Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California

---

Doctorado en Ciencias
en Ciencias de la Tierra con orientación en Geofísica Aplicada

---

Cooperative gravimetric and full waveform inversion

Tesis
para cubrir parcialmente los requisitos necesarios para obtener el grado de
Doctor en Ciencias

Presenta:

Raúl Ulices Silva Avalos

Ensenada, Baja California, México
2020

Tesis defendida por

Raúl Ulices Silva Avalos

y aprobada por el siguiente Comité

Dr. Jonas De Dios De Basabe Delgado	Dr. Mrinal K. Sen
Codirector de tesis	Codirector de tesis

Dr. Mario González Escobar

Dr. Enrique Gómez Treviño

Dra. Selene Solorza Calderón

Dr. Alejandro González Ortega
Coordinador del Posgrado
en Ciencias de la Tierra

Dra. Rufina Hernández Martínez
Directora de Estudios de Posgrado

Resumen de la tesis que presenta Raúl Ulices Silva Avalos como requisito parcial para la obtención del grado de Doctor en Ciencias en Ciencias de la Tierra con orientación en Geofísica Aplicada.

Inversión cooperativa gravimétrica y de forma de onda completa

Resumen aprobado por:

Dr. Jonas De Dios De Basabe Delgado
Director de tesis

En los últimos años ha habido un creciente interés en la inversión de un conjunto de múltiples datos geofísicos para obtener un modelo subterráneo consistente con las mediciones para fines de exploración y explotación de recursos naturales. Desarrollamos un esquema cooperativo basado en la inversión de forma de onda completa (FWI) y las relaciones petrofísicas que minimizan el desajuste entre los datos observados y sintéticos medidos en la superficie en estaciones gravimétricas y sismogramas. Esta combinación particular está motivada por el hecho de que la resolución horizontal de un modelo puede resolverse mediante inversión de gravedad, mientras que la resolución vertical puede estimarse mejor a partir de los datos sísmicos. Este algoritmo utiliza el método de estado adjunto para el cálculo del gradiente necesario para FWI y utiliza un método de mínimos cuadrados de gradiente conjugado restringido para la inversión gravimétrica sujeto a las discrepancias entre los modelos de densidad y velocidad utilizando relaciones petrofísicas entre estas propiedades. El algoritmo propuesto ajusta los datos sísmicos en medios heterogéneos 2D acústicos y elásticos. Para medios acústicos, se exploran ejemplos en 3D conjuntamente con inversión gravimétrica. Probamos nuestro algoritmo en varios modelos sintéticos basados en diferentes estructuras geológicas. En todos los ejemplos, pudimos ajustar los datos y lograr la convergencia iterativa, recuperando la interfaz entre las capas y la parte superior y la forma de los cuerpos de mayor velocidad. Lo comparamos con la inversión separada y la inversión conjunta convencional. Los ejemplos numéricos demuestran que el método propuesto se puede utilizar para combinar con éxito conjuntos de datos gravimétricos y sísmicos para obtener un modelo subsuperficial consistente sin incurrir en el costo computacional de los métodos tradicionales de inversión conjunta.

Palabras clave: Inversión de forma de onda, Propagación de onda, FWI, Inversión Gravimétrica

Abstract of the thesis presented by Raúl Ulices Silva Avalos as a partial requirement to obtain the Doctor of Science degree in Earth Science with orientation in Applied Geophysics.

Cooperative gravimetric and full waveform inversion

Abstract approved by:

Dr. Jonas De Dios De Basabe Delgado
Thesis Director

There has been an increasing interest in recent years in the inversion of multiple geophysical data sets to obtain a consistent subsurface model for exploration and exploitation purposes. We develop a cooperative scheme based on Full Waveform Inversion (FWI) and petrophysical relations that minimizes the misfit between the observed and synthetic data measured at the surface in gravimetric stations and seismograms. This particular combination is motivated by the fact that the horizontal resolution of a model can be resolved by gravity inversion while the vertical resolution can be better estimated from the seismic data. This algorithm uses the adjoint-state method for the computation of the gradient needed for FWI and uses a constrained Conjugate Gradient Least Squares method for gravimetric inversion subject to the discrepancies between the density and the velocity models using petrophysical relationships between these properties. The proposed algorithm solves for seismic data in acoustic and elastic 2D heterogeneous media. For acoustic media, 3D examples are explored jointly with gravimetric inversion. We tested our algorithm on several synthetic models based on different geological structure. In all the examples, we were able to fit the data and achieve iterative convergence, recovering the interface between layers and the top and shape of the higher-velocity bodies. We compare with separated inversion and conventional joint inversion. The numerical examples demonstrate that the proposed method can be used to successfully combine gravimetric and seismic data sets to obtain a consistent subsurface model without incurring the computational cost of traditional joint-inversion methods.

Keywords: Full Waveform Inversion, Wave Propagation, Gravimetric Inversion

Dedications

A mi padre Raúl Silva Rivera, a mi madre Jacoba Avalos Cárdenas y a mis hermanos Madai, Daniel y Luz Elena por apoyarme siempre y en todo momento, siempre tendrán mi amor y mi respeto. Las palabras me quedan cortas para expresar mi eterna gratitud, amo a mi familia.

Al amor de mi vida Hilda Yolanda Fernández Parga por su muy apreciado apoyo sentimental y moral a pesar de la distancia. Gracias por estar a mi lado y por enseñarme a ser feliz. Te amo Yolanda.

A Dios padre por conducir mi camino durante toda mi vida y por concederme el conocimiento, la salud y las fuerzas necesarias para triunfar hasta donde he llegado.

En memoria de Cruz Silva Torres, Juan Pablo Avalos y María Esther Cárdenas.

Acknowledgements

A mi director de tesis, Dr. Jonás de Dios de Basabe Delgado por sus invaluables enseñanzas, conocimientos, orientaciones y por su excelente atención a lo largo de este trabajo de Doctorado y el de Maestría. Gracias por su paciencia con mis problemas constantes en el Cluster. Estaré siempre agradecido con usted.

I would like to thank Dr. Mrinal K. Sen for giving me the opportunity to work under his tutelage, especially at my research stay at The University of Texas At Austin; thank you for the teachings in Inverse Theory and Full Waveform Inversion.

A mis sinodales de tesis Dr. Mario González Escobar, Dr. Enrique Gómez Treviño y Dra. Selene Solorza Calderón por sus valiosos comentarios y por toda la disposición durante los procesos de avance, la redacción del artículo y la revisión de tesis.

Al Centro de Investigación Científica y de Educación Superior de Ensenada por abrirme las puertas para realizar mis estudios de Maestría y Doctorado, en especial al posgrado y profesores de Ciencias de la Tierra. Al CeMIE-Geo por permitirme el uso del cluster Lamb. Al CONACyT por brindarme el apoyo económico para realizar mis estudios de Maestría y Doctorado.

A mis amigos de Zacatecas David, Becerra, Cuco, Daniel, Francisco, Edgar, Alejandro, Genaro y Omar por el apoyo y por mantener una fuerte amistad a lo largo de mi ausencia. A toda mi familia y amigos del Obraje Pinos Zacatecas por todas sus palabras de apoyo y estímulo en mi estancia en Ensenada, nunca olvido mi lugar de origen.

A mis amigos y compañeros, Adrián, Paúl, Fernando, Elí, Eliana, Rafa, Jaziel, Pedro Alejandro y todos aquellos amigos con los que compartí buenos momentos en taquizas y reuniones en Ensenada. Finalmente gracias al grupo de trabajo de *Rock Physics* por sus valiosos consejos y enseñanzas a lo largo de nuestras reuniones.

Table of contents

Page

Spanish abstract ii

English abstract iii

Dedications iv

Acknowledgements v

List of figures viii

List of tables xiv

Chapter 1 Introduction

1.1 Hypothesis 3
1.2 Objectives 4
1.3 Work plan and material 4
1.4 Thesis organization 5

Chapter 2 Forward Problem

2.1 Newton’s law of universal gravitation 7
  2.1.1 Gravimetric forward modeling 8
2.2 Elastodynamics 10
  2.2.1 Wave equation for an isotropic medium 11
  2.2.2 Acoustic wave equation 13
  2.2.3 Forward modeling for acoustic media 14
    2.2.3.1 Wave propagation example 17
  2.2.4 Forward modeling for elastic media 19

Chapter 3 Inverse Problem

3.1 Inverse theory 22
3.2 Gravimetric inversion 25
  3.2.1 Gauss-Newton optimization 25
  3.2.2 Conjugate Gradient Least Squares 25
3.3 Acoustic Full Waveform Inversion 26
  3.3.1 Adjoint method for acoustic media 27
  3.3.2 Gradient construction 29
    3.3.2.1 Acoustic banana-doughnut kernel 31
    3.3.2.2 Example: 2D Marmousi model gradient 35
    3.3.2.3 Gradient construction: Receiver analysis 37
    3.3.2.4 Gradient construction: Source analysis 41
    3.3.2.5 Gradient construction: Frequency analysis 42
  3.3.3 Gradient-based optimization 44
  3.3.4 AFWI Example: Marmousi model 46
3.4 Elastic Full Waveform Inversion 49
  3.4.1 Adjoint method for elastic media 49
  3.4.2 Elastic banana-doughnut kernel 51

Table of contents (continuation)

3.5 Conventional joint inversion 52
3.6 Cooperative inversion 55

Chapter 4 Results: Acoustic media

4.1 Test layered Model 59
  4.1.1 Comparison with respect to independent inversions 63
4.2 Texas-shape model I 66
4.3 Texas-shaped model II 72
4.4 Texas-shaped model III with noisy data 75
4.5 SEG SEAM phase I model 76
4.6 CICESE-shape model with an over-trust fault 83
4.7 Baja-shape model with domino faulting 87
4.8 3D Acoustic FWI: 3D CICESE-BAJA model 92
4.9 Texas-shaped 3D Model. Cooperative inversion 99

Chapter 5 Results: Elastic media

5.1 EFWI: Marmousi model 105
5.2 EFWI: Texas-shape model 108
5.3 Cooperative Inversion: Marmousi model in elastic media 109
5.4 Cooperative Inversion: Texas-shape model I in elastic media 112

Chapter 6 Conclusions

References 118

List of figures

Figure Page

1 Observation vector **r** and position vector **r′** for each differential volume element $d\vec{r}'$ for a continuous of density ρ in cartesian coordinates system. 8

2 Rectangular prism of constant density ρ. The coordinates x_i, y_i, z_i are the corners of the prism for $i = 1, 2$ 9

3 Visual representation of a standard grid discretization for a 2D acoustic media for the pressure field P. 16

4 Ricker wavelet function for peak frequencies 2,5,10,15 and 25 Hz. The function is centered at $t_0 = 0$ 17

5 Forward pressure field $P(x,z;t)$ at different times t for homogenous acoustic media. The white star represents the source position. 18

6 Visual representation of a staggered grid discretization for a 2D elastic media in terms of displacements (u_x and u_z) stresses (τ_{xx}, τ_{zz} and τ_{xz}). . . 21

7 Illustration of the concept of forward and the inverse problems. 22

8 L_p norm for some values of p corresponding to the fit of a straight line $y = ax + b$. 24

9 Forward pressure and backward adjoint wave field for gradient computation were each field is propagated in the opposite direction and accesed in different times. 30

10 Ricker wavelet for the regular wavefield (blue) and adjoint wave field (red) with peak frequency of 15 Hz used for the contruction of the acoustic banana-doughnut kernel. 32

11 3D wave propagation for regular wavefield (left) and adjoint wave field (right) and construction of the gradient at times 0,0.4,0.8,1.2, 1.6 and 2.0 seconds for an homogenous media. 34

12 Modified Marmousi velocity model (left) and starting velocity model (right). The velocity was shortened to 3500 m/s. 35

13 Synthetic seismic data acquisition for the Marmousi model at shots number 20, 40, 60, 80 and 100 corresponding to 200 receivers equally spaced at the surface. 37

14 Adjoint wave-field (left), pressure field (center) and the velocity gradient (right) construction for $t = [0.0, 0.5, 1.0, 1.5, 1.7, 2.0]$ seconds. The edges of the Marmousi model are displayed in the gradient for illustration purposes. The white start represents the source positions and the downsided black-triangle the receiver. 39

15 Velocity gradient for a source located at the origin for several receivers. The white start represents the source positions and the downsided black-triangle the several receivers used. For the bottom figure on the right the whole stream of receivers is used for the gradient. 40

List of figures (continuation)

Figure Page

16 Velocity gradient for a several source locations $s = [0, 500, 1000, 1500, 2000]$ using the whole stream of receivers (200). The white start represents the different source positions. 41

17 Seismic accuracy in terms of frequency. Taken from Yao & Wu (2017). . . . 42

18 Velocity gradient for frequencies $f = [5, 8, 10, 15, 25, 50]$ Hz. 43

19 Arbitrary cost function in terms of steps (α) for the ilustration of local and global minima for $V = V_0 + \alpha_i \partial V$. 44

20 Step line search method using 3 test points (color green) and a parabolic fit $f(x) = ax^2 + bx + c$ (color blue). The red point correspond to the minimum desired and the violet point is the actual value found. 45

21 Bad test points choosen for step line search method using quadratic fit using 3 points (color green). The red point correspond to the minimum desired, the violet point is the value found and the cyan point is the actual value in the function (blue line). 46

22 Velocity Marmousi model after some iterations of FWI. The true velocity model is at the bottom-right. 47

23 Seismogram comparisson for starting (red) and final (blue) synthetic data with respect to the observed data corresponding to a single source and a single receiver for the Marmousi model FWI example. 48

24 Objective funtion (cost, misfit) reduction for 228 iterations of FWI for the Marmousi model example. 48

25 Source - receiver geometry for the computation of the elastic kernels. Taken from Tromp et.al 2005. 51

26 Regular displacement u_x and adjoint displacement $u_x^\dagger$ wave propagation for 52 seconds of recording time for the construction of the p-wave velocity kernel. 52

27 Visual representation of our iterative inversion scheme for gravity and seismic data. From a starting velocity model, we perform FWI to update the velocity model, then, using Gardner’s density-velocity relationship, we perform constrained gravimetric inversion to update the density model, finally, using Gardner’s velocity-density relationship, a velocity model is obtained which will be the starting model to solve FWI. This process is performed iteratively. 58

28 (a) Test layered true velocity model and (b) the starting model. (c) The velocity model after 1000 iterations of cooperative inversion. 60

List of figures (continuation)

Figure Page

29 (a) Test layered true density model and (b) the starting model. (c) The density model after 1000 iterations of cooperative inversion. 61

30 Normalized misfit for seismic (red) and gravity (blue) data in logarithmic scale. 62

31 Comparison of observed (black), initial synthetic (red) and final synthetic traces (blue) measured at 560 *m* due to a source applied at 450 *m* on the surface. 63

32 Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 1000 iterations of our scheme. 63

33 Model result after separated data inversions for conventional FWI (a) and gravimetric inversion (b). 64

34 Geophysical response after separated data inversions for conventional FWI (a) and gravimetric inversion (b). 65

35 Texas-shaped true velocity model I (a) and its smoothing set as a starting model (b). 66

36 Texas-shaped true density model I (a) and its smoothing set as a starting model (b). These models are obtained using Gardner density-velocity. . . . 67

37 Texas-shape velocity model I results for (a) conventional FWI using adjoint method, (b) petrophysical joint inversion, and (c) petrophysical cooperative inversion. 69

38 Texas-shaped density model I results for (a) conventional gravimetric inversion, (b) petrophysical joint inversion using Gauss-Newton and (c) the cooperative inversion proposed. 70

39 Normalized misfit for seismic (red) and gravity (blue) for the Texas-shaped model. 71

40 Texas-shaped true velocity model (a) and its smoothing set as a starting model (b). 73

41 Texas-shaped model results using cooperative inversion (a) and conventional FWI (b). 74

42 Normalized misfit for seismic (red) and gravity (blue) for the Texas-shaped model. 75

43 Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 1000 iterations of our scheme. 77

List of figures (continuation)

Figure Page

44 Texas-shaped true velocity model (a) and its smoothing set as a starting model (b) for an example with noise. 77

45 Velocity gradients using noise-free data (a) and noisy data (b) and the difference in percentage. Both gradients are muted at the water layer. . . . 78

46 Texas-shaped model results using conventional FWI with noise in the data. 79

47 Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 300 *m* due to a source applied at 200 *m* on the surface. 79

48 Normalized misfit for seismic (red) and gravity (blue) for the Texas-shaped model. 80

49 Modified SEG SEAM phase I velocity model and its smoothing set as a starting model (b). 81

50 Modified SEG SEAM phase I velocity model results after 300 iterations for typical FWI (a) and Join Inversion (b). 82

51 Normalized misfit for seismic (red) and gravity (blue) for the SEG model. . 83

52 CICESE-shape true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after a joint inversion progress 84

53 Normalized misfit for seismic (red) and gravity (blue) for the CICESE-shaped model. 86

54 Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 250 *m* due to a source applied at 200 *m* on the surface for CICESE-shape model. 87

55 Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 85 iterations of cooperative scheme for the CICESE-shape model. 88

56 Baja-shape true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after cooperative inversion. 89

57 Normalized misfit for seismic (red) and gravity (blue) for the Baja-shaped model. 90

58 Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 250 *m* due to a source applied at 200 *m* on the surface for Baja-shape model. 90

59 Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 85 iterations of cooperative scheme for the Baja-shape model. 91

List of figures (continuation)

Figure Page

60 3D CICESE-BAJA true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after FWI, corresponding to a constant y-distance of 350 meters. 93

61 3D CICESE-BAJA true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after FWI, corresponding to a constant of y-distance 1000 meters. 94

62 3D CICESE-BAJA true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after FWI, corresponding to a constant y-distance of 1600 meters. 95

63 3D CICESE-BAJA starting velocity gradient for distances (a) y=350, (b) y=100, and (c) y=1600 meters. 96

64 Normalized misfit for seismic (red) and gravity (blue) for the Baja-shaped model. 98

65 Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 250 *m* due to a source applied at 200 *m* on the surface for Baja-shape model. 98

66 3D velocity model for cooperative inversion, where (a) shows the four objectives and (b), (c) and (d) 2D views at constant y. (e) The layered model in the $x-z$ plane (it is constant in the y direction). (f) The starting model is a smooth version of the layered model. 100

67 3D velocity models for different $y-$ views for the true model (left) and the cooperative inversion results (right). 101

68 Normalized misfit for seismic (red) and gravity (blue) for the 3D example. . 102

69 Starting gravimetric residual (a) and final gravimetric residual (b) after cooperative inversion. Both residuals are normalized with respect to the starting residual and the starting contours are displayed in the final to illustrate the regions where the residual is reduced. 102

70 Seismic data fit for different sources and receivers for observed pressure (black dashed), starting synthetic pressure (red) and final synthetic pressure (blue) after 100 iterations of cooperative inversion. 103

71 Marmousi true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after elastic FWI (c). 106

72 Normalized misfit for seismic for the Marmousi model result after EFWI. . . 107

List of figures (continuation)

Figure Page

73 Seismogram comparisson for starting (red) and final (blue) synthetic data with the observed data corresponding to a single source and a single receiver for EFWI. 107

76 Marmousi true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after cooperative inversion in elastic media. 110

77 Marmousi final velocity model for comparisson between conventional elastic FWI (a) the cooperative inversion proposed (b) after 48 iterations. 111

78 Texas-shape final velocity model I for comparisson between conventional elastic FWI (a) the cooperative inversion proposed (b) after 48 iterations. . 112

79 Normalized misfit reduction for seismic data for cooperative inversion (red) and separated inversion (blue). 113

74 Texas-shape true velocity model II (a) and its smoothing set as a starting model (b). The final velocity model after elastic FWI. 114

75 Normalized misfit for seismic data for the Texas-shape model I after elastic FWI. 115

List of tables

Table Page

1 Stiffness tensor reduction under stress-strain symmetries. 12

2 Elastic parameter as function of the Lamè's parameters and density. 13

3 Central differences coefficients for second order derivative for accuracies of $2, 4, 6$ and 8th order with uniform grid spacing. 15

4 Central differences coefficients for first order derivatives for accuracies of $2, 4, 6$ and 8th order with uniform grid spacing corresponding to staggered grid. 20

5 Inverse theory formulation, where N is the number of data and M the number of model parameters. 23

6 Parameters used for the construction of the gradient for the Marmousi model. 36

7 Cost of the objective function and computational cost for 50 iterations for each inversions method discussed. The computational cost is normalized with respect to the cost of conventional FWI. 71

8 Parameters used for 3D CICESE-BAJA acoustic FWI. 92

9 Source and receiver position for the three seismograms selected to compare the data fit. 103

10 Parameters used for elastic FWI for the Marmousi model. 105

Chapter 1. Introduction

The challenges of geophysical exploration for the exploitation of natural resources and reservoir characterization are increasing and require the incorporation of more information collected on the surface. In order to be able to obtain more detailed subsurface models, we need to be able to combine the data from different geophysical methods. In this work, we study the acoustic or elastic properties of the subsurface layers using seismic observations and density variation through the measurement of gravity anomalies.

Full Waveform Inversion (FWI) (Tarantola, 1984, 1986) is a powerful seismic-imaging method used to estimate a velocity model (P-wave velocity model for acoustic FWI) to minimize the discrepancies between observed and synthetic seismograms using a gradient-based optimization method. FWI has become a popular method (Virieux and Operto, 2009) and it has improved significantly throughout the years, reducing the computational cost and improving the resolution of the seismic image.

FWI consists of three main steps performed iteratively for the inversion. The first step is the forward modeling to compute the synthetic data starting from an initial model and obtain the residual subtracting the observed data. Several authors have used the Finite Difference Method (FDM) (Alford *et al.*, 1974; Virieux, 1986) for waveform modeling, however, the Finite Element Method (FEM) (Marfurt, 1984) and the Spectral Element Method (SEM) (Komatitsch and Tromp, 1999) are becoming increasingly popular. The second step is to back-propagate the residual wave field (adjoint field). Then, a cross-correlation between the forward and the adjoint wavefield is computed and, adding over all times and all sources, a velocity gradient is obtained. This is the well-known adjoint method (e.g., Plessix, 2006), which reduces significantly the computational cost because only two forward modelings are required in each iteration of the inversion process. In the final step, the velocity model is updated by adding to the starting model the scaled velocity gradient using a line-search method to determine the increment. If the observed and synthetic data do not match, these steps are repeated until a stopping criterion is reached. This methodology has provided good results for stratigraphic and predominantly horizontal layered models. Despite the good results both in acoustic and elastic media, density variations have largely been ignored (Virieux and Operto, 2009). Some recent works attempted to employ global optimization methods (Datta and Sen, 2016). Our work, however, focuses on the adjoint-based FWI.

The study of gravimetric data is important for estimating density variations of the subsoil measuring the gravity or the gravity gradient tensor on the surface (Zhdanov *et al.*, 2004). Several forward modeling methods exist to compute gravity anomalies by solving Poisson's equation for the gravitational potential. Among the best-known methods is the analytical solution for prismatic bodies (e.g., Nagy, 1966; Bhattacharyya and Leu, 1977), however, solutions for other geometries are readily available (Talwani, 1965; Johnson and Litehiser, 1972; Werner, 1994; García-Abdeslem, 2005). Analogous to waveform modeling, gravity modeling has been also explored using FEM recently (Martin *et al.*, 2017). In this work, we will use the solution for uniform rectangular prisms to be congruent with the grid used on finite differences for waveform modeling.

Gravimetric inversion for density estimation is a linear problem. This method is well known for estimating structures with horizontal changes of mass distribution. The solution is straightforward using Gauss-Newton minimization (e.g., Sen and Stoffa, 2013) to obtain a density model inverting the square matrix on a single step. This method is widely used among geophysicists because of its fast convergence, however, it is computationally expensive and infeasible for large-scale problems, because a square matrix needs to be stored and inverted. One alternative to this problem is to use the Conjugate Gradient Least Squares method (CGLS). This solves the inverse problem without the need to form and store the square matrix (e.g., Sen and Stoffa, 2013).

Nowadays, the study of a region of interest for geophysical exploration or explorations requires the measurements of several geophysical method which needs to be interpreted for a better characterization. Joint inversion allows integrating different geophysical data sets into a consistent Earth - property model. Usually the strategy consists in combining all the methods into one single inverse problem. Vozoff and Jupp

(1975) were the first to perform joint inversion for different geophysical data sets, namely resistivity and magnetotelluric data. Following this, numerous methodologies and different geophysical data-inversion schemes emerged for the reduction of non-uniqueness and ambiguity in the interpretation of the Earth model. Depending on the constraints in the optimization problem, the joint inversion schemes can be classified into petrophysical, structural, or statistical. Petrophysical joint inversion is subject to empirical relationships of the model parameters (Menichetti and Guillen, 1983; Lees and VanDecar, 1991; Zeyen and Pous, 1993), structural joint inversion seeks to minimize the cross product of the gradient of each model parameter (Gallardo and Meju, 2003, 2004) and statistical joint inversion tries to solve the problem attaching to each grid cell of the model a mean point (fuzzy c-mean) depending on the number of c-means (Paasche and Tronicke, 2007; Romero and Gallardo, 2015).

One of the most examined joint geophysical interpretation is the cooperative inversion of seismic and gravimetric data since this method complements each other and both theories depend on the density of the medium. One such example is the work of Roy *et al.* (2005) which performs first-arrival travel time inversion jointly with gravity data using very fast simulated annealing. Works using seismic and gravity data were presented later by Tondi *et al.* (2009); Lin *et al.* (2012); Lin and Zhdanov (2017); Colombo and Rovetta (2018). On the other hand, Blom *et al.* (2017) stress the importance of density in geological processes and present a study for the role of density using seismic and gravimetric data, concluding that density estimation requires a strong a priori model to be able to determine it as an independent parameter.

1.1 Hypothesis

Taking into consideration the sensitivity and resolution of each method, a complex geological environment can be investigated by exploiting the advantages of both methods using a cooperative inversion scheme of seismic and gravimetric data sets. Also, the cooperative integration could fix the disadvantages that each method encounters.

We propose a novel cooperative inversion scheme using gravimetric inversion, a petrophysical relation and FWI. We perform a gravimetric inversion constraining the density model obtained using Gardner's equation from the velocity model provided after FWI. Such constraint will be strong enough to avoid shallower models due to the nature of the gravimetric potential method. We propose the use again of a petrophysical relation to convert the density model into a velocity model and apply one iteration of FWI to obtain a new velocity model that better fits the seismic data. This process will be performed iteratively until FWI converges to a solution, ensuring a data fit in seismic and gravity data. The importance of performing these steps sequentially is that it is more robust than trying to invert for velocity and density jointly, it has a lower computational cost and it allows for better control of each problem separately.

1.2 Objectives

The main objective of this work is to present the inverse problem, solve it and apply it for different geological environments to obtain velocity and density models consistent with measurements at receivers of displacement (or pressure) and gravimetric measurements on the surface for elastic or acoustic 2D/3D media. Therefore three fundamental objectives will be discussed sequentially; solve for the forward problem for each geophysical method, solve for the individual inverse problem for each method and combine both methods into a cooperative inversion algorithm to obtain a consistent Earth-property model. This thesis describes in detail the theoretical framework for forward and inverse problems and presents the synthetic results for applying cooperative inversion for different models.

1.3 Work plan and material

In order to analyze the objective previously shown, the following list of activities is presented in chronological order

1. Develop the gravimetric data forward modeling.
2. Develop the acoustic wave propagation forward modeling.
3. Develop the inversion for gravity data.
4. Develop an acoustic FWI scheme using the adjoint method.
5. Develop the join inversion scheme (Acoustic media).
6. Test the software on synthetic data.
7. Develop the elastic wave propagation forward modeling.
8. Develop an elastic FWI scheme using the adjoint method.
9. Develop the join inversion scheme (Elastic media).
10. Test the software on synthetic data.

This work plan will be supported with personal material (Laptop) for simple examples, developing from scratch functions, subroutines and toolboxes in programming languages like Fortran90 and Python (data visualization). Once the simple tasks are developed and performed successfully, the experiment will be executed on the cluster Lamb of the supercomputing lab at the Specialized Labs System of the Earth Sciences Division of CICESE on Open MP and MPI (28×20 MPI cores capacity).

1.4 Thesis organization

This is divided into six chapters, Chapter 1 is the introduction. Chapter 2 presents the forward modeling framework and is divided in two parts for gravimetric and seismic modeling. For gravity, Newton's law of universal gravitation is discussed and the forward modeling based on the response of a rectangular body is presented. For seismic wave propagation, a brief introduction on elastodynamic theory is presented and the forward modeling for elastic and acoustic media is covered.

Chapter 3 discusses the inverse problem and follows the same organization as chapter 2 for each geophysical method. Firstly, the general basis on inverse theory is presented. Then the separated inversion for each method is discussed: Conjugate Gradient Least Squares method for gravimetric inversion and adjoint method for FWI. Finally, the joint inversion and the sequential inversion frameworks are presented.

The results will be divided for acoustic and elastic media. For chapter 4 several synthetic models for separated, conventional and cooperative inversion on an acoustic media will be discussed and compared, for bi- and three-dimensional models. Chapter 5 covers the results for elastic media using similar synthetic models as acoustic media, adding their respective density and shear velocity models. For last, the conclusions for the work are discussed in chapter 6.

Chapter 2. Forward Problem

This chapter presents the theoretical framework for the geophysical methods used in this work: gravimetric and seismic. For the gravity data, the Newton's Law of gravitation is presented and solved for a parallelepiped of constant density. For seismic data, the wave equation for elastic and acoustic media are discussed, solving them using typical Finite Difference Methods.

2.1 Newton's law of universal gravitation

Newton's law of gravitation (Blakely, 1996) provides the gravitational potential ϕ at an observation point $\mathbf{r}$ due to a body on Earth with density distribution ρ (Figure 1) as

$$\phi(\mathbf{r}) = \int_{\Omega} \gamma \frac{\rho(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} \mathrm{d}V, \tag{1}$$

where $\gamma = 6.672 \times 10^{-11} \mathrm{m}^3 \mathrm{kg}^{-1} \mathrm{s}^{-2}$ is the universal constant of gravitation, $\mathbf{r}'$ is the position for each differential element of density over the volume Ω and $\|.\|$ denotes the distance between vectors. Given a scalar potential, a conservative field $\mathbf{g}$ is given by the gradient of such potential as

$$\mathbf{g}(\mathbf{r}) = \nabla \phi(\mathbf{r}), \tag{2}$$

where the field $\mathbf{g}$ is the gravity acceleration. Consider and arbitrary continuous body of density ρ (Figure 1) in Cartesian coordinates, the components of the gravity acceleration is given by

$$g_x(x,y,z) = \frac{\partial \phi}{\partial x} = \gamma \int_{\Omega} \rho \frac{x - x'}{\left[(x-x')^2 + (y-y')^2 + (z-z')^2\right]^{3/2}} \mathrm{d}x' \mathrm{d}y' \mathrm{d}z', \tag{3}$$

$$g_y(x,y,z) = \frac{\partial \phi}{\partial y} = \gamma \int_{\Omega} \rho \frac{y - y'}{\left[(x-x')^2 + (y-y')^2 + (z-z')^2\right]^{3/2}} \mathrm{d}x' \mathrm{d}y' \mathrm{d}z', \tag{4}$$

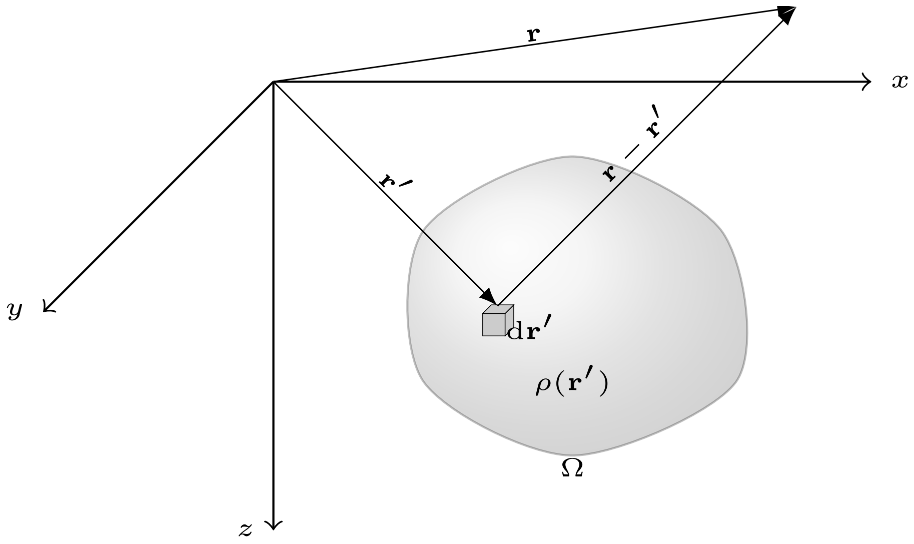


Figure 1. Observation vector **r** and position vector **r′** for each differential volume element d$\vec{r}'$ for a continuous of density ρ in cartesian coordinates system.

$$g_z(x,y,z) = \frac{\partial \phi}{\partial z} = \gamma \int_{\Omega} \rho \frac{z-z'}{\left[(x-x')^2+(y-y')^2+(z-z')^2\right]^{3/2}} \mathrm{d}x'\mathrm{d}y'\mathrm{d}z'. \tag{5}$$

In this work, only the vertical component of the gravity acceleration g_z will be considered.

2.1.1 Gravimetric forward modeling

In order to compute the gravimetric response at any observation point on surface, the discretization of the Earth model is necessary. Given that Equation 5 is valid for a continuous body of arbitrary shape and density distribution and taking advantage of the superposition theorem for Newton's law of gravitation, the Earth model is discretized as a set of rectangular prism for constant density individually (Figure 2).

For the case of a rectangular prism of constant density (Banerjee and Das Gupta, 1977), the analytic solution of Equation 5 is given by

$$g_z = \left\{ \gamma \left[z \tan^{-1}\left(\frac{xy}{z|\Delta r|}\right) - x\ln\left(y+|\Delta r|\right) - y\ln\left(x+|\Delta r|\right)\right] \Bigg|_{\Delta x'_1}^{\Delta x'_2} \Bigg|_{\Delta y'_1}^{\Delta y'_2} \Bigg|_{\Delta z'_1}^{\Delta z'_2} \right\} \rho, \tag{6}$$

where $|\Delta r| = \sqrt{x^2+y^2+z^2}$, $\Delta x'_k = x - x'_k$, $\Delta y'_k = y - y'_k$ and $\Delta z' = z - z'_k$ $k = 1,2$ for the prime coordinates (corners of the prism). This expression corresponds to the gravity measurement at the point (x,y,z) due to the prism and the part within the braces is the gravity kernel.

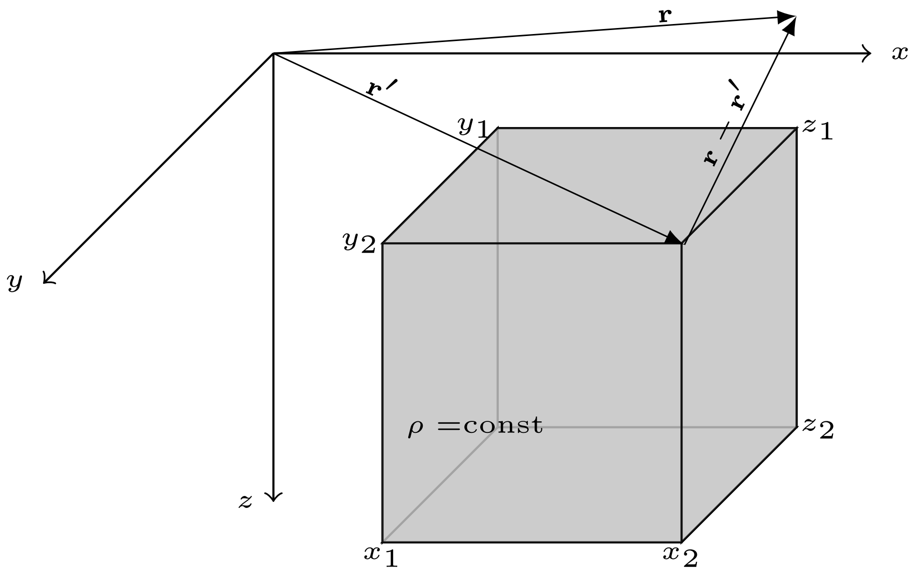


Figure 2. Rectangular prism of constant density ρ. The coordinates x_i, y_i, z_i are the corners of the prism for $i = 1,2$

Typically data acquisition is done on the surface for N_s observation points (gravimetric stations), hence

$$\mathbf{g_z} = \begin{bmatrix} g_{z_1} \\ g_{z_2} \\ \vdots \\ g_{z_{N_s}} \end{bmatrix}, \tag{7}$$

where $\mathbf{g_z} \in \mathbb{R}^{N_s}$. Considering a $\rho(x,y,z)$ parametrization of $M = n_x \times n_y \times n_z$ prisms where n_x, n_y and n_z are the number of prisms for x, y and z directions respectively, a

1D vector model $\mathbf{m}_{\boldsymbol{\rho}}$ can be arranged as

$$\mathbf{m}_{\boldsymbol{\rho}} = \begin{bmatrix} \rho_{111} \\ \rho_{211} \\ \vdots \\ \rho_{n_x 11} \\ \vdots \\ \rho_{n_x n_y n_z} \end{bmatrix}, \tag{8}$$

where $\mathbf{m}_{\boldsymbol{\rho}} \in \mathbb{R}^M$. Given this vector notation, the corresponding matrix for the kernel will be

$$\mathbf{A} = A_{i,j} = \gamma \left[z_i \tan^{-1}\left(\frac{x_i y_i}{z_i |\Delta r_i|}\right) - x_i \ln\left(y_i + |\Delta r_i|\right) - y_i \ln\left(x_i + |\Delta r_i|\right) \right] \Bigg|_{\Delta x'_{1_j}}^{\Delta x'_{2_j}} \Bigg|_{\Delta y'_{1_j}}^{\Delta y'_{2_j}} \Bigg|_{\Delta z'_{1_j}}^{\Delta z'_{2_j}}, \tag{9}$$

where $\mathbf{A} \in \mathbb{R}^{N\times M}$, thus the vector of gravity data can be represented in a matrix form as

$$\mathbf{g}_z = \mathbf{A}\mathbf{m}_\rho, \tag{10}$$

corresponding to the forward modeling for gravity data. This is a linear problem with respect to density.

2.2 Elastodynamics

An elastic body is one that is governed by the generalized Hooke's law, for small deformations and ignoring attenuation the stress and the strain are directly proportional as

$$\boldsymbol{\tau} = \mathbf{c} : \boldsymbol{\epsilon}, \tag{11}$$

where $\boldsymbol{\tau}$ is the stress tensor, $\boldsymbol{\epsilon}$ the strain tensor, $\mathbf{c}$ represents the fourth-order stiffness tensor and : is the double dot product for tensor. In index notation Equation 11 can be represented as

$$\tau_{ij} = c_{ijkl}\epsilon_{kl}. \tag{12}$$

for $i, j, k, l = 1, 2, 3$. Taking in consideration that the strain is proportional to the gradient of the displacement, $\boldsymbol{\epsilon} = \frac{1}{2}\left[\nabla u + (\nabla u)^T\right]$, (Aki and Richards, 2002), Equation 12 can be written as

$$\tau_{ij} = c_{ijkl}\frac{\partial u_k}{\partial x_l}, \tag{13}$$

where $u_k = \left\{u_x(x,t), u_y(x,t), u_z(x,t)\right\}$ is the displacement vector from an equilibrium state to a deformed state. In order to describe the behavior of the elastic body trough the time a wave equation is required. Following Aki and Richards (2002) and assuming that the elastic body is subject to Newton's second law ($F = ma$) normalized over a volume, an equation relating displacement and stresses can be obtained

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \tau_{ij}}{\partial x_j} + f_i, \tag{14}$$

where f_i represent an external force per unit volume, ρ is the density and the acceleration is written as the second derivative of the displacement u_i. Combining equations 13 and 14 the elastodynamic wave equation is obtained

$$\rho\frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j}\left[c_{ijkl}\frac{\partial u_k(x,t)}{\partial x_l}\right] = f_i, \tag{15}$$

valid for heterogeneous, elastic and anisotropic media, ignoring attenuation or viscoelastic effects.

2.2.1 Wave equation for an isotropic medium

The 4th rank stiffness tensor c_{ijkl} contains the constants which characterize the elastic properties of the solid. Since $i, j, k, l = 1, 2, 3$, there are $3^4 = 81$ components that can be reduced considering the following symmetries:

Table 1. Stiffness tensor reduction under stress-strain symmetries.

Stress-strain relationship	$\tau_{ij} = c_{ijkl}\epsilon_{kl}$	81 terms
Symmetry of the strain tensor	$\epsilon_{kl} = \epsilon_{lk}$ $\Rightarrow \tau_{ij} = c_{ijkl}\epsilon_{kl} = c_{ijlk}\epsilon_{lk} = c_{ijlk}\epsilon_{kl}$ $\Rightarrow (c_{ijkl} - c_{ijlk})\epsilon_{kl} = 0$ $\Rightarrow c_{ijkl} = c_{ijlk}$	54 terms
Symmetry of the stress tensor	$\tau_{ij} = \tau_{ji}$ $\Rightarrow c_{ijkl} = c_{jikl}$	36 terms
Strain energy restrictions	$c_{ijkl} = c_{lkij}$	21 terms

The last symmetry relationship of Table 1 is based on thermodynamic assumptions of internal energy (For more discussion see Sen (2016)). In summary, the stiffness tensor is reduced to 21 independent terms corresponding for any elastic-anisotropic material.

Anisotropy refers to the variations of elastic properties as a function of directions. There are different types of anisotropy, e.g. Vertically Transverse Isotropic (VTI), Horizontally Transverse Isotropic (HTI), Tilted Transverse Isotropic media, etc. If there is no preferred direction for the measurement and the stiffness tensor is invariant under reflection and rotation, the solid is considered to be isotropic. For the purpose of this work, only isotropic media will be discussed. In this case, the stiffness tensor can be written as

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), \tag{16}$$

where λ and μ are the Lamè parameters and δ_{ij} is the Kronecker delta function

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{17}$$

Substituting Equation 16 into 15 and reducing indexes, the elastic wave equation for isotropic media is obtained as follows

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_i}\left(\lambda\frac{\partial u_j}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right] + f_i. \tag{18}$$

The independent parameters of the stiffness tensor were reduced to 2: the Lamè's first parameter λ and the shear modulus μ. The parameter λ does not have a direct physical interpretation, for instance the bulk modulus $\kappa = \lambda + \frac{2}{3}\mu$ is commonly used. In general, these parameters can be expressed in terms of other parameters (Table 2).

Table 2. Elastic parameter as function of the Lamè's parameters and density.

p-wave velocity	$v_p = \sqrt{\frac{\lambda+2\mu}{\rho}}$
s-wave velocity	$v_s = \sqrt{\frac{\mu}{\rho}}$
Bulk modulus	$\kappa = \lambda + \frac{2}{3}\mu$
Young's modulus	$Y = \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}$
Poisson ratio	$\nu = \frac{\lambda}{2(\lambda+\mu)}$

For inverse theory purposes we can use this table to change to the model parameter desired, for example, it is common to work with v_p and v_s velocities and in some cases with the density.

2.2.2 Acoustic wave equation

The Elastodynamic wave equation can be simplified considering the wave propagation through acoustic media (fluids, melted bodies, liquid bodies) where there is absent of shear forces and therefore the modulus of rigidity μ is equal to zero. Setting $\mu = 0$, defining $P = \lambda \nabla \cdot \mathbf{u}$ and substituting in Equation 15, the following equation is obtained

$$\frac{1}{\lambda}\frac{\partial^2 P(\mathbf{x},t)}{\partial t^2} - \nabla \cdot \left[\frac{1}{\rho}\nabla P(\mathbf{x},t)\right] = f(\mathbf{x},t), \tag{19}$$

where the scalar field P is the pressure propagated in the media due to an external force f. For constant density, this expression is simplified to the well-know acoustic wave equation

$$\frac{1}{v_p^2}\frac{\partial^2 P(\mathbf{x},t)}{\partial t^2} - \nabla^2 P(\mathbf{x},t) = f(\mathbf{x},t), \tag{20}$$

with the p-wave velocity $v_p^2 = \frac{\lambda}{\rho}$. Despite the fact that this equation is valid for acoustic media, it is still being used for elastic media studies for forward modeling, FWI and

RTM since it is less expensive computationally with respect to the Elastodynamic wave equation and more importantly, the results are acceptable in many applications.

2.2.3 Forward modeling for acoustic media

In order obtain the synthetic seismograms for displacement, velocity or pressure, Equations 18 and 20 need to be solved under some initial conditions. Among the most used techniques for wave propagation, we have Finite Differences Method (FDM) for acoustic (Alford *et al.*, 1974) or elastic media (Kelly *et al.*, 1976), Finite Element Method (FEM) for acoustic (Cohen 2002) and elastic media (Cohen, 2012; De Basabe and Sen, 2007) and Finite Difference using Staggered Grids (SGFD) for elastic media. For this work, the acoustic wave propagation will be solved using FDM and the elastic wave propagation will be solved using SGFD.

