition systems and robust 3D FWI codes.

Within this chapter, I first demonstrate the potential of FWI to reconstruct 3D speed-of-sound models of an *in silico* ultrasound data of the breast. I then investigate and emphasise the benefits of acquiring low-frequency ultrasound data (below $0.5\,MHz$) to obtain highly detailed and accurate recovered models of the breast, and study the consequences of ignoring density, absorption and 3D effects. Finally, I apply the same FWI algorithm to a 2D *in vivo* ultrasound dataset with a limited knowledge on the acquisition geometry and the source wavelet, which results in a high-resolution model of a breast and the possible detection of breast cancer. Furthermore, I stress again the importance of acquiring low-frequency data to obtain a highly accurate and resolved model with FWI for an *in vivo* dataset, and I briefly discuss the impact that neglecting 3D effects has on the recovered model. In the last section of this chapter I summarise the main findings and discuss how these could be improved.

5.2. 3D FWI of a numerical model of the breast

The feasibility and potential of 3D FWI for breast imaging is demonstrated on the numerical model of the breast shown in Figure 5.2. This model of speed-of-sound simulates realistic measured velocities for a human breast (Huthwaite, 2012). The bulk of the model consists of a glandular region represented by a random medium, in which several spherical inclusions that simulate cancer and fat masses with different average velocities have been added. The bulk of the breast is surrounded by an irregular subcutaneous fat layer and a thin outer skin layer. The model is discretised so that its dimensions are $95 \times 431 \times 431$ cells with a grid spacing of $0.29\,mm$ – the latter ensures there is not significant dispersion of P-waves below the maximum inverted frequency of $1.4\,MHz$ for all tests in this section.

The velocity model in Figure 5.2 is used to generate ultrasound data and perform several tests, which can be divided in two sets of tests depending on the acquisition

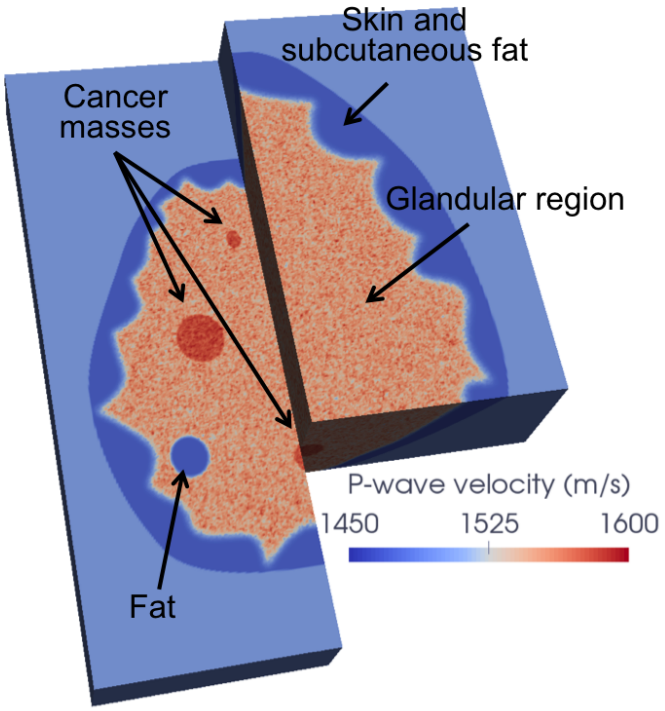


Figure 5.2.: 3D speed-of-sound numerical model of the breast used to generate ultrasound data.

geometry. For the first set, I use a 3D acquisition geometry in which sources and receivers are distributed across different acquisition rings that have their center at the centre of a different coronal view, whereas the data for the second set of tests is generated and acquired with a single ring array of transducers. For both sets of tests, I model $120\,\mu s$ of data using a three-cycle tone-burst source wavelet with a centre frequency of $750\,kHz$ and a time sampling of $0.11\,\mu s$, the latter to ensure numerical stability of the FD code. In particular, the first set of tests is designed to investigate the potential of 3D FWI of the breast, the importance of having low-frequency data and the consequences of ignoring density and absorption during the inversions. On the other hand, the second set of tests is designed to demonstrate that ignoring 3D effects can negatively affect the quality of the recovered model, and that 3D data acquisition and inversion are needed to obtain optimum results.

For the first set of tests, a total of 350 sources are distributed across 3 parallel

rings, and the data is recorded by a total of 1280 receivers distributed across 5 rings with 256 receivers in each ring. The centre of each ring is at the centre of a coronal view, and these are at different planes uniformly distributed in the vertical direction for sources and receivers (a vertical plane corresponds to a fixed z in Figure 5.3b). Ultrasound data are then generated for the first test assuming constant density and absorption, equal to those of water – i.e. a density of $1000\,kg/m^3$ and a very large Q value of $Q = 10^3$. 3D FWI is then performed using the modelled data as the observed data, inverting the data up to a frequency of $1.4\,MHz$ and starting the inversions at different frequencies and using different starting models in order to study the importance of having low-frequency data. To reduce the computation time and number of resources used for the inversions, only 35 sources of the 350 available are used at each iteration, which are different for each iteration so that all the sources are used within an iteration block.

Figure 5.3 displays pairs of coronal and sagittal views for the true, starting and recovered models for two particular slices. The true model is shown in Figures 5.3a and 5.3b, in which the dotted lines indicates the planes at which the complementary view has been extracted: these planes are the same for all models in Figure 5.3. Figures 5.3c and 5.3d correspond to the first starting model used for the inversions, which is just a model with water and no other information on the breast properties: this is the simplest starting model. The latter is used to perform 3D FWI of the observed data starting at a frequency of $700\,kHz$ and iterating 70 times, which results in the recovered model in Figures 5.3e and 5.3f. This result shows that starting the inversions at high frequency and from a starting model that is not very accurate leads to a poor quality recovered model, which is not representative of the true model. This behaviour has been previously documented in geophysics, and is known as cycle skipping (Virieux and Operto, 2009).

Cycle skipping takes place when the modelled data is delayed or advanced more than half a cycle with respect to the observed data, and can lead to completely erroneous models (Guasch and Warner, 2014; Warner and Guasch, 2014). Although several methods and objective functionals have been proposed to address cycle skipping in the FWI context, this can also be prevented by having low frequency data

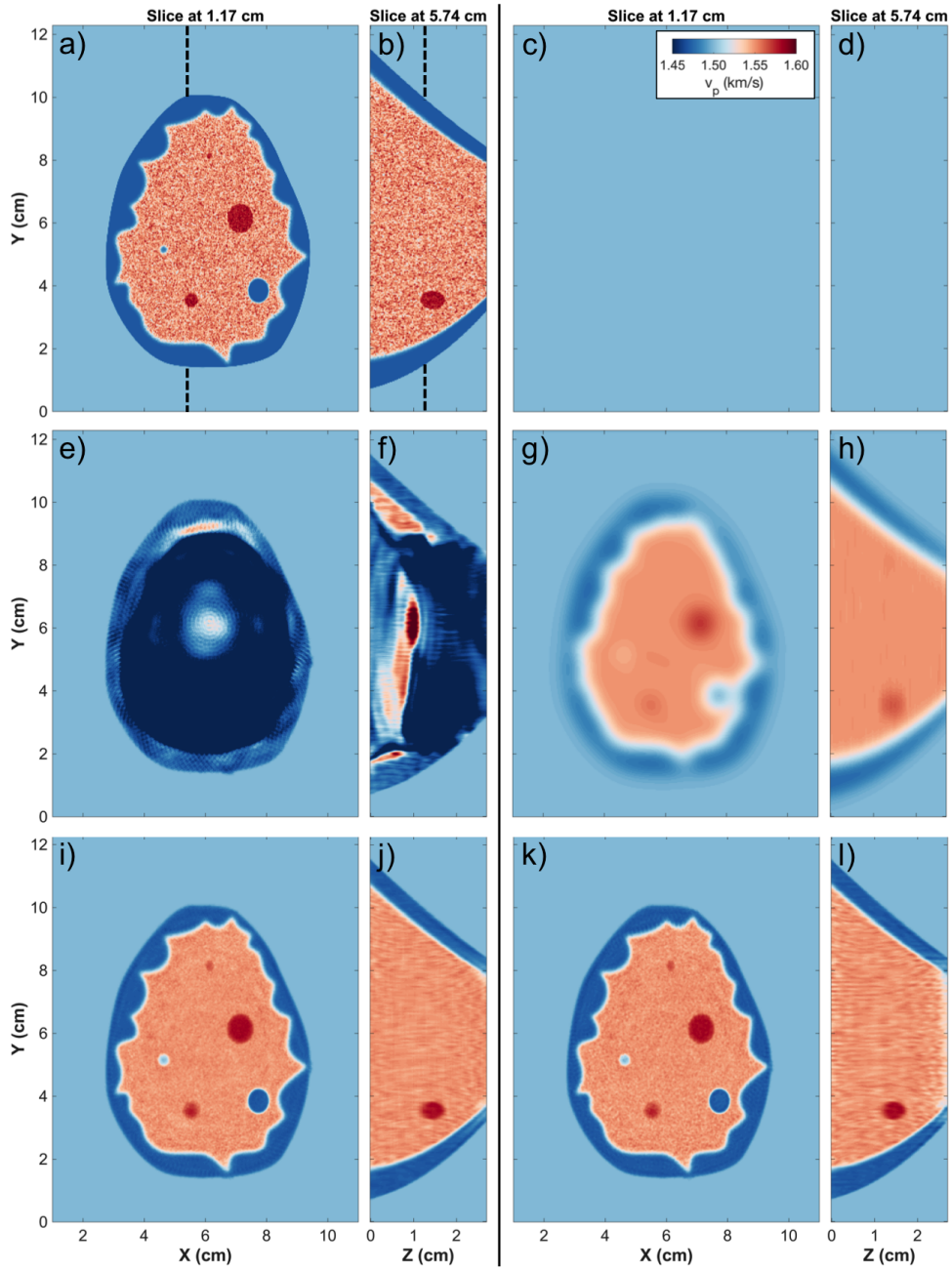


Figure 5.3.: Pairs of coronal and saggital views, respectively, for the numerical model of speed-of-sound of the breast. Panels a) and b) correspond to the true model, c) and d) to a starting model of just water, e) and f) are the result of 3D FWI using the water starting model and a starting frequency of $700\,kHz$, g) and h) are obtained from smoothing the true model with a Gaussian function in 3D, i) and j) are obtained after 3D FWI using the starting model in panels g) and h) and a starting frequency of $700\,kHz$ and k) and l) are obtained after 3D FWI using the water starting model and a minimum frequency of $300\,kHz$. Density and absorption are kept fixed during modelling and inversion, equal to those of water.

(Virieux and Operto, 2009). In that case, the data contains signal with larger periods and is therefore more difficult to be cycle-skipped. Alternatively, a more accurate starting model can be used to overcome cycle skipping, given that in this case the modelled data from the starting model is closer to the observed data, thus reducing any cycle skipping. To test this hypothesis in the medical imaging context, I repeat the previous inversion using the starting model in Figures 5.3g and 5.3h, which has been obtained by smoothing the true model with a 3D Gaussian smoothing filter. The high resolution and accuracy of the recovered models in Figures 5.3i and 5.3j demonstrate that 3D FWI performs well if the starting model is accurate, even though the starting frequency for the inversion is high. This is possible because the difference between the modelled data from the starting model and the observed data at the lowest inverted frequency is less than half a cycle. The inverted model successfully recovers the fat and cancer masses with the correct average properties, as well as the irregular shape of the skin and subcutaneous fat, thus enabling breast cancer detection.

Nevertheless, accurate starting models are rarely available and may be difficult to obtain and be user-dependant. Instead, having and using low frequency data should also reduce cycle skipping without the need of a very accurate starting model or user intervention. Thus, I test the effect of starting the inversion at a lower frequency of $300\,kHz$ and using the simple water model in Figures 5.3c and 5.3d, which results in the recovered model in Figures 5.3k and 5.3l. The quality of the latter is very similar to that obtained in Figures 5.3i and 5.3j, but does not require an accurate starting model. However, it requires starting the inversion at a lower frequency, thus increasing the total number of iterations, and the presence of low frequency signal in the data. Both recovered models show that it is possible to use FWI of ultrasound data to obtain high quality recovered models of the breast to be used for breast cancer diagnostics, with a maximum resolution of about $0.55\,mm$ ($\lambda = v_p/f \approx 1.09\,mm$) for the maximum inverted frequency of $1.4\,MHz$ and an average speed-of-sound in breast of $1520\,m/s$. In other words, two masses separated $0.55\,mm$ could be separated using these recovered models, and this distance could be decreased if acquiring and using data at higher frequencies.

In terms of compute cost and time, the inversions are performed using 36 nodes with one core per node, resulting in an average compute time of 24 h for inversions starting at 300 kHz, in which 110 iterations are performed, and an average time of 16.5 h for inversions starting at 700 kHz and a total of 70 iterations. Although methods to further reduce the compute time exist (e.g. source encoding as implemented by Wang et al. (2015)), the previous tests demonstrate the feasibility and potential of 3D FWI for breast cancer diagnostics.

However, realistic density and absorption for breast tissue have not been considered in the previous tests. Thus, I now generate ultrasound observed data using the model of speed-of-sound in Figure 5.2 and the density and Q models in Figures 5.4a and 5.4b – the latter have been designed using the structure of the velocity model and measured values for density (Huthwaite, 2012) and absorption (Iacono et al. (2015) and references and supplementary information therein).

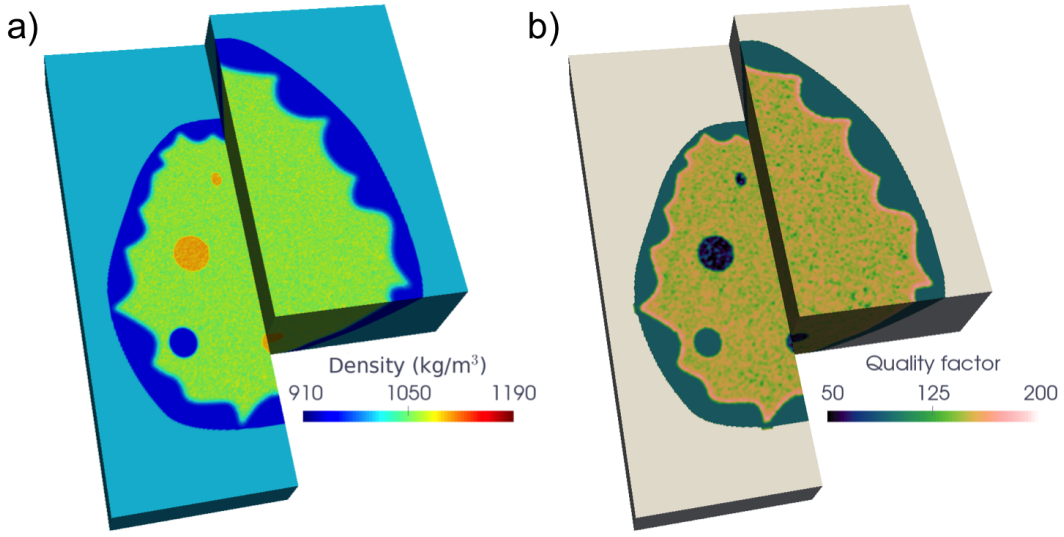


Figure 5.4.: 3D numerical models for the a) density and b) quality factor for the breast.

First, forward modelling is performed using a visco-acoustic code to generate observed ultrasound data, and the data for a representative shot gather recorded at a single ring of receivers are compared to that obtained with constant density and absorption equal to that of water in Figure 5.5. From these gathers, several events

can be identified that correspond to the direct wave, the reflection at the skin of the breast and the transmitted waves through the breast – recorded at transducers opposite to the source–, which are indicated in Figure 5.5a for the shot gather in which density and absorption are those of water. I want to highlight that these type of plots are similar to shot gathers typically obtained for seismic data, but the interpretation is different as the receiver array has a circular shape rather than a straight line, which causes the direct arrival to bend.

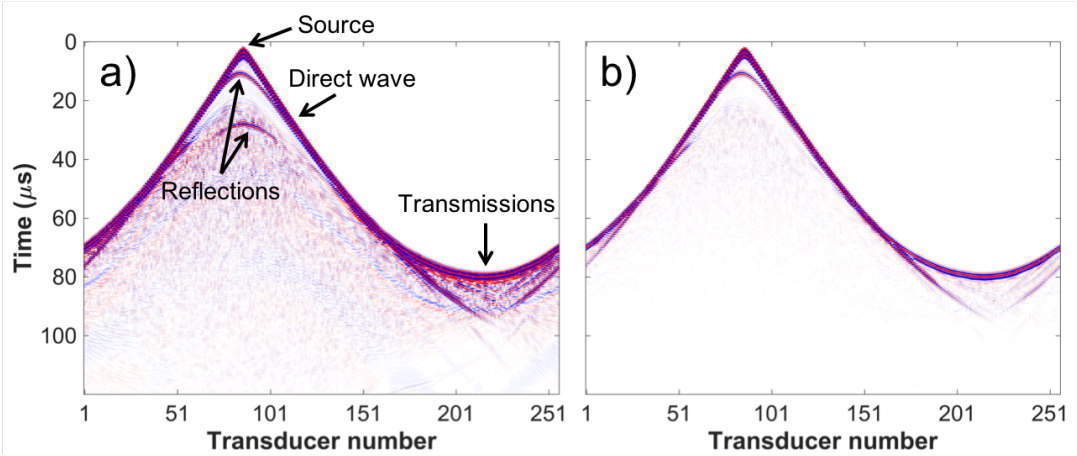


Figure 5.5.: Shot gather recorded at a single receiver ring array in an a) acoustic medium with fixed density and quality factor equal to those of water and b) in a visco-acoustic medium obtained with the variable density and Q models in Figure 5.4.

By comparing the shot gather in Figure 5.5a with that obtained with visco-acoustic modelling with variable density in Figure 5.5b, there is a visible difference in terms of amplitude, caused both by viscous effects and also by the the different changes in acoustic impedance due to different density values. This has a visible consequence on the amplitude of the reflected arrivals, especially the reflection far from the source as the associated wave has travelled within the breast, and also on the transmitted part of the data. The delay introduced by viscous effects is not visible at this scale, but can be observed by comparing two traces for the same receiver.

In order to study the impact of ignoring these effects in the inversions, I first perform acoustic 3D FWI of the acoustic data obtained with constant density and Q values using a multi-scale approach by inverting the data from 300 kHz up to 1 MHz,

the latter being lower than for the tests in Figure 5.3 to reduce the computation cost, yielding the results in Figures 5.6a and 5.6b. I then perform acoustic 3D FWI of the visco-acoustic data with the same inversion parameters and keeping density and absorption fixed and equal to those of water during the inversion, i.e. neglecting density and intrinsic attenuation, resulting in the models in Figures 5.6c and 5.6d.

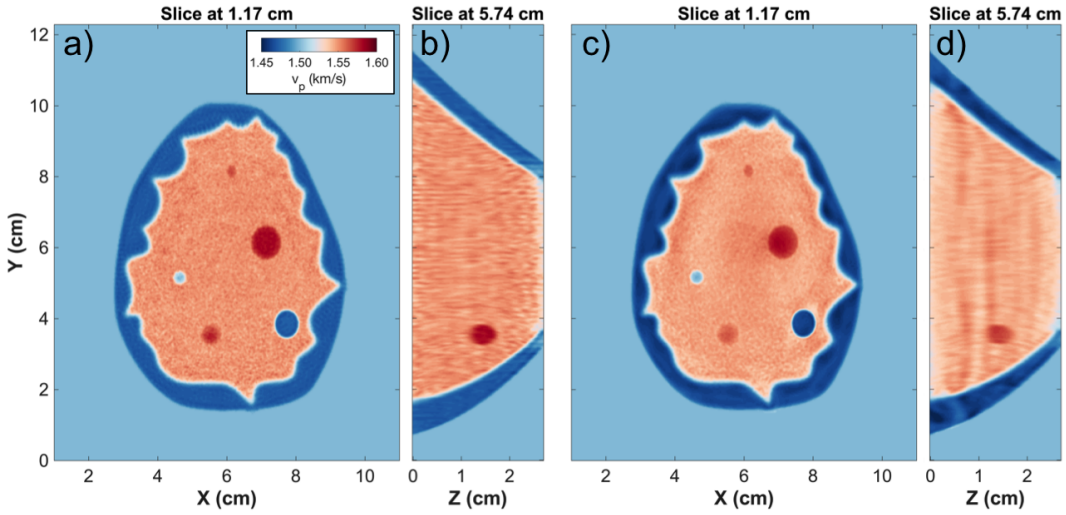


Figure 5.6.: Pairs of coronal and sagittal views for the recovered models of the breast after 3D acoustic FWI of a) the acoustic data with constant density and absortion equal to those of water and b) of the visco-acoustic data ignoring density and absorption during the inversion. The inversions are performed using a water starting model and inverting the data from $300\,kHz$ up to $1\,MHz$.

Ignoring density and absorption results in a model that is overall slower – mainly due to the delay introduced by viscous effects– and with less level of detail – both due to ignoring amplitude attenuation caused by absorption and not accounting for the correct dynamics of the wavefield caused by variations in density. Nevertheless, the outer irregular fat layer and the spherical inclusions are well recovered, and a cancer mass can be differentiated from a mass of fat. Hence, I neglect density and absorption effects for the remaining of this chapter, but this might be relevant for dense breasts. Additionally, the use and development of multi-parameter strategies would be beneficial in this field, as recovering realistic models of tissue density and sound absorption would help better characterise tissue properties – for instance, to differentiate benign and malign tumours.

The last of the first set of tests is the robustness of 3D FWI against random noise in the ultrasound observed data. To test this, I repeat the previous acoustic inversion by adding a high level of random noise to the visco-acoustic data – a different random value is added to each sample, which is proportional to 15% of the average amplitude value between the minimum and the maximum amplitude values for all traces in a shot. Figures 5.7a and 5.7b respectively show the noise-free and noisy shot gathers.

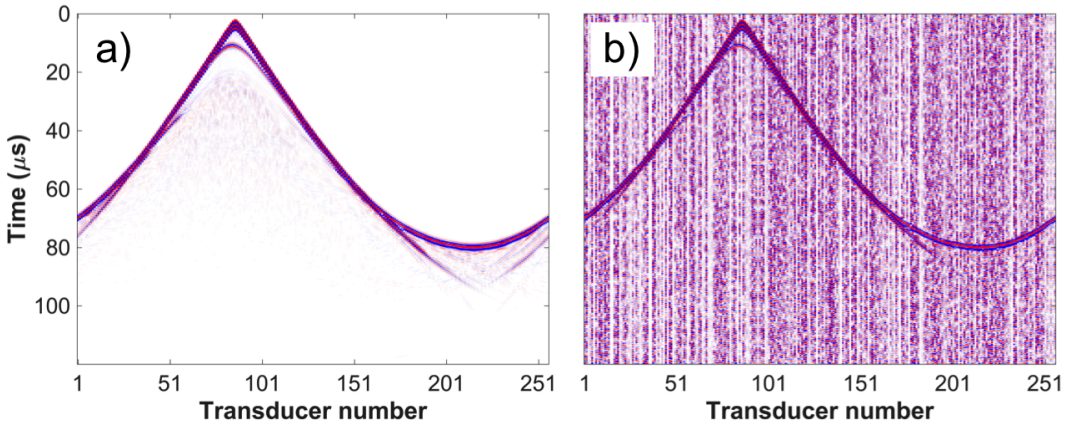


Figure 5.7.: Shot gathers recorded at a single ring array of receivers for the a) visco-acoustic data obtained with the density and absorption models in Figure 5.4 and b) the previous visco-acoustic data with added random noise.

Repeating the inversion for the noisy visco-acoustic data with the same inversion parameters as before, leads to the recovered model in Figures 5.8c and 5.8d.

When compared to the result of inverting the noise-free visco-acoustic data in Figures 5.8a and 5.8b, the inversion of noisy data is able to recover all the features present in the true model, but the result is noisier both in the glandular region but also within the subcutaneous fat layer. This causes the sagittal view in Figure 5.8d to be have more vertical stripes, and thus stronger regularisation might be needed. Although the effect of noise is not relevant for the detection of tumours as these can be clearly observed, the detection of their outer boundary is somewhat more difficult, which could lead to some errors on the estimated volume of the tumour. Nevertheless, the level of noise is very high and FWI still provides acceptable results. Thus, I conclude that FWI is very robust against random noise in the observed ultrasound data.

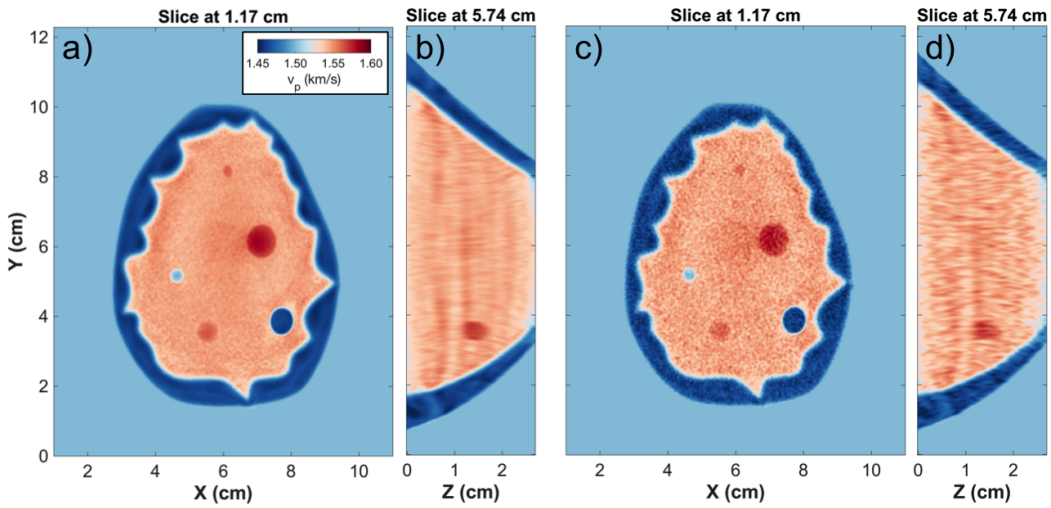


Figure 5.8.: Pairs of coronal and sagittal views for the recovered models of the breast after 3D acoustic FWI of the visco-acoustic data a) without and b) with random noise added in the data. The inversions are performed ignoring density and absorption effects, using a water starting model and inverting the data from $300\,kHz$ up to $1\,MHz$.

Now, the second set of tests are performed. These are designed to analyse the impact of performing 2D FWI of data acquired for the 3D *in silico* breast model. Thus, in contrast to the first set of tests, data are now acquired using a single ring array that contains 50 emitters and 256 receivers uniformly distributed along the circular array, with its centre in the centre of a coronal view and at a depth of $z \approx 1.66\,cm$. Ultrasound data are modelled both considering a 2D slice of the 3D model centered at the position of the ring array and also for the full 3D model of the breast. A three-cycle tone-burst source wavelet is used, which has a central frequency of $500\,kHz$ to avoid dispersion at frequencies above $1\,MHz$ – this is to make the results of 2D and 3D modelling using different kernels comparable. Figures 5.9a and 5.9b display the ultrasound data acquired for a single shot gather when the modelling is performed in 2D across a slice of the model and in 3D across the full 3D model, respectively.

The 2D and 3D modelled data are similar in the earliest arrivals, but differ for later arrivals. For instance, there is a reflection in the 3D data that is not present on the 2D data (indicated by a black arrow), which comes earlier in time and has a different

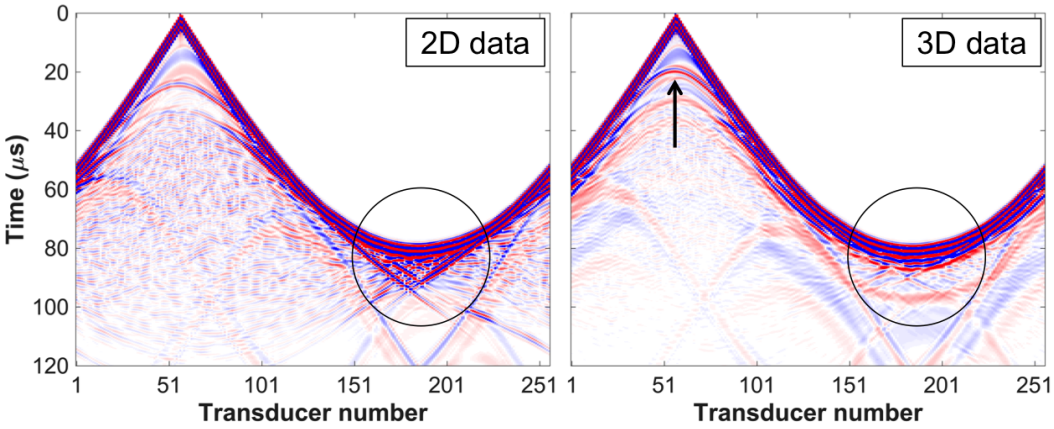


Figure 5.9.: Representative ultrasound shot gathers obtained from a source in a single ring array of transducers surrounding a a) 2D slice of the *in silico* breast model and b) the full 3D model.

curvature than the first reflection in the 2D modelled data. Moreover, there are also differences in the transmitted part of the signal (black circle). The latter two are due to off-plane effects, i.e. waves that travel across different planes and reach the receivers when modelling data for a 3D model, which cannot exist for a 2D model.

