Wide-Angle, Full Waveform Inversion with a Sparse Ocean-Bottom Seismometer Dataset, Imaging the Cyprus Arc from the Eratosthenes Seamount to the Hecataeus Rise

by

© **Christopher Williams**

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science.

Department of Earth Sciences

Memorial University

June 2020

St. John's, Newfoundland and Labrador, Canada

Abstract

Tectonophysicists working with wide-angle seismic reflection and refraction data generally rely on tomographic inversion and forward modelling techniques to produce P-wave velocity models of the subsurface. Only recently, researchers have produced high-resolution P-wave velocity models using full-waveform inversion (FWI). However, they use ocean-bottom seismometer (OBS) datasets with a dense instrument spacing. This study revisits a sparsely acquired OBS dataset from the geologically complex Eastern Mediterranean to explore the feasibility of applying FWI to such a dataset. With frequency domain viscoacoustic FWI, we observe an adequate decrease in the objective function, suggesting that a high-resolution velocity model for the upper 15 km is obtainable through FWI with a typical OBS dataset. This decrease in the objective function does not preclude convergence to a local minimum. The recovered FWI model suffers from uncertainties in the starting model related to a complex geological environment. Cognizant of model uncertainties, we interpret an accretionary prism between the southern Eratosthenes continental block and the northern Hecataeus Rise from the final FWI model.

For my great Nan Jessie, who celebrates her 100^{th} birthday this year.

Acknowledgements

I want to acknowledge family and friends for providing unconditional support. I want to thank members of Dr. Alison Malcolm's and Dr. Kim Welford's research groups for persistent positive feedback. I would also like to thank Dr. Romain Brossier from the Université Grenoble Alpes, Grenoble, France, for help compiling TOY2DAC and useful discussion through e-mail and in person. I would also like to thank Andrzej Górszczyk from the Université Grenoble Alpes, Grenoble, France, for assistance regarding the implementation of Laplace-Fourier domain full-waveform inversion in TOY2DAC. I acknowledge Dr. Oliver Stueker, Paul Sherren, and Compute Canada for providing and assisting me with the computational resources required to complete this research. We thank Captain Friedhelm v. Staa and Captain Matthias Günther, the officers and crew of RV Maria S. Merian, for their support during cruise MSM14/3, which was funded by Deutsche Forschungsgesellschaft (DFG). We thank the Geological Survey of Canada for the loan of the airguns and Dr. A. Aksu for designing their mobilization system for this cruise.

Finally, I thank my supervisors Dr. Alison Malcolm and Dr. Kim Welford for their scientific input, patience, and good jokes.

List of tables

List of figures

List of symbols

A	Complex-valued impedance matrix
B	Gauss-Newton approximation to the Hessian matrix
c	Velocity
C_{ijkl}	Hookean tensor
d	Observed data
f	Objective function
G	Displacement
H	Hessian matrix
J	Jacobian matrix
m	Model parameters
O	Maximum offset where reflections are observed
P	Pseudo-Hessian matrix
Q	Attenuation
R	NLCG accounting for previous estimations
S	Seismic source
T	Traction
u	Seismic wavefield
v_p	P-wave velocity
v_s	S-wave velocity
V	l-BFGS term
W	Data-weighting matrix for least-squares misfit formulation
y	l-BFGS term
α	Step-length
β	l-BFGS term
γ	l-BFGS term
ΔX	OBS spacing
ϵ	Damping in the Pseudo-Hessian matrix
κ	Bulk modulus
λ	Lamé parameter
μ	Lamé parameter
ν	Hessian vector

ρ Density
χ Adjoint operator
ω Angular frequency

List of abbreviations

ECB	Eratosthenes Continental Block
FAT	First Arrival Tomography
FTB	Fold-Thrust Belt
FWI	Full Waveform Inversion
GSC	Geological Survey of Canada
LU	Lower-Upper
NAF	North Anatolian Fault
NLCG	Non-Linear Conjugate Gradient
NRMS	Noise Root Mean Square
OBS	Ocean-Bottom Seismometer
QC	Quality Control
TTI	Tilted Transverse Isotropy
HTI	Horizontal Transverse Isotropy
RTM	Reverse-time migration
SSZ	Supra-Subduction Zone
VTI	Vertical Transverse Isotropy
WARR	Wide-Angle Reflection and Refraction

Chapter 1

Introduction and Overview

1.1 Introduction

The Eastern Mediterranean Sea is the site of a significant incipient collision between the African and Eurasian plates (Kempler, 1998). The area is geologically complex as convergent, transpressive, and transtensional plate tectonic processes interplay. The geological framework of the area is determined by the relative motions of four tectonic plates: the African plate to the south, the Anatolian plate to the north, the Arabian plate to the east, and the European plate to the north-west (Figure 1.1). Fault plane solutions have been used by various authors to determine the regional sense of plate motions (McKenzie, 1972; Papazachos & Papaioannou, 1999; Pilidou et al., 2004), and recent GPS measurements also aid in determining the nature of the plate boundaries (Reilinger et al., 2006; Le Pichon & Kreemer, 2010). Nocquet (2012) provides a review of recent GPS results. These studies conclude that the African and Arabian plates are converging northward towards the Anatolian and Eurasian plates. In the southern Eastern Mediterranean Basin, the northern portion of the African plate remains a passive margin (Figure 1.1).

The evolution of the Eastern Mediterranean region generally follows that of the entire Mediterranean, where Gondwanan-derived microcontinents are transferred to the southern Eurasian plate at various times both before Pangean amalgamation (Stampfli & Borel, 2002; Stampfli et al., 2013) and following the formation of Pangea (Garfunkel, 1998; Stampfli & Borel, 2002; Golonka, 2007; Frizon de Lamotte et al.,

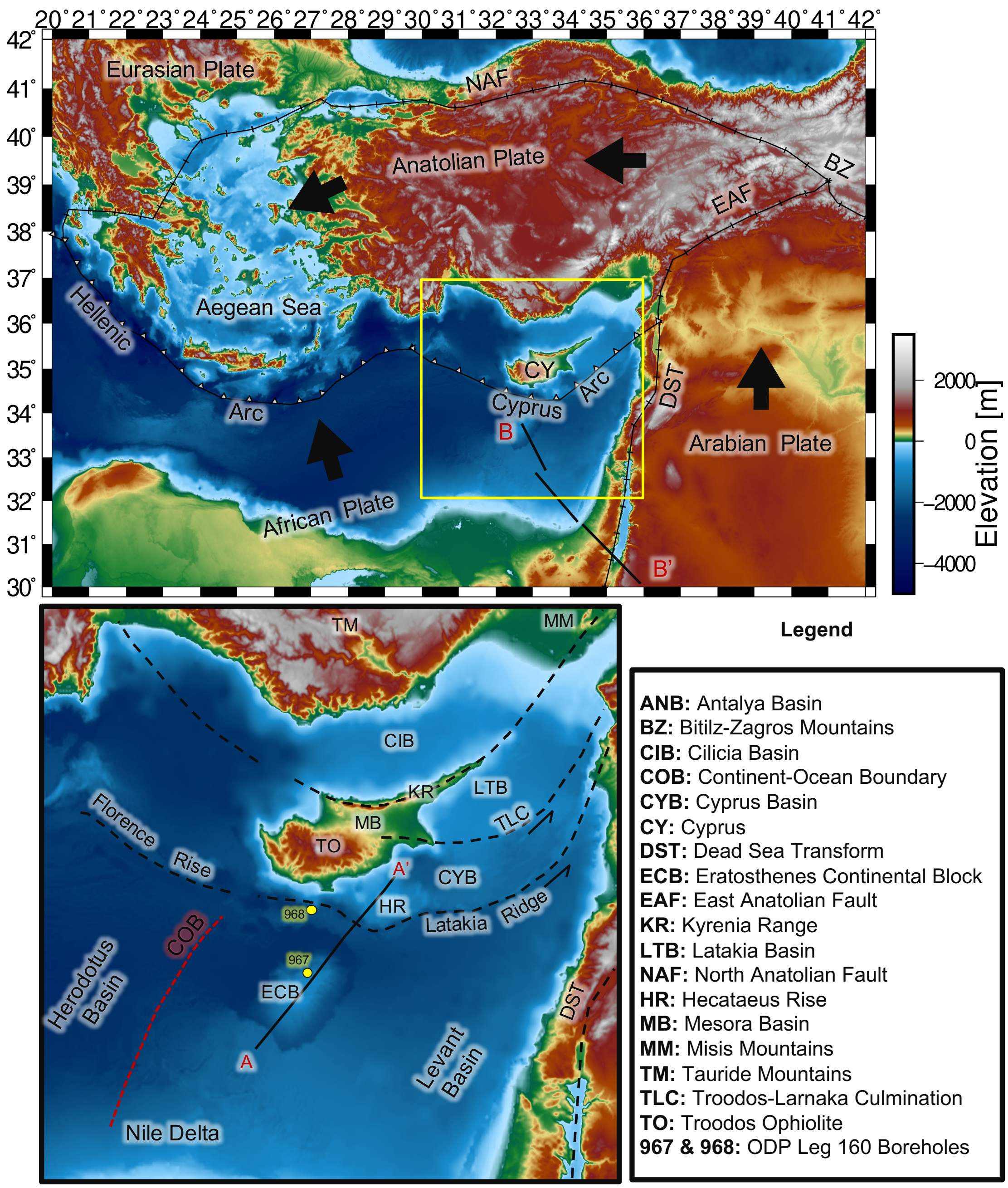

Figure 1.1: A regional (top) and local (bottom) elevation map defines the study area and denotes the main tectonic elements. Tectonic plate boundaries in the regional map are from Bird (2003). In the local map, the profile A-A' is Eastern Mediterranean line 1 from Welford et al. (2015a), shown in Figure 1.2. Profile B-B' in the regional map is the location of a mega transect from Netzeband et al. (2006), shown in Figure 1.4. The COB denotes the approximate region separating continental and oceanic crust (Granot, 2016).

2011; van Hinsbergen et al., 2020). The Eratosthenes Continental Block (ECB), shown in Figure 1.1, is proposed to have been rifted off the Gondwana margin, but was not transferred to the Eurasian Plate (Robertson, 1998). The ECB is commonly termed the Eratosthenes Seamount in the published literature, but this name is geologically incorrect. A seamount is a bathymetric high that formed through past volcanism. As recent seismic investigations (Welford et al., 2015a; Symeou et al., 2018) and plate reconstructions van Hinsbergen et al. (2020) reveal that volcanic processes are not primarily responsible for the formation of this bathymetric feature, it is defined to be the ECB in this thesis. The timing and method of emplacement of the ECB are contentious, as its current position in the Cyprus Arc represents the onset of a continent-continent collision between the African and Eurasian plates. Furthermore, the convergence history between the ECB, the Cyprus Arc, and the Hecataeus Rise is contentious, along with the formation of the Cyprus Arc itself (Segev et al., 2018).

Three wide-angle seismic reflection and refraction profiles were acquired in 2010 to focus on the region of convergent continental tectonics between the ECB and the Cyprus Arc (Welford et al., 2015a,b). These crustal-scale profiles show P-wave velocities produced from tomographic and forward ray-tracing techniques. Welford et al. (2015a) discusses two lines, line 1 and line 1a. Line 1a is angled to the east of line 1 and both lines extend from the ECB to the Hecataeus Rise. The P-wave crustal velocity model for line 1 is shown in Figure 1.2. Line 2 from Welford et al. (2015b) extends from the Levant Basin in the south-east to the Hecataeus Rise in the north-west.

The wide-angle reflection and refraction profiles over the ECB from Welford et al. (2015a) are the first investigations of the crustal structure of the area since Makris et al. (1983). Welford et al. (2015a) interpret the Moho to be at 27 km beneath the ECB, and 20 km beneath the Hecataeus Rise constrained by 2D gravity forward modelling. The crustal affinity of the ECB and the crustal affinity of the Hecataeus Rise are both interpreted to be continental Welford et al. (2015a,b). Summarizing the interpretations from Eastern Mediterranean line 1 (Figure 1.2) after Welford et al. (2015a), a transform zone is inferred between the ECB and Hecataeus Rise. Also, two high-velocity lower crustal blocks are present to the north of the ECB, and its crustal structure varies laterally as the middle layer of the continental crust thickens toward the north. The modelled high-velocity blocks within the Cyprus Arc are either a remnant of Tethyan oceanic crust or igneous intrusions (Welford et al., 2015a).

While traditional crustal-scale model building techniques are compelling, full-waveform inversion (FWI) has become an increasingly popular technique for generating high-resolution velocity models at the crustal-scale. The increase in resolution in models made with FWI over those made with traditional tomographic techniques are due to FWI modelling more of the seismic wavefield than just a first arrival time Virieux & Operto (2009). Some recent studies are successful in inverting wide-angle reflection and refraction data (WARR) (Kamei et al., 2012; Davy et al., 2017; Górszczyk et al., 2017). These studies all successfully perform FWI on a WARR dataset but use unusually dense OBS datasets. Typical WARR datasets are acquired with an OBS spacing greater than 5 km, whereas the studies of Kamei et al. (2012); Davy et al. (2017), and Górszczyk et al. (2017) all use an OBS spacing of less than 5 km.

Eastern Mediterranean Line 1 from Welford et al. (2015a) has an OBS spacing of approximately 12 km. As such this dataset is sparse in comparison to WARR datasets from which FWI successfully generates a high-resolution velocity model (Kamei et al., 2012; Davy et al., 2017; Górszczyk et al., 2017). For example, Eastern Mediterranean

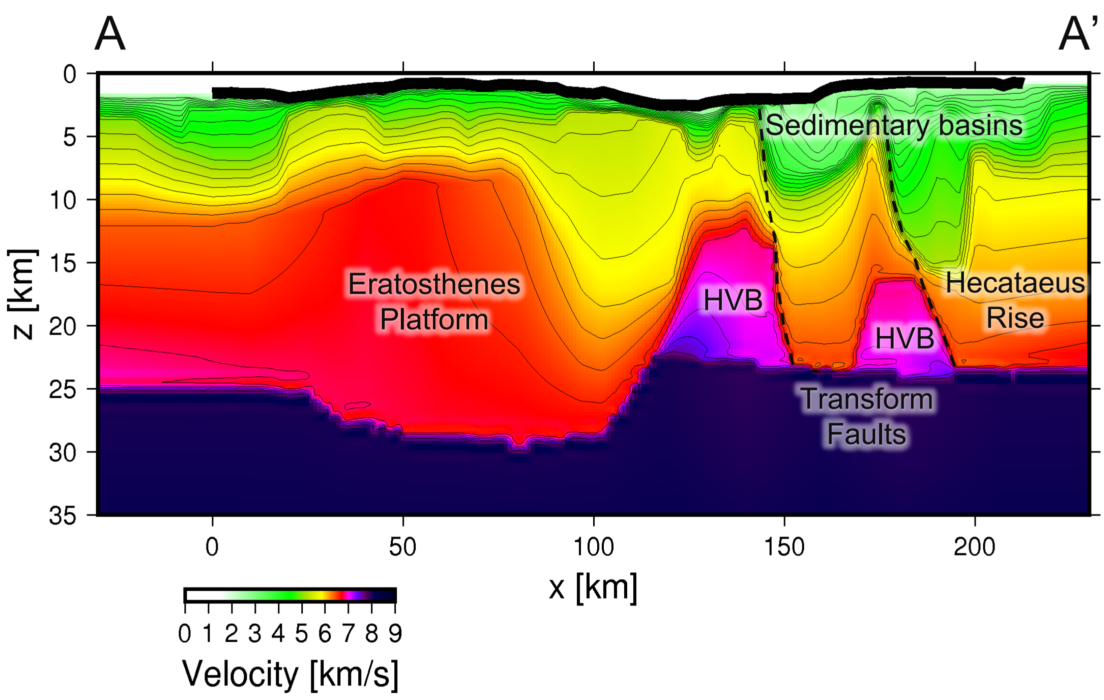

Figure 1.2: Eastern Mediterranean line 1 after Welford et al. (2015a) is shown with a 2:1 vertical exaggeration and labelled with crustal interpretations. The location of Line 1, denoted by A-A', is shown in Figure 1.1. HVB, high-velocity body.

Line 1 from Welford et al. (2015a) spans 260 km in length and contains 21 OBS record signals from approximately 10,000 shots. In comparison, Górszczyk et al. (2017) completes FWI on a model spanning 100 km with 100 OBS and more than 120,000 first break picks. We propose to image the crustal-scale velocity structure from the ECB, across the Cyprus Arc, to the Hecataeus Rise using FWI for Eastern Mediterranean Line 1 (Figure 1.1). If successful on this Eastern Mediterranean dataset this would be the first example of such a sparse application of FWI to the best of our knowledge.

The research goals for this study may be broadly defined as follows.

- Investigate whether FWI is possible on such a sparse crustal-scale dataset.
- Compare a recovered FWI model with the velocity model from Welford et al. (2015a).
- Make any additional geologic interpretations on the recovered FWI model.
- Provide a recommendation for performing FWI on similar datasets.

Many of these research goals rely on the ability to apply full-waveform techniques to this sparse Eastern Mediterranean dataset. FWI will only converge to the global minimum if the starting model is close enough to the global minimum. This problem is more critical for a sparse dataset, such as the one studied here than it would be for a denser dataset. Additionally, the recovered FWI model may be too similar to that produced by Welford et al. (2015a), such that new geologic interpretations may not be possible.

1.1.1 How to Read This Thesis

This thesis is written in a manuscript format (thesis by paper(s)), thus differing from the traditional thesis format. **Chapter 1** is the introduction. The introduction addresses the scientific problem, reviews pertinent literature, and introduces FWI methodology. Any readers keen on learning more about Eastern Mediterranean geology, or FWI are encouraged to read Chapter 1. **Chapter 2** is a co-authorship statement for the following manuscript. This research is published in two conference abstracts as well. Chapter 2 further provides a statement of contribution for both conference abstracts and the longer manuscript that forms the core of the thesis. **Chapter 3** contains a single manuscript written to summarize the scientific results of this study. A reader looking for a concise presentation of the scientific methods and results for this project are encouraged to read this chapter and need not read the rest of the thesis. **Chapter 4** is the summary which consists of a discussion interpreting and suggesting improvements for the scientific results. A reader looking for ideas to improve their FWI results or looking for ideas to improve the results of this study are encouraged to read the summary. Finally, the Appendices contain information on data processing, tomography, synthetic tests, and real-data tests. Readers interested in reproducing the results of this study, or completing an FWI study of their own are encouraged to read the Appendices.

1.2 Geological Review of the Eastern Mediterranean Basin

We begin by outlining some geological disagreement and geological consensus in the region. A probe into the geological literature at the regional, crustal, and basin scales will follow. Teleseismic surveys, seismic refraction surveys, and seismic reflection surveys are discussed for each scale of investigation.

1.2.1 Introduction

The ECB is a geological structure interpreted to be a continental fragment rifted off the northern margin of Gondwana. Its stratigraphic evolution is constrained to the Early Cretaceous through Leg 160 of the Ocean Drilling Program (Robertson & Shipboard Scientific Party, 1996; Kempler, 1998; Robertson, 1998; Robertson et al., 1998). The ECB currently resides within the Cyprus Arc, and its location within the arc has played a role in altering the tectonic configuration of the Eastern Mediterranean (Schattner, 2010; Schildgen et al., 2014). The timing and mechanism of rifting the ECB from the northern Gondwanan margin are controversial. Some reconstructions initiate rifting in the Levant as early as the Permian (Garfunkel, 1998; Stampfli & Borel, 2002; Netzeband et al., 2006; Gardosh et al., 2010; Robertson et al., 2012), and a recent study initiates rifting in the Cretaceous (Segev et al., 2018). Some authors use an intermediate Early-Middle Mesozoic (Triassic-Jurassic) age for the initiation of rifting in the Levant (Frizon de Lamotte et al., 2011; Montadert et al., 2014; van Hinsbergen et al., 2020). Rifting of the ECB from the Gondwanan margin has been described as related to the Tauride Block (Garfunkel, 1998; Gardosh et al., 2010; van Hinsbergen et al., 2020) or back-arc extension (Segev et al., 2018). The Tauride Block composes much of southern Anatolia, and is currently the Tauride Mountains (Figure 1.1). While the subduction of oceanic lithosphere accommodates convergence between the African and Eurasian plates, the position of the ECB relative to the subduction zone is uncertain. In one scenario, oceanic crust is consumed by the Cyprus Arc, generally in the area of the of the Misis-Kyrenia fold-thrust belt (FTB), since the onset of convergence in the Cretaceous (Robertson, 1998; Kempler, 1998; Calon et al., 2005a,b; Frizon de Lamotte et al., 2011). An alternate scenario accretes the Troodos ophiolite to the northern margin of the African plate during the Cretaceous while subduction of oceanic crust occurs north of the Cyprus Arc in the Cilicia Basin (McPhee & van Hinsbergen, 2019). A Miocene compressional event formed the present-day structure of the Cyprus Arc as extended continental crust from the northern African plate had under-thrust Anatolia (McPhee & van Hinsbergen, 2019; van Hinsbergen et al., 2020). However, the Cilicia Basin is described by Schattner (2010) as being formed by back-arc extension. The simultaneous existence of a subduction zone to the north and south of the Troodos microplate in the Upper Cretaceous is suggested by Robertson et al. (2012) as well.

Convergence between the coupled African and Arabian plates and the Eurasian plate began in the Upper Cretaceous, denoted by the widespread obduction of ophiolites (Garfunkel, 1998; Gardosh et al., 2010; Frizon de Lamotte et al., 2011; Robertson et al., 2012; van Hinsbergen et al., 2020). Relevant to this study, the Troodos ophiolite, located on the Island of Cyprus (Figure 1.1), formed in a supra-subduction zone (SSZ) setting, or by seafloor spreading above an oceanic subduction zone (Bailey et al., 2000). There is also general agreement that there are two stages of deformation within the Cyprus Arc: a Miocene compressional stage (Montadert et al., 2014; McPhee & van Hinsbergen, 2019) and a Pliocene to recent transpressional stage (Calon et al., 2005a,b; Hall et al., 2005a,b; Montadert et al., 2014; Reiche & Hübscher, 2015; Symeou et al., 2018). Earlier Eocene deformation is associated with the more northerly Troodos-Larnaka culmination and the Misis-Kyrenia fold-thrust belt (FTB) (Calon et al., 2005a,b). This stage of deformation may be related to the onset of the Bitlis-Zagros slab retreat (Frizon de Lamotte et al., 2011) or the formation of a forebulge (McPhee & van Hinsbergen, 2019) during the onset of eastern continent-continent collision between the Arabian and Eurasian plates. The southerly Latakia Ridge may have formed recently during the compressional phase in the middle Miocene (Symeou et al., 2018).

The ECB may be partially responsible for the uplift of the island of Cyprus (Robertson, 1998), Central Anatolia (Schildgen et al., 2014), and reactivation of compressive structures through transpressive tectonics in the Latakia Basin and Latakia Ridge regions (Hall et al., 2005a,b; Schattner, 2010). The timing of this tectonic regime change along the Latakia Ridge and Latakia Basin correlates with the onset of the counter-clockwise directed tectonic escape of the Anatolian microplate (Robertson, 1998; Hall et al., 2005a; Faccenna et al., 2006; Schattner, 2010; Schildgen et al., 2014).

1.2.2 Regional Scale: A Review of Teleseismic Surveys

Representative teleseismic studies presented in this section provide a general understanding of Eastern Mediterranean geodynamics. The region of investigation spans from the Hellenic Arc in the west to the Bitlis-Zagros Orogen in the east (Figure 1.1). This region represents a temporal tectonic evolution from east to west.

The formation of the North Anatolian Fault (NAF) may be related to a past rupture in the subducting Tethyan oceanic slab beneath the Anatolian plate (Faccenna et al., 2006). Faccenna et al. (2006) propose this based on a P-wave tomographic model of the mantle beneath the Alpine-Mediterranean region from Piromallo & Morelli (2003), and laboratory modelling. The tomographic models from Piromallo & Morelli (2003) suggest a non-continuous slab from the Bitlis-Zagros Orogen to the Hellenic Arc, where the subducting slab becomes progressively ruptured from the island of Cyprus to the Bitlis-Zagros Orogen. Through laboratory modelling, they conclude that the progressive rupture of the western subducting slab would allow for conditions that permit western tectonic escape accommodated by strike-slip motion along the NAF (Faccenna et al., 2006). Faccenna et al. (2006) propose three stages of tectonic evolution. Eastern Bitlis collision and western Aegean extension define the lower to middle Miocene and the NAF forms during the Miocene to Pliocene. Since the upper

Pleistocene, transpressive motion to the east and transtensional motion to the west characterize the NAF.

Schildgen et al. (2014) link regional uplift experienced across Anatolia to slab break-off. In general agreement with Faccenna et al. (2006), the subducting slab beneath the Bitlis-Zagros Orogen is interpreted by Schildgen et al. (2014) to have detached in the middle Miocene, related to continent-continent collision between the Arabian and Eurasian plates. Schildgen et al. (2014) interpret shallow slab break off beneath Cyprus during the Pliocene-Pleistocene. Numerical modelling shows that shallow slab detachment produces topographical uplift similar to the present-day elevation profile across central Anatolia. The presence of the ECB within the Cyprus Arc contributes to Pliocene-Pleistocene shallow-slab and, therefore, regional uplift in Anatolia (Schildgen et al., 2014).

A final regional study from Portner et al. (2018) uses P-wave tomographic techniques to construct a regional velocity model that spans the Anatolian microplate. The authors subdivide the Anatolian region into three temporal stages of tectonic evolution, the Bitlis-Zagros, Cyprus, and Aegean domains. The complexity of the subducting slab increases with eastward progression, as does the temporal tectonic evolution of the region (Portner et al., 2018). Regional interpretations for the slab geometries from Portner et al. (2018) are shown in Figure 1.3, and are in general agreement with the geodynamical models proposed by Faccenna et al. (2006) and Schildgen et al. (2014). The Aegean slab to the west is simplistic and interpreted to be completely intact. The Cyprus slab has a more complex structure, and is intact to the west, but broken to the east. The Bitlis-Zagros slab is severely deformed. The increased deformation of the Bitlis-Zagros slab is due to the ongoing continental collision of the Arabian and Eurasian plates, the increased complexity of the Cyprus slab is related to the incipient continent-continent collision between the African and Anatolian plates, and the minimal complexity of the Aegean slab is related to the present-day subduction of oceanic crust beneath the Hellenic Arc (Portner et al., 2018).

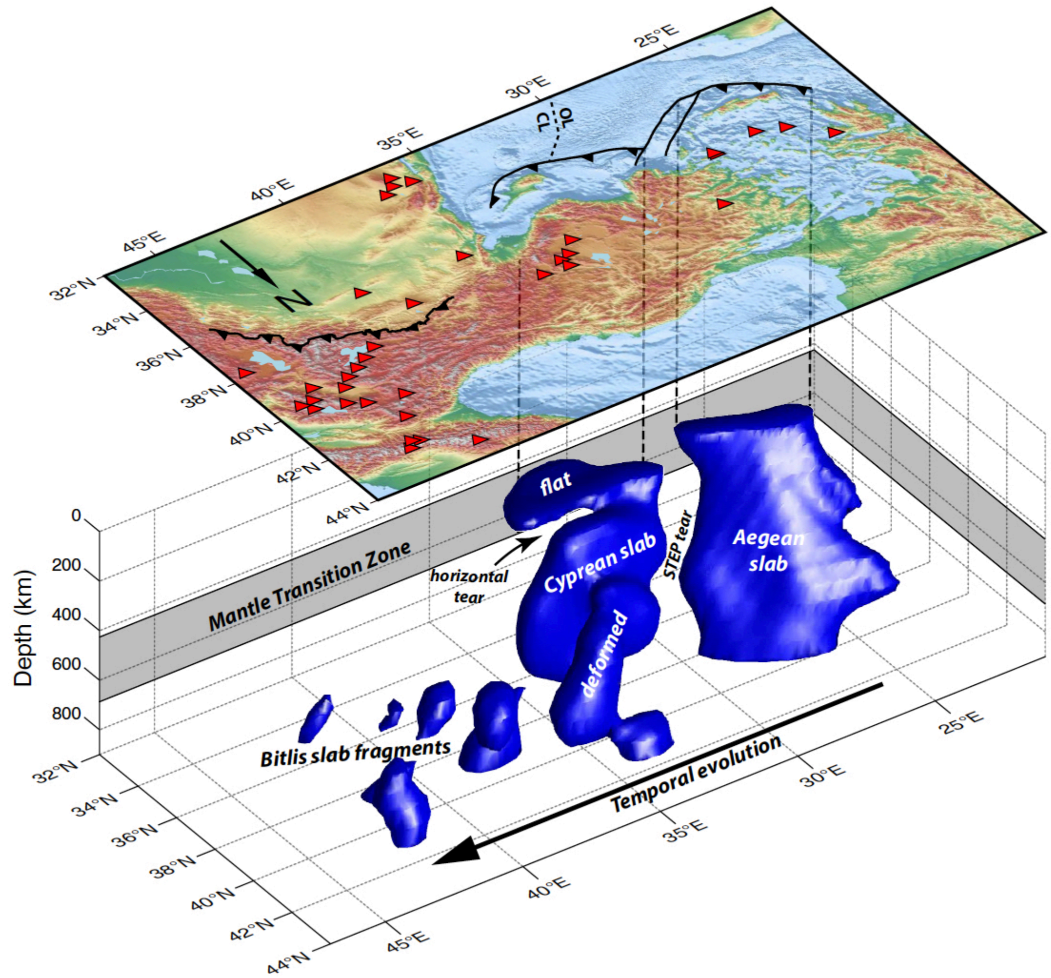

Figure 1.3: The subducting slab geometries in the Eastern Mediterranean after Portner et al. (2018). Red triangles in this image represent Holocene-aged volcanic activity and the blue surface represents an isovelocity contour. OL, oceanic lithosphere; CL, continental lithosphere; STEP, subduction-transform edge propagator system.

1.2.3 Crustal Scale: A Review of Seismic Refraction Surveys

This section presents a review of previous crustal-scale WARR surveys in the Eastern Mediterranean Basin. These surveys provide constraints on the crustal architecture in the region, which is the scale of investigation for this FWI study.

Makris et al. (1983) conduct the first wide-angle seismic refraction survey in the Eastern Mediterranean Basin. They construct a model of the region between Cyprus and Israel using five velocity layers, where the lowermost 6.7 km/s layer represents 8-10 km of oceanic crust beneath the Levant Basin. This layer is underlain by an 8.0

km/s layer that represents the upper mantle. Makris et al. (1983) state that there is a major crustal boundary south of the ECB, and attribute this crustal boundary to the continental to oceanic crust transition.

Ben-Avraham et al. (2002) acquire two WARR surveys in the Levant Basin to build upon the results of Makris et al. (1983). These WARR seismic lines are oriented approximately east-west in the Levant Basin. Ben-Avraham et al. (2002) agree with Makris et al. (1983) that sediments in the Levant Basin lie on top of oceanic crust. The authors model a sedimentary package 10-14 km thick for the Levant Basin. This sedimentary package consists of Pliocene-Pleistocene sediments with a velocity of 2.5 km/s, an underlying pre-Miocene package with a velocity of 3.8 km/s, and a possibly Mesozoic carbonate package with velocities of 4.5 km/s. Ben-Avraham et al. (2002) model the Levant continental margin as thinning toward the Levant Basin with a continental velocity of 6.3 km/s. Oceanic crust composes the floor of the Levant Basin with a velocity of 6.7 km/s. Ben-Avraham et al. (2002) further interpret that the oceanic crust under the Levant Basin is a remnant of a large ocean basin that is consumed during northern convergent tectonics. Furthermore, magnetic and gravity modelling support the presence of oceanic crust with a magnetization of 1000-2500 mA/m and a density of 2800-2900 kg/m^3.

Following the work of Ben-Avraham et al. (2002), two seismic refraction surveys were acquired on land investigating the Dead Sea Transform (Weber et al., 2004; ten Brink et al., 2006). Both profiles are oriented orthogonal to the Levant margin. The results from Weber et al. (2004) and ten Brink et al. (2006) provide insight into the gradual change in crustal structure over hundreds of kilometres onshore. Both profiles show southeast to northwest crustal thinning over 250 km. The northern line experiences a thinner upper crust, from approximately 14 km thick in the east to approximately 8 km thick in the west. The lower crust thins from 15 km in the east to approximately 11 km in the west. The Moho depth subsequently decreases from 34 km in the east to 24 km in the west Weber et al. (2004). Lower crustal seismic velocities are similar to the velocities reported as oceanic crust by Ben-Avraham et al. (2002), ranging from 6.7 to 7.0 km/s. Similar thickness trends are observed onshore to the south by ten Brink et al. (2006) as the lower crust thins to a consistent 8 km to the northwest. The lower crust here has a uniform velocity of 6.7 km/s (ten Brink et al., 2006).

The onshore observations of crustal thicknesses motivated a re-evaluation of the Levant Basin crustal structure and origin from Netzeband et al. (2006). The authors interpret the 6.7 km/s lower crustal layer observed onshore by Weber et al. (2004) to extend beneath the Levant Basin. The location of the southern seismic refraction line from Netzeband et al. (2006) lines up with the DESERT 2000 line from Weber et al. (2004). These two refraction models are combined with a portion of the line from Makris et al. (1983) to generate a crustal mega transect shown as a single section in Figure 1.4. The line segments for this megatransect are shown in Figure 1.1. Along this profile, the continuity of velocities and thicknesses from the DESERT 2000 line appears to extend into the Levant Basin. Sediments have velocities from 1.9-5.0 km/s, and an upper crustal layer has velocities of 6.0-6.4 km/s. A lower crustal layer underlies the upper crustal layer with velocities from 6.5-6.8 km/s (Netzeband et al., 2006).

Feld et al. (2017) present a north-south oriented line that spans 650 km from the ECB to the south, over the island of Cyprus, and into Turkey to the north. Feld et al. (2017) interprets the Eratosthenes Seamount to be 28-37 km thick, with crustal P-wave velocities of 6.5 km/s in contrast to the 27 km thickness interpreted by Welford

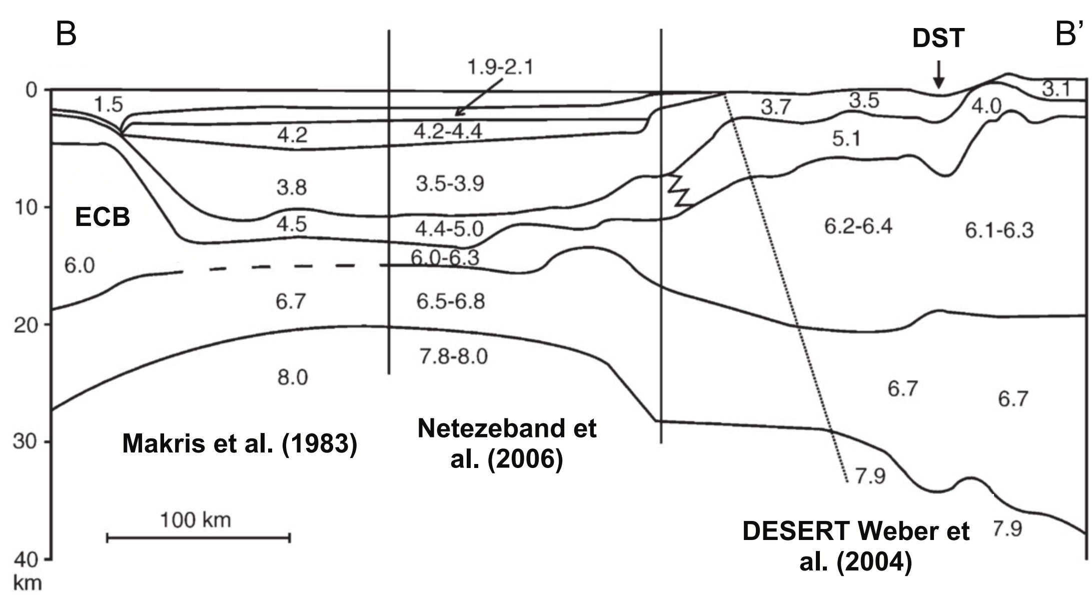

Figure 1.4: A mega transect spans from the Dead-Sea Transform (DST) in the southeast to the ECB in the northwest after Netzeband et al. (2006). Figure 1.1 shows the approximate location of the line segments for this mega transect as B-B'.

et al. (2015a). Oceanic crust is interpreted south of the ECB by Feld et al. (2017), and like Welford et al. (2015a), Feld et al. (2017) interpret a transform zone to the north. The land portion of the acquisition suggests a 12 km thick ophiolite sequence on Cyprus with reflected phases showing a dipping subducting slab beneath Cyprus. Data acquisition over Turkey reveals Moho-depths that vary between 38 and 45 km (Feld et al., 2017).