Consider the following standard grid discretization for the position (x, y, z) and time t

$$\begin{aligned} t &= i_t \Delta t \qquad n = 0, 1, 2, \ldots, n_t, && (21)\\ x &= i_x \Delta x \qquad i_x = 1, 2, 3, \ldots, n_x, && (22)\\ y &= i_y \Delta y \qquad i_y = 1, 2, 3, \ldots, n_y, && (23)\\ z &= i_z \Delta z \qquad i_z = 1, 2, 3, \ldots, n_z, && (24) \end{aligned}$$

where n_x, n_y and n_z are the total number of grid points and n_t is the number of time steps, with Δ the spatial and time step respectively. Firstly, let us considerer the acoustic problem. The discrete form for the spatial and time derivatives is given by (Alford *et al.*, 1974)

$$P^{i_t+1}_{i_x,i_y,i_z} = 2P^{i_t}_{i_x,i_y,i_z} - P^{i_t-1}_{i_x,i_y,i_z} + v_p^2 \Delta t^2 \left(D_x^2 P^{i_t}_{i_x,i_y,i_z} + D_y^2 P^{i_t}_{i_x,i_y,i_z} + D_z^2 P^{i_t}_{i_x,i_y,i_z} \right), \qquad (25)$$

where D_x^2, D_y^2 and D_z^2 are the discrete operators for the second derivative. For example, the second-order discrete operator for the second derivative centered respect to t is

Table 3. Central differences coefficients for second order derivative for accuracies of $2, 4, 6$ and 8th order with uniform grid spacing.

	Approximation at $x = 0$ Central differences								
Order of accuracy	-4	-3	-2	-1	0	1	2	3	4
2				1	-2	1			
4			$-\frac{1}{12}$	$\frac{4}{3}$	$-\frac{5}{2}$	$\frac{4}{3}$	$-\frac{1}{12}$		
6		$\frac{1}{90}$	$-\frac{3}{20}$	$\frac{3}{2}$	$-\frac{49}{18}$	$\frac{3}{2}$	$-\frac{3}{20}$	$\frac{1}{90}$	
8	$-\frac{1}{560}$	$\frac{8}{315}$	$-\frac{1}{5}$	$\frac{8}{5}$	$-\frac{205}{72}$	$\frac{8}{5}$	$-\frac{1}{5}$	$\frac{8}{315}$	$-\frac{1}{560}$

given by

$$\left(\frac{\partial^2 P}{\partial t^2}\right)_n = \frac{P^{n-1}_{i,j,k} - 2P^{n}_{i,j,k} + P^{n+1}_{i,j,k}}{\Delta t^2} + \mathcal{O}(\Delta t^2) \tag{26}$$

with $\mathcal{O}(\Delta t^2)$ the truncation error. For this order, only 3 grid points in time are required for the computation of the second derivative of the pressure. Since the resolution depends on the parametrization of the velocity model in space, it is preferable the use of more grid points for x, y and z, as seen in table 3 for second derivatives for different orders of precision.

```
Data: Velocity model $v_p[n_x, n_z]$
Input : $n_t$ time steps, $n_r$ receivers, $n_s$ sources, $n_x$, $n_z$ grid nodes.
Forward modeling;
for $i_s = 1 : n_s$ do
    for $i_t = 1 : n_t$ do
        for $i_z = 1 : n_z$ do
            for $ix = 1 : n_x$ do
                $P[i_x, i_z, i_t + 1] = 2P[i_x, i_z, i_t] - P[i_x, i_z, i_t - 1]$
                        $+ v_p^2 \Delta t^2 \left(D_x^2 P^{i_t}_{i_x, i_z} + D_z^2 P^{i_t}_{i_x, i_z}\right)$
            end
        end
        for $i_r = 1 : n_r$ do
            $P_{\text{rec}}[i_s, i_r, i_t] = P[i_r, 1, i_t]$
        end
    end
end
```

Algorithm 1: Forward

The visual representation for the discretization of the acoustic wave equation is shown in Figure 3 for a 2D media example.

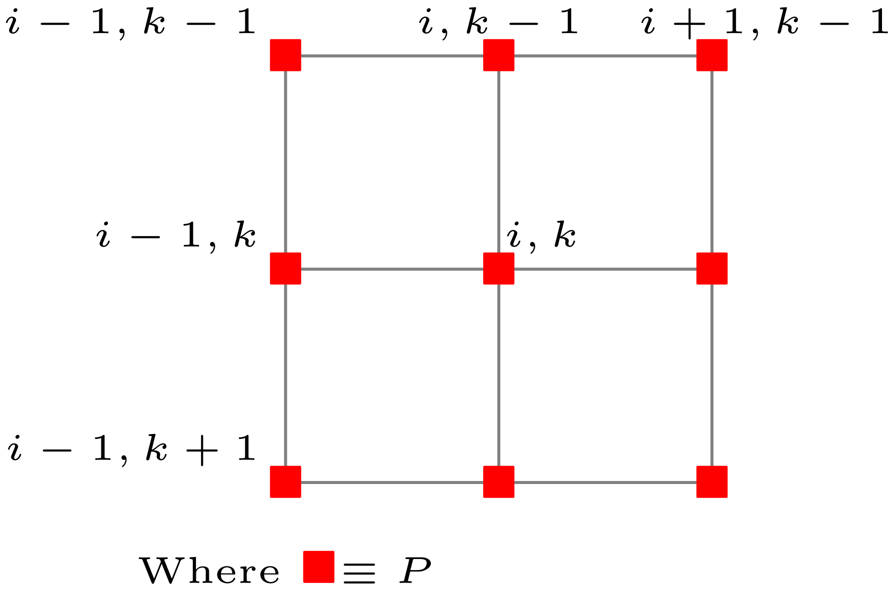


Figure 3. Visual representation of a standard grid discretization for a 2D acoustic media for the pressure field P.

Stability condition

The numerical simulation of Equation 26 involves the recursive computation of the pressure P over the time steps n_t. However, this recursive computation can present incremental error over time because of the truncation of the approximated solution or because of the machine rounding error, this issue can be treated using Von Neumann stability analysis. Based on the work of Virieux (1986), consider a plane-wave solution

$$P = e^{i\left(-\omega i_t \Delta t + i_x \Delta x + i_y \Delta y + i_z \Delta z\right)} \tag{27}$$

corresponding to a solution with Δ spacing in time and space. Substituting Equation 27 into 26 and after doing some reduction the following stability condition is obtained

$$\Delta t \leq \frac{\Delta x}{v_{\text{MAX}} \sqrt{n_D}} \left(\sum_{i=1}^{M} a_m \right)^{-1/2} \tag{28}$$

where v_{MAX} is the maximum value of the velocity model, $\sum_{i=1}^{M} a_m$ is the sum over the coefficients of the Table 3 for each order of precision excluding the centered point , and $n_D = 1, 2, 3$ for the dimension of the problem $(1D, 2D, 3D)$. This condition is very important for the inverse problem, given that it depends on the maximum velocity, therefore the velocity model obtained has to be inspected in every iteration for stability.

2.2.3.1 Wave propagation example

In order to simulate the wave propagation in time a source has to be applied at any point of the space. In this example and in all the following results for this work, a Ricker wavelet is utilized, given by

$$f(t) = 2(\pi\nu)^2 \left\{1 - 2\left[\pi\nu\,(t - t_0)\right]^2\right\} e^{-[\pi\nu(t-t_0)]^2} \tag{29}$$

where ν is the peak frequency of the pulse and t_0 is the time shift. Because of the shape of Equation 29, the Ricker wavelet is also called Mexican hat wavelet, as seen in figure 4 for $t_0 = 0.0$ and ν = [2,5,10,15,25] Hz.

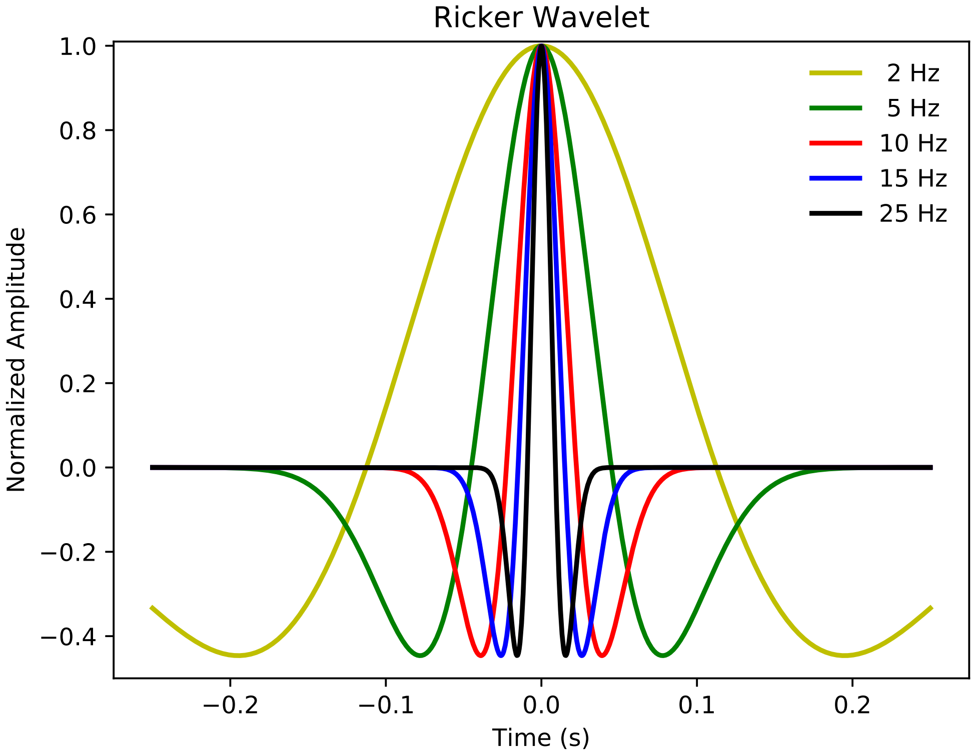


Figure 4. Ricker wavelet function for peak frequencies 2,5,10,15 and 25 Hz. The function is centered at $t_0 = 0$

For low frequencies the wavelet becomes wider and vice versa for high frequencies. Lets consider an homogeneous 2D model with velocity of 1000 km/s covering a distance $x = [0:2000]$ m and $z = [0:1000]$ m. For this example, $t_0 = 0.1$ and the peak frequency is 15 Hz. The pressure field for 2 seconds of simulation due to a source applied at (500,500) is shown in figure 5 for some snapshot.

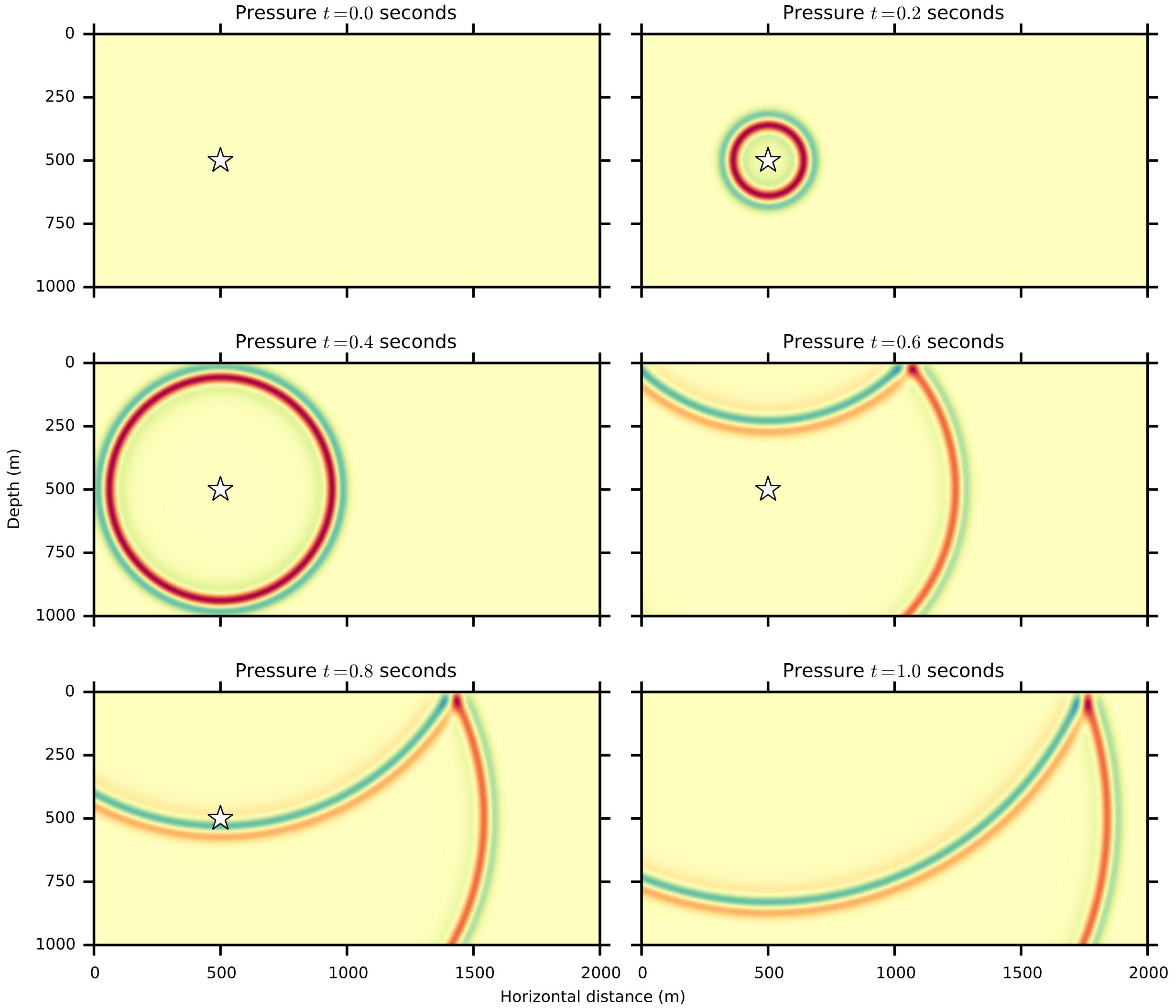


Figure 5. Forward pressure field $P(x, z; t)$ at different times t for homogenous acoustic media. The white star represents the source position.

In the following chapter, the forward modeling for heterogeneous media and 3D examples will be presented.

2.2.4 Forward modeling for elastic media

Explicitly the isotropic wave equation can be expressed as the following set of equations

$$\begin{aligned} \rho\partial_{tt}u_x &= \partial_x\tau_{xx} + \partial_y\tau_{xy} + \partial_z\tau_{xz} + f_x, && (30)\\ \rho\partial_{tt}u_y &= \partial_x\tau_{xy} + \partial_y\tau_{yy} + \partial_z\tau_{yz} + f_y, && (31)\\ \rho\partial_{tt}u_z &= \partial_x\tau_{xz} + \partial_y\tau_{yz} + \partial_z\tau_{zz} + f_z, && (32) \end{aligned}$$

for the displacements and

$$\begin{aligned} \tau_{xx} &= (\lambda + 2\mu)\partial_x u_x + \lambda(\partial_y u_y + \partial_z u_z), && (33)\\ \tau_{yy} &= (\lambda + 2\mu)\partial_y u_y + \lambda(\partial_x u_x + \partial_z u_z), && (34)\\ \tau_{zz} &= (\lambda + 2\mu)\partial_z u_z + \lambda(\partial_x u_x + \partial_y u_y), && (35)\\ \tau_{xy} &= \mu(\partial_y u_x + \partial_x u_y), && (36)\\ \tau_{xz} &= \mu(\partial_z u_x + \partial_x u_z), && (37)\\ \tau_{yz} &= \mu(\partial_z u_y + \partial_y u_z). && (38) \end{aligned}$$

for the stresses. These equations can be rewritten in the velocity-stress formulation absorbing a derivative in time ($\partial_t^2 u_i = \partial_t v_i$) and taking the derivative in time of Equations 33-38. For elastic media FDM with standard grid presents problems where there is a significant contrast of properties (De Basabe & Sen 2015) therefore this forward modeling will be performed using Staggered Grid Finite Difference (SGFD). The discretization of the Elastodynamic wave equation in the displacement-stress formulation is given by

$$[u_x]^{n+1}_{i+1/2,j,k} = 2[u_x]^{n}_{i+1/2,j,k} - [u_x]^{n-1}_{i+1/2,j,k} + \Delta t^2 \left[b_x(D_x\tau_{xx} + D_y\tau_{xy} + D_z\tau_{xz} + f_x)\right]^{n}_{i+1/2,j,k}, \tag{39}$$

$$[u_y]^{n+1}_{i,j+1/2,k} = 2[u_y]^{n}_{i,j+1/2,k} - [u_y]^{n-1}_{i,j+1/2,k} + \Delta t^2 \left[b_y(D_x\tau_{xy} + D_y\tau_{yy} + D_z\tau_{yz} + f_y)\right]^{n}_{i,j+1/2,k}, \tag{40}$$

Table 4. Central differences coefficients for first order derivatives for accuracies of $2, 4, 6$ and 8th order with uniform grid spacing corresponding to staggered grid.

	Approximation at $x = 0$ x—coordinates at nodes							
Order of accuracy	$-\frac{7}{2}$	$-\frac{5}{2}$	$-\frac{3}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{7}{2}$
2				-1	1			
4			$\frac{1}{24}$	$-\frac{9}{8}$	$\frac{9}{8}$	$-\frac{1}{24}$		
6		$-\frac{3}{640}$	$\frac{25}{384}$	$-\frac{75}{64}$	$\frac{75}{64}$	$-\frac{25}{384}$	$\frac{3}{640}$	
8	$\frac{5}{7168}$	$-\frac{49}{5120}$	$\frac{245}{3072}$	$-\frac{1225}{1024}$	$\frac{1225}{1024}$	$-\frac{245}{3072}$	$\frac{49}{5120}$	$-\frac{5}{7168}$

$$[u_z]^{n+1}_{i,j,k+1/2} = 2[u_z]^{n}_{i,j,k+1/2} - [u_z]^{n-1}_{i,j,k+1/2} + \Delta t^2 \left[b_z(D_x\tau_{xz} + D_y\tau_{yz} + D_z\tau_{zz} + f_z)\right]^{n}_{i,j,k+1/2}, \tag{41}$$

for the displacement calculated on midpoints of the grid. This time D_x, D_y and D_z are the discrete operators for the first derivative and $b = 1/\rho$. For stresses

$$[\tau_{xx}]^{n+1}_{i,j,k} = \left[(\lambda + 2\mu)D_x u_x + \lambda(D_y u_y + D_z u_z)\right]^{n+1/2}_{i,j,k}, \tag{42}$$

$$[\tau_{yy}]^{n+1}_{i,j,k} = \left[(\lambda + 2\mu)D_y u_y + \lambda(D_x u_x + D_z u_z)\right]^{n+1/2}_{i,j,k}, \tag{43}$$

$$[\tau_{zz}]^{n+1}_{i,j,k} = \left[(\lambda + 2\mu)D_z u_z + \lambda(D_x u_x + D_y u_y)\right]^{n+1/2}_{i,j,k}, \tag{44}$$

$$[\tau_{xy}]^{n+1}_{i+1/2,j+1/2,k} = \left[\mu_{xy}(D_y u_z + D_x u_y)\right]^{n+1/2}_{i+1/2,j+1/2,k}. \tag{45}$$

$$[\tau_{xz}]^{n+1}_{i+1/2,j,k+1/2} = \left[\mu_{xz}(D_z u_x + D_x u_z)\right]^{n+1/2}_{i+1/2,j,k+1/2}. \tag{46}$$

$$[\tau_{yz}]^{n+1}_{i,j+1/2,k+1/2} = \left[\mu_{yz}(D_z u_y + D_y u_z)\right]^{n+1/2}_{i,j+1/2,k+1/2}. \tag{47}$$

Again the reduction for 2D from 3D media is straightforward ignoring y—dependent elements. The finite difference coefficients for staggered grid are shown in the Table 4.

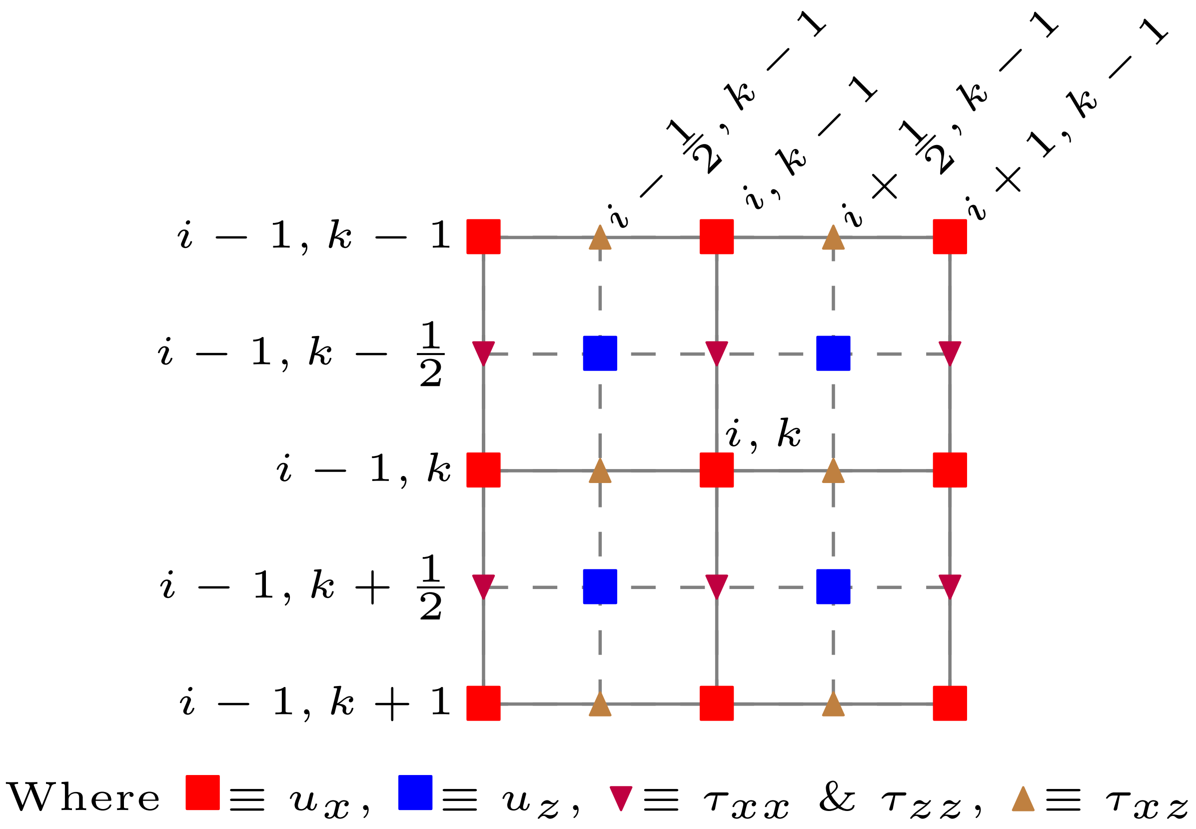


Figure 6. Visual representation of a staggered grid discretization for a 2D elastic media in terms of displacements (u_x and u_z) stresses (τ_{xx}, τ_{zz} and τ_{xz}).

Chapter 3. Inverse Problem

Acquisition of geophysical data at the surface consists of indirect measurements that need to be processed and interpreted to infer an Earth properties model with the main objective to explore or exploit natural resources of the subsoil and academic purposes. As seen in the previous chapter, these properties can be parametrized for the region of interest for the numerical computation of the geophysical data using its corresponding theory. On the other hand, the inverse problem aims to determine the properties that explain observations in seismograms for seismic data and measurements in gravimeters for gravity data.

This chapter presents the basic concepts of inverse theory, providing the theoretical framework for Gravimetric Inversion (GI) and Full Waveform Inversion (FWI) for heterogeneous acoustic and elastic media, putting much more emphasis on adjoint-state method for FWI.

3.1 Inverse theory

As stated above, starting from an Earth model, the forward problem computes theoretical data which will be compared or compute to match real data observations using inverse theory techniques. A simple illustration of this statement is shown in Figure 7.

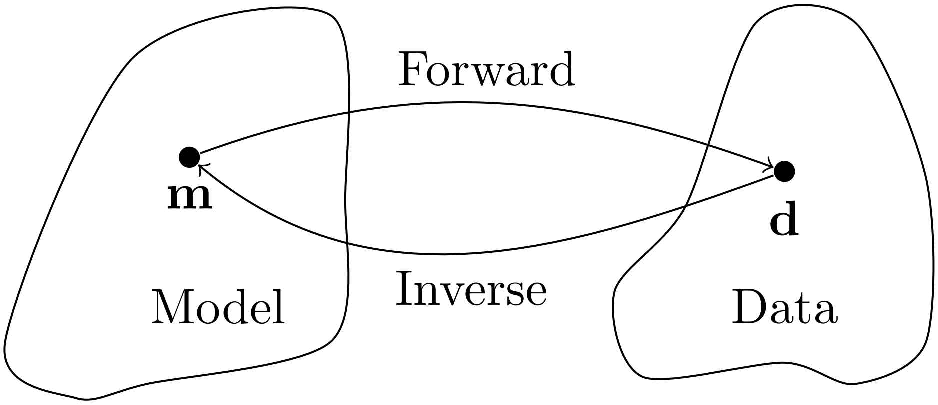


Figure 7. Illustration of the concept of forward and the inverse problems.

Inverse problems can be solved using challenging methodologies improved over the past years, however, most of the time the interpretation can be insufficient, the theory can be inaccurate and the data can be inconsistent or noisy (Jackson, 1972). In order to solve inverse problems, the following elements are essential in its formulation

- Data
- Model parameters
- Forward problem
- Cost/Objective/Error/Misfit function
- Optimization method

Since optimization methods are implied, the use of computational resources is usually required. Let us define the formulation needed for inverse theory:

Table 5. Inverse theory formulation, where N is the number of data and M the number of model parameters.

Data vector	$\mathbf{d} = [d_1\ d_2\ d_3 \cdots d_N]^T$
Model vector	$\mathbf{m} = [m_1\ m_2\ m_3 \cdots m_M]^T$
Forward modeling	$\mathbf{d} = F(\mathbf{m})$
In general	$N \neq M$

The function (F) that involves such elements needs to be stated. The objective function (also known as cost or misfit function) compares the differences between the observed data vector and the synthetic data vector thought the function

$$Q(\mathbf{m}) = \|\mathbf{d} - F(\mathbf{m})\|, \tag{48}$$

where $\|\cdot\|$ is the L_p norm and Q the objective function. The most general form of the L_p norm (Menke, 2018) is defined as

$$\|\cdot\| := \left[\sum_{i=1}^{N} |d_i - F_i(\mathbf{m})|^p\right]^{1/p}, \tag{49}$$

where N is the number of data points and p determines the norm order. Some typical values for p are

$$L_1 := \left[\sum_{i=1}^{N} |d_i - F_i(\mathbf{m})|\right] \tag{50}$$

$$L_2 := \left[\sum_{i=1}^{N} |d_i - F_i(\mathbf{m})|^2\right]^{1/2} \tag{51}$$

$$L_\infty := \max_i |d_i - F_i(\mathbf{m})|. \tag{52}$$

Since the objective function depends on the differences between the data (error), L_p norms for large values of p give a larger weight on the larger errors (Figure 8). Typically L_2 norm is used among the geophysics community, however, L_1 is also largely studied, highlighting that this norm has a discontinuity on the derivative.

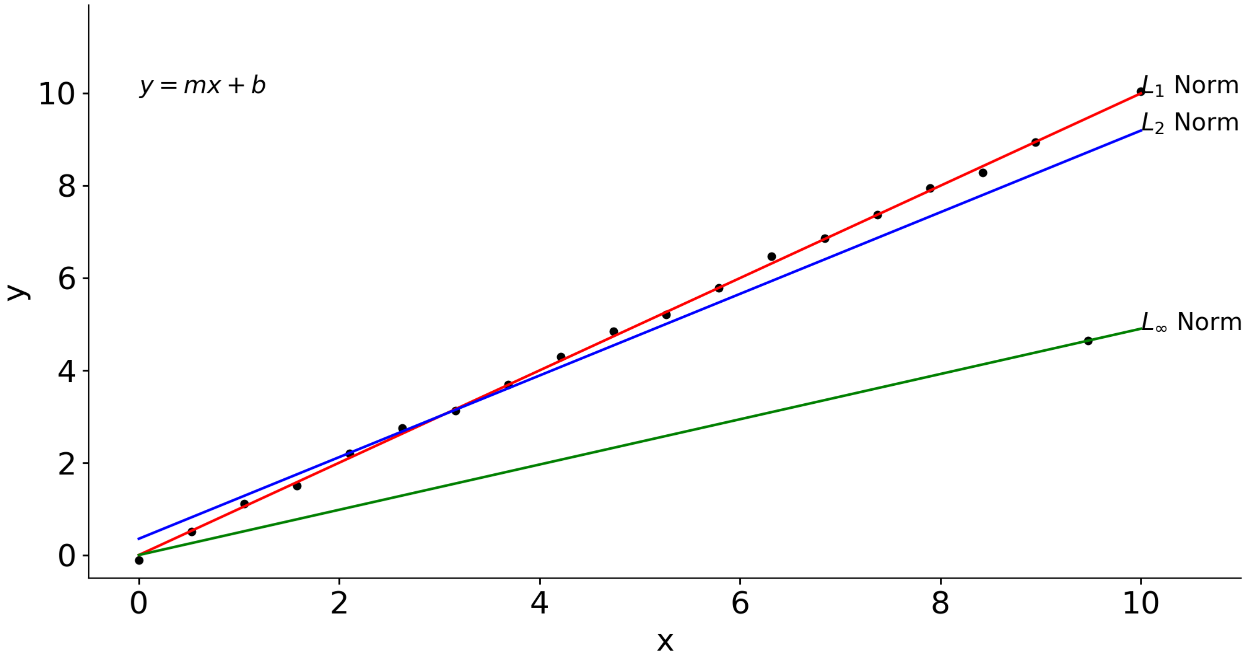


Figure 8. L_p norm for some values of p corresponding to the fit of a straight line $y = ax + b$.

3.2 Gravimetric inversion

The objective function for density estimation due to measurements of the vertical component of the acceleration ($\mathbf{g}_z^{\text{obs}}$) using the L_2 norm is given by

$$Q(\mathbf{m}_\rho) = \sum_{i=1}^{N_s} \left\| \frac{\mathbf{g}_{z_i}^{\text{obs}} - \mathbf{g}_{z_i}^{\text{cal}}}{\sigma_{\mathbf{g}_{z_i}}} \right\|^2 + \alpha_{\text{reg}}^2 ||\mathbf{D}\mathbf{m}_\rho||^2, \tag{53}$$

where α_{reg} is the regularization parameter, $\mathbf{D}$ is the discrete operator for the gradient and $\sigma_{\mathbf{g}_{z_i}}$ is the standard deviation of the data.

3.2.1 Gauss-Newton optimization

Solving this least-squares problem from Equation 53 using Gauss-Newton method (Sen and Stoffa, 2013) an estimated model $\mathbf{m}_\rho$ can be obtained as

$$\mathbf{m}_\rho = \left[\mathbf{A}^T \mathbf{C}_{dd}^{-1} \mathbf{A} + \alpha_{\text{reg}} \mathbf{D}^T \mathbf{D}\right]^{-1} \mathbf{A}^T \mathbf{C}_{dd}^{-1} \mathbf{g}_z^{\text{obs}}, \tag{54}$$

where $\mathbf{C}_{dd}^{-1}$ is the diagonal covariance matrix. This least-squares implementation requires to store and invert the square matrix in equation 54 with dimensions depending on the discretization of the model, i.e., $M \times M$. For a very fine discretization, we encounter storage problems due to this implementation, whereas we need a fine discretization in the model to achieve a good resolution for the seismic inversion and therefore the joint inversion.

3.2.2 Conjugate Gradient Least Squares

An alternative to solving Equation 54 is the use of the Conjugate Gradient Least Squares (CGLS) method. This method minimizes the objective function of equation 53 without the need to form and store the square matrix from equation 54 (Sen and

Stoffa, 2013) using a typical conjugate gradient technique. This method requires as an input $\mathbf{G}$ and $\mathbf{d}$ to find a solution, in this case, the density model ($\mathbf{m}_\rho$) for $\mathbf{G}\mathbf{m}_\rho = \mathbf{d}$, these matrices are given by

$$\mathbf{G} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2}\mathbf{A} \\ \alpha_{\text{reg}}\mathbf{D} \end{bmatrix} \tag{55}$$

$$\mathbf{d} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2}\mathbf{g}_z^{\text{obs}} \\ 0 \end{bmatrix}, \tag{56}$$

in this case, the matrix $\mathbf{G}$ will be large and sparse due to the discrete operations for the Tikhonov regularization, the model vector $\mathbf{m}_\rho$ is not modified.

3.3 Acoustic Full Waveform Inversion

Firstly, let us solve for acoustic FWI. The least-squares functional for minimizing the misfit between the observed and the synthetic pressure due to a single shot ($n_s = 1$) is given by the L_2 norm of the residual

$$Q = \frac{1}{2}\sum_r \int_0^T \left[P_r^{\text{obs}} - P_r^{\text{cal}}\right]^2 \mathrm{d}t, \tag{57}$$

where P_r^{obs} is the observed pressure and P_r^{cal} is the synthetic pressure computed using equation 20. T is the total recording time and r denotes the index for the receiver. Implicitly the P_r^{cal} depends on the model parameter $\mathbf{m}$ as $P_r^{\text{cal}} = P_r^{\text{cal}}(\mathbf{m})$. This model needs to be found in such a way Equation 57 is minimized. Taking the derivative with respect to a model perturbation

$$\partial Q = -\sum_r \int_0^T \int \left[P_r^{\text{obs}} - P_r^{\text{cal}}\right] \delta P(\mathbf{x}, t)\mathrm{d}t, \tag{58}$$

where the quantity δP is the Frèchet derivatives and represents the sensibility for each data point for each model parameter. Roughly this derivative is computed making a small perturbation in the model m_j, then performing a i—th forward modeling the derivative is obtained, therefore $i \times j$ forward modelings are needed to obtain the derivative, which is impractical to implement even with the advances in computational resources, therefore, additional methods for minimizing the problem are required.

3.3.1 Adjoint method for acoustic media

A more efficient way to minimize of Equation 57 relies on the use of the adjoint state method for the acoustic waveform. Let us minimize the augmented misfit function subject to the wave equation multiplied by an arbitrary, well behaved and derivable Lagrange multiplier $\lambda := \lambda(\mathbf{x}, t)$ remaining to be defined (Plessix, 2006) then

$$Q = \frac{1}{2}\sum_r \int_0^T \int_\Omega \left[P_r^{\text{obs}} - P_r^{\text{cal}}\right]^2 \delta(\mathbf{x}-\mathbf{x}_r)\mathrm{d}^3x\mathrm{d}t - \int_0^T \int_\Omega \lambda \left[\frac{1}{v_p^2}\frac{\partial^2 P}{\partial t^2} - \nabla^2 P - f\right]^2 \mathrm{d}^3x\mathrm{d}t, \tag{59}$$

notice that the last term of Equation 59 is zero, corresponding to the wave equation acting as constriction, therefore the problem is consistent. Taking the total derivative

$$\begin{aligned}\delta Q = &-\sum_r \int_0^T \int_\Omega \left[P_r^{\text{obs}} - P_r^{\text{cal}}\right] \delta(\mathbf{x}-\mathbf{x}_r)\delta P\mathrm{d}^3x\mathrm{d}t \\ &- \int_0^T \int_\Omega \lambda \left[-\frac{2}{v_p^3}\frac{\partial^2 P}{\partial t^2}\right] \mathrm{d}^3x\mathrm{d}t \\ &- \int_0^T \int_\Omega \lambda \left[\frac{1}{v_p^2}\frac{\partial^2 \delta P}{\partial t^2} - \nabla^2 \delta P\right] \mathrm{d}^3x\mathrm{d}t,\end{aligned} \tag{60}$$

where the source is considered as independent of the model parameter perturbation. Again the Frèchet derivatives δP appears on the first and last term. In the last term, the linear operator of the wave equation ($\mathcal{L} = \frac{1}{v_p^2}\frac{\partial^2}{\partial t^2} - \nabla^2$) is acting over δP which is a computation that we are looking to avoid. For this, let us first integrate by parts two times for t as

$$\int_0^T \lambda \left[\frac{\partial^2 \delta P}{\partial t^2}\right] \mathrm{d}t = \left(\lambda \frac{\partial \delta P}{\partial t}\right)\Big|_0^T - \left(\frac{\partial \lambda}{\partial t}\delta P\right)\Big|_0^T + \int_0^T \left[\frac{\partial^2 \lambda}{\partial t^2}\right] \delta P\mathrm{d}t. \tag{61}$$

Without loss of generality, let us set $\lambda(x,t=T)=\frac{\partial\lambda}{\partial t}(x,t=T)=0$, thus

$$\int_0^T \lambda\left[\frac{\partial^2\delta P}{\partial t^2}\right]\mathrm{d}t=\int_0^T\left[\frac{\partial^2\lambda}{\partial t^2}\right]\delta P\mathrm{d}t, \tag{62}$$

this means that the second derivative is a self-adjoint operator ($\mathcal{L}=\mathcal{L}^*$). For the Laplacian operator ∇ the same procedure can be done, setting the correct boundary conditions for space. Consider

$$\int_\Omega \lambda\left[\nabla^2\delta P\right]\mathrm{d}^3x. \tag{63}$$

Taking into consideration the following identity

$$\psi\nabla^2\phi-\phi\nabla^2\psi=\nabla\cdot(\psi\nabla\phi-\phi\nabla\psi), \tag{64}$$

then

$$\int_\Omega \lambda\left[\nabla^2\delta P\right]\mathrm{d}^3x=\int_\Omega \delta P\left[\nabla^2\lambda\right]\mathrm{d}^3x+\int_\Omega \nabla\cdot\left[\lambda\nabla\delta P-\delta P\nabla\lambda\right]\mathrm{d}^3x. \tag{65}$$

Applying Gauss’ theorem on the right side of the equation

$$\int_\Omega \nabla\cdot\left[\lambda\nabla\delta P-\delta P\nabla\lambda\right]\mathrm{d}^3x=\int_{\partial\Omega}\left[\lambda\nabla\delta P-\delta P\nabla\lambda\right]\cdot\hat{n}dS, \tag{66}$$

where the integral was changed from volumetric to superficial. Setting homogeneous Dirichlet and Neumann boundary conditions for the Lagrange multiplier

$$\int_{\partial\Omega}\left[\lambda\nabla\delta P-\delta P\nabla\lambda\right]\cdot\hat{n}dS=0, \tag{67}$$

which corresponds to the spatial condition for the multiplier, i.e., the same free surface condition. In this way Equation 60 can be rewritten as

$$\begin{aligned}\delta Q=&-\sum_r\int_0^T\int_\Omega\left[P_r^{\mathrm{obs}}-P_r^{\mathrm{cal}}\right]\delta(\mathbf{x}-\mathbf{x}_r)\delta P\mathrm{d}^3x\mathrm{d}t\\&-\int_0^T\int_\Omega\lambda\left[-\frac{2}{v_p^3}\frac{\partial^2P}{\partial t^2}\right]\mathrm{d}^3x\mathrm{d}t\\&-\int_0^T\int_\Omega\left[\frac{1}{v_p^2}\frac{\partial^2\lambda}{\partial t^2}-\nabla^2\lambda\right]\delta P\mathrm{d}^3x\mathrm{d}t,\end{aligned} \tag{68}$$

where the first and last term can be compacted as

$$\begin{aligned} \delta Q = & -\int_0^T \int_\Omega \left\{ \frac{1}{v_p^2}\frac{\partial^2 \lambda}{\partial t^2} - \nabla^2 \lambda + \sum_r \left[P_s^{\text{obs}} - P_s^{\text{cal}} \right] \delta(x - x_r) \right\} \delta P \mathrm{d}^3 x \mathrm{d}t \\ & -\int_0^T \int_\Omega \lambda \left[-\frac{2}{v_p^3}\frac{\partial^2 P}{\partial t^2} \right] \mathrm{d}^3 x \mathrm{d}t. \end{aligned} \tag{69}$$

Let us define Lagrange multiplier λ in such a way that the first term of Equation 69 is canceled and therefore, the Frèchet derivatives are nullified. Then

$$\frac{1}{v_p^2}\frac{\partial^2 \lambda}{\partial t^2} - \nabla^2 \lambda = -\sum_r \left[P_r^{\text{obs}} - P_r^{\text{cal}} \right] \delta(\mathbf{x} - \mathbf{x}_r), \tag{70}$$

which corresponds to another wave equation using the residuals as a source at the seismogram locations. The importance of this result relies on the computation of the gradient without the need to compute the Frèchet derivatives, instead, a single additional forward modeling needs to be performed using the same wave propagation method but with the residuals as a source. Finally, to give more meaning to the Lagrange multiplier lets define $\lambda(x,t) \equiv P^\dagger(x, T-t)$, thus the gradient

$$\partial V = \frac{2}{v_p^3} \int_0^T P^\dagger(\mathbf{x}, T-t) \frac{\partial^2 P(\mathbf{x},t)}{\partial t^2} \mathrm{d}t, \tag{71}$$

which is a convolution of the pressure and adjoint wave fields.

3.3.2 Gradient construction

In the past section, the gradient due to a single source was obtained. Using multiple sources requires a summation as follows

$$\partial V = \frac{2}{v_p^3} \sum_s^{n_s} \int_0^T P^\dagger(\mathbf{x}, T-t) \frac{\partial^2 P(\mathbf{x},t)}{\partial t^2} \mathrm{d}t. \tag{72}$$

where n_s is the total number of shots. Notice that the pressure and adjoint wavefields are computed in opposite directions for the time stepping: $P(\mathbf{x},t)$ is going forward in time meanwhile $P^{\dagger}(\mathbf{x}, T-t)$ is going backward in time, as illustrated in Figure 9.

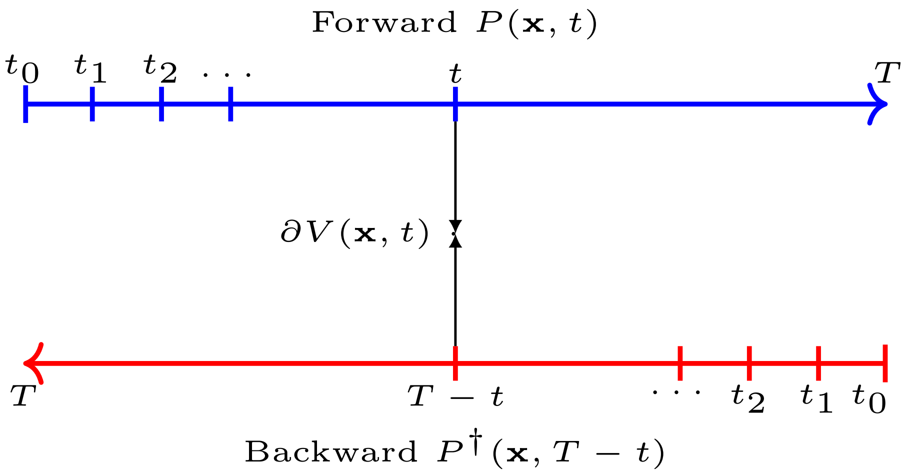


Figure 9. Forward pressure and backward adjoint wave field for gradient computation were each field is propagated in the opposite direction and accesed in different times.

At the time t the pressure and adjoint pressure have to be accessed simultaneously to compute Equation 72. This could be done by saving all the time steps for each wavefield, i.e., a variable $P(\mathbf{x}, T-t)$ with a size of $n_t n_x n_y n_z$ elements needs to be stored for each source to compute the gradient. The algorithm 2 presents the basic steps to obtain a 3D velocity gradient ∂V for n_s sources, where the change to 2D media is straightforward considering y-component equal to zero and therefore the spatial loop $(i_y = 1 : n_y)$ is not performed

Saving all the time slices is impractical for 3D media and high-resolution problems because of large memory requirement. One solution for this issue is going forward in time for the wave propagation P_{n_t}, save the two final slices P_{n_t} and P_{n_t-1} and then perform a wave back-propagation in time to the starting point P_0 as

$$P^{n-1}_{i,j,k} = 2P^{n}_{i,j,k} - P^{n+1}_{i,j,k} + v_p^2 \Delta t^2 \left(D_x^2 P^{n}_{i,j,k} + D_y^2 P^{n}_{i,j,k} + D_z^2 P^{n}_{i,j,k} \right), \tag{73}$$

where the pressure P is computed at the slice $n-1$. In this way, both P and $P^{\dagger}$ are accessed at the same time, however, an additional forward modeling is performed in the total computational cost. This solution has been used by several authors, however, it could present issues when visco-elastic effects are considered (Tromp *et al.*, 2005).

```
Data: Observed pressure $P^{\text{obs}}[n_t, n_r, n_s]$, starting velocity model $v_p[n_x, n_y, n_z]$
Input : $n_s$ time steps, $n_r$ receivers, $n_s$ sources, $n_x$,$n_y$, $n_z$ grid nodes.
for $i_s = 1 : n_s$ do
    Forward modeling;
    for $i_t = 1 : n_t$ do
        for $i_x = 1 : n_x$ do
            for $i_y = 1 : n_y$ do
                for $i_z = 1 : n_z$ do
                    Computes regular field $P[i_x, i_y, i_z, i_t + 1]$, Eq.26
                end
            end
        end
        for $i_r = 1 : n_r$ do
            Measure at receiver location $P^{\text{cal}}[i_t, i_r, i_s] = P[i_{xr}, i_{yr}, i_{zr}, i_t]$;
            Compute data residual $R[i_t, i_r, i_s] = P^{\text{obs}}[i_t, i_r, i_s] - P^{\text{cal}}[i_t, i_r, i_s]$
        end
    end
    Adjoint wave field and velocity gradient;
    for $i_t = 1 : n_t$ do
        for $i_x = 1 : n_x$ do
            for $i_y = 1 : n_y$ do
                for $i_z = 1 : n_z$ do
                    Compute adjoint field $P^{\dagger}[i_x, i_y, i_z, i_t + 1]$;
                    Compute gradient Eq.72, $\partial V_{i_s}[i_x, i_z]$;
                end
            end
        end
    end
    Stack gradient $\partial V \leftarrow \partial V + \partial V_{i_s}$
end
Output: Velocity gradient $\partial V$, data residual $R[n_t, n_r, n_s]$
```

Algorithm 2: 3D gradient construction using the adjoint method.