To investigate the impact of these differences on the recovered velocity models, I perform 2D and 3D FWI of the modelled data for the 3D *in silico* breast model, and a comparison of the results with the true model for the same slice in which the ring of transducers is located are shown in Figure 5.10. The inversions are performed by inverting the data from $150\,kHz$ up to $1\,MHz$ in a total of 100 iterations, using a homogeneous starting model of just water and using 25 sources per iteration to reduce the computational cost.

From these recovered models I make two important observations: first, 2D FWI of 3D modelled data results in a model that is less well resolved, less accurate and cannot recover the irregular shape of the subcutaneous fat layer and the skin, as well as one of the tumours (indicated by a black arrow in Figure 5.10b and, secondly, 3D FWI of data recorded with a single acquisition ring results in a model that contains off-plane effects. The second conclusion is made by observing that two spherical inclusions (numbered as 1 and 2 in Figure 5.10c) appear in the recovered model that are not present in the true model for the same slice, and that the mass of fat denoted as 3 in Figure 5.10c is larger than the same inclusion in the true model. There is

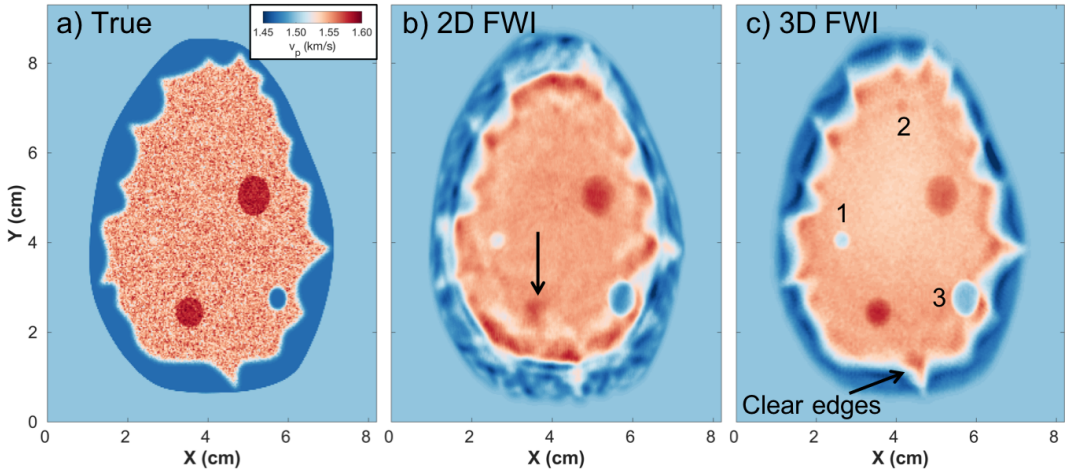


Figure 5.10.: Coronal views of a) the true model and the recovered models after b) 2D FWI and b) 3D FWI of the 3D data modelled for a single ring array of transducers around the full 3D model.

also a visible enlargement and loss of resolution at the edges of the skin and subcutaneous fat when compared to the true model. All these are due to waves that travel at different planes, and the fact that 3D FWI of the data recorded at a single ring array is unable to separate the latter due to the lack of spatial resolution in the vertical direction. To minimise these effects, two immediate solutions are possible: 1) record the data at multiple rings and invert the data in 3D, as in the first set of tests, and 2) perform 3D FWI of data generated with focused ultrasound transducers.

In summary, the analysis performed in this section show that 3D FWI of ultrasound data for breast imaging and breast cancer detection is nowadays possible. To obtain high-quality recovered models, low-frequency data is required, density and absorption can initially be ignored for normal density breasts and 3D effects as well as 3D data acquisition are necessary to separate off-plane effects. This is now corroborated and put in practice on a 2D *in vivo* breast dataset in the next section.

5.3. Applicability to a 2D *in vivo* dataset

The *in vivo* dataset was acquired at the Karmanos Cancer Institute (Wayne State University, United States) using the computer ultrasound risk evaluation (CURE)

system (Littrup et al., 2002). The latter consists of 256 transducers in a ring array of 20 cm diameter that is submerged in water and designed to generate and acquire both reflected and transmitted ultrasound for breast imaging. No other information on the *in vivo* dataset or the experiment was provided, such as the calibrated location of the transducers (i.e. their exact location), a water shot, the variations in temperature of the water bath during the experiment (temperature variations change the speed-of-sound) or the source wavelet. Thus, these may affect the quality of the recovered model if not properly considered.

A representative shot gather of the raw *in vivo* dataset is shown in Figure 5.11a. Five different features and events can be distinguished, as indicated in the figure: 1) the direct wave, 2) reflections, 3) transmissions, 4) a wave that propagates along the ring array (a straight line in a shot gather given the circular array of transducers means the associated wave takes the same time to reach two consecutive transducers and, hence, must propagate along the ring) and 5) coherent noise at early times.

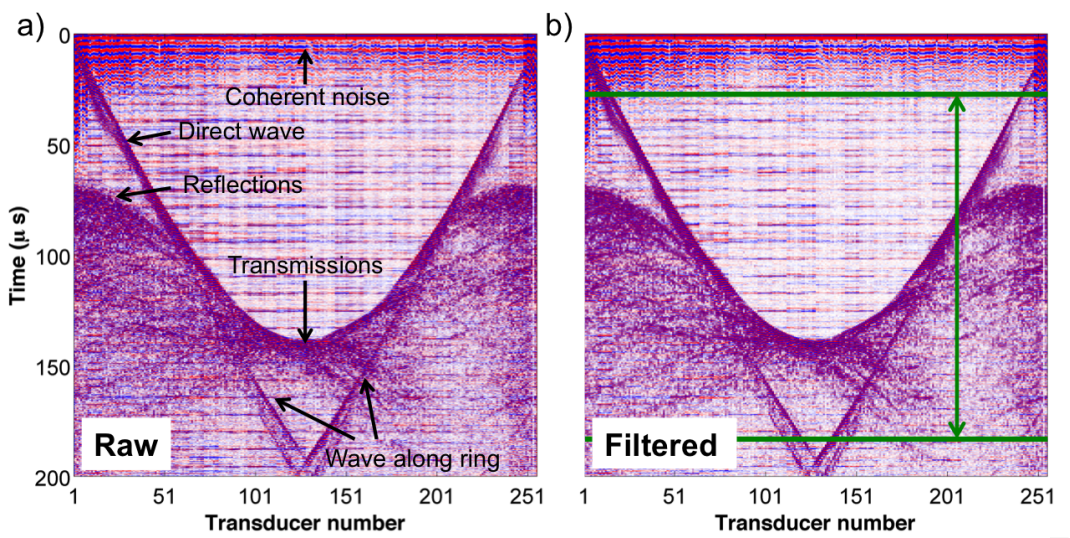


Figure 5.11.: Representative shot gather for the *in vivo* data a) before and b) after band-pass filtering using a zero-phase Butterworth band-pass filter that rejects data below 200 kHz and beyond 1.75 MHz.

The first three events are expected given the acquisition geometry, but the latter two are not and may adversely affect the result. The impact of the wave propagating along the array on the recovered model may introduce a velocity change close to the sources without changing the target significantly, but the coherent noise at early times

may change the model considerably as this corresponds to a signal that propagates infinitely fast compared to the other events, as it reaches all transducers simultaneously – hence, this may be electrical noise. In addition, the dataset is contaminated with random noise, as can be seen by the random signal prior to the first arrivals. To mitigate the latter and also reduce the strength of coherent noise, I filter the raw data with a zero-phase Butterworth band-pass filter to remove signal below $200\,kHz$ and beyond $1.75\,MHz$, which results in the shot gather in Figure 5.11b. The filtered shot gather is overall less noisy before the first arrivals, but is also cleaner for reflections. However, coherent noise is still present – even after careful FK-filtering. Thus, only part of the data is used for the inversions and the first $30\,\mu s$ are not included, which leads to optimal results. The data is also cropped at $180\,\mu s$ to reduce the computation time and also because most of the useful information is contained in the selected part of the data (the selected data is shown by a green arrow and lines in Figure 5.11b).

The filtered and cropped dataset is then used as the observed dataset for the inversions. Before, some more minimal pre-processing is performed: this includes sinc interpolation of the data from $0.16\,\mu s$ to $0.04\,\mu s$ to avoid numerical dispersion, removal of a shot gather with little energy so that the total number of sources (and hence, shot gathers) is 255, calibration of the transducers and extraction of a source wavelet from the data.

Calibration of the transducers is typically done by acquiring a water shot, i.e. acquiring data without a target, and adjusting the location of the transducers depending on the measured arrival times of the direct wave. However, a water shot was not available and hence, calibration was done by Peter Huthwaite (Mechanical Engineering Department, Imperial College London) – and reproduced by myself– using the direct arrivals in the *in vivo* dataset for those traces in which the latter are not interfering with reflected or transmitted events. To extract the exact location of the transducers, automatic TOF measurements of the direct arrivals were computed – for instance, using the Akaike information criterion (AIC) (e.g. Zhang et al. (2003); Li et al. (2009))– and these were used as an input to calibrate the location

of transducers using the method proposed by Roy et al. (2011) to calibrate a breast ultrasound system, which is based on the multi-dimensional scaling (MDS) algorithm.

Once the system is calibrated, a source wavelet needs to be extracted to perform forward modelling, which is necessary for FWI. Several methods exist in geophysics to extract a source wavelet from seismic data (e.g. see Walden and White (1998), Yi et al. (2013) or Warner et al. (2013)). Here I use a method similar to that of Warner et al. (2013) to extract a source wavelet. The latter consists first in selecting those traces for each shot in which the direct arrival through the water does not interact with reflections and transmissions, and muting the data below these arrivals. The data is then arbitrarily shifted in time (this is removed at the end of the process) and a linear move-out (LMO) correction is applied to each trace to remove the dip of the direct arrivals and, thus, every sample is shifted in time given the calibrated distances between transducers and the speed-of-sound in water. Figure 5.12a shows the filtered and muted first arrivals for the selected traces of a shot gather after the LMO correction, in which the nominal (i.e. without calibration) locations of the transducers are used, whereas Figure 5.12b displays the same data after using the calibrated positions.

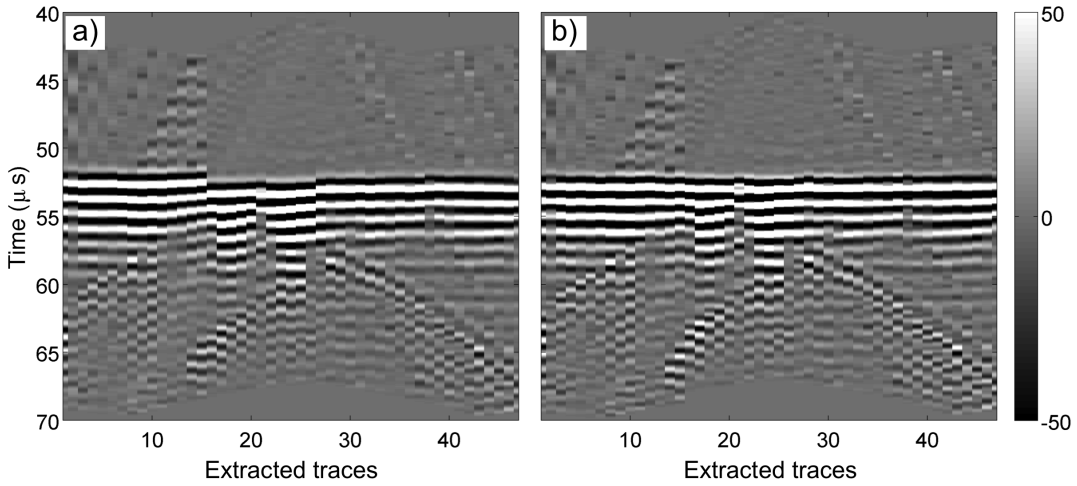


Figure 5.12.: Selected traces used for extracting a source wavelet from the *in vivo* data for a single shot gather after LMO correction using a) the nominal and b) the calibrated location of the transducers.

Note the flatness of the direct arrival when using the calibrated location of trans-

ducers (Figure 5.12b) with respect to the discontinuity visible around traces 15 and 16 when using the nominal positions (Figure 5.12a). This already indicates that without calibration FWI could lead to incorrect recovered models caused by cycle-skipping due to mis-positioning of the transducers (note there is more than half a cycle jump around traces 15 and 16 in Figure 5.12a). The next step in order to extract a source wavelet from the data is to repeat this process for the other shot gathers (this may not be necessary if the noise level is low) and stack all the traces, which results in the stacked trace plotted in Figure 5.13a. This is not yet the source wavelet, as the inaccuracies and effects of modelling data in 2D rather than in 3D have not taken into account. Thus, I model ultrasound data with a guess source wavelet – I choose a Ricker wavelet initially, which is band-pass filtered with the same filter as that applied to the data– for a model with just water, and I repeat the process of selecting a few traces and stacking the LMO corrected traces. After stacking all the traces for all shots, Figure 5.13b shows the resulting stacked trace.

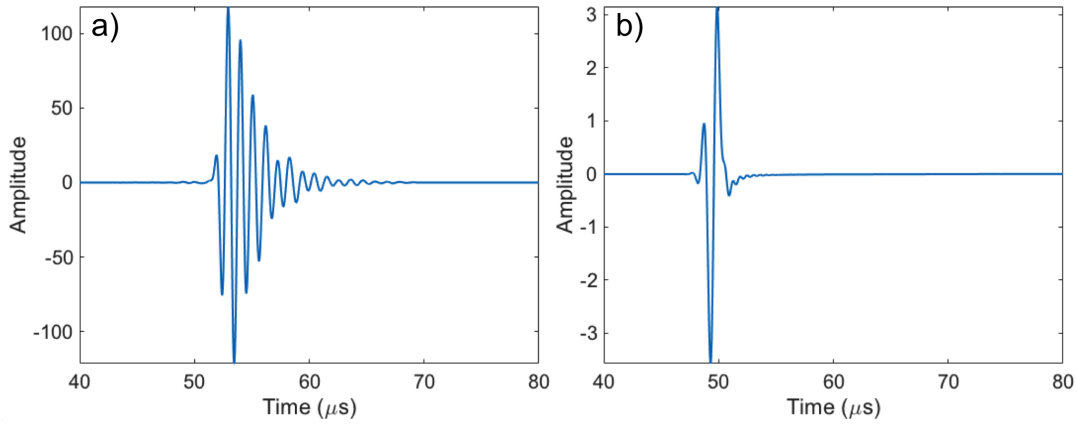


Figure 5.13.: Stacked traces obtained after LMO correction of the direct arrivals of a) the band-pass filtered and muted *in vivo* data and b) the modelled data for a water shot using a band-pass filtered Ricker wavelet as a first guess.

The two stacked traces in Figure 5.13 are not similar, but are now used to extract the source wavelet that will be used as an input for 2D FWI. To do so, a Wiener filter that matches the stacked trace in Figure 5.13b to the stacked trace in Figure 5.13a is computed, and this is applied to the source guess used to model the data (i.e. the band-pass filtered Ricker wavelet). This results in a source wavelet that, when used in

the 2D modelling code, generates similar direct arrivals to those in the *in vivo* data. Finally, the arbitrary time shift introduced earlier is removed, and Figures 5.14a and 5.14b show the final extracted wavelet as well as its amplitude spectrum.

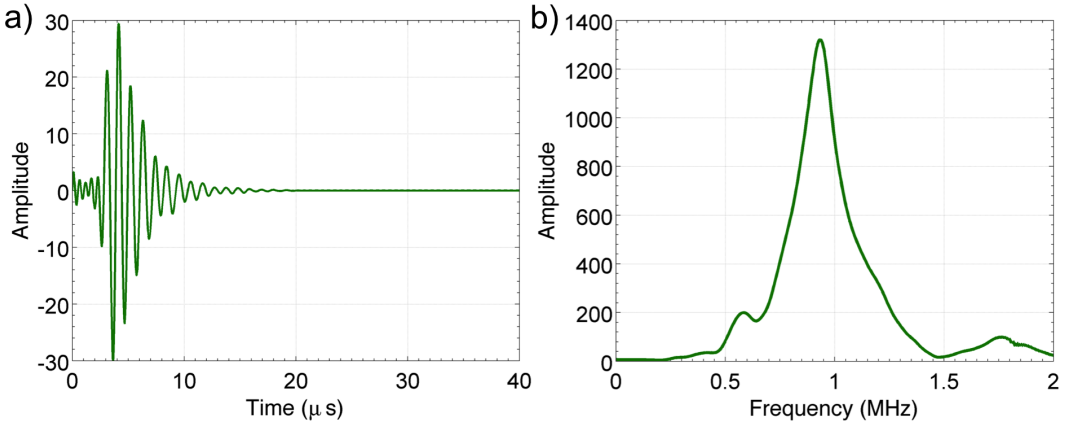


Figure 5.14.: Plot of the a) extracted source wavelet in time and b) its amplitude spectrum used for 2D FWI.

The extracted wavelet has a central frequency of around $950\,kHz$ and little useful energy below $350\,kHz$ (the amplitude is around 130 times smaller at this frequency compared to the peak frequency). Thus, the data does not contain very low-frequencies and cycle-skipping might be expected when inverting the data from an inaccurate starting model.

The source wavelet in Figure 5.14 and the position of transducers obtained after calibration are used now as inputs for 2D FWI. The *in vivo* dataset is inverted using a multi-scale approach up to a frequency of $1.75\,MHz$ using a grid spacing for the models of $0.15\,mm$, which ensures there is little dispersion of P-waves. In order to fit all transducers into the model, its size is 1464×1464 in cells. The inversions are performed using 85 sources per iteration to speed up the inversions and a total of 22 nodes and 8 cores per node are used, resulting in an average compute time of around $16\,h$. All inversions are in 2D to reduce the compute time and given that only a single acquisition ring was available, although 3D would be possible – the compute time could be reduced by increasing the grid spacing, thus performing the inversion up to a lower maximum frequency without compromising excessively on the quality

of the result. Additionally, only a 2D TOF model was provided, which is used as a starting model for one of the inversions due to the lack of low-frequency data.

The first inversion test is performed to analyse the effect of a good calibration and also the impact of including the coherent noise at early times. Thus, a model of just water is used as a starting model. In order to reduce the computation time for this initial test, I use a grid spacing of $0.20\,mm$ (the dimensions of the model for this test are 1251×1251 in cells) and perform the inversion from $400\,kHz$ up to $1.4\,MHz$. Moreover, the coherent noise at early times is also included for this test, so that the observed data starts at $0\,\mu s$. Figures 5.15a and 5.15b show the recovered models of speed-of-sound obtained using the nominal position for the transducers and the calibrated position, respectively.

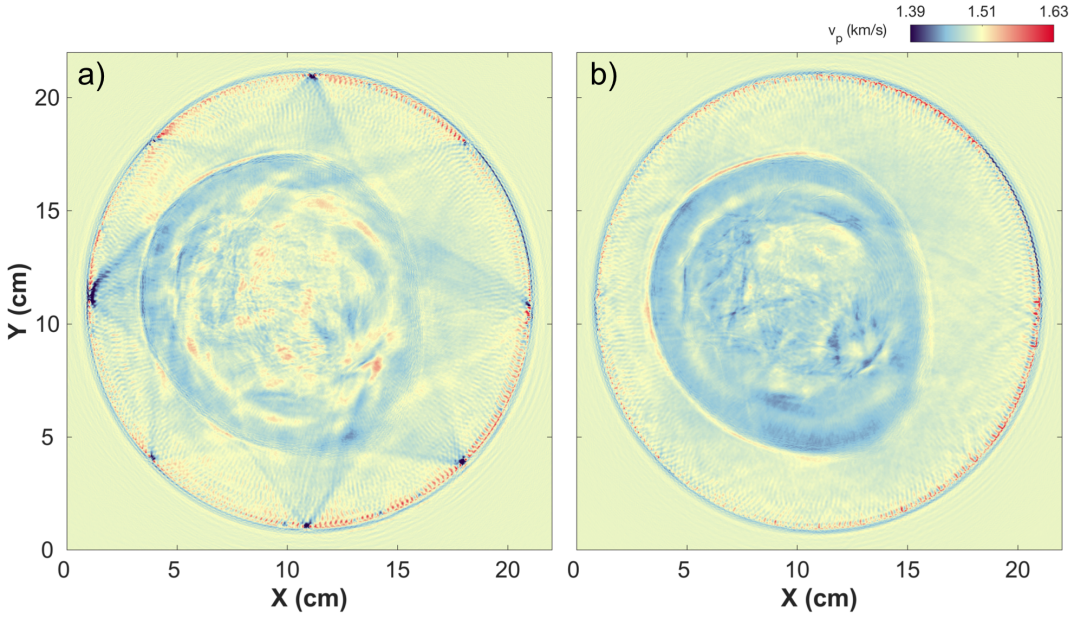


Figure 5.15.: Recovered models of speed-of-sound obtained with a starting model of just water and a) the nominal and b) the calibrated position of the transducers. The inversions are performed from $400\,kHz$ to $1.4\,MHz$ in a total of 100 iterations and using 85 sources per iteration. Data from $0\,\mu s$ to $180\,\mu s$ are used, hence including coherent noise.

The first observation from these results is that using the nominal location of transducers introduces artefacts in the model in the form of constant low velocity triangles with one of their vertices touching a transducer. These may be created by different

parts of the ring being mis-positioned: in fact, the ring is made of 8 segments, which fits with the 8 triangles observed in Figure 5.15a. To compensate for these low velocity areas, the recovered model in Figure 5.15a contains high velocity zones that could be incorrectly associated with tumours. On the other hand, the recovered model in Figure 5.15b does not contain these artefacts as the location of the transducers is calibrated, resulting in a more homogeneous recovered model of the breast. However, 2D FWI of the *in vivo* data is not capable of retrieving a highly detailed and accurate model of the breast, as the starting model is not accurate enough and the acquired frequencies are not low enough, thus causing cycle-skipping effects. A third observation is that there is a velocity perturbation that follows the geometry of the ring, which is caused by the presence of a wave propagating along the ring array, as indicated in Figure 5.11a. Moreover, in both Figures 5.15a and 5.15b there are very high velocity and small perturbations near the transducers towards the centre of the model, which are caused by the fast and coherent noise at early times.

Thus, I discard the first $30\,\mu s$ of data, and perform new inversions up to a frequency of $1.75\,MHz$ using a finer grid spacing of $0.15\,mm$. The results for different tests are shown in Figure 5.16 – note I apply a mask to the results to better observe features in the breast.

First, I perform 2D FWI of the *in vivo* data with a starting model of just water and a starting frequency of $500\,kHz$, which results in the velocity model in Figure 5.16a. Similar to the tests in Figure 5.15, the lack of low-frequency data results in a poor quality recovered velocity model, in which only the outline of the breast can be identified.

To overcome cycle skipping I use a more accurate starting model, provided lower frequencies cannot be recorded. The latter was obtained by Peter Hutwaite using bent-ray tomography (Huthwaite, 2012), and is shown in Figure 5.16b. This is a smooth model in which several features can be identified, such as the edges of the breast, a low velocity region at the bottom part of the breast and a high velocity ellipsoidal feature that could be associated with the presence of a breast tumour. However, the level of detail in the model in Figure 5.16b is low and no other features

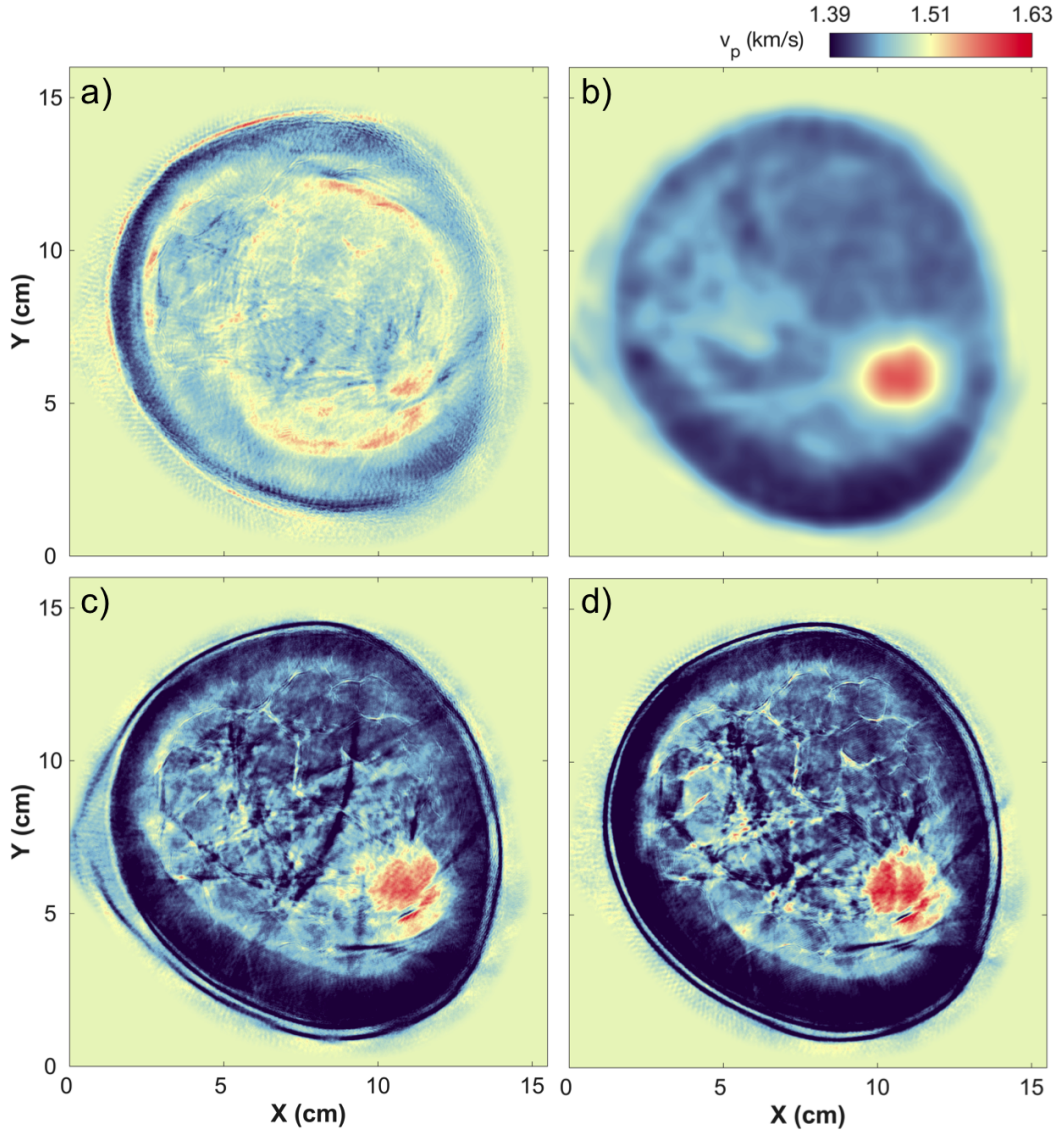


Figure 5.16.: Recovered models of speed-of-sound for the *in vivo* data obtained after a) 2D FWI with a starting model of just water and a minimum frequency of $500\,kHz$, b) bent-ray tomography and 2D FWI using the bent-ray tomography model as a starting model and c) a minimum frequency of $800\,kHz$ and d) a minimum frequency of $500\,kHz$. Data before $30\,mus$ and after $180\,mus$ is excluded and 85 sources per iteration are used for the inversions, which are performed up to $1.75\,MHz$.

can be distinguished.

Hence, I use the latter recovered model as a starting model for the inversions, as this should result in a better conditioned inverse problem and reduce cycle-skipping effects. Figure 5.16c displays the recovered model obtained with the latter as a starting model and starting the inversion at $800\,kHz$. Many differences can be observed

in comparison to the previous model, especially in terms of details: the boundaries of the high-velocity mass associated with a tumour are better defined, the subcutaneous fat layer and skin layers can be distinguished at the edges of the breast and additional high resolution features have appeared inside the breast, which may be vessels. Nevertheless, starting the inversion at a high frequency introduces some artefacts in the recovered model, such as the stretching of breast skin on the left-hand-side of the model or a semi-circular low-velocity stripe that crosses the breast through the middle.

For this reason, I finally repeat the inversion but with a lower starting frequency of $500\,kHz$, as there is still energy in the data with this frequency. Figure 5.16d shows the final high-quality recovered model, which is much more detailed than that in Figure 5.16b and does not show the presence of any obvious artefacts due to cycle-skipping. An interesting observation from this model is that not many details have been recovered within the low-velocity zone near the breast skin and below the tumour that was already visible in the bent-ray tomography model. This might be due to the starting model being too slow in this region, and hence leading to cycle-skipping, or perhaps due to a denser and more absorbent tissue area.

To further observe the benefits of FWI when compared to ray-based methods such as TOF or bent-ray tomography, I show a section at the top right corner of the models in Figures 5.16b and 5.16d in Figures 5.17a and 5.17b, respectively.