1.2.4 Basin Scale: A Review of Seismic Reflection Surveys

Reflection studies are generally basin focused and reveal the tectonic, structural, and stratigraphic histories of sedimentary basins. Interpretations toward the regional tectonic evolution are often obtained by correlating interpretations from basin to basin. Coincident reflection studies often compliment FWI crustal-scale velocity models to correlate geological structure revealed by the refraction line with structure observed on the reflection line (Kamei et al., 2012; Davy et al., 2017; Górszczyk et al., 2017).

To the southeast of the Hecataeus Rise in the northern portion of the Levant Basin, Vidal et al. (2000) present deep seismic reflection lines oriented orthogonal to the Latakia Ridge (Figure 1.1). They conclude that the Northern Levant Basin most likely formed atop thinned or transitional continental crust, which has been subsiding with relatively little deformation since the mid-Cretaceous. Within the basin, 6 km of Mesozoic sedimentary rocks and 4 km of Cenozoic sedimentary rocks overlie Paleozoic basement rocks. Vidal et al. (2000) observe an eastward transition from a sharp boundary of tectonic deformation at the Hecataeus Rise to the development of positive flower structures along the Cyprus Arc.

The seismic profiles interpreted by Hall et al. (2005b,a) indicate that compressional tectonics dominated during the Miocene, whereas strike-slip tectonics dominated during the Pliocene-Quaternary. The Hecataeus Rise transitions into the Latakia Ridge to the east, where the ridge is a positive flower structure that formed during the reactivation of a south-directed thrust Hall et al. (2005b). Similar south-directed thrusts are interpreted to the north-west by Calon et al. (2005a). North of the Latakia Ridge, the Larnaka Ridge is interpreted to be a southern-directed thrust sheet with deep roots linked to African-Anatolian convergence (Calon et al., 2005a). The northern Kyrenia FTB is structurally related to the Larnaka thrust as an imbrication of this main south-directed thrust, and the Latakia Basin between the two is a sizeable piggy-back basin (Calon et al., 2005a,b; Hall et al., 2005a). The western onshore continuation of the northern Kyrenia FTB and the southern Troodos-Larnaka culmination bound the Mesoria Basin on Cyprus as it is the structural equivalent of the Latakia Basin. A tectonic synthesis of the region reveals three main episodes, the first being the onset of thrusting activity generating the Troodos-Larnaka culmination in the Eocene (Calon et al., 2005a,b). The second event is a Miocene compressional event, and the third is a Pliocene transition to strike-slip tectonics (Calon et al., 2005a,b; Hall et al., 2005b,a).

Montadert et al. (2014) provides a review of a recent dense 2-D industry seismic reflection survey in a region directly south of Cyprus. The study area encompasses a vast region spanning the Herodotus Basin to the west, the Latakia Ridge to the east, the Cyprus Arc and Hecataeus Rise to the North, and the ECB to the south. Montadert et al. (2014) uses the separation of the ECB from the northern Gondwanan margin as the method of opening the Levant and Herodotus basins. Rifting begins in the Triassic to Mid-Jurassic, with an additional phase of rifting occurring from the Upper Jurassic to Lower Cretaceous. The Cyprus Arc is a compressional feature formed in the Upper Cretaceous. Its generation is related to the widespread obduction of ophiolites in the Cretaceous, and the Cyprus Arc is active since the Upper Cretaceous to recent times. Subduction of Tethyan oceanic crust occurred beneath Cyprus throughout the Paleogene until the lower Miocene, where the ECB underwent compression, and the Latakia Ridge begins to exhibit strike-slip faulting (Montadert et al., 2014).

Klimke & Ehrhardt (2014) interpret seismic reflection profiles from the ECB into the Western Levant Basin. In the Western Levant Basin, they map approximately 10 km of Mesozoic to Quaternary sedimentary rocks and interpret a lack of deformation in a supposedly compressive regime. This interpretation is consistent with that further to the east in the Levant Basin from Vidal et al. (2000). These observations lead Klimke & Ehrhardt (2014) to interpret that the Hecataeus Rise, Eratosthenes Seamount, and Western Levant Basin all encountered the Cyprus Arc as a single tectonic unit. The Hecataeus Rise experienced the most deformation associated with northward convergence, leaving the Western Levant Basin and Eratosthenes Seamount relatively undeformed. This interpretation leads to a unique plate boundary where the contact point between the African plate and Eurasian microplate is drawn along the Cyprus Arc and Latakia Ridge, and north of the Hecataeus Rise.

Closely related to the work from Klimke & Ehrhardt (2014), Reiche & Hübscher (2015) interpret pre-Messinian deformation to have occurred along the western and southern portions of the Hecataeus Rise. This deformation consists of the reactivation of NE-SW trending anticlines during post-Messinian convergence. A Pliocene-Quaternary unconformity is present throughout the structural high, and this unconformity is related to the collision between the Eratosthenes Seamount and the Cyprus Arc Reiche & Hübscher (2015). The unconformity temporally correlates with the Pliocene-Pleistocene tectonic transition throughout the Mediterranean, which may be related to the ECB entering the Cyprus Arc (Schattner, 2010; Schildgen et al., 2014; Reiche & Hübscher, 2015). Vertical uplift due to the ECB entering the Cyprus Arc affected the island of Cyprus more than the Hecataeus Rise, and the continuity of onshore structural lineaments offshore suggests that the two are related(Reiche & Hübscher, 2015).

An additional seismic study investigates the region along the Cyprus Arc between the ECB and the Hecataeus Rise (Reiche et al., 2016). The location and orientation of some of these seismic reflection lines align with the WARR line we investigate in this study. Along the Cyprus Arc, mobile anhydrite units thicken in the direction of African plate convergence. These units are allochthonous to the west and autochthonous to the east. The western allochthonous anhydrites of Messinian age overlie sediments of younger stratigraphic age. The lateral change in behaviour of the anhydrite wedge from west to east suggests an increased amount of shortening to the west. A seismic line from Reiche et al. (2016) is shown in the same orientation as the seismic refraction line considered for this study in Figure 1.5. In Figure 1.5, Reiche et al. (2016) correlates the interpreted seismic units to onshore Cyprus geology after Robertson et al. (1991). In addition, Reiche et al. (2016) summarize site 967 and 968 from ODP Leg 180 (Robertson, 1998; Robertson et al., 1998).

A lateral variability in crustal composition along the Cyprus Arc influences the tectonic style along the African-Eurasian plate boundary (Symeou et al., 2018). Oceanic crust floors the Herodotus Basin (Granot, 2016), the ECB is composed of thick continental crust, and transitional crust underlies the Levant Basin. Convergence between the African and Eurasian plate occurs along the Cyprus Arc, which is defined by Symeou et al. (2018) to span inwards toward the Misis-Kyrenia Range, and outward to the Latakia Ridge. The southern Latakia Ridge remains inactive until the

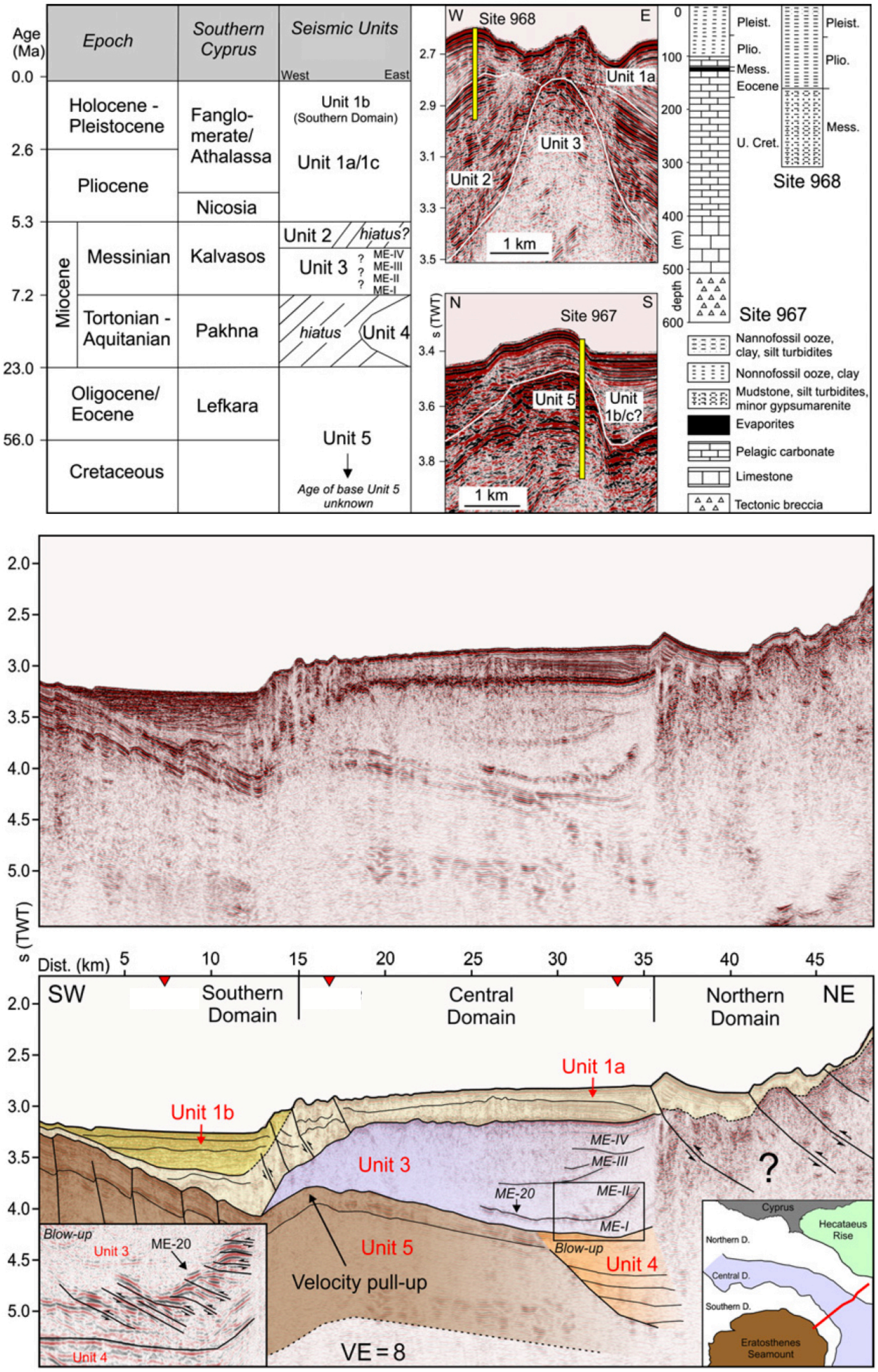

Figure 1.5: A seismic reflection line from Reiche et al. (2016), along with a stratigraphic column correlating the seismic horizons to geological boundaries. Two ODP Leg 160 wells and their associated lithofacies aid the correlation. Figure 1.1 provides the location of these wells south of Cyprus.

Miocene. The main proposition from Symeou et al. (2018) is that a two-stage deformation model exists, where compressional deformation is dominant in the Miocene, and a transpressional tectonic regime is dominant from the Pliocene to recent times. The increased dominance of transpressional tectonics toward the east since the Pliocene is interpreted to be related to a change in the crustal regime.

1.2.5 Summary

First, it will be challenging to constrain the nature and timing of rifting between the ECB and the Levant margin due to the uncertainty of which structures, if any, will be imaged by FWI. In addition, structures imaged in the ECB may be overprinted by Miocene to recent deformation. Finally, the NE-SW orientation of the seismic refraction line available for this study is oblique to the NW-SE perceived direction of extension (Gardosh et al., 2010) and the direction of eastward subduction beneath the ECB Segev et al. (2018).

The NNE-SSW orientation of Eastern Mediterranean Line 1 (Welford et al., 2015a) is well equipped to test various models for Cretaceous to recent convergence between the Eurasian and African plates. There seems to be some confusion in the literature as to where the subduction of Tethyan oceanic crust took place (Subsection 1.2.1). This study may be able to suggest if it is likely that a long-lived subduction zone existed along the Cyprus Arc. To answer this question, the three major tectonic elements for consideration are the ECB, the Cyprus Arc, and the Hecataeus Rise.

1.3 A Review of Full Waveform Inversion as Applied to Field Data

The overarching goal of this study is to get FWI to work on a sparse, wide-angle reflection and refraction (WARR) dataset. In this literature review section, pioneering works pertinent to the development of the FWI methodology are presented, with emphasis on algorithmic considerations for performing FWI. Following this introduction, we review select case studies in detail to understand what steps result in successful applied FWI studies.

1.3.1 Introduction

Lailly (1983) and Tarantola (1984a) suggest an inversion of seismic data to image the subsurface as opposed to seismic migration. Tarantola (1984b) works under the acoustic approximation and shows similarities between seismic inversion and seismic migration. Tarantola (1984a) again works under the acoustic approximation, but develops an iterative algorithm for minimizing the discrepancy between observed and predicted seismic data. Modelling the predicted data under the acoustic approximation requires subsurface models of density and bulk modulus. To make the predicted seismic data better resemble the observed seismic data, the model update is derived by correlating the forward-propagated wavefield with the back-propagated data residuals. This inverse seismic problem posed by Tarantola (1984a) is setup as an optimization problem where the misfit between the modelled and field data is minimized using the least-squares method with gradient descent techniques. Tarantola (1984a) further draws a comparison between his inversion algorithm and migration based on the imaging principle (Claerbout, 1971). The most direct comparison is between FWI and reverse-time migration (RTM). Both FWI and RTM require the forward propagation of the source wavefield. RTM combines this field with the back-propagated data, whereas FWI combines the wavefield with the back-propagated data residuals (Virieux & Operto, 2009).

FWI remained an inefficient imaging technique for seismic reflection data despite its promising initial formulation. FWI is a highly non-linear inverse problem (Tarantola, 1984a), and early disappointments with the approach correlate with an absence of methodology to mitigate non-linearity and convergence to a local minimum (Virieux & Operto, 2009). Researchers now overcome non-linearity in FWI by generating highly-accurate starting models and employing multiscale inversion strategies (Brenders & Pratt, 2007; Górszczyk et al., 2017). The seismic inverse problem is also computationally expensive (Tarantola, 1984a). The number of forward wave propagation problems to be solved for a single iteration of the optimization problem is (at least) two times the number of shots due to the required computation of the source wavefield and back-propagated data residuals. A final problem is related to the typical acquisition geometry of seismic reflection data in the late 20th century. These surveys recorded short offsets. Therefore FWI was limited by the absence of long and intermediate wavelengths. Near offset reflection datasets typically do not contain long wavelengths, which results in an incomplete reconstruction of the FWI model parameters (Virieux & Operto, 2009). While the technology to improve computational power and seismic acquisition evolved naturally over time, a vast amount of research is required to mitigate non-linearity in FWI.

Mora (1987) describes an additional problem with FWI. The acoustic approximation to the wave equation inadequately models the propagation of a real elastic wavefield through the subsurface. For example, energy partitioning for P- to S-wave conversions is unaccounted for by the acoustic approximation. Mora (1987) then argues that the elastic formulation of the wave equation should be used for FWI as it accounts for more of the waveforms observed in real data, and elastic FWI allows for the recovery of P-wave, S-wave, and density subsurface parameters. Mora (1988) completes 2D synthetic FWI using the elastic formulation of the wave equation. P- and S-wave velocity models are recovered from this synthetic inversion along with a density model. Reflected phases recover the short-wavelength subsurface structure, while refracted waveforms recover long-wavelength subsurface structure (Mora, 1988). However, the elastic formulation of the wave equation increases the computational cost of forwarding modelling in comparison to the acoustic formulation of the wave equation. Computational costs and a low wavelength starting model are issues for elastic FWI (Mora, 1987), although the requirement of a low-wavelength starting model is not unique to elastic FWI (Virieux & Operto, 2009).

The work of Tarantola (1984a,b) and Mora (1987, 1988) is formulated in the time domain. Frequency domain methods are introduced by Pratt & Worthington (1990), and further elaborated upon by Pratt et al. (1996, 1998), and Pratt (1999). Pratt & Worthington (1990) apply forward modelling directly to the frequency domain wave equation under the acoustic approximation using finite differences. The authors apply this method to cross-hole seismic data, and Pratt et al. (1996) investigate this method at the crustal-scale with synthetic wide-angle data. The frequency-domain formulation of FWI is advantageous as it allows for a natural progression from low to high frequencies (Pratt et al., 1996) and allows for straightforward incorporation of inelastic attenuation to the forward modelling procedure (Toksöz & Johnston, 1981). The progression from low to high frequencies mitigates the tendency of FWI to converge to a local minimum (Pratt & Worthington, 1990; Pratt et al., 1996). Pratt et al. (1998) formalize the matrix-based frequency domain inversion approach. The FWI problem formulated in the frequency domain requires three forward modelling steps, one to compute the data residuals, another to backpropagate these data residuals, and a third to compute the step length of the local-descent optimization technique. While an additional forward model is required, the computational expense of solving the 2D forward problem in the frequency domain is much cheaper than in the time domain. Finally, Pratt (1999) suggests that FWI success requires a robust estimation of the source signature, careful processing of modelled and predicted data amplitudes, and a selection of frequencies that proceed from low to high values. A method for source wavelet estimation in the frequency domain is provided by Pratt (1999) that only requires one iteration and one forward model per source and frequency. Modelling inelastic attenuation allows for better replication of the observed amplitude-versus-offset (AVO) signature in the modelled dataset. Frequency domain methods allow for a natural progression from low to high frequencies mitigating cycle-skipping or convergence to local minima. Cycle-skipping occurs when the initial model results in arrival times that are more than half a wavelength from the observed traveltimes. By starting FWI at low frequencies, we have longer wavelengths and consequently, a lesser risk of cycle-skipping. Pratt (1999) suggests using traveltime tomography to generate the initial velocity model used at the lowest frequencies.

Following the development of frequency-domain techniques, there have been many convincing FWI applications in the frequency domain (Ravaut et al., 2004; Operto et al., 2006; Brenders & Pratt, 2007). However, it is not straightforward to implement temporal constraints in the frequency domain where data-windowing is available for time domain applications (Shipp & Singh, 2002). The development of so-called Laplace-Fourier domain approaches allowed for frequency-domain methods to more easily implement multiscale strategies that resemble time-domain layer stripping approaches (Shin et al., 2002; Brenders & Pratt, 2007; Brossier et al., 2009; Shin & Cha, 2009). Laplace-Fourier domain waveform inversion (Shin & Cha, 2009) differs from Laplace domain waveform inversion (Shin & Ha, 2008) through the inclusion of complex-valued frequencies. This method is further developed by Brossier et al. (2009) to incorporate the first arrival picks to weigh the application of damping in the frequency domain such that it resembles time-domain damping from the first break picks, rather than from zero time. Laplace-Fourier domain waveform inversion strategies allow for frequency-domain methods to develop multiscale inversion strategies competitive with those used in the time domain (Górszczyk et al., 2017).

Virieux & Operto (2009) states that the best way to mitigate non-linearity in

FWI is to design a multiscale FWI approach which aims at progressively inverting shorter wavelengths. Despite this, convergence to the global minima is uncertain due to the starting model quality, an absence of low frequencies, noise, and limitations posed by wavefield modelling approximations (Virieux & Operto, 2009). In addition to the limitations of the acoustic approximation (Mora, 1987, 1988), the effects of anisotropy and attenuation are relevant. Prieux et al. (2011) presents a sensitivity study between anisotropic and isotropic FWI, revealing that FWI fits the data in both cases but produces different velocity models. Kurzmann et al. (2013) presents a sensitivity analysis for viscoacoustic FWI using attenuation models with varying degrees of accuracy. In their study, attenuation is modelled as a passive parameter and is not continuously updated during FWI. They conclude that not considering seismic attenuation in FWI will lead to the recovery of a poor velocity model. However, Kurzmann et al. (2013) note an improvement in their results even when FWI uses a homogeneous, but reasonable attenuation model.

As it is often difficult to generate an attenuation model (Winkler et al., 1979), seismic attenuation may be inverted for as an active parameter under the viscoacoustic approximation of the wave equation (Malinowski et al., 2011). Anisotropy may be inverted as an active parameter as well, along with velocity and density under the acoustic approximation (Plessix et al., 2013). FWI applications recovering more than a single parameter are called multiparameter inversions. The simultaneous recovery of P- and S-wave velocities is potentially possible under the elastic formulation of the wave equation (Mora, 1987; Sears et al., 2010). If done successfully, recovered P- and S-wave velocity models allow for a highly quantitative description of the subsurface. However, Operto et al. (2013) warn of cross-talk between recovered parameters in a multiparameter inversion scheme. This phenomenon is effectively related to some parameters having a similar contribution to the seismic response, therefore updating one parameter would unintentionally adjust the other.

To summarize, three domains for waveform inversion are the time-domain, frequency-domain, and Laplace-Fourier domain. Within one of these domains, a synthetic acoustic or elastic wavefield is modelled through isotropic or anisotropic media. Complex velocities may incorporate inelastic attenuation in the frequency-domain, and the method of finite differences is a popular method to solve the wave equation numerically in any of these domains. Furthermore, a single or multiple parameters may be recovered by FWI, however multiparameter inversions often generate cross-talk. In the following subsections, select case studies are explored using various combinations of the listed FWI algorithmic considerations above. With this review we aim to develop an understanding of which methodologies result in successful applications of FWI to real seismic data.

1.3.2 Frequency Domain Acoustic FWI Examples

An application of FWI to a real dataset is presented by Ravaut et al. (2004). The land-based data acquisition located in a geologically complex region consisted of a 14.2 km long line sampled by 160 single component geophones spaced every 90 m. The authors implement viscoacoustic waveform modelling with the method of finite differences in the frequency domain. They then use a gradient descent technique to minimize the misfit between the observed and predicted data. The diagonal approximation of the Hessian matrix during optimization provides second-order information to the first order gradient-descent technique. This preconditioning matrix is damped, which has a prewhitening effect on the optimization. A 2-D Gaussian smoothing operation regularizes the inversion where the correlation lengths of the Gaussian are a function of wavelength. The inversion perturbs an initial tomographic velocity model by proceeding from low to high frequencies through 16 frequency steps beginning at 5.4 Hz and finishing at 20 Hz. Well-data in the region validate the recovered P-wave velocity model (Ravaut et al., 2004).

The study from Operto et al. (2006) is applicable to this study as it is the first application of FWI to a real WARR dataset acquired using OBS. A benefit of applying full-waveform techniques to datasets with large aperture angles is the increased redundancy of low wavenumbers. A dense OBS spacing is required to mitigate spatial aliasing, and the dataset used was acquired across the Eastern Nankai Trough in central Japan using 93 OBS spaced at approximately 1 km. Like Ravaut et al. (2004), they use a viscoacoustic formulation of the wave equation and model the wavefield using finite differences (Hustedt et al., 2004). In addition, Operto et al. (2006) applies preconditioning to their gradient-descent optimization algorithm. They damp near offset high amplitude arrivals and incorporate regularization as Gaussian smoothing. Operto et al. (2006) exploit a reciprocity relationship to allow for the modelling of vertical component geophones as pressure receivers, and the pressure sources as vertical component receivers. Doing so increases the computational speed as the number of OBS is much less than the number of shots. In processing the observed data, the authors mute the time domain signature at the water bottom multiple as first-order gradient descent techniques often misinterpret water bottom multiples as scattering events.They acknowledge that this time-domain muting is not reproducible during the frequency-domain modelling of the data. They invert a total of 13 frequencies from 3-15 Hz, and the results are carefully checked through ray tracing, comparing modelled and predicted frequency-domain data, and finally, inverse Fourier transforming the data to the time domain for direct comparison of the wavefields.

Brenders & Pratt (2007) apply frequency-domain FWI to realistic, synthetic data generated by a third party. The data are generated with viscoelastic wavefield modelling from a model unknown to the authors until FWI is complete. Frequency-domain viscoacoustic FWI is used to reproduce the observed seismic data, in turn recovering a velocity model that highly resembles the truth. This synthetic study encompasses a 240 km long profile with long offsets (>100 km), and low starting frequencies enabling the recovery of the long-wavelength structure before the recovery of shorter wavelengths. They use a frequency continuation strategy to accomplish this by progressing from 0.8 Hz to 7.0 Hz frequencies in groups of three. Unlike Operto et al. (2006), an amplitude correction is applied to the data to account for the discrepancies between pressure wavefields and vertical component geophones. Brenders & Pratt (2007) also apply exponential time-domain damping to the data residuals in the frequency domain, which favours linear model updates. This approach is mathematically identical to Laplace-Fourier inversion (Shin & Cha, 2009). The degree of damping is progressively relaxed at higher frequencies. Brenders & Pratt (2007) highlight several aspects of the strategy that are crucial for the success of their study: 1) a highly accurate starting model, generated through traveltime tomography, which mitigates the risk of convergence to a local minimum in the objective function, known as cycle-skipping, 2) frequency-domain implementation, 3) using complex-valued frequencies to simulate time damping in the frequency domain, 4) a multiscale approach incorporating higher frequencies with progressive inversions, 5) matching of amplitudes between the observed and modelled datasets, and 6) an effective seismic data processing workflow for the observed dataset.

1.3.3 Elastic FWI Examples

Shipp & Singh (2002) attempt FWI using the elastic formulation of the wave equation. They model the wavefield using finite differences in the time domain. Throughout their study, Shipp & Singh (2002) make it evident that there are computational issues with both the elastic formulation and performing FWI in the time domain. Only a subset of the data are inverted, elastic parameters such as S-wave velocity and density are linked to the P-wave velocity update, and the elastic response is computed from velocity and stress wavefields. Their inversion scheme cannot take anisotropic or attenuation effects into account. Following two 1-D synthetic tests, they perform FWI on a marine streamer dataset acquired in the Faroe-Shetland Basin. A dense 20 m mesh discretizes the subsurface allowing for frequencies up to 15 Hz to be included in the inversion. They complete the inversion in five segments, and each inversion segment targets a different region in the subsurface. First, wide aperture offsets are windowed in the time-domain to recover the shallow velocity structure, <2.5 km. Their offset selection then moves from far, to near-far, and finally near offsets recovering different portions of the model. A final inversion phase includes all wide aperture data. Throughout their inversion scheme, input and modeled data are time windowed to recover select portions of the wavefield.

Brossier et al. (2009) present a 2-D elastic frequency-domain FWI problem for the SEG/EAGE overthrust model. The authors discretize the elastic wave equation using discontinuous Galerkin triangular meshing (Brossier et al., 2008). Brossier et al. (2009) uses the l-BFGS optimization technique to minimize the misfit between the observed and predicted data. The l-BFGS method uses the gradient to build approximations of the Hessian matrix (Byrd et al., 1995). The inversion strategy of Brossier et al. (2009) consists of an exterior loop that progresses through frequencies and an inner loop that decreases the damping applied through complex-valued frequencies. Damping is applied by using complex frequencies that are weighted from the first break picks. The inversion fails to converge to the solution without weighted damping. In regards to frequency structure, the authors test three techniques for frequencies between 1.7 and 7.2 Hz. The frequencies are first inverted sequentially, where the frequency groups only contain a single frequency (Ravaut et al., 2004; Operto et al., 2006; Brenders & Pratt, 2007). Second, a Bunks frequency strategy (Bunks et al.,

1995) progressively adds the next desired frequency to the group of inverted frequencies. Finally, a simultaneous frequency strategy inverts two separate frequency groups of three with overlapping frequencies. The simultaneous strategy provides superior results and Brossier et al. (2009) additionally improve on the results by increasing the number of frequencies in each frequency group from 3 to 5. Doing so increases data redundancy in FWI, but it comes with a five thirds increase in computational cost.

1.3.4 Recent FWI Studies Applied to WARR Data

OBS data in the western Nankai Trough are inverted by Kamei et al. (2012). They model the frequency-domain visco-acoustic wave equation numerically by finite differences, and use a logarithmic least-squares method to compute the misfit function. This technique for computing the misfit function better separates the phase and amplitude in the OBS dataset (Kamei et al., 2012, 2014). They recover a model 60 km in length and 15 km in depth using 54 OBS and 285 air-gun sources. The starting model for FWI used by Kamei et al. (2012) is a P-wave velocity model obtained using traveltime tomography with an average misfit of 60 ms. They update the starting velocity model by inverting frequencies between 2.25 and 8.5Hz. They then verify inversion output by observing decreases in the logarithmic least-squares misfit functional and visually comparing the observed and modelled datasets in the time-domain. More detailed discussions for waveform inversion strategies are provided by Kamei & Pratt (2013) and Kamei et al. (2013), and a discussion on misfit functionals for OBS data is provided by Kamei et al. (2014).

A time-domain acoustic application of FWI to an OBS dataset is given in Davy et al. (2017). Their approach in the time domain follows that of Warner et al. (2013), who developed a 3D finite-difference algorithm to perform FWI in the time domain. Davy et al. (2017) assume isotropic media despite that the algorithm of Warner et al. (2013) is able to model tilted transverse isotropy (TTI). The data consist of 19 OBS spanning a 63 km wide model. A top mute applied at 0.1 s above and a bottom mute applied at 1.8 s below the first break picks are used to window the time domain data. The maximum observable offsets in the OBS data range from 13 to 23 km, but complex near-offset waveforms require the first 5km of offset to be muted. Davy et al. (2017) validate their results by forward modeling the wavefield through the final FWI model, and through checkerboard tests.

A final crustal-scale OBS FWI study to discuss is that of Górszczyk et al. (2017) who revisits the Eastern Nankai Trough dataset previously investigated by Operto et al. (2006). They solve a frequency-domain visco-acoustic formulation of the wave equation using finite difference methods in this study, and the data are damped using complex-valued frequencies weighted from the first breaks. Like Brossier et al. (2009), Górszczyk et al. (2017) utilize the Laplace-Fourier FWI technique as an inner loop accompanied by an external progressive increase in frequencies to mitigate cycle skipping. Higher data-damping values at shorter offsets will artificially broaden the frequency spectrum allowing the inversion to start at 1.5 Hz. Górszczyk et al. (2017) adopt a different progressive strategy for increasing frequencies in the external multiscale loop. Their quasi-progressive frequency strategy retains a subset of the previously inverted-for low frequencies. This frequency strategy is implemented because long-wavelength structure recovered at low frequencies will degrade once higher frequencies are introduced in the inversion and the previously inverted-for frequencies are absent. A third, inner loop progressively increases the maximum offset in the inversion and is supplemented by a progressive decrease in the damping term for gradient preconditioning (Shin et al., 2001), which will progressively favor deeper velocity updates (Ravaut et al., 2004; Górszczyk et al., 2017). Górszczyk et al. (2017) provide an assessment of their results with a source estimation, synthetic seismogram modelling, ray tracing, dynamic image warping, and checkerboard tests.

Table 1.1: Summary of different FWI algorithmic considerations.

FWI Domain	**Pros**	**Cons**
Time-Domain	Able to window desired arrivals, and a lower degree of computational complexity	Increasing frequencies requires progressive filters, computationally expensive
Frequency-Domain	Natural environment for multi-scale approach to FWI, faster computational speed	Strong non-linear effects at low frequencies reflecting the inability to window events
Laplace-Fourier Domain	Data damping applied in the frequency domain, able to attenuate complex seismic waveforms	Laplace constants need to be defined, depth penetration dependent on offset
Physical Models	**Pros**	**Cons**
Acoustic approximation	Simplistic formulation	Does not account for attenuation, unable to model complex waves such as converted waves
Viscoacoustic	Simplistic formulation, accounts for attenuation	Unable to model complex waves
Elastic	Able to model almost all seismic arrivals	Difficult to implement, high computational cost, parameter trade-offs, and not able to model attenuation
Viscoelastic	Able to model almost all seismic arrivals as well as attenuation	Most difficult to implement, high computational cost, and parameter trade-offs
Anisotropy	**Pros**	**Cons**
Isotropic	Easy to implement	Inaccurate estimation of subsurface velocities with variable aperture sizes
Anisotropic	Accounts for a directional velocity dependence in rocks	Requires model(s) that describe the subsurface anisotropy
Inversion Parameters	**Pros**	**Cons**
Single Parameter	Reliable, stable inversion	Other parameters must be held constant
Multi-Parameter	Able to retain multiple geological subsurface parameters at once	Risk of cross-talk between inversion parameters, expensive computationally
Optimization Techniques	**Pros**	**Cons**
Gradient-descent	Robust and fast	Doesn't take into account the Hessian matrix without preconditioning
l-BFGS	A robust and fast quasi-Newton approach that approximates the Hessian matrix	Hessian matrix is not fully reconstructed
Truncated Newton and Gauss-Newton	A better approximation of the Hessian matrix	Computationally expensive

1.3.5 Considerations for Successful FWI

A highly non-linear response is likely for this FWI application due to the sparsity of the OBS. Reciprocity processes the OBS as seismic sources to reduce computation time, meaning that the Eastern Mediterranean dataset is sparse due to limited seismic sources. It is the sparse shot spacing with reciprocity that will limit resolution (Operto et al., 2006) and introduce non-linearity. Mitigating non-linearity and maximizing the probability of converging to the local minimum will require a sufficiently accurate starting model, an effective data processing strategy, and a multiscale strategy that progressively decreases the recovered wavelength and mitigates non-linearity are required. FWI is successful for many formulations of the inverse problem (i.e. frequency-domain or time-domain elastic, acoustic or elastic, etc.). The formulation of FWI that would best be suited for the available dataset and the goals of this study must also be considered.

Table 1.1 provides a summary of the pros and cons of different FWI formulations. It summarizes the FWI domain, wave equation formulation, directional velocity dependencies, inverted parameters, and optimization techniques. There are two noteworthy omissions from Table 1.1, which warrant further discussion. The first of which is the formulation of the objective function. The least-squares technique is popular, but Kamei et al. (2012) uses a log-formulation. Another relevant omission from Table 1.1 is the numerical technique for solving the forward problem. The method of finite differences is the most popular, but other methods are available such as discontinuous Galerkin (Brossier et al., 2008) and finite element discretization techniques (see Virieux & Operto (2009) for further discussion). We now review the various FWI formulations listed in Table 1.1.

1.4 FWI Methodology

In this section various FWI methods introduced in section 1.3 are discussed, beginning with a generalized FWI workflow. The goal of this section is to narrow down the preferred FWI algorithm for this study.

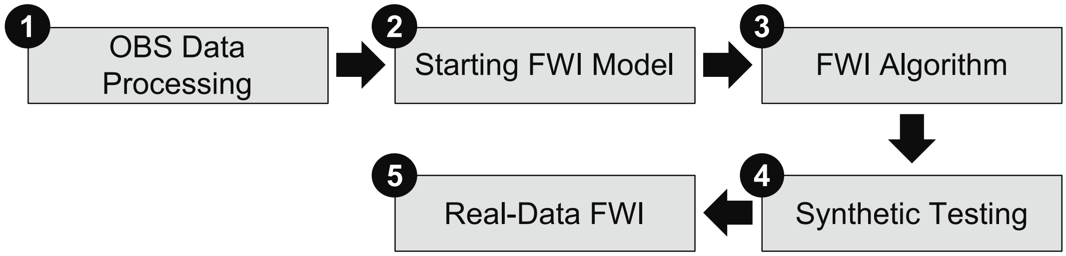

Figure 1.6: A general workflow for tackling the FWI problem.

1.4.1 Full-Waveform Inversion Workflow

Figure 1.6 shows the FWI workflow used in this study. This workflow consists of five general stages, where OBS data processing is the first stage, and building a sufficiently accurate starting model is the second stage. The third stage requires an algorithm to perform FWI, the fourth stage involves synthetic FWI tests, and the final stage performs FWI on the real Eastern Mediterranean dataset. The synthetic models used in the fourth stage are the well-known Marmousi2 model (Martin et al., 2006), and a full-scale synthetic Eastern Mediterranean model based on the final model from Welford et al. (2015a).

Full waveform inversion iteratively reduces the misfit between observed and forward modelled data. Figure 1.7 shows a simplified FWI algorithm. As evident in this figure, at least two forward solves of the wave equation are required for each source, which enable the computation of the gradient, and then the FWI model update. An arbitrary stopping criteria is then defined, which may consist of a target objective function value, a minimum required model update, or a maximum number of iterations. Once met, the final FWI model is output or the multiscale FWI strategy will proceed to the next inversion group. For the example shown in Figure 1.7, the multiscale strategy is only a frequency progression loop. The inversion will continue to iterate at the same frequency should this stopping criteria not be met.