3.3.2.1 Acoustic banana-doughnut kernel

In Equation 72 a regular wavefield $P(\mathbf{x}, t)$ propagates forward in time t with starting condition $P(\mathbf{x}, 0) = 0$ and an adjoint wavefield $P^{\dagger}(\mathbf{x}, T - t) = 0$ propagates backward in time with terminal condition $P^{\dagger}(\mathbf{x}, T) = 0$. This means that, at an early time t, there is no interaction between both wave fields for the convolution. In order to seek for the zone of influence between the fields considerer the 3D forward propagation due to a source located at $(500, 500, 500)$ meters and a receiver at $(1500, 500, 500)$ m for a homogeneous velocity model of 1000 km/s. For the regular wavefield a Ricker wavelet with peak frequency of 15 Hz shifted 0.1 seconds is applied with a time recording $T = 2$ s (Figure 10). For this example lets consider the same Ricker wavelet shifted $T - 1.0 = 1.9$ s for illustration purposes.

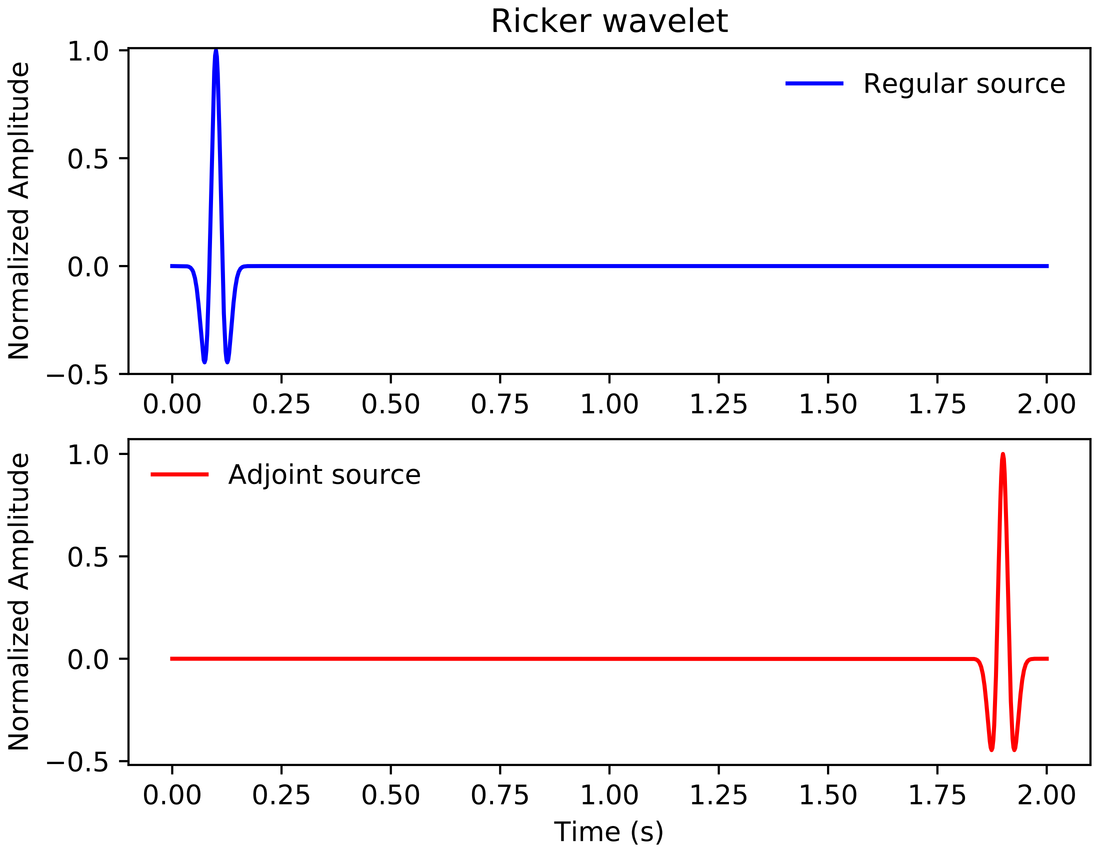
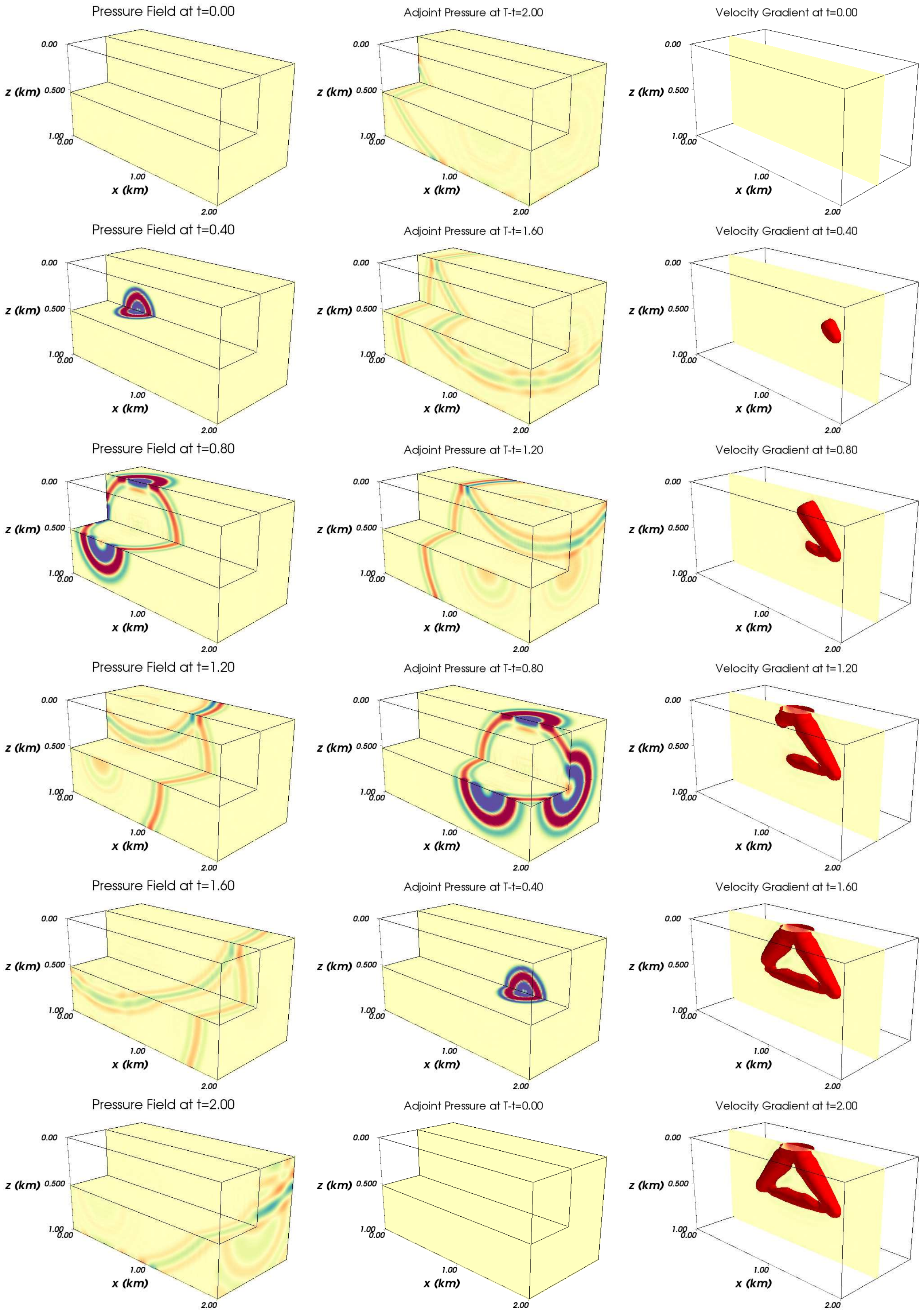


Figure 10. Ricker wavelet for the regular wavefield (blue) and adjoint wave field (red) with peak frequency of 15 Hz used for the contruction of the acoustic banana-doughnut kernel.

Figure 11 shows the propagation for the regular (P) and adjoint ($P^{\dagger}$) wave field and the gradient construction (Red surface) for 2 seconds of recording time. This banana-doughnut shape (Dahlen *et al.*, 2000) exhibits were both wavefields have the most interaction, we can identify two primary zones: the region between the source and the receiver (cigar shape) and the region where the ray path reflects on the surface (banana shape), where both regions have the source and receiver locations as starting and ending point. The second region of influence has the same direction as the ray path, indicating the sensitivity between source and receiver due to a reflection, in this case, the reflection due to free surface boundary condition, therefore when reflections due to stratigraphic changes appear, the sensibility will be presented at this region and between the source and the receivers, which will be analyzed later for an heterogeneous 2D media.



Pressure Field at t=0.00

Adjoint Pressure at T-t=2.00

Velocity Gradient at t=0.00

Pressure Field at t=0.40

Adjoint Pressure at T-t=1.60

Velocity Gradient at t=0.40

Pressure Field at t=0.80

Adjoint Pressure at T-t=1.20

Velocity Gradient at t=0.80

Pressure Field at t=1.20

Adjoint Pressure at T-t=0.80

Velocity Gradient at t=1.20

Pressure Field at t=1.60

Adjoint Pressure at T-t=0.40

Velocity Gradient at t=1.60

Pressure Field at t=2.00

Adjoint Pressure at T-t=0.00

Velocity Gradient at t=2.00

Figure 11. 3D wave propagation for regular wavefield (left) and adjoint wave field (right) and construction of the gradient at times 0,0.4,0.8,1.2, 1.6 and 2.0 seconds for an homogenous media.

3.3.2.2 Example: 2D Marmousi model gradient

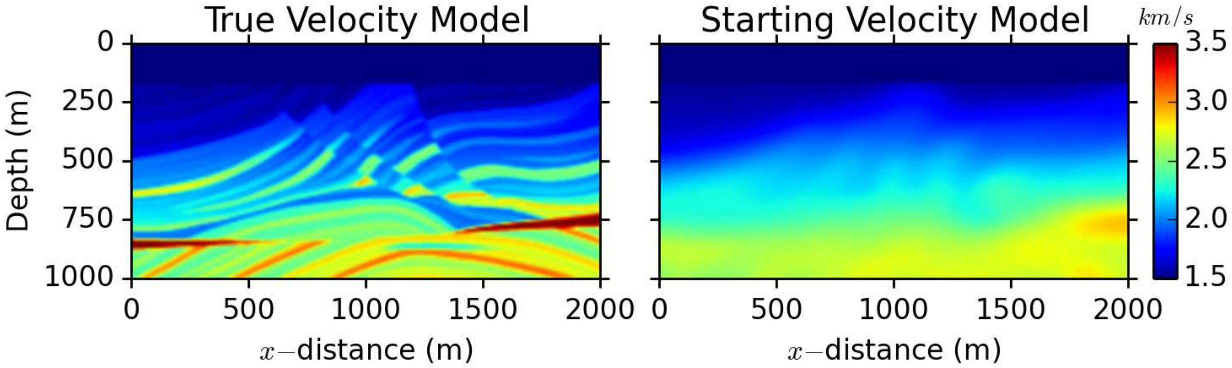


Figure 12. Modified Marmousi velocity model (left) and starting velocity model (right). The velocity was shortened to 3500 m/s.

Consider the modified Marmousi and a starting 2D velocity model of Figure 12. This model was created by the *Institut Francais du Petrole* for depth-migration and velocity estimation in seismic imaging (Versteeg, 1994) and has become a standard example in FWI and RTM for the scientific community. It involves slightly folded layers similar to a bookshelf sliding fault system and a discordance event at the bottom. The velocity range was shortened to $1500 - 3500$ *km/s* covering a depth of 1000 meters and a horizontal distance of 2000 meters on a grid of $n_x = 200$ and $n_z = 100$ grid nodes. The starting model is a smoothed (Gaussian smoothing) version of the true velocity model and the water layer is considered to be known in both models. The following table summarizes the parameters used for the forward modeling and the construction of the gradient for this example.

Table 6. Parameters used for the construction of the gradient for the Marmousi model.

Parameter	Value	Unit	Description
x_{MIN}	0	m	Starting horizontal distance
x_{MAX}	2000	m	Final horizontal distance
z_{MIN}	0	m	Starting depht
z_{MAX}	1000	m	Final depth
n_x	200		Grid nodes for x
n_z	100		Grid nodes for z
T	2	s	Recording time
n_t	1500		Grid nodes for t
v_{MAX}	3500	m/s	Maximum velocity
f	15	Hz	Ricket wavelet's peak frequency
n_r	200		Number of receivers
n_s	100		Number of sources

The distances and the grid nodes used correspond to a spacing of $\Delta x = \Delta z = 10$ meters and a time stepping of $\Delta t = \frac{20}{15} \times 10^{-4} \approx 1.33 \times 10^{-4}$ seconds, which along the maximum velocity satisfies the stability condition

$$\begin{aligned} \Delta t \leq & \frac{\Delta x}{v_{\text{MAX}}\sqrt{2}} \frac{1}{\sqrt{\sum_{i=1}^{M} a_m}} \\ 1.33 \times 10^{-4} \leq & \frac{10}{v_{\text{MAX}}\sqrt{2}} \frac{1}{\sqrt{\sum_{i=1}^{M} a_m}}, \end{aligned} \tag{74}$$

for a 10th order FDM in a 2D media. With respect to the receivers and sources, both locations are equally spaced along the surface, 10 meters spacing between seismograms and 20 meters spacing for sources (shots). For this geometry, the seismic traces are shown in Figure 13 for some shots. This data acquisition correspond to the observed data vector $\mathbf{P}^{\text{obs}} \in \mathbb{R}^{n_t n_r n_s}$ which based on the parameters of Table 6 gives a vector of $1500 \times 200 \times 100 = 30$ Million data elements.

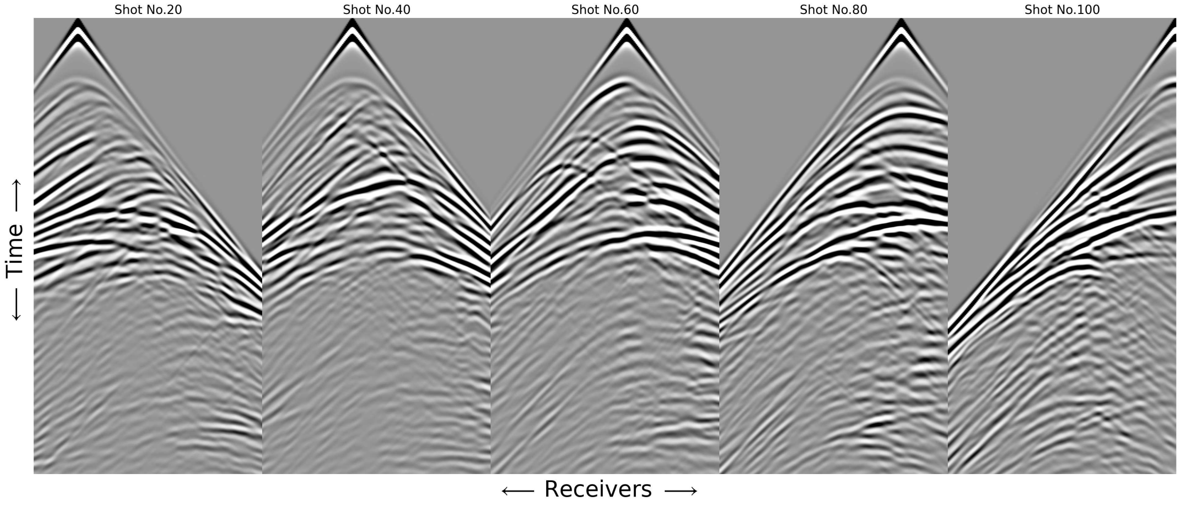


Figure 13. Synthetic seismic data acquisition for the Marmousi model at shots number 20, 40, 60, 80 and 100 corresponding to 200 receivers equally spaced at the surface.

Next, the gradient construction is discussed explaining the contribution of receivers and sources.

3.3.2.3 Gradient construction: Receiver analysis

The gradient construction can be done using the elements of Table 6 and algorithm 2 considering the starting model of Figure 12. To illustrate how the velocity gradient is built using the adjoint method, let us consider a single seismogram located at $(x = 500, z = 0)$ for a source at $(x = 500, z = 0)$. The adjoint field (left), pressure field (center) and the gradient (right) construction over time are shown in Figure 14, adding the footprint of the Marmousi model in the adjoint field for illustration purposes. Although the starting velocity model looks close to the true Marmousi model, the wave propagation of the pressure field behaves almost as a homogeneous model because the velocity contrasts have been smoothed. The wave propagation of the adjoint field corresponds to the propagation of the residual, hence several impulses in the source are presented.

It is evident that the total contribution to the gradient is due to zones where the pressure and adjoint field interact. At earlier iterations, there is no interaction because of the initial conditions, while at later times the gradient exhibits the final aspect of Figure 14 at 2 s. Taking a close look in the final gradient we observe some wavefronts at regions where there is a major velocity contrast and in the region of interaction between the source and the receivers.

Since the interaction zone occurs mostly at the region between the source and receiver let us explore the contribution for receivers located at $(0, 0)$, $(500, 0)$, $(1000, 0)$, $(1500, 0)$ and $(2000, 0)$ as well as the contribution of the whole seismic stream ($f = \sum_r R$) as shown at Figure 15.

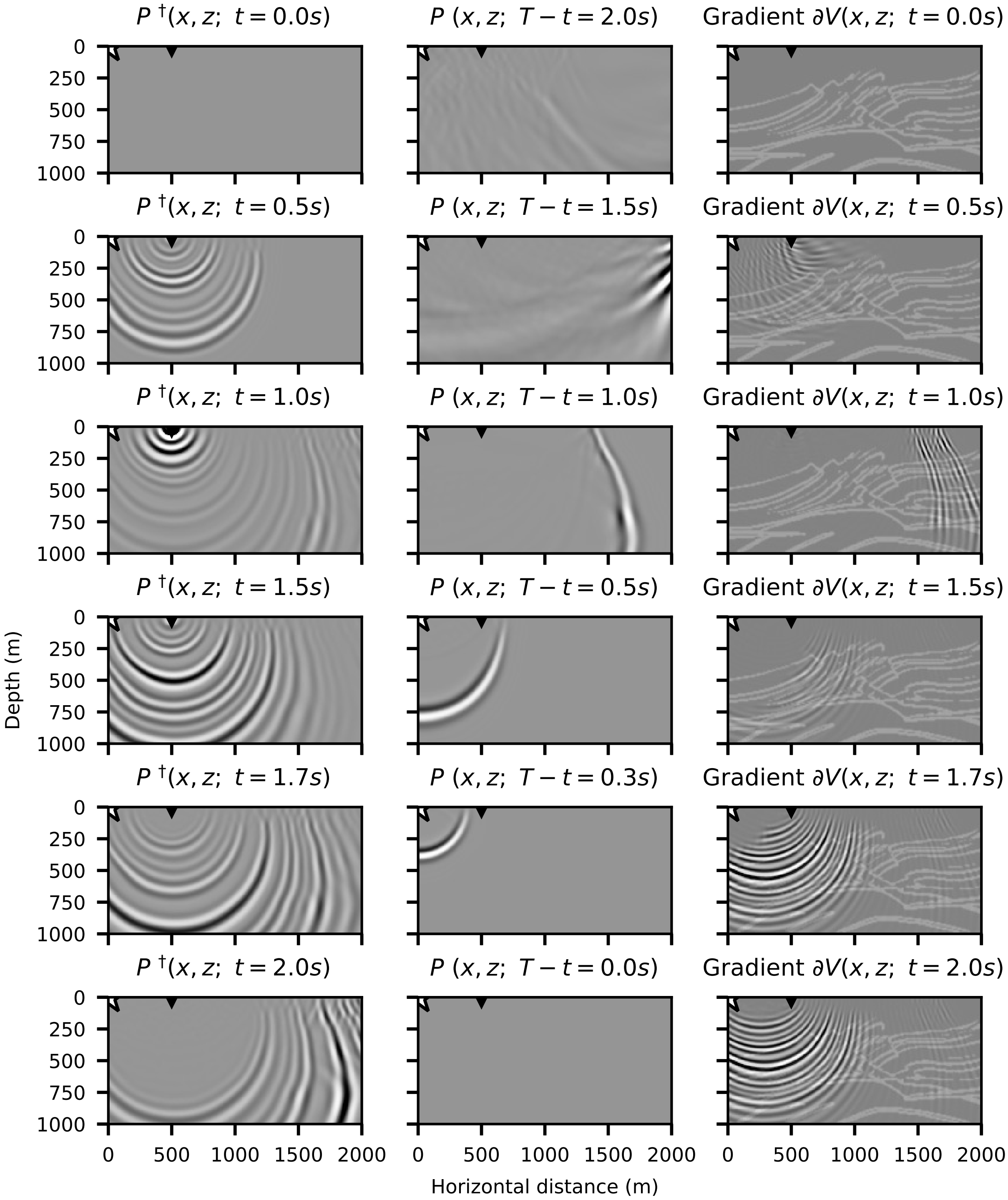


Figure 14. Adjoint wave-field (left), pressure field (center) and the velocity gradient (right) construction for $t = [0.0, 0.5, 1.0, 1.5, 1.7, 2.0]$ seconds. The edges of the Marmousi model are displayed in the gradient for illustration purposes. The white start represents the source positions and the downsided black-triangle the receiver.

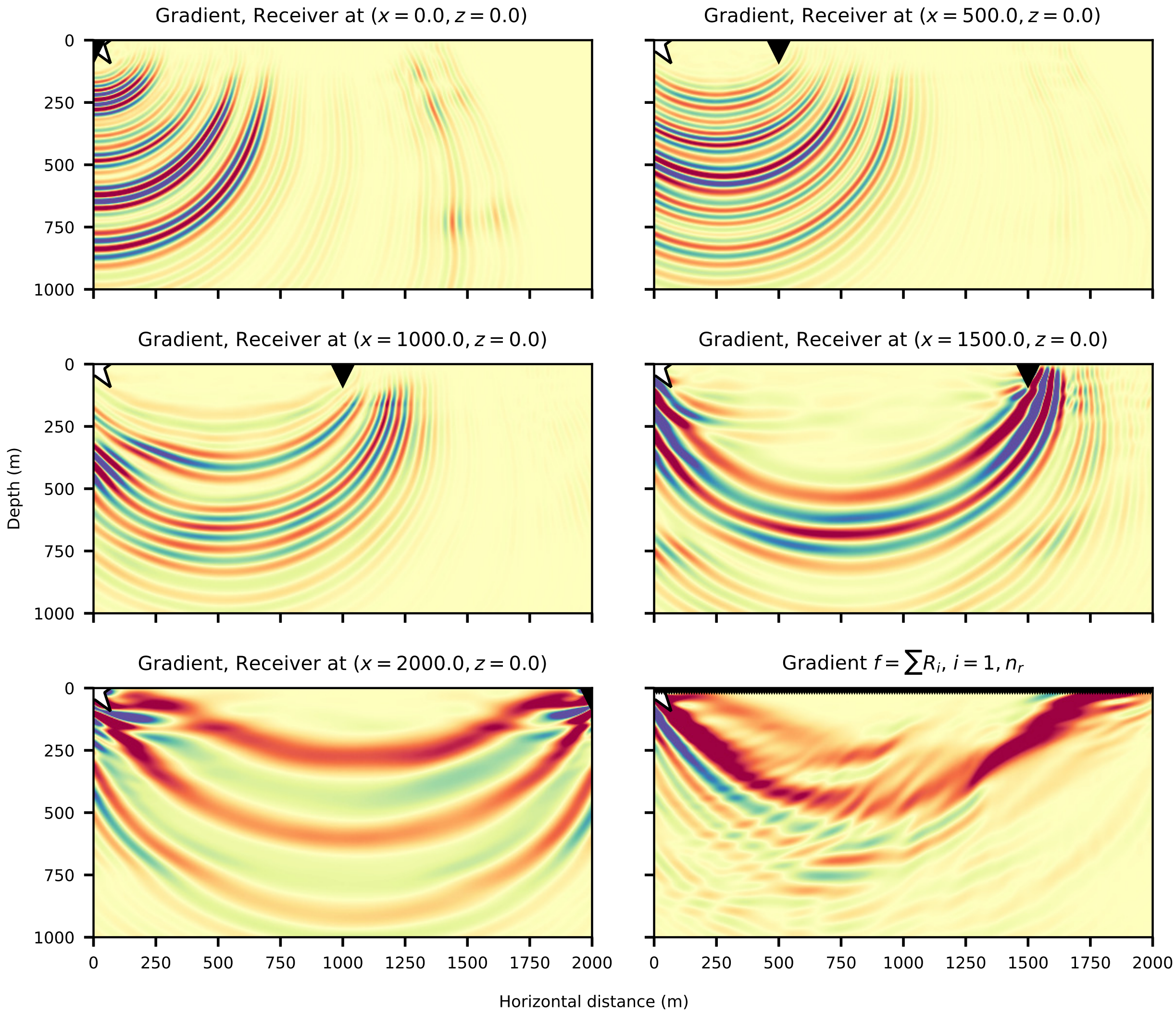


Figure 15. Velocity gradient for a source located at the origin for several receivers. The white start represents the source positions and the downsided black-triangle the several receivers used. For the bottom figure on the right the whole stream of receivers is used for the gradient.

Again the gradient in the function of the receiver position is more sensitive when it is located closer to the source and loses sensitivity far away. The final gradient covering the whole data residual ($f = \sum R_i$) adds all the contributions for each receiver over the surface, obtaining the image of Figure 15 (bottom right) with a better resolution between 0 to 1000 meters because of the source position. Performing FWI with a single source will solve only for velocities beneath and nearby to the source position, thus the role of the source position has to be explored.

3.3.2.4 Gradient construction: Source analysis

Now let us consider several source positions at the surface, $s = [0, 500, 1000, 1500, 2000]$ meters for a depth $z = 0$. The gradient for each source as well as the gradient stacked for all sources ($\partial V = \sum \partial V_i$) is shown at Figure 16

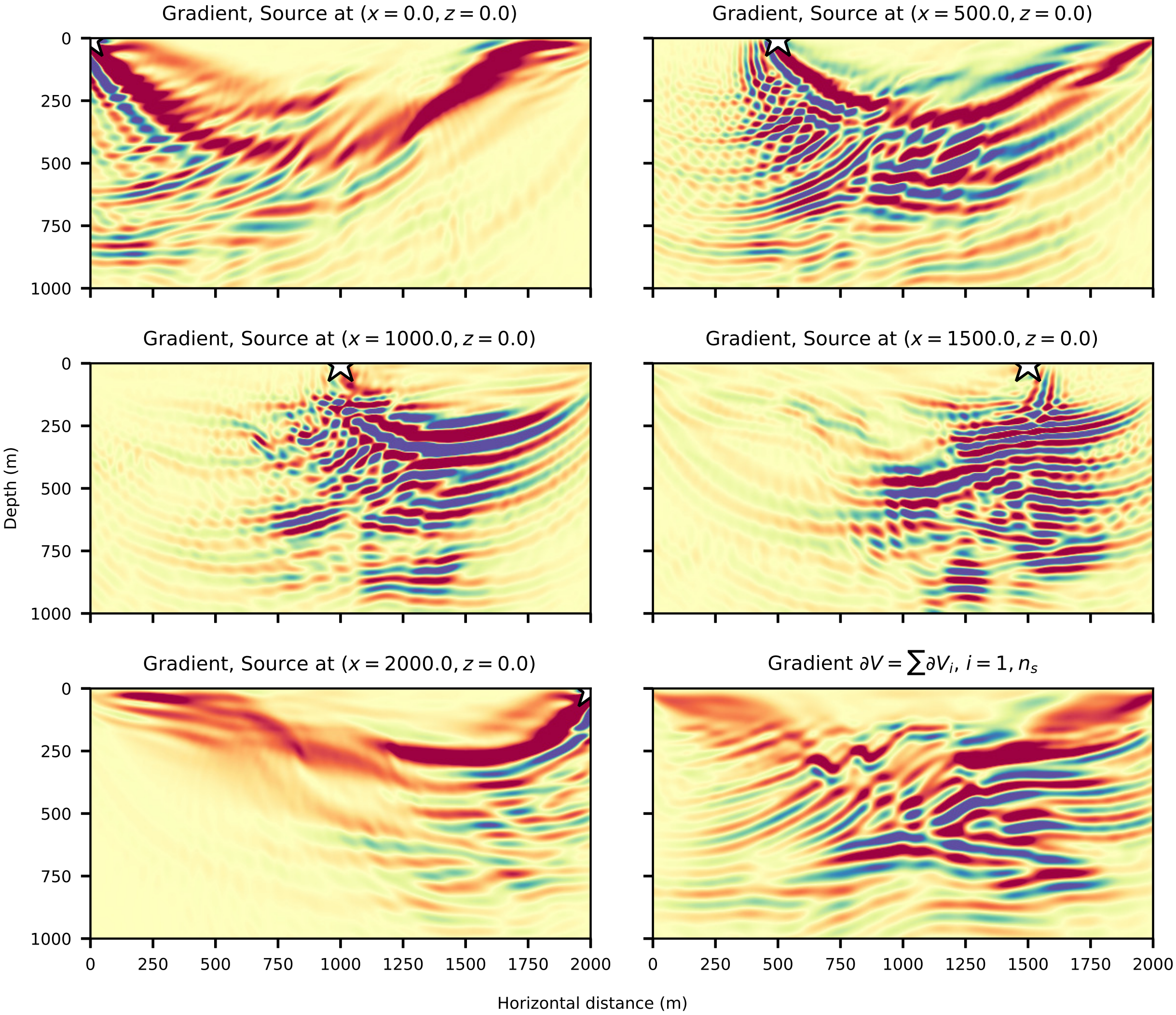


Figure 16. Velocity gradient for a several source locations $s = [0, 500, 1000, 1500, 2000]$ using the whole stream of receivers (200). The white start represents the different source positions.

Each gradient exhibits more sensibility beneath its position at the surface, even though the surface is fully covered with receivers. While the image is not clear at all for each one, the addition of all gradients ($\sum \partial V_i$) into a single one produces a velocity gradient with fine resolution. The image (Figure 16 bottom right) resembles the footprint of the layers for the Marmousi model which roughly looks like typical seismic migration (and RTM). This velocity model will be added to the starting model with the correct amount (step line-search) for an iterative process for FWI. Based on these previous analyses for the number of receivers and sources, we conclude that the use of more receivers and sources is important for the capability to obtain a high-resolution image for FWI.

3.3.2.5 Gradient construction: Frequency analysis

Finally, the role of the frequency will be discussed. Recalling the seismic scale used in seismic processing in the industry (Figure 17), the gradients using frequencies of $f = [5, 8, 10, 15, 25, 50]$ Hz are shown in the Figure 18.

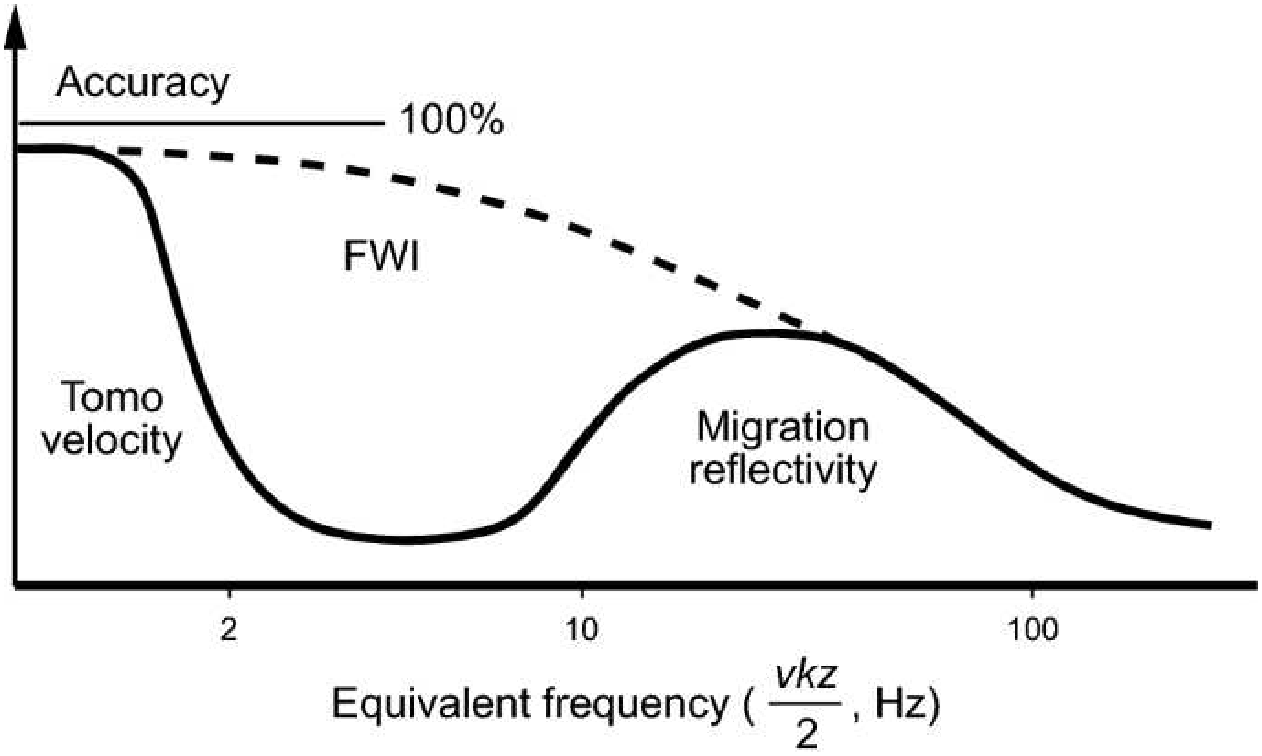


Figure 17. Seismic accuracy in terms of frequency. Taken from Yao & Wu (2017).

Notice that for low frequencies the events on the gradient become wider, decreasing its thickness when the frequency increments. This result is attractive when it comes to FWI: solving for low frequencies will recover large amounts of velocity meanwhile high frequencies will focus on small details on the velocity model. This technique is well-known within the geophysicist community, named as multi-scale FWI. The main goal of this methodology is to avoid local minima and the mitigation of phenomena like cycle skipping, which will be discussed later. The results of this work will be on the frequency range of 5 to 15 Hz

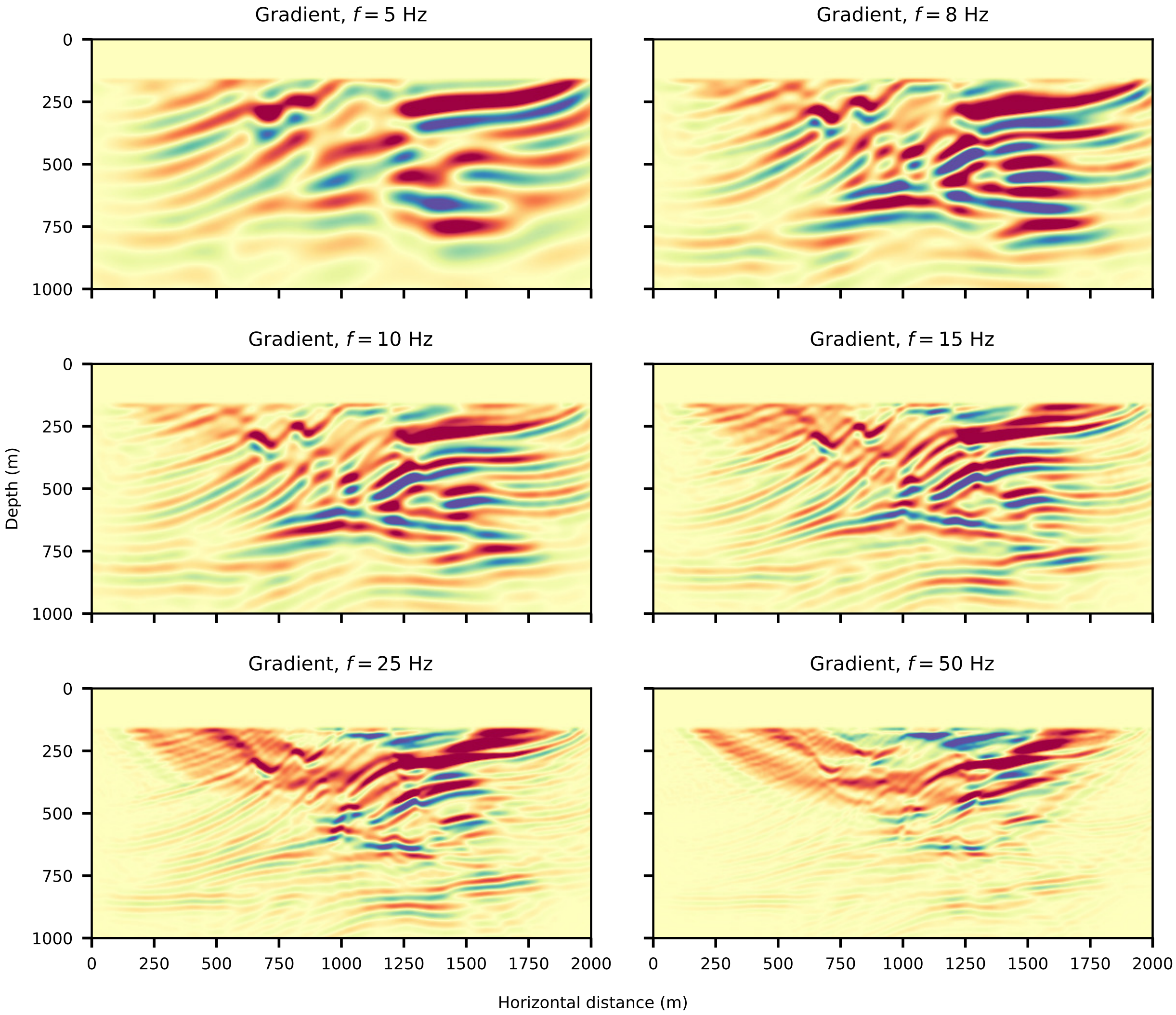


Figure 18. Velocity gradient for frequencies $f = [5, 8, 10, 15, 25, 50]$ Hz.

3.3.3 Gradient-based optimization

Gradient-based optimization minimizes Equation 57 by updating the velocity model V iteratively,

$$V_{n+1} = V_n + \alpha_n \partial V, \tag{75}$$

where the scalar α_n is the step length which represents how much the current model V_n moves along the direction ∂V at the n-th iteration. The efficient minimization depends on an effective choice of the step α which can lead to local minima or global minima as seen in Figure 19.

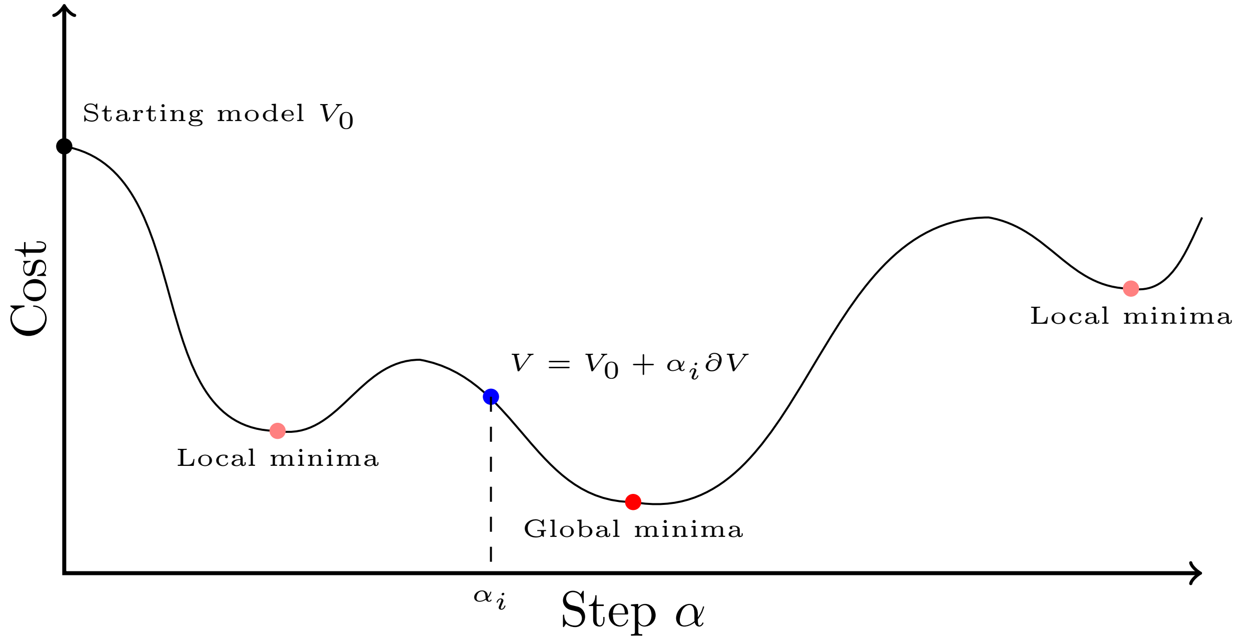


Figure 19. Arbitrary cost function in terms of steps (α) for the ilustration of local and global minima for $V = V_0 + \alpha_i \partial V$.

There are several algorithms for the search of the optimal step length α (Nocedal and Wright, 2006). For this work two techniques were explored. The first consists of picking some test points for the step, for example, see Figure 20 for an arbitrary cost function curve.

With these test points (green color) a parabolic fit $f(x) = ax^2 + bx + c$ for alpha is performed (blue line). The minimum value of the quadratic function will be at $x = -\frac{b}{2a}$

(color violet), which could be close enough to the minimum (color red) depending on the choice of the test values. On the other hand, bad test points (Figure 21) yield bad values of α (color violet) which corresponds to the actual value (color cyan) on the step line curve. It is worth mentioning that each test point represents an additional forward modeling evaluation, thus more test points require more computational cost.

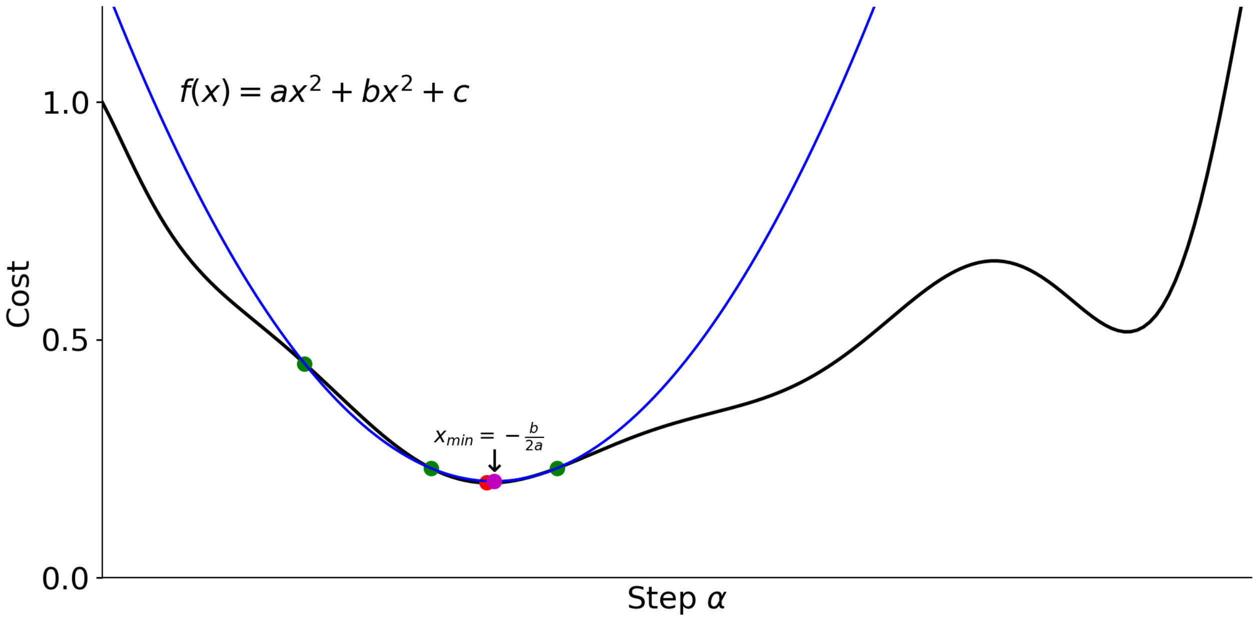


Figure 20. Step line search method using 3 test points (color green) and a parabolic fit $f(x) = ax^2 + bx + c$ (color blue). The red point correspond to the minimum desired and the violet point is the actual value found.