The comparison in Figure 5.17 clearly shows the benefits of using FWI to obtain a higher quality and more detailed model of the breast, in which structures that resemble blood vessels, milk ducts or lymphatic vessels can be observed. This may be relevant as higher resolution models of the breast could be used to detect small tumours, quantify the total volume of a tumour as well as better define the boundaries of larger tumours, and possibly detect small clusters of micro-calcifications, especially in dense breasts, which are indicators of breast cancer (Littrup et al., 2018).

Nevertheless, the observed features in the recovered models may belong to other planes, given the 2D acquisition geometry and the use of 2D FWI. Thus, I strongly

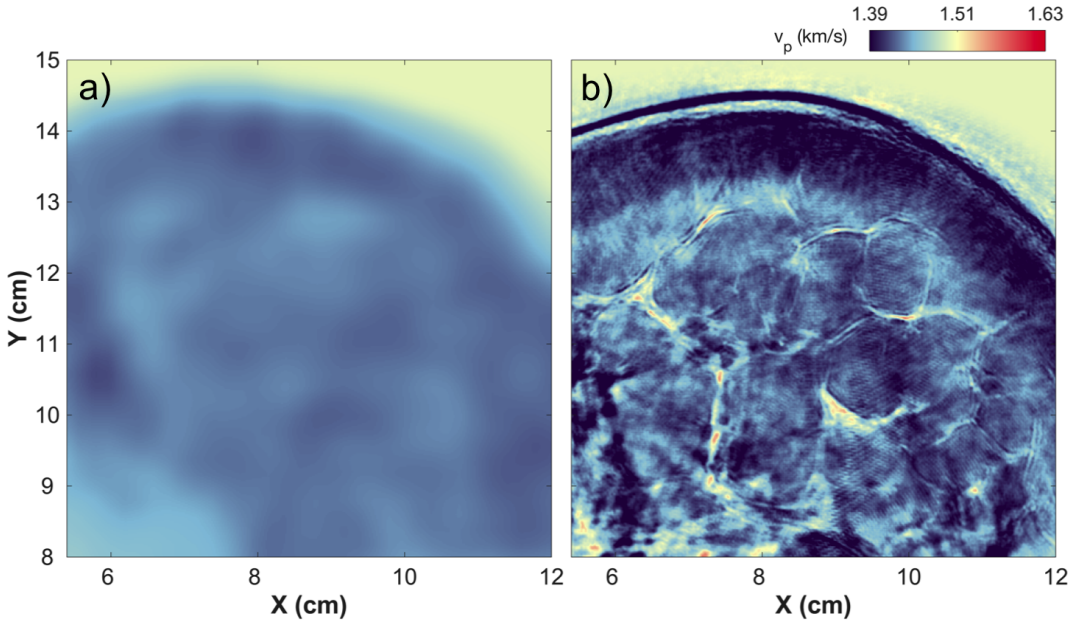


Figure 5.17.: Zoomed in sections of the recovered models of speed-of-sound after a) bent-ray tomography and b) 2D FWI with a minimum frequency of $500\,kHz$ shown in Figures 5.16b and 5.16d.

recommend the use of FWI to reconstruct models of physical properties of the breast, but also the use of a more sophisticated 3D device capable of acquiring and emitting ultrasound below $500\,kHz$.

In summary, the experiments performed in this section demonstrate the feasibility and potential of FWI for imaging breast cancer in *in vivo* experiments.

5.4. Summary

Breast imaging techniques are used to examine the breast, and are essential for breast cancer detection and diagnosis. In this chapter I have demonstrated the benefits of using FWI for breast imaging using ultrasound over more conventional methods that are based on ray theory. The use of FWI and ultrasound to reconstruct high resolution models of the breast also represents an improvement over standard mammography scans, given that the former use non-ionising radiation, they are pain-less, they can be made portable and provide quantitative images of the physical properties of the breast at an affordable cost.

The tests performed both on *in silico* and *in vivo* data demonstrate the feasibility of FWI for breast imaging, also in 3D. This is mainly due to advances on compute power over the last decade, the use of parallel computing and the robustness of existing FWI codes. Although not explored here in detail, several strategies can be used to further reduce the compute cost. These include reducing even more the number of sources used per iteration, acquiring lower frequencies to better condition the inversion problem – and hence using a coarser grid spacing – using objective functionals that minimise cycle-skipping effects or using source encoding methods, amongst others.

In order to assess the importance of several factors on the quality of recovered models of the breast using FWI, I have performed tests on *in silico* 3D data to analyse the effects of having low-frequency data, ignoring density and absorption during the inversion, having strong random noise in the observed ultrasound data and ignoring 3D effects.

The first conclusion obtained from these tests is that the use of low-frequency transducers – a few hundreds of kHz– leads to more accurate high-fidelity recovered models as it reduces the non-linearity of the inversion problem. Moreover, having low-frequency data enables FWI of the breast without the need of an accurate starting model obtained via the intervention of the user and ray-based methods, thus making the image reconstruction process more robust, reliable and automatic.

Secondly, ignoring density and absorption does not have strongly detrimental consequences on the quality of the results for normal density breasts, even in the presence of strong random noise, but might need to be considered for low or high density breasts. This could be addressed using multi-parameter approaches to obtain density and absorption models and invert for them during the inversion, or could also be addressed by using empirical relationships between speed-of-sound and density and absorption.

The last conclusion drawn from the *in silico* experiments is that 3D acquisition as well as 3D FWI are necessary to obtain highly resolved models and avoid off-plane

effects. In order to minimise these, more focused transducers could be used. However, this would require the recording of data in more planes and the inversion of more data, thus increasing the total cost.

These observations also apply to the results obtained for the *in vivo* data. Despite the lack of prior information on the acquisition calibration and the experimental conditions, I have demonstrated that FWI successfully recovers a model of the speed-of-sound of the breast with a high level of detail for *in vivo* data, and that requires a starting model – obtained from bent-ray tomography– in order to avoid cycle-skipping. However, the latter would not be required if lower-frequency data were recorded. Close observation of the recovered model reveals internal structures that may be associated with milk ducts or blood or lymphatic vessels. Nevertheless, these may not belong to the same coronal plane due to the lack of a 3D data acquisition system as well as the use of 2D rather than 3D FWI. The latter could be easily implemented, but requires the estimation of a 3D starting model (for instance, by bent-ray tomography) due to the lack of low-frequency data.

Another aspect that has not been considered in the tests on *in vivo* data is the finite size of transducers – of a few millimeters– compared to the minimum wavelength of the emitted signal. Instead, I have used that sources and emitters are point sources and emitters, as typically done in seismic experiments. This should have an effect on the resolution of the recovered model. In order to properly address this, more numerical modelling experiments should be carried out by using, for instance, conglomerates of point elements or more realistic finite size sources and receivers so that the modelled data fits the experimental data for a water shot more accurately. This has not been further addressed due to the lack of more experimental data.

Overall, 3D FWI of ultrasound data could become a key method in medical imaging over the coming years, as it provides reliable and quantitative models of the body with a higher degree of detail than standard techniques while using non-ionising ultrasound. Future experiments should focus on implementing 3D FWI on *in vivo* data

to image a whole breast efficiently – this could perhaps include running FWI on the cloud– in order to better characterise the volume and shape of breast tumours and separate off-plane effects. Additionally, multi-parameter strategies to recover absorption and density during the inversion would allow better characterisation of malignant breast tumours, thus leading to more accurate diagnostics. The now popular machine learning algorithms could also be used to segment the resulting images and perhaps reduce modelling computation costs.

Finally, I would like to emphasize the fact that both the geophysical and the medical imaging communities could enormously benefit by sharing knowledge and adapting technologies such as FWI, thus fostering the collaboration between two scientific disciplines that share many similarities.

6. Towards 3D imaging of the brain with ultrasound

Imaging the brain in three dimensions (3D) using ultrasound is nowadays not possible mainly because of the large impedance contrast at the skull and also due to the lack of techniques that model wave propagation correctly. Here I investigate how FWI can be effectively used to tackle these problems.

Due to my research on this topic I have received a silver award from the European Institute of Innovation and Technology (EIT), under the EIT Health PhD Transition Fellowship, worth £15000. I have spent the latter on material for the laboratory that will be used to perform *ex vivo* experiments. My conclusions and results obtained from this work have also been successfully utilised by my supervisor, Michael Warner, and our collaborator, Mengxing Tang (Bioengineering Department, Imperial College London), to obtain the prestigious Excellence Fund for Frontier Research award, worth £250000. The latter will also be used to proof these ideas in the laboratory. In addition, the results presented in this chapter have also been recently submitted to a prestigious scientific journal. Finally, I have filed a patent related with this topic, which is currently pending.

I would like to acknowledge Lluis Guasch, who has converted the 3D segmented model of the head into a model with realistic acoustic speeds, and has also designed the acquisition geometry used for the 3D *in silico* experiments.

6.1. Introduction

Neuro-imaging, or imaging the brain, consists in obtaining images of the properties of the brain without being able to access it. These images provide very useful information to professionals studying the functionality of the brain, and are also frequently used by clinicians to diagnose diseases or anomalies in the brain that could be life-threatening.

Two neuroimaging modalities exist: 1) structural imaging and 2) functional imaging. Those techniques that fall in the first category study the structure of the brain at a large scale, such as anomalies caused by trauma or tumours, whereas those in the second category are used to detect anomalies at a finer scale, for instance to perform neurological and cognitive studies (i.e. the study of mental processes related to memory, attention, thinking, etc.) (Zani et al., 2002). Here I focus on the first type only.

In particular, several structural imaging techniques are currently available to image the brain, and each of them provides different information at a different cost, with different spatial resolution and with or without any associated health risk for the patient. The most popular ones include computed tomography (CT) scans and magnetic resonance imaging (MRI) and, to a lesser extent, positron emission tomography (PET). Computed tomography (CT) scans consist of the absorption of X-rays in different directions, and hence the use of ionising radiation. On the contrary, MRI scanners obtain images by processing the electromagnetic signal emitted by protons in the body after the application of strong magnetic fields, which pose a smaller risk to the patient. In practice, CT scans are preferred to diagnose acute problems, for instance after a life-threatening head injury or a neurological disease that requires immediate treatment. This is mainly because CT scanners are widely available, they can be applied to patients with ventilation support or aggited, they are cost-effective and require less acquisition and processing time than MRI scanners (Imtiaz, 2016). Instead, MRI is often used to diagnose neurological diseases that do not require immediate treatment. Nevertheless, researchers have shown that, unlike CT scans, MRI scans provide high resolution images in which gray matter can be distinguished from

white matter (Bunge and Kahn, 2009), and result in more accurate stroke diagnostics (Moreau et al., 2013; Jeena and Kumar, 2013).

Alternatively, PET scanners can also be used to image the structure of the brain at a much lower resolution. This imaging method is based on the detection of gamma rays after the injection of a radionuclide in the body. Despite its advantadges in providing information on the metabloism with respect to the other two methods (Zhang et al., 2017), PET scanners are infrequently used for structural imaging as they use ionising radiation, they provide low resolution images of the brain and are expensive to operate given that these require a cyclotron that generates the (short-lived) radioactive tracers (Bunge and Kahn, 2009).

Therefore, there is a need for an imaging technique that is safe, portable, cost-effective, and that can provide reliable and quantitative information of the structure of the brain at a resolution comparable to MRI. With this aim, I propose to use reflection and transmission ultrasound through the head.

Conventional ultrasound scanners are safe, affordable and portable. However, these fail in reconstructing an image of the brain of an adult patient through the skull at high resolution (Ylitalo et al., 1990; Smith et al., 2009). This is mainly due to three reasons: 1) the skull-bone of an adult is highly reflective, which dominates the reflection signal typically measured in conventional ultrasound, and masks back-scattered reflections within soft tissues in the brain, 2) high frequencies are typically used in conventional ultrasound, which are more easily absorbed than low frequencies, especially within the skull, and 3) conventional tomographic reconstruction methods are based on ray theory and do not account for the right physics of wave propagation.

These difficulties are less important when imaging the neonatal brain and, in fact, conventional ultrasound is routinely used for this purpose. This is mainly possible because ultrasound waves are directly transmitted through the fontanelle[1] (Correa et al., 2004; Ecury-Goossen et al., 2015) and, therefore, these travel directly into the

[1]The fontanelle is the space in the bones of the skull in an infant or fetus before the skull is completely formed.

brain without crossing the skull.

In this chapter I propose and demonstrate that transmission and reflection of low-frequency ultrasound (below $1\,MHz$) through an adult head, combined with the use of 3D FWI, allows non-invasive imaging of the brain. This could have a high impact within the medical community if proofed in the laboratory, as it could be adapted to replace CT scans for rapid diagnostics of life-threatening and acute brain injuries and diseases. The main advantages with respect to the latter are the portability of ultrasound systems – thus enabling imaging at the bed-side – the use of safe non-ionising radiation and the higher spatial resolution for soft tissue, comparable to that of MRI, at an affordable cost.

An immediate application of this neuro-imaging method is imaging the brain for stroke diagnosis. Stroke was the second most common cause of death worldwide in 2013, and the third leading cause of disability (Feigin et al., 2017). The latter are classified in two categories: ischaemic and haemorrhagic. Whereas the first is caused by an occlusion of an artery in the brain that prevents normal blood flow, the second is produced by a rupture of a blood vessel or as a result of trauma (Donnan et al., 2008). Because the treatment for these two types of stroke is very different – nearly the opposite – an incorrect diagnostic can have fatal consequences (Donnan et al., 2008). Furthermore, rapid diagnostic of stroke and the consequent immediate treatment improves the outcome (Belt et al., 2016).

Here I suggest the current method for neuroimaging with ultrasound could be used to differentiate between stroke types at the ambulance due to its portability, thus enabling faster recovery and improved survival rates. I demonstrate the potential of the current method in differentiating stroke types for an *in silico* dataset within this chapter.

Other applications of this imaging method are possible, and may include detection and imaging of brain tumours.

This chapter is divided in three main sections. Within the first section, I demon-

strate the potential of FWI of low-frequency ultrasound data for imaging a realistic *in silico* adult head, both in 2D and 3D, and I show it can potentially be used to differentiate stroke types. The first part of this first section is dedicated to demonstrate that reflections as well as transmissions are necessary to reconstruct a model of the brain, that a relatively accurate and smooth model of the skull is necessary for a successful inversion, and that having noise in the data as well as ignoring density and absorption effects do not compromise significantly the quality of the recovered models of the brain. I then demonstrate that it is possible to distinguish between stroke types with ultrasound due to the different acoustic properties of blood and a blood clot by generating and inverting data for a modified 2D section of the brain, which includes both a haemorrhage and a blood clot. In the last part of the first section I show that 3D FWI is necessary to successfully image the brain, and that ignoring 3D effects results in an incorrect and inaccurate recovered model.

Within the second section, I explore two methods to determine the skull prior to the inversions. As demonstrated throughout this chapter, this is a key element that needs to be addressed for a successful reconstruction of the brain.

Finally, the results are then summarised and further discussed in the last section of this chapter, and future lines of research are examined.

6.2. 3D brain phantom

All experiments in this chapter are based on the MIDA 3D human head model (Iacono et al., 2015). In order to use this segmented head model, which is obtained from MRI images, realistic acoustic velocities are assigned to each segment based on experimental values when available (Hasgall et al. (2015) and references therein). For those regions in which the acoustic velocity has not been previously measured, estimates of their velocity are used based on the properties of surrounding tissues and tissues with similar characteristics. Figure 6.1a shows three planes from a 3D view of the resulting 3D velocity model of the head, which is used for the 3D experiments in this chapter. On the other hand, Figure 6.1b displays the 2D transverse section used within this chapter for the 2D experiments.

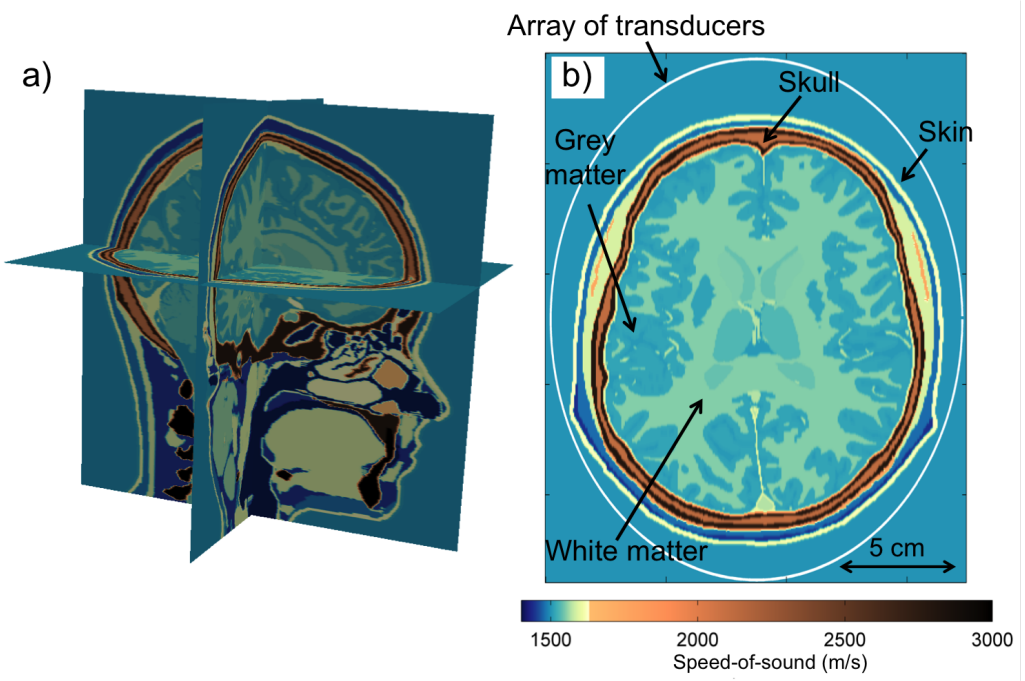


Figure 6.1.: *In silico* models of the speed-of-sound within the head used for a) 3D and b) 2D experiments.

Three main regions can be distinguished in the latter 2D section: 1) the brain, mainly formed of white and gray matter, 2) the skull, formed of three layers (the inner and outer tables, which have similar acoustic properties, and the diploë, which is the spongy cancellous bone between the inner and outer tables) and 3) the outer skin.

The 3D model is discretised so that the grid spacing is $0.8\,mm$, resulting in a model of dimensions $231 \times 321 \times 321$ in cells. This allows a relatively fast data modelling and inversion computation with little dispersion of acoustic waves below the maximum inversion frequency of $650\,kHz$ (a highly accurate kernel that only requires 3 grid-points per wavelength is used in 3D). Regarding the 2D model, this is discretised with a grid spacing of $0.5\,mm$, which ensures there is no dispersion below $550\,kHz$ with the 2D kernel (this requires at least 5 grid-points per wavelength to avoid numerical dispersion).

Various tests are now performed using these models within this section. First, the feasibility of FWI to recover a model of the brain is tested using the *in silico* 2D model in Figure 6.1a, and different aspects that may lead to an incorrect and innacurate

result are assessed. Next, the 2D *in silico* model is modified to simulate a stroke caused both by a hemorrhage and a blood clot, and the potential of FWI to recover these anomalies is tested. Finally, 3D FWI of 3D ultrasound data is performed and discussed, and the importance of accounting for 3D effects is analysed in this section.

6.2.1. 2D inversion

The feasibility and reliability of FWI for brain imaging is now tested on the realistic 2D *in silico* model of the head in Figure 6.1b for a healthy brain, and also on a modified model that includes a haemorrhage and a blood clot.

For both models, a grid spacing of $0.5\,mm$ ensures there is little dispersion of P-waves below $550\,kHz$. This results in a 2D model of dimensions 350×480 in cells. Ultrasound data are generated with a densely sampled array of ultrasound transducers around the head, as indicated by the white ellipse in Figure 6.1b, which contains 512 uniformly distributed transducers that can act both as emitters and receivers. The source wavelet used is a the three-cycle toneburst, as shown in Figure 6.2a, with a centre frequency of $350\,kHz$ (see amplitude spectrum in Figure 6.2b).

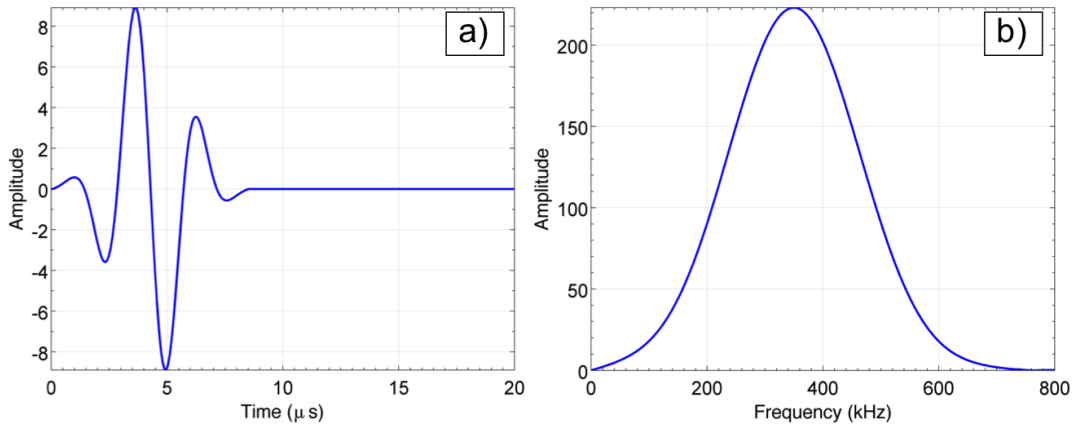


Figure 6.2.: Three-cycle tone-burst source wavelet a) in time and b) its amplitude spectrum used to model 2D *in silico* data for a transverse section of the 3D head model.

A total of $300\,\mu s$ of data are then recorded with a time sampling of $0.08\,\mu s$ to ensure numerical stability.

6.2.1.1. *In silico* healthy brain

Figure 6.3a shows a representative shot gather generated by a single ultrasound transducer for the 2D model in Figure 6.1b, and Figure 6.3b displays the same gather with added random noise.

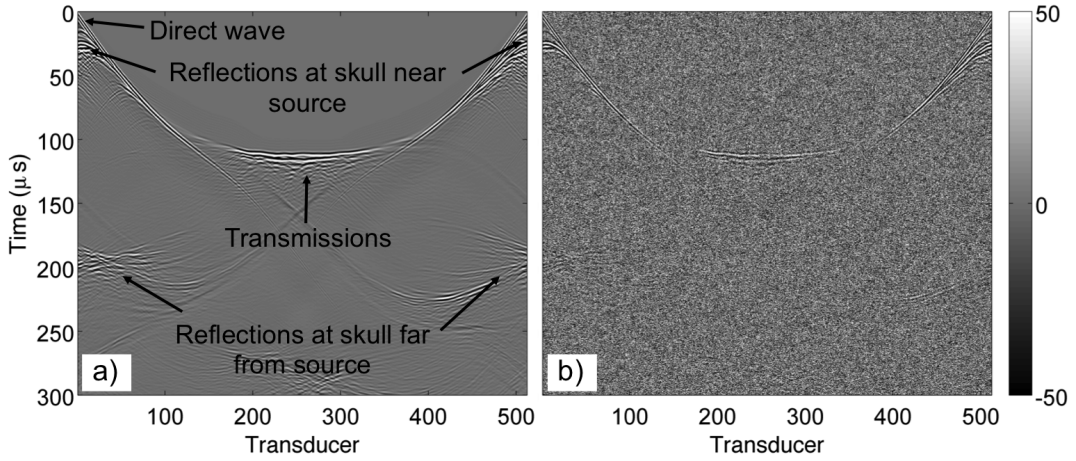


Figure 6.3.: Representative shot gather obtained for the 2D model of the head in Figure 6.1b a) before and b) after adding random noise.

Four distinctive features can be observed in the data (as indicated on Figure 6.3): 1) the direct wave that propagates through the water bath, 2) reflections near the skull, 3) reflections at the skull but far from the source and 4) transmissions through the head.

In conventional ultrasound tomography, only near-incidence reflections are used, and information that is transmitted through the head is not recorded. This limits the capacity of imaging the brain considerably, as will be shown later in this section, mainly due to the large impedance contrast between the surrounding water and the outer table of the skull, as well as the contrast between the inner table and the cerebrospinal fluid (CSF). This is due to the faster acoustic speed of ultrasound within the skull and larger densities than the soft tissue in the brain. As a result, reflections at the skull are much larger than those happening within soft-tissue variations in the brain, and are therefore obscured by the presence of the former. However, if transmissions are recorded and low frequencies are used – which are less attenuated than higher frequencies within the skull – the recorded data on the other side of the

head does not contain any strong reflection that dominates the signal. If FWI is then performed on the full recorded dataset, a high-fidelity high-resolution recovered model can be obtain. This is shown in Figure 6.4 for three different starting models.

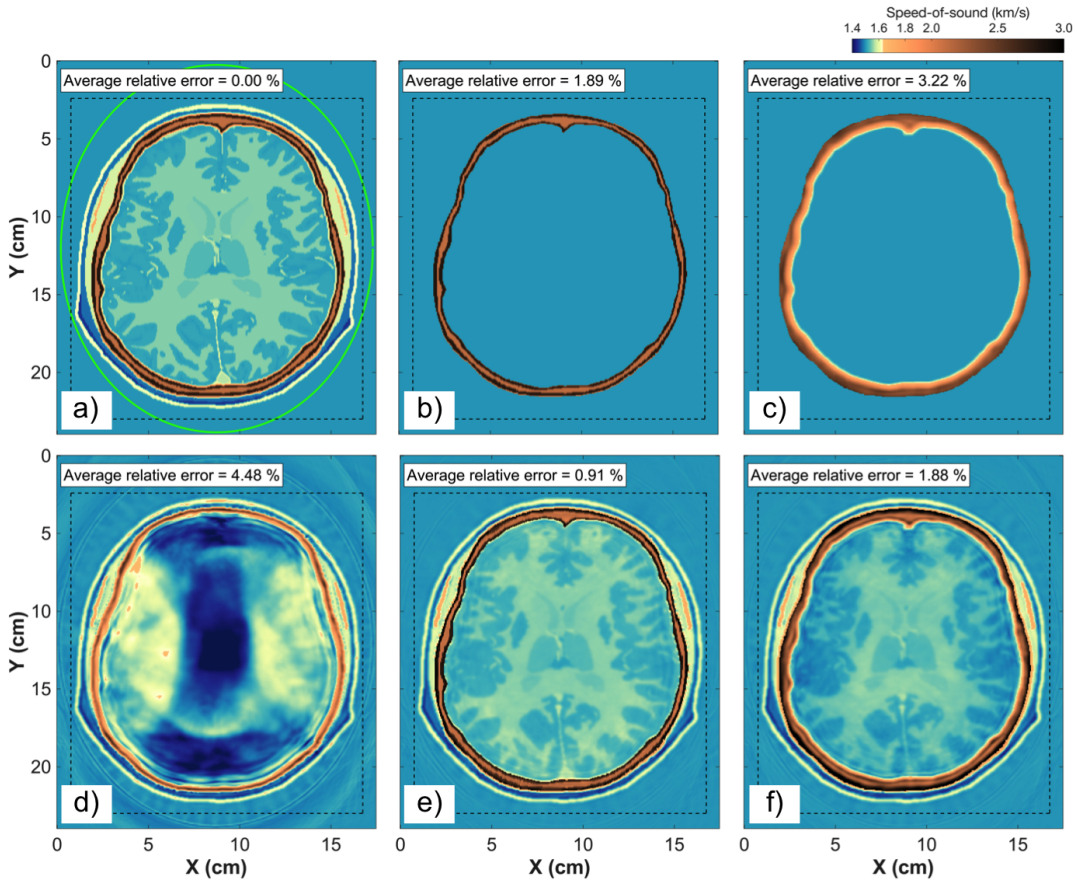


Figure 6.4.: Transverse sections of the head for the a) true and starting models with b) the exact skull and c) a smooth skull. The bottom pannels show the recovered models of the head obtained with starting models of d) just water, e) water and the exact skull as in b) and f) water and the smooth skull as in c).

The true model is displayed in Figure 6.4a as a reference, in which the green ellipse indicates the acquisition geometry. For all the models in this figure and subsequent ones, the average relative error is computed within the black dotted box with respect to the true model: this gives an indication of the difference with the true model, such that a model close to the true model has smaller relative error values, and is used to assess the quality of the results. Figures 6.4b and 6.4c show two starting models, which consist of the exact true skull and the latter after having applied smoothing towards the centre of the model only, respectively. In both starting models, the in-

terior of the head is filled with a homogeneous water model. These two models are designed to evaluate the importance of having an accurate model of the skull in the starting model, given that the latter shows the largest impedance contrasts within the model and, thus, could lead to unwanted cycle-skipping effects if not properly considered. On the contrary, the soft tissue in the brain has similar properties to water and, therefore, a homogeneous water starting model might be accurate enough to recover the brain structure.

To test these hypotheses, 2D FWI of the observed data is performed using the latter two starting models and a model with just water and no information on the skull characteristics. For all these inversions, the observed dataset is inverted in a multi-scale approach from $100\,kHz$ up to $550\,kHz$ in a total of 110 iterations, and using 65 sources per iteration. This results in a computation time of only $1.5\,h$ across 11 nodes with 6 tasks per node.