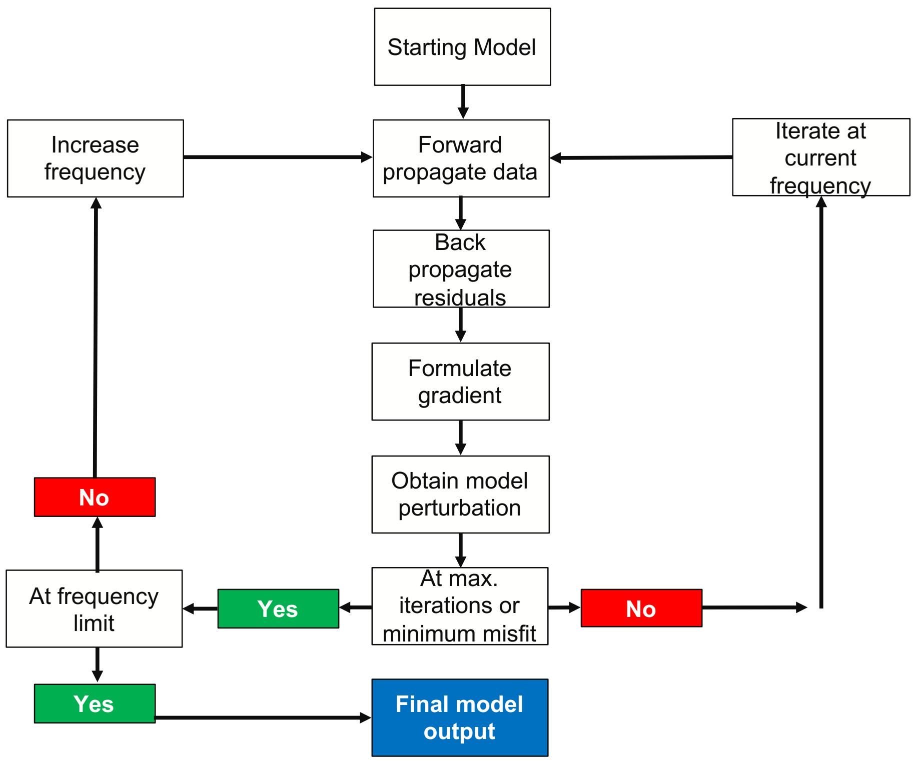

Figure 1.7: A simplified FWI algorithm.

1.4.2 OBS Data Processing

The seismic data acquisition will limit what type of FWI is possible. Offset angles must be long enough to propagate a full suite of seismic wavelengths through the subsurface. Diving waves generally have longer wavelengths than near-offset reflections, and are preferably recorded at long offsets (Operto et al., 2006). Because of the wide-angle crustal survey configuration, this will not be an issue. The instruments used to record seismic data may limit what type of FWI is possible. Hydrophones and multi-component geophones record seismic signals, where the hydrophone will record a pressure wavefield, and geophones record planar displacements. A vertical component geophone or hydrophone may be used for acoustic modeling, but three-component geophones are required for elastic modeling. Operto et al. (2006) shows that in the absence of quality hydrophone data, the modelled data may be processed (within the FWI algorithm) in a reciprocal manner. The receivers are treated as vertical sources and the shots are processed as pressure sensors. Given the dependency of computational time on the number of shots (Pratt et al., 1998), processing the OBS as sources will significantly reduce the computation time as the number of receivers is much less than the number of shots.

Operto et al. (2006) mention that it is difficult to make rules as to how dense a survey must be with respect to OBS spacing. However, they do suggest that the fold of the survey must be greater than one to avoid spatial aliasing. They formulate the inequality,

$$1 < \frac{O}{2\Delta X}, \tag{1.1}$$

where O is the maximum offset with observable reflections, and ΔX is the OBS spacing. The fold of a WARR survey is defined to be $\frac{O}{2\Delta X}$. Due to a variety of factors including noise and nature of the reflecting boundary, Operto et al. (2006) suggests that the theoretical fold is generally much smaller than the real fold. Nevertheless, Equation 1.1 dictates that for our dataset, reflections must be present beyond at least 24 km offsets in the OBS gathers with a 12 km OBS spacing. Wide-angle reflections are interpreted in the data at much further offsets by Welford et al. (2015a). Despite this, given the factors mentioned above and a variable maximum offset with discernible signal for each OBS, correctly imaging the subsurface reflections is still a concern for this study.

Following various practical considerations toward data quality, processing of the seismic data should, in general, mitigate the limitations of approximate wavefield modelling and increase the signal to noise ratio. The three-stage approach of (Górszczyk et al., 2017) provides a general idea of the required data processing steps for FWI. In the first stage, amplitude balancing is applied to correct for seafloor coupling differences at each OBS. This ensures that the dynamic amplitude range for all of the OBS are similar. Moreover, amplitude-versus-offset (AVO) effects are preserved for each OBS when this amplitude balancing correction is applied. Effectively the same processing step is completed by Brenders & Pratt (2007) and Kamei et al. (2012) who match the amplitudes of the predicted data with the modelled data while preserving AVO relationships. In addition, a final amplitude correction accounts for the amplitude discrepancy between a real 3D wavefield and a modelled 2D wavefield (Pratt, 1999).

In the second stage, Górszczyk et al. (2017) improves the signal-to-noise ratio by despiking high amplitude events. Analyzing the noise-RMS (NRMS) amplitude window above the first breaks allows for the identification of noisy traces. These traces with high NRMS amplitudes have their amplitudes progressively scaled-down. Górszczyk et al. (2017) claim to obtain superior FWI results by attenuating noisy traces rather than muting them as done by Ravaut et al. (2004) and Operto et al. (2006). Górszczyk et al. (2017) apply a coherency filter following trace-amplitude scaling, which is also applied by Ravaut et al. (2004) and Operto et al. (2006), albeit using different coherency filters. The final stage of pre-processing includes spiking deconvolution to whiten the amplitude spectrum. A white amplitude spectrum corresponds to a balanced cost function over frequency assuming these frequencies carry signal (Ravaut et al., 2004; Operto et al., 2006; Kamei et al., 2012; Górszczyk et al., 2017). The data are then band-pass filtered before a Fourier transform to the frequency domain.

In Appendix A, a detailed account of the OBS data processing flow used for this study is provided, and in Chapter 3, this processing flow is summarized. The OBS processing flow in this study is based upon that of Górszczyk et al. (2017), but differs in the details of the application of spectral whitening, filtering, and coherency.

1.4.3 Building the Initial FWI Model

First arrival tomography (FAT) is a popular method to build a starting model for waveform inversion (Ravaut et al., 2004; Operto et al., 2006; Brenders & Pratt, 2007; Kamei et al., 2012; Górszczyk et al., 2017). Davy et al. (2017) utilize tomographic imaging to build a starting model as well, but the authors additionally constrain the velocities in the sedimentary layers by forward modelling. Furthermore, waveform inversion in the Laplace domain is an interesting alternative method to build a starting model (Shin & Ha, 2008). FAT is ideal for wide-angle crustal surveys as the technique generates a velocity model that can reproduce first arrivals picked on OBS gathers within uncertainty. TOMO2D is a popular FAT algorithm used in this study to build a starting model with suitable accuracy for FWI (Korenaga et al., 2000).

TOMO2D (Korenaga et al., 2000) performs refraction and reflection tomography to build 2D velocity models. A hybrid combination of the graph method and the ray-bending method solves the forward problem. The graph method serves as an initial guess to the ray path, which is then refined by the ray-bending method. Traveltime residuals computed along a ray path generate velocity perturbations through the tomographic inverse problem. A covariance matrix incorporates first break pick uncertainties, and we use Gaussian smoothing to regularize the inverse problem. Two 1D correlation lengths in the x- and z-dimensions constrain the wavelength of velocity perturbations as a function of depth. The inverse problem is solved using LSQR, an optimization technique similar to the conjugate gradient method for large sparse matrices (Paige & Saunders, 1982).

Often a more sophisticated approach to FAT is required to get a starting model for FWI with an appropriate degree of accuracy. Górszczyk et al. (2017) perform three iterations of FAT to obtain an optimal starting model where the first break picks are adjusted in uncertain regions following each iteration. The authors adjust the first break picks by assessing travel time and phase residuals for the uncertain areas after each update. If one omitted these complex portions of the data, it would impede the illumination of complex structures in the data. During FWI, this can generate cycle-skipping of wide-angle diving waves (Górszczyk et al., 2017).

As alluded to throughout Section 1.3, a sufficiently accurate starting model is required for FWI. Formally, FWI will fail to converge to a local minimum if more than a 180-degree phase shift separates the observed and predicted data (Virieux &

Operto, 2009). Figure 1.8 shows a 1D example of cycle skipping, where a damped cosine function is plotted with two phase shifts to illustrate how an inadequately modelled first arrival time results in FWI converging to a local minimum. When using local optimization algorithms, the waveform that has been phase-shifted by more than 180 degrees will converge to the side lobe of a similar phase. This undesired outcome is convergence to a local minima. The green trace that is shifted by less than 180 degrees will converge to the peak of the observed waveform, therefore converging to a global minimum. There are two methods to allow for the red trace to converge to the black trace without shifting it. First, if the frequency of all traces are lowered, the phase shift between the red and black traces will decrease until it becomes less than 180 degrees. Second, global optimization algorithms search the entire model space. Therefore the red trace will still theoretically converge to the black trace. However, global optimization techniques are computationally infeasible when applied to FWI problems (Virieux & Operto, 2009). It is therefore paramount that the starting model for FWI is accurate enough to prevent cycle-skipping and convergence to a local

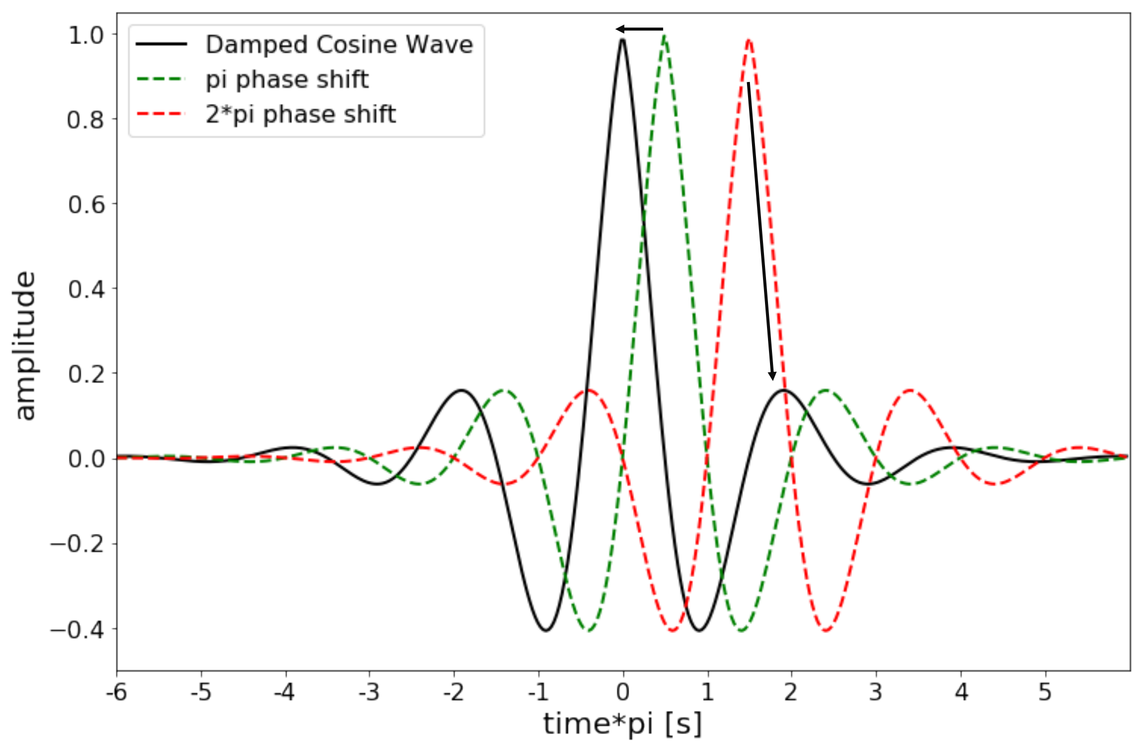

Figure 1.8: A 1D visual representation of cycle-skipping using a damped cosine function with a $\pi/2$ and $3\pi/2$ phase shift. Arrows show what the phase-shifted waveforms will converge to using local-descent techniques.

minimum. That said, convergence will be to a local minimum that hopefully is close enough to the global minimum to be useful.

Appendix B details the starting velocity model building procedure for this study, and Chapter 3 summarizes this procedure. The multistage model building procedure from Górszczyk et al. (2017) is used with TOMO2D (Korenaga et al., 2000) to build a starting model with a comparable RMS traveltime misfit to Kamei et al. (2012); Davy et al. (2017), and Górszczyk et al. (2017).

1.4.4 Full Waveform Inversion: The Forward Problem

The most significant consideration for the forward problem in FWI is what formulation of the wave equation to implement. Different formulations have different implications for FWI, as summarized in Table 1.1. This subsection will present the acoustic and elastic formulations of the wave equation in both the time and frequency domains. Both frequency-domain formulations may include inelastic attenuation, effectively formulating the viscoacoustic and viscoelastic wave equations.

The time-domain acoustic wave equation after Tarantola (1984a) is formulated as,

$$\left[\frac{1}{\kappa(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla \cdot \left(\frac{1}{\rho(\mathbf{r})}\nabla\right)\right] u(\boldsymbol{r}, t) = S(\boldsymbol{r}, t), \tag{1.2}$$

where $\kappa(\mathbf{r})$ is the bulk modulus and $\rho(\mathbf{r})$ is the density. The time-domain source term is $S(\boldsymbol{r}, t)$, and the wavefield is $u(\boldsymbol{r}, t)$. The value $\mathbf{r}$ represents a two-dimensional parameterization of [x,z] model-space. Given the following relationship, $\boldsymbol{v_p}^2 = \frac{\boldsymbol{\rho}}{\boldsymbol{\kappa(r)}}$, Equation 1.2 may be reformulated in terms of P-wave velocity, v_p, under a constant density assumption. The strength of this equation is its simplicity, but a downside is that this equation is invalid for wave-propagation through solids. To consider a more complete model for wavefield propagation in solids, one must consider the elastic formulation of the wave equation.

The time domain elastic formulation of the wave equation is presented after Mora (1987) who summarize from Aki & Richards (1980) as,

$$\begin{aligned} \rho \ddot{u}_i - \partial_j C_{ijkl} \partial_l u_k &= G_i \\ C_{ijkl} \partial_l u_k n_j &= T_i. \end{aligned} \tag{1.3}$$

In this compact notation the wavefield, $u_i = u_i(\mathbf{x}_s, \mathbf{x}, t)$, is the ith component of displacement from a source that is represented by displacement, G_i, and traction, T_i. This seismic source is located at $\mathbf{x}_s$ and is computed at position $\mathbf{x}$ for times t. The second derivative of the wavefield is denoted by $\ddot{u}_i$, and C_{ijkl} is a Hookean tensor that describes the elastic properties of the medium. Under the isotropic assumption, the Hookean tensor C_{ijkl} contains the two Lamé parameters, λ and μ. P- and S-wave velocities or impedances may be derived from the Lamé parameters and density to parameterize the subsurface. The strength of the elastic formulation lies in its ability to correctly handle multiple parameters in the forward model. Also, predicted amplitudes better model observed seismic amplitudes in this formulation.

To transition from the time to frequency-domain requires a Fourier transform of the wave equation. The frequency domain formulation of the acoustic wave equation is modified after Ravaut et al. (2004) as,

$$\left[\frac{-\omega^2}{\kappa(\mathbf{r})} + \nabla \cdot \left(\frac{1}{\rho(\mathbf{r})}\nabla\right)\right] u(\boldsymbol{r}, \omega) = S(\boldsymbol{r}, \omega). \tag{1.4}$$

Ravaut et al. (2004) writes the expression for seismic attenuation after Toksöz & Johnston (1981) as,

$$\tilde{c}(\mathbf{x}, \mathbf{z}) = \frac{c(\mathbf{x}, \mathbf{z})}{\left(1 + \frac{i}{2Q} sign(\omega)\right)}. \tag{1.5}$$

In Equation 1.5, the seismic attenuation is Q, i is an imaginary number, c is velocity, and $\tilde{c}$ is the complex velocity. The *sign* function extracts the sign of a real number. The equation indicates that smaller Q values attenuate the acoustic wavefield more than larger Q values. If the density model is variable, attenuation must be included in the formulation through a complex bulk-modulus as, $\tilde{k}(\mathbf{x}, \mathbf{z}) = \rho(\mathbf{x}, \mathbf{z})\tilde{c^2}(\mathbf{x}, \mathbf{z})$.

The frequency-domain formulation of the elastic wave equation is obtained by taking the Fourier transform of Equation 1.3. Using the compact notation from Aki & Richards (1980), the frequency-domain elastic wave equation is

$$\begin{aligned} -\rho\omega^2 u_i - \partial_j C_{ijkl} \partial_l u_k &= G_i, \\ C_{ijkl} \partial_l u_k n_j &= T_i. \end{aligned} \tag{1.6}$$

The Hookean tensor C_{ijkl} that describes the elastic properties of the medium using the two Lamé parameters, λ and μ. The Lamé parameters may be cast as complex values to model seismic attenuation. Acknowledging the link between complex velocity and attenuation in Equation 1.5, λ and μ may be related to complex velocities as,

$$\begin{aligned} \tilde{\lambda} = \rho(\tilde{v_p}^2 - 2\rho\tilde{v_s}^2), \\ \tilde{\mu} = \rho\tilde{v_s}^2. \end{aligned} \tag{1.7}$$

Complex P- and S-wave velocities are denoted as $\tilde{v_p}$ and $\tilde{v_s}$ above.

It is clear from the above equations that more computational resources are required to solve the elastic wave equation than the acoustic approximation.. The elastic wave equation requires solving two partial differential equations, whereas the acoustic wave equation only solves one. This computational cost is 2-3 orders of magnitude higher for elastic FWI than acoustic FWI (Operto et al., 2013). It also becomes evident why the frequency domain formulation of the wave equation is computationally cheaper than the time domain. The frequency-domain wave equation may solve for a single frequency, ω. The time-domain formulations must consider a range of discretized time values over which the wavefield propagates. Furthermore, frequency-domain formulations may easily incorporate seismic attenuation into forward modelling. It is based on these discussion points that we opt for the frequency-domain formulation of the acoustic wave equation for FWI.

While the multi-component OBS data are available to perform elastic FWI, this study is attempting to invert a sparsely sampled OBS dataset with instruments spaced at approximately 10 km intervals. The sparseness of this dataset may be an ideal opportunity to use a more computationally intensive method, such as the viscoelastic method. However, the most recent successful FWI applications to crustal-scale OBS surveys use the acoustic approximation (Davy et al., 2017), and the viscoacoustic approximation (Kamei et al., 2012; Górszczyk et al., 2017). A preference toward the more simplistic frequency domain viscoacoustic formulation over the viscoelastic formulation is based on its proof of success. The sparsity of the OBS dataset in this study can be exploited elsewhere with more computationally intensive approaches.

Two formulations of the method of finite differences are available to consider isotropic or anisotropic materials. Hustedt et al. (2004) provide a frequency-domain finite-difference solution for isotropic materials, and Operto et al. (2009) provide a finite difference solution for TTI materials. The modelling of materials under the

TTI approximation requires a more rigorous solution to the wave equation, and also requires a subsurface description of the three Thomsen parameters (Thomsen, 1986). A poor description of the subsurface Thomsen parameters may harm FWI, and recent crustal-scale studies do not consider anisotropic materials (Kamei et al., 2012; Davy et al., 2017; Górszczyk et al., 2017). Furthermore, Prieux et al. (2011) suggest that the changes in velocity between wide-aperture and short-aperture angles are mainly related to the vertical transverse isotropy (VTI). These anisotropic effects are due to wide-aperture seismic arrivals preferably travelling horizontally, and short-aperture seismic arrivals preferably traveling vertically. However, due to the configuration of any crustal-scale OBS survey, the vast majority of seismic arrivals are wide-aperture. Hence the isotropic approximation is sufficent. Shorter offset data are attenuated through a data-weighting scheme (Operto et al., 2006; Górszczyk et al., 2017) or a near-offset mute (Ravaut et al., 2004; Brenders & Pratt, 2007; Kamei et al., 2014; Davy et al., 2017).

Laplace-Fourier domain FWI closely resembles frequency-domain FWI. Therefore Equation 1.4 is still applicable with seismic attenuation as described in Equation 1.5. What the Laplace-Fourier approach does is damp the modelled data and apply a weight derived from the first break picks to the observed and modelled data (Brossier et al., 2009). From Shin & Cha (2009), the Fourier-Laplace transform of a damped signal $u(t)$ is,

$$TF(u(t)e^{\sigma t}) = \int_{-\infty}^{\infty} u(t)e^{-i\omega t}e^{-\sigma t}dt, \tag{1.8}$$

where σ is the Laplace constant, which governs the degree of damping, and TF is the Fourier transform. Brossier et al. (2009) adjust Equation 1.8 by weighting the application of damping from the first breaks. Equation 1.8 then becomes,

$$TF(u(t)e^{\sigma t_0}) = \int_{-\infty}^{\infty} u(t)e^{-i\omega t}e^{-\sigma(t-t_0)}dt, \tag{1.9}$$

where t_0 is the damping time. This time is preferably the first arrival, but it may be any arbitrary time.

Performing FWI in the Laplace-Fourier domain has two benefits. The first is that a “mirage-like reconstruction” allows for the recovery of ultra-low frequencies (Shin & Cha, 2009). Time-domain data are muted above carefully picked first arrivals to prevent the amplification of noise through this process. Secondly, the application of a weighted Laplace-Fourier domain inversion from the first break picks allows for the development of a powerful multiscale inversion strategy (Brossier et al., 2009; Górszczyk et al., 2017). This technique is as close to time-windowing as one may get in the frequency domain. Laplace-Fourier domain waveform inversion is incorporated into the methodology of this study as lower frequencies mitigate cycle skipping, and data-damping is a powerful multiscale inversion strategy (Górszczyk et al., 2017).

In this subsection, the forward problem is set up, but not presented in the context of an FWI algorithm. The following section will expand on our preferred forward modelling formulation and link forward modelling to the inverse problem.

1.4.5 Full Waveform Inversion: Implementation

This study utilizes the open-source TOY2DAC FWI code from the SEISCOPE Consortium (available at `https://seiscope2.osug.fr/TOY2DAC,82?lang=en`) to perform FWI. TOY2DAC utilizes the frequency domain viscoacoustic approximation of the wave equation shown in Equation 1.4. It is also able to model isotropic (Hustedt et al., 2004) or anisotropic media (Operto et al., 2009), and it can invert for multiple parameters. However, as explained in the previous subsection, we will use the isotropic formulation, and recover only the P-wave velocity. Laplace-Fourier domain FWI is implemented in TOY2DAC as well as described by Górszczyk et al. (2017). This FWI algorithm satisfies the desired formulation of the forward problem discussed in subsection 1.4.4. Using TOY2DAC aligns with Górszczyk et al. (2017), who used TOY2DAC to implement a three-loop multiscale inversion that produced a robust recovered P-wave velocity model.

Any inverse problem must minimize an objective function. TOY2DAC formulates the FWI objective function using the least-squares method, written as the sum of squared differences between the modelled and predicted data,

$$minimize \frac{1}{2} \sum_{s} |u_s(\mathbf{m}) - d_s|^2. \tag{1.10}$$

The modeled data as a function of the model parameters, $\mathbf{m}$, is $u_s(\mathbf{m})$, and d_s is the observed data. The summation is over all sources. The term *minimize* refers to a broad suite of techniques that search the objective function for a minimum. The global minimum corresponds to the most likely combination of model parameters that result in our model best fitting the data.

A predicted dataset is required to compute the cost function. TOY2DAC generates this predicted seismic dataset using the frequency-domain viscoacoustic formulation of the wave equation in Equation 1.4 with complex velocities as in Equation 1.5. This wave propagation problem is solved numerically using a fourth-order discretization scheme for the isotropic case (Hustedt et al., 2004). Following discretization, the 2D viscoacoustic wave equation written in matrix format is,

$$\mathbf{AU} = \mathbf{S}, \tag{1.11}$$

where $\mathbf{U}$ is the wavefield, $\mathbf{S}$ is the source, and $\mathbf{A}$ is a complex valued "impedance" matrix. Both $\mathbf{U}$ and $\mathbf{S}$ are column vectors assuming a single source (Pratt et al., 1998). The impedance matrix will take into account the subsurface model parameters,

$$\mathbf{A} = \boldsymbol{\kappa} - \omega^2 \boldsymbol{\rho} + i\omega \mathbf{Q}. \tag{1.12}$$

The stiffness matrix, $\boldsymbol{\kappa}$, is effectively the bulk modulus in Equation 1.4. The mass matrix, $\boldsymbol{\rho}$, corresponds to the density, and the damping matrix, $\mathbf{Q}$, is the subsurface attenuation Q in Equation 1.5.

Equation 1.11, is a sparse linear system. Lower-upper (LU) factorization of the complex impedance matrix recovers its inverse and solves for the forward modelled seismic wavefield. The inverted impedance matrix obtained through LU factorization additionally solves the adjoint problem and recovers the gradient, this leads to a model update (Plessix, 2006). For each source, the use of a single LU factorization to solve the forward and adjoint problem further reduces computational time in the frequency domain compared to the time domain (Métivier et al., 2014). This benefit assumes that the LU factorization of the impedance matrix can be stored in memory. TOY2DAC utilizes MUMPS (Amestoy et al., 2000) as a LU factorization algorithm.

To fully solve the wave propagation problem in Equation 1.11 after inverting the impedance matrix, an estimation of $\mathbf{S}$ is required. TOY2DAC utilizes the method of Pratt (1999) for the estimation of the source signature in the frequency domain.

Here, the source terms are multiplied by an unknown complex scalar, σ,

$$\mathbf{U} = \mathbf{A}^{-1}\sigma\mathbf{S}. \tag{1.13}$$

Each frequency and source combination requires a new value of sigma that estimates the source signature. This is formulated as,

$$\sigma = \frac{u_o^T(\mathbf{m})d^*}{u_o^T(\mathbf{m})u_o^*(\mathbf{m})}, \tag{1.14}$$

where * denotes the complex conjugate (Pratt et al., 1998; Ravaut et al., 2004). This source estimation technique requires a forward solution to obtain the initial wavefield, $u_o(\mathbf{m})$. In TOY2DAC, a single source estimation for each OBS or the average of these signatures may be defined to estimate the seismic source.

Now that the predicted wavefield is computed, data residuals are computed as shown in Equation 1.10. Górszczyk et al. (2017) formulates the least-squares model misfit in matrix notation as,

$$f(\mathbf{m}) = \Delta d^T W \Delta d. \tag{1.15}$$

The data misfit Δd is the difference between the predicted and observed data for each source, $\Delta d_s = u(\mathbf{m}) - d$. The cost function is $f(\mathbf{m})$, and W is a data-weighting function. The primary purpose of the data weighting function is to account for lower amplitude, far offset arrivals contributing less to the cost function. Typically a linear gain function accomplishes this (Operto et al., 2006; Górszczyk et al., 2017). Moreover, undesirable offsets, such as near offsets influenced by the direct arrivals, can be muted using this data-weighting function.

Local descent algorithms minimize the cost function, and are structured as,

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha_k \Delta\mathbf{m}_k, \tag{1.16}$$

where k is the iteration, α_k is the step-length, and $\Delta\mathbf{m}_k$ is the descent direction for the current model update. TOY2DAC utilizes the SEISCOPE Optimization Toolbox (Métivier & Brossier, 2016) to perform numerical optimization. The step length is consistently computed using a line-search that satisfies the Wolfe conditions. The

Wolfe conditions are,

$$\begin{aligned} f(\mathbf{m}_k + \alpha\Delta\mathbf{m}_k) &\leq f(\mathbf{m}_k + b_1\alpha\nabla f(\mathbf{m}_k)^T\nabla(\mathbf{m}_k), \\ \nabla f(\mathbf{m}_k + \alpha\Delta\mathbf{m}_k &\geq b_2\nabla f(\mathbf{m}_k)^T\Delta\mathbf{m}_k. \end{aligned} \tag{1.17}$$

Constants b_1 and b_2 above are set to 10^{-4} and 0.9 respectfully (Nocedal & Wright, 2006). The two Wolfe conditions ensure that the choice of α will generate a sufficient decrease in the cost function. The descent direction, $\Delta\mathbf{m}_k$, is required to compute the step length. A Newton model update formulates the descent direction as,

$$H(\mathbf{m}_k)\Delta\mathbf{m}_k = \nabla f(\mathbf{m}_k). \tag{1.18}$$

The gradient, $\nabla f(\mathbf{m}_k)$, may be efficiently computed using the adjoint state method (Plessix, 2006). The Hessian matrix is denoted by $H(\mathbf{m}_k)$. It is a matrix of second-order derivatives for the model parameters, and it is infeasible to compute given the size and sparsity of FWI (Virieux & Operto, 2009). Where the gradient defines the slope of the objective function, the Hessian matrix defines its curvature (Nocedal & Wright, 2006). Pratt et al. (1998) describes the Hessian matrix as predicting defocusing events in the subsurface. Therefore its inverse sharpens subsurface velocity perturbations and better recovers the multiply scattered wavefield.

While a significant consideration for FWI is how to model the seismic wavefield, an equally significant consideration is how to find the minimum of the objective function (Equation 1.15). The optimization technique used to find the local minimum will have implications on computational cost and efficiency for the inverse problem (see Table 1.1 for an overview).

1.4.6 Full Waveform Inversion: The Inverse Problem

The SEISCOPE Optimization Toolbox is a vast numerical optimization library for large-scale nonlinear optimization problems. Métivier & Brossier (2016) provides an introduction to the software and descriptions of the optimization routines available. Within the toolbox, ten numerical optimization routines are readily available for FWI in TOY2DAC. They are preconditioned and non-preconditioned versions of steepest descent, nonlinear conjugate gradient (NLCG), l-BFGS, truncated Gauss-Newton, and truncated Newton. Steepest descent and NLCG are first-order optimization routines, meaning they only utilize the information contained in the gradient to reach a local minimum. The l-BFGS algorithm is a quasi-Newton method, and truncated Newton and Gauss-Newton methods directly estimate the Hessian matrix.

Preconditioned algorithms speed up the convergence of the optimization routine (Métivier et al., 2014; Métivier & Brossier, 2016), and introduce valuable second-order information to first-order optimization routines (Ravaut et al., 2004; Operto et al., 2006). Preconditioned algorithms utilize the pseudo-Hessian matrix, $P(\mathbf{m}_k)$. This matrix is effectively the diagonal elements of the Gauss-Newton approximation of the Hessian matrix,

$$\tilde{P}(\mathbf{m}_k) = \left(\frac{1}{diag(J_k(\mathbf{m})^T J_k(\mathbf{m})}\right), \tag{1.19}$$

where $J_k(\mathbf{m})$ is the Jacobian matrix and $J_k(\mathbf{m})^T J_k(\mathbf{m})$ is the Gauss-Newton approximation to the Hessian. The Gauss-Newton approximation to the Hessian matrix, $B(\mathbf{m}_k)$, is,

$$B_k(\mathbf{m}) = J(\mathbf{m}_k)^T J(\mathbf{m}_k). \tag{1.20}$$

The pseudo-Hessian matrix is $diagB(\mathbf{m}_k)$. The preconditioner shown in Equation 1.19, $\tilde{P}(\mathbf{m}_k)$, is from Shin et al. (2001). An alternate formulation is provided by Métivier et al. (2014) to account for the rapid amplitude decrease in the wavefield with depth by adding a damping term, ϵ, to Equation 1.19,

$$P_\epsilon(\mathbf{m}_k) = \left(\frac{1}{diag(B_k(\mathbf{m})) + \epsilon D}\right). \tag{1.21}$$

In the formulation of (Métivier et al., 2014), $D = max(diag((B_k(\mathbf{m})))$. However, Górszczyk et al. (2017) defines D as the identify matrix, I. Including the damping term ϵ in a multiscale FWI strategy along with offset continuation benefits FWI by implementing a layer-stripping approach. Smaller values of ϵ weigh the gradient toward deeper velocity perturbations, and larger values favour shallower velocity updates (Ravaut et al., 2004; Górszczyk et al., 2017). Górszczyk et al. (2017) also include regularization by introducing Gaussian smoothing within the formulation of the preconditioner as wavelength-dependent smoothing (Ravaut et al., 2004).

The descent direction, $\Delta\mathbf{m}_k$ in Equation 1.16, is desired from each optimization routine. The preconditioned steepest descent algorithm estimates the descent direction as,

$$\Delta \mathbf{m}_k = P_\epsilon(\mathbf{m}_k)\nabla f(\mathbf{m}_k). \tag{1.22}$$

This algorithm solves for the descent direction, but the inclusion of the preconditioner ensures that the optimization routine accounts for second-order information (Nocedal & Wright, 2006; Métivier & Brossier, 2016).

The NLCG algorithm is similar to the steepest descent algorithm but takes into account the descent direction computed at the previous model update. The NLCG method is a popular, cheap alternative to quasi-Newton techniques for FWI problems. The preconditioned NLCG algorithm used by the Optimizaiton Toolbox is,

$$\begin{aligned} \Delta \mathbf{m}_k &= P_\epsilon(\mathbf{m}_k)\nabla f(\mathbf{m}_k), k = 0, \\ \Delta \mathbf{m}_k &= P_\epsilon(\mathbf{m}_k)\nabla f(\mathbf{m}_k) + R(\mathbf{m}_k)\Delta \mathbf{m}_{k-1}, k \geq 1. \end{aligned} \tag{1.23}$$

The first iteration of the NLCG algorithm is equivalent to the steepest descent algorithm (Equation 1.22). Subsequent iterations then account for the previous descent direction, and $R(\mathbf{m}_k)$, which is computed in the Toolbox after Dai & Yuan (1999) as,

$$R(\mathbf{m}_k) = \frac{||\nabla f(\mathbf{m}_k||^2}{(\nabla f(\mathbf{m}_k) - f(\mathbf{m}_{k-1})T\Delta \mathbf{m}_{k-1}}. \tag{1.24}$$

Using the formulation of Dai & Yuan (1999) convergence toward the local minimum is ensured as soon as the step-length satisfies the Wolfe conditions in Equation 1.17 (Métivier & Brossier, 2016). Should computational time be an issue in FWI problems, the preconditioned NLCG algorithm would be a satisfactory optimization algorithm. Otherwise, FWI should use algorithms that better account for the effect of the Hessian matrix.

The first quasi-Newton algorithm to be discussed is the l-BFGS method (Byrd et al., 1995). The l-BFGS algorithm extracts curvature information from the gradient while only retaining the information from the most recent iterations. The Hessian estimated through the l-BFGS method, $H_a(\mathbf{m}_k)$, is written after Nocedal & Wright (2006) as,

$$H_a(\mathbf{m}_{k+1}) = V(\mathbf{m}_k)^T H_a(\mathbf{m}_k) V(\mathbf{m}_k) + \gamma(\mathbf{m}_k) s(\mathbf{m}_k) \beta(\mathbf{m}_k)^T. \tag{1.25}$$

The terms $\gamma(\mathbf{m}_k)$ and $V(\mathbf{m}_k)$ are defined as,

$$\begin{aligned} \gamma(\mathbf{m}_k) &= \frac{1}{y_k^T \beta_k}, \\ V(\mathbf{m}_k) &= I - \gamma(\mathbf{m}_k) y(\mathbf{m}_k) \beta(\mathbf{m}_k)^T. \end{aligned} \tag{1.26}$$

Finally, the pairs $\beta(\mathbf{m}_k)$ and $y(\mathbf{m}_k)$ are,

$$\begin{aligned} \beta(\mathbf{m}_k) &= \mathbf{m}_{k+1} - m_k, \\ y(\mathbf{m}_k) &= \nabla f(\mathbf{m}_{k+1}) - \nabla f(\mathbf{m}_k). \end{aligned} \tag{1.27}$$

A more detailed derivation of the l-BFGS method is provided by (Byrd et al., 1995). The limited memory component of the BFGS algorithm (hence l-BFGS) governs how many past curvature approximations will build the Hessian. Nocedal & Wright (2006) and Métivier & Brossier (2016) provide a good example of how this iterative Hessian estimation works within an algorithm. The preconditioned l-BFGS algorithm adjusts the term $H_a(\mathbf{m}_k)$ in Equation 1.25. For the first l-BFGS estimation of the Hessian matrix the approximate Hessian matrix will be used, $H_a^0(\mathbf{m}_k) = P_\epsilon(\mathbf{m}_k)$. Preconditioning provides a more accurate initial approximation of the Hessian matrix, which the l-BFGS algorithm builds upon to approximate the Hessian matrix fully. As this computation is relatively cheap, the pseudo-Hessian may be repeatedly re-estimated for each iteration, k.