Another alternative is step line search method using interval reduction. Consider the range of values of steps $[\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_k]$ with k the number of test points for $\alpha_1 < \alpha_2 < \alpha_3 < \cdots < \alpha_k$ with their respective cost $[\text{cost}_1, \text{cost}_2, \text{cost}_3, \cdots, \text{cost}_k]$. In this method, we select the value of α which correspond to the minimum cost. If the minimum cost correspond to the first test value then for the next iteration a *zoom in* is performed for the test points $[\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_k] \times \text{zoom}$ with $\text{zoom} < 1$, on the other hand if the optimal step correspond to the final point test a *zoom out* is performed as $[\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_k]/\text{zoom}$. A typical value of zoom is $\frac{\sqrt{5}-1}{2}$ corresponding to the reciprocal of the Golden ratio. Evidently, this method is more expensive computationally speaking, however it has a compensation on effectiveness due to the optimal choice of α.

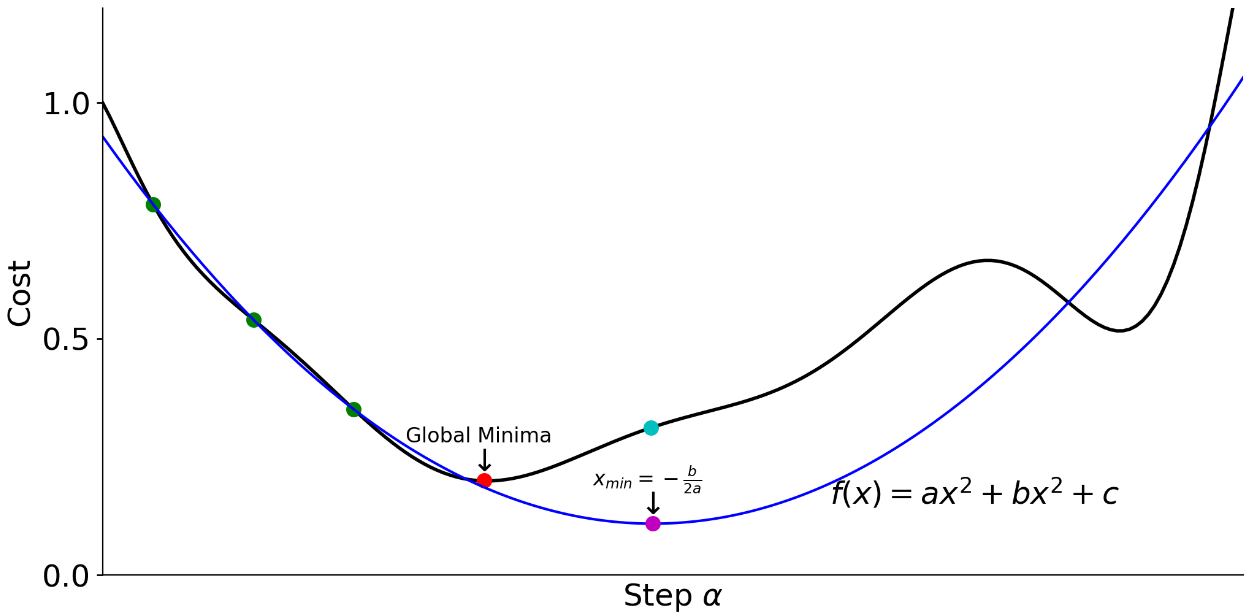


Figure 21. Bad test points choosen for step line search method using quadratic fit using 3 points (color green). The red point correspond to the minimum desired, the violet point is the value found and the cyan point is the actual value in the function (blue line).

3.3.4 AFWI Example: Marmousi model

Once exposed all the essential elements for AFWI (Acoustic Full Waveform Inversion) such as forward modeling, gradient build, step-line search, the iterative minimization can be performed as the following algorithm summarizes

```
Data: Observed pressure $P^{\text{obs}}[n_t, n_r, n_s]$, starting velocity model $V$
Input : FWI iterations $n_{\text{iter}}$ .
for $i_{iter} = 1 : n_{iter}$ do
    Computes forward modeling and adjoint method (Alg. 2);
    Performs step-line search for $\alpha$;
    Update velocity model $V_{\text{iter+1}} = V_{\text{iter}} + \alpha \partial V$;
end
Output: Velocity model $V$, data residual $R[n_t, n_r, n_s]$
```

Algorithm 3: Typical AFWI proccess.

Using the same parameters of the exercise for gradient construction using the Marmousi model (Table 6) and 10 test points at the step line search method with interval reduction, the FWI process for some iterations is shown in Figure 22. At early iterations, the first aspect to recover is the stratigraphic information because of the shape of the gradient (Figure 16). For the next iterations, the velocity starts recovering its value at each point of the model, with more resolution on the central part of the survey. The final velocity after 228 iterations (Figure 22 bottom left) resembles quite good the true model.

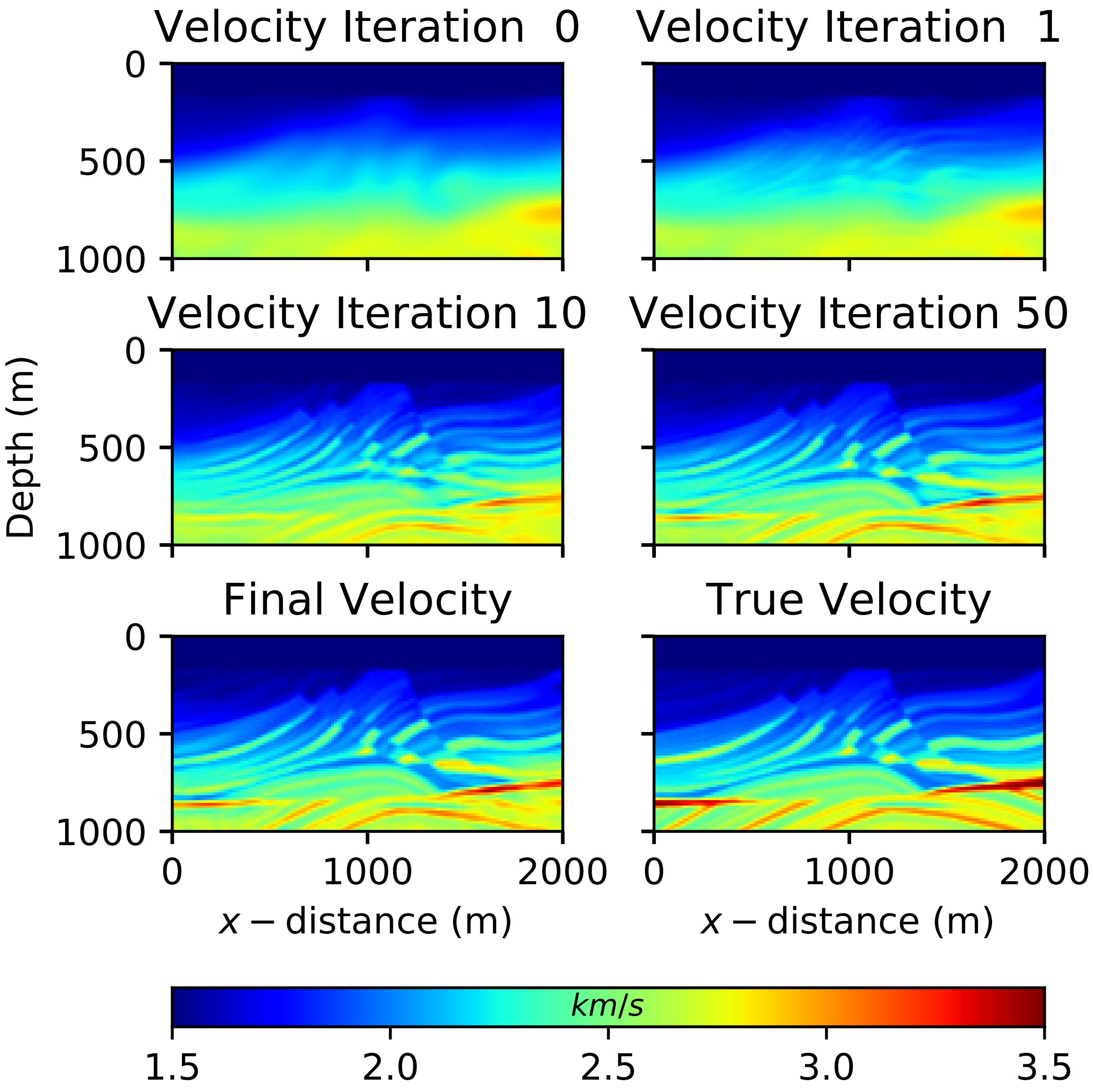


Figure 22. Velocity Marmousi model after some iterations of FWI. The true velocity model is at the bottom-right.

Since FWI seeks to minimize the differences between synthetic and observed data, let us take a look at a single seismogram located at (1000,0) meters on the domain. As seen in Figure 23 the synthetic data (blue line) matches the observed data (black line) even when Gaussian noise has been added (5%).

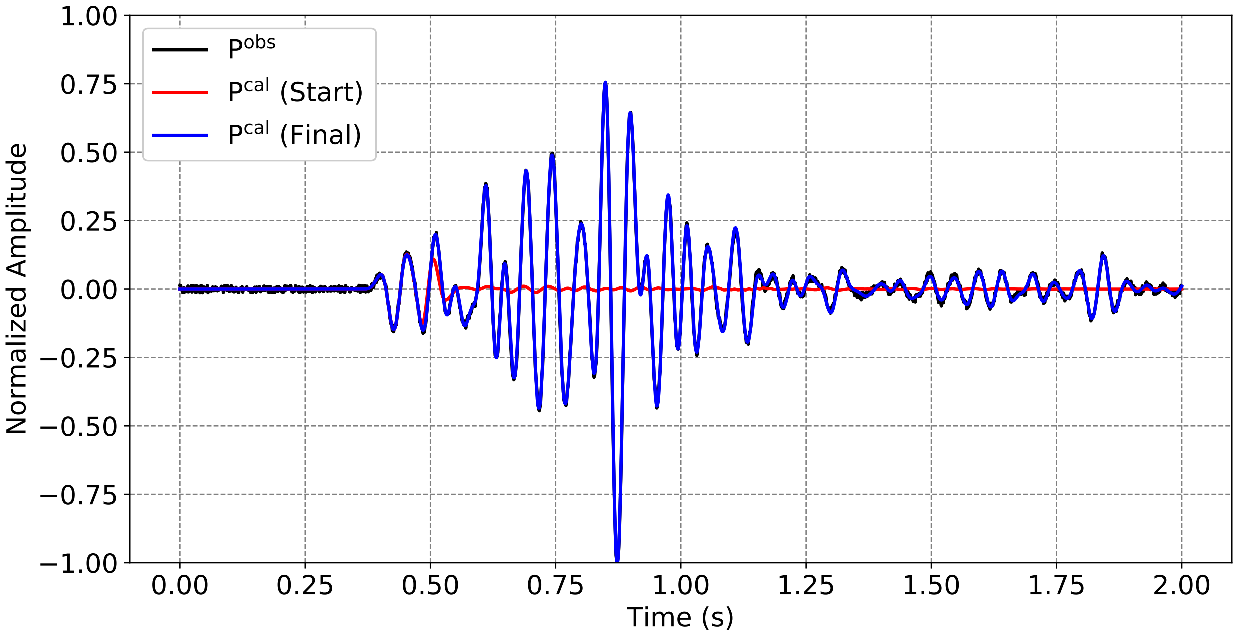


Figure 23. Seismogram comparisson for starting (red) and final (blue) synthetic data with respect to the observed data corresponding to a single source and a single receiver for the Marmousi model FWI example.

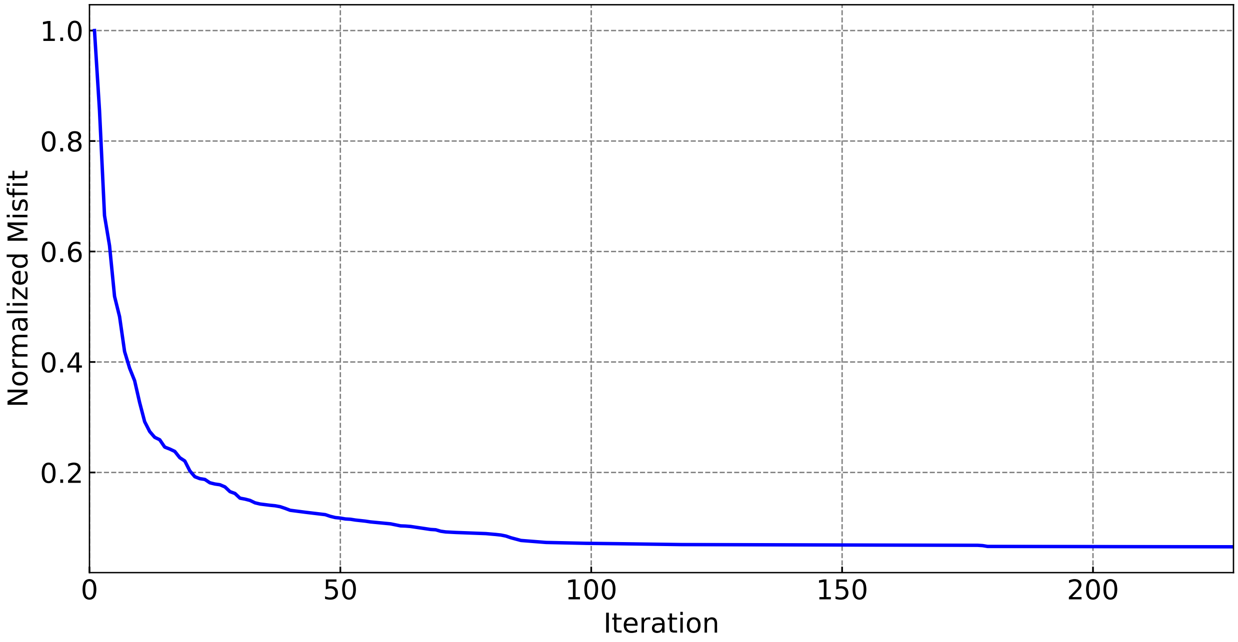


Figure 24. Objective funtion (cost, misfit) reduction for 228 iterations of FWI for the Marmousi model example.

A more accurate indicator for the quality of the FWI iterative process is the analysis of the objective function for each iteration (Figure 24). The objective function for this example is reduced fast at earlies iterations and becomes slower for later iterations, because the stratigraphic information has been recoverer first and, at the end of the process, only the velocity value is getting recovered slowly.

The analysis and results of this section show the concept and elements involved in FWI. This analysis of receivers, sources, frequencies and step line-search on acoustic FWI can be extrapolated to an elastic FWI since only the forward modeling changes but algorithm is the same.

3.4 Elastic Full Waveform Inversion

The objective function for Elastic Full Waveform Inversion (EFWI) is given by

$$Q(m) = \frac{1}{2}\sum_{s}\sum_{r}\int_{0}^{T} ||\mathbf{u}_{r,s}^{\mathrm{obs}} - \mathbf{u}_{r,s}^{\mathrm{cal}}||^{2}\mathrm{d}t, \tag{76}$$

where $\mathbf{u}_{r,s}^{\mathrm{obs}}$ is the observed displacement and $\mathbf{u}_{r,s}^{\mathrm{cal}}$ is the synthetic displacement computed using the Elastodynamic wave equation. T is the total time of recording, r is the receiver index and s is the source index. The displacements can be u_x, u_y and/or u_z (or velocities v_x, v_y, v_z) for a model m which depends on the Lame parameters and density (or velocities v_p and v_s).

3.4.1 Adjoint method for elastic media

Again, the direct minimization of Equation 76 involves the computation of the Frechet derivatives, which increase even more the computational cost for elastic media because the displacement (or velocity) fields are vectors, $u_i = \{u_x, u_y, u_z\}$. The same procedure as AFWI can be pursued using the adjoint-state method and the mathematical deduction of the gradients will not be detailed, since second derivatives, which are self-adjoint operators, are involved. For further details of the adjoint method for elastic media, see Tromp *et al.* (2005).

For an isotropic media 3 gradients can be found. The gradients δ_ρ (density), δ_μ (shear modulus) and δ_κ (bulks modulus) are given by (Tromp *et al.*, 2005)

$$\delta_\rho(\mathbf{x}) = -\sum_r \int_0^T \rho(\mathbf{x})\mathbf{u}^\dagger(\mathbf{x}, T-t)\cdot\partial_t^2\mathbf{u}(\mathbf{x}, t)dt \tag{77}$$

$$\delta_\mu(\mathbf{x}) = -\sum_r \int_0^T 2\mu(\mathbf{x})\mathbf{D}^\dagger(\mathbf{x}, T-t) : \mathbf{D}(\mathbf{x}, t)dt \tag{78}$$

$$\delta_\kappa(\mathbf{x}) = -\sum_r \int_0^T \kappa(\mathbf{x})[\nabla\cdot\mathbf{u}^\dagger(\mathbf{x}, T-t)][\nabla\cdot\mathbf{u}(\mathbf{x}, t)]dt \tag{79}$$

$$\tag{80}$$

where **:** is a double dot product operator between tensors, **D** denotes the deviatoric strain defined by

$$\mathbf{D} = \frac{1}{2}\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right] - \frac{1}{3}(\nabla\cdot\mathbf{u})\mathbf{I}, \tag{81}$$

notice that these computations involve more complex operations than in the case of acoustic media. For the adjoint deviatoric strain $\mathbf{D}^\dagger$ correspond the same equation of **D** but using $\mathbf{u}^\dagger$. The elastic gradient can be expressed in terms of the shear-wave velocity

$$\delta_{v_s}(\mathbf{x}) = 2\left(\delta_\mu - \frac{4}{3}\frac{\mu}{\kappa}\delta_\kappa\right), \tag{82}$$

and the compressional-wave velocity

$$\delta_{v_p}(\mathbf{x}) = 2\left(1 + \frac{4}{3}\frac{\mu}{\kappa}\delta_\kappa\right), \tag{83}$$

then, step line search could be used to obtain the model parameters iteratively.

3.4.2 Elastic banana-doughnut kernel

Following the work of Tromp *et al.* (2005), the source-receiver geometry for an isotropic elastic media with homogeneous properties (Figure 25) is used

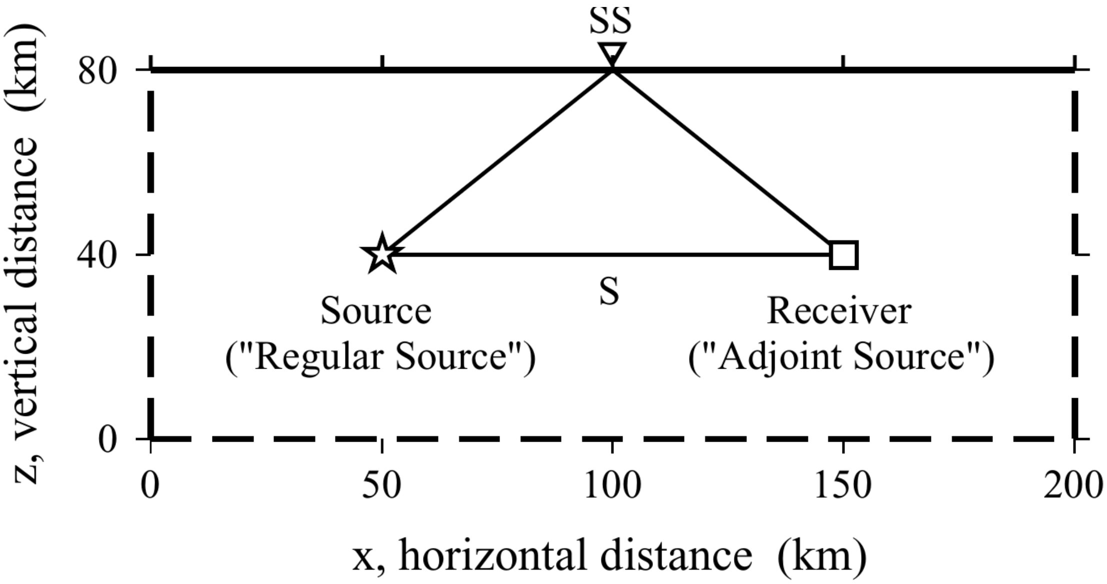


Figure 25. Source - receiver geometry for the computation of the elastic kernels. Taken from Tromp et.al 2005.

Following the same procedure as the previous section for acoustic media, the wave propagation for the horizontal displacement and the back-propagation for the adjoint horizontal displacement is shown in Figure 26 for 52 seconds of recording time. For illustration purposes, the *p*-wave velocity kernel is shown in this example. These results of the gradient construction also present the shape of a banana-doughnut with the same direction of the ray path through the free surface.

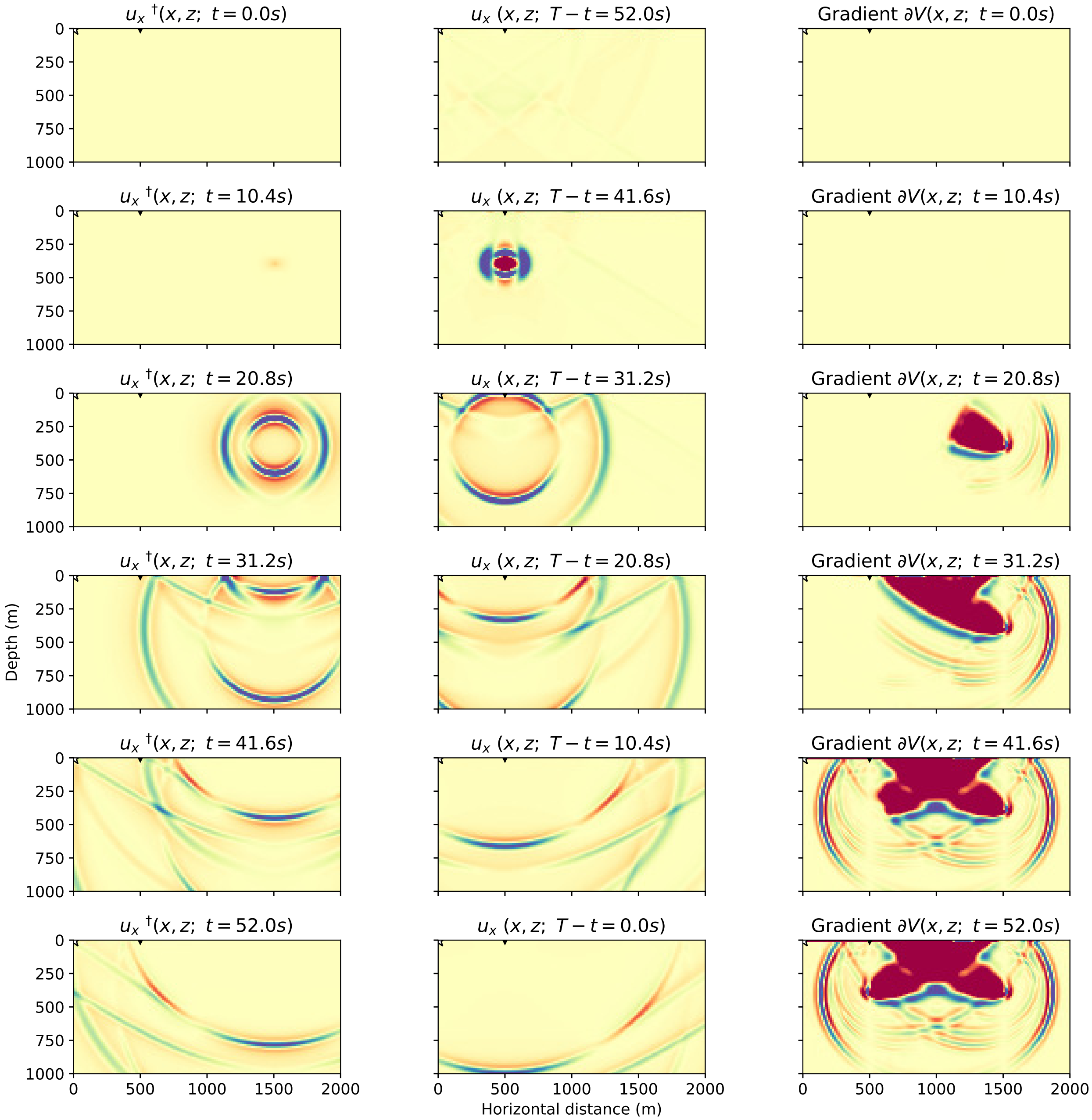


Figure 26. Regular displacement u_x and adjoint displacement $u_x^{\dagger}$ wave propagation for 52 seconds of recording time for the construction of the *p*-wave velocity kernel.

3.5 Conventional joint inversion

In a conventional joint-inversion scheme, different geophysical forward problems are simultaneously solved to obtain a consistent Earth-property model that matches the respective data sets measured at the surface. Usually, the strategy consists of combining all the parameters into one objective function, leading to a large system of often disparate parameters (e.g. Roy *et al.*, 2005).

Let us consider two arbitrary geophysical data sets $\mathbf{d}_m^{\text{obs}}$ and $\mathbf{d}_n^{\text{obs}}$ for the models $\mathbf{m}$ and $\mathbf{n}$. The generalized objective function involving two geophysical methods is given by

$$\begin{aligned} Q_{\text{Total}}(\mathbf{m},\mathbf{n}) &= ||\mathbf{d}_m^{\text{obs}} - \mathbf{d}_m^{\text{cal}}||^2 + \alpha_m^2 ||\nabla \mathbf{m}||^2 + \beta_m^2 ||\mathbf{m} - \mathbf{m}_{\text{apr}}||^2 \\ &= ||\mathbf{d}_n^{\text{obs}} - \mathbf{d}_n^{\text{cal}}||^2 + \alpha_n^2 ||\nabla \mathbf{n}||^2 + \beta_n^2 ||\mathbf{n} - \mathbf{n}_{\text{apr}}||^2 \\ &+ \gamma^2 Q_{\text{Joint}}(\mathbf{m},\mathbf{n}), \end{aligned} \tag{84}$$

where α_m and α_n are the Tikhonov regularization parameters, β_m and β_n are the *a priori* model parameter, $\mathbf{d}_m^{\text{cal}}$ and $\mathbf{d}_n^{\text{cal}}$ are the synthetic data and $Q_{\text{Joint}}(\mathbf{m},\mathbf{n})$ is the joint inversion constraint between both models with its respective weight γ.

There are mainly three types of joint inversion techniques, depending on the functional Q_{Joint}:

- Petrophysical joint inversion, where the models are constrained by an empirical relationship (e.g. Roy *et al.*, 2005; Lin and Zhdanov, 2017; Blom *et al.*, 2017), in this case $Q_{\text{Joint}} = ||\mathbf{m} - \mathbf{H}(\mathbf{n})||^2$, where $\mathbf{H}$ is the petrophysical relation;
- Structural joint inversion (Gallardo and Meju, 2003, 2004), where the functional is used to match the structure for both models trough the cross gradient, $Q_{\text{Joint}} = ||\nabla \mathbf{m} \times \nabla \mathbf{n}||^2$; and
- Statistical joint inversion, *e.g.* using the fuzzy *c*-means technique (Paasche and Tronicke, 2007; Romero and Gallardo, 2015).

We will focus on the petrophysical joint inversion. Structural and statistical joint inversion techniques are beyond the scope of this work and will be explored in future work.

In order to compare conventional joint inversion with the cooperative inversion proposed in this work, let us minimize equation 84 subject to an arbitrary petrophysical relationship $\mathbf{m} = \mathbf{H}(\mathbf{n})$. Using a Gauss-Newton optimization we obtain the following system.

$$\begin{bmatrix} \mathbf{A}_m^T\mathbf{A}_m + \alpha_m^2\mathbf{D}^T\mathbf{D} + \beta_m^2\mathbf{I} + \gamma^2\mathbf{I} & -\Theta\gamma^2 \\ -\Theta\gamma^2 & \mathbf{A}_n^T\mathbf{A}_n + \alpha_n^2\mathbf{D}^T\mathbf{D} + \beta_n^2\mathbf{I} + \Theta^2\gamma^2 \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ \mathbf{n} \end{bmatrix} = \begin{bmatrix} \gamma^2\left[\mathbf{H}(\mathbf{n}_0) - \Theta\mathbf{n}_0\right] + \beta_m^2\mathbf{m}_{\text{apr}} + \mathbf{A}_m^T\left[\mathbf{d}_m^{\text{obs}} - \mathbf{d}_m^{\text{cal}}(\mathbf{m}_0) + \mathbf{A}_m\mathbf{m}_0\right] \\ -\Theta\gamma^2\left[\mathbf{H}(\mathbf{n}_0) - \Theta\mathbf{n}_0\right] + \beta_n^2\mathbf{n}_{\text{apr}} + \mathbf{A}_n^T\left[\mathbf{d}_n^{\text{obs}} - \mathbf{d}_n^{\text{cal}}(\mathbf{n}_0) + \mathbf{A}_n\mathbf{n}_0\right] \end{bmatrix}, \tag{85}$$

where $\mathbf{A}_m$ and $\mathbf{A}_n$ are the Frechet derivatives for each methods, $\mathbf{D}$ is the discrete operator for the gradient ∇, $\mathbf{m}_0$ and $\mathbf{n}_0$ are the starting models (obtained from the linearized problem) and $\Theta = \frac{\partial \mathbf{H}}{\partial n}$ is the derivative of the petrophysical relationship evaluated at the starting model $\mathbf{n}_0$. Notice that if $\gamma = 0$, equation 85 is reduced to

$$\begin{bmatrix} \mathbf{A}_m^T\mathbf{A}_m + \alpha_m^2\mathbf{D}^T\mathbf{D} + \beta_m^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_n^T\mathbf{A}_n + \alpha_n^2\mathbf{D}^T\mathbf{D} + \beta_n^2\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ \mathbf{n} \end{bmatrix} = \begin{bmatrix} \beta_m^2\mathbf{m}_{\text{apr}} + \mathbf{A}_m^T\left[\mathbf{d}_m^{\text{obs}} - \mathbf{d}_m^{\text{cal}}(\mathbf{m}_0) + \mathbf{A}_m\mathbf{m}_0\right] \\ \beta_n^2\mathbf{n}_{\text{apr}} + \mathbf{A}_n^T\left[\mathbf{d}_n^{\text{obs}} - \mathbf{d}_n^{\text{cal}}(\mathbf{n}_0) + \mathbf{A}_n\mathbf{n}_0\right] \end{bmatrix}, \tag{86}$$

which corresponds to the separated inversion of each data set, since there is no constraint in the objective function, and the inverse of the matrix can be obtained by blocks. For this work, we consider that $\mathbf{m}$ is the seismic model and $\mathbf{n}$ is the gravimetric model. The Frechet derivatives of the seismic functional are given by $A_m(i,j) = \frac{\partial P_i}{\partial m_{V_j}}$, which correspond to the i-th derivative for the data P_i for the j-th model parameter m_{V_j}. Notice that this results in large computational cost given that it implies performing forward modelings for each grid cell of the model, as opposed to the two forward modelings required by the adjoint method for FWI. For the gravimetric inversion, $\mathbf{A}_n = \mathbf{A}_g$. Finally, for the joint inversion, the derivative of the petrophysical function is given by

$$\Theta = 4\left(\frac{1}{0.31}\right)^4 \mathbf{m}_\rho^3\Big|_{\mathbf{m}_\rho=\mathbf{m}_{\rho 0}}, \tag{87}$$

which corresponds to the derivative of Gardner's equation with respect to the density.

3.6 Cooperative inversion

In cooperative inversion, we propose a sequential approach in which we solve at different stages for the densities and velocities, the resulting system is, therefore, more manageable and there is more control over the parameters at each stage. We call this a cooperative strategy to distinguish it from the joint strategies that solve all the geophysical parameters together at every iteration. Unlike conventional joint inversions, where the problem is to minimize a two-part objective function (seismic and gravity errors), this cooperative inversion is based on alternately minimizing the errors in seismic and gravity data iteratively. The main reasons to perform these sequentially are to increase robustness and reduce the computational cost, and to keep always a strong control in the gravimetric inversion, avoiding the natural behavior of this potential method to yield shallower models.

Furthermore, in the proposed scheme we do not need to impose depth-dependent weights or constraints to the gravimetric inversion to avoid shallower models, this is achieved instead by using the velocity model after FWI as the *a priori* gravimetric model.

We seek to minimize the gravimetric data constrained with the velocity model obtained after an FWI process using the following objective function

$$Q(\mathbf{m}_\rho) = \sum_{i=1}^{N_s} \left\| \frac{\mathbf{g}_{z_i}^{\text{obs}} - \mathbf{A}\mathbf{m}_\rho}{\sigma_{\mathbf{g}_{z_i}}} \right\|^2 + \alpha_{\text{reg}}^2 ||\nabla \mathbf{m}_\rho||^2 + \beta^2 ||\mathbf{m}_\rho - \mathbf{m}_{\rho(V)}||^2, \tag{88}$$

where $\mathbf{m}_\rho$ is the density model obtained using a petrophysical relationship as a function of the velocity model obtained from iterative AFWI or EFWI. β is the parameter that weights the roll on the inversion of seismic versus gravimetric inversion; Higher values of β obeys more the seismic contribution and vice versa. We use the following relationship from Gardner *et al.* (1974) as petrophysical constraint,

$$\rho = \rho_0 V_P^{k_0}, \tag{89}$$

with ρ_0 = 0.31 g/cm^3 and k_0 = 0.25. Brocher (2005) compute a polynomial fit and compares the curve with Gardner relationship for density in function of velocity,

$$\rho(\text{g/cm}^3) = 1.6612 V_P(\text{km/s}) - 0.4721 V_P^2 + 0.0671 V_P^3 - 0.0043 V_P^4 + 0.000106 V_P^5, \tag{90}$$

and velocity as a function of density

$$V_P = 39.128\rho - 63.064\rho^2 + 37.083\rho^3 - 9.1819\rho^4 + 0.8228\rho^5, \tag{91}$$

valid for densities between $2.0 < \rho < 3.5$(g/cm^3) and velocities in the range $1.5 < V_P < 8.5$(km/s) respectively. Since both Brocher's equations are based on polynomial fit and are not inversely related, an iterative procedure using equations 90 and 91 will not lead to the same velocity-density value. As an example, starting from a velocity of 3500 m/s, using equation 90 a density of 2.318 g/cm^3 is computed, then using equation 91 to recover the initial velocity we obtain a value of 3692.34 m/s, a change of 192.34 m/s (5.49%) only for a single performance. Therefore, since we requiere that the two functions be inverse of each other, we will use Gardner's equation.

Concerning the β parameter, our results will focus more on the velocity model for FWI to avoid bad shallower models due to a weakly-restricted gravimetric inversion. Then the density model will give feedback to the velocity model using an empirical relationship for the next FWI iteration. The CGLS method is implemented in a straightforward way modifying **G** and the **d** matrix from equations 55 and 56 as follows

$$\mathbf{G} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2}\mathbf{A} \\ \alpha_{\text{reg}}\mathbf{D} \\ \beta\mathbf{I} \end{bmatrix} \tag{92}$$

$$\mathbf{d}^{\text{CG}} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2}\mathbf{g}_z^{\text{obs}} \\ 0 \\ \beta\mathbf{m}_{\rho(V_P)} \end{bmatrix}, \tag{93}$$

where **I** is the identity matrix. Once again, the matrix **G** is stored in a large sparse format to compute the constrained CGLS. It is worth recalling that regardless of the model obtained when properly fitting a gravity anomaly, the total mass is uniquely recovered as implied by Gauss' theorem (Grant and West, 1965). This means that although gravity is a low-resolution geophysical tool, it does provide unique information linked to the velocity model.

```
**Data:**
$m_0$, $d_0 = d^{\text{CG}}$, $r_0 = G^T d^{\text{CG}}$
$p_0 = r_0$, $t_0 = Ap_0$, $n$-iterations
**Result:** model $m_k$
while $k < n$ or stop criteria is satisfied do
    $\alpha_k = ||r_{k-1}||^2 / ||t_{k-1}||^2$
    $m_k = m_{k-1} + \alpha_k p_{k-1}$
    $d_k = d_{k-1} - \alpha_k t_{k-1}$
    $r_k = G^T d_k$
    $\beta_k = ||r_k||^2 / ||r_{k-1}||^2$
    $p_k = r_k + \beta_k p_{k-1}$
    $t_k = G p_k$
    $k \leftarrow k + 1$
end
```

Algorithm 4: CGLS algorithm to iteratively solve the problem $Gm = d^{\text{CG}}$

The algorithm to solve the system $\mathbf{Gm} = \mathbf{d}^{\text{CG}}$ is presented in algorithm 4 (Sen and Stoffa, 2013). Notice that an efficient implementation of this algorithm requires that all the matrices be stored in a sparse representation. The resulting density model will give feedback to FWI using Gardner's density-velocity relationship to obtain a velocity model which will be the starting model to solve the next iteration of FWI. In practice, including the gravimetric data in the inversion process yields smoother models with fewer artifacts.

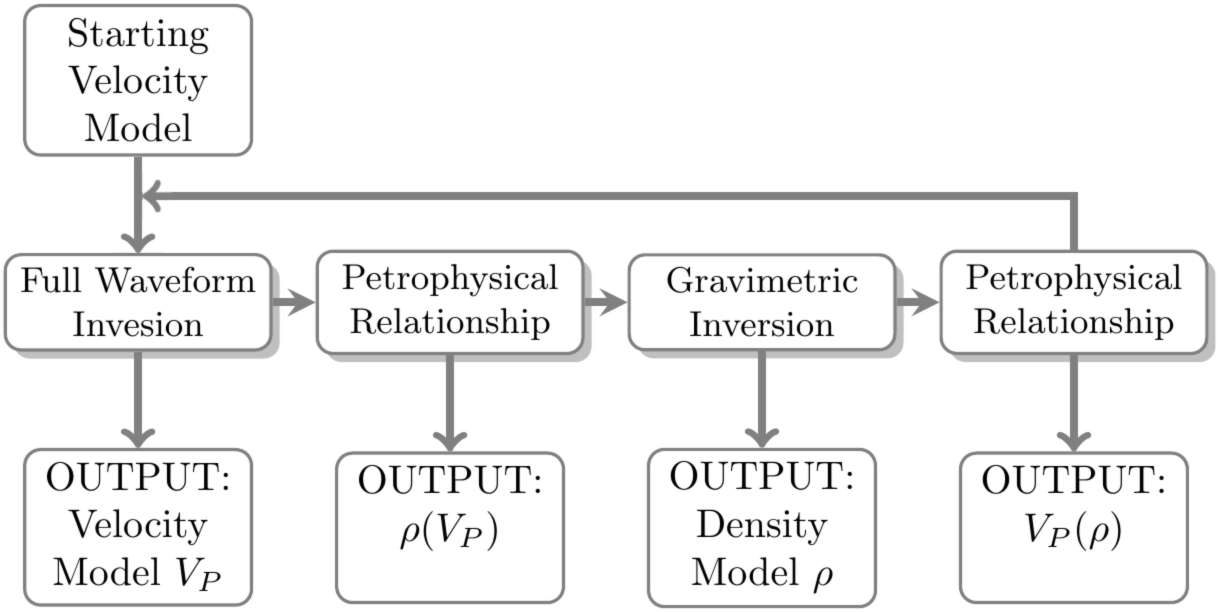


Figure 27. Visual representation of our iterative inversion scheme for gravity and seismic data. From a starting velocity model, we perform FWI to update the velocity model, then, using Gardner's density-velocity relationship, we perform constrained gravimetric inversion to update the density model, finally, using Gardner's velocity-density relationship, a velocity model is obtained which will be the starting model to solve FWI. This process is performed iteratively.

In summary, this cooperative inversion scheme for gravity and seismic data consist on an iterative scheme as shown in the diagram of Figure 27 containing the following steps: From a starting velocity model, we perform FWI to update the velocity model, then, using Gardner's density-velocity relationship, we perform constrained gravimetric inversion to update the density model, finally, using Gardner's velocity-density relationship, a velocity model is obtained who will be the starting model to solve FWI, performing iteratively this steps.

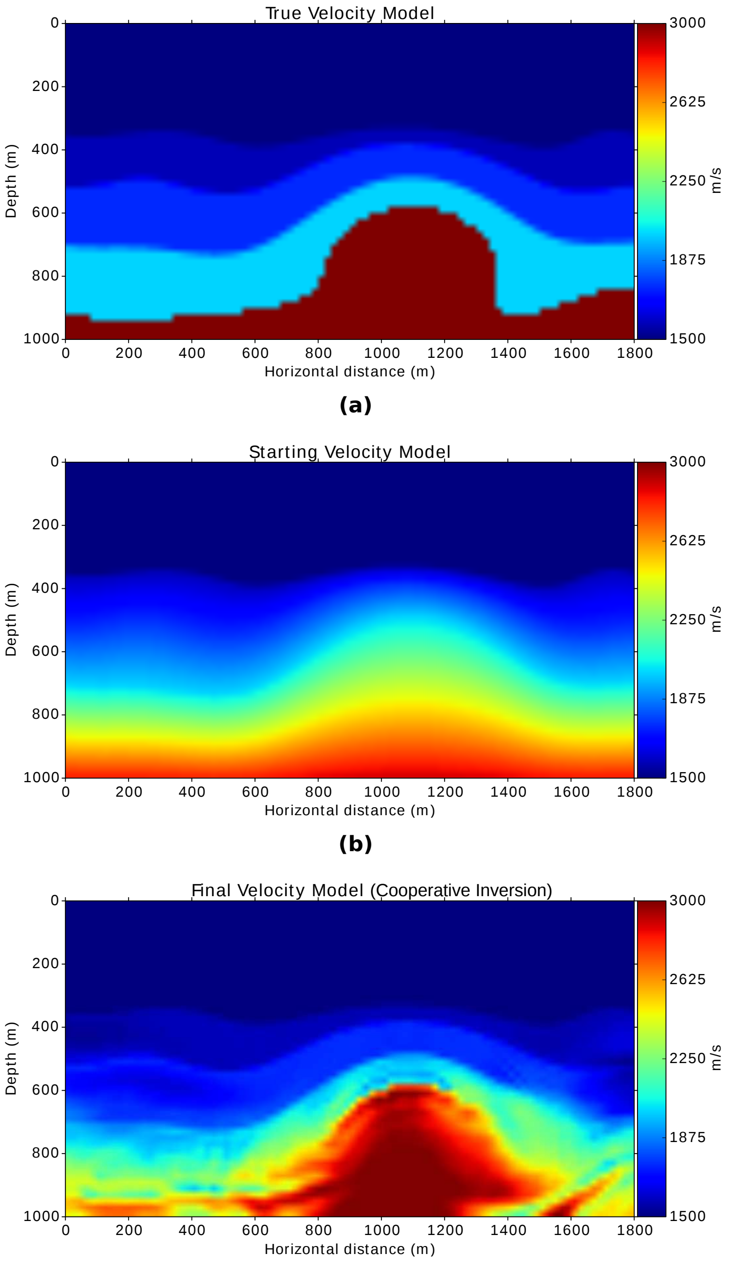
Chapter 4. Results: Acoustic media

In order to test the proposed cooperative inversion algorithm and show its advantages, we apply this method on several synthetic data sets for gravimetric and seismic data for acoustic media. The first example is a simple layered model with an anticline where here is a comparison with conventional methods. The second model is a Texas-shaped model; in this example, the separated, conventional joint and sequential inversion is discussed. The third model has a similar structure as model 2 with more resolution. In the fourth example, we add random noise to the data and demonstrate that the method can recover the model under these conditions. In the fifth example, we apply the method to the SEG SEAM phase 1 model. Then, 2 additional synthetic models are explored based on the geological structure of over-thrust and dominoes fault systems. Finally an example of 3D acoustic FWI and an example of 3D cooperative inversion are explored.

4.1 Test layered Model

The first model contains a layered part on top of the model and an anticline at the bottom of the domain (Figure 28a). The shallower layers have lower velocity (1500-2000 *m/s*) and the anticline has a velocity of 3000 *m/s*. The true density model is presented in Figure 29a, this model satisfies Gardner's relation (equation 89). We discretize the synthetic geophysical model covering a horizontal distance of 1800 *m* and a depth of 1000 *m* for both seismic and gravimetric data. For test purposes, we use a small grid in this example ($n_x = 90$ and $n_z = 50$). We modeled 20 sources equally spaced along the surface and recorded at 90 receivers also along the surface for the seismometers and gravimetric stations. The total recorded time for the traces is 2 *s* and the time sampling depends on the stability condition of the forward modeling. The source time function is a Ricker wavelet with a peak frequency of 15 Hz distributed spatially with a Gaussian function.