Figure 6.4d shows the result after using a homogeneous water starting model, whereas Figures 6.4e and 6.4f respectively show the results obtained with the starting models in Figures 6.4b and 6.4c. I make the following observations:

1. Despite starting at low frequency, using just a homogeneous water model is not enough to reconstruct a model of the brain due to the large impedance contrast at the skull, and only the outer skin and the shape of the skull are recovered.

2. FWI with an exact model of the skull results in a very accurate reconstruction of the brain with high resolution (at $550\,kHz$ and for an average speed-of-sound of the brain of $1550\,m/s$, the wavelength is $2.8\,mm$, so that the resolution is $1.4\,mm$), only limited by the limited bandwidth in the data and the range of frequencies in the inversion.

3. Using an approximate skull starting model is not critical, as all structures are correctly recovered but with slightly slower average velocities. This is due to the starting skull model being also slower, which is modified during the inversion.

These results demonstrate that, unlike conventional ultrasound tomography meth-

ods and systems, FWI can be used with low frequency reflected and transmitted ultrasound to image the brain at a resolution close to a millimetre, which is sufficient to distinguish different structures within the brain. Furthermore, it is clear that prior information of the structure and characteristics of the skull need to be included to obtained accurate recovered models of the brain with FWI. Two ways on how to tackle this will be discussed in section 6.3.

However, the previous tests need to be considered with care, as different factors that may adversely affect the inversions have not yet been taken into account. Instead, I have assumed data are noise-free and that the density and absorption of the head are the same as water. Thus, I now test how these influence the quality of the results, and what is the recovered model that can be obtained when only reflections are considered, similar to conventional ultrasound tomography.

First, I test what is the effect of having strong random noise in the data. Figure 6.3b displays the same shot gather as in Figure 6.3a after adding white noise, which dominates over the signal for most transducers. Nevertheless, the recovered models in Figures 6.5a and 6.5b for the exact skull and smooth starting models in Figures 6.4b and 6.4c, respectively, illustrate that random noise does not significantly affect the quality of the results, and it only introduces some noise in the inversion results.

Thus, this confirms that FWI is robust against random noise and, consequently, data acquired in future laboratory experiments will likely be sufficiently clean for FWI to work.

Next, I perform FWI of short-offset data only, thus excluding transmissions. This is analogous to conventional ultrasound tomography. For this purpose, I exclude the data for those transducers that are further than $3\,cm$ from the emitter, such that only reflections are considered (see Figure 6.4 for an idea of what this distance represents in comparison to the dimensions of the model and the acquisition geometry), and perform 2D FWI. Figure 6.6 compares the result of inverting the full dataset (Figure 6.6a) to only inverting for reflections (6.6b).

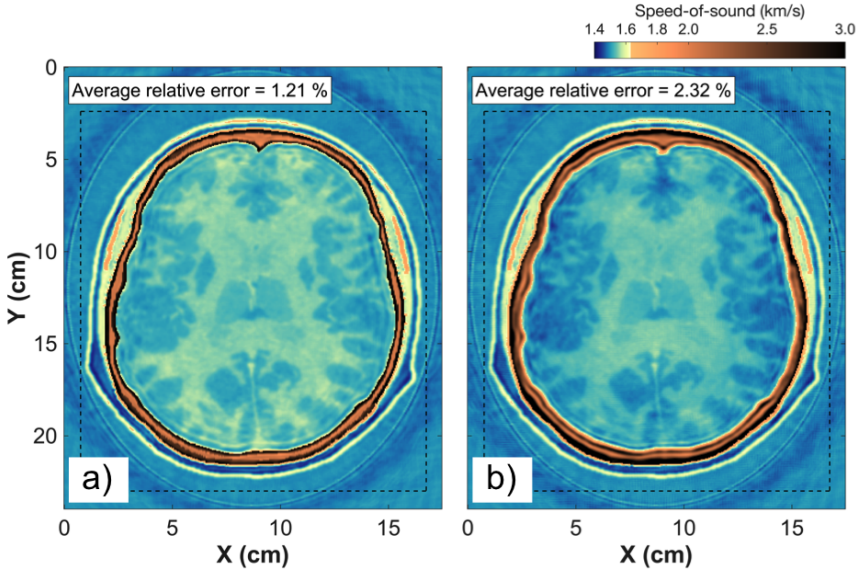


Figure 6.5.: Recovered transverse sections of the head with FWI using a) the exact skull starting model in Figure 6.4b and b) the smooth skull model in Figure 6.4c for the noisy observed data.

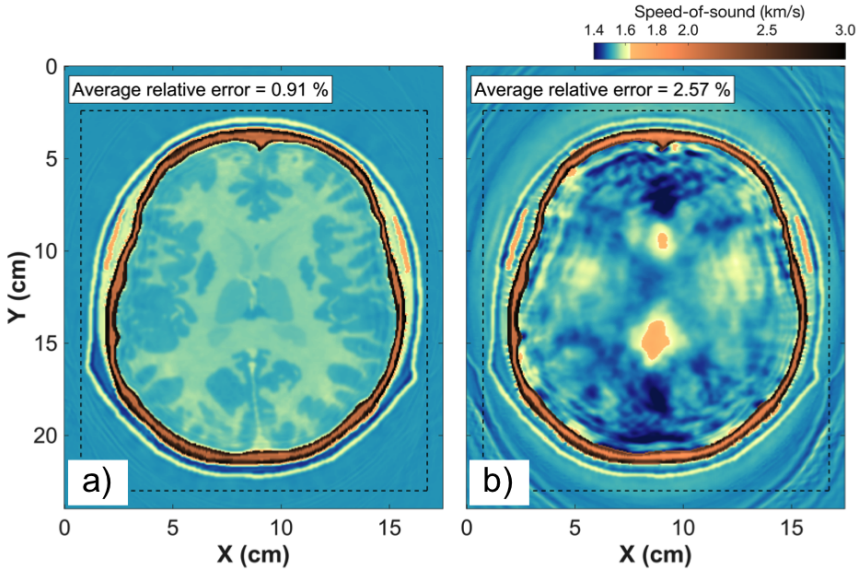


Figure 6.6.: Recovered transverse sections of the brain obtained with the exact skull for a) the full dataset and b) only the nearest offsets (below 3 cm source-receiver offset).

As the reflection data is dominated by reflections at the skull and reflection within brain soft-tissue variations are much weaker, FWI of reflection data is unable to recover the brain. This demonstrates that acquisition of transmission data as well as reflections is essential to image the structure of the brain.

Nevertheless, the previous tests have been performed assuming a constant density and negligible signal absorption for the brain and skull tissue, keeping them constant and equal to those in water. Thus, I now investigate the impact of modelling data using the same speed-of-sound model but realistic models for density and absorption, which are respectively shown in Figures 6.7a-c, and inverting the data ignoring these effects. The latter have been obtained from Hasgall et al. (2015) and references therein when possible, and estimated from other available measurements of similar tissues when not possible (appendix A shows how to transform absorption coefficients from the units used in medical imaging to those used in geophysics).

With the latter models, I generate three observed datasets. The first is generated with the true acoustic velocity model and true density, but ignoring absorption, whereas the second is also generated with the true acoustic speed model but with constant density equal to that of water, and the absorption model in Figure 6.7c. Finally, the third dataset is generated with the true models for acoustic speed, density and absorption shown in Figures 6.7a-c. 2D FWI is then performed for each of these datasets, and the results are respectively displayed in Figures 6.7d-f. In order to minimise the impact of density and absorption in the result, I use an amplitude-balanced functional, which produces optimal results.

Figure 6.7d shows the result of modelling the data with realistic density and inverting it assuming water density. As a result, some areas of the brain are brighter, given that the impedance contrast during the inversion is incorrect, leading to incorrect wave dynamics. However, all the structures are fully recovered, and the average relative error is slightly above that obtained in Figure 6.4e when data was modelled and inverted with constant density. Thus, I conclude that density effects are not relevant.

The effect of modelling data with realistic absorption and fixed density and ignoring absorption during the inversion is shown in Figure 6.7e. Again, the brain structure is well recovered, but the average velocities are now slower. As in the experiments in chapter 4 and 5, this is due to the time delay and amplitude attenuation caused by absorption. Its impact on the average relative error is similar to that of neglecting density, and it also does not compromise the quality of the recovered model.

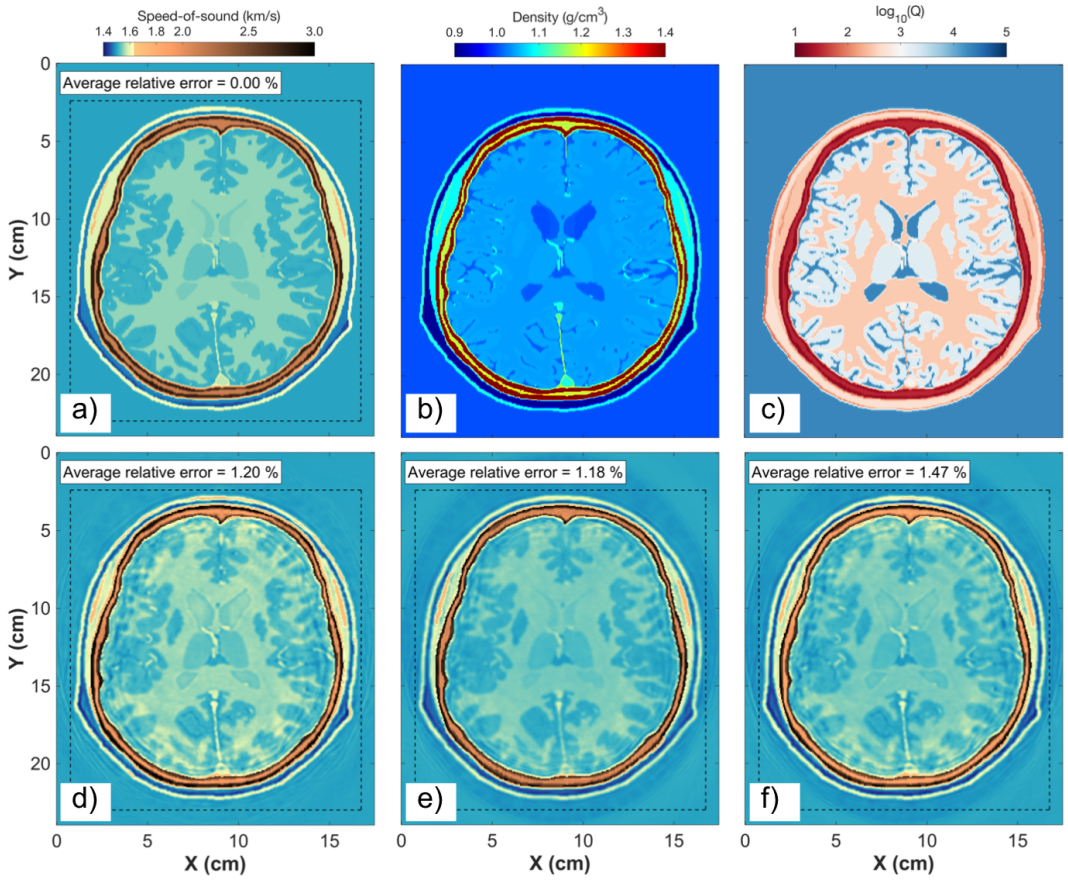


Figure 6.7.: Transverse sections of the brain for the true a) acoustic speed, b) density and c) absorption models. The bottom panels show the recovered models of acoustic speed obtained with FWI ignoring density and absorption effects for data generated d) with variable density and no absorption, e) with fixed density and variable absorption and f) with variable density and absorption. An amplitude-balanced functional is used to minimise the effect of the latter, and the exact skull starting model is used throughout.

Finally, Figure 6.7f confirms that the combined effect of neglecting absorption and density for the third dataset still yields an acceptable and accurate recovered model of the brain, in which white and gray matter can clearly be differentiated and the structure of the brain is clearly visible.

In summary, the results in Figure 6.7 illustrate that neglecting the effect of density and absorption does not compromise the quality of the recovered models of the brain significantly, especially when using an amplitude-balanced functional. More accurate results could be obtained if multi-parameter inversion was performed or if an experimental relationship between acoustic speed and density was used, equivalent

to Gardner's law in geophysics. As an example of the latter, Figure 6.8 shows that P-wave velocity and density may be fitted with a polynomial function with known coefficients when considering experimental values for these properties for different body tissues. Nevertheless, this relationship should be more accurately determined by measuring these properties for a larger number of tissue samples.

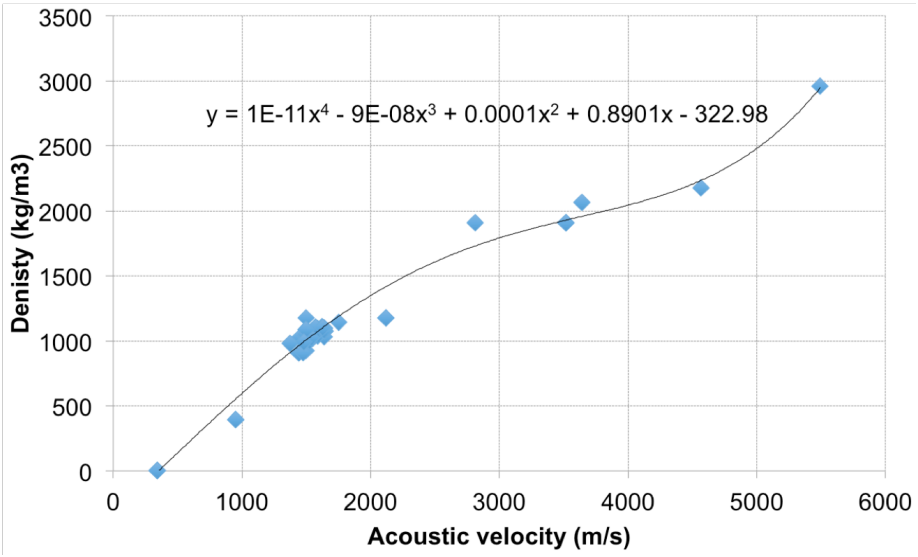


Figure 6.8.: Cross-plot of experimental values for density and acoustic velocity for different body tissues, which have been obtained from Iacono et al. (2015). A polynomial function of fourth order is fitted to the data, which can potentially be used to derive density from acoustic velocity in FWI.

The impact of density and absorption on the recovered models are due to differences in the observed data with that obtained with fixed density and absorption. This is shown in Figure 6.9 for the three datasets corresponding to variable density and no absorption, fixed density and variable absorption and variable density and absorption (Figures 6.9a, 6.9c and 6.9e). The panels in Figures 6.9b, 6.9d and 6.9f show the differences of the latter with the dataset obtained from fixed density and absorption equal to that of water.

The differences in Figure 6.9b are mainly due to amplitude differences, given that impedance contrasts are different than those when the density is fixed. These are therefore minimised when using and amplitude-balanced functional (Warner et al., 2012). On the other hand, the differences in Figure 6.9d are principally due to the time

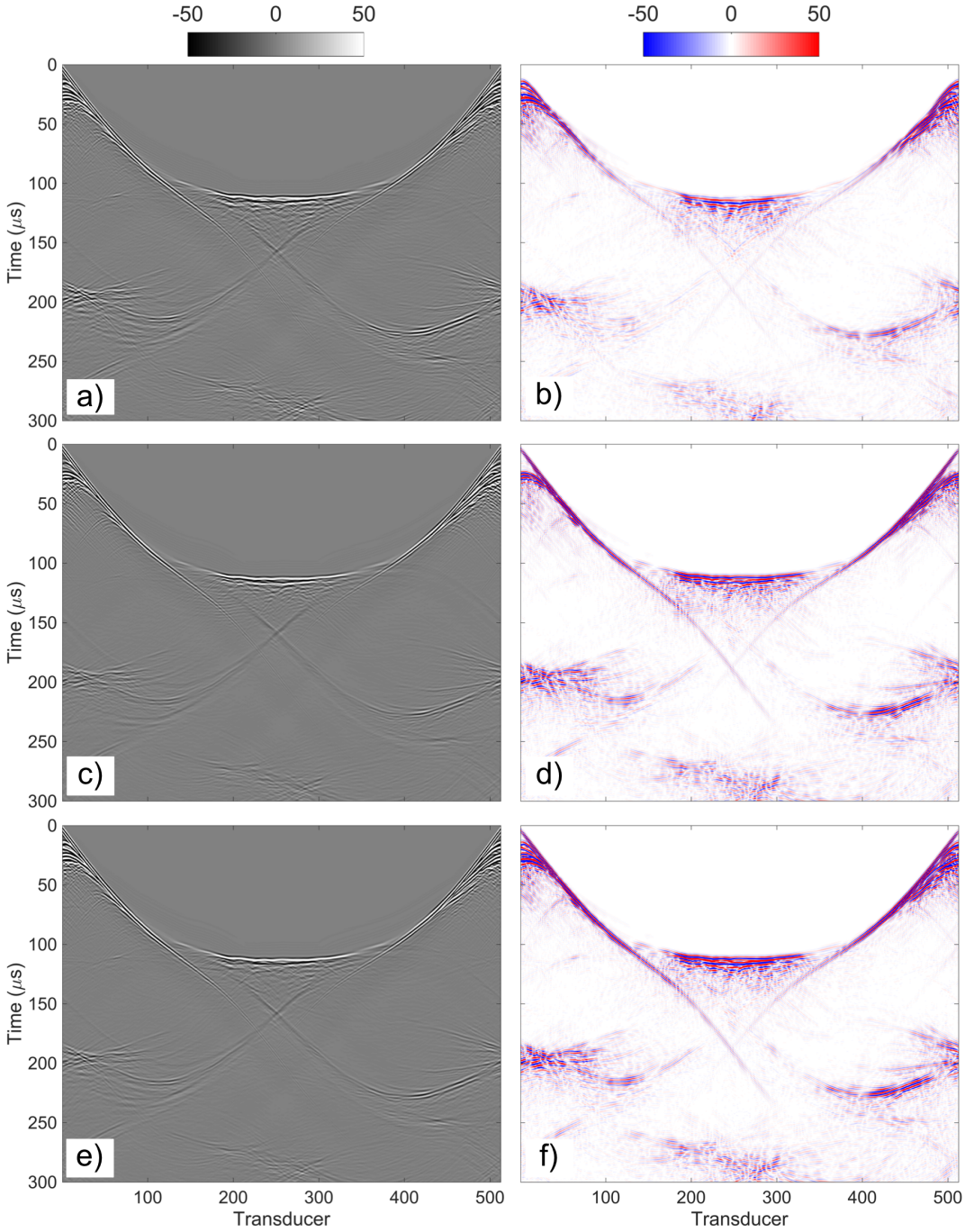


Figure 6.9.: Representative shot gathers obtained after modelling data with a) variable density and no absorption, c) fixed density and variable absorption and e) variable density and absorption. The panels on the right display the difference of the panels on the left with the data generated with fixed water density and no absorption shown in Figure 6.3a.

delay introduced by absorption, and to a lesser extend due to amplitude attenuation (as shown in chapter 4). The combination of the latter two results in larger overall

differences in Figure 6.9f. Nevertheless, these do not halt FWI from obtaining high quality recovered models of the brain, and are therefore neglected in the remaining tests of this chapter.

6.2.1.2. *In silico* brain with blood clot and haemorrhage

In this section I test the capacity to differentiate between ischaemic and haemorrhagic strokes using FWI of transmitted and reflected ultrasound. With this aim, I modify the true acoustic model in Figures 6.1b and 6.4a by adding a haemorrhage and a blood clot in two separate areas of the brain, as indicated in Figure 6.10a. To simulate the latter two, I have used acoustic speeds of blood-filled soft tissue as in Mourad and Kargl (2000), and for thrombus as in Huang et al. (2011). The remaining panels in Figure 6.10 illustrate the recovered models obtained after applying FWI on the dataset generated with this modified head model.

As in Figure 6.4, Figures 6.10b and 6.10c show the exact skull and smooth skull starting models, but now the average relative error is computed with respect to the model in Figure 6.10a. Figures 6.10d-f then show the recovered models obtained with a model of just water, the exact skull starting model and the smooth skull starting model, respectively. As in Figure 6.4, an approximate starting model of the skull is necessary to recover an accurate model of the brain. Furthermore, when using the exact or smooth starting models, both the blood clot and the haemorrhage can be clearly identified and differentiated. This confirms the potential of FWI of ultrasound for stroke diagnosis. The latter also holds in the presence of strong random noise, which results in the models shown in Figures 6.10g and 6.10h for the exact and smooth skull starting models, respectively.

The shot gather in Figure 6.11a and the latter after adding random noise in Figure 6.11b are similar to those in Figure 6.3 obtained for a healthy patient. However, there are differences, and these can be observed in Figure 6.11c, in which the difference between the observed data for the head model with a blood clot and a haemorrhage (Figure 6.11a) and the observed data for a stroke-free patient (Figure 6.3a) is displayed.

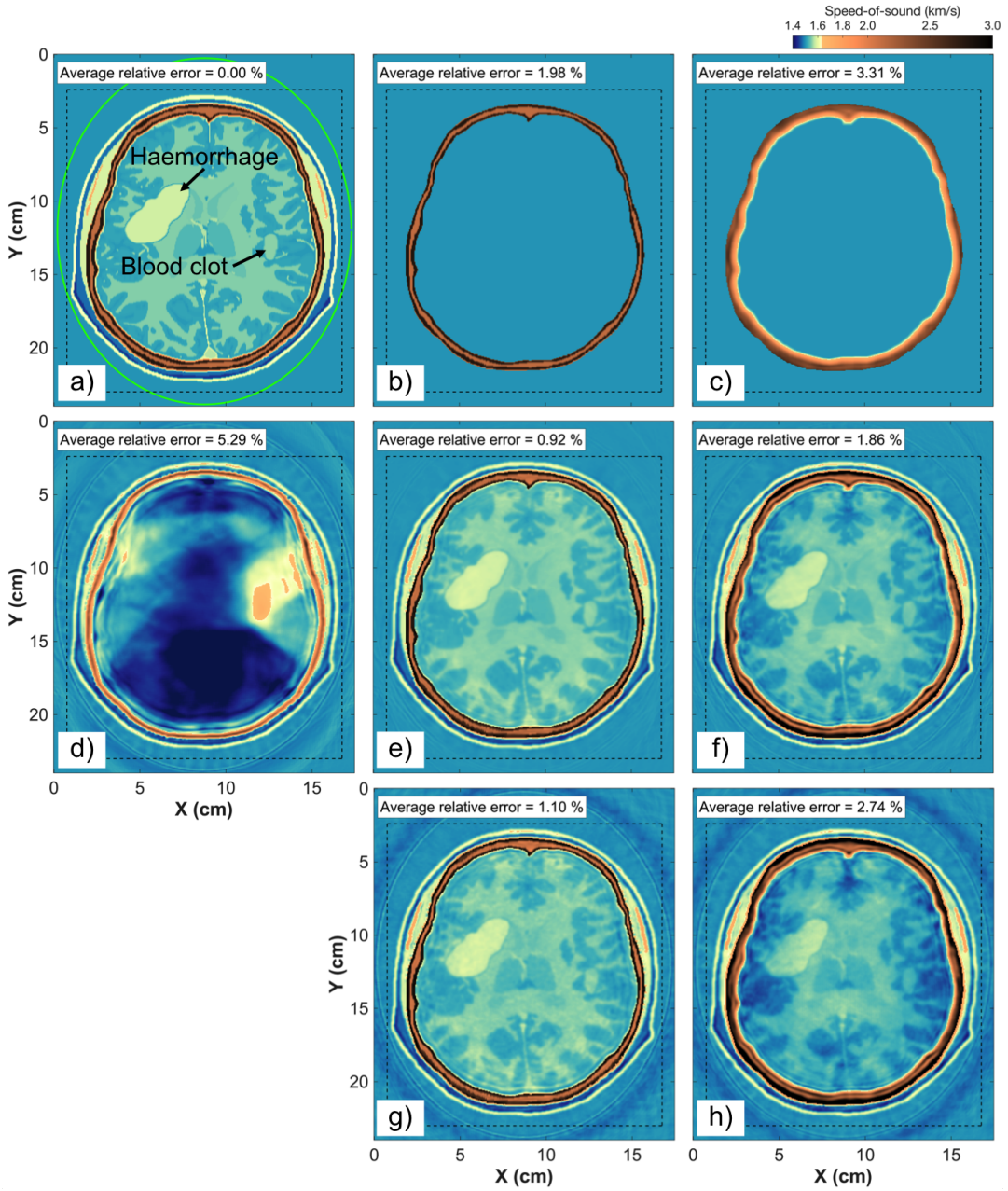


Figure 6.10.: Transverse sections of the brain for a) the modified true model containing a blood clot and a hemorrhage, b) the exact skull starting model, c) the smooth skull starting model and the recovered models after 2D FWI of the noise-free data in d), e) and f) using a water starting model, the exact skull starting model and the smooth skull starting model, and of the noisy data using g) the exact skull starting model and h) the smooth starting model.

There are visible differences for reflections, but also for the transmitted part of the signal – note the difference in scale between Figure 6.11c and the gathers above. These differences are successfully considered by the FWI algorithm, thus resulting in

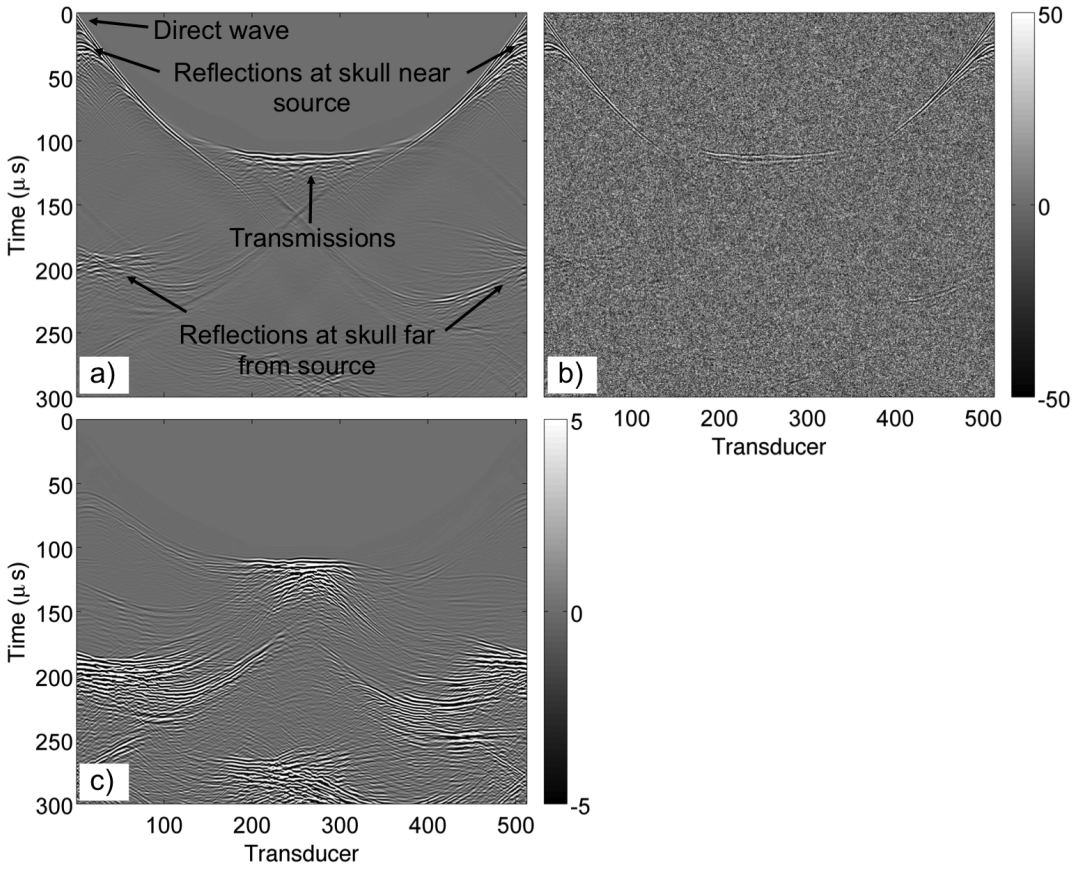


Figure 6.11.: Representative shot gather obtained for the modified head model to simulate a stroke patient a) without and b) with noise. Panel c) shows the difference of the gather in a) with that obtained for a healthy head model, as displayed in Figure 6.3a.

the recovered models in Figure 6.10.

In summary, the latter *in silico* experiments demonstrate the potential of FWI of ultrasound data for stroke diagnosis.

6.2.2. 3D inversion

The feasibility and applicability of 3D FWI to image the entire brain is now tested.

The first experiment is designed to assess the consequences of performing 2D FWI of 3D data, rather than 3D FWI. This is relevant as a stack of 2D inversions could result in a 3D image of the brain at an inferior cost than 3D FWI if 3D effects were not important. For this purpose, I generate two *in silico* datasets obtained with the same acquisition ring as for the 2D experiments, but modelling the data in 2D using

the acoustic model in Figure 6.4a and modelling the data in 3D with the full 3D model of the head in Figure 6.1a. Figures 6.12a and 6.12b respectively show the data modelled in 2D and in 3D, in which amplitudes have been normalised by the average RMS of each shot.