While the l-BFGS method can extract second-order information by curvature estimations from the gradient, the resultant Hessian matrix is an approximation and depends on the number of stored gradients. A large number of stored gradients will impose computational limitations, but may improve the curvature estimation. Nocedal & Wright (2006) suggests that the l-BFGS memory parameter or the number of gradients to be stored should range between 3 and 20. Noisy real-data inversions should have a lower memory parameter and noise-free synthetic inversions should have a larger memory parameter. While the l-BFGS algorithm is shown to be more efficient with preconditioning (Métivier et al., 2014; Métivier & Brossier, 2016), the reconstruction of the Hessian matrix is still imperfect.

Truncated Newton and Gauss-Newton methods are intriguing for FWI studies as they provide a better estimation of the Hessian matrix than l-BFGS (Métivier et al., 2013, 2014). The truncated Newton algorithm solves the Newton system of equations, i.e. Equation 1.18, by computing the Hessian vector products which are denoted as $H(\mathbf{m})\nu$ where ν is a vector in the model space. The truncated Gauss-Newton method is similar but solves a Gauss-Newton system of equations, i.e. Equation 1.20. The computation of vector products enables the Newton or Gauss-Newton linear system to be solved using a matrix-free version of the conjugate gradient algorithm. Preconditioning may be incorporated in this linear system, however it must be symmetric (Métivier & Brossier, 2016), therefore the preconditioned Newton system of equations becomes,

$$P_{\epsilon}(\mathbf{m}_k)H(\mathbf{m}_k)\Delta\mathbf{m}_k = P_{\epsilon}(\mathbf{m}_k)\nabla f(\mathbf{m}_k). \tag{1.28}$$

An approximate solution to the Newton or Gauss-Newton system of equations is obtained by limiting the number of conjugate gradient iterations and defining stopping criteria (Métivier et al., 2014).

The Hessian vector products are computed by Métivier et al. (2014) using the second-order adjoint state method under no assumption of linearity, whereas Pratt et al. (1998) and Métivier et al. (2013) assume linearity. The three term Hessian vector product formulae following Métivier et al. (2014) is,

$$\begin{aligned} H(\mathbf{m}_k)\nu = -&\mathbf{Re}\left[\left(\frac{\partial^2 F(\mathbf{m}_k, u_k)^T}{\partial \mathbf{m}_k^2}\chi_k\right)^T \mu_{1_k}\right] \\ +&\mathbf{Re}\left[\left(\frac{\partial^2 F(\mathbf{m}_k, u_k)^T}{\partial u \partial m}\chi_k\right)^T \mu_{2_k}\right] \\ +&\mathbf{Re}\left[\frac{\partial^2 F(\mathbf{m}_k, u_k)^T}{\partial m}\mu_{3_k}\right]. \end{aligned} \tag{1.29}$$

In the above equation, u is still the forward modeled wavefield, μ_{1_k} is the adjoint wavefield, μ_{2_k} is an additional wavefield computed through the forward problem, and μ_{3_k} is an additional adjoint wavefield. The term χ is an adjoint operator (see Plessix (2006), defined as λ there), and **Re** extracts the real part of its argument. Under the Gauss-Newton approximation Equation 1.29 becomes

$$B(\mathbf{m}_k)\nu = \mathbf{Re}\left[\frac{\partial^2 F(\mathbf{m}_k, u_k)^T}{\partial m}\mu_{3_k}\right]. \tag{1.30}$$

The Hessian vector products for the Gauss-Newton approximation in Equation 1.30 are simply the first order terms in Equation 1.29.

Both of the truncated methods require two forward problems and two adjoint problems to be solved, where the steepest descent, NLGC, and l-BFGS techniques require the solution of one forward and one adjoint problem. The Gauss-Newton approximation only requires the incident wavefield to be stored, whereas the truncated Newton method requires the forward and adjoint wavefields to be stored in memory (Métivier et al., 2014).

It is evident throughout this discussion that an accurate representation of the Hessian matrix in optimization problems aids in efficiently minimizing the cost function. However, a more accurate approximation of the Hessian matrix is a result of an increasingly computationally expensive inverse problem. Relevant WARR studies have been successful using gradient-descent techniques (Operto et al., 2006; Davy et al., 2017), l-BFGS (Kamei et al., 2012), and preconditioned l-BFGS (Górszczyk et al., 2017). Métivier et al. (2014) tests the truncated Newton technique by inverting near-offset seismic data.

A computationally expensive optimization technique that formulates approximations of the Hessian matrix such as l-BFGS, truncated Newton, or truncated Gauss-Newton should exploit the sparsity in the observed data. It is anticipated that the cost function for this FWI problem will be plagued by local minima due to nonlinearity. Therefore, a reasonable curvature estimate will aid in convergence to a minimum. We choose to use the preconditioned l-BFGS algorithm as it is proven robust (Górszczyk et al., 2017). In Appendix C.3, the preconditioned steepest descent and NLCG algorithms are tested on a full-scale synthetic Eastern Mediterranean model and compared to the results with a preconditioned l-BFGS algorithm. In practice, the truncated Newton and Gauss-Newton methods are unable to achieve acceptable convergence with the sparse Eastern Mediterranean dataset.

1.5 Summary

This study aims at re-investigating the Cyprus Arc in the Eastern Mediterranean Basin by applying high-resolution full-waveform methods to OBS data after Welford et al. (2015a). This chapter provides a general review of the FWI method with a focus on applied studies. The FWI algorithm for this study is a 2-D frequency-domain, viscoacoustic, isotropic algorithm able to perform FWI in the Laplace-Fourier domain.

We use TOY2DAC from the SEISCOPE Consortium to implement this formulation of FWI. For numerical optimization we use the preconditioned l-BFGS algorithm from the SEISCOPE Optimization toolbox (Métivier & Brossier, 2016). This FWI strategy is developed by considering the methodologies of relevant successful FWI real-data applications (Ravaut et al., 2004; Operto et al., 2006; Kamei et al., 2012; Davy et al., 2017; Górszczyk et al., 2017).

If FWI is successful with this sparse Eastern Mediterranean dataset, a geological interpretation is to follow. Perhaps the best geological issue to address, as alluded to in Subsection 1.2.5, is to better define the role of the southernmost Cyprus Arc in the convergent region between the ECB and Hecataeus Rise. Following the geological review in Section 1.2, it is unclear exactly where subduction took place in the Eastern Mediterranean to accommodate the northward motion of the African and Arabian plates since the Cretaceous.

Chapter 2

Co-authorship Statement

2.1 Manuscript

The manuscript is is written by Christopher Williams, and edited by Alison Malcolm and Kim Welford. Christopher Williams completed the work presented in this thesis under the guidance of Alison Malcolm and Kim Welford. Chapter 3 of this thesis is being prepared for submission to Geophysical Journal International.

2.2 Conference Proceedings

Christopher Williams presented the intermediate results from this study at the SEG Annual Meeting in San Antonio, Texas, in poster format (Williams et al., 2019a). Additionally, Christopher Williams presented the intermediate results from this study in oral format a the AGU Fall Meeting in San Francisco, California (Williams et al., 2019b). Attending both conferences proved valuable in gaining knowledge that contributed to the successful completion of this thesis.

Chapter 3

Wide-Angle, Full Waveform Inversion with a Sparse Ocean-Bottom Seismometer Dataset: A Case Study from the Eastern Mediterranean

3.1 Summary

Full waveform inversion (FWI) is an increasingly popular method to recover high-resolution velocity models of the subsurface. Recently, wide-angle reflection and refraction (WARR) datasets have been successfully inverted to produce crustal-scale P-wave velocity models. Traditionally, forward modelling or traveltime tomography are used to produce these crustal-scale models. However, examples thus far have looked only at datasets with unusually dense ocean-bottom seismometer (OBS) spacings. This study revisits a sparsely acquired WARR dataset in the geologically complex Eastern Mediterranean Basin. The dataset spans from the Eratosthenes Seamount to the south, across the Cyprus Arc, and onto the Hecataeus Rise. This dataset consists of 16 OBS spanning a profile of 190 km in length. We implement FWI in the frequency domain with a viscoacoustic approximation to the wave equation for synthetic and

real-data inversions. We additionally use an adapted multiscale inversion strategy to mitigate the impact of data sparsity on FWI. After using traveltime tomography to build a starting model, we use FWI to invert the OBS from 2.0 to 7.25 Hz using a true progressive frequency strategy. We assess the resultant FWI model by analyzing the observed and predicted frequency domain traces and comparing the predicted frequency domain traces to the starting traces. We interpret that the simplified attenuation model used results in a avoidable mismatch between observed and predicted data amplitudes at far offset, resulting in a poorly constrained deeper section of the recovered model. A geological interpretation with the aid of a coincident seismic reflection line reveals that the inversion recovers a high-velocity carbonate horizon over the Eratosthenes seamount that dips northward beneath a 2-3 km thick evaporite wedge where the horizon becomes absent. The recovered FWI model reveals that continental crust of the Eratosthenes Continental Block underthrusts an interpreted accretionary prism and the Hecataeus Rise. Based on the results of this study, we encourage the used of FWI on other typical WARR datasets, but perhaps in simpler geological environments.

3.2 Introduction

Wide-angle seismic reflection and refraction (WARR) surveys provide insight into local crustal structure and composition, enabling constraints on the regional tectonic evolution. Seismic data are typically acquired using variably spaced ocean-bottom seismometers (OBS) in marine settings. These seismic datasets are then used to produce crustal-scale velocity models by reproducing the traveltime information in the recorded data (for example see Wu et al. (2006); Klingelhoefer et al. (2009); Eakin et al. (2014) or Watremez et al. (2015)). Forward modelling techniques or tomographic inversion are popular methods to produce these crustal-scale velocity models Zelt & Smith (1992); Korenaga et al. (2000). However, there are issues with both techniques. Forward modelling is subjective and time consuming. First-arrival tomography (FAT) inverts for the first arrivals in an OBS gather and has a theoretical resolution limit of the first Fresnel zone (Williamson, 1991). Both techniques only consider travel time information in the seismogram, ignoring the seismic waveforms.

Full waveform inversion (FWI) methods account for the phase and amplitude of the seismic data. These techniques are then able to produce a higher resolution model than FAT with a half-wavelength theoretical resolution limit (Lambare et al., 2003; Virieux & Operto, 2009). FWI is formulated in the time domain under the acoustic approximation by Tarantola (1984a), where the cross-correlation of a forward propagated wavefield and the backpropagated data-residuals provides the model update. Therefore, the method is similar to reverse-time migration (Claerbout, 1971). Pratt & Worthington (1990) formulate FWI in the frequency domain and introduce a computationally efficent approach consisting of inverting increasing monochromatic frequencies (Pratt et al., 1996).

The objective functions for FWI problems contain many local minima as a consequence of the cyclic seismic waveform and the ill-posedness of the sparsely constrained inverse problem. Beginning FWI at ultra-low frequencies would remedy this issue, but controlled source methods in practice are unable to generate low enough frequencies that overcome background noise (Virieux & Operto, 2009). Typically a highly accurate starting velocity model is required to overcome this problem in FWI applications (Brenders & Pratt, 2007). Specifically, the modelled data must be within half a wavelength of the observed data to avoid convergence to a local minima (Sirgue & Pratt, 2004). Multiscale inversion strategies may also avoid convergence to a local minima (Górszczyk et al., 2017). Implementing a layer-stripping approach by spatial and temporal windowing in the time domain is practical as it progressively inverts more complex waveforms from deeper portions of the model (Shipp & Singh, 2002). A layer-stripping approach is inherently difficult to apply in the frequency domain, but Laplace-Fourier domain waveform inversion (Brossier et al., 2009; Shin & Cha, 2009), in combination with offset partitioning (Górszczyk et al., 2017), has a similar effect. Real-data FWI applications must additionally process the observed dataset to account for the approximate wave-physics in the predicted data, as well as amplitude discrepancies between the observed and predicted data in addition to unmodelled signal or noise (Ravaut et al., 2004; Operto et al., 2006; Brenders & Pratt, 2007).

With careful data processing, an accurate starting model, and an effective multiscale FWI strategy, there are recent successful applications of 2D FWI to WARR datasets (Operto et al., 2006; Kamei et al., 2012; Davy et al., 2017; Górszczyk et al., 2017). The first synthetic applications of FWI to a WARR dataset are completed by Pratt et al. (1996), and the first successful application of FWI to a crustal-scale WARR dataset is presented by Operto et al. (2006). More recent studies include an

OBS dataset located in the western Nankai Trough that is inverted using frequency domain viscoacoustic FWI (Kamei et al., 2012), and a dataset from the Galicia Margin offshore Spain that is inverted using time-domain acoustic FWI (Davy et al., 2017). Górszczyk et al. (2017) revisit the WARR dataset previously inverted by Operto et al. (2006) and apply a three-stage multiscale inversion strategy to recover a P-wave crustal velocity model. However, these WARR datasets are densely spaced with OBS placed at approximately one (Kamei et al., 2012; Górszczyk et al., 2017) and three (Davy et al., 2017) km intervals. Therefore, there has not yet been an FWI investigation on a WARR dataset with a typical OBS spacing.

This study aims to re-investigate a WARR dataset from the Eastern Mediterranean by applying full-waveform techniques to this relatively sparse OBS dataset. Welford et al. (2015a) produce a crustal-scale velocity model for this dataset using a combination of tomographic and forward modelling techniques. The resultant velocity model utilizes 21 OBS positioned along a 260 km long transect that spans from the Eratosthenes Seamount to the Hecataeus Rise across the Cyprus Arc (Figure 3.1). Welford et al. (2015a) and Welford et al. (2015b) interpret a continental crustal affinity for both the Eratosthenes Seamount and Hecataeus Rise. The Cyprus Arc is a transform zone (Feld et al., 2017), and two high-velocity lower crustal bodies are either igneous intrusions or remnant Tethyan oceanic crust (Welford et al., 2015a,b). This study uses the term Eratosthenes continental block (ECB) as the historical name (Eratosthenes Seamount) is misleading.

3.3 Background

3.3.1 Geological Review of the Eastern Mediterranean

The primary tectonic controls on the Eastern Mediterranean are the northwestern counterclockwise motion of the Anatolian microplate and the northern motions of the African and Arabian plates with respect to the Eurasian plate, as shown in Figure 3.1(Reilinger et al., 2006; Nocquet, 2012). The Levant Basin formed when the ECB was it was rifted off the Gondwanan margin. The timing and nature of this rifting mechanism are disputed. The onset of rifting in the Levant may be as early as the Late Paleozoic (Garfunkel, 1998; Gardosh et al., 2010) or as late as the Cretaceous

(Segev et al., 2018). The rifting mechanism is either related to the Tauride Block (Garfunkel, 1998; Gardosh et al., 2010), back-arc extension (Segev et al., 2018), or kinematically related to the opening of the Atlantic Ocean (van Hinsbergen et al., 2020).

It is better established that the ECB existed as a carbonate platform on the passive northern margin of Gondwana until the Upper Cretaceous (Robertson, 1998). At this time, the widespread obduction of ophiolites represented a regional transition from divergent to convergent tectonics (van Hinsbergen et al., 2020). The subduction of Tethyan oceanic crust results in convergence between the northeastern African margin and Anatolia, however the location of this convergence is contentious. Subduction is interpreted to have occurred along the Cyprus Arc (Robertson, 1998; Schattner, 2010), or at the location of the northerly Cilicia Basin (McPhee & van Hinsbergen, 2019; van Hinsbergen et al., 2020). Calon et al. (2005a) and Calon et al. (2005b) interprets the Troodos-Larnaka culmination to be the centre of a diffuse zone of convergence. The tectonic model employed by McPhee & van Hinsbergen (2019) and van Hinsbergen et al. (2020) obducts the Troodos ophiolite onto the northern margin of the African plate during the Cretaceous. Therefore, there are three general models for where subducting Tethyan oceanic crust accommodated north-south convergence in the Eastern Mediterranean since the Upper Cretaceous.

As the African and Eurasian plates converge, an Eocene deformational event in the northerly Misis-Kyrenia Range (Calon et al., 2005a,b) is kinematically linked to Arabian and Eurasian plate convergence (Frizon de Lamotte et al., 2011; McPhee & van Hinsbergen, 2019). Miocene compression and Pliocene transpression are interpreted across the Cyprus Arc (Hall et al., 2005b; Calon et al., 2005a; Montadert et al., 2014; Reiche & Hübscher, 2015; Symeou et al., 2018). McPhee & van Hinsbergen (2019) attribute the compressional stage to the cessation of subduction north of Cyprus, when extended continental crust from the African margin underthrusted the subduction. This underthrusting represents the early stages of continent-continent collision in the Eastern Mediterranean and contributes to the uplift of Cyprus (Robertson, 1998; Schildgen et al., 2014).

The present-day plate boundary between the African and Eurasian plates lies along the Cyprus Arc, south of the Island of Cyprus, and north of the ECB (Symeou et al., 2018). Reiche et al. (2016) interpret a thick salt wedge between the ECB and

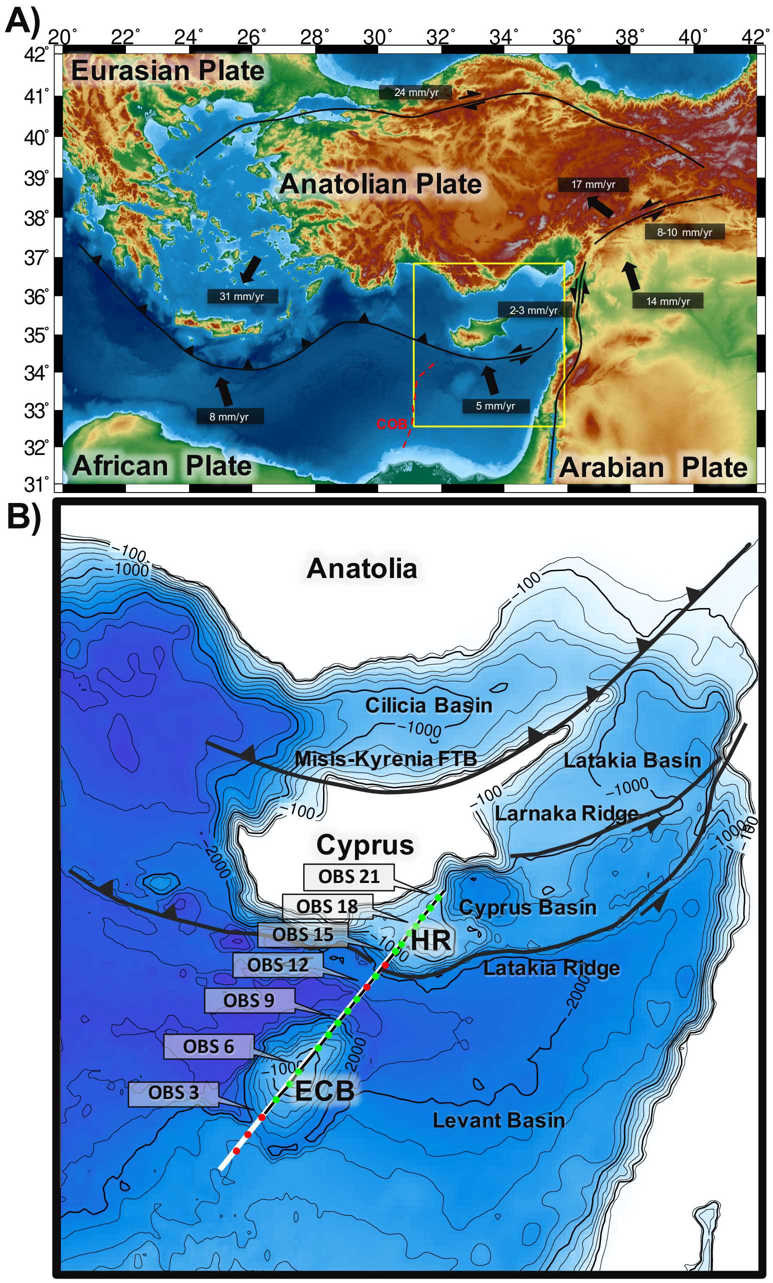

Figure 3.1: A regional (A) and local (B) topography map. The yellow box in the regional map represents the study area. In (A), the plate velocities are from Nocquet (2012) and the continent-ocean-boundary (COB) is from Granot (2016). In (B), ECB is the Eratosthenes Continental Block, and HR is the Hecataeus Rise. The white line is a seismic reflection line parallel to the refraction line shown in black. Red dots are OBS excluded from this study from Welford et al. (2015a) and the green dots are OBS used in this study. Every third OBS is labelled.

Hecataeus Rise along the Cyprus Arc, and Reiche & Hübscher (2015) interpret the continuation of structural lineaments from the island of Cyprus to the Hecataeus Rise. The Hecataeus Rise and ECB remained structurally high during the Messinian salinity crisis, and Reiche & Hübscher (2015) interprets only isolated basins of evaporites on the Hecataeus Rise. Carbonates of Upper-Cretaceous to Miocene age are confirmed present on the ECB by ODP Leg 160 with minor evaporite lenses (Robertson, 1998).

3.3.2 Eastern Mediterranean Seismic Dataset

Three marine seismic profiles we acquired south of Cyprus in 2010 (Welford et al., 2015a,b). Line 1 is recorded on 21 multi-component geophones and hydrophones and is the longest of these WARR profiles. The 21 instruments are from the Geological Survey of Canada (GSC), Dalhousie University, and Germany. The German research vessel R.V. Maria S. Merian recorded the seismic data. The research vessel fired shots using a 98 L airgun array approximately every 130 m for a total of 1812 shots. The OBS data have previously been converted to SEGY, corrected for timing, and relocated using the direct arrivals (Welford et al., 2015a).

A preliminary assessment of the OBS data reveals that the hydrophones are noisy, and the vertical component geophones record less-noisy data. However, the vertical component geophones for OBS 12 and 14 did not record data. Despite the absence of two stations, we use the vertical component geophones as the FWI dataset. Figure 3.1 shows all of the OBS, highlighting the ones used in this study. The omission of OBS 2 and 3 is due to poor data quality, and the removal of a higher quality OBS 1 is a result of the approximately 40 km long void created between OBS 1 and 4 following the omission of OBS 2 and 3. The resultant Eastern Mediterranean dataset for FWI is composed of 16 OBS of German and GSC origin. We decrease the extent of the model from 260 km (Welford et al., 2015a) to 190 km; the resulting average OBS spacing is approximately 12 km. These OBS densely sample the Hecataeus Rise to the north, and sparsely sample the northern edge of the ECB and Cyprus Arc. The most geologically relevant region between the ECB and the Hecataeus Rise is now the primary focus of this study.

3.4 Frequency-Domain FWI

Pratt & Worthington (1990) and Pratt (1990) introduce frequency-domain waveform inversion under the acoustic approximation of the wave equation and the elastic wave equation respectfully. With the frequency-domain formulation, seismic attenuation may be modelled through complex velocities (Toksöz & Johnston, 1981), yielding the viscoacoustic frequency-domain formulations of the wave equation. The model space is then parameterized by P-wave velocity, density, and attenuation. Frequency domain FWI is a computationally efficient alternative to time-domain FWI due to the discretization of the forward problem by frequency. It also provides a natural environment to implement multiscale frequency progression strategies which aid in convergence to the global minimum (Pratt et al., 1996). Frequency domain viscoacoustic FWI is popular, but there are fundamental issues with the methodology. It is pointed out by Mora (1987) that wave propagation in the subsurface is fundamentally elastic, and the acoustic approximation is unable to reproduce the recorded seismic response. Moreover, Prieux et al. (2011) perform FWI sensitivity tests, demonstrating the importance of correctly modelling anisotropic media in the subsurface.

This study uses the open-source TOY2DAC algorithm from the SEISCOPE Consortium (available at `https://seiscope2.osug.fr/TOY2DAC,82?lang=en`) to implement frequency-domain FWI on the Eastern Mediterranean dataset. TOY2DAC implements 2D viscoacoustic FWI in the frequency domain and can model isotropic or anisotropic (VTI/HTI/TTI) media (Hustedt et al., 2004; Operto et al., 2009). The discretization of the forward problem uses a fourth-order finite difference scheme (Hustedt et al., 2004), and the resultant sparse linear system is solved using lower-upper (LU) factorization as proposed for FWI by Pratt et al. (1998). Solving for the complex-valued "impedance" matrix in the frequency domain allows for it to be stored in memory and then re-used to solve the adjoint problem and recover the gradient (Plessix, 2006; Métivier et al., 2014). The MUMPS package (Amestoy et al., 2000) is used by TOY2DAC to perform LU factorization and solve the forward problem.

Within TOY2DAC, data misfits use the least squares method where a data-weighting function is included to balance the contribution of lower amplitude far-offset arrivals to the objective function (Górszczyk et al., 2017). TOY2DAC uses the SEISCOPE Optimization Toolbox, which contains a suite of numerical optimization routines tailored for large, sparse inverse problems (Métivier & Brossier, 2016).

Here we use the l-BFGS algorithm for the FWI optimization problem, which incorporates information contained within the pseudo-Hessian matrix as a preconditioner (Shin et al., 2001). Incorporating the preconditioner improves the efficiency of the l-BFGS algorithm, which is repeatedly re-estimated at each iteration for little additional computational cost (Métivier et al., 2014; Métivier & Brossier, 2016). Damping the preconditioner stabilizes its implementation, and decreasing the damping value favours deeper velocity perturbations. Therefore, consistently decreasing the damping value after each inversion, we may incorporate decreasing the gradient preconditioner in a multiscale inversion strategy (Ravaut et al., 2004; Górszczyk et al., 2017).

This study utilizes the multiscale FWI approach from Górszczyk et al. (2017) to invert the OBS data. Their approach consists of a three-loop inversion, with the external loop being a frequency progression from low to high frequencies. This exterior loop is critical for the inversion to avoid convergence to a local minimum (Pratt et al., 1996). A middle loop implements waveform inversion in the Laplace-Fourier domain (Shin & Cha, 2009). The degree of wavefield damping is progressively lowered, which systematically introduces more complex waveforms into the inversion. The Laplace constant governs the amount of damping, and data-weights are incorporated in the frequency domain to apply damping from the first-arrival times (Brossier et al., 2009). An inner loop progressively increases the maximum offset of data to be inverted. As the maximum offset increases, the gradient preconditioner is decreased, resulting in deeper velocity perturbations (Ravaut et al., 2004). Górszczyk et al. (2017) provides additional information about their multiscale approach and a complete description of viscoacoustic FWI using TOY2DAC.

3.5 Synthetic Testing

3.5.1 Sparse-Marmousi Synthetic Tests

We use the well-known Marmousi2 model (Martin et al., 2006) to test the effect of data sparsity and the potential of a multiscale inversion to mitigate these effects. A sparse synthetic dataset is defined here to contain 17 explosive sources, whereas the standard test dataset contains 170 explosive sources. Synthetic data generation for each dataset utilizes 660 pressure receivers spaced at 25 m. The sparsity difference

Table 3.1: This table shows the inner two loops of the multiscale FWI strategy for the Marmousi synthetic test. It depicts different Laplace constant values for loop two, and maximum offset constraints and preconditioning values for loop three. The total number of iterations ran for each inversion group is at the intersection of the Laplace constant at a given row and offset/preconditioner value for a given column. We repeat the inner two inversion loops for each frequency group, and the maximum number of iterations per offset is on the bottom. Iterations is abbreviated as It. in this table.

Loop Three	Offsets [km] Preconditioner	4 10^{-2}	8 10^{-3}	12 10^{-4}	17 10^{-5}
Loop Two	Laplace Constant [s]	**It. per inversion group**			
	5.0	*5*	*N/A*	*N/A*	*N/A*
	2.0	*5*	*5*	*N/A*	*N/A*
	1.0	*5*	*5*	*5*	*N/A*
	0.5	*5*	*5*	*5*	*5*
	It. per offset	20	15	10	5

of an order of magnitude reflects an approximate 10 km OBS spacing in this study versus the approximate 1 km OBS spacing of Górszczyk et al. (2017).

All three synthetic inversions use the same progressive frequency strategy (Bunks et al., 1995). In the progressive frequency approach, each frequency group adds a single frequency to an expanding group of inversion frequencies, such that each inversion includes all previously inverted frequencies. The Marmousi synthetic examples invert within the 2-12 Hz bandwidth and append frequencies in 0.5 Hz increments. The first frequency group contains three frequencies from 2-3 Hz. Additionally, all inversions apply linear data-weighting to the cost function where the nearest offsets are multiplied by 0.5 and the farthest offsets are multiplied by 1.0. A free surface boundary condition speeds up computational time and conceptually provides a secondary wavefield to further decrease the cost function. However, proper estimation of secondary scattering requires a good estimation of the Hessian matrix (Pratt et al., 1998).

The sparse synthetic dataset is inverted with and without the middle and inner loops of the multiscale inversion strategy. Table 3.1 defines the middle Laplace constant loop and an inner offset loop for the synthetic Marmousi study. The Laplace constant loop varies from 5s to 0.5s, a high to a low amount of wavefield damping. Computing the first arrival travel times using TOMO2D (Korenaga et al., 2000) allows for the weighted application of wavefield damping from the first breaks. The inner loop inverts the data at four different maximum offsets with four different gradient preconditioning values.

Table 3.1 additionally shows the maximum number of FWI iterations to be run at each combination of Laplace constant and maximum offset. The first column of this table shows the Laplace constant values for loop two, and the first two rows show the maximum offset and preconditioning values for loop three. Beginning with high Laplace constants, the values of loop three in the first two rows vary from left to right. The values along each Laplace constant row depict how many iterations of FWI to run for each inversion group. For particular Laplace constant and maximum offset combinations, FWI is not run, and N/A denotes these combinations the table. This approach resembles that of Górszczyk et al. (2017), which excludes long offsets from being inverted with large Laplace constants. This structure limits the total number of iterations and also prevents artificial velocity perturbations in the recovered model (Górszczyk et al., 2017). The multiscale strategy will run a maximum of 50 iterations for each frequency group with this approach. Adding together the number of iterations for each inversion group in Table 3.1 produces this value. We desire a baseline inversion to compare this multiscale inversion strategy to an inversion that omits the inner two loops, and inverts the data by only progressively increasing frequencies. To ensure the results of this baseline inversion will be comparable to the multiscale inversion, we run 50 iterations of FWI for each frequency group for the baseline inversion, without the Laplace constant or offset loops. Therefore, both the baseline and multiscale inversions run an equivalent maximum number of FWI iterations per frequency group.

The maximum number of iterations, together with the model update convergence criteria, define how many iterations of FWI to run for each inversion group. The model update convergence criteria is,

$$\epsilon > f(x_{k-1}) - f(x_k). \tag{3.1}$$

This equation describes that the cost function at the current iteration, $f(x_k)$, must have adequately decreased from the cost function at the previous iteration, $f(x_{k-1})$. For the Marmousi synthetic tests, ϵ is 10^{-4}. It is beneficial to define stopping criteria this way to avoid over-fitting the observed data. The choice of ϵ governs the trade-off between resolution and variance in the final inversion result. Of course, choosing this value is not as pressing an issue during the synthetic tests as it is for the real-data inversion.

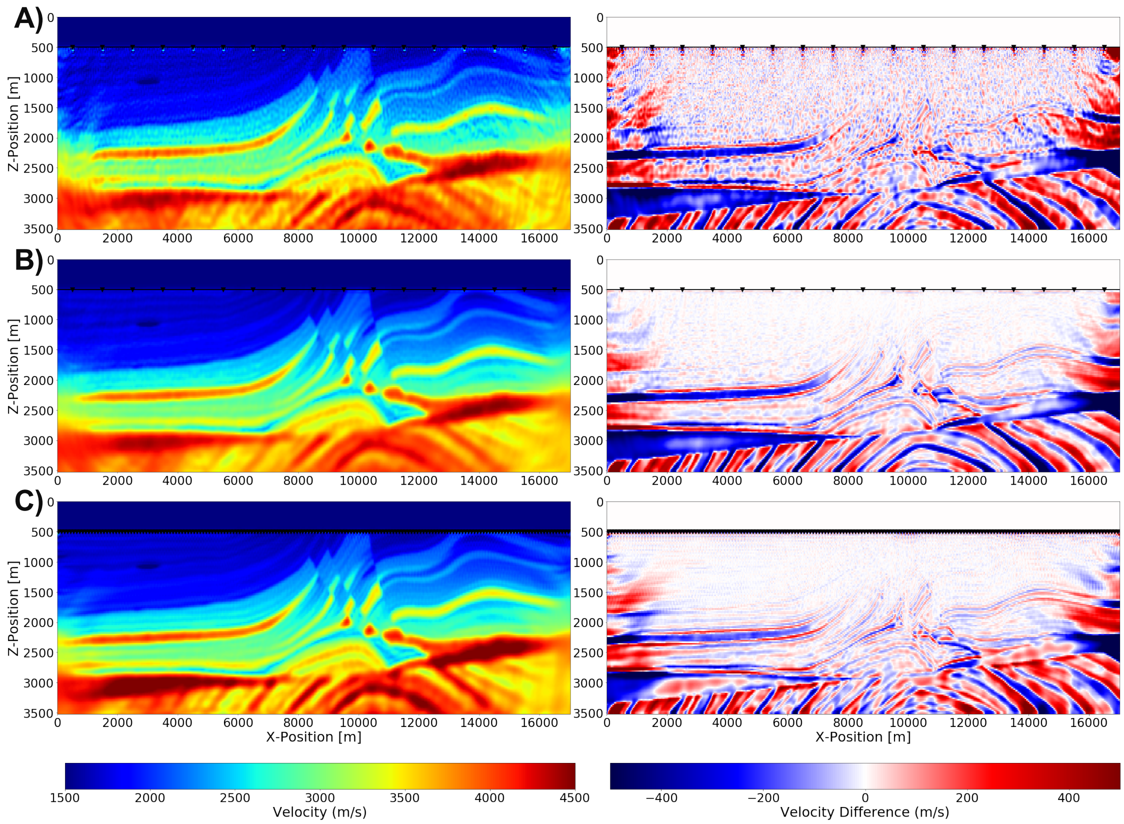

Figure 3.2: A sparse and dense synthetic dataset are inverted using the Marmousi2 model. The recovered velocity model is shown to the left, and the residual velocity field is shown to the right. The residual field is obtained by subtracting the true model from the inverted model. The sparse dataset in (A) is inverted by progressing from low to high frequencies, as is the dense dataset in (C). The sparse dataset in (B) is inverted using a three-loop multiscale strategy. The synthetic geometries for (A) and (B) are identical.

The synthetic inversion result for a baseline inversion with the sparse dataset is in Figure 3.2A, the multiscale inversion with the sparse dataset is in Figure 3.2B and a baseline inversion with the dense dataset is in Figure3.2C. In each figure, the recovered velocity model is on the left-hand side, and the residual velocity field, obtained by subtracting the true velocity model from the inverted model, is on the right-hand side. The residual velocity field effectively measures the accuracy of the inversion, and for a perfect inversion, the residual velocity field would be zero. The 17 shot baseline inversion is still able to recover the main structure of the Marmousi model. The velocity updates are particularly noisy, however, and near-surface high velocity residuals underlie the locations of the shots. These inversion artifacts vanish when we increase the amount of data by an order of magnitude, and by implementing the multiscale inversion strategy. Although a faint vertically oriented acquisition footprint is noticeable in the velocity residuals from the sparse multiscale inversion, both the sparse multiscale inversion and dense baseline inversion adequately recover the Marmousi model. The dense data inversion outperforms the sparse multiscale inversion only slightly in the shallow and central portions of the model.

While the multiscale FWI strategy is useful, it is not as effective as including more data in the inversion. However, at the scale of the Marmousi model, it may be argued that weighting the wavefield damping from the first breaks may not be the best strategy. To image the Marmousi structure, far offset diving waves recover long-wavelength background structure and near offset reflections illuminate the short-wavelength structure. Many of the near offset reflections are attenuated by weighing the application of damping from the first arrivals. While preferably inverting for short-wavelength reflections would improve the resolution, doing so would increase the probability of encountering cycle-skipping problems. A potential solution would be a hybrid approach that will first preferentially recover long-wavelength diving waves by weighing the damping application from the first breaks. Then, the inversion would preferentially recover the short-wavelength reflections by damping from a constant time (i.e. $t = 0$). This hybrid approach would only desirable for smaller-scale seismic reflection experiments such as this Marmousi synthetic test, where the recovery of near offset reflections is desired. In contrast, we wish to recover far offset reflections in WARR experiments and application of data damping weighted from the first arrivals is recommended.

3.5.2 Synthetic Eastern Mediterranean Model

We now test our multiscale strategy on a synthetic version of the model we hope to recover from our field dataset. To perform this synthetic experiment, we use a cropped version of the final model from Welford et al. (2015a) as the true model. To estimate the starting FWI model we use a Gaussian filter of 2 km in the x- and 4.5 km in the z-dimensions to smooth the original model. The mesh spacing for these models is 50 m, and there are 3801 and 601 nodes in the x- and z-dimension, respectfully. Both the true and starting models for this synthetic example are shown in Figure 3.3.