The inversion results and starting model are shown in Figures 28 and 29. The

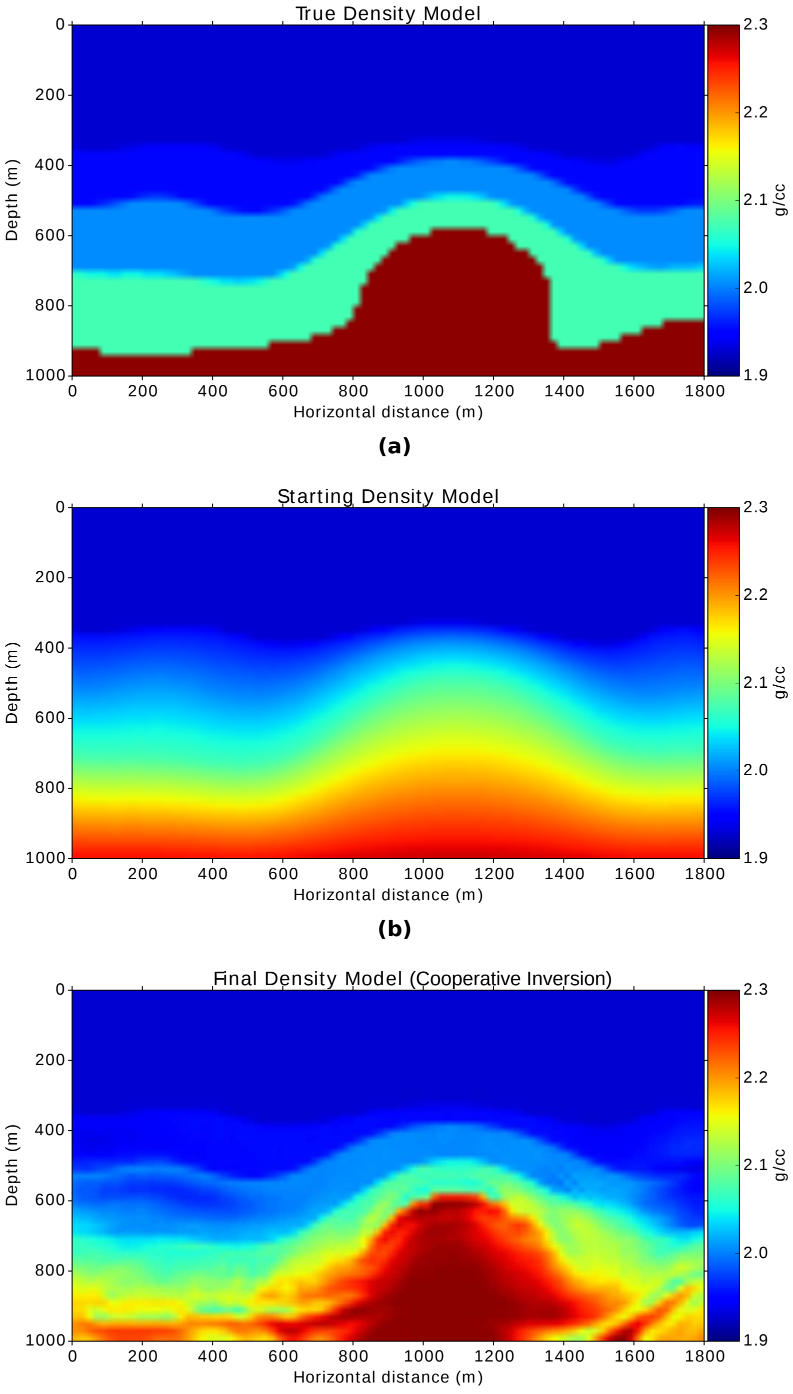


(a)

(b)

(c)

Figure 28. (a) Test layered true velocity model and (b) the starting model. (c) The velocity model after 1000 iterations of cooperative inversion.



True Density Model

Starting Density Model

Final Density Model (Cooperative Inversion)

(a)

(b)

(c)

Figure 29. (a) Test layered true density model and (b) the starting model. (c) The density model after 1000 iterations of cooperative inversion.

starting model (Figure 28b for velocity and Figure 29b for density) is a layered model smoothed in the vertical direction and the water layer is assumed to be known. The inversion parameters were chosen empirically following the L-curve method (Hansen, 1992), and are given by $\alpha_{reg} = 1.96 \times 10^{-8}$ and $\beta = 1.39 \times 10^{-8}$; these values ensure more weight on FWI and a proper smoothness on the geological model. The final results after 1000 iterations using FWI and gravimetric inversion are shown in Figures 28c and 29c for velocity and density models respectively. The results show that the interfaces are recovered correctly, except for small parts near the edges of the domain where the model is expected to be poorly resolved. The convergence of the seismic and gravity misfit is shown in Figure 30. The gravimetric method displays a faster convergence at an earlier stage, but a slower convergence in the late iterations.

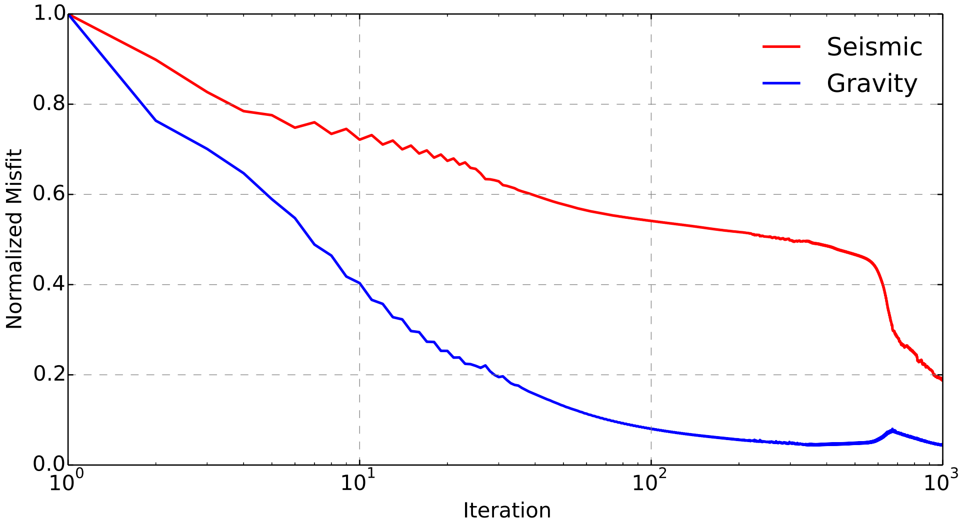


Figure 30. Normalized misfit for seismic (red) and gravity (blue) data in logarithmic scale.

In Figure 31 we plot the pressure seismograms corresponding to the observed data and the initial and final models from a source located at 450 *m* and a receiver at 560 *m* on the surface in the interval 0.5 to 1.5 *s*. The synthetic pressure (red) due to the starting model is not close to the real pressure (black) and some events shown opposite amplitude with respect to the true pressure (For example at $t = 0.93$ *s*). On the other hand, the synthetic pressure (Blue line) as a result of 1000 iterations shows a good agreement with respect to the observed one, the match in phase is very good.

Concerning gravity data, we selected the stations between 600 and 1200 *m* and plot the observed (black), the starting (red) and the final (blue) gravimetric responses

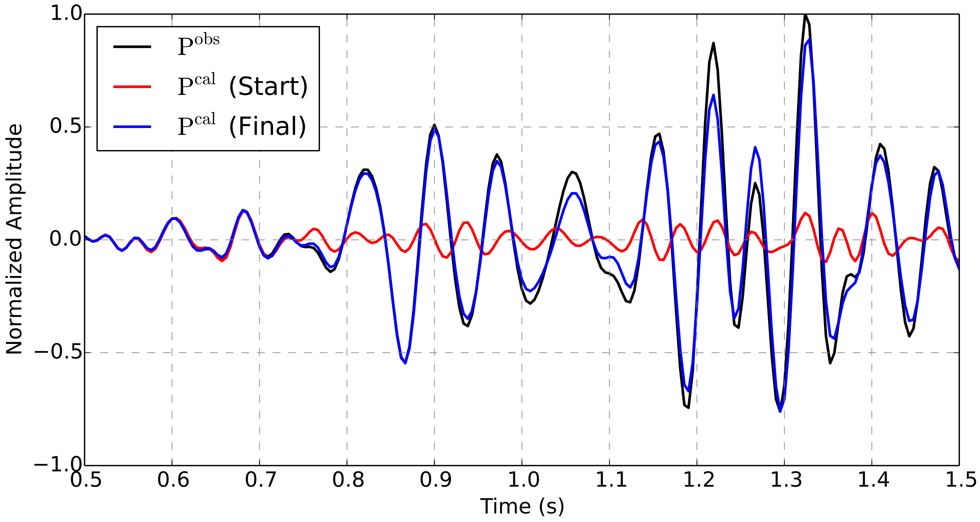


Figure 31. Comparison of observed (black), initial synthetic (red) and final synthetic traces (blue) measured at 560 *m* due to a source applied at 450 *m* on the surface.

in Figure 32. Consistent with the misfit reduction, the gravimetric response after 1000 iterations fits the observed gravity data.

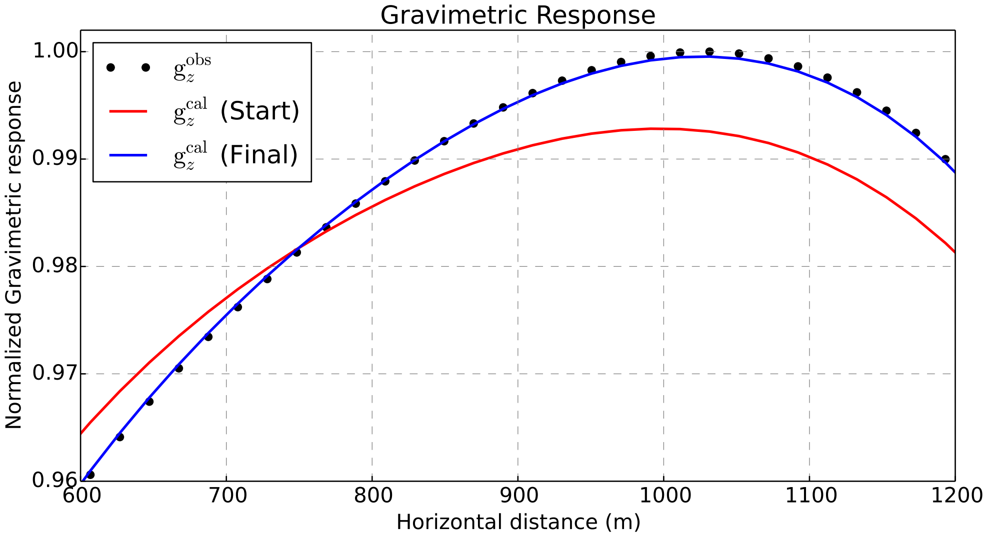
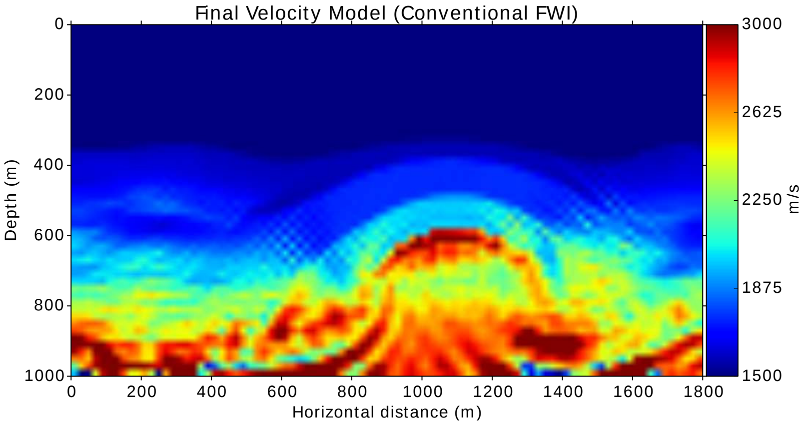


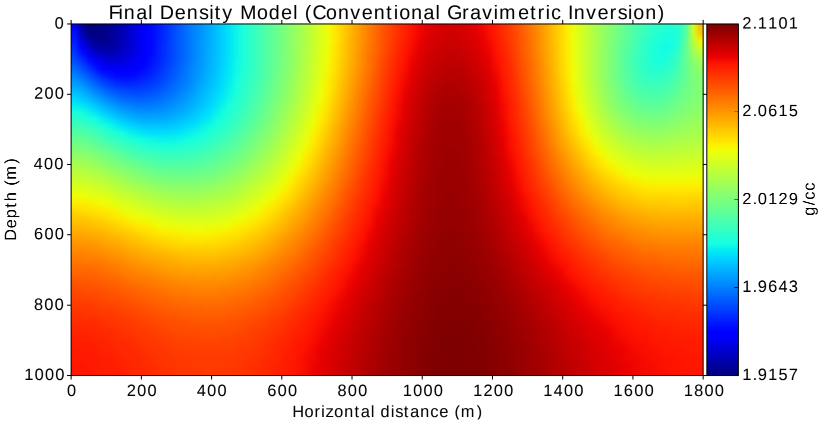
Figure 32. Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 1000 iterations of our scheme.

4.1.1 Comparison with respect to independent inversions

We now proceed to compare the results of the cooperative inversion with those of inverting separately the seismic and gravimetric data. Beginning with the same starting velocity model (Figure 28b), we perform conventional FWI the same number of iterations using the same array of sources-receivers from the previous problem. The final velocity model (Figure 33a) recovers the layer interfaces, but the central region of the anticline lacks resolution.



(a)

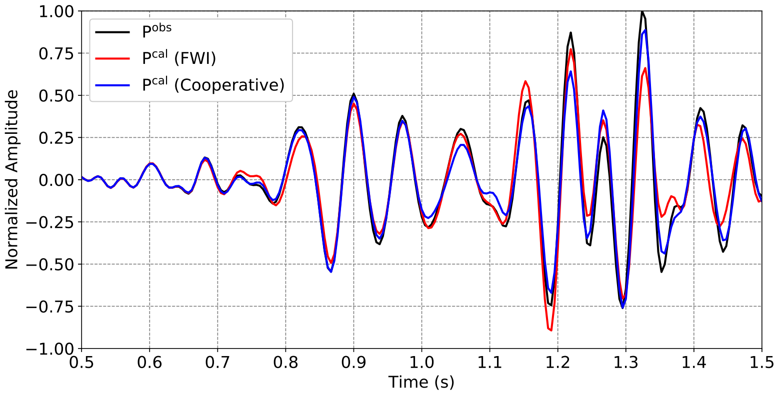


(b)

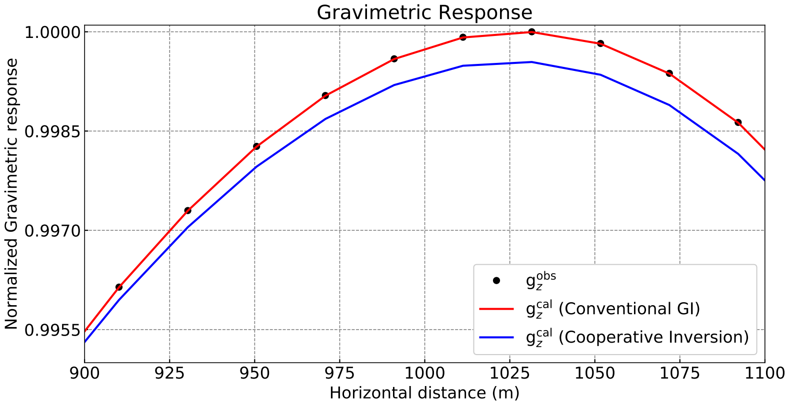
Figure 33. Model result after separated data inversions for conventional FWI (a) and gravimetric inversion (b).

On the other hand, performing conventional gravimetric inversion applying CGLS and using the same parameters as before results in a density model (Figure 33b) which completely mismatches the stratigraphic information of the true model from Figure 29a. This behavior is characteristic and well known for this potential method, due to the natural decay of the kernels. Several papers solve this issue using a depth-weighting function or constrain in the objective function (Li and Oldenburg, 1998; Boulanger and Chouteau, 2001). Instead of this, in our cooperative scheme, the whole velocity model acts as a constrain for the gravimetric inversion.

We finally proceed to analyze the data fitting for each method. First, we select the same receiver position as in Figure 31 and plot the trace due to the velocity model of Figure 33b obtained after conventional FWI, then we compare it with the result obtained from cooperative inversion, as shown in Figure 34a. The phase and amplitude match is very good, however, the model reached after 1000 iterations do not resemble the true velocity model as good as the result using cooperative inversion. Finally, we compare the gravimetric response using conventional gravimetric inversion and our cooperative inversion scheme. For this purpose, we chose the stations located in a range of 900 to 1100 *m* since the previous results were already close to the observed data (see Figure 32). As expected, the response using only gravimetric inversion fits better the observed data in comparison to the response from the cooperative inversion (Figure 34b), although the resulting model does not resemble the true model.



(a)

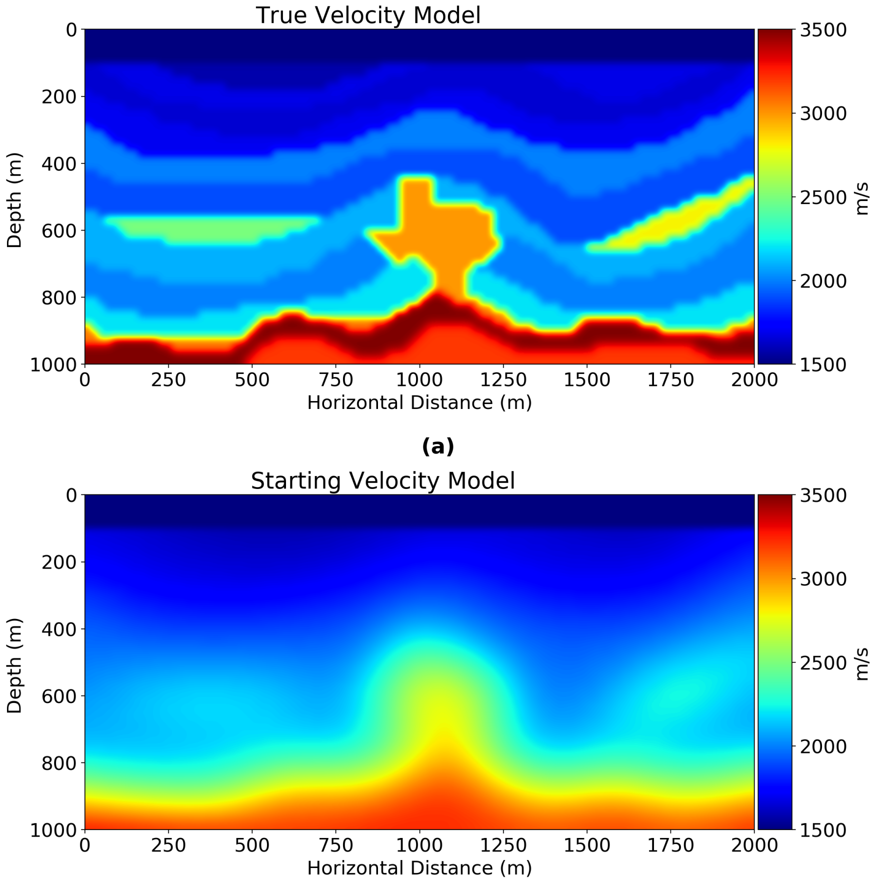


(b)

Figure 34. Geophysical response after separated data inversions for conventional FWI (a) and gravimetric inversion (b).

4.2 Texas-shape model I

For this example, we created a laterally heterogeneous layered model. Beneath the low-velocity layers, we place a structure with the shape of Texas, as shown in Figure 35a. The shallow layers have lower velocities (between 1500 and 2000 *m/s*) with respect to the deepest layer ($\sim$ 3400 *m/s*). In addition to the high-velocity body, we placed two targets at medium depth to the left and the right of the body. We used Gardner's density-velocity relationship to obtain the density model for gravity data as shown in Figure 36a. Notice that the maximum velocity used for this example is 3500 *m/s*, hence Gardner's equation applies to this example.

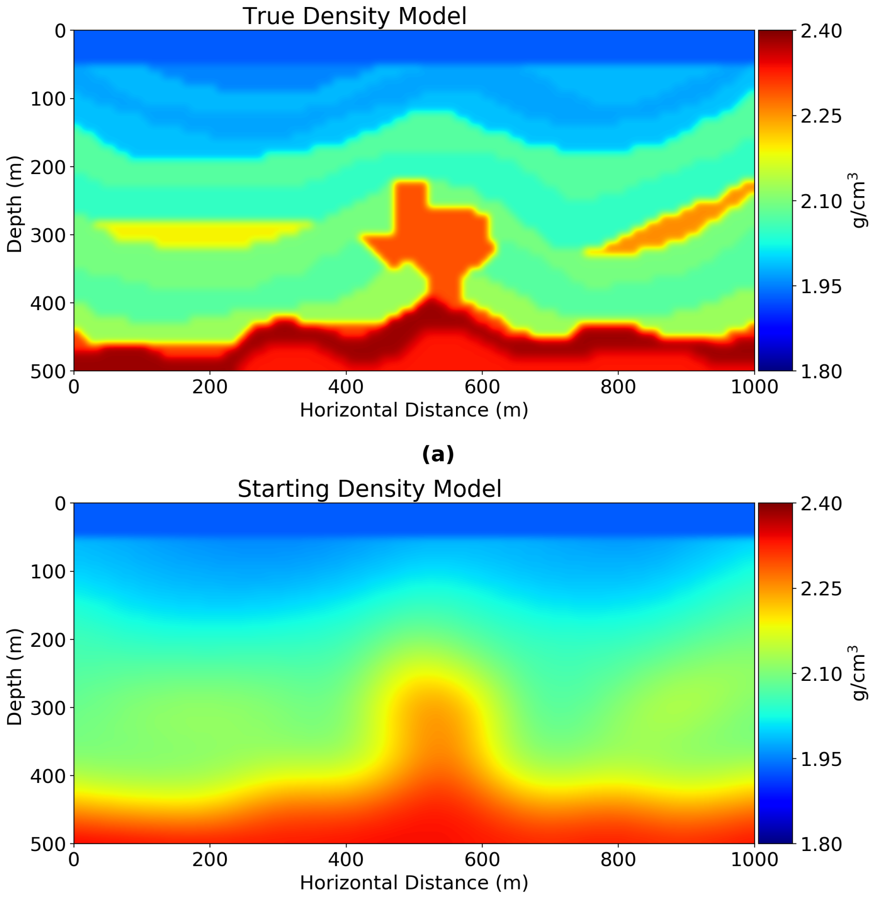


(a)

(b)

Figure 35. Texas-shaped true velocity model I (a) and its smoothing set as a starting model (b).

The synthetic geophysical model covers a horizontal distance of 2000 *m* and a depth of 1000 *m* for both seismic and gravimetric data. To be able to compare the cooperative and joint inversion strategies, we use a small mesh for this model



(a)

(b)

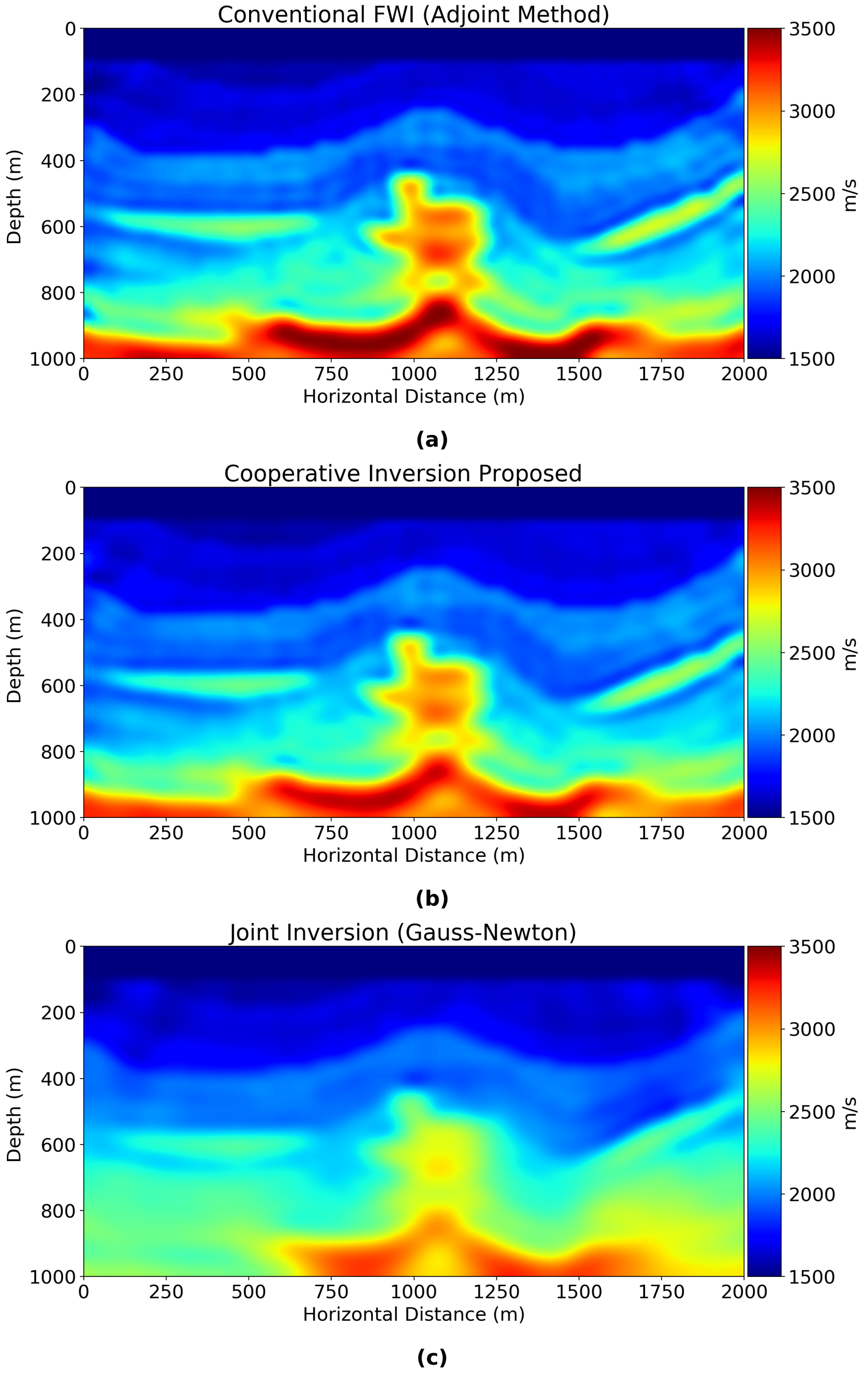
Figure 36. Texas-shaped true density model I (a) and its smoothing set as a starting model (b). These models are obtained using Gardner density-velocity.

($n_x = 100$, $n_z = 50$). We modeled $n_s = 10$ sources equally spaced along the surface and recorded at $n_r = 100$ receivers along the surface for the seismic and gravimetric stations. The total record time for the seismograms is 2 seconds and the time sampling depends on the stability condition of the forward modeling, resulting in $n_t = 750$ samples. The source time function is a Ricker wavelet with a peak frequency of 8 Hz. In summary, we have $n_t \times n_r \times n_s = 750\ 000$ seismic data points and $n_r = 100$ gravimetric data points, and are computing a model of $n_x \times n_z = 5\ 000$ cells for this simple example, therefore, the seismic problem is over-determined whereas the gravimetric problem is under-determined. In this case, the number of seismograms and gravimetric stations is equal. We emphasize that the method does not require that the number of stations has to be the same, however, we use the same location for the receivers for simplicity. Often in practice, the number of gravimetric stations is less than the number of seismic stations, the method can accommodate for this but the reduction in computational cost would be negligible.

For comparison purposes, we use the same parameters and starting models in all the methods and perform 50 iterations. Furthermore, we also include the results of the separate inversions for seismic (FWI using the adjoint method) and gravity (CGLS unconstrained). For the particular case of the conventional joint inversion, the parameter γ was chosen empirically in a similar way as in the L-curve method.

First, let us compare the results for the velocity model (Figure 37). The conventional FWI process recovers the stratigraphic information including the medium-depth targets, as shown in Figure 37a. The shape and velocity values of the high-contrast body are also approximately recovered. This result is similar to the one obtained using the proposed cooperative inversion (Figure 37b). For the conventional joint inversion process (Figure 37c), the general geometry is recovered, however, the basement, the Texas-shaped body and the medium-depth targets have lower velocities and lower resolution.

Figure 39 shows the normalized seismic misfits of the three methods with respect to the iteration number. The joint-inversion scheme stagnates after the 10th iteration, but FWI and the cooperative scheme achieve a smaller misfit.



Conventional FWI (Adjoint Method)

(a)

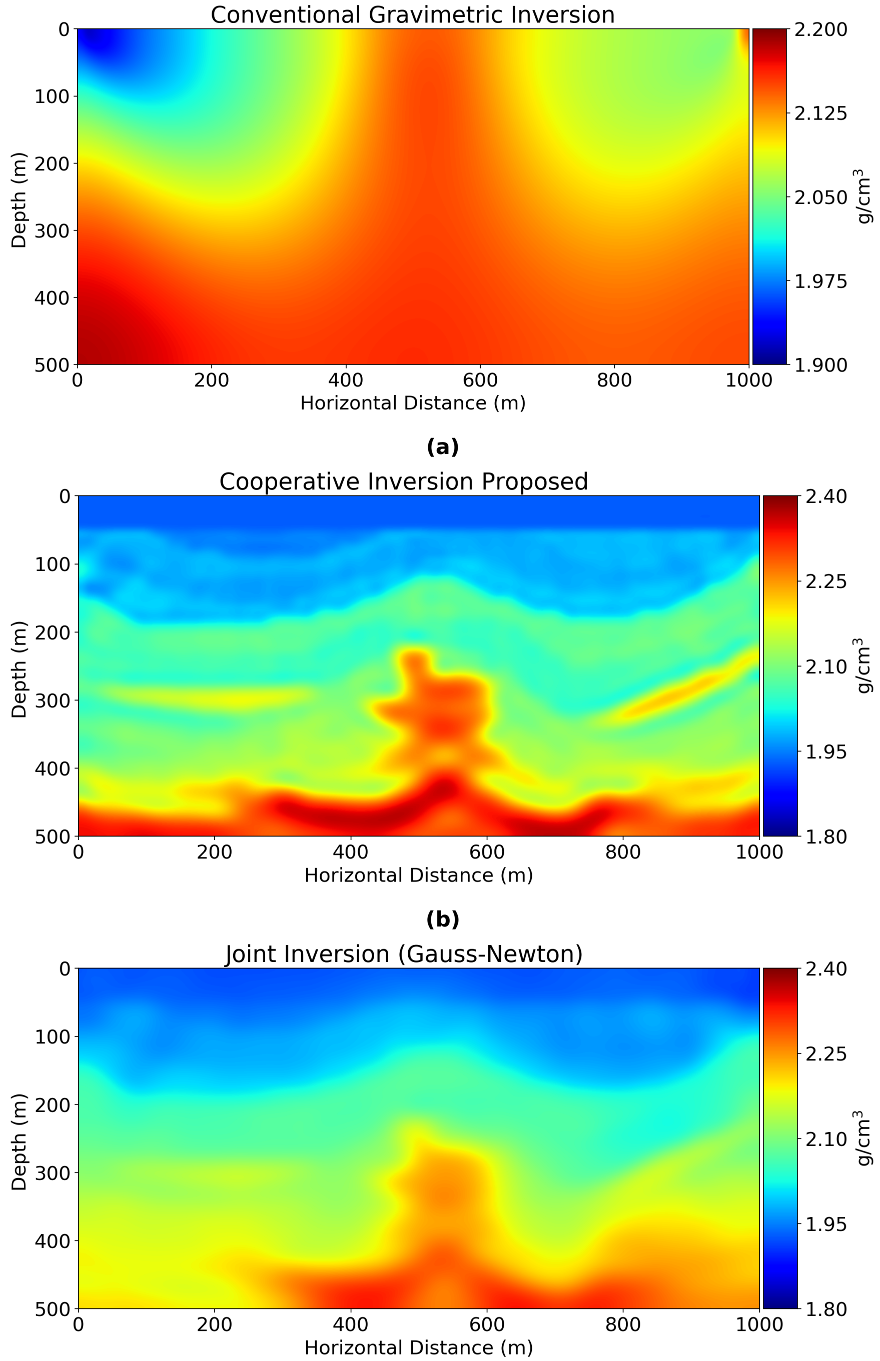
Cooperative Inversion Proposed

(b)

Joint Inversion (Gauss-Newton)

(c)

Figure 37. Texas-shape velocity model I results for (a) conventional FWI using adjoint method, (b) petrophysical joint inversion, and (c) petrophysical cooperative inversion.



(a)

(b)

(c)

Figure 38. Texas-shaped density model I results for (a) conventional gravimetric inversion, (b) petrophysical joint inversion using Gauss-Newton and (c) the cooperative inversion proposed.

The results of the computed densities are shown in Figure 38. Once again, the model losses stratigraphic information when conventional gravimetric inversion is employed (Figure 38a), resulting in a shallower density model. When the conventional joint inversion is applied, the shape and position of the central body are obtained, however, the two targets at the left and right of the model are not identified (Figure 38c). On the other hand, using cooperative inversion (Figure 38b), all the targets are recovered better than with the other two methods.

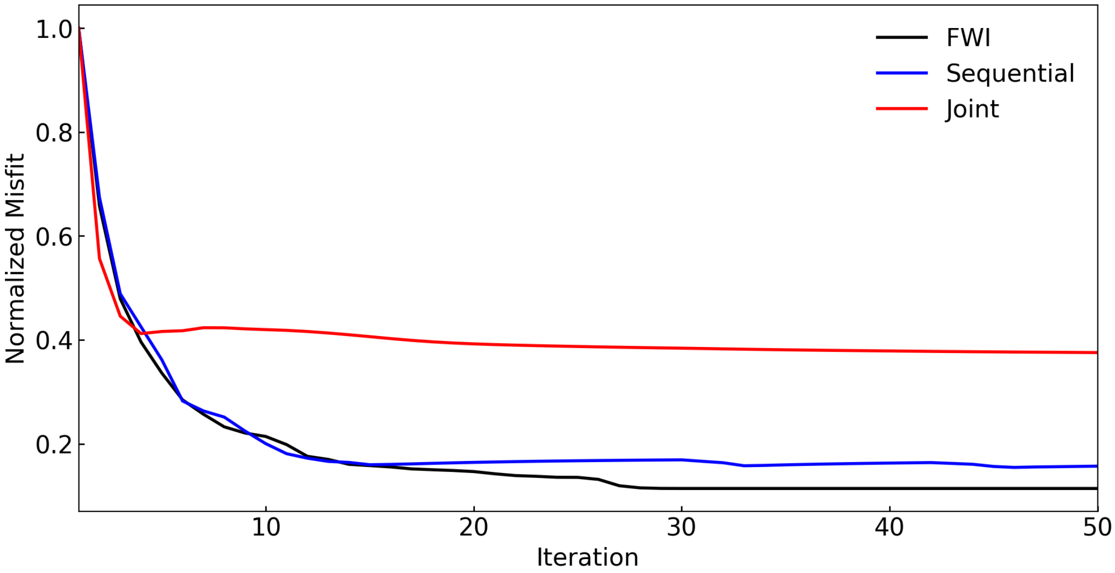


Figure 39. Normalized misfit for seismic (red) and gravity (blue) for the Texas-shaped model.

Table 7. Cost of the objective function and computational cost for 50 iterations for each inversions method discussed. The computational cost is normalized with respect to the cost of conventional FWI.

Inversion Method	Objective Function Cost Reduction	Computational Time (50 Iterations)
Conventional Joint	37.5%	681.63
Conventional FWI	11.4%	1.00
Cooperative	15.7%	1.02

The main difference between the three methods is in the computational cost. Table 7 shows the execution times for performing 50 iterations with each of the 3 methods, normalized using the time of the conventional FWI.

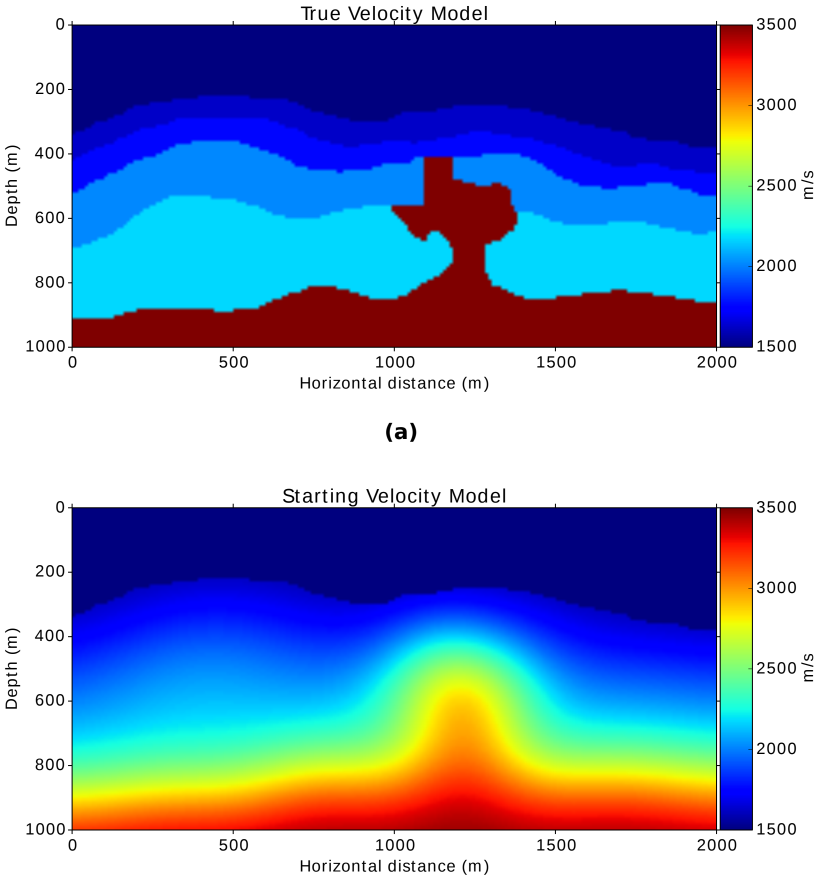
The computational cost for conventional joint inversion using the Gauss-Newton method is significantly higher than that of the other methods, this is mainly due to the computation of the Frechet derivatives. On the other hand, the overhead of the proposed cooperative scheme is comparatively small, amounting to 2% in this example, and the computational cost of inverting the gravimetric data is almost negligible (0.3% of the cost of FWI).

4.3 Texas-shaped model II

For this example, we modify the layered model previously described. Besides the three low-velocity layers, we replace the round body with a structure with the shape of Texas, as shown in Figure 40a. The shallower layers have lower velocity (between 1500 and 2000 *m/s*) with respect to the deepest layer (3500 *m/s*). Henceforth the density model will not be presented to avoid redundancies since it can be obtained from the velocity model by applying Gardner's equation.

We discretize the synthetic geophysical model covering a horizontal distance of 2000 *m* and a depth of 1000 *m* for both seismic and gravimetric data. In this example, we modeled 100 sources equally spaced along the surface and recorded at 182 receivers along the surface for synthetic seismometers and gravimetric stations. The total record time for the seismograms is 2 *s* and the time sampling depends on the stability condition of the forward modeling. The source time function is a Ricker wavelet with a peak frequency of 15 *Hz*.

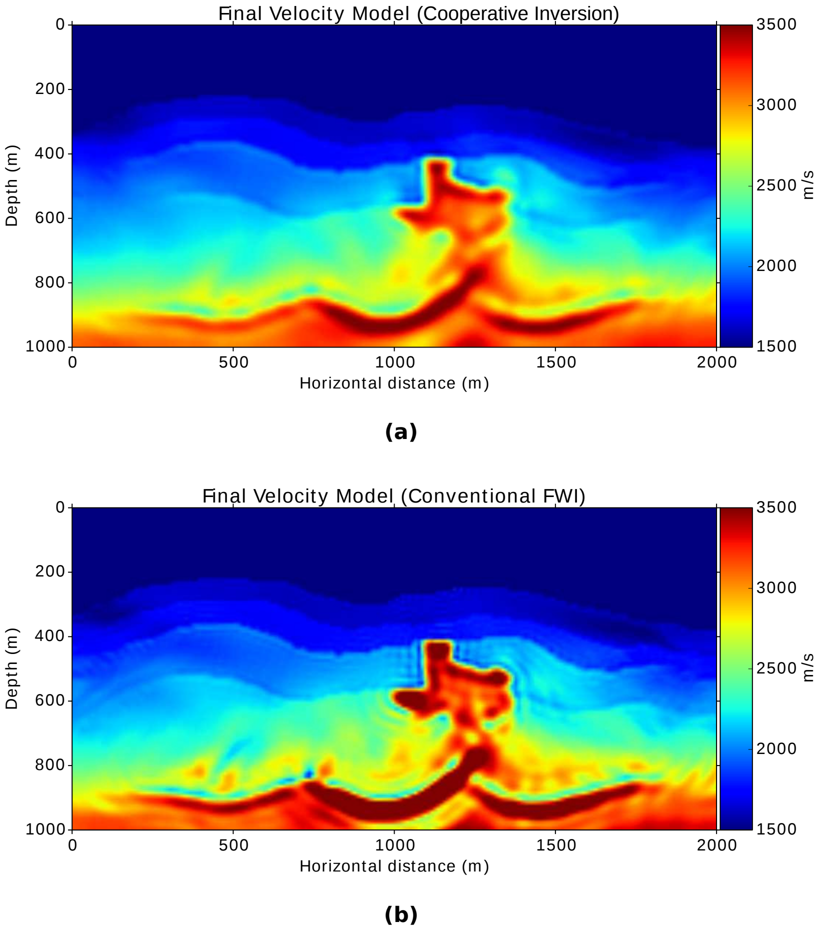
We perform the cooperative inversion algorithm for a maximum number of 250 iterations. The starting model (Figure 40b) is a smooth layered model and the water layer is assumed to be known. We use the same inversion parameters as in the previous example. The final velocity model is shown in Figure 41a. It shows that the interfaces are recovered correctly, except for small parts near the edges of the domain. The highest density-velocity zone of the model is partially recovered including the small curve below the high-velocity structure. We solve the same test model using conventional FWI, the result is shown in Figure 41b. Even though FWI reaches a good velocity model highlighting the stratigraphic part, it presents some artifacts around the high-velocity structure, which are not present in the result of our cooperative inversion. The



(a)

(b)

Figure 40. Texas-shaped true velocity model (a) and its smoothing set as a starting model (b).



(a)

(b)

Figure 41. Texas-shaped model results using cooperative inversion (a) and conventional FWI (b).

convergence of the cooperative inversion is shown in Figure 42. The gravimetric inversion shows a volatile behavior that is very obvious in the first 45 iterations because the algorithm gives more weight to the seismic part, however, the two curves reach convergence.

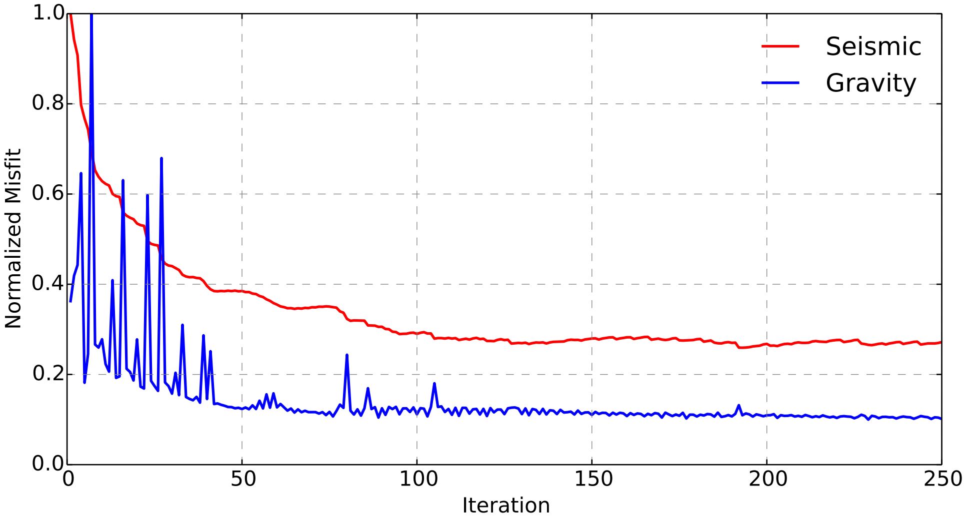


Figure 42. Normalized misfit for seismic (red) and gravity (blue) for the Texas-shaped model.

4.4 Texas-shaped model III with noisy data

For this example, we modify the Texas-shaped model from the previous example by increasing its complexity. We added more heterogeneities to the low-velocity layers and faulting in some regions of the model, keeping the Texas-shaped body (see Figure 44a). Furthermore, we add Gaussian random noise of 5% of the standard deviation to the seismic and gravimetric data.

To study the effect of random noise, we show in figure 45c the velocity gradients for clean and noisy data, as well as the difference. Qualitatively, there seems to be no difference between both gradients, however, Figure 45c shows that there are discrepancies of 0.638% percent of the maximum amplitude of the free noise gradient. For example, at a single iteration of step line search instead of adding 100.00 km/s of velocity FWI will add 100.64 km/s in the worst-case scenario. Although we do not show the gradients and contaminated traces, we did the same analysis with more noisy data: Noise of 20% produces differences of 3.0% percent of the maximum amplitude of the free noise gradient, value still acceptable despite the fact of being very noisy data. With respect to the form of Figure 45c.

Starting from the velocity model of Figure 44b we run the iterative scheme for 300 iterations obtaining the velocity model of Figure 46a. Compared with the true velocity model, the shape of the high-velocity body is recovered as well as some structural information on the layers: the faults are detected properly with their respective velocities values. In the true model (Figure 44a) we put five objectives at certain depths with an appreciable velocity contrast with respect to the layers where they are located, these objectives simulate geological traps. After 300 iterations of our cooperative scheme, the medium depth objectives are recovered. The deepest objective is not getting recovered since it is located too far from the seismic survey and too close to a strong velocity-density contrast at the bottom of the model. At early iterations, the gravimetric inversion had been solved, but we keep performing it sequentially to help FWI with the regularization embedded in the gravimetric inversion.

The seismic traces of Figure 47 show an excellent match in phase between the observed noisy data and the final synthetic data, starting from the initial synthetic trace (Red line). With respect to the amplitude, some events exhibit good recovery, the events at the end of the seismogram lack of recovery because correspond to deepest regions of the model. With respect to the misfit evolution (Figure 48), both seismic and gravimetric inversion converges harmonically and well behaved, with the slight difference that the gravimetric inversion converges faster. This because the mass distribution due to the starting model (Figure 44b) amenable for gravimetric inversion.