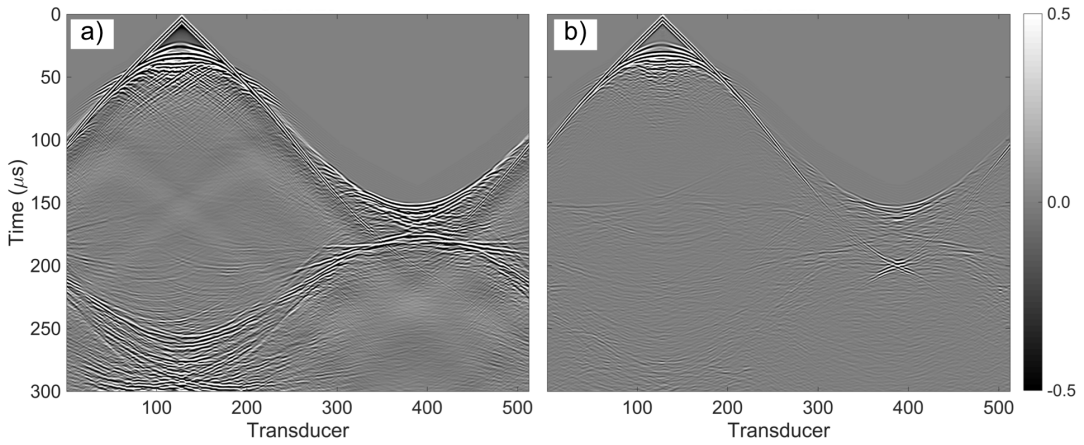


Figure 6.12.: Representative shot gather generated by a) modelling data in 2D across a transverse section of the head and b) modelling data in 3D with the full model of the head. In both cases, the transducers are in a single ellipsoidal array in the same plane as the transverse section shown in Figure 6.4a, and the amplitudes in both gathers are normalised with the average RMS of each shot for comparison.

The comparison between these gathers reveals that differences are important, both qualitatively and quantitatively. These are due to reflections and transmitted arrivals in the 3D dataset generated at the skull and through the brain off-plane, i.e. at planes that are parallel to the acquisition geometry but that are not present in the 2D model.

In order to assess whether these effects are important, 2D FWI of both datasets is performed. The resulting recovered models, respectively shown in Figures 6.13a and 6.13b for the 2D inversion of 2D and 3D data, illustrate that accounting for 3D effects is crucial.

Whereas 2D FWI of 2D data produces a high-quality recovered model, 2D FWI of 3D data results in a model with a discontinuous outer skin, a faster skull and nearly a homogeneous brain with no visible coherent structure. Therefore, I conclude that accounting for 3D effects is essential for a successful recovery of the acoustic speed of the brain with FWI.

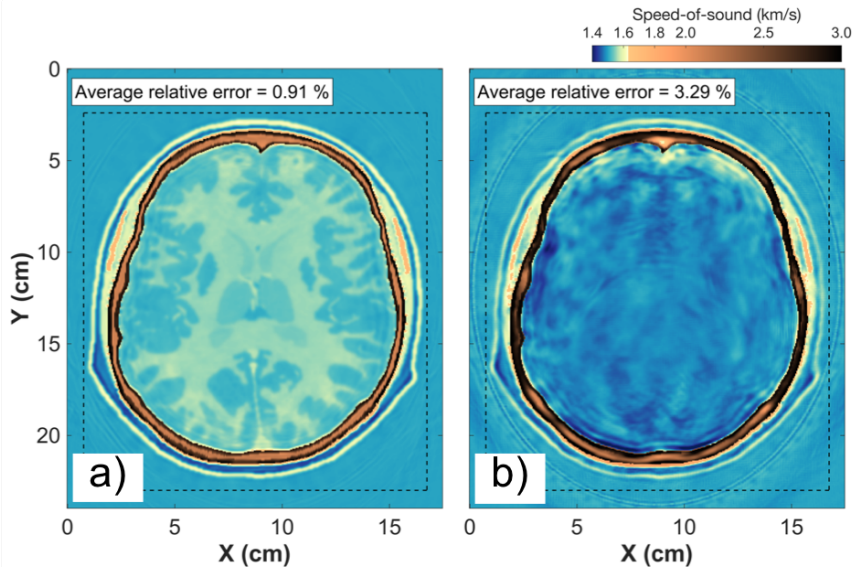


Figure 6.13.: Recovered models obtained by 2D FWI of a) the 2D data and b) the 3D data. Both inversions have been performed using the exact skull starting model and inverting the data from $100\,kHz$ to $550\,kHz$.

Next, I perform 3D FWI of 3D data to image not just a section, but the entire brain. In order to have a good azimuth coverage across different parts of the head, an array of transducers that surrounds the head as if these were in a helmet is designed, taking into account that no sensors can be below the forehead or below the neck. The array of transducers contains 1320 elements in total, being all of them able to emit and receive ultrasound. The source wavelet is again a three-cycle tone-burst, but with a slightly higher centre frequency of $400\,kHz$. 3D ultrasound data are then generated for all the emitters, and 3D FWI is performed by inverting the data from $100\,kHz$ up to $600\,kHz$ in 110 iterations, and using only 88 sources per iteration to reduce the number of resources required.

Figures 6.14 and 6.15 show the results for sagittal and transverse views of the head. In these figures, three inversions are performed using different starting models: a homogeneous water starting model (resulting in Figures 6.14d and 6.15d), the exact skull starting model shown in 6.14b and 6.15b, yielding the results in Figures 6.14e and 6.15e, and a smooth skull starting model as shown in Figures 6.14c and 6.15c, leading to the recovered models in Figures 6.14f and 6.15f.

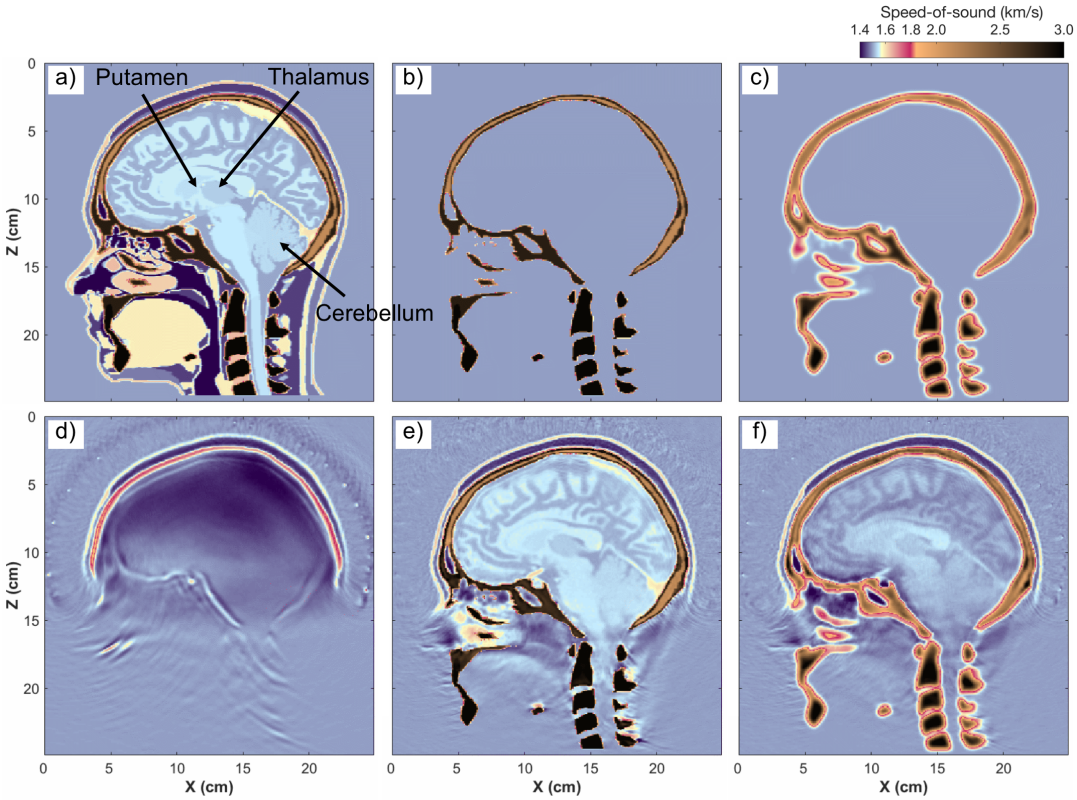


Figure 6.14.: Sagittal views of a) the true and starting speed-of-sound models of the head with b) the exact skull and c) the skull after smoothing is applied. The bottom panels show the recovered models of acoustic speed obtained with 3D FWI using a d) homogeneous water starting model, e) the exact skull starting model and f) the smooth skull starting model.

All the inversions shown in Figures 6.14 and 6.15 have been performed using 89 nodes and 1 task per node, with a total elapsed time of 30 h. The initial aim of these experiments was not to perform fast inversions, but to demonstrate the potential of 3D FWI to safely image the brain. More research should be done in the future to implement this at the speed that the brain demands.

The results in Figures 6.14 and 6.15 demonstrate that 3D FWI of low-frequency ultrasound with a relatively accurate starting model yields a high-quality model of the brain, in which sub-millimetre structures can be differentiated. For instance, the putamen and thalamus are successfully recovered on the sagittal views, and gray and white matter can easily be differentiated in both sections.

Analogously to the 2D tests performed in the previous section, having a homogeneous water starting model is not enough to recover the brain, and only the shape of

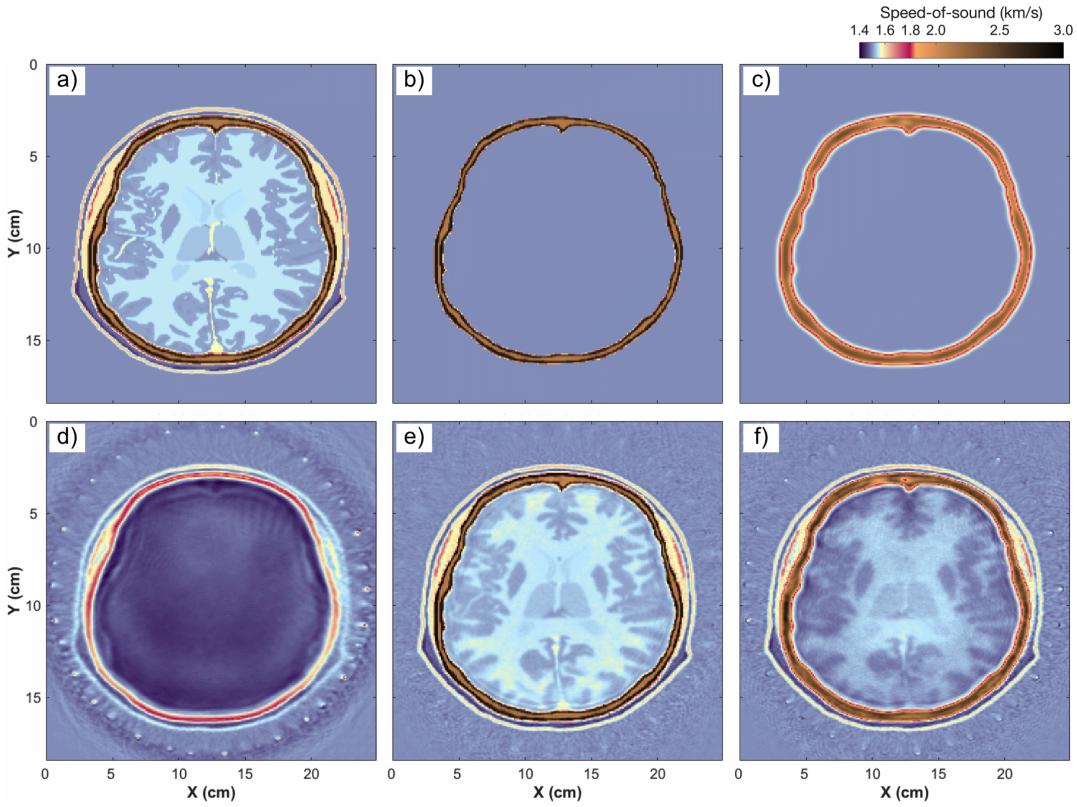


Figure 6.15.: Transverse views of a) the true and starting speed-of-sound models of the head with b) the exact skull and c) the skull after smoothing is applied. The bottom panels show the recovered models of acoustic speed obtained with 3D FWI using a d) homogeneous water starting model, e) the exact skull starting model and f) the smooth skull starting model.

the skull is recovered with incorrect properties. This is because, from the start, the data modelled with a water starting model is cycle-skipped due to not accounting for the large impedance contrast at the skull.

On the other hand, the use of the exact skull as a starting model gives excellent results, only limited by the range of frequencies inverted – hence the resolution could be enhanced by inverting the data up to higher frequencies provided the latter were in the data, but at an added cost – and also by the distribution and number of transducers. Note, for instance, that the base of the cerebellum and anything below that is not well recovered. This is because there is a limited coverage within this region, i.e. waves that travel below this region are not recorded by any transducer.

Finally, the structure of the brain is also recovered when using a smooth starting model, but with slower average acoustic speeds and less level of detail. This is because

the inversion tries to initially modify the properties of the skull, as these are too slow, and that affects the rest of the model. In fact, note that velocities within the three layers of the skull are reversed in the smooth starting model, i.e. the diploë has higher speed-of-sound than the outer and inner tables. Thus, as for the 2D tests, a more accurate result could be obtained by using a starting model in which the velocities of the different layers are closer to the true model.

It is therefore clear that having an approximate but accurate enough starting model of the skull is necessary to obtain reliable recovered models of the speed-of-sound in the brain, both in 2D and 3D. Hence, I now explore two different ways of obtaining approximate starting models of the skull to be used for FWI.

6.3. Estimation of a skull model

I now propose and implement two methods to estimate a model of the skull, which can then be used during FWI. Both methods are based on the use of reflection data to estimate the shape and boundaries of the skull, but they differ in the methodology used to find the inner and outer tables. Whereas in the first method I only use pure reflection data and I implemented it in 2D, in the second method I use near-offset data and I implement it in 3D. Nevertheless, the first method could potentially be extended to tackle 3D problems and to include near-offset data after making some assumptions.

The results obtained for these two methods show that it is possible to obtain prior information of the skull provided reflection data are recorded, and that this leads to improvements in the recovered speed-of-sound models of the brain when compared to using a homogeneous water starting model. However, these are still preliminary methods with certain limitations, which are also discussed.

6.3.1. Method 1

The first method uses pure reflection data assuming that each transducer is able to emit and receive right after emitting ultrasound. In order to make it more realistic,

method 1 is applied to the 2D data generated for the true transverse section of the speed-of-sound model of the head with variable density shown in Figures 6.7a, 6.7b and 6.9a.

The main steps of this method are summarised as follows:

1. Extract zero-offset data (i.e. normal-incidence reflection data).
2. Automatically pick the onset time of the reflection originated at the outer table of the skull.
3. Generate a circle from each source location – assuming these are known– with radius equal to the distance a wave has travelled from the source to the outer table, using the extracted times in 2.
4. Find the points tangent to two consecutive circles and closer to the centre of the model. All points together should define the outer table.
5. Remove outliers.
6. Smooth the position of the clean tangent points and interpolate.
7. Define a polygon across all interpolated points.
8. Repeat steps 2 to 7 for the inner table (it can also be performed simultaneously with steps 2 to 7).
9. Flood the region between the polygons defining the outer and inner tables with typical acoustic velocities for the skull.

In practice, this workflow can be applied in less than a minute on a single node, but it requires careful explanation and some assumptions. This is further elaborated in appendix B.

As a result of this method, two starting models of the skull are obtained making different assumptions to find the inner table: for the first strategy, an assumption is made based on the fact that the skull has three distinct layers and that the wavelet used is a three-cycle tone-burst, whereas for the second strategy the inner table is computed after finding the outer table and flooding the skull with constant thickness.

These two starting models are displayed in Figures 6.16b and 6.16c, whereas the exact starting model is plotted in Figure 6.16a as a reference (see appendix B for more details). Additionally, two more starting models are derived from the model in Figure 6.16c by smoothing the latter towards the centre of the model and in all directions, and these are respectively shown in Figures 6.16d and 6.16e.

The right panels in Figure 6.16 are the result of applying 2D FWI of the observed data generated with variable density and no absorption using the starting models on the left column.

Several conclusions can be drawn from these results. First, using an estimated starting model of the skull results in a better recovered model of the brain than that obtained without prior information (e.g. see Figure 6.4d for data with fixed density). Second, using the starting model obtained with the second strategy assuming a constant thickness of the skull results in more accurate results – note the reduction in the average relative error between Figures 6.16g and 6.16h. Third, the use of a starting model that is smooth in all directions leads to a high-resolution recovered model of the brain (Figure 6.16j), which is more accurate than with any other guess of the skull. The differences of the latter with the recovered model obtained with the exact skull in Figure 6.16f are mainly due to velocity differences within the skull, which translates into a poorer resolution at the borders of the brain and an average slower velocity of the brain. The latter, however, are not significant.

The previous results demonstrate that estimating the skull prior to FWI of the full dataset allows a high quality reconstruction of the brain, and that only an approximate and smooth model of the skull is necessary for that.

As discussed in appendix B, extension of this method to 3D problems is in theory possible, but I have not yet implemented it. Instead, I now present a second method that is directly applied to retrieve prior information of the skull in 3D before a full 3D inversion.

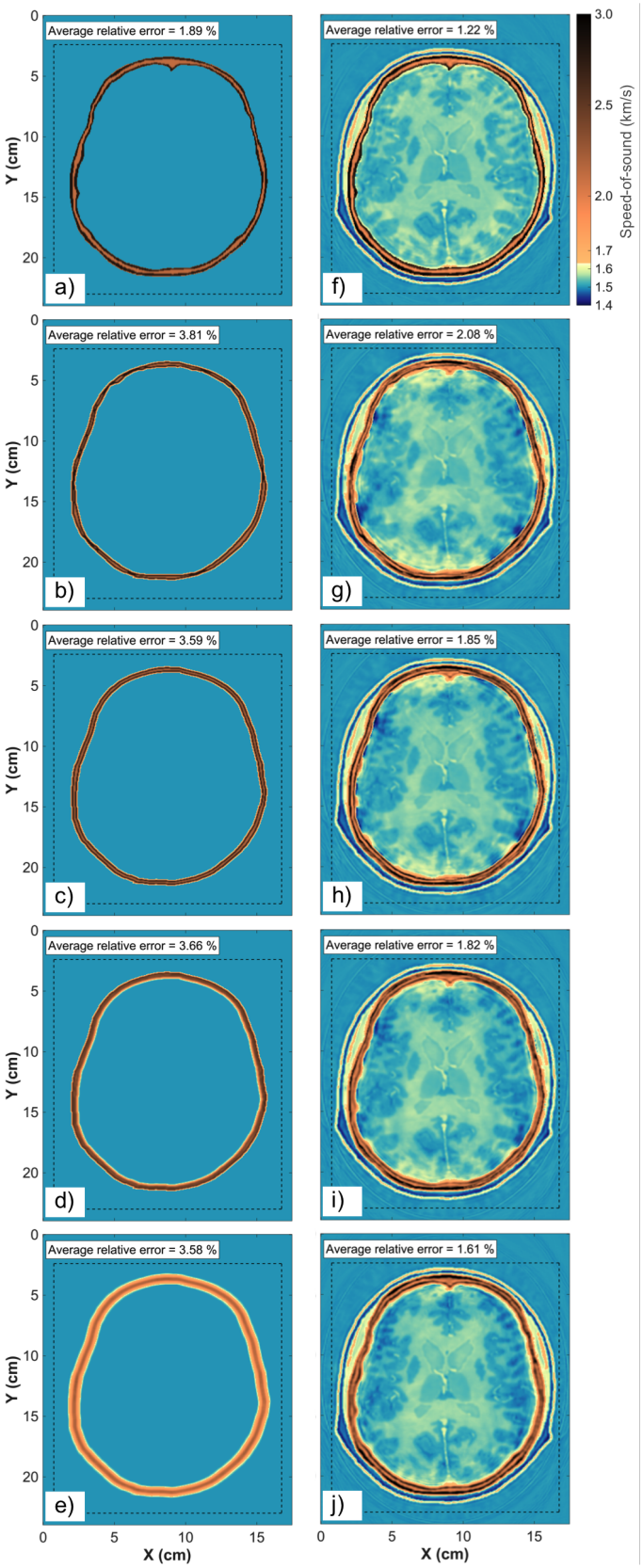


Figure 6.16: Transverse sections of estimates of the skull (left column) and of the recovered models with 2D FWI of data generated with variable density and no absorption using the estimates on the left column as starting models. The estimates of the skull correspond to a) the exact skull and the estimated skulls obtained with method 1 and b) the first strategy, c) the second strategy (flooding with constant thickness), d) inwards smoothing of the model in c) and e) smoothing in all directions of the model in c).

6.3.2. Method 2

The second method to find the skull is based on an initial inversion of near-offset data, i.e. by including normal incidence and near-incidence reflections only. The steps to obtain a starting model are simpler than for method 1, but rely on the use of an efficient 3D FWI code.

The basis of this method is to perform only a few iterations of 3D FWI of near-offset reflection data using a homogeneous water starting model, followed by flooding the recovered model of the skull with typical acoustic velocities. The advantage of doing this before a full inversion is that 3D FWI of near-offset data is fast to compute, and a smooth model of the skull can still be obtained mainly based on reflections. However, the recovered velocities are not correct, as the initial model is cycle-skipped because the velocities of ultrasound within the skull are much larger than those in water. Thus, flooding the recovered skull with higher velocities, closer to those of a typical skull, should result in a starting model for the full inversion that is less cycle-skipped.

Figures 6.17a and 6.17b illustrate the steps of this method for three sagittal and three transverse views of the head, respectively.

Panels a), b) and c) in these figures show the true speed-of-sound models for the different views. In the first step of this method, 3D FWI of near-offset data that corresponds to receivers closer than $14\,cm$ from the emitter are performed, and only the first $80\,\mu s$ of data are used for the inversion. The latter is enough to include reflections at the skull, while decreasing the computation time considerably. A total of 20 iterations are performed from $100\,kHz$ to $150\,kHz$, which results in an elapsed time of $1.2\,h$ when using 89 nodes and 88 sources per iteration. This is a relatively short time when compared to the full 3D inversion.

The recovered models displayed in Figures 6.17d-f and 6.17d-f show the expected recovery of part of the skull with incorrect velocities, which is limited in coverage given the acquisition system so that anything below the eyes or the back of the head is not recovered. Moreover, no information of the brain is recovered given that trans-

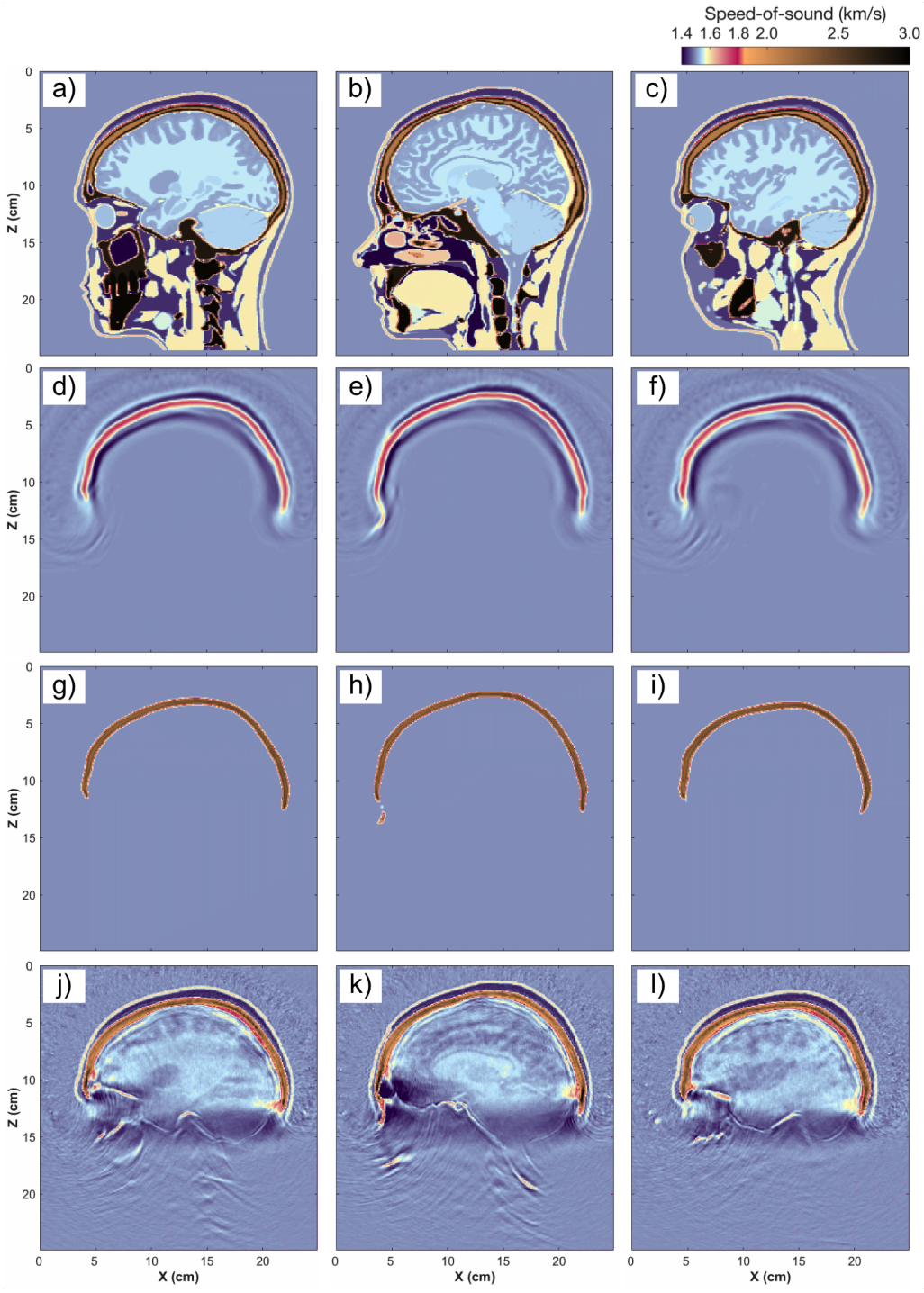


Figure 6.17.: Each column corresponds to a different sagittal section at $X = 17.0\,cm$, $X = 14.2\,cm$ and $X = 11.4\,cm$ of the true model of speed-of-sound (first row), the recovered model obtained by 3D FWI of near-offset data (second row), estimated model of speed-of-sound of the skull (third row) and 3D FWI of the full dataset using the estimated model of the skull.

missions and later reflections are excluded from the inversion.

The second step of method 2 consists of flooding the recovered surface of the skull with typical skull velocities, knowing that this always has three layers. In practice, I first remove values of speed-of-sound that fall below a threshold to eliminate any noise before the skull, followed by an increase of the velocities within the skull while keeping values below the threshold unchanged. I then eliminate the outer skin in order not to assign too large velocities to this – skin is also easily recovered during the inversion–, invert the polarity of the velocities in the skull so that it is slower in the middle and apply smoothing to the recovered skull model. This step is performed automatically, and it takes less than a minute for the entire model.

The estimated models of the skull obtained with method 2 are shown in Figures 6.17g-i and 6.17g-i.

Finally, the last step of method 2 is to perform 3D FWI of the entire dataset using the estimated model of the skull as a starting model, which results in the speed-of-sound models in Figures 6.17j-l and 6.18j-l (the same inversion parameters as for the 3D inversions in section 6.2.2 are used). These recovered models show that it is possible to recover a model of the brain using this method, as the recovered structures and cortical folds coincide with those of the true model. Although the recovered models are accurate for upper regions in the brain, in which there is good coverage, this method is unable to recover anything below a certain depth (see Figures 6.17j-l for depths larger than $12\,cm$). This is because the starting model does not include parts of the skull at the base of the brain, as these are not recovered from the inversion of near-offset data at short times during the first step.

Thus, these results also indicate that further research is necessary to better image the entire brain, including the cerebellum. Nevertheless, the latter are also encouraging, as they demonstrate there is potential to use FWI of ultrasound data to image the brain provided a method to deal with the skull is in place. Some of these possible strategies are now briefly discussed in the next section, and conclusions from the results presented within this chapter are made.