We then generate a synthetic observed dataset based on the real-data geometry

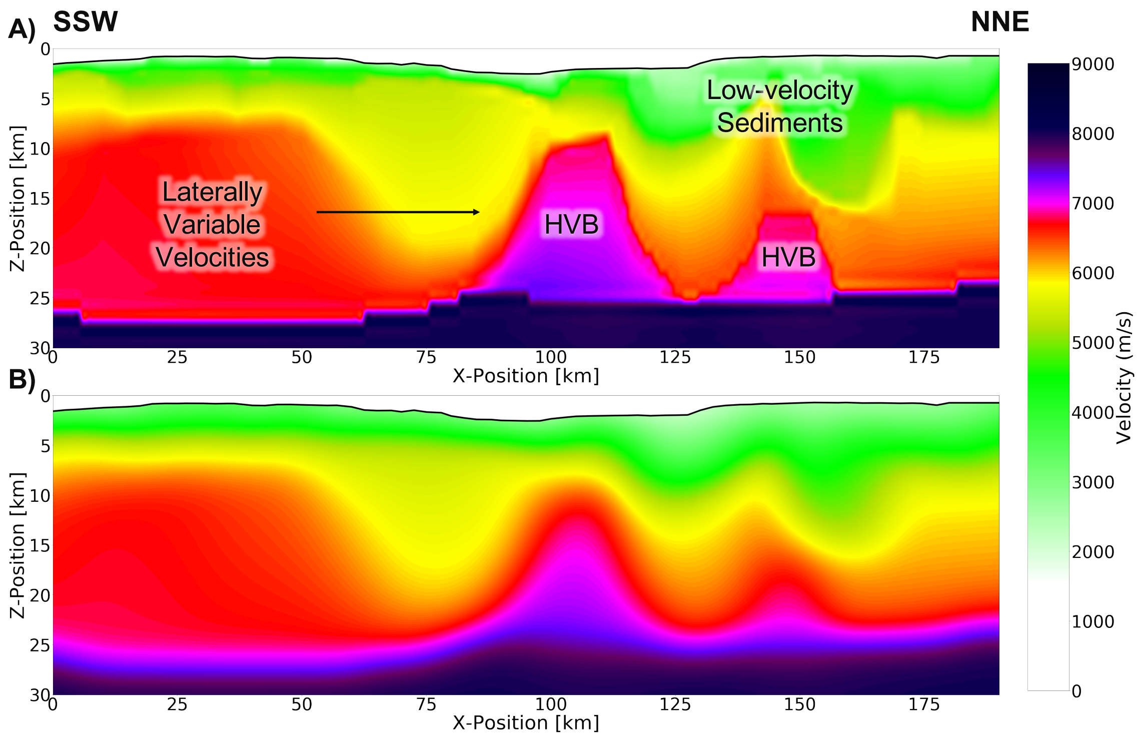

Figure 3.3: The true synthetic Eastern Mediterranean model in (A) is a cropped version of the final model from Welford et al. (2015a). The starting model for FWI is a smoothed version of the true model. HVB, high-velocity body.

using 16 OBS and 8763 first break picks. The density model for this synthetic inversion is generated by applying Gardner's relation (Gardner et al., 1974) to the starting velocity model. The attenuation model is identical to the model used in the real data inversion as well, with a subsurface attenuation value of 50 and a water column attenuation value of 10,000.

We use the multiscale inversion strategy of Górszczyk et al. (2017) once again for this example, but there are two main differences between this inversion and the Marmousi synthetic example. Forward modelling travel times through the starting model and the true model reveals that some traces are not matched to within half a wavelength at approximately 0.75 Hz, meaning that we have to start from lower frequencies. The first frequency group for this inversion will thus contain 0.25 Hz, 0.5 Hz, and 0.75 Hz frequencies. The 0.25 to 5.25 Hz bandwidth inversions use the same progressive frequency strategy as the Marmousi example, but with a 0.5 Hz frequency step size outside of the first frequency group.

Table 3.2: This table shows the inner two loops of the multiscale FWI strategy for the synthetic Eastern Mediterranean test. It depicts different Laplace constant values for loop two, and maximum offset constraints and preconditioning values for loop three. The total number of iterations ran for each inversion group is at the intersection of the Laplace constant at a given row and offset/preconditioner value for a given column. We repeat the inner two inversion loops for each frequency group, and the maximum number of iterations per offset is on the bottom. Iterations is abbreviated as It., and inversion is abbreviated as inv. in this table.

Loop Three	Offsets [km] Preconditioner	25 10^{-2}	50 10^{-4}	100 10^{-6}
Loop Two	Laplace Constant [s]	**It. per inv. group**		
	4.0	*8*	*N/A*	*N/A*
	1.0	*8*	*12*	*N/A*
	0.5	*8*	*12*	*24*
	It. per offset	24	24	24

The second difference in this synthetic inversion in comparison to the Marmousi inversion is the iteration structure. Table 3.2 shows a logical partitioning of FWI iterations such that the maximum number of iterations is constant for each offset. This technique ensures adequate illumination of the deep structure at far offsets. This inversion additionally encourages a higher number of total iterations by setting the model update convergence criteria in equation 3.1 to 5×10^{-5}. A small degree of smoothing as a function of wavelength is introduced (Ravaut et al., 2004), and the data weighting function now weights the nearest offset data as zero and the furthest offset data as one.

The synthetic Eastern Mediterranean inversion results are shown in Figure 3.4A, and the updated and residual velocity fields in Figures 3.4B and 3.4C, respectfully. Subtracting the starting velocity model from the inverted velocity model produces the updated velocity model, and this model should resemble the main structural elements of the true model. The recovered velocity model from FWI does a good job of approximating the true velocity model shown in Figure 3.3A. The inverted model is a good approximation of the shallow subsurface velocity structure above approximately 10 km depth. From 10-20 km, the outline of the larger high-velocity body, centred at x=105 km, is well represented in the velocity update model. However, there is a poor recovery of this high-velocity body's interior structure. Below 20 km depth, the inverted model is not a good representation of the true velocities. A second high-velocity body at x=145 km is not resolved by the inversion, as evidenced in the

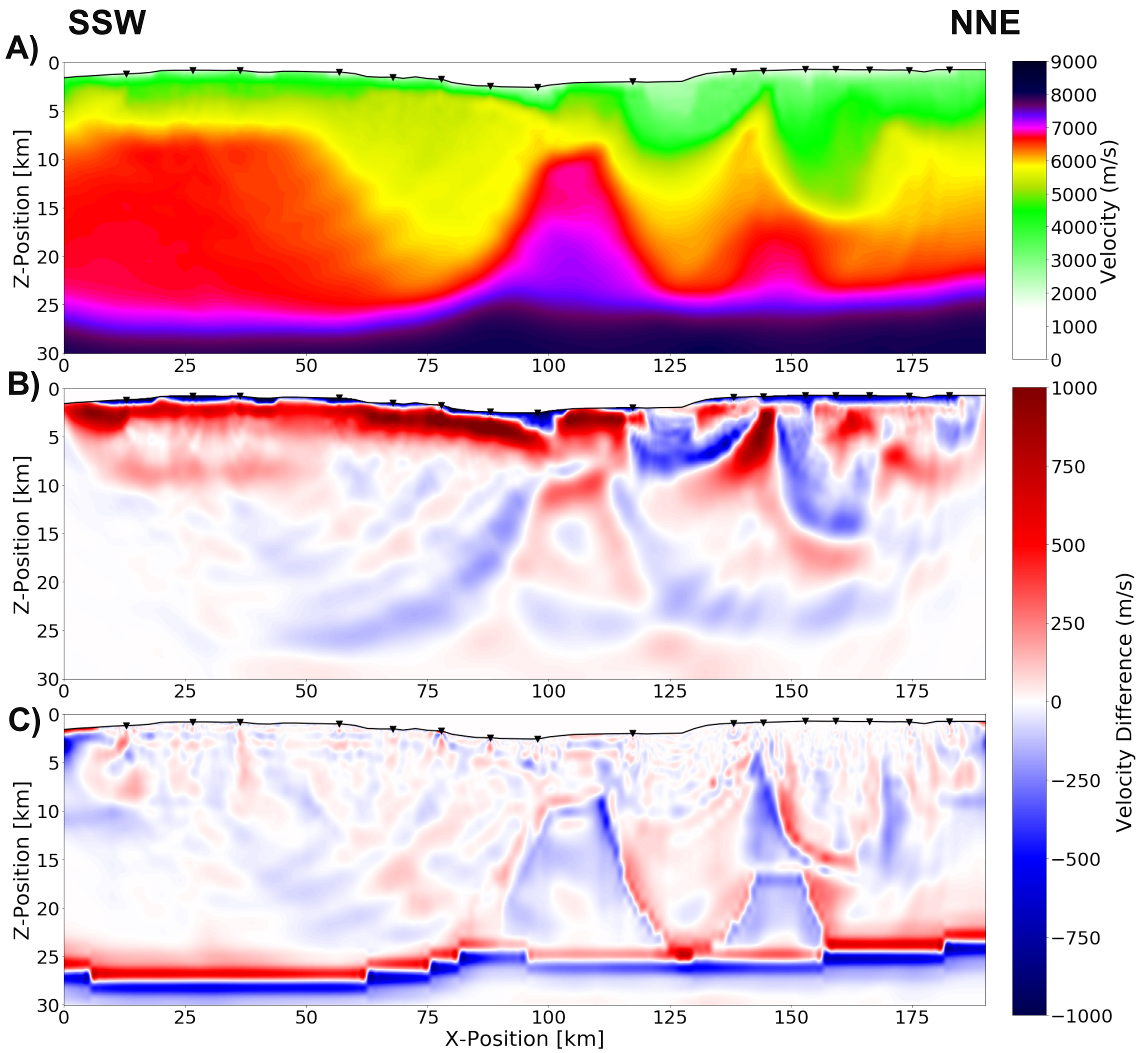

Figure 3.4: The inverted synthetic Eastern Mediterranean is in (A), the updated velocity field is in (B), and (C) shows the residual velocity field. The updated and residual velocities fields show an encouraging recovery of the correct velocity structure down to 10 km depth. The main structural elements from 10-20 km depth are visible in the velocity update, but the recovery of these elements deteriorates, as shown in the residual velocity field.

velocity update.

The recovered velocity model suffers from the sparsity of the Eastern Mediterranean acquisition. The same vertically oriented artifacts beneath the modelled source locations are present in the residual velocity model in Figure 3.4C as in the multiscale Marmousi inversion in Figure 3.2B. The velocity update model in Figure 3.4B shows velocity updates that appear to be angled corresponding to the propagating ray paths at depth, especially in the 10-20 km range at x>80 km. The low starting frequencies in this synthetic example mitigate their presence, but the sparse dataset contributes to generating these artifacts.

Despite the effects of data sparsity being evident in this synthetic inversion, it converges acceptably to a good approximation of the true synthetic model. However, this inversion utilizes unrealistically low frequencies to achieve this result. Based on these results, it is likely that the real data inversion will be limited to the recovery of structure in the upper 10 km, with a possibility of recovering the crustal structure down to 20 km. Vertically oriented model artifacts beneath the OBS locations are expected in the real-data inversion as they are prominent in both synthetic inversions. Additionally, velocity artifacts along the ray path are likely to be an issue given data sparsity, and the frequency bandwidth of the real data.

3.6 Real-Data Inversion

3.6.1 OBS Data Processing

The OBS data processing flow employed in this study requires three phases, and is built by considering relevant FWI studies (Ravaut et al., 2004; Brenders & Pratt, 2007; Operto et al., 2006; Kamei et al., 2012; Górszczyk et al., 2017). Following the data QC described in subsection 3.3.2, the second stage consists of standard data processing steps. The first of these steps consists of a spherical divergence correction applied by multiplying the data by $\sqrt{t}$, which corrects for the 3D-2D conversion of wavefield amplitudes (Pratt, 1999). The second step consists of whitening the amplitude spectrum to remove notches in the spectra and balance the relative contributions of different frequencies to the objective function (Kamei et al., 2012; Górszczyk et al., 2017). While low frequencies are essential to mitigate cycle-skipping, the better signal is present at higher frequencies in the Eastern Mediterranean dataset. This strategy emphasizes the contribution of superior signal in the FWI objective function while smoothing the amplitude spectra of each OBS. We next apply a filter to match a time-domain estimation of the wavelet, which was close to zero-phase, to an idealized zero-phase Ricker wavelet for each shot. This processing step smooths the amplitude spectra of notches and minimizes the risk of FWI fitting low-frequency noise. The application of this filter acts as a band-pass filter within the bandwidth of the designed Ricker wavelet, which is approximately 0-18 Hz for each OBS. The final step for data processing is to re-sample the OBS data to 4 ms. No conventional coherency filter is applied to the data as it alters the AVO relationships in the data.

While the first phase of data processing pertains to noise removal, the second phase processes the amplitudes. The RMS amplitudes for each gather are first bulk-shifted to one, which accounts for differences in instrument coupling to the seafloor. Rather than muting noisy traces, this study follows the approach of Górszczyk et al. (2017), where the noise root mean square (NRMS) window is analyzed, and traces with high NRMS values have their amplitudes progressively scaled-down. The normalization of OBS amplitudes before this processing step is required to automate it for all OBS. Another bulk-shift will now match the amplitudes of the observed dataset with those of the modelled dataset. The modelled data are generated by forward modelling the viscoacoustic wavefield through the starting FWI model. The amplitude discrepancy is minimized between the modelled and observed data through the application of a single constant, ensuring that AVO relationships are unchanged for the observed data.

The final data processing step consists of generating the frequency-domain observed data for FWI. The time-domain data are damped and muted above the first break picks before a Fourier transform and FWI frequency extraction. This OBS data processing flow is not independent of the FWI starting model generation, where multiple processing steps require first break picks. Starting FWI velocity, density, and attenuation models are required to obtain the correction for the amplitude discrepancy between observed and predicted data. This data processing flow then requires a set of starting FWI models, and first break picks to complete, so building the starting model is done in parallel with OBS data processing.

3.6.2 FWI Starting Model

Viscoacoustic FWI requires a highly accurate starting P-wave velocity model. While the crustal-scale velocity model produced by Welford et al. (2015a) is a good representation of the data and the subsurface geology, it is not sufficiently accurate to be a starting model for FWI. Welford et al. (2015a) report a total RMS misfit of 143 ms considering reflected and refracted phases for their model. Successful crustal-scale

FWI studies report total RMS misfits between 48 and 60 ms for their starting velocity models (Kamei et al., 2012; Davy et al., 2017; Górszczyk et al., 2017). We use first arrival tomography (FAT) using TOMO2D (Korenaga et al., 2000) to build our starting velocity model.

The first break picks from Welford et al. (2015a) are revisited for this study as well. Re-picking first breaks on the Eastern Mediterranean dataset attempts to in-fill gaps in the picks due to weak signal. Additionally, the first break picks are adjusted to be zero-phase. We do this iteratively. First, the tomographic inversion parameters are adjusted to produce a velocity model that best reproduces the first break picks and is geologically reasonable. Should convergence be unsatisfactory, we readjust the first break picks in problematic noisy regions to improve the result of the inversion. For this dataset, traveltime tomography must be implemented as a multiscale inversion to converge to the minimum. Traveltime data are inverted in two-second increments up to 20 seconds. The inversion incorporates uncertainty in first break picks according to computed signal-noise ratios (Welford et al., 2015a).

The velocity model shown in Figure 3.5A is the starting velocity model for FWI. Its upper crustal structure resembles that of Welford et al. (2015a) shown in Figure 3.3A, however, the deeper structure in the tomographic model is missing the two high-velocity bodies. Both velocity models show a laterally variable velocity structure throughout the ECB and low-velocity sediments occupying the region north of the Cyprus Arc. The Eastern Mediterranean is geologically complex, which results in complexities during traveltime tomography. There are substantial lateral velocity variations, high-velocity salt, high-velocity carbonates, low-velocity sediments, and high-velocity igneous intrusions interpreted in this study area (Reiche et al., 2016; Reiche & Hübscher, 2015; Welford et al., 2015a,b; Feld et al., 2017). The geological complexity and data sparsity complicate traveltime tomography.

In principle, the starting model should be able to forward model the wavefield to within half a wavelength of the observed waveforms (Sirgue & Pratt, 2004). A threshold pick residual may be defined for a given frequency where residual travel times greater than this value are subject to cycle-skipping, and less than this value are not,

$$\frac{1}{2} > f\left|t_{pick} - t_{model}\right|, \tag{3.2}$$

where f is frequency in cycles per second, $|t_{pick} - t_{model}|$ is the pick residual for the modeled and observed first break picks. This inequality assesses all seismic arrivals for each OBS, as shown in Figure 3.5B. This model assessment technique assumes that the first break picks are perfect. Of course this is likely not the case, but assessing the starting velocity model for FWI in such a way is quick and practical. This starting model assessment reveals that most of the data are modelled to within half a wavelength at 2.25 Hz, but one problematic region is evident near x=150 km. A histogram of the pick residuals is shown in Figure 3.5C. It emphasizes the importance of low frequencies in the data as the reciprocal nature of Equation 3.2 results in the acceptance of poorly recovered arrival times at lower frequencies. The total RMS misfit is also plotted at 57.8 ms, which is sufficient based on previous WARR studies.

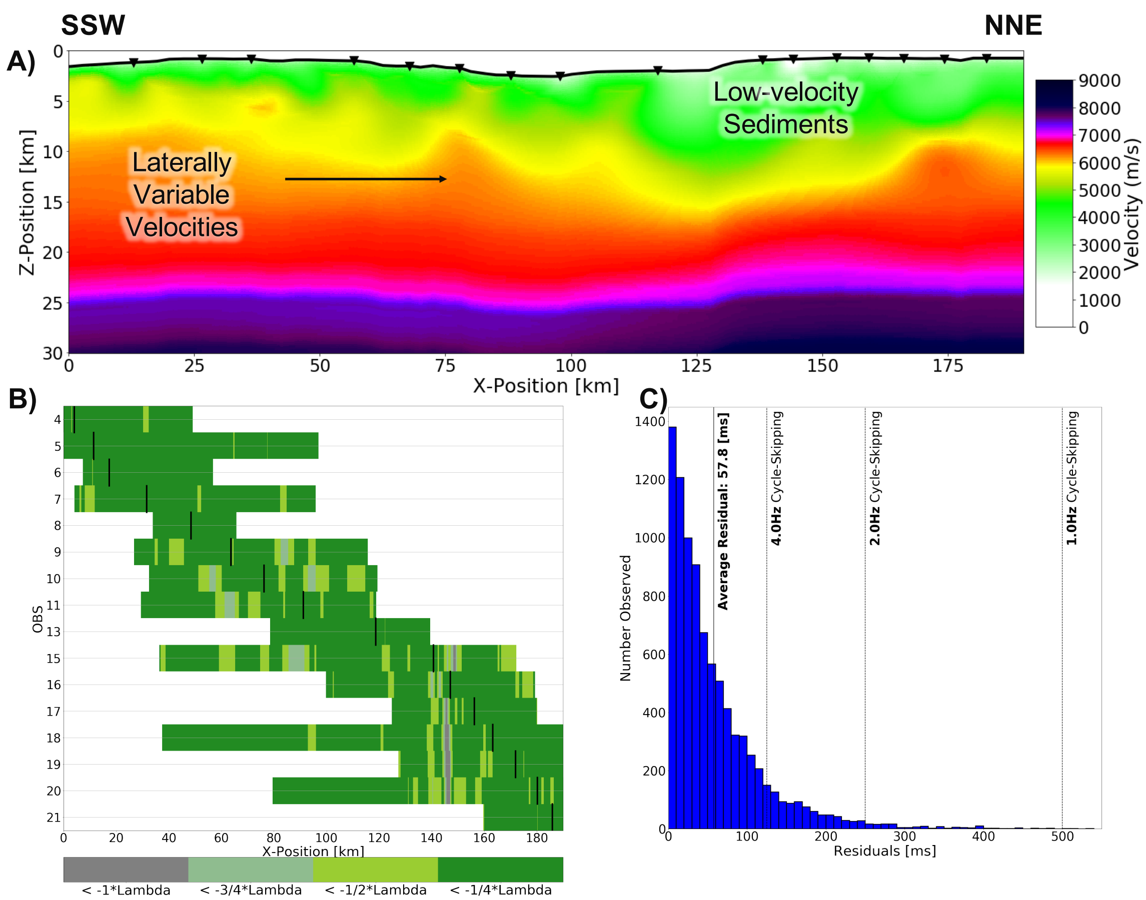

Figure 3.5: The starting velocity model for FWI is in (A). Two assessments of model accuracy are in (B) and (C). In (B), the traveltime residuals computed for cycle-skipping given a frequency of 2.25 Hz. In (C), the pick residuals in (B) are shown with cycle-skipping cutoffs at 1.0, 2.0, and 4.0 Hz, as well as the RMS pick residual. The bin size for the histogram is 10 ms.

We use FAT to invert a model with a mesh spacing of 200m, which is far too sparse for FWI at meaningful frequencies. The recovered tomographic model must be interpolated to a node spacing of 50 m for which we use nearest neighbour interpolation. The starting density model is generated by applying Gardner's relation (Gardner et al., 1974) to the FAT velocity model, like the approach used for the synthetic Eastern Mediterranean inversion in subsection 3.5.2. By comparing RMS amplitude decay curves between observed and modelled data with various subsurface attenuation values, the attenuation model has a subsurface value of 50. This is a high degree of attenuation in comparison to a subsurface value of 200 used by Górszczyk et al. (2017). The attenuation model is built this way as even a simplistic approximation of attenuation improves the results of FWI (Kurzmann et al., 2013; Górszczyk et al., 2017)

3.6.3 Eastern Mediterranean Inversion

The Eastern Mediterranean dataset is inverted using the same multiscale FWI strategy as in both the synthetic examples. We use a more computationally intensive frequency strategy than previous frequency-domain inversions due to the sparsity of the OBS dataset. The progressive frequency strategy from Bunks et al. (1995) is used to invert the data from 2.0 to 7.25 Hz, where the first frequency group is 2.0, 2.25, and 2.5 Hz, and each subsequent frequency group adds an additional frequency 0.25 Hz higher to the group. This frequency strategy results in 20 frequency groups and 250 total frequencies to invert for throughout the inversion. We test both an overlapping frequency strategy (Brossier et al., 2009) and a quasi-progressive frequency strategy (Górszczyk et al., 2017) as well as varying densities of frequency sampling as small as 0.05 Hz. The progressive strategy outperforms the other strategies, and there is a diminishing improvement with progressively smaller frequency spacing beyond 0.25 Hz.

The recovered velocity model after 257 iterations of FWI is shown in Figure 3.6A. Figure 3.6B shows the velocity model update from the starting tomographic model, and the velocity update from a detrended background model is shown in Figure 3.6C. We generate this detrended background model based on the methodology of Kamei et al. (2012). A three-term polynomial is fit to each column of the starting velocity model before increasing the node spacing from 200 m to 50 m. Subtracting the

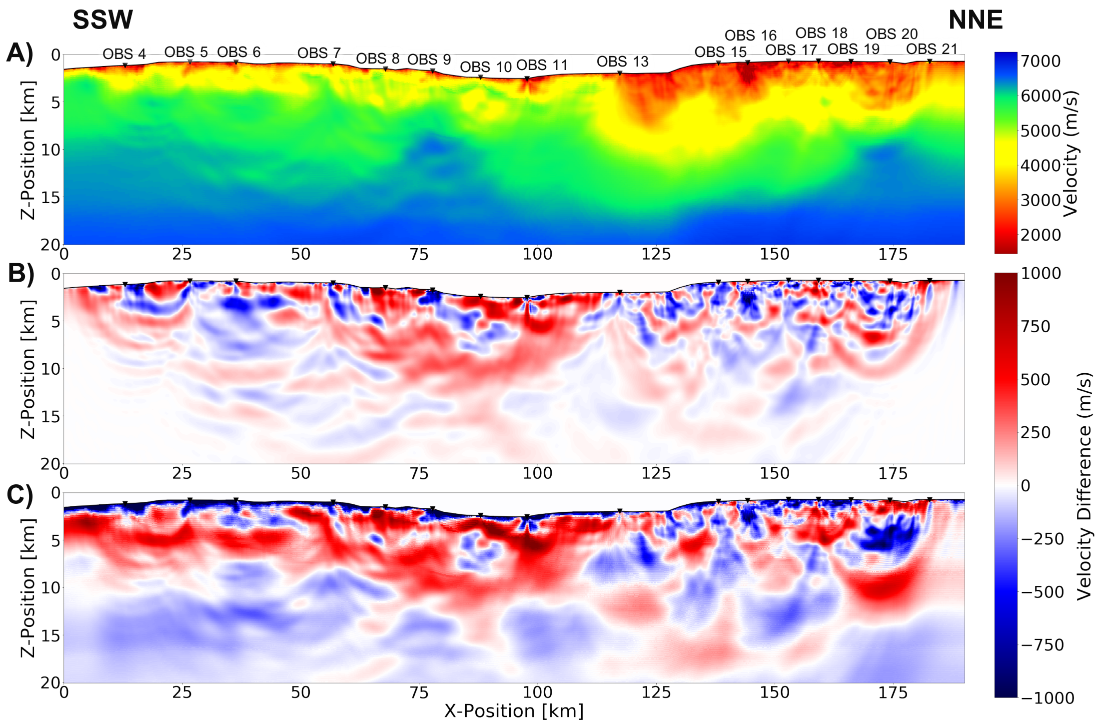

Figure 3.6: (A) The recovered Eastern Mediterranean velocity model, in (B), the updated velocity field from FWI, and in (C), the updated velocity from a detrended background velocity model.

detrended tomographic model from the recovered FWI model is useful for geologic interpretations as it will include long-wavelength geologic trends within the tomographic model.

The real-data inversion parameters for the inner two loops are given in Table 3.3. The offset strategy is the same as the synthetic Eastern Mediterranean inversion, but this configuration is highly sensitive as FWI did not converge with five offset groups. Furthermore, we decrease the first Laplace constant from 4.0s to 2.0s, as FWI did not converge with greater data-damping. Also, we decrease the number of iterations to a maximum of five per inversion group, consistent with the Marmousi synthetic example in subsection 3.5.1. While running a consistent number of iterations per offset improved the recovery of deep velocity structure for the full-scale synthetic example, it mainly resulted in the generation of near-surface velocity artifacts in the real data example.

In practice, the sparse data inversion does not acceptably converge without a carefully designed data weighting function. Here we use a function that mutes the first 5 km of offset, then uses linear data weight increases from zero to two at the maximum offset. Muting the first 5 km of data is critical as FWI generates model perturbation artifacts due to the high amplitude direct wave and near offset reflections. Brenders & Pratt (2007) applies a near offset mute in their study as well, but the authors apply a gain function to the gradient to achieve a similar effect. Regularization in the form of Gaussian smoothing applies a constant, moderately high degree of smoothing weighted twice as much in the x-dimension as the z-dimension to the approximate Hessian and gradient. The model convergence update criteria in Equation 3.1 is set to $\epsilon = 10^{-4}$. If this parameter is too large, there will not be any significant velocity perturbations, and if it is too small we will over-fit the data. As vertical component geophones record the observed dataset with an explosive source, the data are modelled reciprocally with vertical component sources and pressure receivers (Operto et al., 2006).

The FWI velocity update indeed contains more structure than the tomographic model. A high-velocity feature appears in the recovered model at depth between x=60 and x=115 km. There is a lateral discontinuity in the velocity updates at approximately x=115 km, and the primary velocity updates are within the upper 15 km. There are also several artifacts in the recovered model, for example, the velocity update just beneath OBS 11. Also, the presence of a low-velocity section beneath OBS 16 is unrealistic. We further assess the model in the next subsection.

Table 3.3: This table shows the inner two loops of the multiscale FWI strategy for the Eastern Mediterranean inversion. It depicts different Laplace constant values for loop two, and maximum offset constraints and preconditioning values for loop three. The total number of iterations ran for each inversion group is at the intersection of the Laplace constant at a given row and offset/preconditioner value for a given column. We repeat the inner two inversion loops for each frequency group, and the maximum number of iterations per offset is on the bottom. Iterations is abbreviated as It., and inversion is abbreviated as inv. in this table.

Loop Three	Offsets [km]	25	50	100
	Preconditioner	10^{-2}	10^{-4}	10^{-6}
Loop Two	Laplace Constant [s]	**It. per inv. group**		
	2.0	*5*	*N/A*	*N/A*
	1.0	*5*	*5*	*N/A*
	0.5	*5*	*5*	*5*
	It. per offset	15	10	5

3.6.4 FWI Model Assessment

To assess the quality of our recovered model we begin by assessing the data fit. We begin in the frequency domain, where the data are forward modelled through all frequencies in the final frequency group and with the final Laplace constant. The linear data weighting function used during FWI is used to then weight the predicted data, and then the frequency-domain damping weights from the first breaks are applied. By completing the same workflow for the observed and starting data, frequency domain traces of the observed, starting, and predicted data can be fairly compared. We plot these data at 5.0 Hz for OBS 20 in Figure 3.7A. The data weighting function is effective at balancing the amplitudes from near to far offsets. However, while we see clear signal until approximately 55 km offset, at that point the frequency-domain trace becomes much noisier. The predicted data generally fit the observed dataset better than the starting model but is not a perfect fit. However, for OBS 20 at 5.0 Hz there is a clear improvement over the initial tomograpic model and given the sparsity of the dataset and geological complexity of the region we consider this a good result. In Figure 3.7B, only the predicted and observed data are plotted to highlight the fit better. The fit is the poorest toward the positive offsets. The location of OBS 20 (Figure 3.6) may explain this poor fit as the model parameters at the exterior of the model are poorly constrained.

Figure 3.7 is useful, but it only shows the FWI result for a single shot-frequency combination. We now employ a second strategy to quickly assess all shot-frequency combinations and compare the fit of the predicted data from the final FWI model to the predicted data from the traveltime tomography model. The RMS difference is computed for each shot and modelled frequency between the observed data and both the predicted and starting data. Figure 3.8A summarizes how the misfit of the predicted dataset decreases compared to the starting dataset. Specifically, blue colours indicate that the predicted data fit the observed data better than the starting data at a specific frequency and shot. The results indicate that the predicted dataset is a better approximation of the observed data for 85 % of the shot-frequency combinations. This image also provides useful information about where FWI is struggling to decrease the overall misfit. In Figure 3.8B, the relative misfit reduction for each OBS is weighted by normalizing its number of first break picks by the OBS with the most first break picks. Doing so highlights which shots are most influential in updating the model parameters, and how well these shots are fit.

The best recovered station is OBS 16, shown in Figure 3.9. As shown in Figure 3.8, OBS 16 is well recovered across all frequencies. This is evident by comparing the predicted data in Figure 3.9B to the observed data in Figure 3.9C. A wide-angle reflection highlighted in both the predicted and observed data further supports the quality of this fit. However, to match the observed data FWI greatly reduces the amplitudes in the predicted data compared to those in starting dataset shown in Figure 3.9A. This amplitude reduction is required to match the observed data, but the concern is that FWI is compensating for a bad attenuation model by introducing spurious velocity perturbations that do not reflect the real geology.

The two stations that cause the most concern in Figure 3.8 are OBS 9 and OBS 11. The results from OBS 11 are due to the station recording minimal low-frequency signal. It is a German OBS, and in general, the GSC stations record better low-frequency signal than the German stations. The German stations are OBS 8, 9, 10, 11, 12, and the remainder are GSC. The reduction in misfit observed in Figure 3.8 is then reasonable given its lack of low frequencies. However, recalling Figure 3.6B, there is likely a vertically oriented spurious model update directly beneath OBS 11.

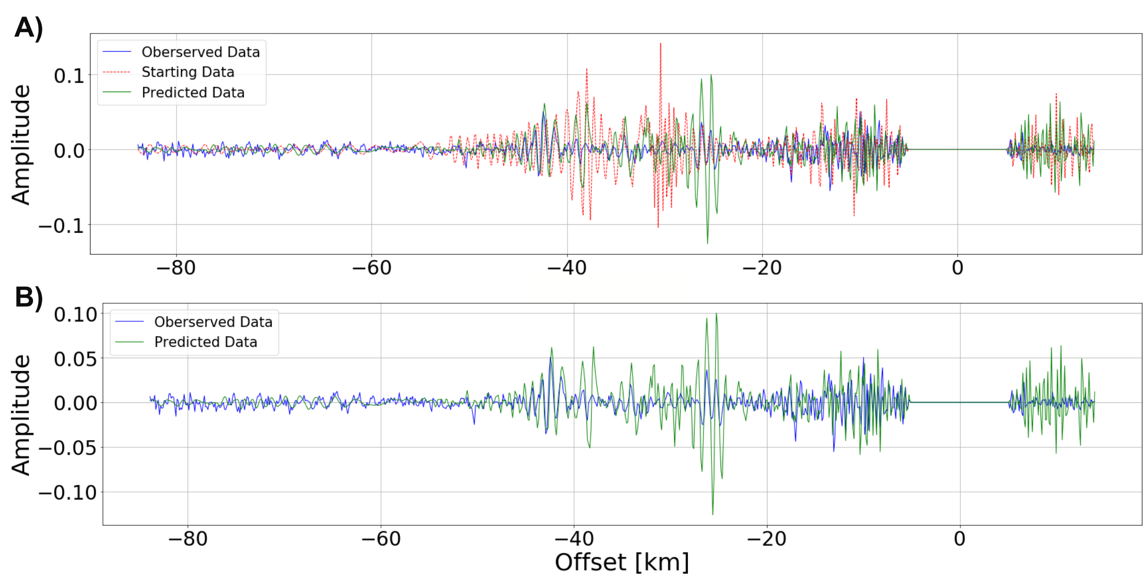

Figure 3.7: (A) The frequency-domain traces for the observed, predicted, and starting data for OBS 20 at 5.0 Hz, in (B) only the observed and predicted traces for OBS 20 at 5 Hz.

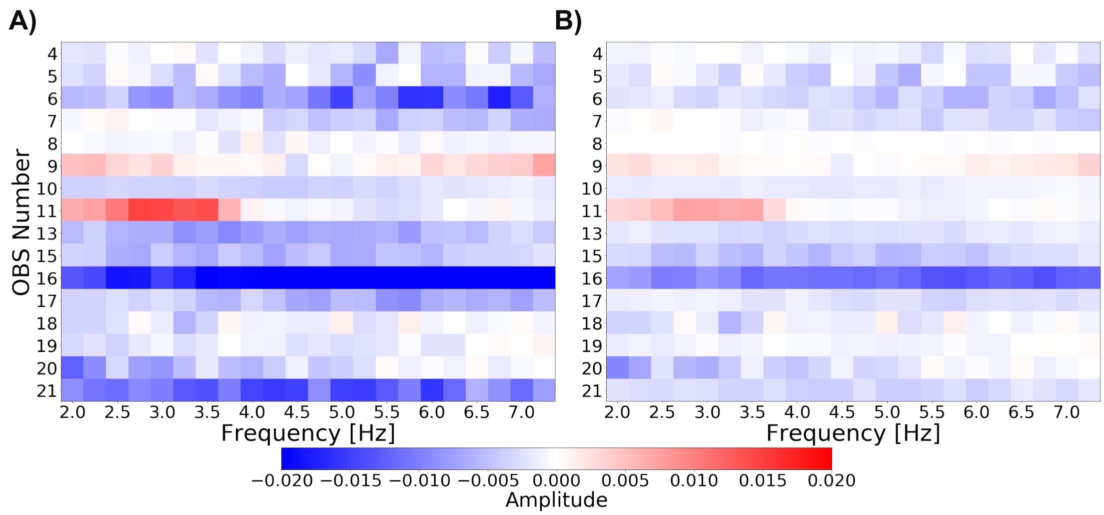

Figure 3.8: An assessment of the FWI predicted data against the starting model. Negative values indicate that the FWI model better matches the observed dataset at a particular shot-frequency combination than the starting model. Red values suggest that the starting model better models the observed dataset than the FWI velocity model at a particular shot-frequency combination. (A) The difference, and (B) the weighted difference with the weight based on the number of first break picks for that particular OBS is applied.

The inversion likely fit low-frequency noise to generate this artifact, and it is not subsequently able to correct this update.

We further investigate OBS 9 in the time-domain following forward modelling in the frequency-domain and inverse Fourier transform. Time-domain data modelled through the starting model are shown in Figure 3.10A, the data are modelled through the final FWI model in Figure 3.10B, and the observed dataset is shown in Figure 3.10C. The observed data are band-pass filtered between 2.0 and 7.25 Hz, the frequency bandwidth we use for FWI. Observations from Figure 3.10 include that the FWI model is better at reproducing the complex waveforms punctuated by shingling (Mereu et al., 1990) and wide-angle reflections. The observed dataset also contains low-frequency noise, which may explain why the relative misfit reduction appears so poor at short offsets but does not explain why it is poor at higher frequencies. Looking through the frequency domain traces for OBS 9, it turns out that the best signal is present from 4-5 Hz, and both the low and high frequencies are noisy with minimal seismic signal. A complete presentation of the FWI results for all OBS is provided by Williams (2020).