4.5 SEG SEAM phase I model

The SEG SEAM phase I model is a complex earth model designed for sub-salt imaging in tertiary basins, typical for the Gulf of Mexico. It was modified from the original by shortening the velocity range to 1500-3500 *m/s* and by adding a water layer on the top

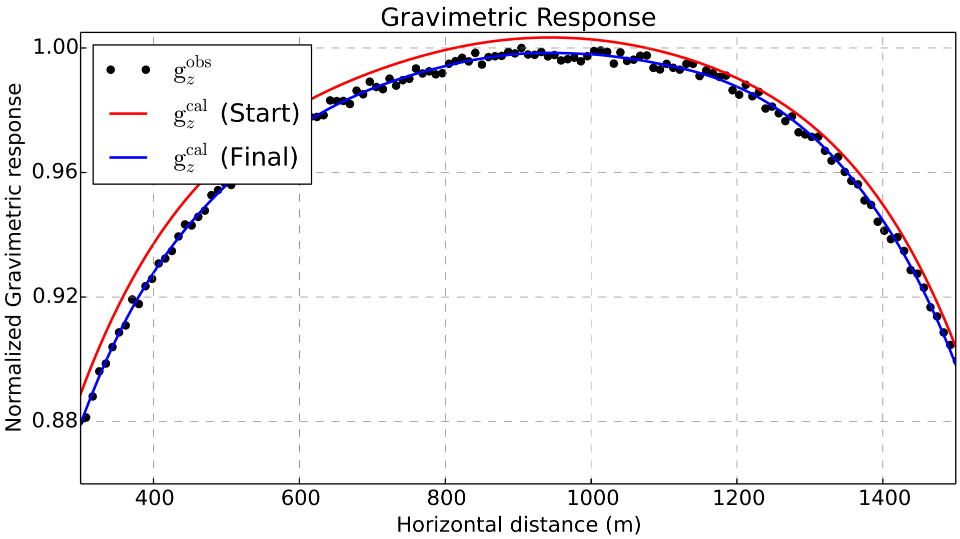
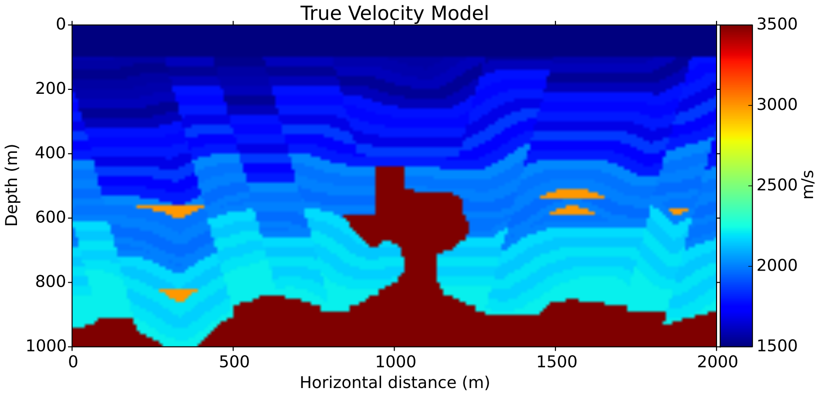
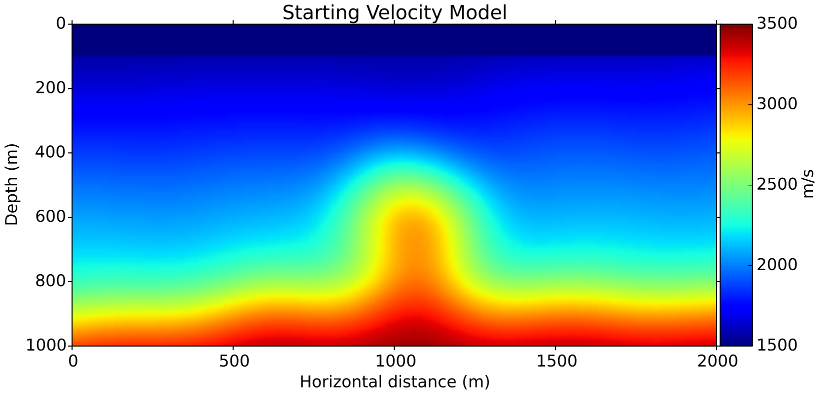


Figure 43. Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 1000 iterations of our scheme.

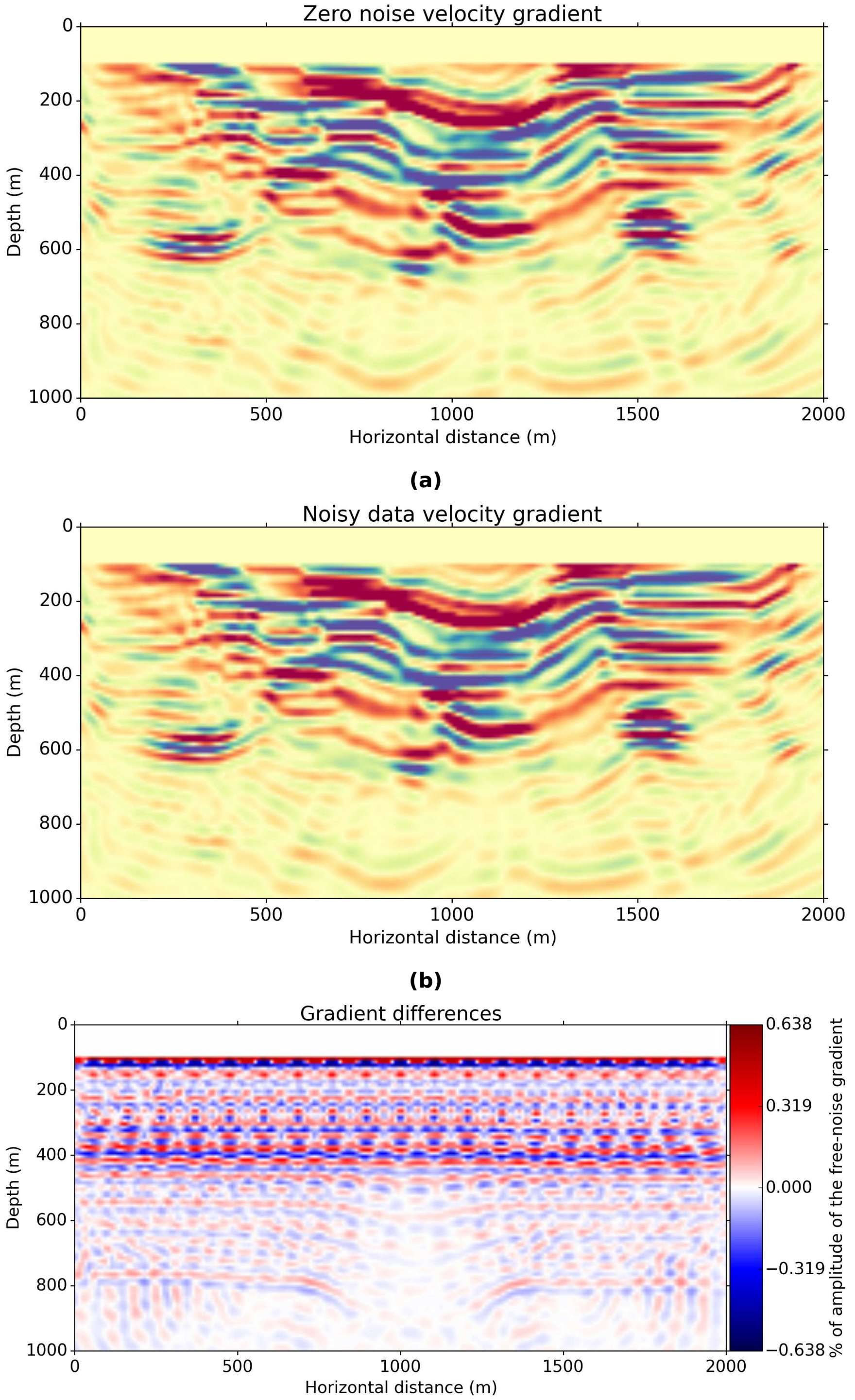


(a)



(b)

Figure 44. Texas-shaped true velocity model (a) and its smoothing set as a starting model (b) for an example with noise.

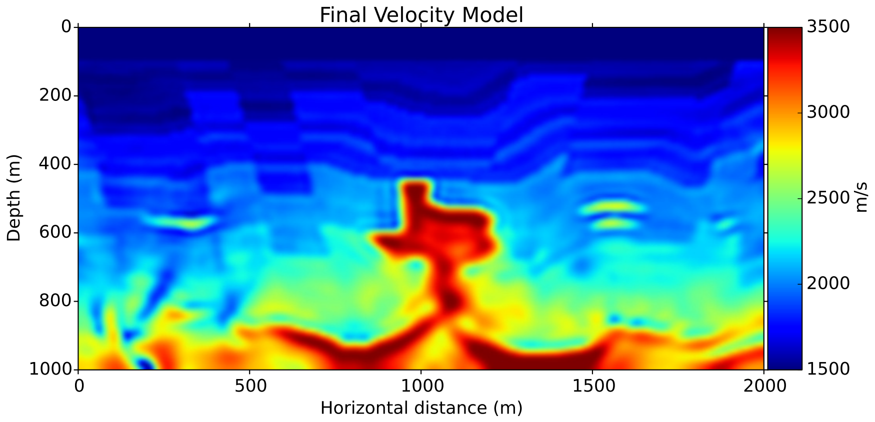


(a)

(b)

(c)

Figure 45. Velocity gradients using noise-free data (a) and noisy data (b) and the difference in percentage. Both gradients are muted at the water layer.



(a)

Figure 46. Texas-shaped model results using conventional FWI with noise in the data.

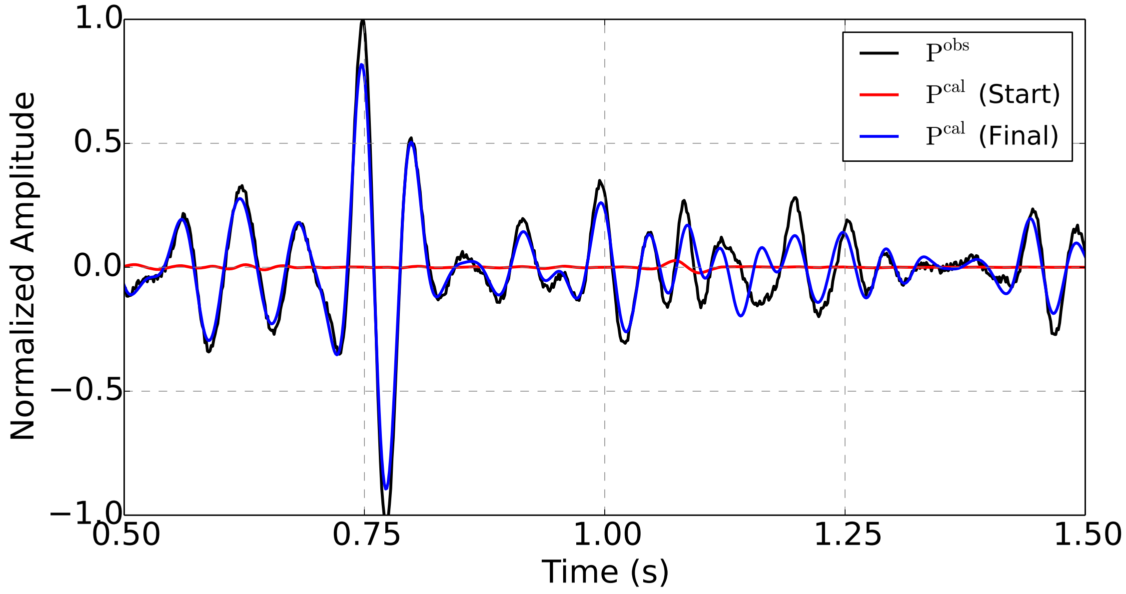


Figure 47. Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 300 *m* due to a source applied at 200 *m* on the surface.

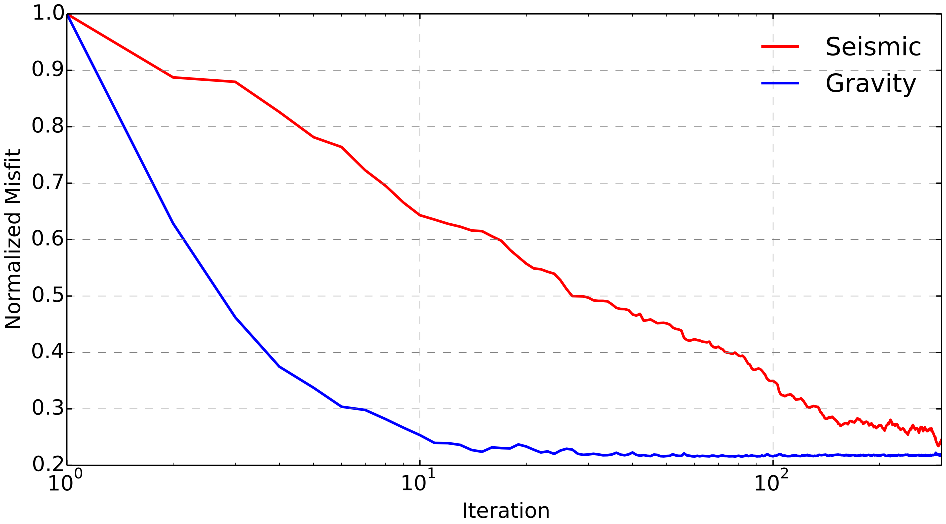
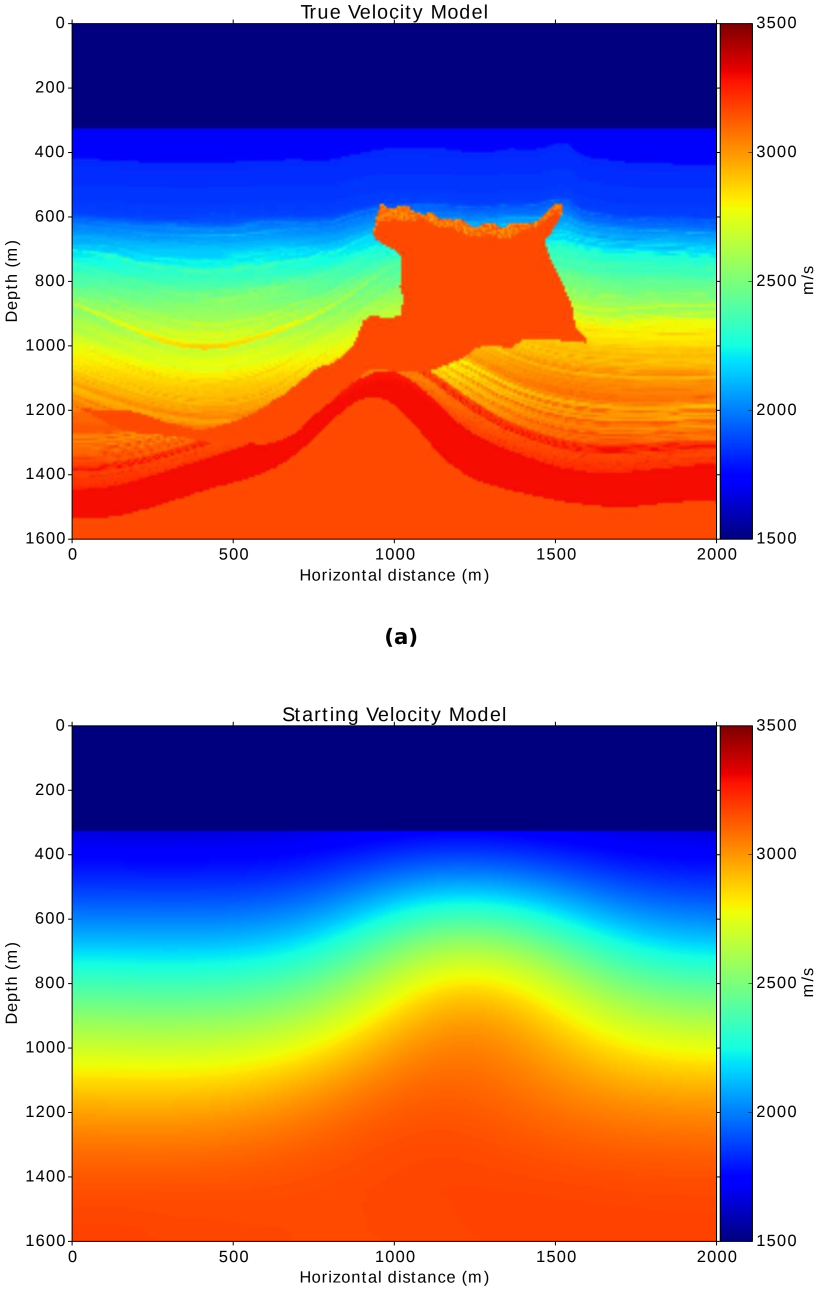


Figure 48. Normalized misfit for seismic (red) and gravity (blue) for the Texas-shaped model.

as shown in Figure 49a. Since this model have the same velocity range that previous ones, the parameters used model (Sources, receivers, inversion constants) will be also used. In this case, we increase the depth to 1600 m. We perform 300 iterations of the cooperative inversion scheme using the starting model of Figure 49b.

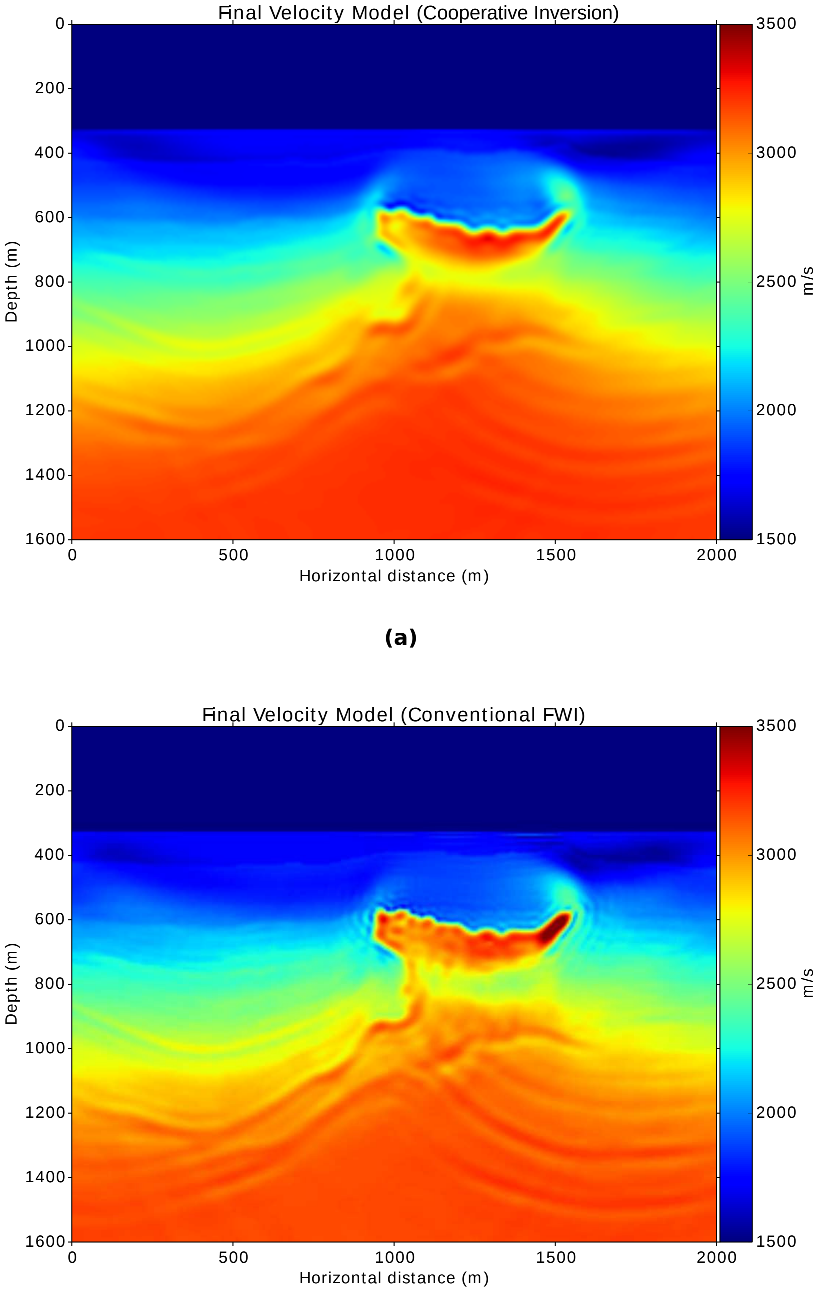
Similar to the previous results, pure FWI focuses more on the layered part of the model as shown in Figure 50b; meanwhile, cooperative inversion (Figure 50a) leads to more mass addition on the inner part and does not produce spurious effects around the high-velocity body. The misfit (Figure 51) shows a good and relatively fast convergence at early iterations, however, after the 100th iteration, the misfit starts showing a slow convergence. The final velocity model recovers the stratigraphic part and the shape and top of the body. It is clear that the slow convergence at the last iterations is due to the moderate addition of velocity in the inner part of the body.



(a)

(b)

Figure 49. Modified SEG SEAM phase I velocity model and its smoothing set as a starting model (b).



(a)

(b)

Figure 50. Modified SEG SEAM phase I velocity model results after 300 iterations for typical FWI (a) and Join Inversion (b).

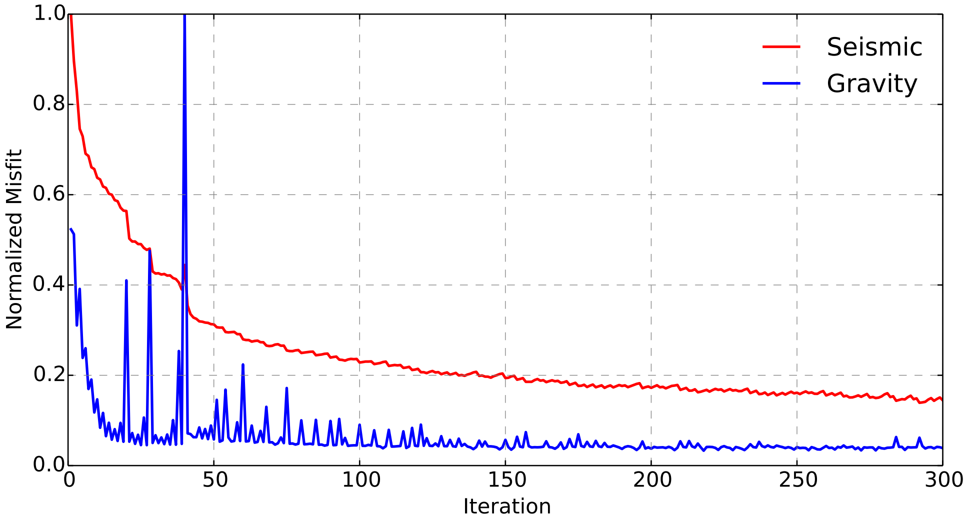


Figure 51. Normalized misfit for seismic (red) and gravity (blue) for the SEG model.

4.6 CICESE-shape model with an over-trust fault

The following results correspond to a synthetic model based on the geological environment of a thrust fault. This type of fault occurs when the lower layers of the crust are folded and pushed upwards above younger layers, creating a reverse fault that has a dip less than 45 degrees with respect to the horizontal (Serra, 1977). If the dip is approximately 15 degrees, the fault is called an over-thrust fault. The true velocity model created and based on the previous description is shown in figure 52a. The domain covers a horizontal distance of 2000 meters and a depth of 1000 meters. On the shallower layers, the velocity is low (1500-2000 m/s) and the stratigraphic resembles the logo of CICESE (pelican) on the left corner, that's why this model is named CICESE-shaped model.

On medium depth layers on the right side of the model, five faults are placed with a dip of 60 degrees. On the deepest layers with higher velocity (around 3500 m/s), the trusting occurs with dipping around 15 degrees. As previously explained, the water later is considered to be known. It is important to mention that this synthetic model is created only with purposes of testing the inversion algorithm on geological environments similar to thrust faults, neglecting its diagenesis and interpretation.

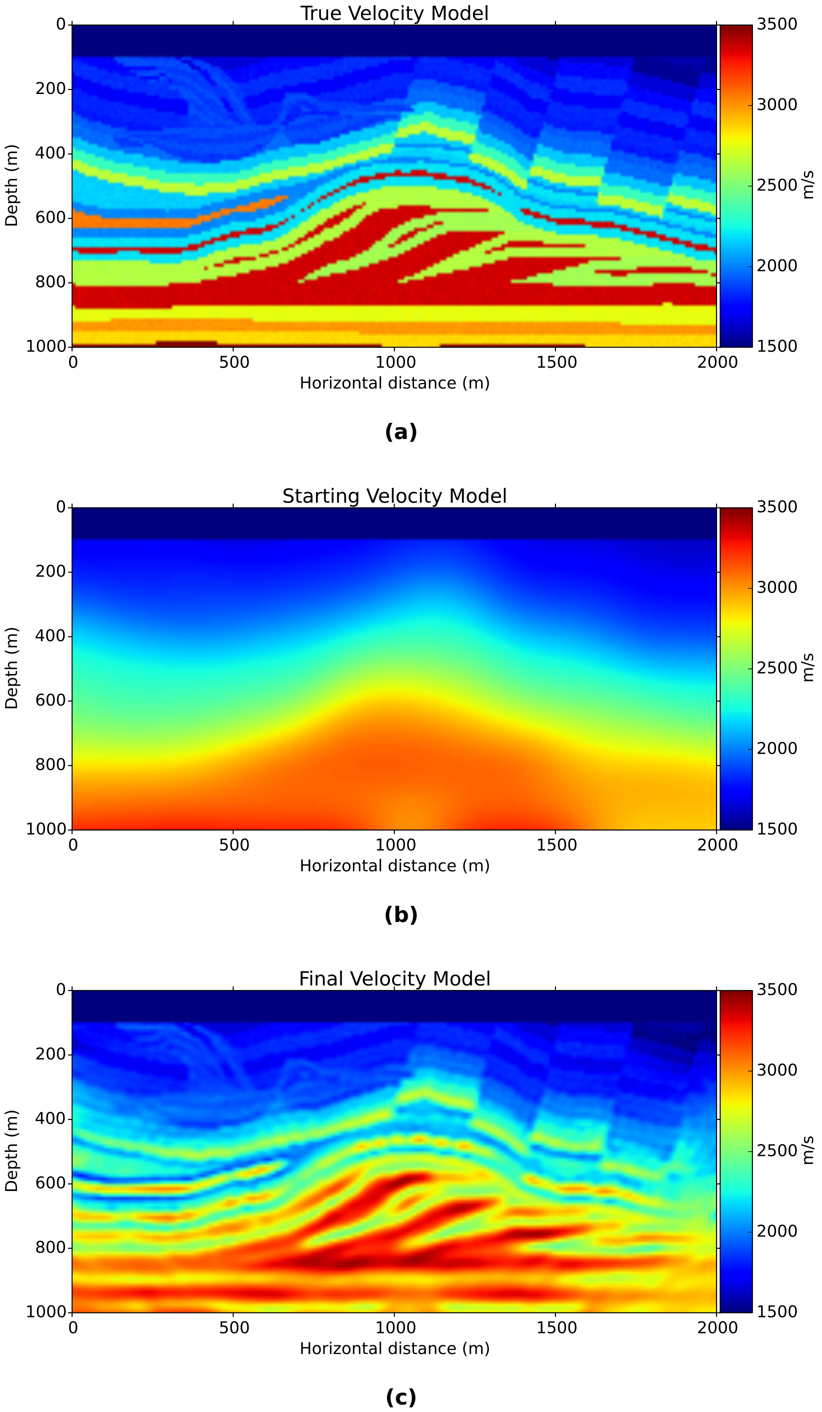


Figure 52. CICESE-shape true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after a joint inversion progress

The starting model is a Gaussian smoothing of the true velocity model which results in a loss of stratigraphic and fault information (Figure 52b). Since the domain and velocity range is the same as the Texas-model results, the same parameter is used in this experiment (sources, receivers, gravimetric stations, regularization, recording time, data noisy). Again, the density model is not shown since it is obtained applying the Gardner $v_p - \rho$ petrophysical relationship to the models of Figure 52.

The final velocity model after 85 iterations of cooperative inversion is shown in Figure 52c. As explained before, the true model has mainly 3 different regions of interest to compare with the final result. Firstly, let us compare the upper-left side of the model where the logo of CICESE is presented in the low-velocity layers. The final velocity model recovers perfectly the shape of the pelican, the position and the values of its layers. This means that the joint inversion process can detect and discern shallower structures with low contrast. Beneath this region, some horizontal layers of higher velocity (around 2500-3000 m/s) are also recovered. Secondly, consider the upper-right part where some faults are present. The five 60 degrees faults are recovered in the same position and the same dipping. With respect to the velocity values, the result is acceptable, presenting some difficulties on the farthest right side because of the dipping towards the right: this part does not have a reflection of the seismic wave onto the surface where the seismograms are located. Also, the layers (without faulting) beneath the 60 degrees faults are recovered. With respect to the third region, the bottom part of the mode where the trusting is presented, the three layers that thrust at the middle of the model are recovered in the same way as the small layers and element around the trusting. It is expected that the velocity values are not recovered since the deepest is the one that provides less information at the surface, however, the stratigraphic footprint is accurate. In summary, this kind of environment (thrust fault, 60 degrees faults, shallower structures) can be studied using this methodology.

Concerning the behavior of the cooperative inversion process, the cost of the objective function is presented in Figure 53. It shows that the convergence is fast and well-behaved for seismic and gravity data. For the gravimetric inversion, the inversion is immediately solved but it is continuously solved to help the seismic inversion even though the gravity data has been fitted. With respect to the seismic fit it seems that this result can be improved using more iterations

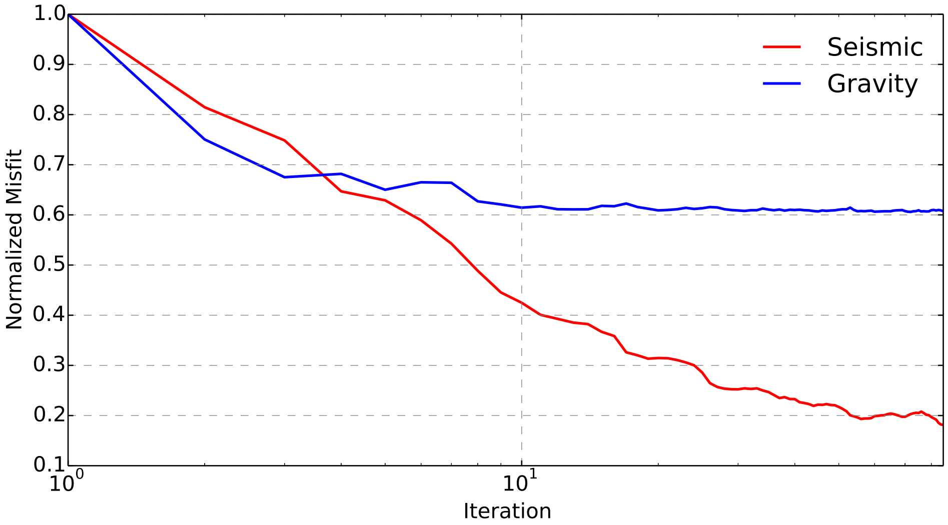


Figure 53. Normalized misfit for seismic (red) and gravity (blue) for the CICESE-shaped model.

With respect to the data fit, let's compare a single receiver (250 meters on the surface) for the real, initial and final seismic trace as shown in Figure 54. The fit in the seismogram is achieved in phase and partially recovered in amplitude, this means that the layer position was recovered successfully and the velocity total value is recovered, even when Gaussian noise has been added to all traces. It is important to recall that the process seeks to adjust all traces at the time (200 receivers by 100 sources). Respect to the gravimetric fit, the gravity data at the 200 stations are shown in Figure 55.

For this example, since most of the density is concentrated at the middle region of the starting model like the true model, the observed and the starting synthetic gravity data match, that is why the gravimetric inversion is achieved faster than the seismic inversion.

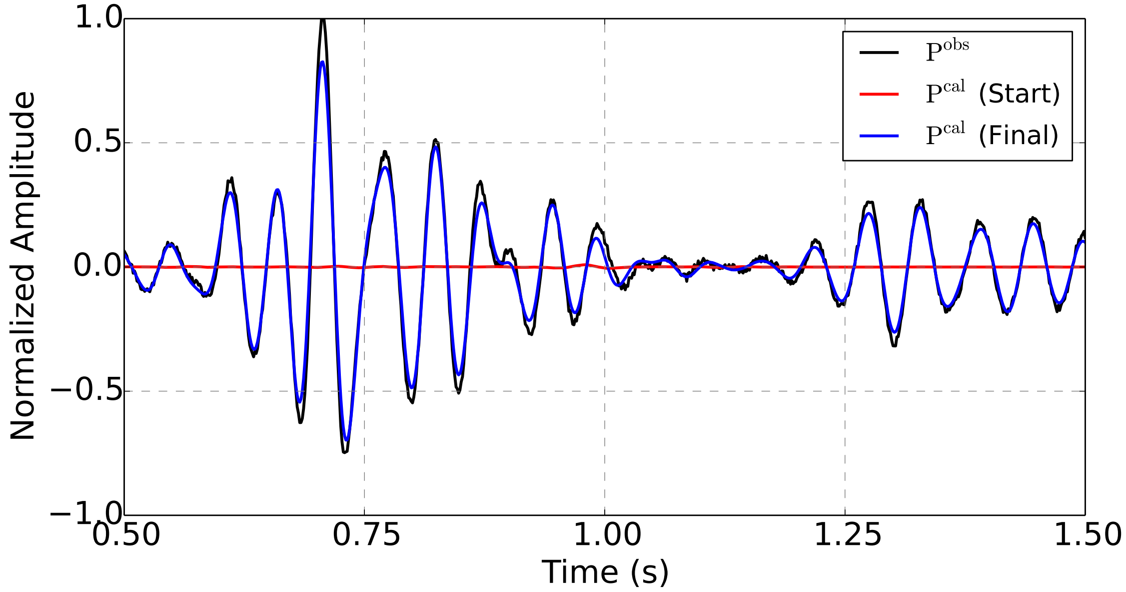


Figure 54. Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 250 *m* due to a source applied at 200 *m* on the surface for CICESE-shape model.

4.7 Baja-shape model with domino faulting

To conclude the section of results using 2D cooperative inversion with gravimetric and acoustic data, a final synthetic model is presented. The main intention of this model is to recreate the geological environment of some tilted block faults to increase possible scenarios where the algorithm can be exploited. This type of fault occurs when the lower crust flows into the upper crust, producing fractures and creating detachment faults. On a local scale, these normal faults appear like tilted blocks on a bookshelf. The true velocity model is shown in figure 56a where the domino faults can be appreciated. The domain covers a horizontal distance of 2000 meters and a depth of 1000 meters. On the shallower layers, the velocity is low (1500-2000 m/s) and the domino faults have a dip of 60 degrees. This model resembles the Marmousi model as shown in the previous chapter, with the difference that the discordance presented has the shape of the peninsula of Baja California at both sides at the bottom of the domain, hence the name of Baja-shape model. As explained before, locally the faults look like tilted dominoes, where each block has lower velocities at the top and 3 layers of higher velocities (>3000 m/s) around medium velocities (2000-3000 m/s). The bottom part is not faulted.

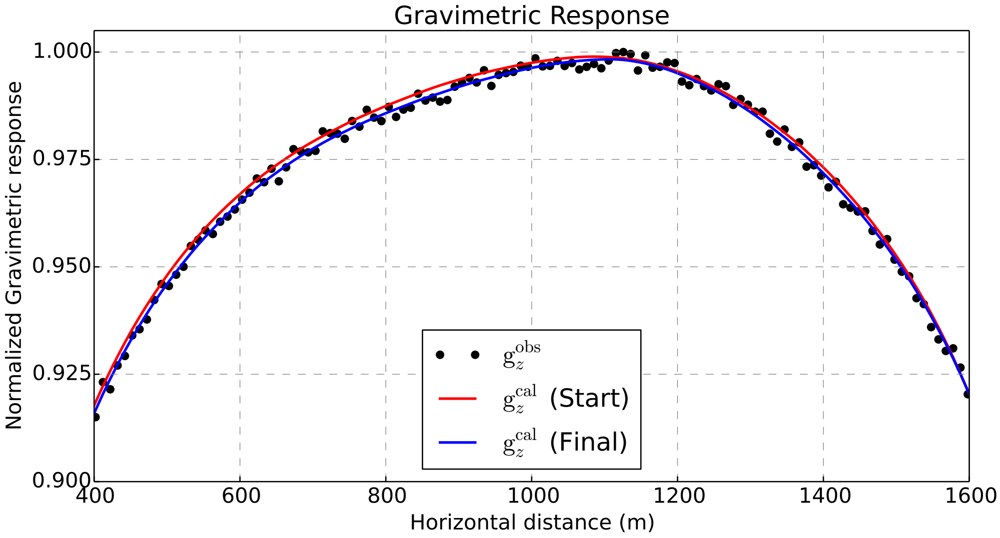
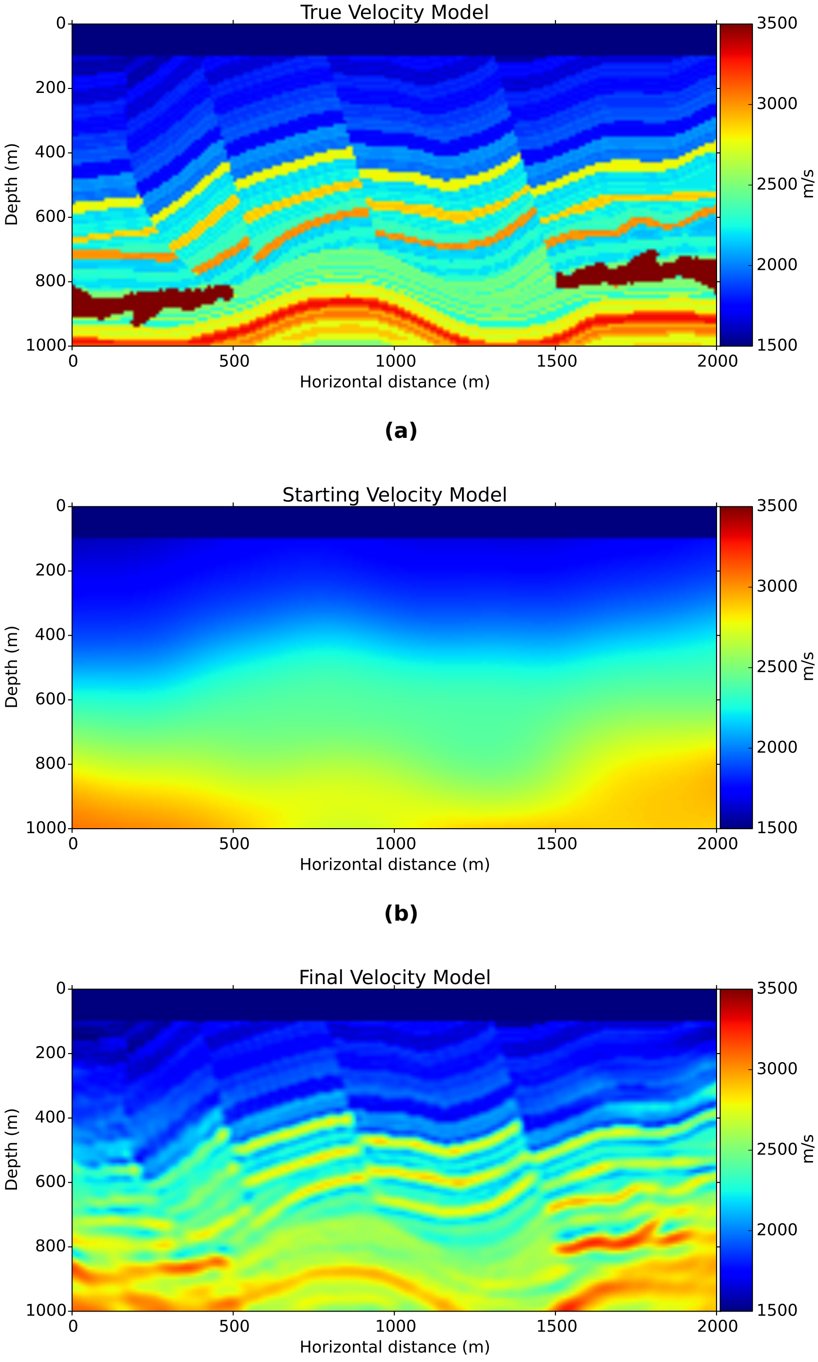


Figure 55. Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 85 iterations of cooperative scheme for the CICESE-shape model.

The starting model is a Gaussian smoothing of the true velocity model, where the faults and block positions are not presented (see Figure 52b). Since the domain and velocity range is the same as previous results, the same parameters are used in this experiment (sources, receivers, gravimetric stations, regularization, recording time, moisey). The density model is not shown.

The final velocity model after 85 iterations of cooperative inversion is shown in Figure 56c. Differently to previous results, this model has not large bodies of high velocity-density, all layers have approximatively the same thickness, that is the reason why the result presents a good recovery in the layers positions, the tilted block faults and the discordance events (Baja shape), however, lacks in the total velocity recovery, for example, the 3 tilted layers with velocities of 2000-3000 m/s do not recover the real value.

For the evolution of the algorithm, the cost of the objective function is presented in Figure 57, where the convergence is fast and well-behaved for seismic and gravity data. The gravimetric inversion has faster convergence than the seismic inversion, meaning that the gravimetric problem has been solved. The seismic convergence is well-behaved but needs more iterations of pure FWI without gravimetric.



(a)

(b)

(c)

Figure 56. Baja-shape true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after cooperative inversion.

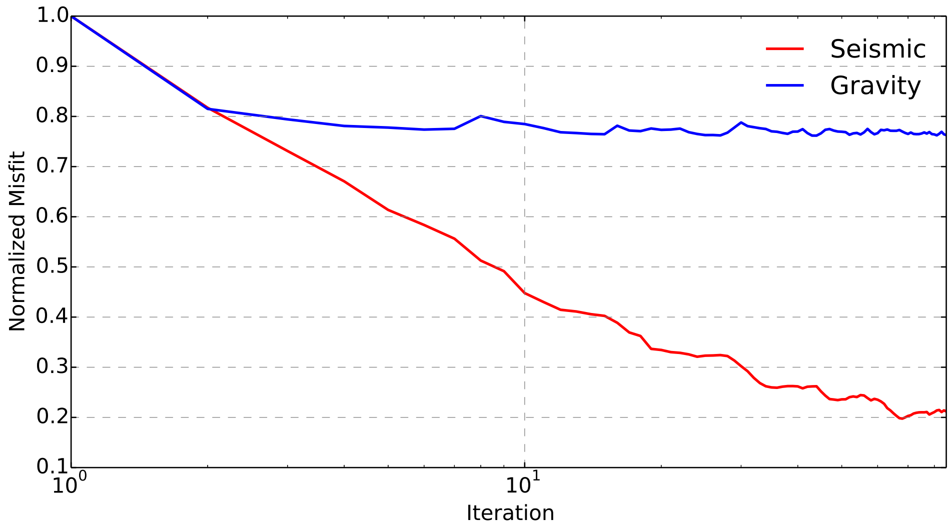


Figure 57. Normalized misfit for seismic (red) and gravity (blue) for the Baja-shaped model.

Respect to the data fit, a single receiver located at 250 meters on the surface is picked and presented in Figure 58 for the real, initial and final seismogram. The fit in the seismogram is mostly matched in phase and, as expected, the amplitude is still partially missing since the velocity values have not been totally recovered as shown in Figure 56c. It is also expected that the gravity data will no provide enough information given the geological environment of this model, as shown in Figure 59 where the observed, the initial and starting data are similar.

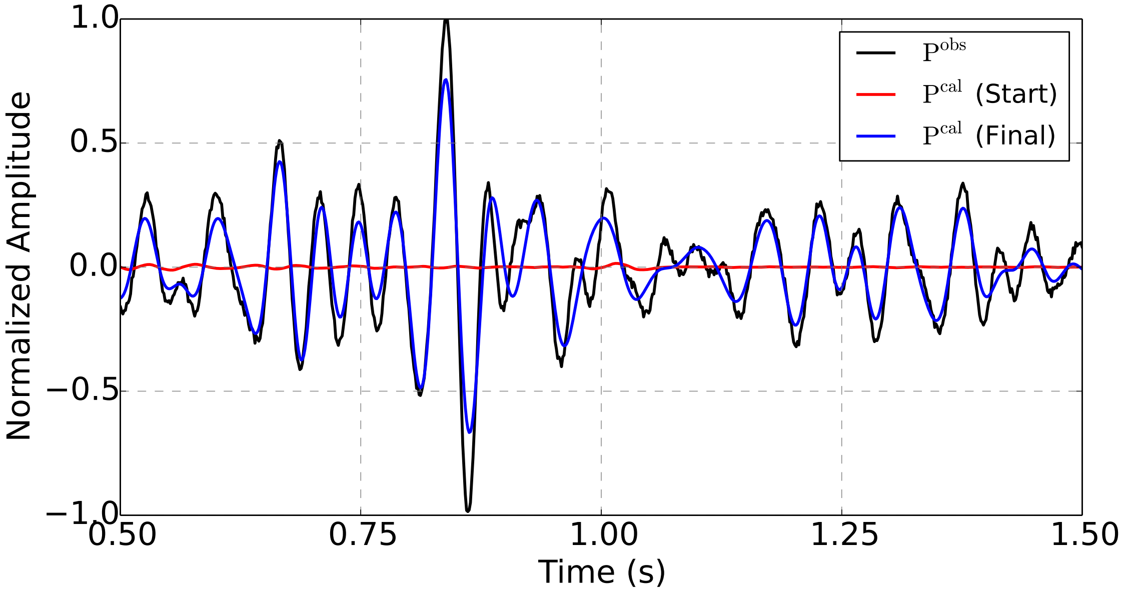


Figure 58. Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 250 *m* due to a source applied at 200 *m* on the surface for Baja-shape model.