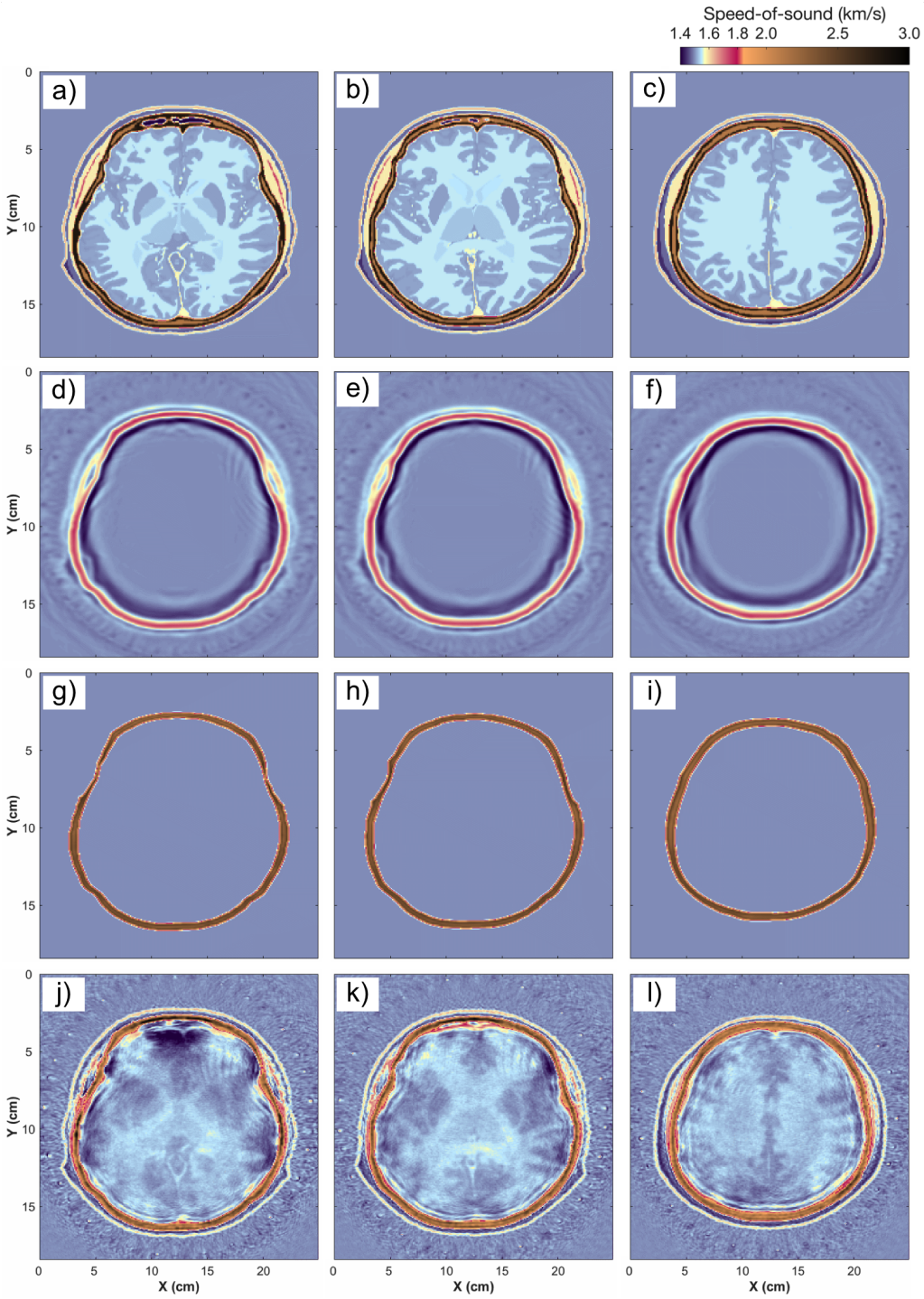


Figure 6.18.: Each column corresponds to a different transverse section at $Z = 8.8\,cm$, $Z = 8.2\,cm$ and $Z = 6.5\,cm$ of the true model of speed-of-sound (first row), the recovered model obtained by 3D FWI of near-offset data (second row), estimated model of speed-of-sound of the skull (third row) and 3D FWI of the full dataset using the estimated model of the skull.

6.4. Summary

Structural imaging of the brain is essential to diagnose acute neurological diseases and brain injuries that can potentially be life-threatening. Although MRI scanners can provide more useful information and images of the brain with better resolution than those obtained with CT scanners, the latter are typically employed as they are cheaper and faster. However, CT scans are only performed once a patient reaches a hospital due to the lack of mobility of the system and the use of ionising radiation.

In this chapter I have suggested the use of low-frequency ultrasound systems to image the structure of the brain. Although this is currently only performed on neonates through the fontanelle and is thought to be impossible on adults due to the acoustic properties of the skull, I have demonstrated that ultrasound can be used to image the brain of an adult. I have proofed this *in silico* by using low-frequency transducers, recording reflected but also transmitted data through the head and, unlike conventional ray-based ultrasound tomography systems, using an accurate reconstruction method such as FWI, which accounts for the right physics of wave propagation. The benefits of using ultrasound is that these are safe, they can provide quantitative information of the tissue properties potentially fast and ultrasound systems are portable and affordable.

Within the first part of this chapter I have analysed what conditions are necessary to obtain a good reconstruction of the brain with ultrasound and FWI. I have done this by analysing the recovered models of speed-of-sound obtained with FWI of low-frequency ultrasound with or without considering transmissions, with or without considering density, absorption or 3D effects and having prior information on the skull characteristics.

The first conclusion from these tests is that having low-frequency data helps reduce the non-linearity of the inverse problem, especially due to the large impedance contrast at the skull, and that recording transmission data is necessary to recover an image of the brain. In fact, conventional reflection ultrasound tomography fails

mainly because reflections are dominated by reflections at the skull, and those taking place at the soft tissue variations within the brain have much smaller amplitudes. Instead, transmissions are not dominated by a strong reflector and they can be successfully used in FWI to recover a model of the brain. The use of low-frequency transducers is also important, as higher frequencies are more affected by absorption and non-linear propagation effects can take place at high frequencies and large amplitudes (Zhao et al., 2014; Pinton et al., 2011).

With respect to density and absorption, ignoring these effects can negatively affect the quality of the recovered models of the brain, but minimally. The main variations in density and absorption with respect to the values for water are at the skull, and these could be considered during the inversion in various way to improve the accuracy of the recovered speed-of-sound models. One possibility is to use an experimental law that relates density with acoustic speed, similarly to geophysics. Other methods might be to invert for density and absorption, or to remove these effects from the data prior to the inversion (e.g. by using matching filters and estimated of the Q and density models). But elastic effects have not been considered here and these may also have an impact on the results. Although the shear modulus of the brain tissue may be negligible, the skull has a non-zero shear wave modulus (White et al., 2006) and, therefore, elasticity may also play a role. However, based on the knowledge from geophysical experiments, elastic effects should not be as relevant as density and absorption effects (Warner et al., 2012), and I have demonstrated that the latter do not prevent FWI from recovering the structure of the brain.

Another aspect that has been considered within this chapter is the importance of accounting for 3D effects. The obtained results show that 2D FWI of 3D data leads to inaccurate recovered models, in which the structure of the brain is not recovered. On the contrary, 3D FWI of 3D data performs well and is able to recover sub-millimetre models of the brain for frequencies lower than $1\,MHz$.

For the experiments performed in this chapter, a full 3D inversion takes roughly $30\,h$ across multiple nodes, but many strategies can be employed to reduce this time,

which have not been tested here. The aim should be to reduce this down to minutes. Currently, this is difficult to achieve, but it should be possible within the next decade thanks to advances in compute power, the use of Graphics Processing Units to perform certain operations in FWI or the use source encoding techniques to reduce the total cost. Once the total time is reduced to just minutes, neuro-imaging with ultrasound could be implemented in the ambulance, or even used intra-operatively or for time-lapse imaging of the brain.

In the meantime, more research is necessary to reduce the non-linearity of the inverse problem, especially by incorporating prior information on the skull characteristics. The experiments performed on this chapter with regards to the starting model demonstrate it is necessary to have some knowledge of the properties of the skull prior or during the inversion, as this is essential to recover the structure of the brain.

I have suggested and implemented two methods, based on the use of reflection data, to address this. Although these methods show some success, also in 3D, more strategies need to be investigated to better recover the bottom part of the brain, in which the skull is poorly estimated due to the poor coverage in this region. This may include improving the proposed methods, applying constrains to the objective functional that take into account the always similar characteristics of a human skull or the use of machine learning techniques to build a guess starting model from a large database – for instance, depending on the characteristics of the patient.

The conclusions drawn from this chapter are now being proofed in the laboratory by performing *ex vivo* experiments on human skulls. This will most likely raise technical-related issues as well as more fundamental problems that would need to be addressed with *in silico* simulations. One of this will be to incorporate the finite size of the transducers in the inversion, as this is typically done by considering point sources: this is valid in geophysics given the emitted frequencies and the dimensions of the sources, but it may lead to inaccuracies at the medical scale, in which the size of the transducers is comparable to the dominant wavelength. The latter and other

aspects are now being investigated.

If successful, this technology could be applied to improve survival and recovery rates of many patients, for instance of stroke patients, but it could also be extended to image other parts of the body (e.g. the digestive system). Although potentially feasible, this will still require years of research.

7. Conclusions and next steps

The aim of this thesis was double: first, to design and implement a methodology to address elastic and viscous effects in FWI of seismic data in order to improve the quality of recovered subsurface P-wave velocity models and, second, to assess the feasibility of 3D FWI for medical imaging problems.

I have achieved these original goals by always utilising the acoustic wave equation during the inversions, demonstrating that the latter can lead to high-quality results if elastic or viscous effects are mitigated prior to a final acoustic inversion, or if they are neglected in the medical imaging context. In both cases, I have tested my hypothesis against numerical simulations first, and used the knowledge obtained from the latter to then invert more realistic or experimental datasets.

Despite the similarities between the two parts of my thesis, there are clear distinctions between the methodologies applied, the problems to be solved and the conclusions obtained from these results. Thus, I now summarise the main conclusions for the application to seismic data and ultrasound data, and discuss possible future lines of research.

7.1. Addressing elastic and viscous effects in acoustic FWI of seismic data

In the first part of this thesis I have assessed the impact of neglecting elastic and visco-acoustic effects on the recovered P-wave velocity models of the subsurface obtained from acoustic FWI. I have concluded that elastic effects are typically small in marine datasets, and can be ignored or addressed with the present strategy, but are important for land datasets. Nevertheless, I have also shown a particular case in which elastic

effects dramatically affect the quality of 3D FWI of a marine field dataset due to the presence of a chalk layer close to the sea-bottom with non-zero shear wave velocity.

On the other hand, I have shown that visco-acoustic effects typically have a larger impact on the quality of recovered models for marine datasets, and need to be addressed.

Next, I have used a strategy to mitigate elastic or visco-acoustic effects that consists in modelling acoustic and elastic or visco-acoustic data prior or after an acoustic inversion, followed by matching the elastic or visco-acoustic data to the acoustic data and application of these filters to the observed data. This data-matching method requires computation of matching filters, as well as estimates of v_p/v_s or Q.

The aim of the proposed data-matching method is to provide improved recovered P-wave velocity models compared to those obtained after acoustic FWI of observed data, and is two to three times (or just one time if using a variation of the approach in the presence of strong elastic or visco-acoustic effects) more expensive than conventional acoustic FWI, but much less so than elastic or visco-acoustic FWI in the time domain, especially for 3D datasets.

The tests carried out in chapters 3 and 4 on synthetic datasets demonstrate that this methodology is effective, and that it is able to improve the accuracy and level of detail of the recovered P-wave velocity models. I have also studied its limitations in the presence of noise or strong elastic or viscous effects, and have suggested additional workflows that can overcome some of these limitations at the same cost or even at a lower cost than the original workflows.

An important conclusion obtained from numerical experiments is that just approximate and smooth models of v_p/v_s and Q that are more accurate at the sea-bottom are required to mitigate most of the elastic and visco-acoustic effects, respectively. This is because the largest contrasts typically take place at the sea-bottom, and using an estimate of these models and the proposed data-matching method is generally better than not taking any action.

These conclusions have been then validated on field datasets, both to address elastic and visco-acoustic effects for a 3D and a 2D marine datasets, respectively.

In the former, a smooth model for v_p/v_s has been estimated using the starting P-

wave velocity model and taking into account the regional geology, which contemplates the presence of a chalk layer. Whereas 3D FWI of the observed data leads to inaccurate results, in which the chalk layered is dramatically thinned due to neglecting elastic effects, the 3D model obtained after 3D FWI of the matched data preserves the chalk layer, and introduces some additional useful information of the layers above the chalk. Inversion of this 3D dataset has been possible because the inversion was acoustic, but it would have been unaffordable to perform elastic FWI due to the large computational burden, both in terms of runtime and memory requirements. However, the results for this dataset are still preliminary results, and further research on the v_p/v_s model and the matching parameters is necessary, as well as migration of the data with the recovered velocity model for quality assessment.

For the second field dataset, an approximate 2D model of Q has been estimated from near-offset data and using a method that is based on Gabor transforms of the data. Subsequent application of the data-matching workflow results in a more detailed subsurface model, in which the continuity of the layers is improved, especially at the bottom of the model. The improvements have been verified by migrating the data, observing the flatness of migrated gathers and the fit between migrated stacked sections and the final recovered velocity models. However, the difference in compute time with respect to visco-acoustic FWI in the time-domain is not large, especially in 2D. Thus, the application of the data-matching workflow to address visco-acoustic effects in 2D can be seen as a step previous to the application to account for visco-elasticity.

There are several possible avenues for further research related with the current data-matching workflow. Initially, I recommend implementing an automatic strategy to choose the matching parameters within the data-matching code, which would result in more reliable recovered models. Then, the present method could be used to mitigate visco-elastic effects, which would require an additional estimate for the quality factor for S-waves. This could lead into research to find more reliable estimates of v_p/v_s, Q_p and Q_s, for instance by using hybrid local-global methods. The latter would be suited if inverting for just a few parameters, as only approximate estimates of the

latter are required.

Alternatively, other ways of mitigating elastic and viscous effects should be investigated. One possibility is the implementation of machine learning techniques for this purpose. For instance, the latter could be used to transform acoustic to elastic data without the need of elastic modelling, which could reduce the computational time significantly.

Nevertheless, more research on multi-parameter inversions should be simultaneously performed. If successful, the latter can be of great importance not only due to improvements on the quality of P-wave velocity models, but also on the recovery of other parameter classes that contain other useful information on the subsurface properties.

7.2. 3D FWI of ultrasound data: breast and brain imaging

The second part of this thesis has been dedicated to analyse the applicability and feasibility of FWI to 2D and 3D ultrasound datasets of the breast and the brain.

The importance of breast imaging with ultrasound is significant, as successfully imaging the breast with ultrasound waves would represent a safe, affordable and reliable alternative to conventional mammographies, which use ionising radiation.

In particular, in this thesis I have shown the first application of 3D FWI of ultrasound data to image the breast, and demonstrated that low-frequency data (below $0.5\,MHz$) is essential to recover a reliable model of the breast in which tumours can be differentiated from masses of fat. In addition, I have shown that ignoring density and absorption results in acceptable recovered models, but not accounting for the 3D nature of the breast in FWI results in a model that contains off-plane effects.

These conclusions have also been validated on an *in vivo* dataset, in which little information on the acquisition geometry, the starting model or the source wavelet was provided. After careful transducer calibration, a source wavelet has been extracted from the data with a relatively high central frequency, of $950\,kHz$, and 2D FWI has been performed ignoring data at early times, as these were contaminated with co-

herent noise. Due to the lack of low-frequency data, a starting model obtained from TOF tomography was used, which resulted in a highly-detailed speed-of-sound model of the breast in which a tumour and finer structure – possibly due to blood vessels, lymphatic or milk ducts– can be observed. The latter shows a much higher level of detail than the model obtained with conventional ray-based methods. Nonetheless, these structures could correspond to other planes, and 3D data acquisition as well as 3D data inversion are required to separate the latter.

On the other hand, to my knowledge this is the first time in which non-invasive ultrasound waves have been successfully used to image an adult human brain through the skull. Unlike conventional ultrasound tomography, in which only reflections are used, I have shown that neuro-imaging is possible if transmitted as well as reflected ultrasound data are recorded, 3D effects are considered, low-frequency transducers are used (below $1\,MHz$) and 3D FWI is performed with some prior information on the skull. Additional tests also indicate that density and absorption can be in principle ignored without compromising the quality of the recovered models.

This research is of special importance as brain imaging with ultrasound could be used to diagnose life-threatening brain injuries and diseases using non-ionising radiation, at an affordable cost, with reliable information and faster than current methods due to the portability of ultrasound systems, which could enable brain imaging at the ambulance in the near future. For instance, I have demonstrated that FWI of ultrasound can be used to differentiate between ischaemic and haemorrhagic strokes, thus enabling a rapid intervention and the administration of the right medicine to the patient. As an end result, this could improve the speed of recovery and the survival rates of stroke patients.

However, there are aspects that still require further research. I have shown the importance of having an approximate starting model of the skull in order to avoid cycle-skipping effects, and I have presented two different methods for estimating a starting model based on reflections at the skull. Although these methods show some success and are encouraging, the latter should be improved and other strategies should also be explored. One possibility is the application of machine learning algorithms to

generate starting models based on the characteristics of the patient.

Another aspect that requires further attention is the optimisation of the inversion code, perhaps using graphical processing units, in order to quickly obtain images of the brain: this is key to improve survival rates. This might not be possible nowadays, but will become a reality as computers get more powerful over the coming years.

More importantly, the hypotheses and conclusions drawn from this research on neuro-imaging should also be validated against *ex vivo* and *in vivo* experiments. This is now in process.

Appendices

A. Relation between absorption coefficients in medical ultrasound and geophysics

The loss of intensity of an ultrasound wave travelling a distance z through a medium is expressed as:

$$I(z) = I_0 e^{-\mu z} \tag{A.1}$$

in which I is the intensity of the wave after travelling through the medium, I_0 is the initial intensity and μ is the attenuation coefficient. Similarly, the pressure variation $p(z)$ within this distance is expressed as follows:

$$p(z) = p_0 e^{-\alpha z} \tag{A.2}$$

being p_0 the initial pressure of the wave and α an attenuation coefficient such that $\mu = 2\alpha$ as $I(z) = p^2(z)$.

According to (Lowrie, 2007, page 135), the amplitude variation $A(z)$ of a seismic wave travelling through a medium is:

$$A(z) = A_0 e^{-\frac{\pi z}{Q\lambda}} \tag{A.3}$$

In this equation, A_0 is the initial amplitude, Q is the quality factor and λ is the wavelength. Assuming we can record pressure data, the amplitude of the wave in equation (A.3) corresponds to the pressure of the seismic waves, $p(z)$. Equating equations (A.2) and (A.3) and taking into account that $\mu = 2\alpha$ and $v = \lambda f$ - being

v the velocity of propagation of the wave and f its frequency - we find the relation between the attenuation coefficient in ultrasound imaging and the quality factor used in seismics:

$$Q = \frac{\pi f}{\alpha v} = \frac{2\pi f}{\mu v} \tag{A.4}$$

The attenuation coefficient's values given in the literature generally refer to μ as they refer to intensity decay (Hoskins et al., 2010, pages 12-13). Moreover, this coefficients are typically expressed in terms of the frequency, in other words:

$$\mu = \mu_f f \tag{A.5}$$

in which μ_f is expressed in $dB \cdot cm^{-1} \cdot MHz^{-1}$. Hence, Q can be easily obtained from the tables in the literature as:

$$Q = \frac{2\pi f}{\mu v} = \frac{2\pi}{\mu_f v} \tag{A.6}$$

Example 1: brain

We take a constant velocity for the brain of $v = 1560\, m/s$ and an attenuation coefficient of ultrasound waves of $\mu_f = 0.6\, dB \cdot cm^{-1} \cdot MHz^{-1}$ according to (Culjat et al., 2010). Taking into account the conversion factor $\mu(dB \cdot cm^{-1}) = 4.343\mu(cm^{-1})$, we have that $\mu_f = 0.1382 \cdot 10^{-4}\, m^{-1}Hz^{-1}$. Substituting these values in equation (A.6) we have $Q = 291.4387$, i.e. $Q^{-1} = 0.0034$.

Recall usual Q values on the earth are between 0.1 (high attenuation) and 100 (low attenuation), i.e. Q^{-1} are between 0.01 (low attenuation) and 10 (high attenuation).

Example 2: skull

We take a constant velocity for the skull of $v = 2681\, m/s$ and an attenuation coefficient of ultrasound waves of $\mu_f = 8.42\, dB \cdot cm^{-1} \cdot MHz^{-1}$ according to (Culjat et al., 2010) (average of cortical and trabecular bones). Taking into account the conversion factor $\mu(dB \cdot cm^{-1}) = 4.343\mu(cm^{-1})$, we have that $\mu_f = 1.9388 \cdot 10^{-4}\, m^{-1}Hz^{-1}$. Substituting these values in equation (A.6) we have $Q = 12.0879$, i.e. $Q^{-1} = 0.0827$.

Recall usual Q values on the earth are between 0.1 (high attenuation) and 100 (low attenuation), i.e. Q^{-1} are between 0.01 (low attenuation) and 10 (high attenuation).

B. Estimating a skull starting model: method 1

The steps performed to estimate a model of speed-of-sound for the skull using method 1 as described in chapter 6 are now further elaborated.

Firstly, zero-offset data is extracted, and is plotted in Figures B.1a and B.1b for all the 512 sources around the head.

The second step consists of picking the arrival time of the reflection corresponding to the outer table of the skull. This can be done visually by inspecting Figures B.1a and B.1b, as the reflection at the skull takes place after the reflection at the skin and it has stronger amplitudes due to the larger impedance contrast at the boundary (both due to velocity and density contrasts). However, doing this manually for various datasets would mean long waiting times, which would delay the imaging of the brain and any subsequent treatment. For this reason, I do this automatically by obtaining the TOF using the Akaike information criterion (AIC) as proposed by Li et al. (2009), which is typically used by seismologists to find the onset of a seismic arrival after an earthquake (e.g. Zhang et al. (2003)). The use of this automatic picking technique is not critical to the method, and other automatic picking techniques could be used, but the former yields results that are accurate enough. After excluding the reverberations due to the injection of the source (see Figure B.1a) and defining a window that is long enough to include all the reflection at the skull, Figure B.1a shows the picked arrival times corresponding to the onset of the reflection originated at the outer table (white crosses).

Simultaneously, the same procedure can be repeated to find the onset of the reflection originated at the boundary between the inner table and the brain – this is

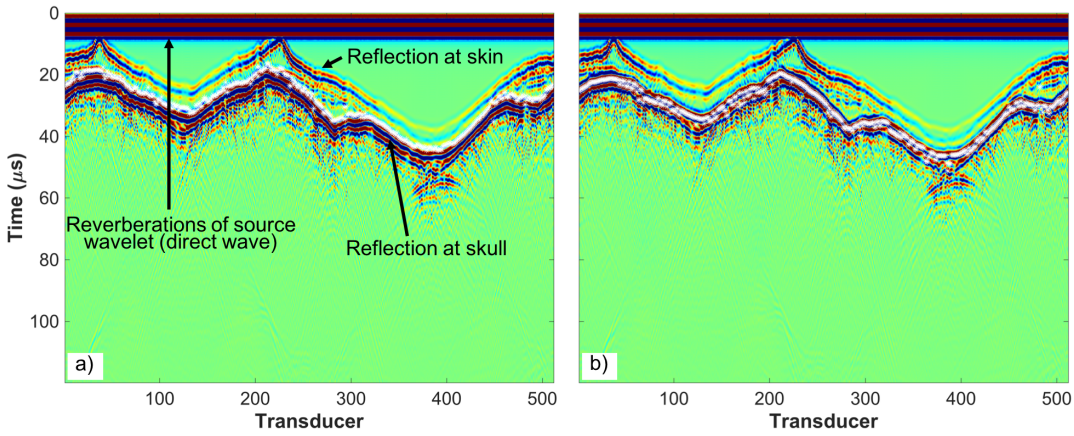


Figure B.1.: Normal incicdence (zero-offset) reflection dataset obtained for the transverse section of the head model with variable density and no attenuation in Figure 6.7, in which travel-time picks for the onset of a) the reflection at the outer table and b) at the inner table are overlaid (white crosses).

step 8 of method 1, but can be performed simultaneous with steps 2 to 7. I suggest two possibilities to find the onset in time for the inner table: 1) utilise the shape of the source wavelet and the fact that the skull always has two layers, and 2) use the onset for the outer table plus a fixed time from an average velocity of the skull and constant thickness.

For the first strategy, I observe that an approximate inner table can be obtained when automatically picking the time corresponding to the inflection point between the secondary maximum and the secondary minimum of the reflections at the skull – these time picks are plotted in Figure B.1b as white crosses. This is due to the fact that the source wavelet has three cycles and that the skull always has three interfaces, but it may be different for a different wavelet. The second strategy does not depend on the source wavelet, and is only based on an average estimate of the skull thickness (which I choose to be $4\,mm$). I test both strategies independently from now on, and conclude that the second strategy gives the best results after applying strong smoothing to the obtained starting model.

The picked arrival times are then used in step 3 to calculate the distance that a wave has travelled from the source to the skull and back to the transducer. This requires an assumption on the velocity through the medium in which the wave has travelled. For the outer table, the reflected wave only crosses water and the outer

skin, reaches the outer table and travels back to the transducer: thus, I assume an average velocity slightly above the water velocity. For the inner skull, I split the calculation in two using the onset for the outer table, and assume an average velocity of $2600\,m/s$ – this is an approximate average of the velocity for the skull– from the time corresponding to the outer table to the time corresponding to the inner table. Neglecting the assumption on the velocities and provided the arrival times are correct, the measured distances describe a circular area in which a reflection is possible, i.e. in any direction. Figure B.2a shows the resulting circles for each source in which a reflection at the outer table may have originated (the results for the inner table are not plotted as the subsequent steps are the same as for the outer table, and only the radii of the circles changes).

In step 4 of method 1, the tangent points between two neighbouring circles that are closer to the centre of the model are computed, and plotted in Figure B.2b. Provided the distance between two consecutive emitters is not large, the line that joints these two tangent points is a good approximation of the outer table within that region. This is because the surface of the skull needs to be tangent to each circle assuming normal incidence and provided the sensors are not far from the skull, and the only possibility to be tangent to two consecutive circles is that the skull is defined by the tangent points between two circles. This assumes normal incidence and that the spacing between transducers is not too large (there are 512 transducers for this example).

Nevertheless, the extracted tangent points in Figure B.2b do not define the outer table perfectly. This is in part due to the use of an automatic picking method, which may introduce some errors. For this reason, Step 5 consists of removing outliers before defining a polygon across all tangent points. This is done in three different stages.

At the first stage, I compute the distance between a single tangent point and all the others and I compute this for all extracted points, resulting in the distance matrix plotted in Figure B.3a. Smoothing is then applied to this matrix, resulting in Figure B.3b, and those points in which the difference between the original distance matrix and the smooth distance matrix is larger than a certain threshold are removed.

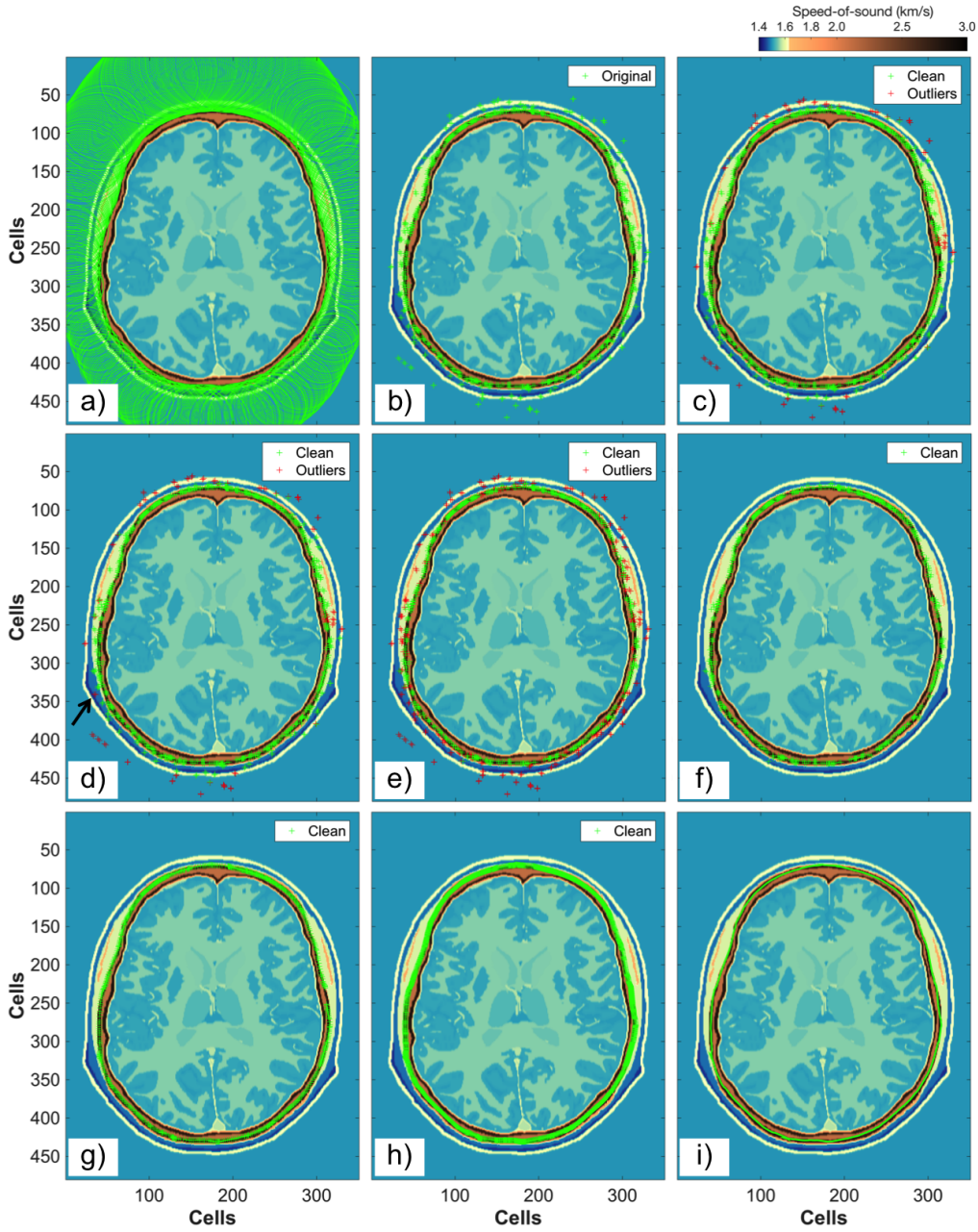


Figure B.2.: Succession of steps to find the outer table of the skull from reflection data. From left to right and top to bottom in green, a) circles with radii depending on the travel-time picks of the onset of the reflection, b) tangent points between pairs of circles closer to the centre of the model, removal of outliers after stages c) 1, d) 2 and e) 3, resulting in the clean points in f), g) clean points after smoothing and h) interpolation and i) polygon that goes through all the points (green line). The true model is overlaid for comparison purposes.