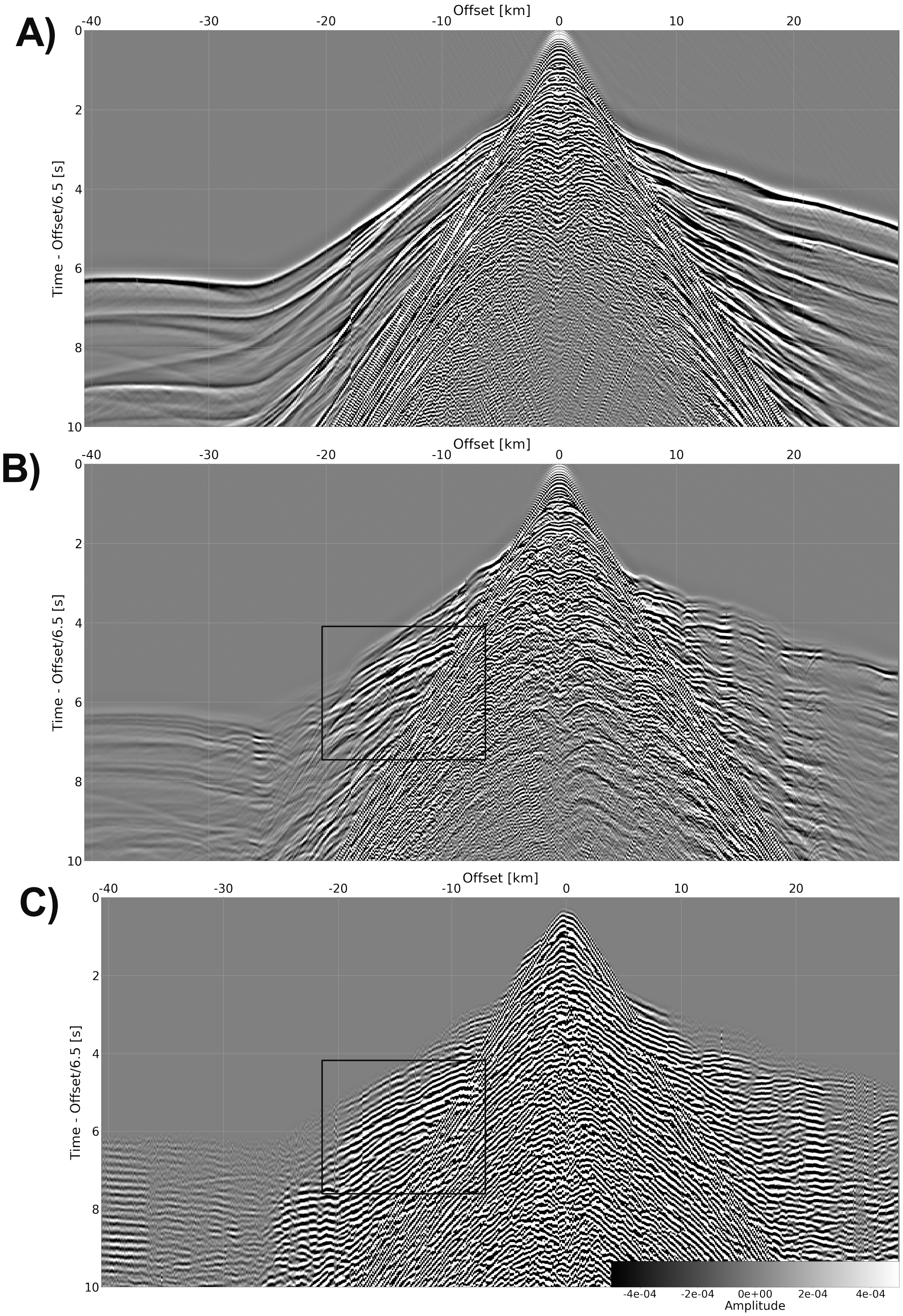

Figure 3.9: For OBS 16, (A) data modelled in the initial model, (B) data modelled in the final FWI model and, (C) the observed seismic data. All data are filtered between 2 and 7.25 Hz as these are the frequencies used in the inversion.

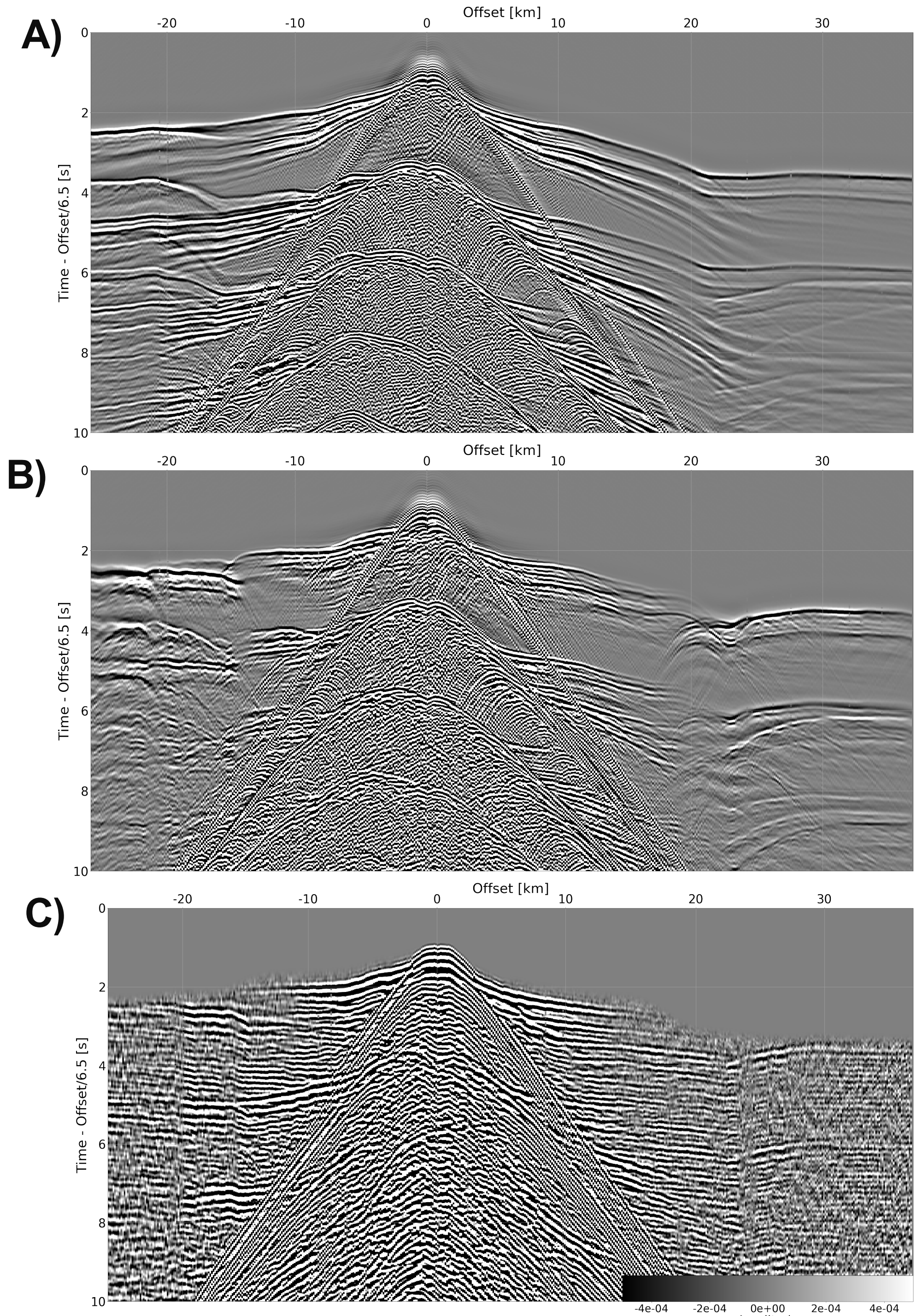

Figure 3.10: For OBS 9, (A) data modelled in the initial model, (B) data modelled in the final FWI model and, (C) the observed seismic data. All data are filtered between 2 and 7.25 Hz as these are the frequencies used in the inversion.

3.7 Geologic Interpretation

While data sparsity adversely affects the recovered FWI model, overall, the inversion fits the seismic data much better than a tomographic model. The FWI model in Figure 3.6 is more detailed than the starting tomographic model in Figure 3.5, and more geological information may be extracted from the seismic dataset. A geological interpretation of the FWI model is constrained by other recent WARR studies (Welford et al., 2015a,b; Feld et al., 2017) and recent seismic reflection studies (Klimke & Ehrhardt, 2014; Montadert et al., 2014; Reiche et al., 2016; Reiche & Hübscher, 2015; Symeou et al., 2018). Additionally, we leverage a near-parallel seismic reflection line for this interpretation to further constrain what the recovered subsurface anomalies may correspond to (Reiche et al., 2016).

Figure 3.11A shows a coincident seismic reflection profile which we convert to depth using the FWI velocity. The reflection profile shows a high amplitude reflection between 100 and 120 km at approximately 3 km depth. This reflection is the top of a sizeable autochthonous salt wedge thickening northward toward the zone of tectonic convergence (Reiche et al., 2016). Beneath this feature, reflectors are dipping to the north where they become less coherent. Toward the south, these reflectors become near-parallel to the bathymetry. Immediately, to the north of the salt wedge, the reflections become chaotic before becoming coherent once again.

Overlaying the FWI velocities on top of the seismic reflections in Figure 3.11B, there is a large lateral velocity transition north of the salt wedge to lower seismic velocities. While the high-velocity evaporites punctuate this transition, there is a regional velocity transition at this location with higher seismic velocities over the ECB and lower seismic velocities over the Hecataeus Rise. The northern Hecataeus Rise generally exhibits lower seismic velocities than the southern ECB and more chaotic velocity updates. A high-velocity body is outlined in Figure 3.11B. This feature is modelled slightly further to the north by Welford et al. (2015a) (Figure 3.3A), and Feld et al. (2017) shows a high velocity anomaly in a similar location. This high-velocity feature likely corresponds to a rift-related igneous intrusion previously interpreted using magnetic data (Ben-Avraham et al., 1976).

In Figure 3.11C the detrended FWI velocity update is overlain on the reflection data. It shows a good match between a large velocity increase and the interpreted salt

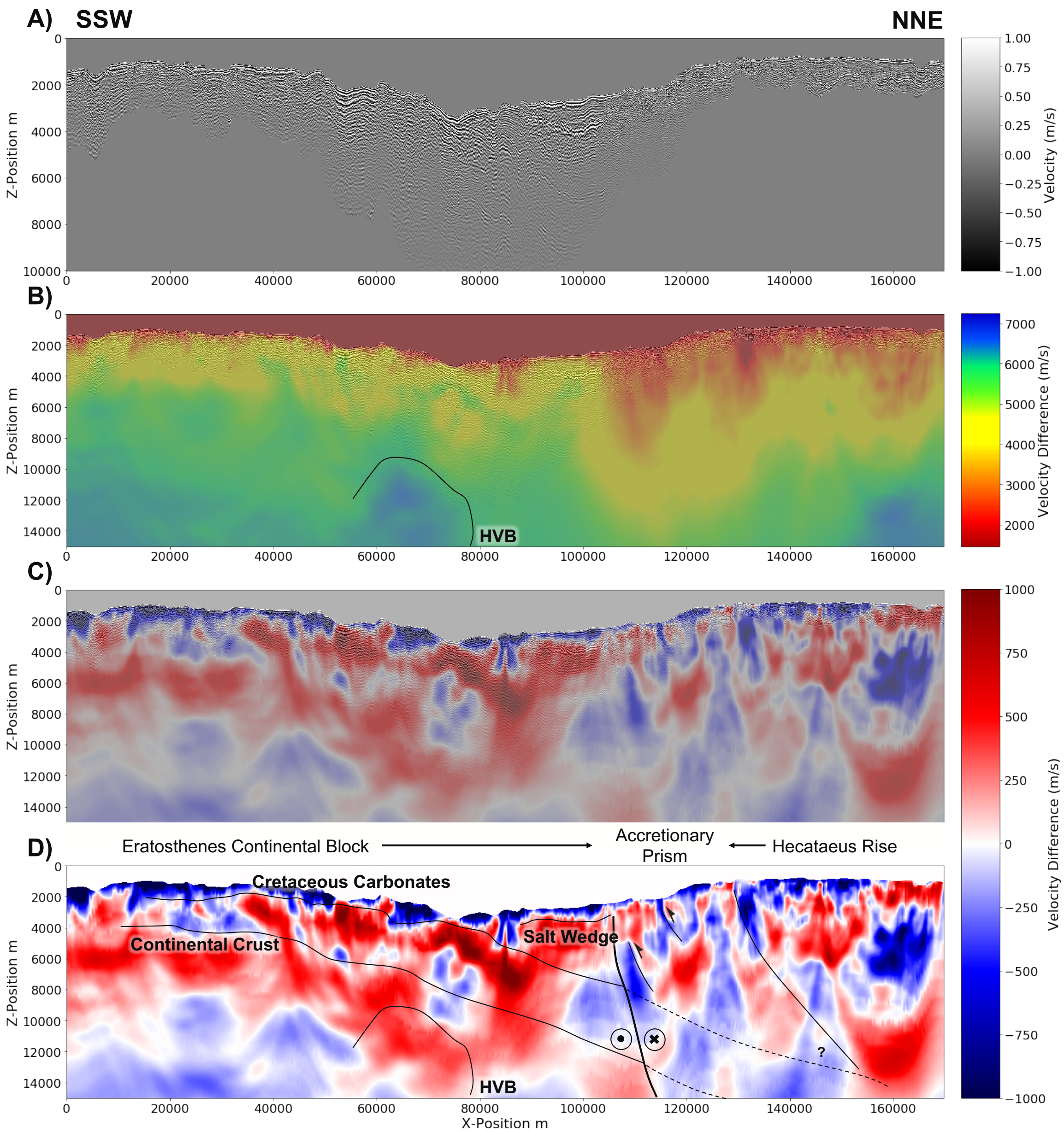

Figure 3.11: The coincident seismic reflection line from Reiche et al. (2016) is in (A). It is cropped above the waterbottom multiple, and it has a vertical exaggeration of 4.5. In (B), this reflection line overlays the FWI velocity model and it is overlain on the velocity update from the detrended background model in (C). A geologic interpretation is shown in (D). These images have a vertical exaggeration of three.

wedge near the middle of the profile. There is also a significant velocity increase that is near-parallel to the bathymetry over the ECB and progressively dips northward toward the Hecataeus Rise. This high-velocity horizon becomes lost beneath the 2-3 km thick salt wedge. It corresponds to northward dipping reflectors observed in the seismic data. This is likely a high-velocity carbonate horizon. Progressing northward, there is a change from lateral velocity updates to inclined velocity updates that steeply dip towards the north. The velocity updates then again become lateral toward the northern end of the profile.

From these observations from the FWI model and the seismic reflection data we propose three tectonic domains in the study area. In Figure 3.11D, the ECB is located to the south, and the present-day transpressional plate boundary is at approximately 120 km. We interpret an accretionary prism 30 km in width further north. Examining the FWI update from the tomographic model in Figure 3.6B, this accretionary prism may be broader. Eakin et al. (2014) interpret accretionary prisms across the Manila subduction zone in the south China Sea. The authors perform traveltime tomography to image the prism. They characterize it by low velocities between 2-4.5 km/s, and locally as low as 1.5-1.8 km/s. These velocities are in agreement with what is imaged by FWI in this study for the interpreted accretionary prism.

The Hecataeus Rise exists to the north of this accretionary prism. Continental crust from the ECB is speculatively interpreted as under-thrusting both the accretionary prism and Hecataeus Rise as high velocities continue to dip northwards from depths of 15-20 km (Figure 3.6C). However, the FWI results become uncertain at depth. The disappearance of an interpreted high-velocity carbonate horizon beneath the salt may correspond to the termination of this horizon, or be an artifact from sub-salt imaging. The high-velocity carbonate horizon is speculatively traced as under-thrusting the accretionary prism and Hecataeus Rise.

3.8 Discussion

Performing FWI on a sparse OBS dataset further reduces the number of constraints on an already ill-posed inverse problem. In order to achieve convergence, we adopt a three-loop multiscale inversion strategy from Górszczyk et al. (2017). Two synthetic inversions test this multiscale inversion strategy and encourage the feasibility of the field data inversion. The Marmousi model proves that the inversion strategy mitigates the effect of data sparsity, and the Eastern Mediterranean model proves that inverting the real-dataset is viable at the scale of the field data acquisition. For the Marmousi synthetic test, employing the multiscale inversion strategy has a similar effect on the inversion as adding more data. The synthetic Eastern Mediterranean inversion proves that the field data inversion is viable, as its geometry produced the synthetic dataset.

Given the success of these synthetic data inversions, the same multiscale approach of Górszczyk et al. (2017) inverts the Eastern Mediterranean dataset. The inversion parameters are noticeably sensitive for this field data inversion. Therefore, we complete a parametric study before finding a set of inversion parameters resulting in the greatest reduction of the objective function. This overall reduction in misfit corresponds to a recovered velocity model that is close enough to the global minima to be useful. Given the sparse dataset and challenging geological environment, it is unreasonable to expect a perfect inversion, but the imperfections must be acknowledged. The complex geological environment and sparse dataset produced a mismatch in the traveltime tomography in shots at approximately x=150 km (see Figure 3.5B). This feature is problematic for FWI as it is subject to cycle-skipping, but it is persistent in the tomographic model following many inversion parameters and first break pick adjustments. This feature is likely a high-velocity carbonate or salt body that traveltime tomography with a sparse OBS dataset is unable to resolve. We cannot preclude anisotropy contributing to this mismatch as isotropic media are assumed. Recovering the field data amplitudes are additionally problematic, especially at far offsets. We disregard elastic phenomena such as P-S converted waves due to our use of the acoustic approximation. Without moving to the elastic formulation of the wave equation, a more realistic attenuation model may improve the fit of the modelled seismic amplitudes.

Excluding wide-angle reflections from traveltime tomography to build the starting velocity poorly constrains the deep crustal structure. Welford et al. (2015a) includes these wide-angle reflections in their model, and is able to constrain the deep crustal structure by including these reflections and an additional gravity forward model. Kamei et al. (2012) and Davy et al. (2017) both present their final interpretations above 15 km depth, as does this study. Operto et al. (2006) and Górszczyk et al. (2017) present final models down to 25 km depth, but their interpretations generally lie above 20 km. FWI provides a high resolution image for the shallow upper crust, but it is perhaps better to forward model wide-angle reflections if the goal of a study is to image from 15-30 km depth.

Cognizant of model uncertainties, a geological interpretation reveals three tectonic domains (Figure 3.11). To high-velocity horizons over the ECB dip to the north beneath an interpreted accretionary prism. We interpret these horizons to underthrust the Hecataeus Rise due to the general northward motion of the African plate with respect to the Eurasian place since the Upper Cretaceous (Robertson, 1998; van Hinsbergen et al., 2020). However, convergence transitioned to transpression since the Pliocene, and it becomes difficult to interpret the high-velocity Horizons beneath the high-velocity salt wedge. We then interpret the present-day plate boundary between the African and Eurasian plates at approximately 110 km in our model. The interpretation of an accretionary prism south of the Hecataeus rise suggests that shortening occurs along the present-day Cyprus Arc, but does not preclude subduction elsewhere.

3.9 Conclusions

This study re-investigated a WARR dataset acquired from the Eratosthenes Continental Block (ECB) in the south, across the Cyprus Arc, and to the Hecataeus Rise in the north. Previously, a combination of tomographic and forward modelling techniques constructed a crustal-scale velocity model with these data (Welford et al., 2015a). Now, FWI methodologies are applied to this sparse crustal-scale OBS dataset using open-source software. The sparse Eastern Mediterranean OBS dataset with an average OBS spacing over 10 km represents the sparsest application of FWI to a real OBS dataset to our knowledge, and it additionally is completed over a complex geological environment. Examination of the predicted frequency domain data exhibits an acceptable decrease in the objective function, exemplified through the examination of predicted time-domain shot gathers. The main conclusions from this study are as follows.

- The multiscale FWI approach of Górszczyk et al. (2017) is effective at minimizing the impact of sparse-data inversions, both for the Marmousi synthetic test and a full-scale Eastern Mediterranean synthetic test.
- The multiscale FWI approach of Górszczyk et al. (2017) is also able to minimize the objective function for a real-data inversion adequately. This real-data inversion uses a sparse OBS dataset acquired in a geologically complex region.

- There is a concern that artificial velocity perturbations exist in the FWI velocity model due to an oversimplified attenuation model. A better attenuation model should alleviate this concern. To generate a better attenuation model, the amplitude decay method of Tonn (1991) may prove useful for WARR data, and joint inversion may be useful as well (Malinowski et al., 2011).

- The recovered FWI velocity model is of higher complexity and theoretically provides a higher resolution image for the upper 10-15 km than the forward model from Welford et al. (2015a). However, the FWI velocity model poorly constrains the deeper crustal structure.

- A geological interpretation reveals three tectonic domains, the southern ECB, an accretionary prism, and the Hecataeus Rise to the north. The plate boundary between the African and Eurasian plates exists between the ECB and the accretionary prism, where a large lateral velocity juxtaposition is a result of the present-day transpressional regime.

Although FWI is challenging to implement for field-data, the ability to achieve convergence with a sparse OBS dataset in a complex geologic environment should warrant future investigations in simpler geologic environments. This study has shown that reworking an OBS dataset with FWI may lead to additional geologic interpretations. Seismic datasets are expensive and challenging to acquire, meaning geophysicists should extract as much information from them as possible. Any future wide-angle surveys should consider the possibility of FWI investigations following initial tomographic and forward modelling analysis. Therefore, we recommend that these surveys are acquired with a dense OBS array that can record low frequencies. Additionally, we favour the generation of tomographic models rather than forward modelling, as tomographic models are good starting models for FWI. FWI presents an exciting opportunity to revisit some WARR datasets.

Chapter 4

Summary

4.1 Introduction

This thesis applies the challenging FWI technique to a sparse OBS dataset from the Eastern Mediterranean. FWI is a complex geophysical imaging technique described in sections 1.3 and 1.4. Furthermore, the Eastern Mediterranean is geologically complex as discussed in section 1.2. The synthetic and real-data FWI results are assessed further in this summary, drawing upon synthetic FWI examples provided in Appendix C. A further assessment of the real data results considers additional presentations of time-domain and frequency-domain assessments of the results in Appendix D. Due to the limited amount of data, all 16 of the modelled and observed shot gathers are in Appendix D. We refer to this as a full-disclosure assessment of the results as we cannot bias the reader's opinion on the inversion result through a selective presentation. In this study, the amount of data allows for this, but we acknowledge that this is impractical in most situations. Suggestions to improve the final FWI velocity model from this study proceed with a full-disclosure assessment of the results. This thesis concludes by revisiting the four research goals defined in section 1.1.

4.2 Analysis of FWI Results

Analyses of the FWI results from Chapter 3 are broken down as the synthetic FWI results and the real-data FWI results. The reader is encouraged to read Appendix C for more information on the additional synthetic FWI studies presented in this thesis. Supplementary information for the real-data inversions are provided in Appendix D, but this primarily consists of inversion parameters and figures with little discussion.

4.2.1 Synthetic FWI Results

In subsection C.2 an additional FWI test using the Marmousi model supplements the results in subsection 3.5.1. Deconstructing the multiscale inversion strategy in subsection C.2 provides insight into each loop of the multiscale inversion. Three different datasets are inverted, a dense dataset with 144 sources, a sparse dataset with 18 sources, and an ultra-sparse dataset with only two sources. The multiscale FWI strategy of Górszczyk et al. (2017) is broken up in to four components, with a baseline inversion only progressing over frequencies like that in subsection 3.5.1. However, in subsection C.2, an overlapping frequency strategy is used (Brossier et al., 2009). The second inversion adds the inner offset loop and the optimization technique changes from l-BFGS to preconditioned l-BFGS. Doing so results in a vast improvement of the dense, sparse, and ultra-sparse synthetic datasets. The next step replaces the overlapping frequency strategy with a quasi-progressive frequency approach that adds more frequencies to the inversion while retaining lower frequencies (Górszczyk et al., 2017). The addition of this step results in further improvements of the recovered FWI models for all datasets. The final step in the multiscale FWI approach is to implement the Laplace-Fourier domain waveform inversion strategy for a full three-loop multiscale inversion (Górszczyk et al., 2017). However, only the ultra-sparse dataset improves by doing so. The velocity models of both the dense and sparse datasets have regressed in comparison to the previous results.

Two factors explain these results, one being the scale of the Marmousi inversion, and the other is related to the inverse crime being committed. The inverse crime is using the same forward modelling technique to perform the inversion and generate the observed data. The same visco-acoustic forward modelling algorithm forward models the observed dataset and performs forward modelling in the context of FWI. As a result the model may replicate all waveforms in the observed dataset. By performing Laplace-Fourier domain FWI, portions of the data are omitted from FWI by design, and the observed dataset gets smaller. The recovered FWI model then worsens as we are constraining it with less data. Thus the inverse crime is one explanation for why the full and sparse data inversions have regressed during Laplace-Fourier domain waveform inversion overall.

At the Marmousi scale, the recovery of near offset reflections is crucial as they reconstruct the short-wavelength structure. Diving waves recover long-wavelength structure and damping the data from the first arrival causes FWI to preferentially reconstruct the diving waves. This is detrimental at the Marmousi scale, but beneficial at the crustal-scale. Far-offset reflections recover short-wavelength structure at the crustal-scale, which is not attenuated as much as near offset reflections through the Laplace-Fourier approach.

The multiscale inversion result shown in Figure 3.2B is not the best inversion result possible. However, it proves the point that a multiscale FWI strategy adequately accounts for artifacts produced by FWI with a sparse dataset. These results furthermore suggest that Laplace-Fourier domain waveform inversion may not be necessary at smaller scales of investigation with a reasonably dense dataset. A quasi-progressive frequency strategy inverted by the preconditioned l-BFGS algorithm and a multiscale offset strategy are sufficient in that case. However, for real data investigations where there are complex elastic waveforms, these should be attenuated through the Laplace-Fourier approach when using an acoustic approximation of the wave equation.

At the crustal-scale, the synthetic Eastern Mediterranean inversion does not acceptably converge without the application of Laplace-Fourier domain waveform inversion. There is a pitfall with Laplace-Fourier domain waveform inversion at the crustal-scale, however, and that is that the application of damping absolutely must be weighted from the first breaks. TOY2DAC implements this weighting to both the observed and predicted datasets within the algorithm.

Figure C.7 shows a FWI result for the Eastern Mediterranean model using the same inversion parameters as the result from Figure 3.4, but without damping weighted from the first breaks. While damping from $t = 0$, the contribution of seismic arrivals at later times to the objective function is damped exponentially (Equation 1.8). The result is that the inversion only recovers the shallowest subsurface velocity structure.

Following from the introduction of the optimization techniques provided by the SEISCOPE Optimization toolbox in subsection 1.4.6, subsection C.3.2 shows the results for FWI using the preconditioned steepest descent and NLCG algorithms. We use the same inversion parameters as the preconditioned l-BFGS inversion in Figure

3.4. A preconditioned steepest-descent algorithm produces the velocity model in Figure C.8, and the result from a preconditioned NLCG algorithm is in Figure C.9. In subsection 1.4.6, the truncated Newton and Gauss-Newton methods are introduced as promising optimization routines that better approximate the Hessian matrix. However, in running the preconditioned version of both algorithms for the synthetic Eastern Mediterranean model, we cannot obtain an acceptable recovered model devoid of high-velocity artifacts without deviating too far from the multiscale FWI strategy.

The preconditioned l-BFGS algorithm outperforms both the preconditioned steepest-descent and NLCG algorithms (subsection C.3.1). The l-BFGS algorithm requires fewer iterations and is quicker than the other gradient-based approaches. The NLCG algorithm is slightly better than the steepest descent algorithm in regards to computational time but requires more iterations. However, the NLCG algorithm provides a much better approximation of the true Eastern Mediterranean model than the steepest descent algorithm. The preconditioned truncated Newton and Gauss-Newton methods recover the general structure of the true model but contain many near-surface high-velocity artifacts. Furthermore, the truncated Newton and Gauss-Newton methods require much more computational time to converge to a minimum. These results convince us that the preconditioned l-BFGS algorithm can most effectively find a minimum in the real-data Eastern Mediterranean inversion.

The difficulty in obtaining convergence from the truncated Newton and Gauss-Newton optimization routines is likely a result of the implementation of a multiscale inversion strategy. In this inversion strategy, the input data are always changing. Because both truncated methods are a better estimation of the Hessian matrix, they may be more sensitive to constantly altering the input dataset than the l-BFGS technique. No tests are completed with the truncated Newton or Gauss-Newton methods without aspects of the multiscale inversion strategy as it is required to deal with the sparsity of the Eastern Mediterranean dataset.

4.2.2 Real FWI Results

Due to the sparsity of the OBS dataset, we model OBS gathers both the FWI and starting velocity models at each OBS. Both modelled OBS gathers may be shown with the observed OBS gather and reasonably presented in this thesis. Furthermore, frequency domain traces for the modelled, predicted, and starting datasets are provided at 5.0 Hz for each OBS other than OBS 20. This information lies within Appendix D, and supplements the FWI model assessment in subsection 3.6.4.

Figure D.1 shows that the final FWI model converged toward a minimum. Paired with the frequency-shot analysis shown in Figure 3.8, FWI has undoubtedly converged in this study. The recovered FWI model better fits the observed seismic data than the starting tomographic model. However, as shown in Figure D.1, the bulk of the cost function reduction occurs within the first 100 iterations. These earlier iterations occur at lower frequencies where there is a lower signal to noise ratio in the data (subsection A.3.5). This observation would cause more concern should the spectrum have been whitened during data processing, as doing so would increase the probability of fitting noise through FWI. However, the spectrum is processed to be smooth and curved with its apex aligned with the peak low-frequency signal in the data. There is a relative increase in the total cost function at approximately iteration 150, followed by a subsequent decrease. Perhaps at higher frequencies, FWI is adjusting some previous model updates at lower frequencies that were fitting the noise. Doing so is possible because the contribution of higher frequencies to the objective function is greater than lower frequencies due to the processing strategy for the amplitude spectra. This data-processing step is discussed in subsection A.3.3.

As discussed in subsection 3.6.4, some OBS have a poor seismic signal at low frequencies, for example, OBS 11. OBS 11 also generates a spurious velocity update immediately beneath it. We try removing OBS 11 from the dataset, and also reducing its influence by dividing its amplitude by an order of magnitude. Neither technique yielded a superior result than the final FWI model presented in Chapter 3, but they both reduced the presence of the anomalous velocity update beneath OBS 11.

Figure D.2 shows the observed and predicted data traces for OBS 4-7 at 5.0 Hz. The observed data at near offsets within 10 km are much more complex than at 10-50 km offsets, and appear to be fit worse. This poor fit corresponds to unmuted complex seismic arrivals within the nearest offsets that viscoacoustic FWI may not replicate. These problematic near offset arrivals include the direct wave, which generates uncertain near-surface velocity perturbations if included. These near-surface velocity perturbations are likely the result of bathymetric features outside of the 2D plane. OBS 4-7 all fit the observed data better than the starting data at 5.0 Hz, recalling Figure 3.8.

The observed and predicted traces in Figure D.3 show OBS 8-11 at 5.0 Hz. The quality of the data fit is variable dependent on the consideration of offsets. All four instruments are German, and typically record a poorer seismic signal than the GSC instruments. The lowest quality OBS that remains in the FWI dataset is OBS 8, with only data observed to a maximum of 10 km offset. There is still a lack of signal for the far negative offsets of OBS 11 at 5.0 Hz. OBS 13, 15, 16, and 17 are shown at 5.0 Hz in Figure D.4 and are all superior fits to the observed dataset than the starting data, supported by Figure 3.8. OBS 13 is a German instrument, but OBS 15-21 are GSC and record higher quality signals at far offsets. In particular, OBS 15 in Figure D.4B records data up to approximately -80 km offset. An analysis of the data fit from near to far offsets for this OBS raises concerns. At near offsets, the predicted data are of higher amplitude than the observed data, and at far offsets, the predicted data have lower amplitudes than the observed. Overall the fit is acceptable, but the amplitude discrepancy is troubling. Not properly modelling the seismic amplitudes at far offsets results in unconstrained deeper portions of the FWI model. Decreasing the subsurface attenuation by increasing Q to 100 helped to some degree, but then the near-offset amplitudes become grossly overestimated. A more complex attenuation model with a high degree of near-surface attenuation and a low degree of deep attenuation is likely required to fix these amplitude mis-matches.

OBS 18, 19, and 21 modelled at 5.0 Hz are shown in Figure D.5. OBS 20, shown in Figure D.5C is modelled at 3.0 Hz as Figure 3.7 shows it at 5.0 Hz. Like OBS 15, there is a large amplitude discrepancy at far offsets for OBS 18. However, OBS 20 appears to be noisier across all offsets as lower frequencies generally contain a poorer seismic signal. The amplitude match at far offsets for OBS 18 is abysmal between 70-100 km offsets. There is seismic signal in the observed data at these far offsets, and although noisy, it would certainly allow for a better constraint on the deep velocity structure. At 20 km offset, OBS 18 exhibits a unique anomaly in the predicted data. This consistently low amplitude feature is present in the observed data but to a lesser degree.

Examining these frequency domain traces suggested that the data are generally fit to a satisfactory degree, however, the observed data may be noisy at certain frequencies and offsets. The time-domain gathers for the observed dataset are now compared to the forward modelled starting and predicted data. To mitigate the length of this discussion, the four OBS gathers with first break picks at the furthest offsets are

QC'd, but all other OBS gathers are in section D.3 with the exception of OBS 9 and 16, shown in subsection 3.6.4. OBS 5 in Figure D.10 contains signal in the observed dataset up to approximately 70 km offset. There appears to be a spurious artifact beyond an offset of -20. However, the velocity model has poorly constrained model updates in this region. From 10-20 km offset, there is an evident recovery of a reflected arrival just beneath a reduction time of 2 seconds. At a reduced time of approximately 3 seconds just beyond 20 km offset, several wide-angle reflections are partially recovered by FWI, one being a multiple. Beyond 30 km offset, the amplitudes of the observed and predicted datasets decrease, and at offsets greater than 50 km, it appears that the match is poor in the time domain.

OBS 15 in Figure D.14C shows a complex observed dataset. The predicted data in Figure D.14B are fitting the overall structure of the data, but poorly in places. Beyond offsets of -40 km, the data are poorly fit principally due to a high degree of attenuation in the modelled dataset. However, the amplitudes within the first 25 km are in general agreement with the observed dataset. An interpreted low amplitude, far-offset reflection at -40 km is well-recovered, as is the complex structure of the modelled data between -40 and -20 km offset. However, there is a poor recovery of high amplitude reflections at -35 km. Several near offset reflections between -20 and 20 km within the predicted data arguably correspond to features in the observed dataset.

For OBS 18 and 20 in Figures D.16 and D.18, respectfully, there an artifact in the time-domain data that must be addressed. Recalling the starting model assessment in Figure 3.5B, at approximately x=150 km, traveltime tomography is unable to reproduce the first breaks well enough to mitigate cycle-skipping. Recalling subsection B.2.2, there is a misfit between the observed and predicted data consistently at this location for each iteration of the first break picks for OBS 15-20. This consistent misfit is partially responsible for the many iterations of first break picks. The artifact is prominent in the starting shot-gathers for OBS 18 and 20, and even more so in the predicted shot-gathers. No such feature exists in the observed dataset, and it is an artifact produced by traveltime tomography.

4.3 Suggested Improvements for Real-Data FWI

Suggested improvements toward data processing, starting model generation, and FWI implementation may improve the recovered FWI velocity model shown in Chapter 3. Resolutions for several issues alluded to in subsection 4.2.2 are provided here.

4.3.1 Improving Seismic Data Processing

The seismic data processing workflow used for this FWI study is fully presented in Appendix A. Changes to four aspects of this data processing workflow may immediately improve the FWI result. They concern the removal of OBS 12 and 14, the omission of coherency filtering, the estimation of the time-domain wavelet for filtering, and the amplitude bulk shift.