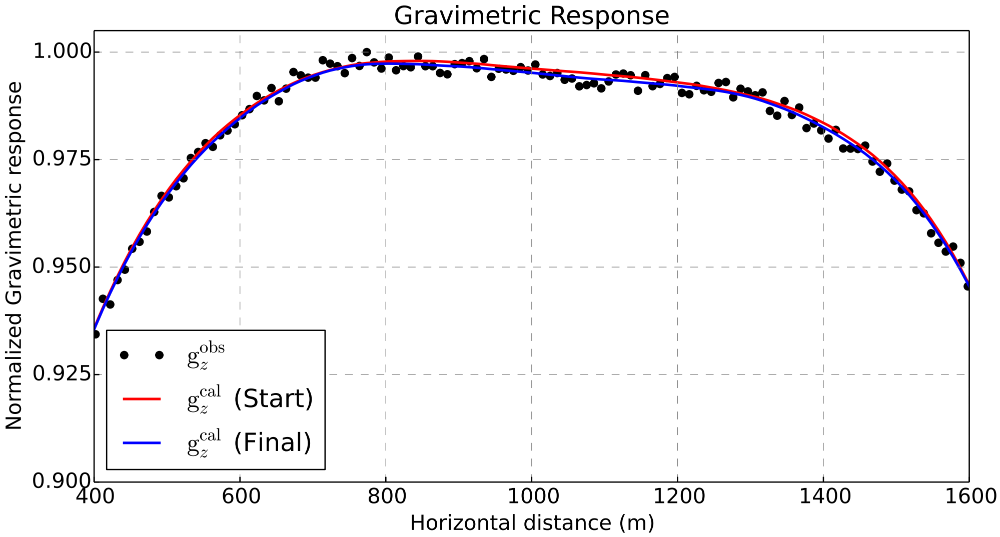


Figure 59. Normalized true (black dots), starting (red line) and final (blue line) gravimetric responses result of 85 iterations of cooperative scheme for the Baja-shape model.

This result is important in order to test the usefulness of the cooperative scheme proposed in this work when different geological scenarios are employed.

4.8 3D Acoustic FWI: 3D CICESE-BAJA model

Although the previous result solves the inversion problem proposed for bi-dimensional media the immediate question comes to the possibility to perform this methodology for three-dimensional models. First, consider only 3D FWI in acoustic media. The synthetic 3D model proposed to solve combines the environments used before: a discordance event with the high velocity with the shape of the peninsula of Baja California and a high-velocity body with the shape of the logo of CICESE beneath lower velocity layers, without any kind of faulting as shown in Figures 60a, 61a and 62a. The domain covered and other parameters are consistent with the previous results for simplicity, with the important detail that the grid points and source-receiver geometry were shortened to reduce computational costs. The following table summarizes the parameter used in this 3D example.

Table 8. Parameters used for 3D CICESE-BAJA acoustic FWI.

Parameter	Value	Unit	Description
x_{MIN}	0	m	Starting x-horizontal distance
x_{MAX}	2000	m	Final x-horizontal distance
y_{MIN}	0	m	Starting y-horizontal distance
y_{MAX}	2000	m	Final y-horizontal distance
z_{MIN}	0	m	Starting depth
z_{MAX}	1000	m	Final depth
n_x	40		Grid nodes for x
n_y	40		Grid nodes for y
n_z	20		Grid nodes for z
T	2	s	Recording time
n_t	400		Grid nodes for t
v_{MAX}	3500	m/s	Maximum velocity
f	5	Hz	Ricket wavelet's peak frequency
n_r	40×40		Number of receivers
n_s	10×10		Number of sources

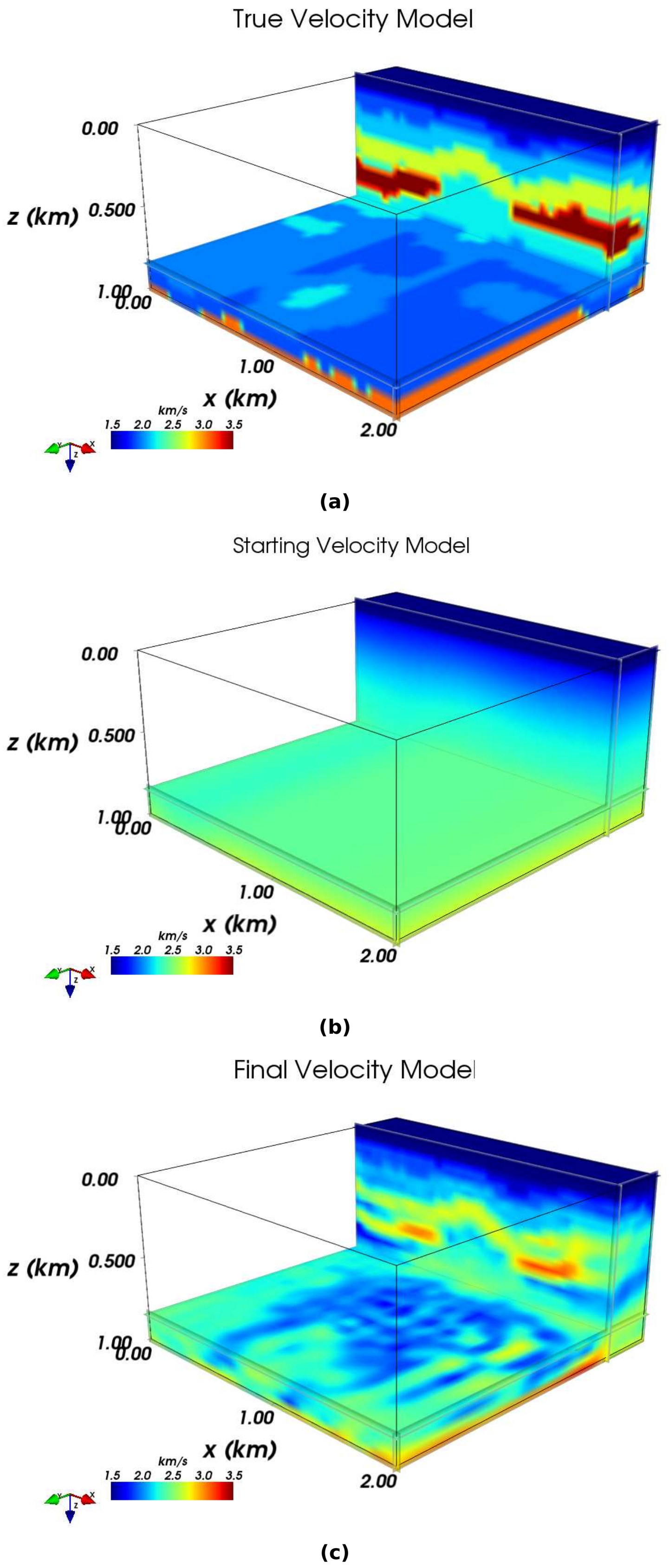
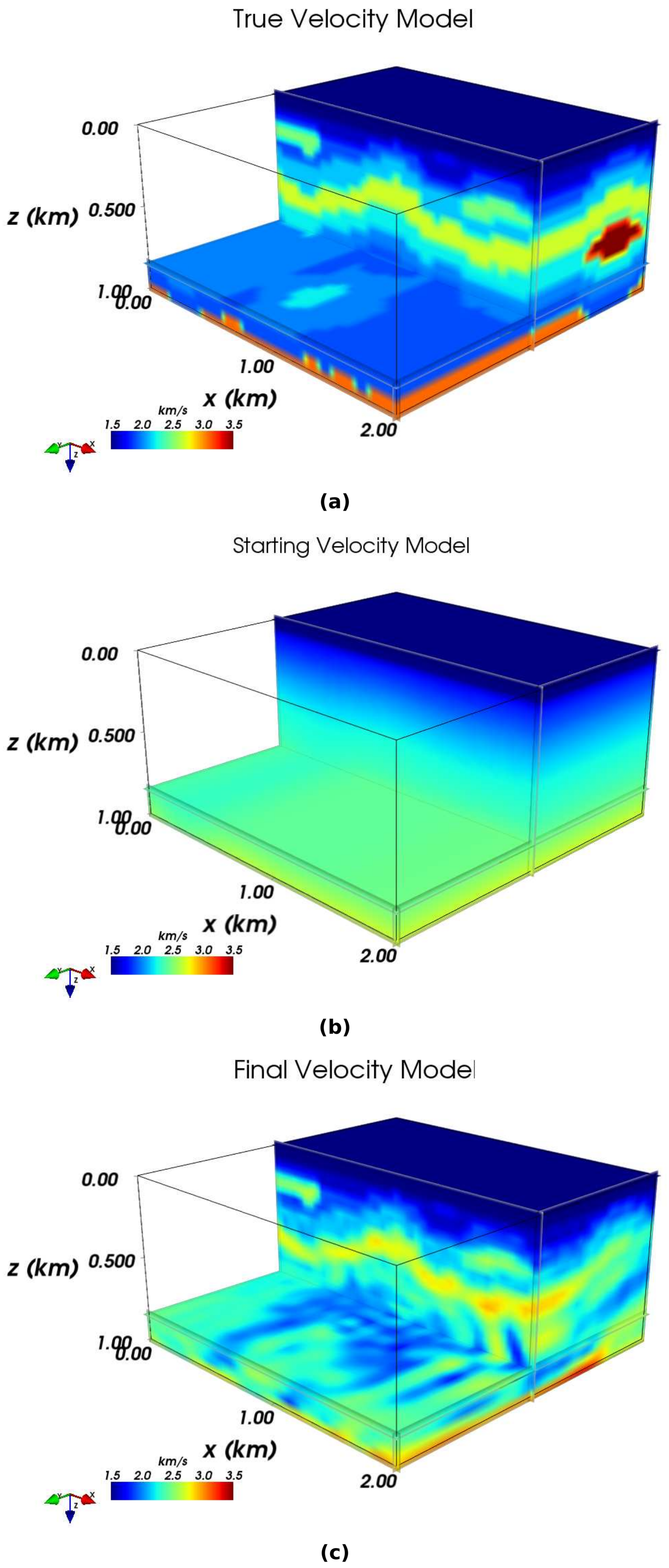


Figure 60. 3D CICESE-BAJA true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after FWI, corresponding to a constant y-distance of 350 meters.


True Velocity Model

(a)

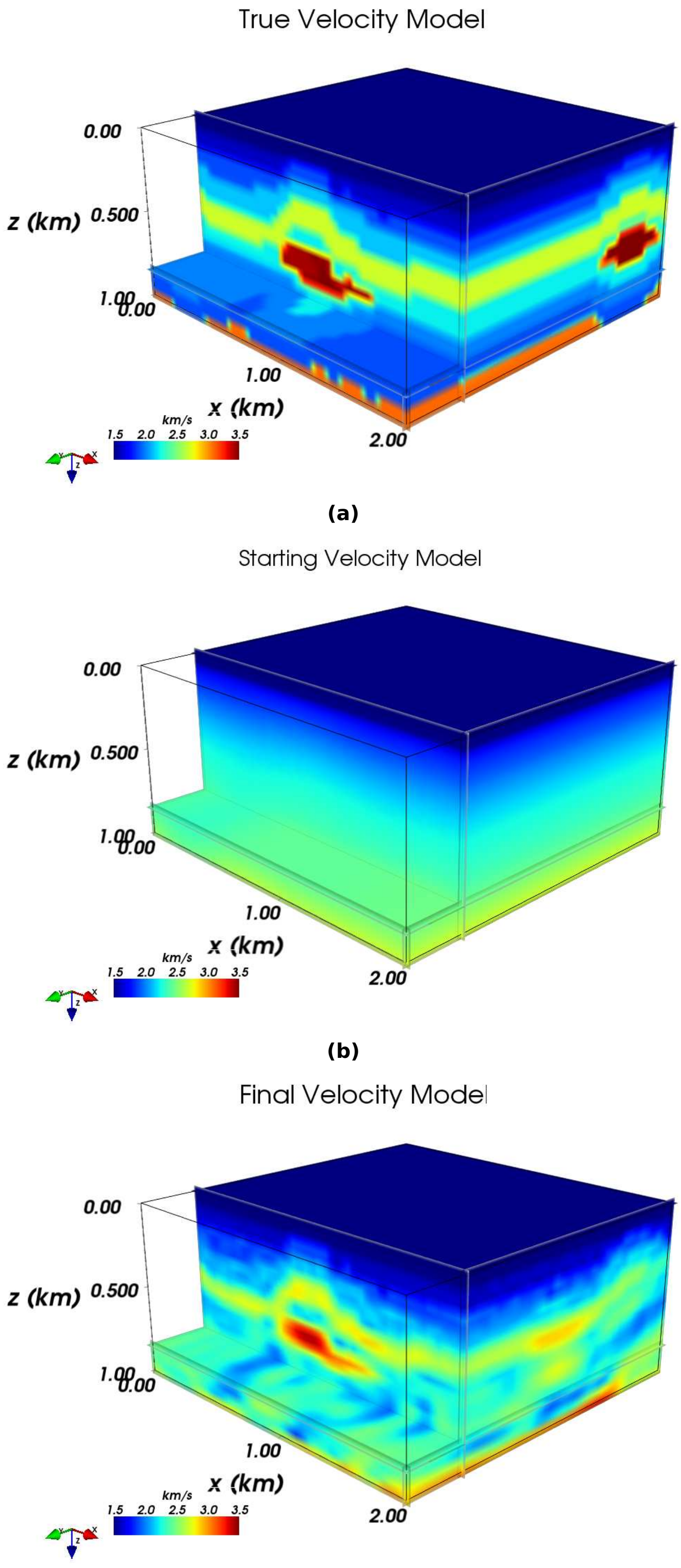
Starting Velocity Model

(b)

Final Velocity Model

(c)

Figure 61. 3D CICESE-BAJA true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after FWI, corresponding to a constant of y-distance 1000 meters.



(a)

(b)

(c)

Figure 62. 3D CICESE-BAJA true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after FWI, corresponding to a constant y-distance of 1600 meters.

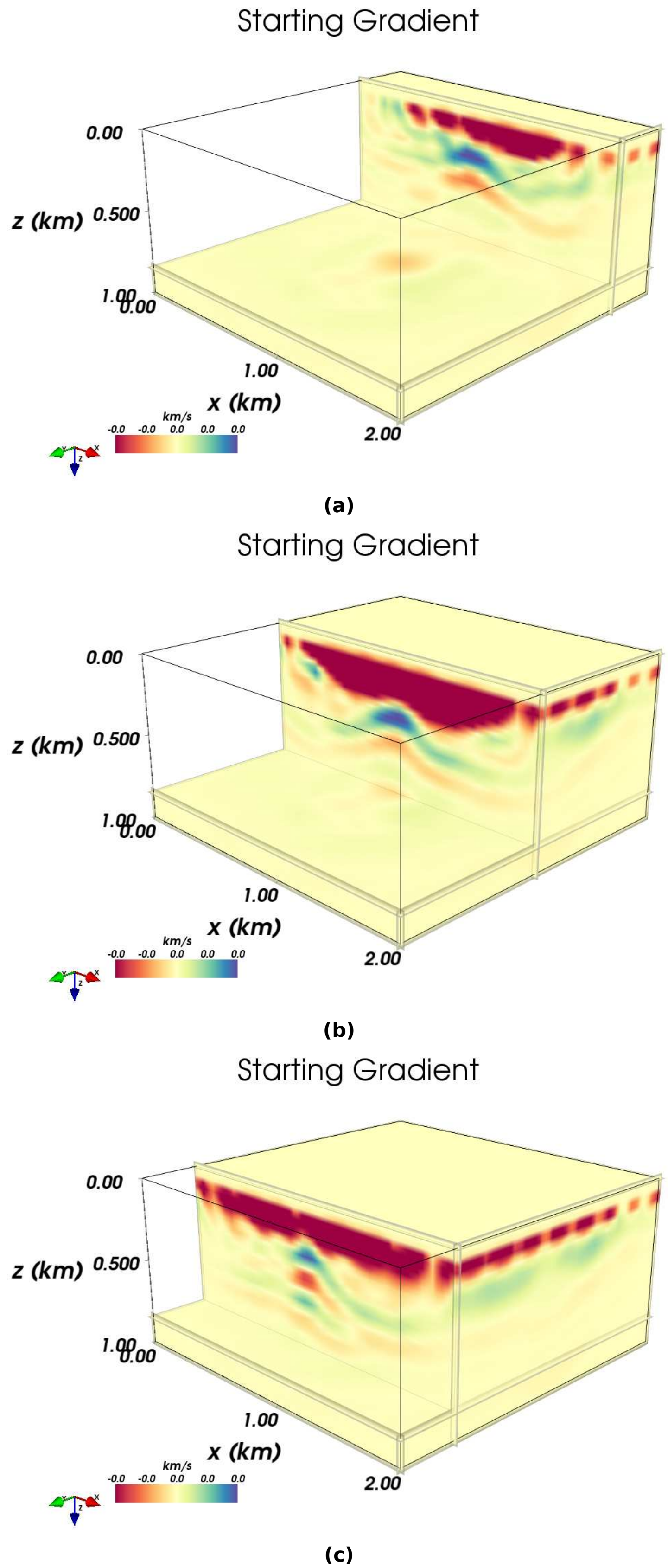


Figure 63. 3D CICESE-BAJA starting velocity gradient for distances (a) y=350, (b) y=100, and (c) y=1600 meters.

Given that it is hard to visualize a 3D models on paper, 3 constant $y-$distances were selected along the domain: $y = 350$ meters, $y = 1000$ meters and $y = 1600$ as shown in Figure 60, 61 and 62 since the most important heterogeneities are presented in these $x-z$ planes. With respect to the values of the Table 8 notice that the frequency has been lowered to 5 Hz. This is mainly because more grid points cannot be added because of the limited computational resources, therefore the frequency needs to be lowered to avoid grid dispersion. Recalling the exercises of Chapter 3, $400 \times 400 \times 200$ spatial grid point by 1000 time grid points were used for the acoustic banana-doughnut kernel (Figure 11) corresponding to 38 hours of computation, for a single source and a single receiver.

The starting model consists of horizontal velocity layers increasing smoothly in depth. i.e., this time a Gaussian smoothing have not been applied as shown in Figures 60b, 61b and 62b. In this case, 40 receivers in the x-distance by 40 receivers in the y-distance were used on the surface, given a total of 1600 receivers for each shot. The source geometry consists of 10 sources in the x-distance by 10 sources in the y-distance, given a total of 100 sources distributed at 100 cores of MPI process. Thus $40 \times 40 \times 10 \times 10 \times 400 = 64$ millions of data points to be adjusted, considering a low resolution modeel, notice the significant increment in cost for 3D problems.

The FWI process performed 300 iterations with a computational cost of 5 hours per iterations (per gradient computation), with a total of 62.5 days for this *simple* example. The starting gradient with this starting model corresponds to Figure 63 where immediately the sensibility due to the layers is presented. The final result of inversion is shown on the slices previously selected corresponding to Figures 60c, 61c and 62c respectively.

As seen in Figure 60c, the low-velocity layers and the medium velocity layer (the yellow one of approximately 2800 m/s) are correctly located for the inversion process. With respect to the event that resembles a discordance (Baja California shape), it is partially recovered presenting the higher velocity on this slice. For the y-distance of 1000 meters, in the true model (Figure 61a) were placed 2 bodies of velocities of 2500 m/s (color green) at the upper right and left of the 3D model. The final inversion recovers these bodies in shape, position and velocity value (Figure 61c) and also recovers the horizontal layer in the middle of the model (color yellow). Finally, it is shown in Figure 62c for y-distance of 1600 meters that FWI recovers a layer of 2800 m/s (color yellow) which appears in all slices. In this case, a body of high velocity was placed in the middle of the true velocity model which is also located on position and shape, especially on the top.

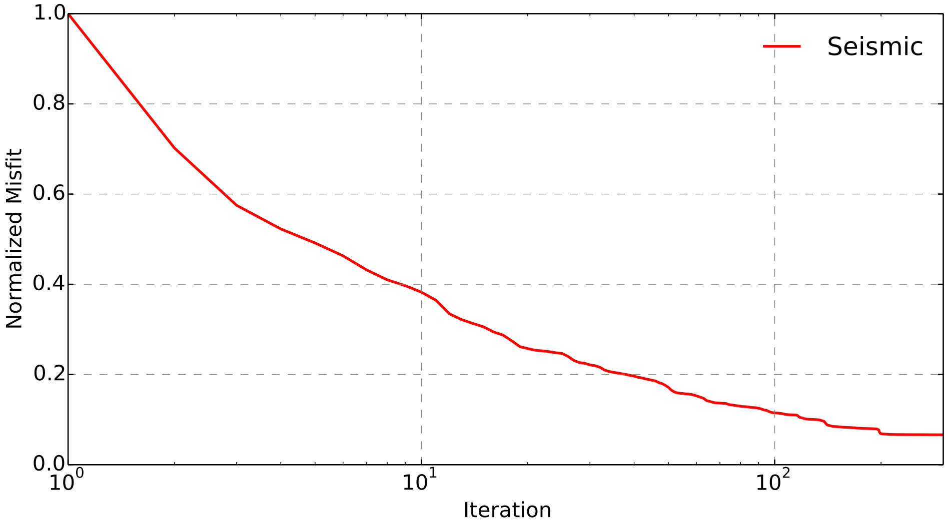


Figure 64. Normalized misfit for seismic (red) and gravity (blue) for the Baja-shaped model.

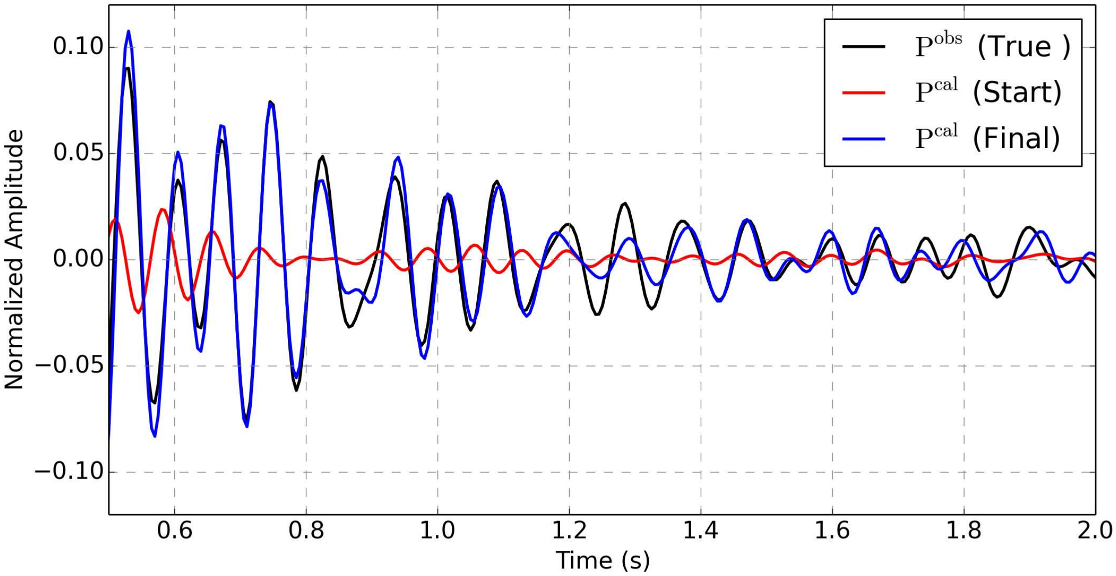


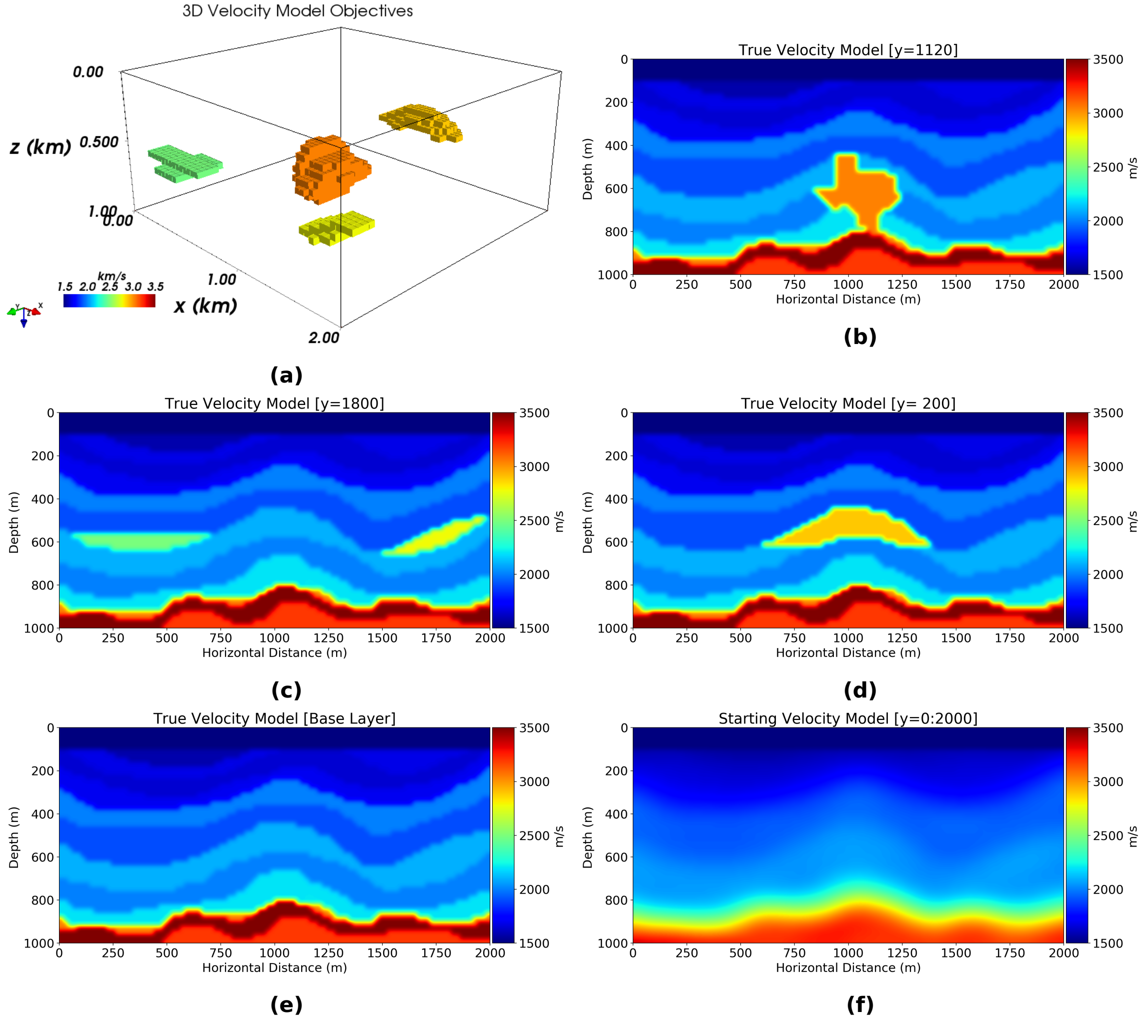
Figure 65. Comparison of observed (black), initial synthetic (red) and final synthetic traces measured at 250 *m* due to a source applied at 200 *m* on the surface for Baja-shape model.

With respect to the evolution of the FWI code developed, the objective function over 300 iterations is shown in Figure 64. The cost reduction is achieved and it is well behaved and slower for later iterations. With respect for the data fit, a single receiver located at (250,250) meters on the surface are selected and presented in Figure 65 for the real, initial and final seismogram. All events match in phase due to the stratigraphic recovering. 3D FWI is easy to implement and solve, however, the computational cost grows exponentially if high-resolution models are intended to solve, or in this case, several tests executions need to be performed to calibrate a cooperative inversion scheme.

4.9 Texas-shaped 3D Model. Cooperative inversion

For this example, we created a 3D model based on the structure of the Texas-shape model I and discretize it with a mesh of size: $n_x = 50, n_y = 50$ and $n_z = 25$. The model consists on 4 primary structures of high velocity (Figure 66a) surrounded by low-velocity layers, constant along the $y-$direction (Figure 66e). The central structure is a 3D Texas-shaped body (Figure 66b), primarily concentrated at $y = 1120$ m with a velocity of 3050 m/s. There are two other structures located approximately in the $x-z$ plane for $y = 1800$ m with velocities of 2500 and 2800 m/s respectively, as shown in figure 66c. Finally, a fourth object is centered at $y = 200$ m with a velocity of 2900 m/s with the shape of a banana. The density model for this example (not shown) is obtained from Gardner's equation. The inversion and modeling parameters for this model are similar to the previous examples, however, we have increased the number of sources to $n_s = n_{s_x} \times n_{s_y} = 10 \times 10 = 100$ and the number of receivers to $n_r = n_{r_x} \times n_{r_y} = 50 \times 50 = 2500$ equally spaced along the surface and we use a peak frequency of 8 Hz given the grid used. We perform 100 iterations of the cooperative inversion scheme using the 3D starting model shown in Figure 66f.

Similar to the previous 2D results, the method successfully recovers the main features of the model (Figure 67). The shape and position of the main object are recovered, nevertheless, the total mass is not fully recovered (compare Figures 67c and 67d). The other 3 bodies and the layered structure are also recovered, as shown in



(a) (b) (c) (d) (e) (f)

Figure 66. 3D velocity model for cooperative inversion, where (a) shows the four objectives and (b), (c) and (d) 2D views at constant y. (e) The layered model in the $x-z$ plane (it is constant in the y direction). (f) The starting model is a smooth version of the layered model.

Figures 67f and 67b.

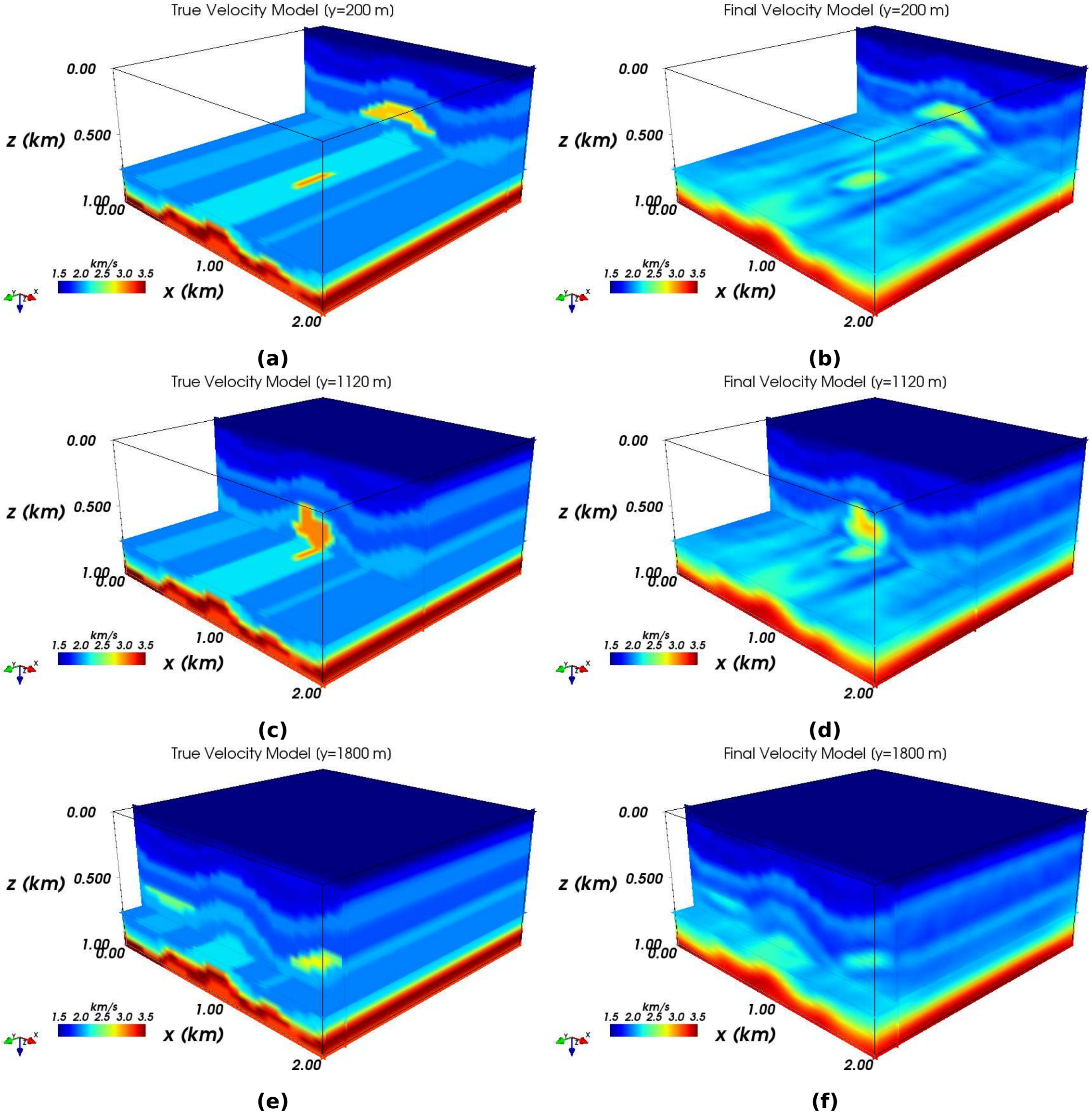


Figure 67. 3D velocity models for different $y-$ views for the true model (left) and the cooperative inversion results (right).

The data misfits (Figure 68) show a relatively fast convergence for the seismic data. The gravimetric misfit shows an increase in the first part because the method gives more weight to FWI, however, it exhibits good convergence after the 6th iteration. In the later iterations, the seismic misfit starts showing a slow convergence whereas the gravimetric misfit shows further improvement. The slow convergence of FWI is due to the moderate accretion of velocity and density in the Texas-shaped body and the other structures.

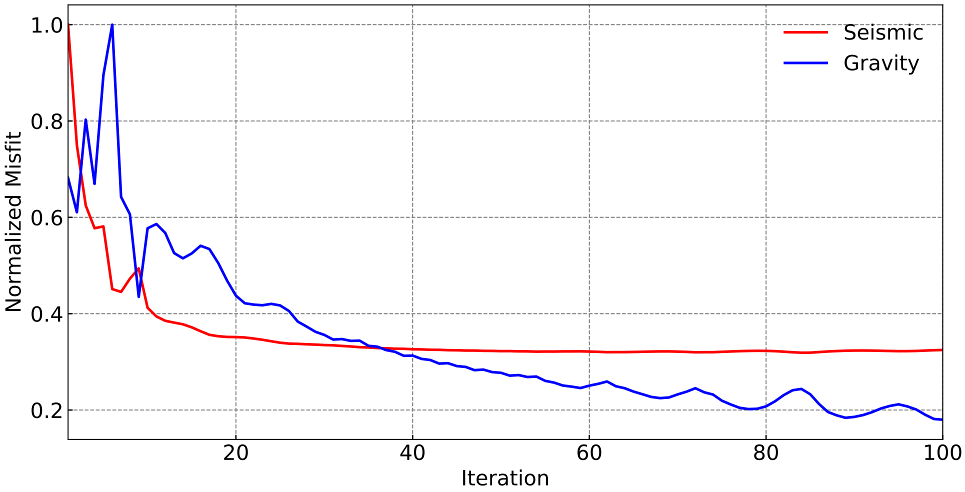
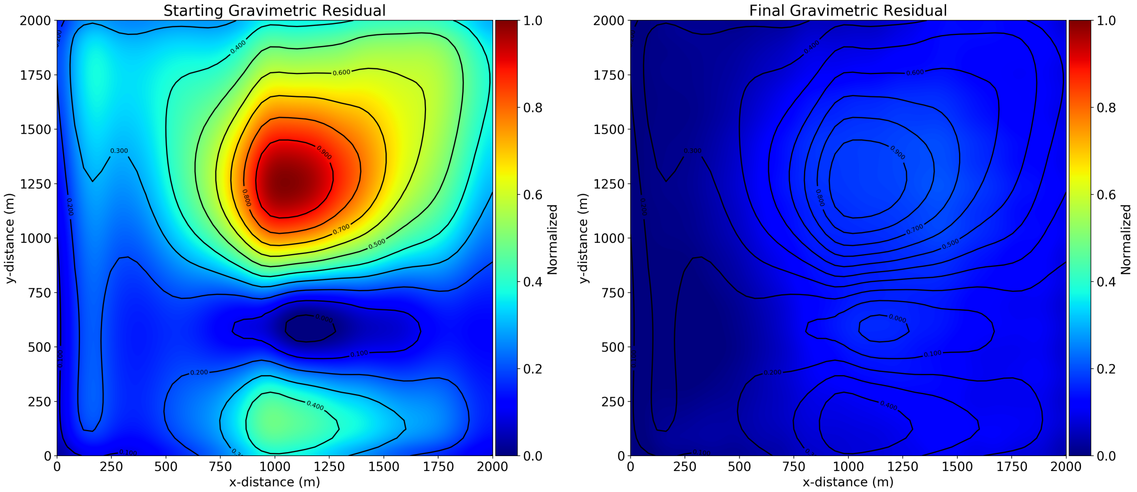


Figure 68. Normalized misfit for seismic (red) and gravity (blue) for the 3D example.

The starting and final gravity residual are shown in Figures 69a and 69b respectively. The starting residual is mainly due to the Texas-shaped body. Besides that, the other objectives also contribute to the residual in their respective regions. The final gravimetric residual has been normalized with respect to the starting residual, the contour of the starting residual is displayed in Figure 69b to facilitate the comparison. Notice that the amplitude of the residual has been reduced one order of magnitude.



(a) **(b)**

Figure 69. Starting gravimetric residual (a) and final gravimetric residual (b) after cooperative inversion. Both residuals are normalized with respect to the starting residual and the starting contours are displayed in the final to illustrate the regions where the residual is reduced.

To compare the fit of the seismic data, we select three receivers with different source positions and show the observed (black dashed), starting (red) and final (blue)

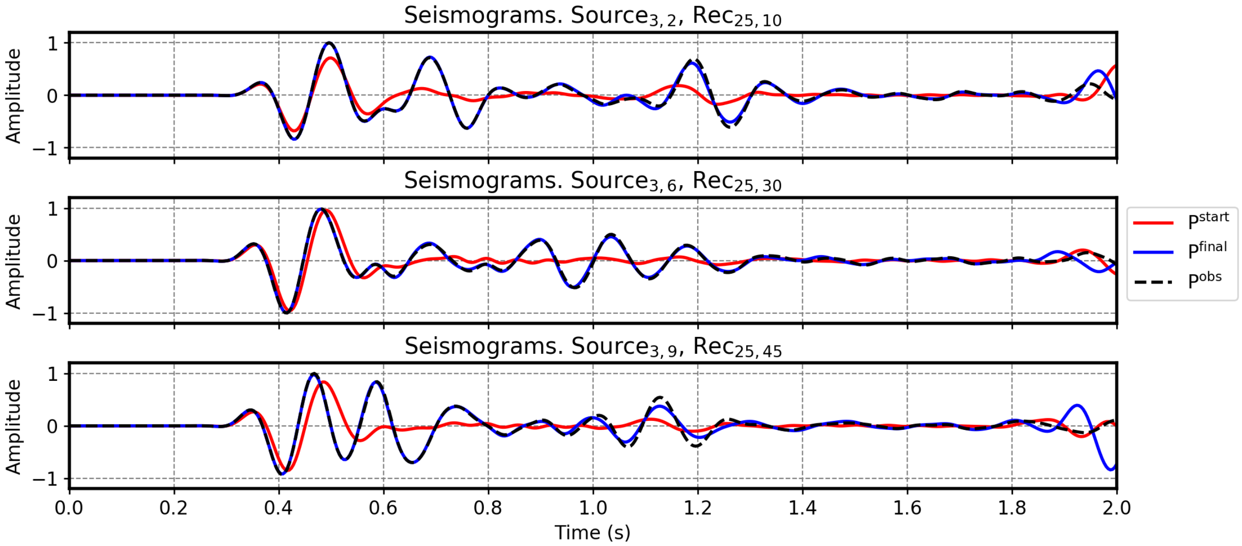


Figure 70. Seismic data fit for different sources and receivers for observed pressure (black dashed), starting synthetic pressure (red) and final synthetic pressure (blue) after 100 iterations of cooperative inversion.

Table 9. Source and receiver position for the three seismograms selected to compare the data fit.

Source Position (x, y) [m]	Receiver Position (x, y) [m]	Nearest Velocity Target
(444.44, 222.22)	(1020.41, 408.16)	Figure 66d
(444.44, 1111.11)	(1020.41, 1224.49)	Figure 66b
(444.44, 1777.78)	(1020.41, 1836.73)	Figure 66c

pressures seismograms in Figure 70. The three seismograms are strategically located near the different velocity objectives (Figure 66a) as shown in table 9. Consistently with the convergence of the seismic misfit (Figure 68), the seismograms from the final iteration match the observed pressure for all receivers as shown in figure 70. All the events are matched with only minor discrepancies. Therefore, we conclude that both geophysical data sets are fitted with proper 3D velocity and density models.

Chapter 5. Results: Elastic media

In the previous chapters, different geological environments using full waveform inversion and 2D cooperative gravimetric with full waveform inversion in acoustic media were presented with successful results, explaining the advantages in each case. The results were compared with individual inversion, showing the benefits of the cooperative inversion scheme proposed in this work.

Although the Earth obeys Hooke's theory of elasticity, the acoustic approximation is an excellent approximation to solve seismic problems in the processing of traces, RTM and FWI. In recent years, many researchers have implemented this methodology in such media in the same successful way as this work, neglecting effects such as elasticity, anisotropy, attenuation or poroelasticity. Then it is imperative to solve this problem on a more realistic medium to close the gap between theory and real seismic problems.

In this chapter, the inversion problem will be applied to the elastic media. Therefore the Elastodinamic wave equation will be modeled and inverted using the formulation described in Chapter 2. Firstly, Elastic Full Waveform Inversion (EFWI) will be solved for 2 synthetic models previously solved in AFWI: the Marmousi model and the Texas Shaped Model. Finally, the same problems will be solved using cooperative inversion, exposing the advantages and disadvantages of the methodology proposed.

Given that in the previous chapter the separated gravimetric inversion was already performed and discussed, those results will not be presented to avoid redundancies. Also, taking into consideration the discussions about the computational cost of 3D inversion for seismic problems, application to 3D elastic media will not be discussed because the computational cost due to the vectorial nature (3 component data) to this problem is prohibitive.

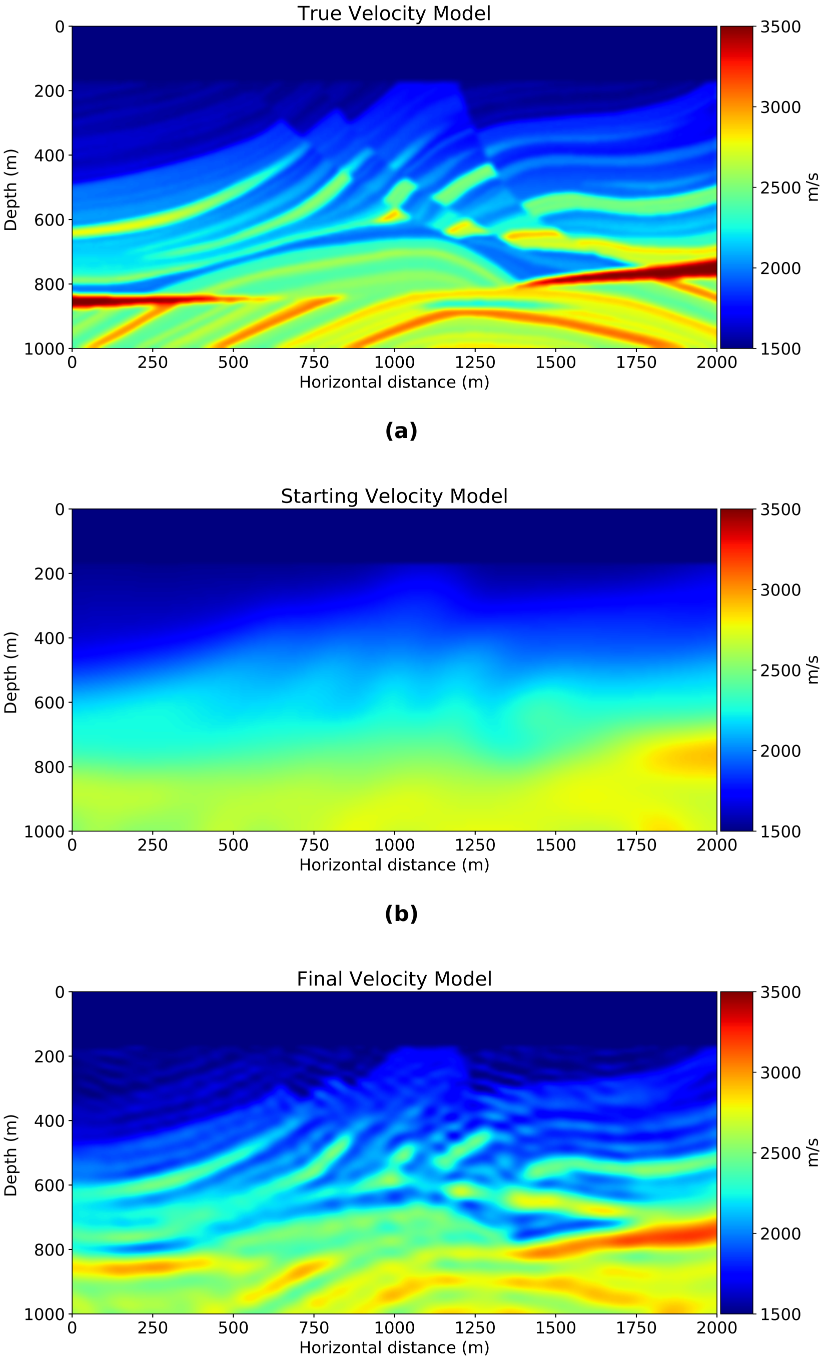
5.1 EFWI: Marmousi model

First, let us consider the previously explored Marmousi model. The geometry and parameters used are the same as in the example of AFWI, as shown in Table 10. The *S*-wave velocity is computed using the ratio $v_p/v_s = \sqrt{3}$ and the density is scaled using Gardner petrophysical relationship $\rho = \rho_0 V_P^{k_0}$, therefore the models for v_s, ρ are not presented for this example. For the step line search, 10 test points are used and only the vertical component of the displacement is considered. The true and starting velocity models are the same as the previous result, as seen in Figure 71a and 71b respectively.