This results in the matrix distance in Figure B.3c, in which a distance of zero has been assigned to the outliers (i.e. blue row or column). This strategy is similar to

that employed by Roy et al. (2011) to de-noise the distance matrix when calibrating ultrasound data. In this case, it is useful to remove those points that are far from neighbouring points. Figure B.2c displays the resulting points after removing outliers (green crosses) as well as the outliers removed at stage 1 (red crosses).

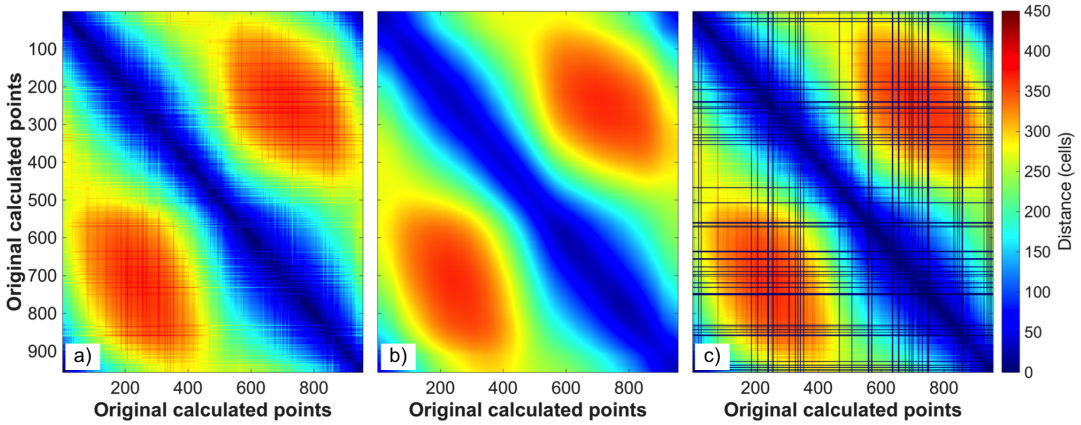


Figure B.3.: Distance matrices obtained during removal of outliers in step 1 a) before and b) after 2D smoothing. Those points in which the average distance across a column or a raw have a difference with the smooth distance matrix than a threshold are removed, and a distance value of zero is given resulting in the distance matrix in c).

At the second stage, the coordinates of all remaining points are plotted as in Figure B.4a, and 1D smoothing is applied to each coordinate, resulting in the panel in Figure B.4b. Again, those points in which the variation in either of the two coordinates is too sudden – i.e. so that the difference between the smooth and raw positions exceeds a threshold– are removed, and a distance of zero is associated to these points (see Figure B.4c). In this case, only two outliers are removed, as indicated in Figure B.2d.

At the third stage, the average and the standard deviation of the distance between neighbouring points are computed, and those pairs of points with distances larger than a certain threshold – which is based on the average and the standard deviation– are removed. This process is repeated twice. Figure B.2e shows the clean points and the outliers after the three stages to remove outliers, whereas Figure B.2f just displays the points after having removed the outliers.

Next, 1D smoothing is applied to the vectors of x and y coordinates for each point to avoid rough transitions, interpolation is applied (step 6) and a polygon that goes

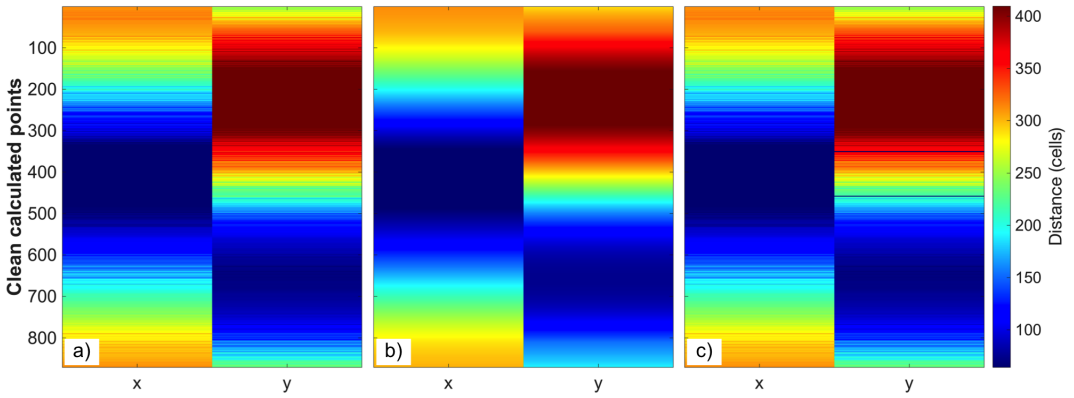


Figure B.4.: 1D plots of the spatial coordinates for the filtered tangent points after step 1 a) before and b) after 1D smoothing. Those points in which the difference between the smooth and raw coordinates is larger than a certain threshold are removed, and assigned a zero distance value. This results in the plot in panel c).

through all the points is defined (step 7). The results after the last three processing steps are shown in Figures B.2g-i, respectively. The latter polygon defines the outer table.

Steps 2 to 7 are then repeated in step 8 in order to find the inner table, using the travel-time picks in Figure B.1b and those obtained from the outer table in Figure B.1a and assuming a constant skull thickness of $4\,mm$, as previously discussed. Figures B.5a and B.5c show the computed circles with centre at the source locations for both strategies used to find the inner table, and the resulting polygons obtained after steps 2 to 7 are respectively displayed in Figures B.5b and B.5d (the same threshold values as for the outer table are used).

Finally, step 9 of method 1 consists of flooding the region between the outer and inner tables with typical speed-of-sound values for the skull. Figure B.6 illustrates how this is done for the two strategies to find the inner table, which leads to two starting models.

First, the region between the outer and inner tables (as plotted in Figures B.6a and B.6d for the two strategies to compute the inner table) is filled with zero values, and anywhere else is set to one, followed by mild smoothing to have a smooth transition from zero in the middle to one outside the skull. This is shown in Figures B.6b and B.6e. Then, a linear interpolation is performed for each cell in the models, so that

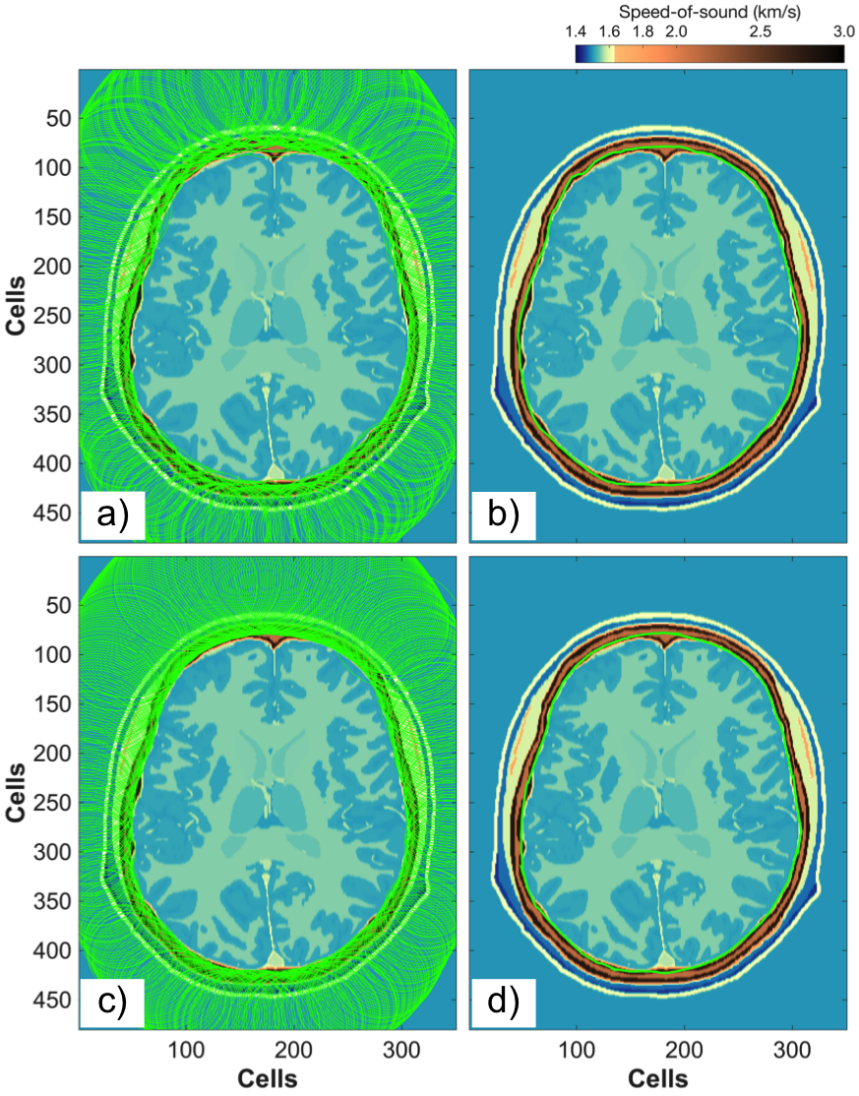


Figure B.5.: Plots for steps 3 and 7 of a) the circles corresponding to the inner table (green cicles) and b) the extracted polygon (green line) for the first strategy to pick the onset of the reflection at the inner table. The bottom panels show the corresponding plots obtained for the second strategy to pick the arrival times, obtained from the outer table plus a constant thickness of the skull of $4\,mm$. The true model of acoustic speed is overlaid.

a cell with a value of zero has a speed-of-sound of $2200\,m/s$, any cell with a value of 0.4 has a speed-of-sound of $2800\,m/s$ and anything below this value is clipped to $1500\,m/s$, which results in the models in Figures B.6c and B.6f. These are the starting models used in Figures 6.16b and 6.16b (note the average relative error is computed within this appendix with respect to the exact skull model, whereas it is computed with respect to the true transverse section of the head in chapter 6).

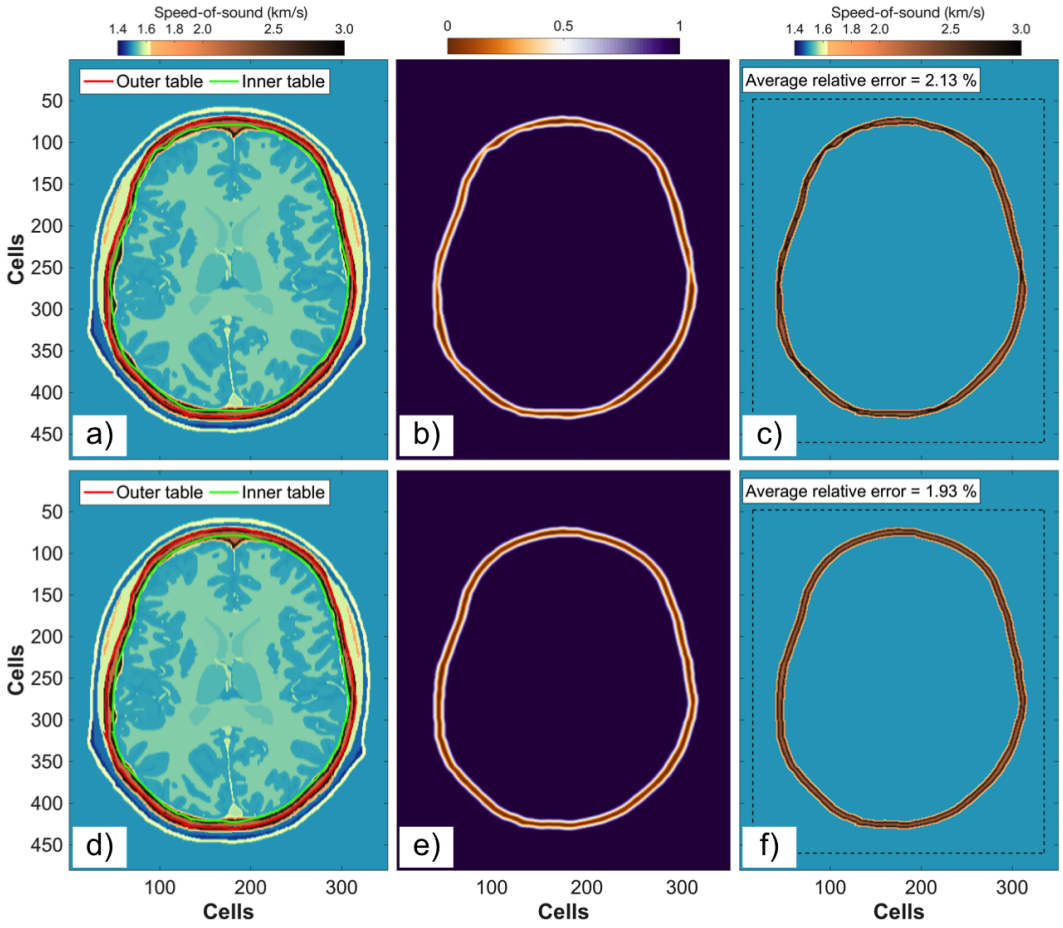


Figure B.6.: Final step to obtain a starting model for the skull using method 1. The region between the outer and inner tables in a) is filled with zeros, ones are assigned everywhere else and smoothing is applied, resulting in b). This is then converted to typical velocities of the skull in c) by assigning velocities to those regions with values between 0 and 0.4. The top panels show the results obtained for the inner table computed with the first strategy, whereas the bottom panels correspond to those obtained with the second strategy (constant skull thickness).

The current method could potentially be extended to account for short offset reflections. This would require finding tangent points between two ellipses with their focus at source and receiver positions. However, this has not yet been implemented.

Additionally, extension to 3D problems is in principle possible. This would require computing tangent planes to spheres, rather than tangent points to circles. This has also not been implemented at this stage.

C. Permissions

This appendix contains documents related to permissions for each reproduced figure in this thesis. Figures in chapters 3 and 5 are reproduced under the Creative Commons Attribution Non-Commercial No Derivatives licence.

9th April, 2018

Dear SEG Publications,

I am completing my PhD thesis at Imperial College London entitled 'Acoustic full-waveform inversion in geophysical and medical imaging'.

I seek your permission to reprint, in my thesis an extract from:

- Tsvankin, I. (2001). Seismic signatures and analysis of reflection data in anisotropic media. Elsevier Science Ltd., first edit edition. The extract to be reproduced is: Figures 1.5 and 1.6.
- Routh, P., Neelamani, R., Lu, R., Lazaratos, S., Braaksma, H., Hughes, S., Saltzer, R., Stewart, J., Naidu, K., Averill, H., Gottumukkula, V., Homonko, P., Reilly, J., and Leslie, D. (2017). Impact of high-resolution FWI in the Western Black Sea: Revealing overburden and reservoir complexity. The Leading Edge, 36(1):60–66. The extract to be reproduced is: Figures 5 and 12.
- Warner, M. and Guasch, L. (2014). Adaptive waveform inversion - FWI without cycle skipping - Theory. In SEG Technical Program Expanded Abstracts 2014, pages 1089–1093. The extract to be reproduced is: Figure 1.

I would like to include the extract in my thesis which will be added to Spiral, Imperial's institutional repository http://spiral.imperial.ac.uk/ and made available to the public under a Creative Commons Attribution-NonCommercial-NoDerivs licence.

If you are happy to grant me all the permissions requested, please return a signed copy of this letter. If you wish to grant only some of the permissions requested, please list these and then sign.

Yours sincerely,

Oscar Calderon Agudo

Permission granted for the use requested above:

I confirm that I am the copyright holder of the extract above and hereby give permission to include it in your thesis which will be made available, via the internet, for non-commercial purposes under the terms of the user licence.

Signed: [signature]

Name: Ted Bakamjian

Organisation: Society of Exploration Geophysicists

Job title: Associate Executive Director, Knowledge Management

Please follow guidelines on — seg.org/Publications/Policies-and-Permissions/Permissions

Figure C.1.: Permission granted for Figures 2.1, 2.4 and 2.5. Signature has been removed for privacy.

Order Completed

Thank you for your order.

This Agreement between Imperial College London -- Oscar Calderon Agudo ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4324851370545
License date	Apr 09, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	Seismic Waves
Licensed Content Author	Hamid N. Alsadi
Licensed Content Date	Jan 1, 2017
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Acoustic full-waveform inversion in geophysical and medical imaging
Instructor name	Joanna Morgan
Institution name	Imperial College London
Expected presentation date	Apr 2018
Portions	Fig. 2.11, Fig. 2.12
Attachment	
Requestor Location	Imperial College London Imperial College London South Kensington Campus London, SW7 2AZ United Kingdom Attn: Oscar Calderon Agudo
Billing Type	Invoice
Billing address	Imperial College London Imperial College London South Kensington Campus London, United Kingdom SW7 2AZ Attn: Oscar Calderon Agudo
Total	0.00 GBP

ORDER MORE CLOSE WINDOW

Copyright © 2018 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Figure C.2.: Permission granted for Figure 2.2.

Order Completed

Thank you for your order.

This Agreement between Oscar Calderon Agudo ("You") and Elsevier ("Elsevier") consists of your order details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License number	Reference confirmation email for license number
License date	Apr, 23 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Elsevier Books
Licensed Content Title	Wave Fields in Real Media
Licensed Content Author	
Licensed Content Date	2015
Licensed Content Pages	60
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier chapter?	No
Will you be translating?	No
Original figure numbers	Figures 2.10, 2.13
Title of your thesis/dissertation	Acoustic full-waveform inversion in geophysical and medical imaging
Publisher of new work	Imperial College London
Author of new work	Joanna Morgan
Expected completion date	Apr 2018
Estimated size (number of pages)	1
Requestor Location	Imperial College London Imperial College London South Kensington Campus London, SW7 2AZ United Kingdom Attn: Oscar Calderon Agudo
Publisher Tax ID	GB 494 6272 12
Billing Type	Invoice
Billing address	Imperial College London Imperial College London South Kensington Campus London, United Kingdom SW7 2AZ Attn: Oscar Calderon Agudo
Total	0.00 GBP

Figure C.3.: Permission granted for Figure 2.3.

Bibliography

Agudo, O. C., da Silva, N. V., Warner, M., Kalinicheva, T., and Morgan, J. (2018a). Adressing viscous effects in acoustic full-waveform inversion. *Geophysics*, (Under review).

Agudo, O. C., da Silva, N. V., Warner, M., and Morgan, J. (2016). Acoustic full-waveform inversion in an elastic world. In *SEG Technical Program Expanded Abstracts 2016*, volume 24, pages 1058–1062. Society of Exploration Geophysicists.

Agudo, O. C., da Silva, N. V., Warner, M., and Morgan, J. (2018b). Acoustic full-waveform inversion in an elastic world. *Geophysics*, 83(3):1–15.

Agudo, O. C., Guasch, L., Huthwaite, P., and Warner, M. (2018c). 3D imaging of the breast using full-waveform inversion. In *Proceedings of the International Workshop on Medical Ultrasound Tomography*, pages 99–110.

Agudo, O. C., Vieira da Silva, N., Warner, M., and Morgan, J. (2017). Addressing viscous effects in acoustic full-waveform inversion. In *79th EAGE Conference and Exhibition*.

Aki, K. and Richards, P. G. (2002). *Quantitative seismology*. University Science Books, second edi edition.

Alford, R. M., Kelly, K. R., and Boore, D. M. (1974). Accuracy of finite-difference modelling of the acoustic wave equation. *Geophysics*, 39(6):834–842.

Alkhalifah, T. (2000). An acoustic wave equation for anisotropic media. *Geophysics*, 65(4):1239–1250.

Alsadi, H. N. (2016). *Seismic hydrocarbon exploration*. Springer International Publishing.

Ames, W. F. (2014). *Numerical methods for partial differential equations*. Academic Press.

Aminzadeh, F., Brac, J., and Kunz, T. (1997). *3-D salt and overthrust models*. Society of Exploration Geophysicists.

Ammar, G. S. and Gragg, W. B. (1988). Superfast solution of real positive definite Toeplitz systems. *SIAM Journal on Matrix Analysis and Applications*, 9(1):61–76.

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.

Auld, B. A. (1990). *Acoustic fields and waves in solids*. Krieger Publishing Company.

Bai, T., Tsvankin, I., and Wu, X. (2017). Waveform inversion for attenuation estimation in anisotropic media. *Geophysics*, 82(4):WA83–WA93.

Barnes, C. and Charara, M. (2009). The domain of applicability of acoustic full-waveform inversion for marine seismic data. *Geophysics*, 74(6):WCC91–WCC103.

Belt, G. H., Felberg, R. A., Rubin, J., and Halperin, J. J. (2016). In-transit telemedicine speeds ischemic stroke treatment: preliminary results. *Stroke*, 47(9):2413–2415.

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

Bernard, S., Monteiller, V., Komatitsch, D., and Lasaygues, P. (2017). Ultrasonic computed tomography based on full-waveform inversion for bone quantitative imaging. *Physics in Medicine & Biology*, 62(17):7011–7035.

Beydoun, W. B. and Tarantola, A. (1988). First Born and Rytov approximations: modeling and inversion conditions in a canonical example. *The Journal of the Acoustical Society of America*, 83(3):1045–1055.

Bhakta, T. and Landro, M. (2013). Comparison of different rock physics models for chalk reservoirs. In *10th Biennial International Conference & Exposition*, page 7.

Blanch, J. O., Robertsson, J. O. A., and Symes, W. W. (1995). Modeling of a constant Q : Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique. *Geophysics*, 60(1):176–184.

Bojanczyk, A. W., Brent, R. P., and de Hoog, F. R. (1995). Stability analysis of a general toeplitz systems solver. *Numerical Algorithms*, 10(2):225–244.

Bording, R. P., Gersztenkorn, A., Lines, L. R., Scales, J. A., and Treitel, S. (1987). Applications of seismic travel-time tomography. *Geophysical Journal International*, 90(2):285–303.

Bourgeois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P., and Versteeg, R. (1990). Marmousi, model and data. In *The Marmousi Experience*, pages 5–16.

Brent, R. P. (1988). Old and new algorithms for Toeplitz systems. *Proceedings of SPIE, Advanced Algorithms and Architectures for Signal Processing III*, 975:2–9.

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

Brossier, R., Operto, S., and Virieux, J. (2015). Velocity model building from seismic reflection data by full-waveform inversion. *Geophysical Prospecting*, 63(2):354–367.

Bunch, J. R. (1985). Stability of methods for solving Toeplitz systems of equations. *SIAM Journal on Scientific and Statistical Computing*, 6(2):349–364.

Bunge, S. A. and Kahn, I. (2009). Cognition: an overview of neuroimaging techniques. *Encyclopedia of Neuroscience*, 2:1063–1067.

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

Butler, D. K. (2005). *Near-surface geophysics*. SEG Books.

Carcione, J. M. (2001). *Wave fields in real media: wave propagation in anistropic, anelastic and porous media*, volume 31. Elsevier Science Ltd.

Carcione, J. M. (2009). Theory and modeling of constant- Q P- and S-waves using fractional time derivatives. *Geophysics*, 74(1):T1–T11.

Carcione, J. M., Kosloff, D., and Kosloff, R. (1988a). Wave propagation simulation in a linear viscoacoustic medium. *Geophysical Journal International*, 93(2):393–401.

Carcione, J. M., Kosloff, D., and Kosloff, R. (1988b). Wave propagation simulation in a linear viscoelastic medium. *Geophysical Journal International*, 95(1988):597–61.

Causse, E., Mittet, R., and Ursin, B. (1999). Preconditioning of fullwaveform inversion in viscoacoustic media. *Geophysics*, 64(1):130–145.

Causse, E. and Ursin, B. (2000). Viscoacoustic reverse-time migration. *Journal of Seismic Exploration2*, 9:165–184.

Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M. (1985). A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, 50(4):705–708.

Cerveny, V. (2001). *Seismic ray theory*. Cambrdige University Press.

Chan, V. and Perlas, A. (2011). Basics of ultrasound imaging. In *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, chapter 2, pages 13–19. Springer New York, New York, NY.

Chapman, C. (2004). *Fundamentals of seismic wave propagation*. Cambridge University Press.

Chapman, C., Hobro, J., and Robertsson, J. (2010). Elastic corrections to acoustic finite-difference simulations. In *SEG Technical Program Expanded Abstracts 2010*, pages 3013–3017. Society of Exploration Geophysicists.

Chapman, C. H., Hobro, J. W. D., and Robertsson, J. O. A. (2014). Correcting an acoustic wavefield for elastic effects. *Geophysical Journal International*, 197(2):1196–1214.

Cheng, X., Jiao, K., Sun, D., and Vigh, D. (2015). A new approach of visco-acoustic waveform inversion in the time domain. In *SEG Technical Program Expanded Abstracts*, pages 1183–1187. Society of Exploration Geophysicists.

Claerbout, J. F. (1985). *Fundamentals of geophysical data processing: with applications to petroleum prospecting*. Blackwell Scientific Publications.

Clayton, R. and Engquist, B. (1977). Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the Seismological Society of America*, 67(6):1529–1540.

Correa, F., Enriquez, G., Rossello, J., Lucaya, J., Piqueras, J., Aso, C., Vazquez, E., Ortega, A., and Gallart, A. (2004). Posterior fontanelle sonography: an acoustic window into the neonatal brain. *American Journal of Neuroradiology*, 25(August):1274–1282.

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.

Culjat, M. O., Goldenberg, D., Tewari, P., and Singh, R. S. (2010). A review of tissue substitutes for ultrasound imaging. *Ultrasound in Medicine and Biology*, 36(6):861–873.

da Silva, N., Yao, G., Warner, M., Umpleby, A., and Debens, H. (2017). Global visco-acoustic full waveform inversion. In *79th EAGE Conference & Exhibition*, pages 12–15.

Dai, N. and West, G. F. (1994). Inverse Q migration. In *SEG Technical Program Expanded Abstracts 1994*, pages 1418–1421. Society of Exploration Geophysicists.

Danecek, P. and Seriani, G. (2008). An Efficient Parallel Chebyshev Pseudo-spectral Method for Large Scale 3D Seismic Forward Modelling. In *70th EAGE Conference & Exhibition*, pages 9–12.

de Hoog, F. (1987). A new algorithm for solving Toeplitz systems of equations. *Linear Algebra and its Applications*, 88-89:123–138.

Debens, H., Warner, M., and Umpleby, A. (2015a). Global anisotropic FWI. In *77th EAGE Conference & Exhibition*, pages 1–4.

Debens, H. A. (2015). *Three-dimensional anisotropic full-waveform inversion*. Phd thesis, Imperial College London.

Debens, H. A., Warner, M., Umpleby, A., and da Silva, N. V. (2015b). Global anisotropic 3D FWI. In *SEG Technical Program Expanded Abstracts 2015*, pages 1193–1197. Society of Exploration Geophysicists.

del Negro, C. (1996). Application of the Wiener filter to magnetic profiling in the volcanic environment of Mt. Etna (Italy). *Annali di Geofisica*, 39(1):67–79.

Dmowska, R. (2010). *Advances in geophysics*. Academic Press.

Donnan, G. A., Fisher, M., Macleod, M., and Davis, S. M. (2008). Stroke. *Lancet*, 371(9624):1612–23.

Doré, G. and Robbins, J. (2005). The Buzzard Field. *Geological Society, London, Petroleum Geology Conference series*, 6(1):241–252.

Ecoublet, P. E., Singh, S. C., Chapman, C. H., and Jackson, G. M. (2002). Bent-ray traveltime tomography and migration without ray tracing. *Geophysical Journal International*, 149(3):633–645.

Ecury-Goossen, G. M., Camfferman, F. A., Leijser, L. M., Govaert, P., and Dudink, J. (2015). State of the art cranial ultrasound imaging in neonates. *Journal of Visualized Experiments*, (96):1–10.

Evans, D. H., Jensen, J. A., and Nielsen, M. B. (2011). Ultrasonic colour Doppler imaging. *Interface Focus*, 1(4):490–502.