This data processing flow emphasizes the importance of preserving the real-data amplitudes to ensure that they match the predicted data amplitudes. As evidenced by Figure 4.2.2, this endeavour is ultimately unsuccessful. In attempting to match the observed and predicted data amplitudes as much as possible, this study removed two OBS and omitted coherency filtering. It is impossible to achieve a perfect match between the observed and predicted data given the use of an acoustic approximation to the wave equation in what is really an elastic media (Mora, 1987). Seeing how poor the fit is between the observed and predicted data amplitudes, despite the cautious approach taken during data processing, is discouraging. In hindsight, it may be more beneficial to include the two hydrophone OBS (OBS 12 and 14) excluded due to the absence of a vertical component geophone. To correct for the hydrophones use, we would then apply an amplitude correction derived by amplitude differences between vertical component geophones and hydrophones from good quality hydrophone and geophone data. However, performing traveltime tomography with OBS 12 and 14 included in the dataset did generate spurious velocity perturbations in the subsurface (subsection B.2.2). The benefit of introducing these omitted OBS into FWI is then uncertain.

A similar argument can be made toward the omission of a coherency filter, discussed in subsection A.3.6. If applied to the observed dataset, it may alter AVO relationships in a manner that acoustic wavefield modelling does not predict. Despite this disadvantage, a coherency filter will increase the signal to noise ratio of the observed dataset at far offsets, which would be advantageous. Moreover, the predicted data are already incorrectly replicating the amplitudes of the observed dataset under the acoustic approximation. The benefit of adding a coherency filter is possibly more influential than the perceived downside.

As discussed in subsection A.3.3, to generate an estimate of the time-domain wavelet, only the nearest offset traces are extracted from the data. Due to seismic data being non-stationary, the wavelet should change with offset (Yilmaz, 2001). An offset dependent convolution estimates the time-domain source wavelet with autocorrelations at multiple offsets. An offset dependent filter will then produce a superior result in smoothing the amplitude spectra and removing reverberations in the autocorrelations at far offsets.

The amplitude bulk shift that matches the observed data amplitudes to the modelled data amplitudes may be improved as well. Trace RMS amplitudes of the observed data are inherently more significant than those of the modelled data, as discussed in subsection A.4.2. The observed data RMS amplitudes are calculated based on the entire seismic trace, but they should only consider a window just beneath the first arrival. Because we are applying FWI in the Laplace-Fourier domain, damping of the seismic wavefield at later times renders these data irrelevant. Moreover, the observed dataset is noisy, resulting in larger trace RMS amplitudes. A better match in the amplitudes of the observed and predicted data would result from only computing RMS amplitudes in a window beneath the first arrivals.

4.3.2 Improving the Starting FWI Model

The two most prominent issues with the real data FWI results are related to the starting FWI model. As discussed in subsection 4.2.2, the first issue is the mismatch in amplitudes at far offsets. The second issue is the inability of traveltime tomography to recover high-velocity horizons generating localized inaccuracies. This subsection discusses improvements to the starting model to address these issues. A better density model will contribute to a better estimation of the real seismic amplitudes, but a superior attenuation model will likely be more influential. Discussion regarding the starting velocity model addresses an artifact present in the forward modelled waveforms in the final FWI model.

While various aspects of the data processing workflow may be adjusted to better match the observed and predicted data amplitudes, the primary control is the attenuation model. This study consistently observes a poor match in amplitude between the observed and predicted datasets at far offsets, and an acceptable match at near offsets. This observation suggests that the subsurface attenuation structure is more complicated than predicted, which is hardly a surprise as we use a homogeneous attenuation value to represent the entire subsurface. Given the geological complexity of the Eastern Mediterranean, this simplistic attenuation model is insufficient. There are effectively three options available to improve this result, as alluded to in subsection 3.9. The first is to build a more accurate attenuation model. An attenuation gradient may be sufficient, where Q increases with depth. Even though this approach is still wrong, it should be less-wrong than the current approach and result in better amplitude estimations of the predicted data. Secondly, Tonn (1991) provides ten different methods for computing Q from the seismic data. One that may apply to crustal-scale seismic refraction data is the amplitude decay method. It involves computing the ratio of amplitudes for two different distance or time measurements. The third option is to perform a joint inversion for velocity and attenuation. Such an inversion has not been attempted at the crustal-scale to the best of our knowledge, but is applied in near-offset studies Malinowski et al. (2011). Although a joint inversion for velocity and attenuation may seem, doing this for the sparse OBS dataset would result in this already challenging inversion becoming even more difficult.

We use Gardner's relation Gardner et al. (1974) to estimate the subsurface densities from the velocities. This is a good approximation of subsurface densities in sediments, but poor outside of sedimentary basins. The interpreted section in figure 3.11D additionally reveals the presence of high-velocity carbonates and evaporites. Gardner's relationship will break down for these sedimentary rocks. The inadequate representation of true rock densities, especially at depth, may also contribute to far-offset amplitude mismatches between the observed and predicted data, but to a lesser degree than the simplistic attenuation model. An improved density model should use a velocity-density relationship better suited for the entire region of investigation. The Nafe-Drake relation (Ludwig et al., 1970) and Brocher's regressive fit (Brocher, 2005) are two alternate empirical velocity-density relationships that formulate a regressive velocity-density relationship by considering more than just sedimentary rocks. However, both relationships will likely fail to account for the densities of the carbonates and the evaporites. An alternate solution would be a joint density-velocity inversion using TOMO2D (Korenaga et al., 2001). Such an inversion would simultaneously generate the starting velocity and density models. It is unclear if this joint inversion would produce a better starting velocity model, or hinder the generation of a velocity model.

Finally, the starting velocity model must be improved. Along with improving the attenuation model, this is the most critical portion of the workflow to revisit. Multiple iterations of first break picking yielded a misfit in observed and predicted data near shot x=180 km, and this large misfit has propagated into the predicted FWI dataset. An analysis in subsection B.2.3 reveals that no single subsurface feature is responsible for this misfit. However, as exhibited by the forward modelled first break picks for OBS 10 in Figure B.10, this mismatch is an overestimation of travel time as it arrives at a later time than the picked first break. It is then likely that this inaccuracy in the starting model corresponds to high-velocity evaporites or carbonates that are not adequately recovered by traveltime tomography. A solution would be to manually place a high-velocity body into the starting model to remedy this artifact. Figure B.12A, which shows the ray paths of arrivals that do not satisfy the cycle skipping criteria in Equation 3.2, provides insight into the locations of these absent high-velocity bodies. Over the Hecataeus Rise, Reiche & Hübscher (2015) interpret small isolated evaporite basins, which are likely what the rays near x=170 correspond to. Revisiting traveltime tomography is required with fixed-high velocity bodies in locations where high-velocity salt or carbonates are interpreted. Of all the proposed solutions to improve the FWI results in this study, this is likely the most difficult to implement, but arguably the most critical.

4.3.3 Improving the FWI Inversion Strategy

As demonstrated through synthetic testing and the ability of FWI to achieve convergence with a sparse dataset in a geologically complex region, we have shown the multiscale inversion strategy of Górszczyk et al. (2017) to be robust. However, given the results, some minor parametric adjustments could be helpful. Additional justifications for the isotropic and acoustic assumptions follow.

Two parametric configurations from the multiscale approach that warrant further investigation are to the starting frequency and maximum offset. Two factors led us to choose the starting frequency of 2.0 Hz. As discussed in subsection A.3.5, it is the lowest frequency where a seismic signal is observable in the OBS dataset. FWI tests starting at 2.25 Hz and 2.5 Hz yield inferior results to those starting at 2.0 Hz. We did not attempt to push the inversion to frequencies below 2.0 Hz as the results from subsection A.3.5 suggested a significant lack of low-frequency signal there.

The maximum offset is an additional parameter that may be considered for further testing. A maximum offset that corresponds to the maximum offset of the picked first breaks is used in all FWI tests. However, given the analysis provided in subsection 4.2.2, FWI may converge better if the far offsets are omitted and the maximum offsets are constrained to 50 km for example. Cropping the maximum offset to 50 km would only result in the omission of approximately 15 % of the available data, and potentially result in better velocity perturbations in the upper 10-15 km.

Modelling anisotropy and elastic waves may additionally improve the results of this study. However, other ideas may be more effective and simplistic. With this sparse Eastern Mediterranean dataset, our preference is to keep FWI as straightforward as possible. For denser datasets resulting in more constrained inversions, it is perhaps worth considering modelling the subsurface anisotropy to improve the result. Modelling anisotropy requires knowledge of subsurface anisotropy, which is not present nor easy to obtain. Modelling elastic waves requires knowledge of a subsurface S-wave velocity structure, which is not known nor easy to obtain. The inversion will become more complex to reproduce complex phenomena in the dataset without considering other simplistic options. However, perhaps the inability of the predicted dataset to produce the complex AVO signature in the observed dataset could be mitigated by invoking the elastic wave equation than a more complex attenuation model. Alternatively, the misfit between the observed and predicted travel times in the starting model could be due to unaccounted for anisotropic effects.

4.3.4 Apriori Information During FWI

We process the dataset in the time-domain as it provides a more intuitive understanding of the seismic waveform. However, frequency-domain FWI, in its essence, attempts to match an observed frequency trace with a predicted frequency trace. In hindsight, there is little consideration toward the appearance and physical meaning of this frequency domain data. Some useful questions to ask during data processing involve what do the seismic signals and noise look like in the frequency domain, and how may one better encourage FWI to fit the signal. Formulating so-called apriori data weights for the observed data may improve the recovered FWI model. For example, if a given frequency is noisy at a particular shot, its amplitude can be decreased to ensure FWI prioritizes other shots at that frequency. Reducing the amplitude of the problematic OBS 11 explored this concept briefly, but a better solution is to decrease the amplitudes of noisy low-frequencies, and leave higher frequencies untouched. A shot with a high-quality low-frequency signal will have its amplitude increased. Apriori signal-to-noise information may be obtained through cross-correlating the frequency domain traces of the observed and starting data. The starting data are noise-free, and seismic signals in the observed dataset with a high correlation should contain less noise. FWI should then prioritize fitting those data, by implementing relative amplitude manipulation. The resultant observed dataset will not have a smooth amplitude spectrum, but its spectrum will reflect the signal to noise ratio in the observed data, therefore (in theory) improving the result of FWI.

4.4 Conclusions

This thesis employs a multiscale FWI strategy from Górszczyk et al. (2017), and open-source FWI software to invert a sparse WARR dataset from the Eastern Mediterranean. Synthetic tests using Marmousi and Eastern Mediterranean models reveal that the multiscale strategy benefits sparse data inversions and that FWI will converge to a solution at the crustal-scale of investigation. The Marmousi inversions suggest that the three-loop inversion of Górszczyk et al. (2017) may not be ideal for a smaller scale of investigation, and the Eastern Mediterranean inversion highlights the importance of weighting the damping in the Laplace-Fourier domain from the first break picks.

Despite some shortcomings discussed in subsection 4.2.2, FWI can adequately converge to a solution that better predicts the observed seismic dataset. This alone is a success not only considering the extreme data sparsity, but also the complex geological environment and multiple OBS sources (GSC and German). However, there are two significant shortcomings in the final FWI crustal-scale velocity model. The first is related to matching the observed and predicted data amplitudes at far offsets. There are sparse constraints on the deep FWI velocity perturbations as a result. Subsection 4.3 provides multiple suggestions to improve far-offset amplitude prediction, but updating the density and attenuation models is the most promising. Subsurface densities that are inaccurate in the deeper portions of the model may contribute to amplitude mismatches at far offsets. Either the Nafe-Drake (Ludwig et al., 1970) or Brocher's regressive fit (Brocher, 2005) may yield improvements. A superior attenuation model will likely provide the most significant improvement in matching the observed and predicted data amplitude trends. A simple attenuation gradient may be sufficient, or a model with high-attenuation sediments and low-attenuation crystalline rocks may use apriori information from this study and Welford et al. (2015a).

A second shortcoming for the FWI results presented in this study is a traveltime mismatch present in the starting velocity model. This misfit is propagated to the FWI model and is due to the inability of traveltime tomography to recover certain high-velocity features in the subsurface. As shown in section B.2.2, an exhaustive seven iterations of first break picking could not resolve this issue. Other techniques must be considered by future works to build the starting model, which include fixing high-velocity bodies in place during traveltime tomography. Davy et al. (2017) build their FWI starting model with a similar approach but for the entire shallow sedimentary column.

A detailed geologic interpretation for this study area is in section 1.2, but the geologic interpretation in section 3.7 is brief due to the uncertainties in the recovered FWI model. Using constraints from a coincident seismic reflection line (Reiche et al., 2016), a correlation between lateral amplitude discontinuities and the resultant FWI model down to 10-15 km depth leads to the interpretation of the geologic structure. An interpreted 30 km wide accretionary prism is present between the southern ECB and the Hecataeus Rise to the north. A carbonate horizon dips beneath a 2-3 km thick evaporite wedge, and a horizon representative of the top continental crust for the ECB is under-thrusting the accretionary prism and Hecataeus Rise.

Now, the four primary research goals from section 1.1 are revisited.

- *Investigate whether FWI is possible on such a sparse crustal-scale dataset.* FWI is possible on such a sparse dataset depending on the meaning of possible. The answer is yes if it means an overall reduction in the data misfit through FWI. However, uncertainties in the recovered FWI model are acknowledged, and they may require a substantial amount of work to address.

- *Compare a recovered FWI model with the velocity model from Welford et al. (2015a).* The final FWI velocity model presents a more detailed shallow (<15 km) reconstruction of the subsurface velocities. The velocity model from Welford et al. (2015a) better recovers a deeper (greater than 15 km) subsurface velocity structure. Both models recover a high-velocity body at approximately 10-15 km depth to the south, and both models recover a high lateral velocity juxtaposition where the interpreted plate boundary exists between the African and Eurasian plates.

- *Make any additional geologic interpretations on the recovered FWI model.* Supported by a coincident seismic reflection line, the FWI model interprets a lateral change in velocity related to a 30 km wide accretionary prism south of the Hecataeus Rise. A high-velocity horizon attributed to a thick carbonate package is dipping beneath an autochthonous salt wedge that is 2-3 km thick, and a high-velocity horizon attributed to the top of the Eratosthenes Continental Block dips northward beneath the accretionary prism and Hecataeus Rise.

- *Provide a recommendation for performing FWI on similar datasets.* For similar datasets, viscoacoustic FWI is a reliable imaging technique to recover a detailed crustal-scale velocity model above approximately 15 km depth, with a proper post-inversion assessment. However, complex Eastern Mediterranean geology produces unavoidable uncertainties in the inversion that a more simplistic geological environment may avoid. Re-visiting previously acquired crustal-scale datasets with FWI techniques will keep a lot of graduate students busy, and inverting challenging datasets will only lead to the development of better FWI techniques.

Appendix A

Seismic Data Processing

A.1 Introduction

This appendix focuses on processing ocean-bottom seismometer (OBS) data for full-waveform inversion (FWI). There are four stages of data processing, the first of which is simply a data quality control (QC) stage. The next two steps are processing stage

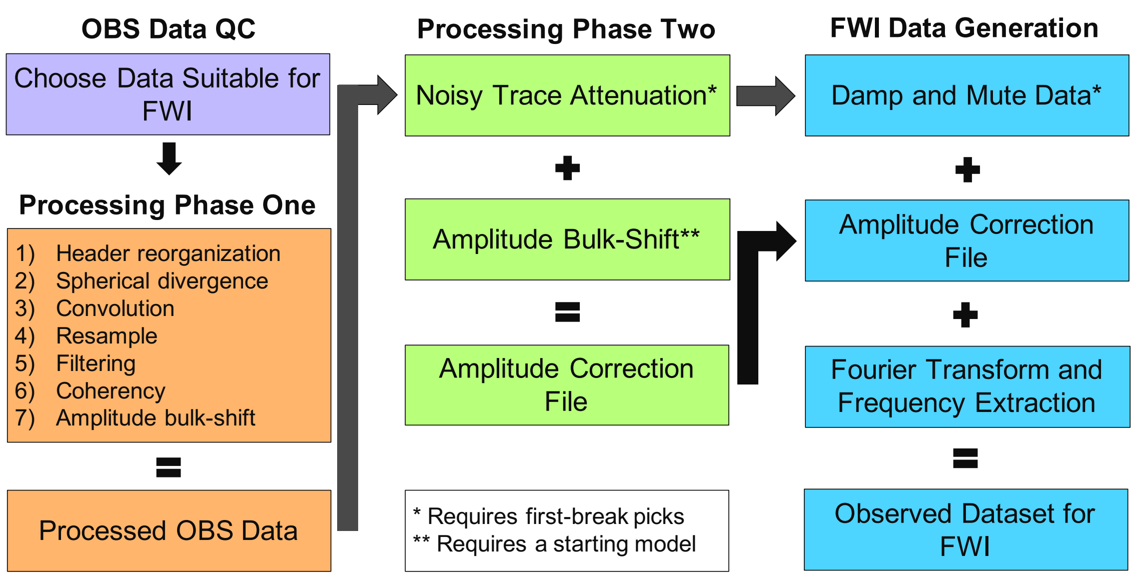

Figure A.1: The final OBS data processing workflow is shown. Each processing phase is color-coded; the QC phase is purple, phase 1 is orange, phase 2 is green, and the FWI data generation phase is blue.

1 and processing stage 2. Processing stage 1 is completed using Globe Claritas data processing software. The remainder of the data processing is completed in Python. The codes used to complete these processing steps are provided at `https://github.com/celw10/TOMO2D_2_TOY2DAC.git`. Processing stage 2 requires first break picks and a starting velocity model. Therefore a portion of the starting model workflow (discussed in Appendix B) must be completed before processing stage 2. Processing stage 1 produces SEGY data, and processing stage 2 produces an amplitude correction file. During the TOY2DAC data-generation stage, the amplitude correction file from processing stage two is applied to the processed SEGY data from processing stage 1. Similar to processing stage 2, this final data-generation stage requires first break picks. The entire OBS data processing workflow is summarized in Figure A.1.

A.2 Data QC

Eastern Mediterranean line 1 is shown in Figure A.2, where 21 OBS span approximately 220 km. The OBS data have previously been converted to SEGY, corrected for clock drift, mechanical delay of the airguns, and relocated (Welford et al., 2015a). There are three different types of OBS instruments used in this acquisition. We refer to them by their origin which is: the Geological Survey of Canada (GSC), Dalhousie University, and Germany. Figure A.2 shows the location and name of each OBS. Dalhousie instruments are located to the southwest, and the German instruments lie between two groups of GSC instruments to the north. The maximum fold, sample rate, and trace length for the vertical component geophones are in Table A.1. Also, a qualitative assessment compares the data condition for both the hydrophones and vertical component geophones. There is an observed correlation between data quality and instrument source. Dalhousie stations OBS 1 and OBS 2 are notably poor, containing minimal signal. There are no vertical component data available for German instruments OBS 12 and OBS 14 as well.

Data processing only considers the vertical component geophone data due to its superior data quality. In Figure A.3A and A.3B we show the vertical component geophone and hydrophone data for OBS 20. There is a significant difference between the OBS gathers, where the geophone has a much higher signal-to-noise ratio (SNR). Much of the noise in the hydrophone data is low-frequency noise, as the amplitude

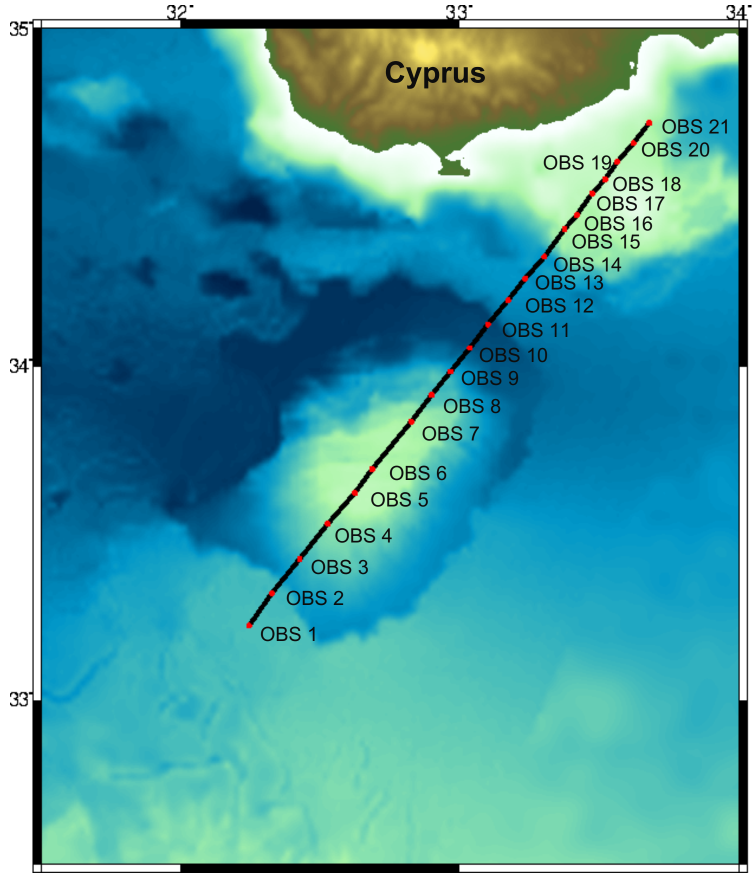

Figure A.2: Eastern Mediterranean line 1 is shown with OBS locations and names.

spectrum in Figure A.3B has a higher amplitude low-frequency component than the amplitude spectrum in Figure A.3A. We choose to omit the hydrophone data due to these observations, and vertical component geophone data are processed going forward.

German OBS 12 and 14 are omitted from the dataset as they only record pressure data (Table A.1). While both OBS are in critical positions in the model, the tomographic inversion improves following the omission of OBS 12 and 14 (Appendix B).

The southern Dalhousie stations are omitted from this study as well. OBS 2 and OBS 3 have a minimal SNR, and would only hinder FWI if included. The higher quality dataset of the two, OBS 2, is shown in Figure A.4B. Besides a faint diving wave at near offsets, data are completely dominated by noise. The omission of these two Dalhousie instruments generates a void between OBS 1 and OBS 4,

Table A.1: A description of the raw OBS data. The data are qualitatively ranked based on maximum observable first arrival time. First arrivals observable at later arrival times correlate with visible data at further offsets. The data quality is excellent with observable head waves between 15-20 s, good if observable head waves lie between 10-15 s, mediocre if head waves are observed between 5-10 s, sub-par if observable head waves lie between 1-5 s, and/or if the OBS has no usable signal.

OBS Num.	Source	Max. Fold	Sample Rate (ms)	Trace Length (ms)	Vertical Cmpt. Quality	Hydrophone Quality
1	Dalhousie	1821	3.999	59981	Excellent	Good
2	Dalhousie	1821	4.0	59996	Poor	Sub-par
3	Dalhousie	1812	3.999	59981	Poor	Sub-par
4	GSC	1813	3.999	59981	Excellent	Mediocre
5	GSC	1812	3.999	59981	Excellent	Good
6	GSC	1813	3.999	59981	Good	Good
7	GSC	1813	4.0	59996	Good	Good
8	German	1794	5.0	49995	Sub-par	Mediocre
9	German	1794	5.0	49995	Mediocre	Mediocre
10	German	1794	5.0	49995	Mediocre	Mediocre
11	German	1794	5.0	49995	Good	Mediocre
12	German	1794	20.0	49980	Absent	Sub-par
13	German	1794	5.0	49995	Mediocre	Sub-par
14	German	1794	20.0	49980	Absent	Mediocre
15	GSC	1813	4.0	59996	Excellent	Good
16	GSC	1812	4.0	59996	Excellent	Excellent
17	GSC	1813	3.999	59981	Good	Good
18	GSC	1813	3.999	59981	Excellent	Good
19	GSC	1812	3.999	59981	Excellent	Sub-par
20	GSC	1812	3.958	59983.5	Excellent	Mediocre
21	GSC	1812	3.999	59981	Excellent	Mediocre

spanning approximately 40 km. This space would add many sparsely constrained model parameters to the inversion, increasing the number of local minima in the objective function, and decreasing the reliability of the results. The region of primary geological interest at the crustal-scale, as identified by Welford et al. (2015a) lies between the Eratosthenes Seamount and the Hecataeus Rise. Despite OBS 1 being a high-quality dataset, as indicated in Table A.1, it is omitted as it lies outside the area of interest.

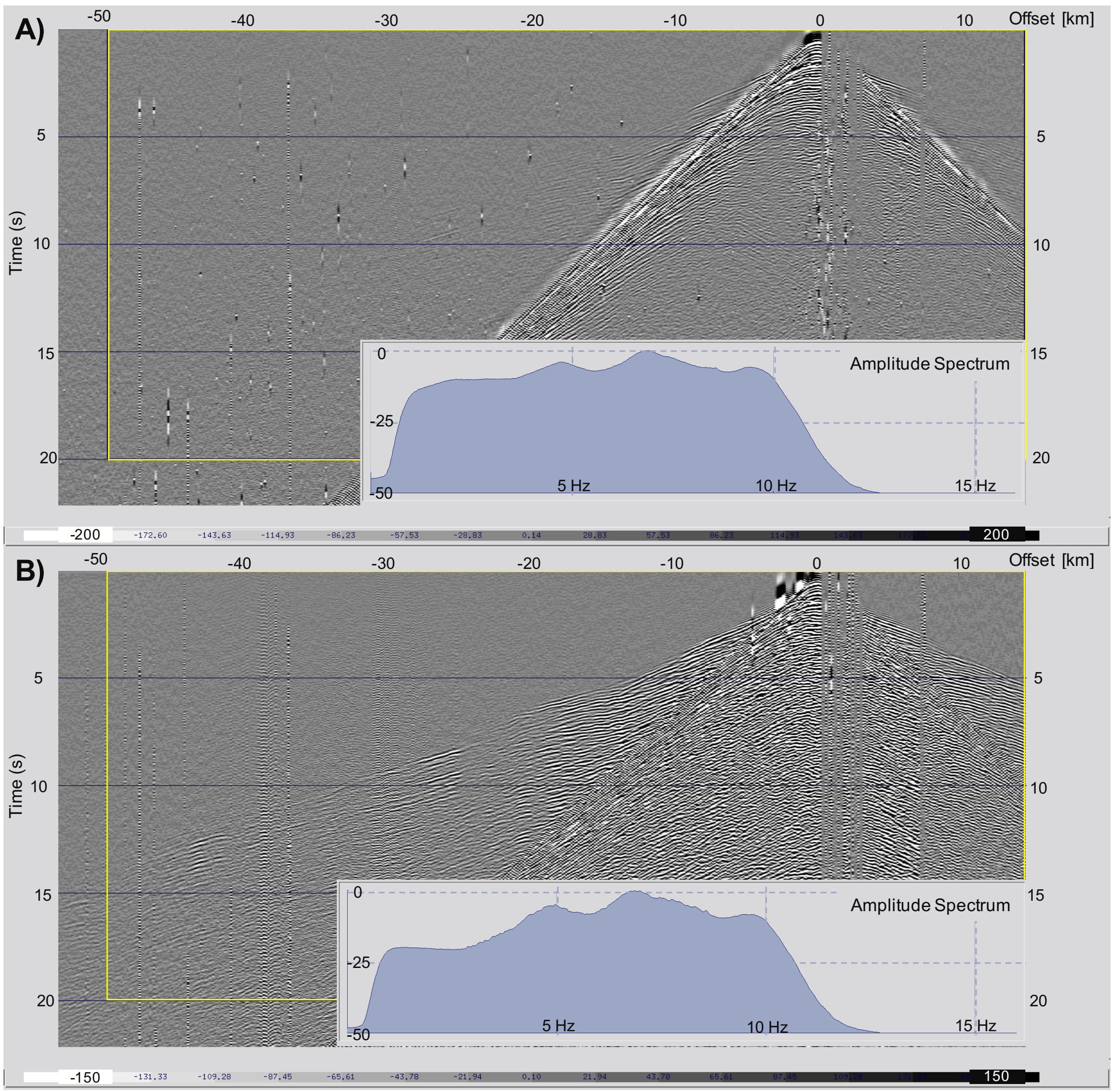

Figure A.3: (A) The hydrophone and the vertical component geophone (B) are shown for OBS 20. An amplitude spectrum is extracted from the region within the box shown for each gather. A Butterworth bandpass filter of 0.5/1.5-9.0/11.0 Hz along with a spherical divergence correction applied as $\sqrt{time}$ is temporally applied to highlight signal in these OBS gathers.

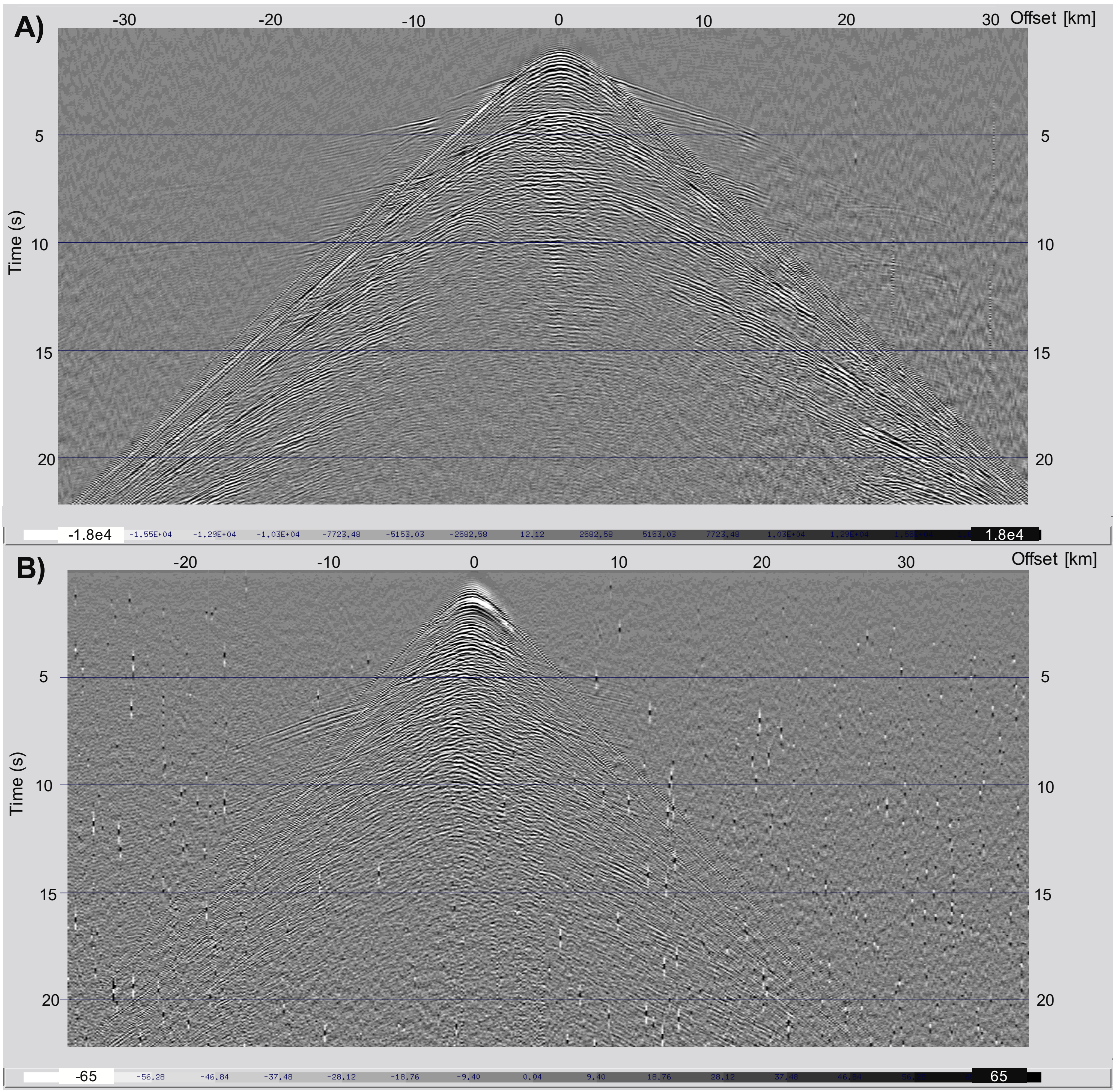

Figure A.4: (A) The omitted OBS 12 and the omitted OBS 2 (B) is shown. A Butterworth bandpass filter of 0.5/1.5-9.0/11.0 Hz along with a spherical divergence correction applied as $\sqrt{time}$ is temporally applied to highlight signal in these OBS gathers.

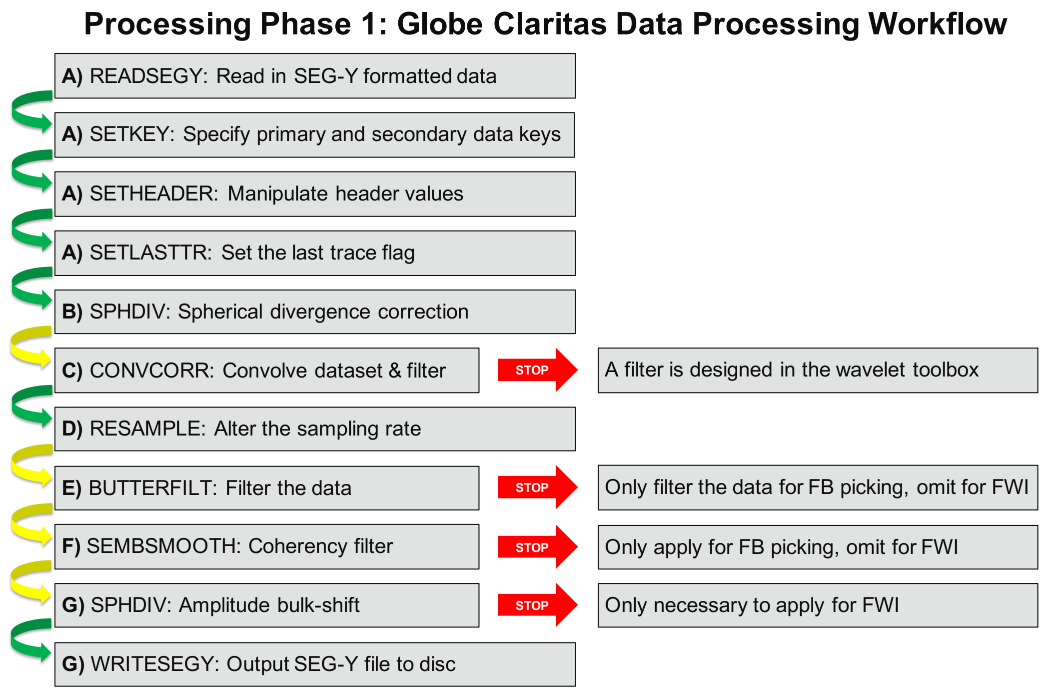

Figure A.5: A summary of the first stage of our OBS processing flow. Each step requires at least one Globe Claritas processing module. Green arrows indicate that this processing step is applied to all output data, while a yellow arrow indicates that this processing step applies to one of the two copies of data. The red arrows imply that the processing step requires additional information or that we apply it to one copy of the OBS data.

A.3 Data Processing Stage 1

The first stage of seismic data processing consists of seven steps following the preliminary data QC stage, as shown in Figure A.1. The omission or inclusion of specific data processing steps, shown in Figure A.5, will generate two different copies of the seismic data. The data requirements for first break picking are different to those for FWI, requiring the processing of a secondary dataset. For first break picking the SNR should be as high as possible. For FWI, the OBS dataset should be aligned with the wavefield modelling physics as much as possible.

A.3.1 Header Reorganization

The OBS data already are in SEG-Y format, OBS relocating and offset calculations are not required (Welford et al., 2015a). The first processing step is then quite simple, and consists of three Globe Claritas modules; SETKEY, SETHEADER, and SETLASTTR. These three modules process all output data (Figure A.5A). Minor trace header re-configurations are required once importing the data with READSEGY. The primary and secondary keys are first set to RECORDNUM and REEL using SETKEY. Second, the RECORDNUM header assigns the proper OBS number to the data. The "trace type" or TRTYPE is set to one as well, meaning that Globe Claritas reads the traces as data. A final header processing step is to set the primary key as RECORDNUM under advanced parameters in the SETLASTTR module. This parameter must be defined to use the REPEAT module despite using single-shot OBS data.

A.3.2 Spherical Divergence

A spherical divergence correction is applied to the data as shown in Figure A.1B. The spherical divergence correction is used in this situation to account for the projection of a 3-D wavefield onto a 2-D plane. A real explosive source produces a 3-D wavefield, but it is only recorded and modelled on a 2-D plane. A 3-D wavefield decays as $\frac{1}{time}$ and a 2-D wavefield will decay as $\frac{1}{\sqrt{time}}$. Therefore a 3-D to 2-D amplitude correction may be approximated by multiplying all traces by $\sqrt{time}$ (Pratt, 1999). An example is shown for OBS 11 in Figure A.6. The spherical divergence correction is not applied in Figure A.6A, but it is applied in Figure A.6B. This correction is completed using the SPHDIV module in Globe Claritas by setting TPOWER1 to 0.5.