Table 10. Parameters used for elastic FWI for the Marmousi model.

Parameter	Value	Unit	Description
x_{MIN}	0	*m*	Starting x-horizontal distance
x_{MAX}	2000	*m*	Final x-horizontal distance
z_{MIN}	0	*m*	Starting depth
z_{MAX}	1000	*m*	Final depth
n_x	200		Grid nodes for *x*
n_z	100		Grid nodes for *z*
T	2	*s*	Recording time
n_t	1500		Grid nodes for *t*
v_{MAX}	3500	*m/s*	Maximum velocity
f	15	*Hz*	Ricket wavelet's peak frequency
n_r	200		Number of receivers
n_s	100		Number of sources

The velocity model obtained after 100 iterations of EFWI resemble the stratigraphic information of the Marmousi model. This particular result shows a spurious artifact in the final velocity model (Figure 71c). This problem is attributed to the match of multiple events on elastic seismograms, for example, in the inversion some *S*-wave events could be treated as P-wave events, therefore the appearance of the artifact describes this match with the presence of a small reflector on the top of the model. The iterative appearance of these artifacts contaminates on later iterations. Several authors mitigate these problems using filters on the gradient or using multiscale EFWI starting from lower frequencies.



(a)

(b)

(c)

Figure 71. Marmousi true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after elastic FWI (c).

A more accurate indicator for the quality of the EFWI iterative process is the analysis of the objective function for each iteration (Figure 72). The objective function for this example converges fast at earlies iterations and becomes slower for later iterations at the point to reach a stagnation, this is mainly due to the presence of such spurious artifacts commented earlier.

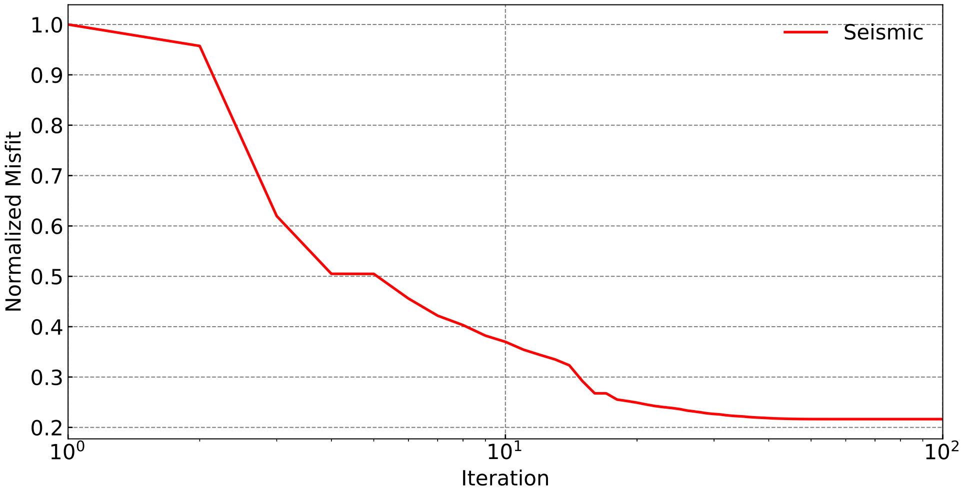


Figure 72. Normalized misfit for seismic for the Marmousi model result after EFWI.

The objective function seeks to minimize only the vertical component of the displacement u_z so let’s take a look at a single seismogram located at (1000,0) meters on the domain. As seen in Figure 73 the synthetic displacement (blue line) matches the observed displacement (black line).

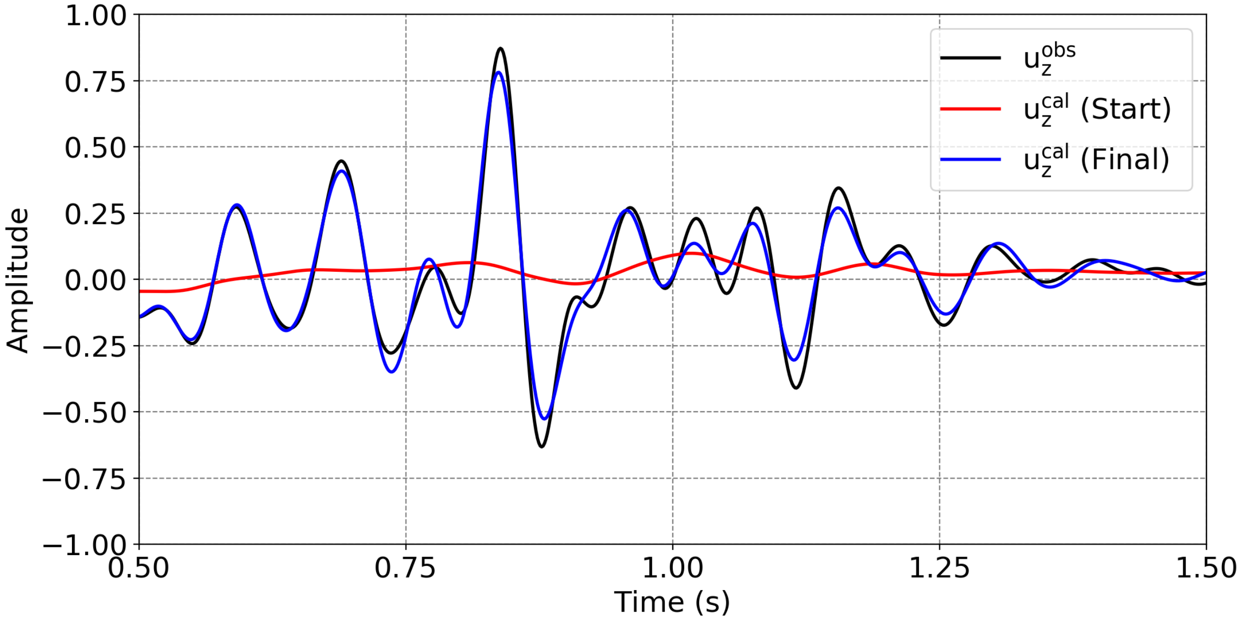


Figure 73. Seismogram comparisson for starting (red) and final (blue) synthetic data with the observed data corresponding to a single source and a single receiver for EFWI.

As seen in the seismogram fit, the match is quite acceptable for all the events. The phase is matched but the amplitude is still missing. Unfortunately, these results can not be improved using more iterations, therefore the implementation of additional techniques has to be performed. It is work mentioning that these results are not contaminated with Gaussian noisy, therefore this issue is to the cause of the spurious artifacts.

5.2 EFWI: Texas-shape model

For this example, let us consider the Texas-shaped model that was introduced in the previous chapter. The challenge in this model is trying to solve the presence of the spurious artifacts of the previous subsection, i.e, isolate the effect at the seismograms due to the existence of heterogeneities in the layers, in contrast with the Marmousi model. The true velocity model and the starting model are the same as the AFWI problem. The S-wave velocity model uses the ratio $v_p/v_s = \sqrt{3}$ and the density model is scaled using Gardner petrophysical relationship $\rho = \rho_0 V_P^{k_0}$ as shown in Figures 74a and 74b. The same forward model, inverse problem and geometric parameters are used as the ones of Table 10.

The result after 48 iterations is shown in Figure 74c. The spurious artifacts also appear on this model, hence the use of more heterogeneities on the layers is discarded as the cause of this issue. When more iterations are employed this problem gets worse, that's the reason why only 48 iterations were performed. The few liability aspects are the recovery of the stratigraphic information of the layers and the detection of the shape and position of the high-velocity body. In summary, this result is the same as acoustic FWI with the artifacts as the main disadvantage.

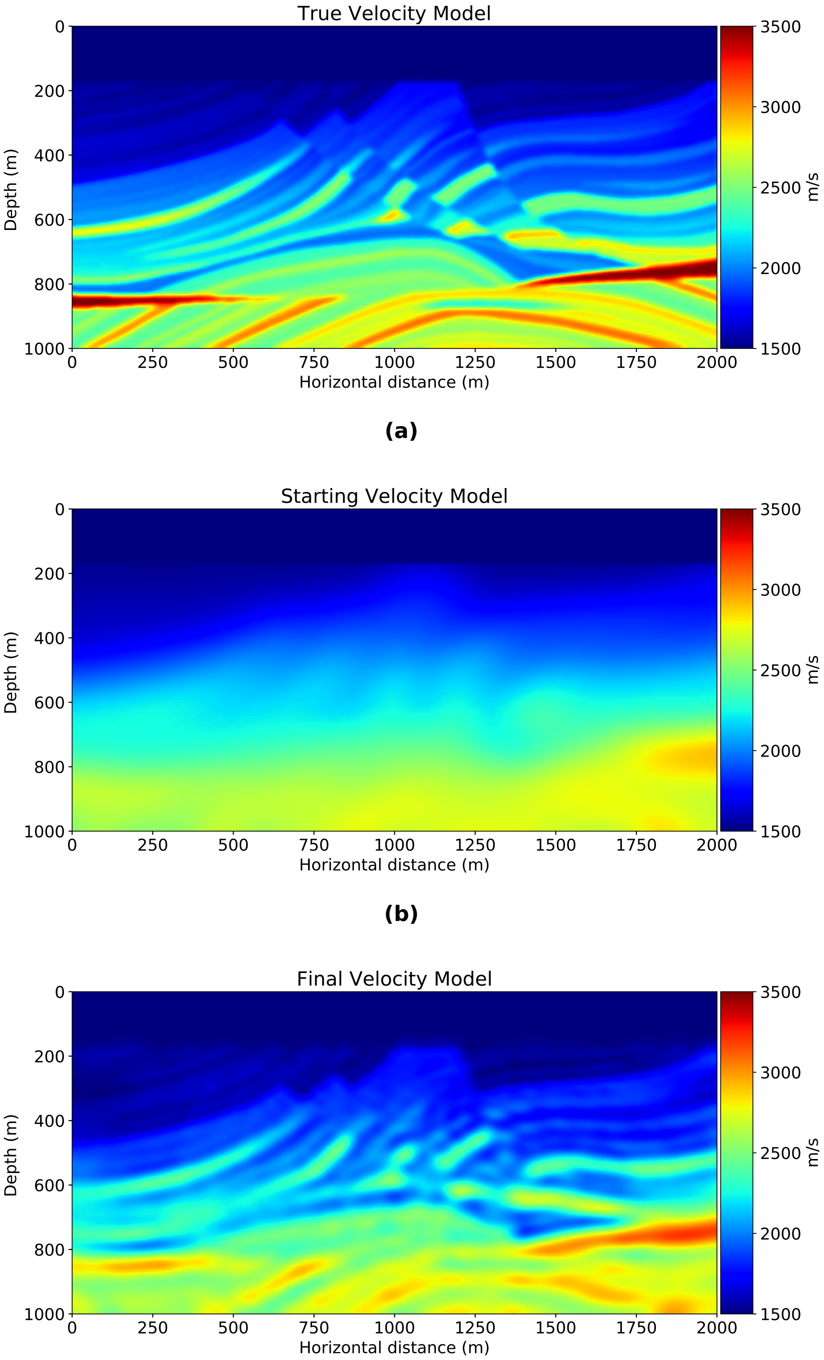
Concerning the data fit, for these 48 iterations the objective function is shown in Figure 75. Notice that the minimization of the objective function is achieved until the 40th iterations where stagnation is reached. As a quick experiment (not presented), elastic FWI setting the v_s-velocity equal to zero provides results without spurious artifacts.

This is expected since the problem is reduced to acoustic FWI. Therefore we conclude that such artifacts are mainly due to shear waves and converted shear waves.

5.3 Cooperative Inversion: Marmousi model in elastic media

Since the artifacts, which contaminate the inversion problem, do not depend on the structural environment of the velocity model and depends on a large degree to shear waves, we need additional process to mitigate them. Several authors employ noise reduction filters such as Gaussian filters, total variation filters and Lanczos filters, among others, to reduce these artifacts. Multiscale EFWI is also an alternative however, these artifacts can still appear for some frequencies. Recently, different norms in the objective function are being studied to solve this issue. Recalling the nature of the Tikhonov regularization (equal to Gaussian filtering) in this subsection the elimination of such effects is studied using the cooperative scheme proposed, where the filter is indirectly embedded in the gravimetric inversion.

First, consider one more time the Marmousi model solution using EFWI. The same parameters are used, with the addition of measurement of 200 gravimetric stations in the surface for gravity data. The true velocity model and the starting model are the same as in the AFWI example. The S-wave velocity model uses the ratio $v_p/v_s = \sqrt{3}$ and the density model is scaled using Gardner petrophysical relationship $\rho = \rho_0 V_P^{k_0}$ as shown in Figure 76a and 76b.



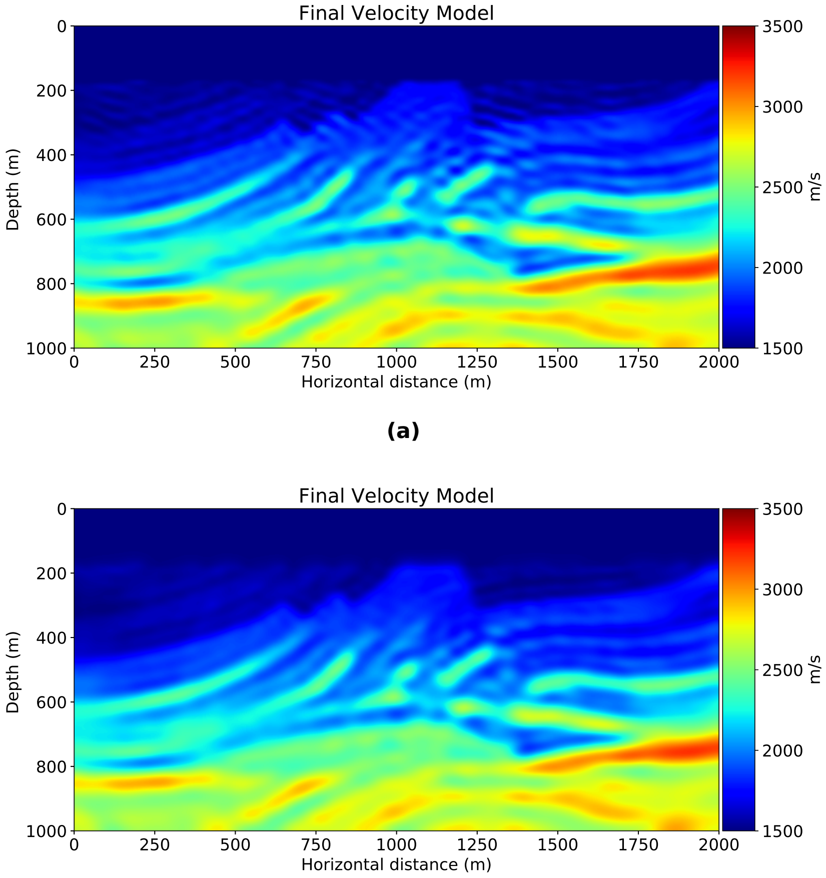
(a)

(b)

(c)

Figure 76. Marmousi true velocity model (a) and its smoothing set as a starting model (b). The final velocity model after cooperative inversion in elastic media.

In order to compare the result of cooperative with separated inversion, 100 iterations were performed showing the final velocity model in Figure 76c. As predicted, the elimination of the spurious artifacts is achieved at the expense of the slight smoothness presented in the model. This is because the cooperative inversion acts as a filter when the gravimetric inversion is performed. To facilitate the comparison of cooperative with separated inversion, the final velocity model is presented for each methodology in Figure 77.



(a)

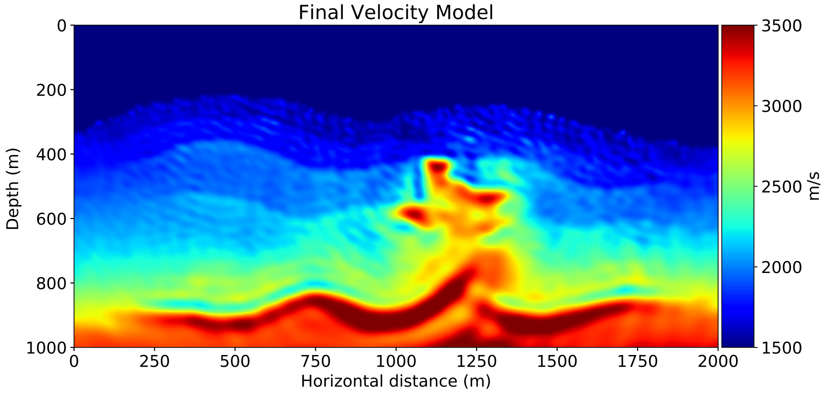
(b)

Figure 77. Marmousi final velocity model for comparisson between conventional elastic FWI (a) the cooperative inversion proposed (b) after 48 iterations.

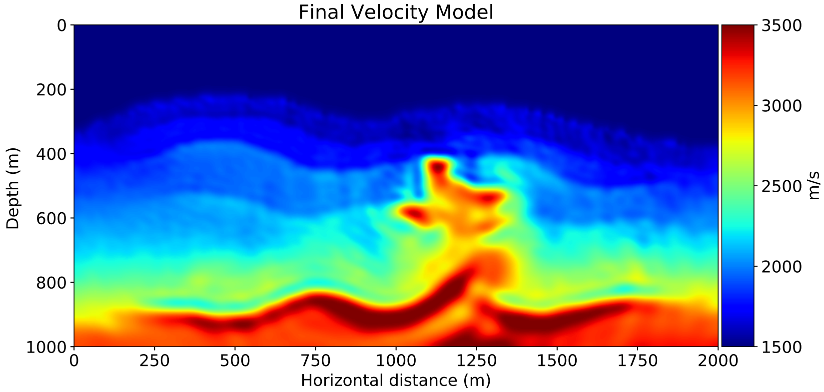
Comparing side by side it is more evident that the cooperative inversion is better as in the top of the velocity model of Figure 77. Notice that the weight of gravimetric inversion has to be controlled mainly by 2 reasons explained in the previous chapters: more weight in gravity data yields to shallower models without stratigraphic information and the Tikhonov regularization could yield a very smooth model.

5.4 Cooperative Inversion: Texas-shape model I in elastic media

To keep consistency with the comparisons, let us apply cooperative inversion to the Texas-shaped model I on an elastic medium. The Marmousi model results provided the insight that the artifacts do not depend on the model structure and the parameters used, therefore the same values and geometries will be used. The true velocity model and the starting model are the same as in the AFWI example and petrophysical considerations will be the same. The final velocity model after 48 iterations is shown in Figure 78.



(a)



(b)

Figure 78. Texas-shape final velocity model I for comparisson between conventional elastic FWI (a) the cooperative inversion proposed (b) after 48 iterations.

It is clear that the artifacts are reduced inside the layers for the result of cooperative inversion (Figure 78b) with respect to the result of separate inversion (Figure 78a).

Finally, the comparison of seismic misfit is presented in Figure 79.

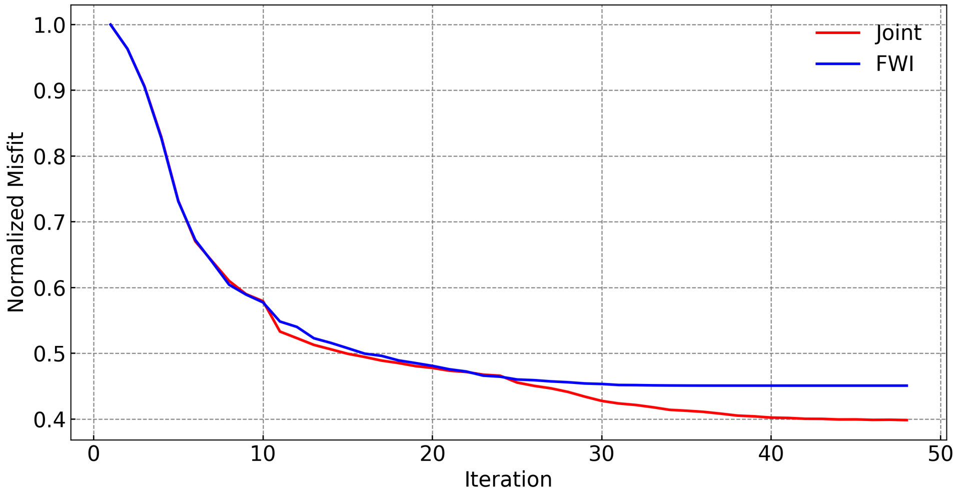


Figure 79. Normalized misfit reduction for seismic data for cooperative inversion (red) and separated inversion (blue).

The misfit exhibits a similar reduction at earlies iterations, however, later the cooperative inversion adjust better the seismic traces given that the elimination of artifacts does not contaminate the data.

This concludes the results section, cooperative scheme works independently of the structure of the model as seen in the results of acoustic FWI. It is worth mentioning that the code is also written for 3D Elastic media, however, currently, we do not have enough computational resources to perform the experiments.

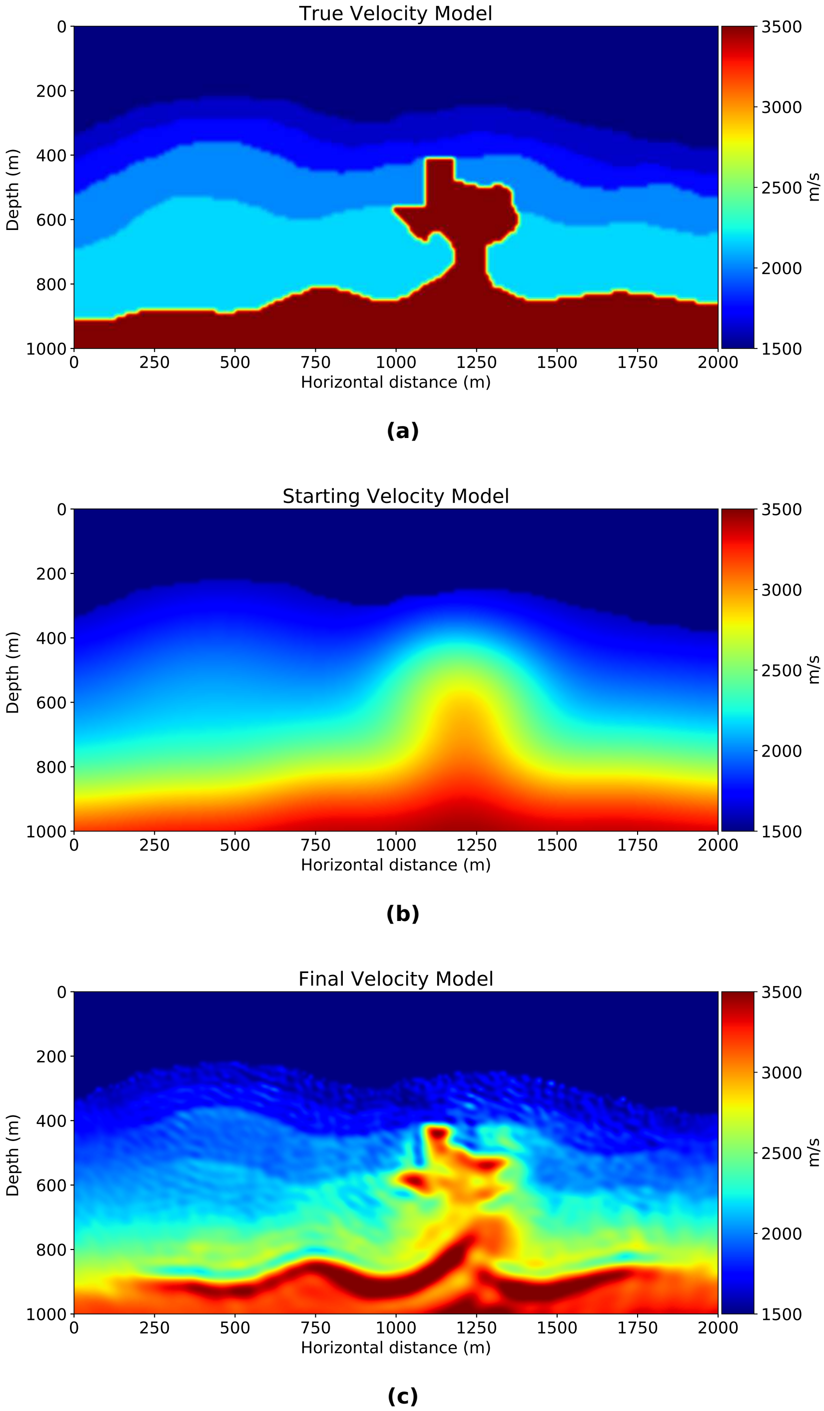


Figure 74. Texas-shape true velocity model II (a) and its smoothing set as a starting model (b). The final velocity model after elastic FWI.

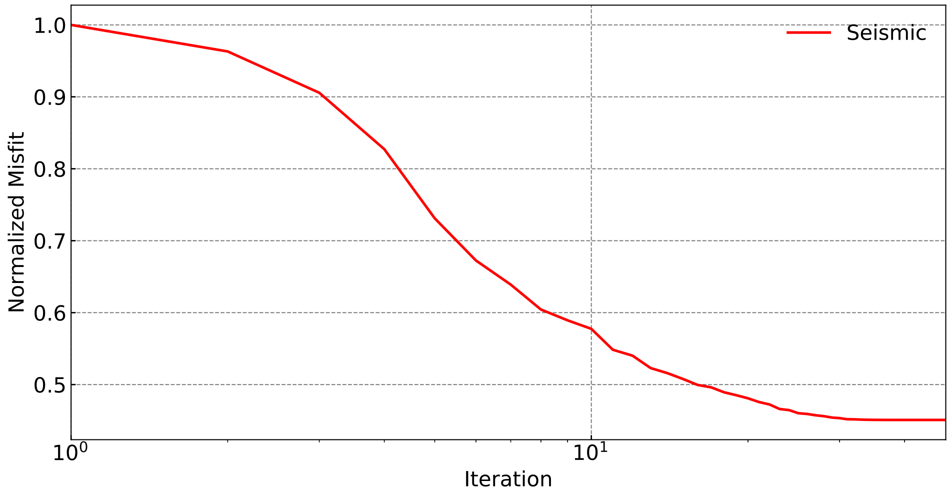


Figure 75. Normalized misfit for seismic data for the Texas-shape model I after elastic FWI.

Chapter 6. Conclusions

Gravimetric and Full Waveform Inversion in a cooperative scheme based on petrophysical relationships can be used to characterize typical geological environments found in real field data such as irregular high velocity-density bodies embedded in complex horizontal layers, thrust and dominoes fault systems, for both acoustic and elastic media.

As long as the starting model is acceptable and FWI has more weight in the cooperative inversion algorithm, both methods converge recovering, for the synthetic models presented, the stratigraphic part, the fault dip, the discordances and the top and shape of the deepest bodies with high velocity (and density). We could accelerate the mass addition and layer recovering on the inner part of the body and the horizontal layers by increasing the weight of gravimetric inversion, however, the recovered stratigraphic part could be severely affected, as shown in the result of conventional gravimetric inversion.

When a cooperative scheme is employed iteratively, the gravimetric inversion is mostly improved when the model is constrained with the information of the model after FWI, i.e., we do not get shallow undesired models. Acoustic and elastic FWI benefit from the sequential application of gravimetric inversion to avoid saturation on some parts of the model (like the inner part of high-velocity bodies) thanks to the embedded Tikhonov regularization. In particular, the elastic FWI reduces the presence of spurious reflectors (artifacts due to shear waves) when the cooperative scheme is performed, since the gravimetric inversion behaves like a filter given the nature of the regularization employed.

We have compared through numerical examples the proposed approach with conventional FWI and joint inversion. The computational cost of the cooperative approach is similar to that of FWI and significantly smaller than that of conventional joint inversion. The main reason for the differences in compute times is that the joint inversion requires computing the Frechet derivatives, whereas FWI and the cooperative approach can be more efficiently implemented using the adjoint method. On the other hand, the cooperative approach successfully incorporates seismic and gravimetric data, yielding a unified model that profits from both geophysical methods.

Another important aspect of our proposed scheme is the use of the same grid structure in gravity and FWI. We employ regularization in gravity inversion. These results are promising and encourage us to continue with more complex and a more realistic model with higher resolution . Note that we employ a petrophysical relationship to relate velocity to density, and use a constant density AFWI. However, our method is also applicable with variable density FWI followed by a soft constraint on velocity-density relationships. Note that some recent FWI methods for salt estimation use constrained model space such as level sets, etc. Our approach can be improved further by using such constrained model spaces and a more accurate petrophysical relationship for salt.

Finally, the cooperative inversion proposed is also coded for three-dimensional models for acoustic and elastic media, as well as for gravimetric problems.

Several ideas emerge through this PhD work which are attractive to implement on a separated investigation, such as: joint inversion using fuzzy c-means or structural constraints for 3D (which is already developed for 2D media) and characterization of 3D salt body structures using additional geophysical methods (i.e. Magnetic data).

Bibliography

Abubakar, A., Gao, G., Habashy, T. M., and Liu, J. (2012). Joint inversion approaches for geophysical electromagnetic and elastic full-waveform data. *Inverse Problems*, **28**(5): 055016.

Aki, K. and Richards, P. G. (2002). *Quantitative seismology*, Vol. 1.

Alford, R., Kelly, K., and Boore, D. M. (1974). Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics*, **39**(6): 834–842.

Banerjee, B. and Das Gupta, S. (1977). Gravitational attraction of a rectangular parallelepiped. *Geophysics*, **42**(5): 1053–1055.

Berdichevskiui, M. N. and Zhdanov, M. S. (1984). *Advanced theory of deep geomagnetic sounding*. Elsevier Science Ltd.

Bhattacharyya, B. and Leu, L.-K. (1977). Spectral analysis of gravity and magnetic anomalies due to rectangular prismatic bodies. *Geophysics*, **42**(1): 41–50.

Blakely, R. J. (1996). *Potential theory in gravity and magnetic applications*. Cambridge university press.

Blom, N., Boehm, C., and Fichtner, A. (2017). Synthetic inversions for density using seismic and gravity data. *Geophysical Journal International*, **209**(2): 1204–1220.

Bosch, M., Meza, R., Jiménez, R., and Hönig, A. (2006). Joint gravity and magnetic inversion in 3d using monte carlo methods. *Geophysics*, **71**(4): G153–G156.

Boulanger, O. and Chouteau, M. (2001). Constraints in 3d gravity inversion. *Geophysical prospecting*, **49**(2): 265–280.

Brocher, T. M. (2005). Empirical relations between elastic wavespeeds and density in the earth's crust. *Bulletin of the Seismological Society of America*, **95**(6): 2081–2092.

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.

Christensen, N. I. and Mooney, W. D. (1995). Seismic velocity structure and composition of the continental crust: A global view. *Journal of Geophysical Research: Solid Earth*, **100**(B6): 9761–9788.

Cohen, G. (2012). *Higher-order numerical methods for transient wave equations*. Springer.

Colombo, D. and Rovetta, D. (2018). Coupling strategies in multiparameter geophysical joint inversion. *Geophysical Journal International*, **215**(2): 1171–1184.

Constable, S. C., Parker, R. L., and Constable, C. G. (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**(3): 289–300.

Crase, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A. (1990). Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, **55**(5): 527–538.

Dahlen, F., Hung, S.-H., and Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes—i. theory. *Geophysical Journal International*, **141**(1): 157–174.

Datta, D. and Sen, M. K. (2016). Estimating a starting model for full-waveform inversion using a global optimization method. *Geophysics*, **81**(4): R211–R223.

De Basabe, J. D. and Sen, M. K. (2007). Grid dispersion and stability criteria of some common finite-element methods for acoustic and elastic wave equations. *Geophysics*, **72**(6): T81–T95.

De Basabe, J. D. and Sen, M. K. (2015). A comparison of finite-difference and spectral-element methods for elastic wave propagation in media with a fluid-solid interface. *Geophysical Journal International*, **200**(1): 278–298.

Edwards, R., Nobes, D., and Gómez-Treviño, E. (1984). Offshore electrical exploration of sedimentary basins: The effects of anisotropy in horizontally isotropic, layered media. *Geophysics*, **49**(5): 566–576.

Fichtner, A. (2010). *Full seismic waveform modelling and inversion*. Springer Science & Business Media.

Fichtner, A. and Trampert, J. (2011). Resolution analysis in full waveform inversion. *Geophysical Journal International*, **187**(3): 1604–1624.

Gallardo, L. A. and Meju, M. A. (2003). Characterization of heterogeneous near-surface materials by joint 2d inversion of dc resistivity and seismic data. *Geophysical Research Letters*, **30**(13).

Gallardo, L. A. and Meju, M. A. (2004). Joint two-dimensional dc resistivity and seismic travel time inversion with cross-gradients constraints. *Journal of Geophysical Research: Solid Earth (1978–2012)*, **109**(B3).

Gao, G., Abubakar, A., and Habashy, T. M. (2012). Joint petrophysical inversion of electromagnetic and full-waveform seismic data. *Geophysics*, **77**(3): WA3–WA18.

García-Abdeslem, J. (2005). The gravitational attraction of a right rectangular prism with density varying with depth following a cubic polynomial. *Geophysics*, **70**(6): J39–J42.

Gardner, G., Gardner, L., and Gregory, A. (1974). Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics*, **39**(6): 770–780.

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

Godfrey, N., Beaudoin, B., and Klemperer, S. (1997). Ophiolitic basement to the great valley forearc basin, california, from seismic and gravity data: Implications for crustal growth at the north american continental margin. *Geological Society of America Bulletin*, **109**(12): 1536–1562.

Grant, F. S. and West, G. F. (1965). Interpretation theory in applied geophysics. *McGraw-Hill Book*.

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.

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

Hermanns, M. (2002). Parallel programming in fortran 95 using openmp. *Universidad Politecnica de Madrid, Spain*.

Hole, J. (1992). Nonlinear high-resolution three-dimensional seismic travel time tomography. *Journal of Geophysical Research: Solid Earth (1978–2012)*, **97**(B5): 6553–6562.

Hole, J. and Zelt, B. (1995). 3-d finite-difference reflection traveltimes. *Geophysical Journal International*, **121**(2): 427–434.

Ikelle, L. and Amundsen, L. (2005). *Introduction to Petroleum Seismology*. Number v. 12 in: Introduction to Petroleum Seismology. Society of Exploration Geophysicists.

Jackson, D. D. (1972). Interpretation of inaccurate, insufficient and inconsistent data. *Geophysical Journal International*, **28**(2): 97–109.

Johnson, L. R. and Litehiser, J. J. (1972). A method for computing the gravitational attraction of three-dimensional bodies in a spherical or ellipsoidal earth. *Journal of Geophysical Research*, **77**(35): 6999–7009.

Kelly, K., Ward, R., Treitel, S., and Alford, R. (1976). Synthetic seismograms; a finite-difference approach. *Geophysics*, **41**(1): 2–27.

Kolb, P., Collino, F., and Lailly, P. (1986). Pre-stack inversion of a 1-d medium. *Proceedings of the IEEE*, **74**(3): 498–508.

Komatitsch, D. and Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical journal international*, **139**(3): 806–822.

Lees, J. M. and VanDecar, J. (1991). Seismic tomography constrained by bouguer gravity anomalies: Applications in western washington. *Pure and Applied Geophysics*, **135**(1): 31–52.

Lelièvre, P. G., Farquharson, C. G., and Hurich, C. A. (2012). Joint inversion of seismic traveltimes and gravity data on unstructured grids with application to mineral exploration. *Geophysics*, **77**(1): K1–K15.

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

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

Lin, F.-C., Schmandt, B., and Tsai, V. C. (2012). Joint inversion of rayleigh wave phase velocity and ellipticity using usarray: Constraining velocity and density structure in the upper crust. *Geophysical Research Letters*, **39**(12).

Lin, W. and Zhdanov, M. (2017). Joint inversion of seismic and gravity gradiometry data using gramian constraints. In: *SEG Technical Program Expanded Abstracts 2017*. Society of Exploration Geophysicists, pp. 1734–1738.

Lines, L. R., Schultz, A. K., and Treitel, S. (1988). Cooperative inversion of geophysical data. *Geophysics*, **53**(1): 8–20.

Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, **66**(3): 639–666.

Marfurt, K. J. (1984). Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, **49**(5): 533–549.

Martin, R., Chevrot, S., Komatitsch, D., Seoane, L., Spangenberg, H., Wang, Y., Dufréchou, G., Bonvalot, S., and Bruinsma, S. (2017). A high-order 3-d spectral-element method for the forward modelling and inversion of gravimetric data—application to the western pyrenees. *Geophysical Journal International*, **209**(1): 406–424.

McGillivray, P. and Oldenburg, D. (1990). Methods for calculating fréchet derivatives and sensitivities for the non-linear inverse problem: A comparative study1. *Geophysical Prospecting*, **38**(5): 499–524.

Menichetti, V. and Guillen, A. (1983). Simultaneous interactive magnetic and gravity inversion. *Geophysical Prospecting*, **31**(6): 929–944.

Menke, W. (2018). *Geophysical data analysis: Discrete inverse theory*. Academic press.

Moorkamp, M., Heincke, B., Jegen, M., Roberts, A. W., and Hobbs, R. W. (2011). A framework for 3-d joint inversion of mt, gravity and seismic refraction data. *Geophysical Journal International*, **184**(1): 477–493.

Mora, P. (1987). Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, **52**(9): 1211–1228.

Nagy, D. (1966). The gravitational attraction of a right rectangular prism. *Geophysics*, **31**(2): 362–371.

Nagy, D., Papp, G., and Benedek, J. (2000). The gravitational potential and its derivatives for the prism. *Journal of Geodesy*, **74**(7-8): 552–560.

Nocedal, J. and Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.

Oldham, R. D. (1906). *The Quarterly Journal of the Geological Society of London*. The Society.

OpenMP Architecture Review Board (2008). OpenMP application program interface version 3.0.

Operto, S., Virieux, J., Amestoy, P., L'Excellent, J.-Y., Giraud, L., and Ali, H. B. H. (2007). 3d finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study. *Geophysics*, **72**(5): SM195–SM211.

Paasche, H. and Tronicke, J. (2007). Cooperative inversion of 2d geophysical data sets: A zonal approach based on fuzzy c-means cluster analysis. *Geophysics*, **72**(3): A35–A39.

Pica, A., Diet, J., and Tarantola, A. (1990). Nonlinear inversion of seismic reflection data in a laterally invariant medium. *Geophysics*, **55**(3): 284–292.

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.

Portniaguine, O. and Zhdanov, M. S. (1999). Focusing geophysical inversion images. *Geophysics*, **64**(3): 874–887.

Portniaguine, O. and Zhdanov, M. S. (2002). 3-d magnetic inversion with data compression and image focusing. *Geophysics*, **67**(5): 1532–1541.

Pratt, R. G. and Worthington, M. (1990). Inverse theory applied to multi-source crosshole tomography. part 1: acoustic wave-equation method. *Geophysical prospecting*, **38**(3): 287–310.

Romero, A. and Gallardo, L. A. (2015). Borehole constrained inversion of geophysical data based on fuzzy clustering. In: *SEG Technical Program Expanded Abstracts 2015*. Society of Exploration Geophysicists, pp. 1792–1796.

Roy, L., Sen, M. K., McIntosh, K., Stoffa, P. L., and Nakamura, Y. (2005). Joint inversion of first arrival seismic travel-time and gravity data. *Journal of Geophysics and Engineering*, **2**(3): 277.

Sen, M. K. (2016). *GEO 390D - Lecture notes in advanced seismology*. University of Texas at Austin, n.d. Web.

Sen, M. K. and Stoffa, P. L. (2013). *Global optimization methods in geophysical inversion*. Cambridge University Press.

Serra, S. (1977). Styles of deformation in the ramp regions of overthrust faults. *Wyoming Geological Association*, 487–498.

Shin, C. and Cha, Y. H. (2008). Waveform inversion in the laplace domain. *Geophysical Journal International*, **173**(3): 922–931.

Shin, C. S. (1988). *Nonlinear elastic wave inversion by blocky parameterization*. PhD. Thesis, University of Tulsa.

Sirgue, L. and Pratt, R. G. (2004). Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics*, **69**(1): 231–248.

Smith, N. C. and Vozoff, K. (1984). Two-dimensional dc resistivity inversion for dipole-dipole data. *Geoscience and Remote Sensing, IEEE Transactions on*, (1): 21–28.

Talwani, M. (1965). Computation with the help of a digital computer of magnetic anomalies caused by bodies of arbitrary shape. *Geophysics*, **30**(5): 797–817.

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8): 1259–1266.

Tarantola, A. (1986). A strategy for non linear inversion of seismic reflection data. *Geophysics*.

Tarantola, A. (1987). Inverse problems theory. *Methods for Data Fitting and Model Parameter Estimation. Elsevier, Southampton*.

Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms including elasticity and attenuation. *Pure and Applied Geophysics*, **128**(1-2): 365–399.

Telford, W. M. and Sheriff, R. E. (1990). *Applied geophysics*, Vol. 1. Cambridge University press.

Thakur, N. K. and Rajput, S. (2010). *Exploration of Gas Hydrates: Geophysical Techniques*. Springer Science & Business Media.

Tikhonov, A. and Arsenin, V. (1977). *Solutions of ill-posed problems*. Scripta series in mathematics. Winston.

Tondi, R., Achauer, U., Landes, M., Daví, R., and Besutiu, L. (2009). Unveiling seismic and density structure beneath the vrancea seismogenic zone, romania. *Journal of Geophysical Research: Solid Earth*, **114**(B11).

Tromp, J., Tape, C., and Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, **160**(1): 195–216.

Tsoulis, D. (2012). Analytical computation of the full gravity tensor of a homogeneous arbitrarily shaped polyhedral source using line integrals. *Geophysics*, **77**(2): F1–F11.

Versteeg, R. (1994). The marmousi experience: Velocity model determination on a synthetic complex data set. *The Leading Edge*, **13**(9): 927–936.

Vidale, J. (1988). Finite-difference calculation of travel times. *Bulletin of the Seismological Society of America*, **78**(6): 2062–2076.

Vidale, J. E. (1990). Finite-difference calculation of traveltimes in three dimensions. *Geophysics*, **55**(5): 521–526.

Vigh, D. and Starr, E. W. (2008). 3d prestack plane-wave, full-waveform inversion. *Geophysics*, **73**(5): VE135–VE144.

Vigh, D., Starr, E. W., *et al.* (2008). Comparisons for waveform inversion, time domain or frequency domain? In: *2008 SEG Annual Meeting*. Society of Exploration Geophysicists.

Virieux, J. (1986). P-sv wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, **51**(4): 889–901.

Virieux, J. and Madariaga, R. (1982). Dynamic faulting studied by a finite difference method. *Bulletin of the Seismological Society of America*, **72**(2): 345–369.

Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**(6): WCC1–WCC26.

Virieux, J., Operto, S., Ben-Hadj-Ali, H., Brossier, R., Etienne, V., Sourbier, F., Giraud, L., and Haidar, A. (2009). Seismic wave modeling for seismic imaging. *The Leading Edge*, **28**(5): 538–544.

Vozoff, K. and Jupp, D. (1975). Joint inversion of geophysical data. *Geophysical Journal International*, **42**(3): 977–991.

Weidelt, P. (1975). Electromagnetic induction in three-dimensional structures. *J. Geophys*, **41**(85): 109.

Werner, R. A. (1994). The gravitational potential of a homogeneous polyhedron or don't cut corners. *Celestial Mechanics and Dynamical Astronomy*, **59**(3): 253–278.

Zeyen, H. and Pous, J. (1993). 3-d joint inversion of magnetic and gravimetric data with a priori information. *Geophysical Journal International*, **112**(2): 244–256.

Zhdanov, M. S. (2002). *Geophysical inverse theory and regularization problems*, Vol. 36. Elsevier.

Zhdanov, M. S. and Fang, S. (1996). Quasi-linear approximation in 3-d electromagnetic modeling. *Geophysics*, **61**(3): 646–665.

Zhdanov, M. S., Ellis, R., and Mukherjee, S. (2004). Three-dimensional regularized focusing inversion of gravity gradient tensor component data. *Geophysics*, **69**(4): 925–937.