Feigin, V. L., Norrving, B., and Mensah, G. A. (2017). Global burden of stroke. *Circulation Research*, 120(3):439–448.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., and Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136(5):E359–E386.

Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J. W., Comber, H., Forman, D., and Bray, F. (2013). Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *European Journal of Cancer*, 49(6):1374–1403.

Fincke, J. R., Feigin, M., Prieto, G. A., Zhang, X., and Anthony, B. (2016). Towards ultrasound travel time tomography for quantifying human limb geometry and material properties. In Duric, N. and Heyde, B., editors, *Proceedings SPIE*, number April, pages 1–11.

Francis, J. G. F. (1961). The QR transformation: A unitary analogue to the LR transformation–Part 1. *The Computer Journal*, 4(3):265–271.

Francis, J. G. F. (1962). The QR transformation–Part 2. *The Computer Journal*, 4(4):332–345.

Gardner, G. H. F. (1974). Formation velocity and density: the diagnostic basics for stratigraphic traps. *Geophysics*, 39(6):770–780.

Gennisson, J. L., Deffieux, T., Fink, M., and Tanter, M. (2013). Ultrasound elastography: Principles and techniques. *Diagnostic and Interventional Imaging*, 94(5):487–495.

Given-Wilson, R., Layer, G., Warren, M., and Gazet, J. C. (1997). False negative mammography: Causes and consequences. *The Breast*, 6(6):361–366.

Golub, G. H. and Van Loan, C. F. (2013). *Matrix Computations (4th Edition)*. The Johns Hopkins University Press, fourth edi edition.

Graves, R. W. (1996). Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. *Bulletin of the Seismological Society of America*, 86(4):1091–1106.

Guasch, L. (2012). *3D elastic full-waveform inversion*. Phd thesis, Imperial College London.

Guasch, L. and Warner, M. (2014). Adaptive waveform inversion - FWI without cycle skipping - Applications. In *76th EAGE Conference and Exhibition*.

Guasch, L., Warner, M. R., Stekl, I., and Umpleby, A. (2010). 3D elastic wavefield inversion in the time domain. In *EAGE Conference & Exhibition*, pages 14–17.

Gubbins, D. (2004). *Time series analysis and inverse theory for geophysicists*. Cambridge University Press.

Guo, P. and McMechan, G. A. (2015). Separation of absorption and dispersion effects in Q-compensated viscoelastic RTM. In *SEG Technical Program Expanded Abstracts*, pages 3966–3971. Society of Exploration Geophysicists.

Hamilton, E. L. (1979). Vp /Vs and Poisson's ratios in marine sediments and rocks. *The Journal of the Acoustical Society of America*, 66(4):1093–1101.

Hasgall, P., Di Gennaro, F., Baumgartner, C., Neufeld, E., Gosselin, M. C., Payne, D., Klingenböck, A., and Kuster, N. (2015). IT'IS database for thermal and electromagnetic parameters of biological tissues. https://www.itis.ethz.ch/virtual-population/tissue-properties/overview/.

Hauge, P. S. (1981). Measurements of attenuation from vertical seismic profiles. *Geophysics*, 46(11):1548–1558.

Heywang-Köbrunner, S. H., Hacker, A., and Sedlacek, S. (2011). Advantages and disadvantages of mammography screening. *Breast Care*, 6(3):199–207.

Hicks, G. J. and Pratt, R. G. (2001). Reflection waveform inversion using local descent methods: Estimating attenuation and velocity over a gas-sand deposit. *Geophysics*, 66(2):598.

Higham, N. J. (2002). *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics, second edition.

Hobro, J. W. D., Chapman, C. H., and Robertsson, J. O. A. (2014). A method for correcting acoustic finite-difference amplitudes for elastic effects. *Geophysics*, 79(4):T243–T255.

Hofvind, S., Ponti, A., Patnick, J., Ascunce, N., Njor, S., Broeders, M., Giordano, L., Frigerio, A., and Tornberg, S. (2012). False-positive results in mammographic screening for breast cancer in Europe: A literature review and survey of service screening programmes. *Journal of Medical Screening*, 19(SUPPL. 1):57–66.

Hoskins, P., Martin, K., and Thrush, A. (2010). *Diagnostic ultrasound*. Cambrdige University Press, 2 edition.

Hu, J. (2004). *Modeling of ground deformation using spatial analysis and elasto-plastic finite strain theory*. PhD thesis, University of Southern California.

Huang, C.-C., Lin, Y.-H., Liu, T.-Y., Lee, P.-Y., and Wang, S.-H. (2011). Review: study of the blood coagulation by ultrasound. *Journal of Medical and Biological Engineering*, 31(2):79.

Huthwaite, P. (2012). *Quantitative imaging with mechanical waves*. PhD thesis, Imperial College London.

Huthwaite, P., Simonetti, F., and Huang, L. (2010). The different structural scales of the breast and their impact on time-of-flight and diffraction tomography. In D'hooge, J. and McAleavey, S. A., editors, *Medical Imaging 2010: Ultrasonic Imaging, Tomography, and Therapy*, volume 7629, pages 1–8.

Iacono, M. I., Neufeld, E., Akinnagbe, E., Bower, K., Wolf, J., Vogiatzis Oikonomidis, I., Sharma, D., Lloyd, B., Wilm, B. J., Wyss, M., Pruessmann,

K. P., Jakab, A., Makris, N., Cohen, E. D., Kuster, N., Kainz, W., and Angelone, L. M. (2015). MIDA: a multimodal imaging-based detailed anatomical model of the human head and neck. *PLOS ONE*, 10(4):1–35.

Ikelle, L. T. and Amundsen, L. (2005). *Introduction to Petroleum Seismology*. Society of Exploration Geophysicists.

Imtiaz, A. M. (2016). Trauma radiology: importance of computed tomography scans in acute traumatic brain injury. *International Journal of Pharmaceutical Sciences and Business Management*, 4(4):7–15.

Jeena, R. S. and Kumar, S. (2013). A comparative analysis of MRI and CT brain images for stroke diagnosis. In *International Conference in Microelectronics, Communication and Renewable Energy*, pages 1–5.

Jeffreys, H. and Jeffreys, B. S. (1972). *Methods of mathematical physics*. Cambrdige University Press, 3rd edition.

Kalinicheva, T., Warner, M., Ashley, J., and Mancini, F. (2017). Two- vs three-dimensional full-waveform inversion in a 3D world. In *SEG Technical Program Expanded Abstracts*, pages 1383–1387. Society of Exploration Geophysicists.

Karsl, H. and Bayrak, Y. (2004). Using the Wiener-Levinson algorithm to supress ground-roll. *Journal of Applied Geophysics*, 55:187–197.

Kelly, K. R., Ward, R. W., Treitel, S., and Alford, R. M. (1976). Synthetic seismograms: a finite-difference approach. *Geophysics*, 41(1):2–27.

Kelvin, L. T. W. (1875). Math. and Phys. papers. *Cambridge University Press*, 3:27.

Kjartansson, E. (1979). Constant Q -wave propagation and attenuation. *Journal of Geophysical Research*, 84(B9):4737–4748.

Kublanovskaya, V. (1962). On some algorithms for the solution of the complete eigenvalue problem. *USSR Computational Mathematics and Mathematical Physics*, 1(3):637–657.

Kurzmann, A., Przebindowska, A., Köhn, D., and Bohlen, T. (2013). Acoustic full waveform tomography in the presence of attenuation: a sensitivity analysis. *Geophysical Journal International*, 195(2):985–1000.

Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. In *Conference on Inverse Scattering: Theory and Application*, pages 206–220. Society for Industrial and Applied Mathematics, Expanded Abstracts.

Lakshmikantham, V., Sen, S. K., Venkaiah, V. C., and Sivasundaram, S. (1994). Stabilizing Trench's algorithm to invert symmetric Toeplitz matrices. *Applied Mathematics and Computation*, 60(2-3):249–263.

Langan, R. T., Lerche, I., and Cutler, R. T. (1985). Tracing of rays through heterogeneous media: An accurate and efficient procedure. *Geophysics*, 50(9):1456–1465.

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

Levinson, N. (1949). The Wiener RMS (root mean square) error criterion in filter design and prediction. In *Extrapolation, interpolation, and smooting of stationary time series with engeneering application*, chapter Appendix B, pages 129–148. M.I.T. Press and John Wiley and Sons.

Li, C., Huang, L., Duric, N., Zhang, H., and Rowe, C. (2009). An improved automatic time-of-flight picker for medical ultrasound tomography. *Ultrasonics*, 49(1):61–72.

Li, F., Zhou, H., Jiang, N., Bi, J., and Marfurt, K. J. (2015). Q estimation from reflection seismic data for hydrocarbon detection using a modified frequency shift method. *Journal of Geophysics and Engineering*, 12(4):577–586.

Li, Y., Wang, C., and Gu, H. (2017). Full-waveform inversion of elastic data with a corrected acoustic wave propagator. In *International Geophysical Conference*, pages 46–49. Society of Exploration Geophysicists and Chinese Petroleum Society.

Lindman, E. L. (1975). "Free-space" boundary conditions for the time dependent wave equation. *Journal of Computational Physics*, 18(1):66–78.

Lines, L. R. and Newrick, R. T. (2004). *Fundamentals of geophysical interpretation*. SEG Books.

Lines, L. R., Slawinski, R., and Bording, R. P. (1999). A recipe for stability of finitedifference waveequation computations. *Geophysics*, 64(3):967–969.

Littrup, P. J., Duric, N., Azevedo, S., Chambers, D., Candy, J. V., Johnson, S., Auner, G., Rather, J., and Holsapple, E. T. (2002). Computerized ultrasound risk evaluation (CURE) system: development of combined transmission and reflection ultrasound with new reconstruction algorithms for breast imaging. In Maev, R. G., editor, *Acoustical Imaging*, pages 175–182. Springer US.

Littrup, P. J., Duric, N., Li, C., Sak, M., Sandhu, G., Bergman, K., Boone, M., and Chen, D. (2018). Current challenges in breast screening and diagnosis: from molecules to peritumoral regions and radiomics the emerging imaging of whole breast stiffness. In *Proceedings of the International Workshop on Medical Ultrasound Tomography*, pages 1–14.

Liu, H.-P., Anderson, D. L., and Kanamori, H. (1976). Velocity dispersion due to anelasticity; implications for seismology and mantle composition. *Geophysical Journal International*, 47(1):41–58.

Løberg, M., Lousdal, M. L., Bretthauer, M., and Kalager, M. (2015). Benefits and harms of mammography screening. *Breast Cancer Research*, 17(1):1–12.

Lowrie, W. (2007). *Fundamentals of geophysics*. Cambrdige University Press, 2 edition.

Lupinacci, W. M. and Oliveira, S. A. M. (2015). Q factor estimation from the amplitude spectrum of the timefrequency transform of stacked reflection seismic data. *Journal of Applied Geophysics*, 114(March 2015):202–209.

Manbachi, A. and Cobbold, R. S. C. (2011). Development and Application of

Piezoelectric Materials for Ultrasound Generation and Detection. *Ultrasound*, 19(4):187–196.

Marfurt, K. J. (1984). Accuracy of finitedifference and finiteelement modeling of the scalar and elastic wave equations. *Geophysics*, 49(5):533–549.

Martin, G. S., Wiley, R., and Marfurt, K. J. (2006). Marmousi2: An elastic upgrade for Marmousi. *The Leading Edge*, 25(2):156–166.

Matthews, T. P., Wang, K., Li, C., Duric, N., and Anastasio, M. A. (2017). Regularized dual averaging image reconstruction for full-wave ultrasound computed tomography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64(5):811–825.

Maxwell, J. C. (1867). On the dynamical theory of gases. *Philosophical Transactions of the Royal Society of London*, 157:49–88.

McDonal, F. J., Angona, F. A., Mills, R. L., Sengbush, R. L., Van Nostrand, R. G., and White, J. E. (1958). Attenuation of shear and compressional waves in Pierre Shale. *Geophysics*, 23(3):421–439.

Mertz, L. (2015). Light plus sound: combination technology delivers a one-two punch to disease. *IEEE Pulse*, 6(3):28–33.

Mittet, R., Sollie, R., and Hokstad, K. (1995). Prestack depth migration with compensation for absorption and dispersion. *Geophysics*, 60(5):1485–1494.

Mora, P. (1987). Nonlinear twodimensional elastic inversion of multioffset seismic data. *Geophysics*, 52(9):1211–1228.

Moreau, F., Asdaghi, N., Modi, J., Goyal, M., and Coutts, S. B. (2013). Magnetic resonance imaging versus computed tomography in transient ischemic attack and minor stroke: the more you see the more you know. *Cerebrovascular Diseases Extra*, 3(1):130–136.

Morettin, P. A. (1984). The Levinson algorithm and its applications in time aeries analysis. *International Statistical Review*, 52(1):83.

Morgan, J., Warner, M., Bell, R., Ashley, J., Barnes, D., Little, R., Roele, K., and Jones, C. (2013). Next-generation seismic experiments: Wide-angle, multi-azimuth, three-dimensional, full-waveform inversion. *Geophysical Journal International*, 195(3):1657–1678.

Mourad, P. D. and Kargl, S. G. (2000). Fluid properties of fluid-saturated blood clots. Technical report, Applied Physics Laboratory,University of Washington.

Mufti, I. R. (1990). Large-scale three-dimensional seismic models and their interpretive significance. *Geophysics*, 55(9):1166–1182.

Müller, P., Schürmann, M., and Guck, J. (2015). The theory of diffraction tomography. http://arxiv.org/abs/1507.00466.

Murphy, I. A. N. G., Dillon, M. F., Doherty, A. N. N. O., and Mcdermott, E. W. (2007). Analysis of patients with false negative mammography and symptomatic breast carcinoma. *Journal of Surgical Oncology*, 96:457–463.

Musicus, B. (1988). Levinson and Fast Choleski Algorithms for Toeplitz and Almost Toeplitz Matrices. Technical Report 538.

Newman, P. (1973). Divergence effects in a layered earth. *Geophysics*, 38(3):481–488.

Nguyen, B. D. and McMechan, G. A. (2015). Five ways to avoid storing source wavefield snapshots in 2D elastic prestack reverse time migration. *Geophysics*, 80(1):S1–S18.

Nolet, G. (2008). *A breviary of seismic tomography*. Cambridge University Press.

O'Dowd, R. J. (1991). The WienerLevinson algorithm and illconditioned normal equations. *Geophysical Journal International*, 106(2):399–406.

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.

Oppenheim, A. V. and Schafer, R. W. (1989). *Discrete-time signal processing*. Prentice Hall.

Othman, M., Ali, M., and Farouk, R. (2014). Normal mode analysis of the nonlinear acoustic wave equation. *Walailak Journal of Science and Technology*, 11(4):341–347.

Pérez-Liva, M., Herraiz, J. L., Udías, J. M., Miller, E., Cox, B. T., and Treeby, B. E. (2017). Time domain reconstruction of sound speed and attenuation in ultrasound computed tomography using full-wave inversion. *The Journal of the Acoustical Society of America*, 141(3):1595–1604.

Pinson, L. J. W., Henstock, T. J., Dix, J. K., and Bull, J. M. (2008). Estimating quality factor and mean grain size of sediments from high-resolution marine seismic data. *Geophysics*, 73(4):G19–G28.

Pinton, G., Aubry, J. F., Fink, M., and Tanter, M. (2011). Effects of nonlinear ultrasound propagation on high intensity brain therapy. *Medical Physics*, 38(3):1207–1216.

Pinton, G., Dahl, J., Rosenzweig, S., and Trahey, G. (2009). A heterogeneous nonlinear attenuating full- wave model of ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 56(3):474–488.

Plessix, R., Stopin, A., Kuehl, H., Goh, V., and Overgaag, K. (2016). Visco-acoustic full waveform inversion. In *78th EAGE Conference and Exhibition*.

Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.

Plessix, R.-É., Baeten, G., de Maag, J. W., ten Kroode, F., and Rujie, Z. (2012). Full waveform inversion and distance separated simultaneous sweeping: a study with a land seismic data set. *Geophysical Prospecting*, 60(4):733–747.

Pratt, R., Song, Z., Williamson, P., and Warner, M. (1996). Two-dimensional velocity models from wide angle seismic data by waveform inversion. *Geophysical Journal International*, 124(December):323–340.

Pratt, R. G. (1990). Inverse theory applied to multi-source cross-hole tomography- Part 2: Elastic wave-equation method. *Geophysical Prospecting*, 38(3):311–329.

Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics*, 64(3):888–901.

Pratt, R. G., Huang, L., Duric, N., and Littrup, P. (2007). Sound-speed and attenuation imaging of breast tissue using waveform tomography of transmission ultrasound data. In Hsieh, J. and Flynn, M. J., editors, *Proceedings SPIE*, volume 6510, pages 1–12.

Pratt, R. G., Shin, C., and Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133(2):341–362.

Pratt, R. G. and Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography- Part 1: acoustic wave-equation method. *Geophysical Prospecting*, 38(3):287–310.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical recipes: the art of scientific computing*. Cambridge University Press, third edition.

Provenzano, G., Vardy, M. E., and Henstock, T. J. (2017). Pre-stack full waveform inversion of ultra-high-frequency marine seismic reflection data. *Geophysical Journal International*, 209(3):1593–1611.

Quan, Y. and Harris, J. M. (1997). Seismic attenuation tomography using the frequency shift method. *Geophysics*, 62(3):895–905.

Raikes, S. A. and White, R. E. (1984). Measurements of Earth attenuation from downhole and surface seismic recordings. *Geophysical Prospecting*, 32(5):892–919.

Raknes, E. B. and Arntsen, B. (2017). Challenges and solutions for performing 3D time-domain elastic full-waveform inversion. *The Leading Edge*, 36(1):88–93.

Raknes, E. B. and Weibull, W. (2016). Efficient 3D elastic full-waveform inversion using wavefield reconstruction methods. *Geophysics*, 81(2):R45–R55.

Ray, F. M., Pinnock, S. J., Katamish, H., and Turnbull, J. B. (2010). The Buzzard Field: anatomy of the reservoir from appraisal to production. *Petroleum Geology Conference series*, 7(1):369–386.

Robinson, E. A. and Durrani, T. S. (1986). *Geophysical signal processing*. Prentice-Hall International.

Robinson, E. A. and Treitel, S. (1980). *Geophysical signal analysis*. Prentice-Hall International.

Routh, P., Neelamani, R., Lu, R., Lazaratos, S., Braaksma, H., Hughes, S., Saltzer, R., Stewart, J., Naidu, K., Averill, H., Gottumukkula, V., Homonko, P., Reilly, J., and Leslie, D. (2017). Impact of high-resolution FWI in the Western Black Sea: Revealing overburden and reservoir complexity. *The Leading Edge*, 36(1):60–66.

Roy, O., Jovanović, I., Durić, N., Poulo, L., and Vetterli, M. (2011). Robust array calibration using time delays with application to ultrasound tomography. *Proceedings of SPIE*, 7968:1–11.

Ruiter, N. V., Hopp, T., Zapf, M., Menshikov, A., Kaiser, C., and Gemmeke, H. (2018). 3D ultrasound computer tomography for breast cancer diagnosis at KIT: an overview. In *Proceedings of the International Workshop on Medical Ultrasound Tomography*, pages 205–216.

Sams, M. and Goldberg, D. (1990). The validity of Q estimates from borehole data using spectral ratios. *Geophysics*, 55(1):97–101.

Scales, J. A. (1987). Tomographic inversion via the conjugate gradient method. *Geophysics*, 52(2):179–185.

Sears, T., Singh, S., and Barton, P. (2008). Elastic full waveform inversion of multi-component OBC seismic data. *Geophysical Prospecting*, 56(6):843–862.

Shen, X. (2010). Nearsurface velocity estimation by weighted earlyarrival waveform inversion. *SEG Technical Program Expanded Abstracts*, pages 1975–1979.

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.

Shin, C., Jang, S., and Min, D.-J. (2001). Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, 49(5):592–606.

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.

Silverton, A. T. T. (2015). *Applied 3D full-waveform inversion: increasing the resolution and depth penetration*. Ph.d. thesis, Imperial College London.

Simonetti, F. and Huang, L. (2008). From beamforming to diffraction tomography. *Journal of Applied Physics*, 103(10):1–7.

Sirgue, L., Barkved, O., Dellinger, J., Etgen, J., Albertin, U., and Kommedal, J. (2010). Thematic Set: Full waveform inversion: the next leap forward in imaging at Valhall. *First Break*, 28(1728):65–70.

Smith, G. D. (1985). *Numerical solution of partial differential equations: finite difference methods*. Clarendon Press.

Smith, S. W., Ivancevich, N. M., Lindsey, B. D., Whitman, J., Light, E., Fronheiser, M., Nicoletto, H. A., and Laskowitz, D. T. (2009). The ultrasound brain helmet: feasibility study of multiple simultaneous 3D scans of cerebral vasculature. *Ultrasound in Medicine and Biology*, 35(2):329–338.

Song, Z., Williamson, P. R., and Pratt, R. G. (1995). Frequencydomain

acousticwave modeling and inversion of crosshole data: Part IIInversion method, synthetic experiments and realdata results. *Geophysics*, 60(3):796–809.

Stork, C. (1988). *Ray trace tomographic velocity analysis of surface seismic reflection data*. Phd thesis, California Institute of Technology.

Stronge, G. (2018). *The limitations of acoustic full-waveform inversion in an elastic world*. PhD thesis, Imperial College London.

Stekl, I. and Pratt, R. G. (1998). Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators. *Geophysics*, 63(5):1779–1794.

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.

Tarantola, A. (1986). A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*, volume 120. Society for Industrial and Applied Mathematics.

Thomas, J. D. and Rubin, D. N. (1998). Tissue harmonic imaging: Why does it work? *Journal of the American Society of Echocardiography*, 11(8):803–808.

Thomsen, L. (1986). Weak elastic anisotropy. *Geophysics*, 51(10):1954–1966.

Tranquart, F., Grenier, N., Eder, V., and Pourcelot, L. (1999). Clinical use of ultrasound tissue harmonic imaging. *Ultrasound in Medicine & Biology*, 25(6):889–894.

Tsvankin, I. (2001). *Seismic signatures and analysis of reflection data in anisotropic media*. Elsevier Science Ltd., first edit edition.

Turnes, C. K., Balcan, D., and Romberg, J. (2014). Superfast Tikhonov regularization of Toeplitz systems. *IEEE Transactions on Signal Processing*, 62(15):3809–3821.

Umpleby, A., Warner, M., and Stekl, I. (2010). Time vs frequency for 3D wavefield tomography. *72nd EAGE Conference and Exhibition*, (June):14–17.

Vaseghi, S. V. (2000). Wiener Filters. In *Advanced Digital Signal Processing and Noise Reduction*, volume 9, chapter 6, pages 178 – 204. John Wiley & Sons Ltd.

Veitch, B., Rickett, J., and Hobro, J. (2012). Imaging elastic media by corrections to acoustic propagation. *SEG Technical Program Expanded Abstracts*, (6):1–5.

Venstad, J. M. (2016). Industry-scale finite-difference elastic wave modeling on graphics processing units using the out-of-core technique. *Geophysics*, 81(2):T35–T43.

Vidale, J. E. (1990). Finitedifference calculation of traveltimes in three dimensions. *Geophysics*, 55(5):521–526.

Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.

Walden, A. and White, R. (1998). Seismic wavelet estimation: a frequency domain solution to a geophysical noisy input-output problem. *IEEE Transactions on Geoscience and Remote Sensing*, 36(1):287–297.

Wang, K., Matthews, T., Anis, F., Li, C., Duric, N., and Anastasio, M. (2015). Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(3):475–493.

Wang, T. and Cheng, J. (2017). Elastic full waveform inversion based on mode decomposition: the approach and mechanism. *Geophysical Journal International*, 209(2):606–622.

Wang, Y. (2004). Q analysis on reflection seismic data. *Geophysical Research Letters*, 31(17):11–14.

Wang, Y. (2009). *Seismic inverse Q filtering*. John Wiley & Sons Ltd.

Wang, Y. (2014). Stable Q analysis on vertical seismic profiling data. *Geophysics*, 79(4):D217–D225.

Wapenaar, C. P. A. and Berkhout, A. J. (1989). *Elastic wave field extrapolation: Redatuming of single- and multi-component seismic data*. Elsevier.

Warner, M. and Guasch, L. (2014). Adaptive waveform inversion - FWI without cycle skipping - Theory. In *SEG Technical Program Expanded Abstracts 2014*, pages 1089–1093.

Warner, M., Morgan, J., Umpleby, A., Stekl, I., and Guasch, L. (2012). Which Physics for full-wavefield seismic inversion? In *EAGE Conference & Exhibition*, pages 4–7.

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C., Conroy, G., and Bertrand, A. (2013). Anisotropic 3D full-waveform inversion. *Geophysics*, 78(2):R59–R80.

Warner, M., Stekl, I., and Umpleby, A. (2008). Efficient and 3ffective 3D wavefield tomography. In *70th EAGE Conference & Exhibition*, pages 1–5.

Warren, L. M., Dance, D. R., and Young, K. (2017). Radiation risk with digital mammography in breast screening. Accessed on 17/03/2018, retrieved from https://www.gov.uk/government/publications/breast-screening-radiation-risk-with-digital-mammography/radiation-risk-with-digital-mammography-in-breast-screening.

White, P., Clement, G., and Hynynen, K. (2006). Longitudinal and shear mode ultrasound propagation in human skull bone. *Ultrasound in Medicine & Biology*, 32(7):1085–1096.

Wiener, N. (1949). *Extrapolation, interpolation, and smooting of stationary time series with engeneering application*. M.I.T. Press and John Wiley and Sons.

Williamson, P. R. (1991). A guide to the limits of resolution imposed by scattering in ray tomography. *Geophysics*, 56(2):202–207.

Xu, S., Chen, F., Lambaré, G., and Zhang, Y. (2013). Full waveform inversion of reflected seismic data. *Journal of Seismic Exploration*, 22(5):449–462.

Xue, Z., Fomel, S., and Zhu, T. (2017). Visco-acoustic full waveform inversion with Q-compensation. In *79th EAGE Conference and Exhibition*.

Xue, Z., Zhu, T., Fomel, S., and Sun, J. (2016). Q-compensated full-waveform inversion using constant-Q wave equation. In *SEG Technical Program Expanded Abstracts 2016*, volume 35, pages 1063–1068. Society of Exploration Geophysicists.

Yao, G. and Wu, D. (2017). Reflection full waveform inversion. *Science China Earth Sciences*, 60(10):1783–1794.

Yi, B. Y., Lee, G. H., Kim, H. J., Jou, H. T., Yoo, D. G., Ryu, B. J., and Lee, K. (2013). Comparison of wavelet estimation methods. *Geosciences Journal*, 17(1):55–63.

Yilmaz, O. (1987). *Seismic data processing*. Society of Exploration Geophysicists.

Ylitalo, J., Koivukangas, J., and Oksman, J. (1990). Ultrasonic reflection mode computed tomography through a skullbone. *IEEE Transactions on Biomedical Engineering*, 37(11):1059–1066.

Zani, A., Biella, G., and Proverbio, A. M. (2002). Brain Imaging Techniques : Invasiveness and Spatial and. In *The Cognitive Electrophysiology of Mind and Brain*, chapter Appendix F, pages 417–422. Elsevier Science (USA).

Zener, C. (1948). *Elasticity and anelasticity of metals*. University of Chicago Press.

Zhang, H., Thurber, C., and Rowe, C. (2003). Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings. *Bulletin of the Seismological Society of America*, 93(5):1904–1912.

Zhang, L., Rector, J. W., and Hoversten, G. M. (2005). Finite-difference modelling of wave propagation in acoustic tilted TI media. *Geophysical Prospecting*, 53(6):843–852.

Zhang, X. Y., Yang, Z. L., Lu, G. M., Yang, G. F., and Zhang, L. J. (2017). PET/MR imaging: new frontier in Alzheimer's disease and other dementias. *Frontiers in Molecular Neuroscience*, 10(November):1–12.

Zhao, J., Li, X., Yu, J., and Wang, Y. (2014). Nonlinear ultrasound simulation based on full-wave model and comparisons with KZK. *International Journal of Bioscience, Biochemistry and Bioinformatics*, 4(5):322–325.

Zhou, J., Wu, X., Teng, K. H., Xie, Y., Lefeuvre, F., Anstey, I., and Sirgue, L. (2013). FWI-guided Q tomography and Q-PSDM for imaging in the presence of complex gas clouds, a case study from offshore Malaysia. In *SEG Technical Program Expanded Abstracts 2013*, pages 4765–4769. Society of Exploration Geophysicists.

Zhu, T. (2017). Numerical simulation of seismic wave propagation in viscoelastic-anisotropic media using frequency-independent Q wave equation. *Geophysics*, 82(4):WA1–WA10.

Zhu, T. and Harris, J. M. (2013). A constant-Q time-domain wave equation using the fractional Laplacian. In *SEG Technical Program Expanded Abstracts 2013*, number 2, pages 3417–3422. Society of Exploration Geophysicists.

Zhu, T., Harris, J. M., and Biondi, B. (2014). Q -compensated reverse-time migration. *Geophysics*, 79(3):S77–S87.