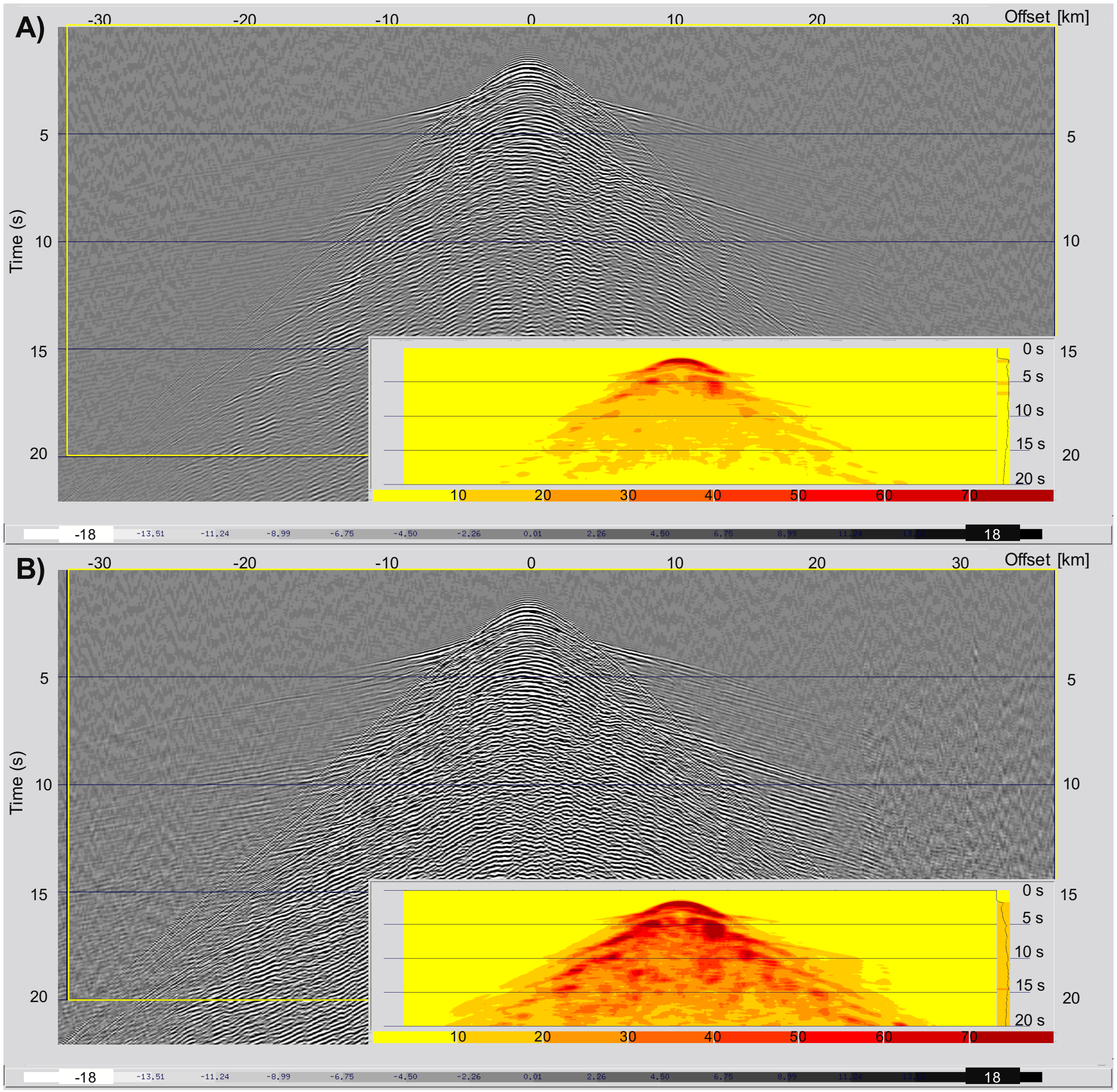

Figure A.6: (A) OBS 11 is shown without, and (B) shows OBS 11 with the application of a spherical divergence correction. A zero-phase Butterworth filter of 0.5/1.5-9.0/11.0 Hz is applied to highlight the signal in this OBS gather. The absolute values of the amplitudes are color-coded in a plot toward the bottom right of each figure.

A.3.3 Convolution

Next, the data are convolved with a filter designed in the Globe Claritas wavelet module (Figure A.1C). This is the first filter to be applied to the data, as previous Figures A.3, A.4, and A.6 are all shown with a Butterworth band-pass filter temporarily applied to attenuate noise for visualization. Convolution of this filter with the OBS dataset should accomplish three goals: remove bubble effects from the source wavelet, balance the amplitude spectra, and consistently shape the source wavelet for each instrument. Accomplishing these goals will result in a more simplistic wavelet that is easier to estimate using the linear inversion technique of Pratt (1999).

We design a filter to shape the source wavelet to an idealized, zero-phase Ricker wavelet, with a peak frequency between 6.0 and 7.0 Hz. The OBS data are close to, but not exactly zero-phase. The average autocorrelation of each OBS is phase-shifted to zero-phase prior to estimating the matched filter between the average autocorrelation and idealized Ricker wavelet. To generate the appropriate filter to be convolved with an OBS gather, the following workflow is completed for each OBS gather:

- A subset of seismic traces must be extracted from the data. Approximately 150 near-offset traces are windowed from 0-10 s and imported into the Globe Claritas wavelet application. The TRPRINT module can output these traces into a suitable format.
- An autocorrelation function (ACF) for each trace is then computed.
- These autocorrelations are stacked to produce an average ACF.

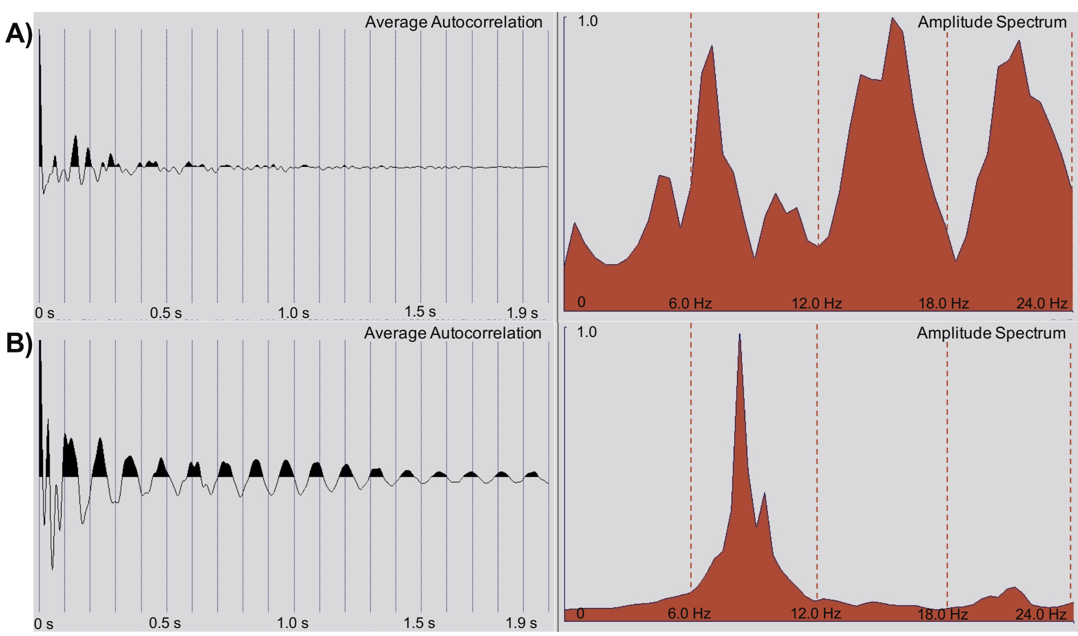

Figure A.7: Stacked, representative autocorrelations for OBS 18 and 9 are shown in (A) and (B), respectively.

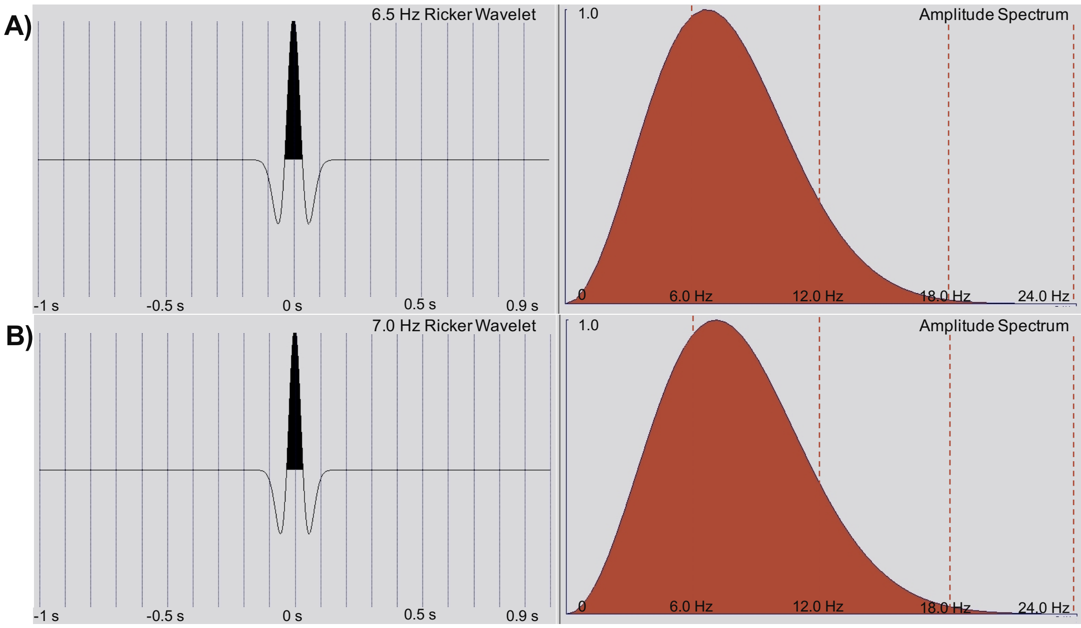

Figure A.8: Two Ricker wavelets with peak frequencies of 6.5 Hz (A) and 7.0 Hz (B) are shown. They are designed for OBS 18 and 9, respectively.

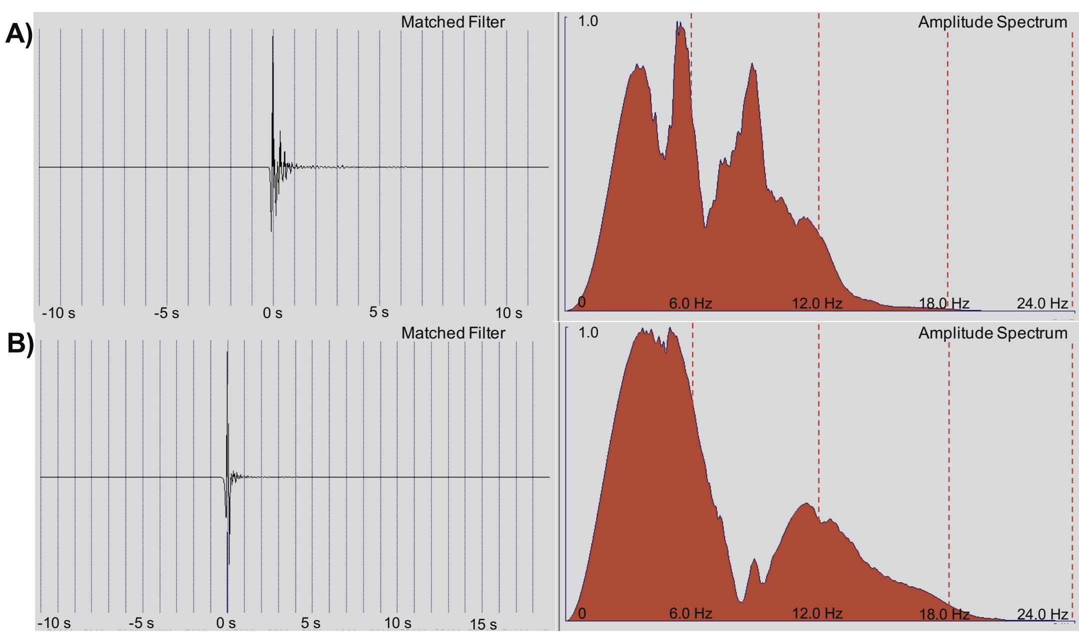

Figure A.9: Designed filters to match the average autocorrelation of the dataset to an idealized Ricker wavelet. (A) The filter for OBS 18, and (B) shows the filter for OBS 9.

- The sampling-rate of this average ACF must be consistent with that of the input data.

- Normalize the amplitude spectrum of the ACF. An average ACF and corresponding amplitude spectra are displayed for OBS 18 and 9 in Figures A.7A and A.7B, respectively. Both estimated wavelets are windowed at 2 seconds to highlight these stacked autocorrelations.

- Shift the average ACF to zero-phase. Only a small time-shift may be required to align the peak energies with t=0.

- Design an idealized Ricker wavelet. A two second Ricker wavelet is designed with a peak frequency slightly lower than the peak frequency of the data.

- Normalize the amplitude spectra of the Ricker wavelet. The idealized wavelets for OBS 18 and 9 are displayed in Figures A.8A and A.8B, where 6.5 Hz and 7.0 Hz wavelets are generated for each dataset, respectively. The reason why these Ricker wavelets aren't identical is that OBS 9 is a German instrument, and it is deficient in low-frequency signal. The peak frequency for the Ricker wavelet is slightly higher here in an attempt to reduce increasing the amplitude of low-frequency noise. However, we stress that this difference is slight, and both wavelets share a high degree of similarity.

- Match the estimated wavelet to the Ricker wavelet. This process is essentially a frequency domain division (i.e. a deconvolution). These matched filters for OBS 4 and 13 are displayed in Figures A.9A and A.9B, respectively.

- Output the matched filter from the Globe Claritas wavelet module, saving it as a *.wts file.

- Convolve the output filter with each trace in its respective OBS gather.

To avoid confusion, the filter designed in this subsection will now be referred to as the matched filter as it replaces the estimated wavelet for each OBS with an idealized zero-phase Ricker wavelet. In Figure A.10A, OBS 18 is shown without the application of the matched filter, and in Figure A.10B, it is shown with the application of the matched filter. Likewise, OBS 9 in Figure A.11A is without the application of the matched filter, and in Figure A.11B the matched filter is applied. All figures show

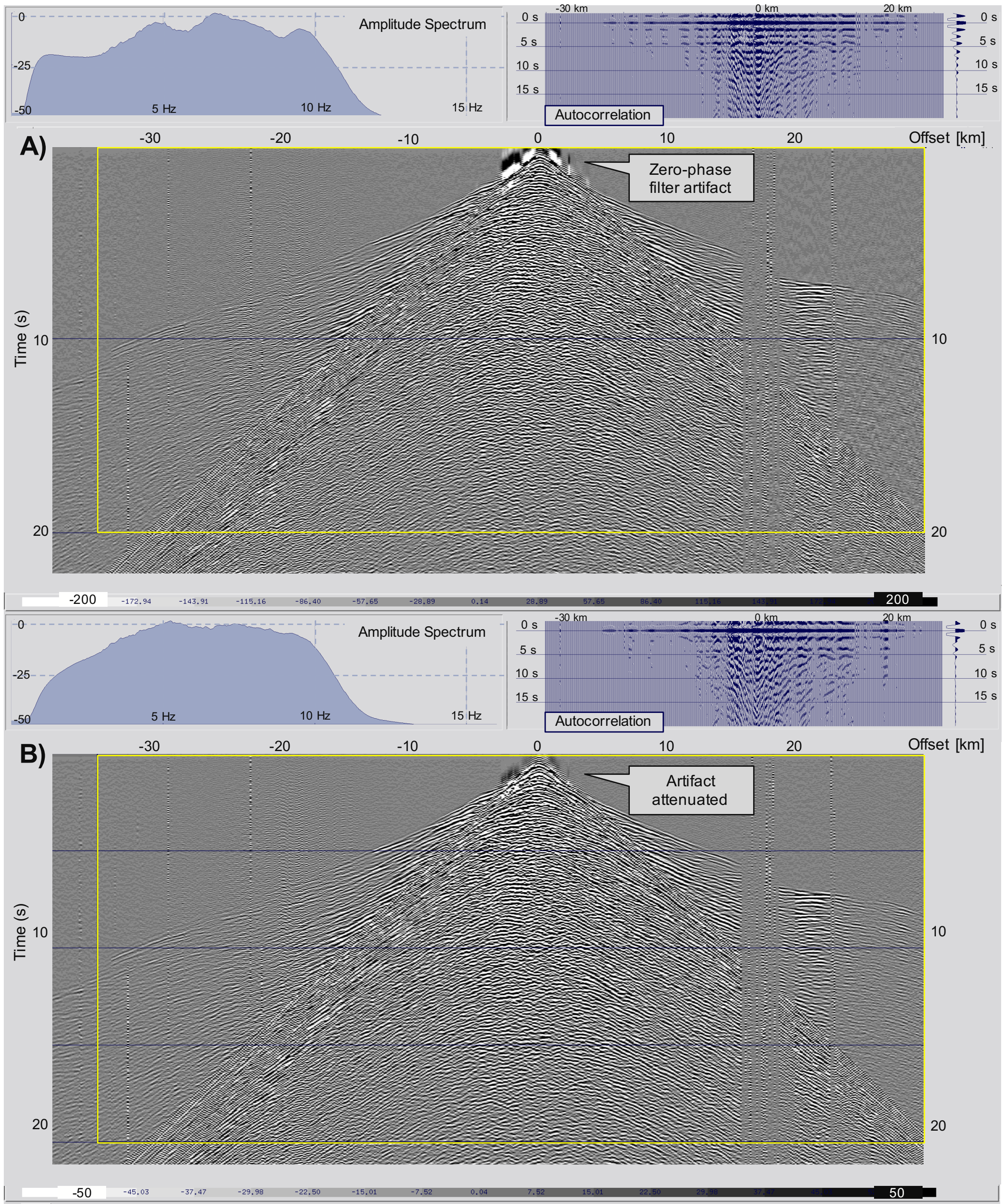

Figure A.10: (A) OBS 18 before, and (B), OBS 18 after convolving a matched filter with the data. An amplitude spectrum and autocorrelation are shown to the top left and right respectfully. The amplitude spectrum and autocorrelation are computed in the region defined by the yellow box.

a shot gather, amplitude spectrum, and autocorrelation. A zero-phase Butterworth band-pass filter is applied in all figures with corner frequencies 0.5/1.5-9.0/11.0 Hz for visualization purposes. This Butterworth band-pass filter is applied to remove noise in Figures A.10A and A.11A, where the matched filter has not been applied. For Figures A.10B and A.11B, this Butterworth band-pass filter is applied along with the matched filter.

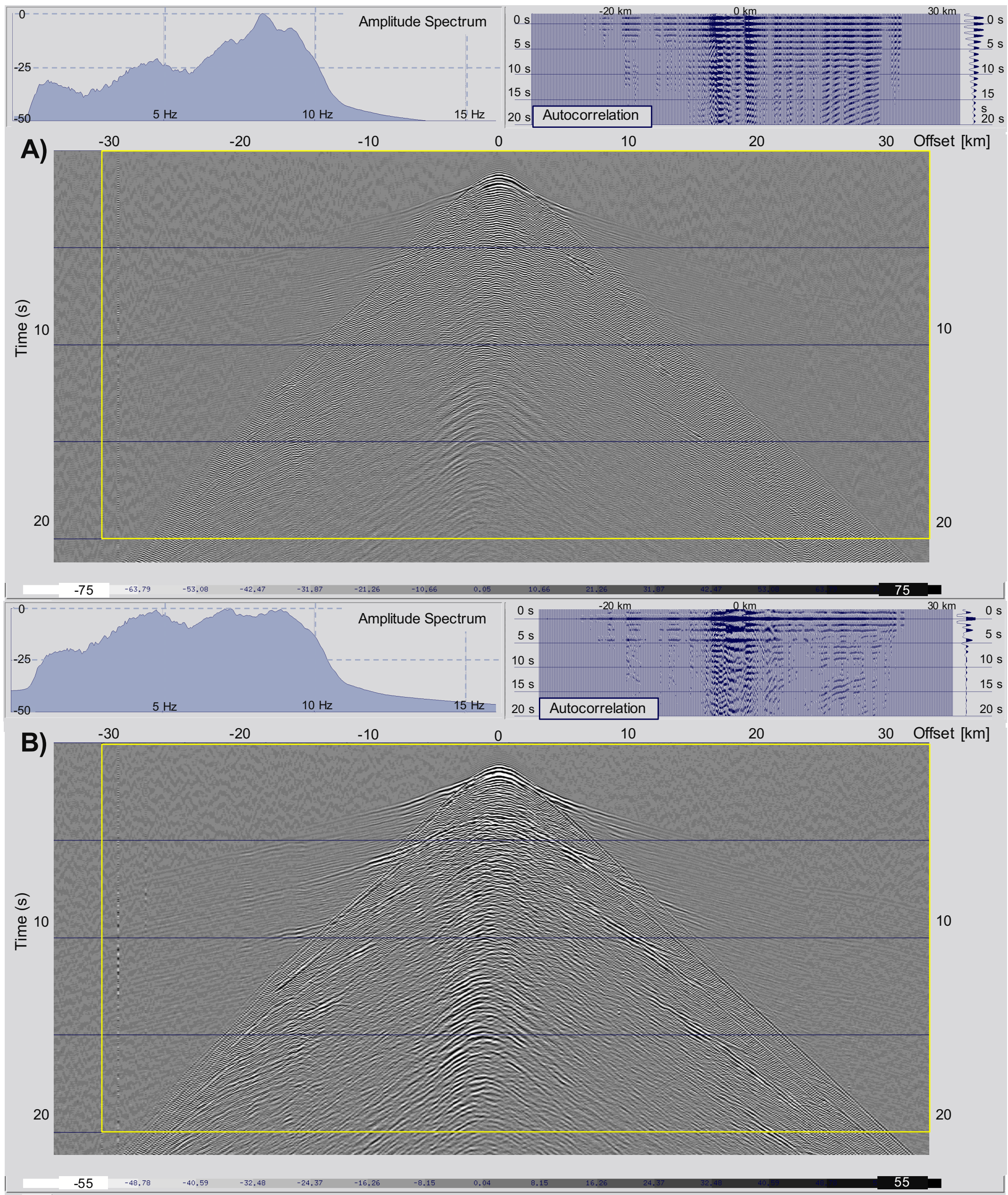

Figure A.11: (A) OBS 9 before, and (B), OBS 9 after convolving a matched filter with the data. An amplitude spectrum and autocorrelation are shown on the top left and right respectfully. The amplitude spectrum and autocorrelation are computed in the region defined by the yellow box.

Convolving the matched filter shown in Figure A.9A with OBS 18 improves the quality of this OBS gather. The amplitude spectrum is more balanced in Figure A.10B with the matched filter applied than it is in Figure A.10A without the matched filter applied. The bubble effect is evident in the autocorrelation shown in Figure A.10A, but it has been attenuated in Figure A.10B. The average ACF in Figure A.10B appears to be more simplistic as well.

OBS 9 shown in Figure A.11 is recognized to be of poorer quality than OBS 18. The low frequency content in OBS 9 shown in Figure A.11A exhibits a notched, unbalanced amplitude spectrum. This amplitude spectrum is improved by convolving the matched filter shown in Figure A.9B to OBS 9. While the new amplitude spectrum in Figure A.11B for OBS 9 is not as balanced as the spectrum for OBS 18, it is a significant improvement in comparison to the spectrum without the matched filter. Reverberations in the autocorrelation function shown in Figure A.11A are so noisy that they are likely to be caused by more than just a bubble effect. The application of the matched filter for OBS 9 significantly attenuates these reverberations manifested in the autocorrelation, as shown in Figure A.11B. The stacked ACFs for OBS 18 and 9 convolved with their matched filters share more similarities than they do without the matched filter.

A.3.4 Re-sampling

As illustrated in Table A.1, there is a variable sample rate for each OBS. This sample rate must be consistent for all datasets before any Fourier Transforms required to convert the OBS data into the frequency-domain for FWI. As shown in Figure A.5D, seismic data re-sampling is completed using the RESAMPLING module in Globe Claritas. The user must specify whether the dataset is zero or minimum-phase in this module. Zero-phase re-sampling is applied to the data as the data are processed to be zero-phase in subsection A.3.3. The target sampling rate for the new data is 4 ms.

A.3.5 Filtering

This section will focus on the frequency content in the data, with emphasis on the lowest frequency available to mitigate the risk of cycle skipping. As illustrated in Figure A.5E, two volumes of data are generated from this processing step. One

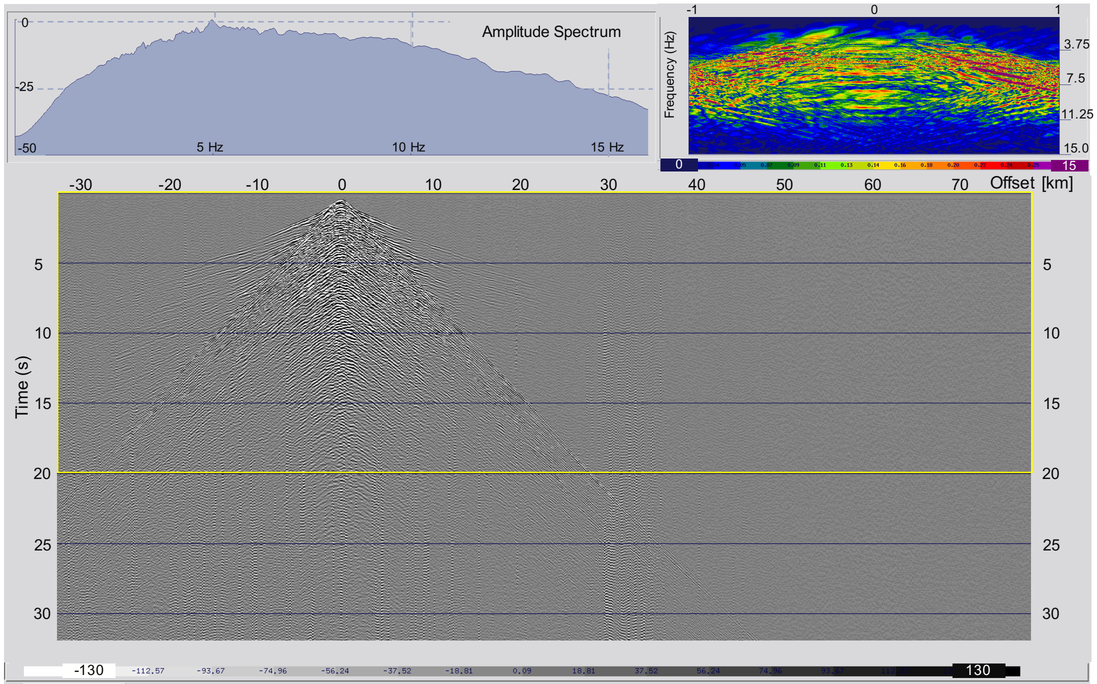

Figure A.12: OBS 5 is shown with corresponding amplitude and F-K spectra. Both spectra are computed within the region defined by the yellow box.

volume of data has the Butterworth band-pass filter applied to it. This volume of OBS data is for first break picking. Another volume of data intended for FWI does not have the Band-pass filter applied to it.

Figure A.12 shows OBS 5 along with its amplitude and F-K spectra. The amplitude spectrum shown here for OBS 5 is smooth, and it approximately correlates with the peak frequency content in the original data before convolution. The F-K spectrum shows that the majority of the seismic energy lies within the first 11.25 Hz, and a spatial alias occurs in the direct wave between 7.5 and 11.25 Hz. The amplitude spectra shown here is representative of other GSC instruments, whereas the amplitude spectra for German OBS generally have a notched amplitude spectrum (recall Figure A.11).

We investigate the frequency content of OBS 5 in Figures A.13, A.14, and A.15. The goal of this investigation is to correlate the waveforms in the x-t domain to a frequency range. Doing so will provide a rough understanding of what the starting frequency should be for FWI. The conclusions for this investigation are summarized in Table A.2.

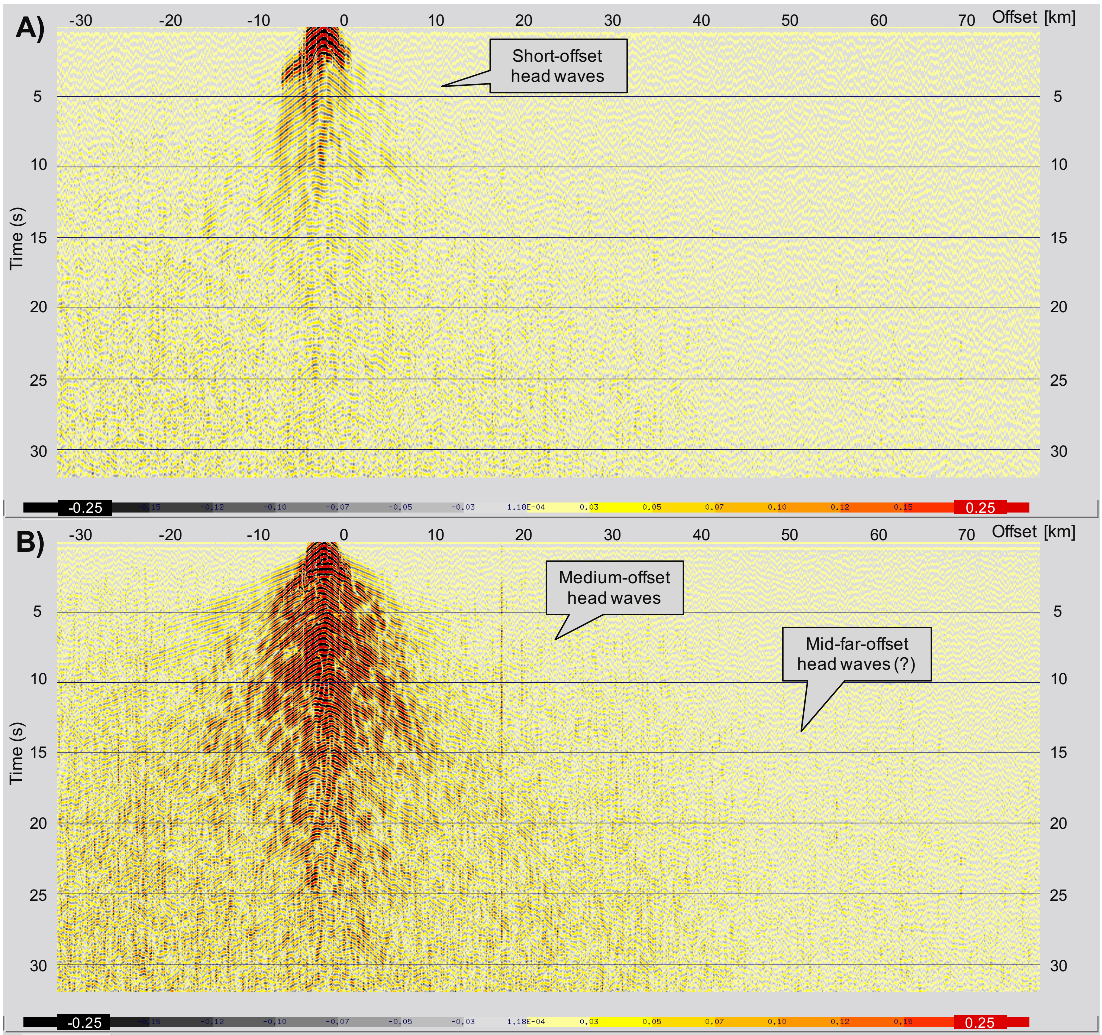

Figure A.13: OBS 5 is shown with two different Butterworth band-pass filters applied. A 1.0 Hz taper is added to each end of the band-pass filter. (A) shows a frequency content of 1.5-2.0 Hz, and (B) shows a frequency content of 2.0-2.5 Hz.

In Figure A.13A first break signal is only observed up to approximately 10 km offset for the 1.5-2.0 Hz frequency band. However, due to the use of 1.0 Hz tapers, some

Table A.2: Summary of a sequential filtering test for OBS 5.

Filter	Figure	Maximum Offset With Undisputed Signal
0.5/1.5-2.0/2.5 Hz	A.13A	10 km
1.0/2.0-2.5/3.5 Hz	A.13B	25 km
1.0/2.0-4.0/5.0 Hz	A.14A	45 km
3.0/4.0-6.0/7.0 Hz	A.14B	$>$ 70 km
5.0/6.0-8.0/9.0 Hz	A.15A	$>$ 70 km

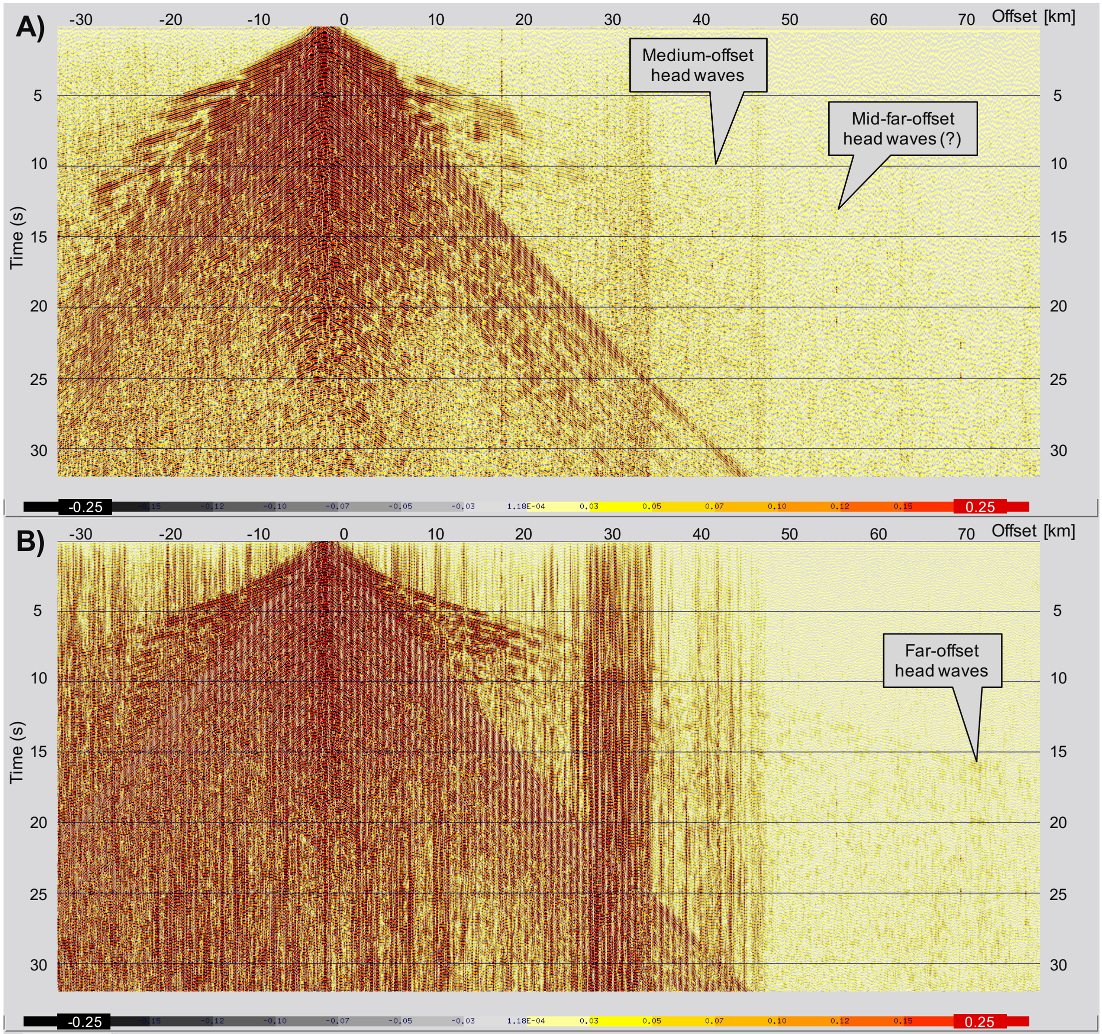

Figure A.14: OBS 5 is shown with two different Butterworth band-pass filters applied. A 1.0 Hz taper is added to each end of the band-pass filter. (A) shows a frequency content of 2.0-4.0 Hz, and (B) shows a frequency content of 4.0-6.0 Hz.

of this signal may lie outside this frequency band. There is a big difference between the 1.5-2.0 Hz and 2.0-2.5 Hz frequency bands when comparing Figures A.13A and A.13B. In addition to first breaks being observed at approximately 25 km offset, a faint seismic signal might be present at upwards of 50 km, as highlighted in Figure A.13B. The 2.0-2.5 Hz frequency band is interpreted to be a better starting frequency bandwidth for FWI.

Figures A.14A, A.14B, and A.15A correspond to the 2.0-4.0, 4.0-6.0, and 6.0-8.0 Hz frequency bands respectfully. Each increasing frequency band has some signal present

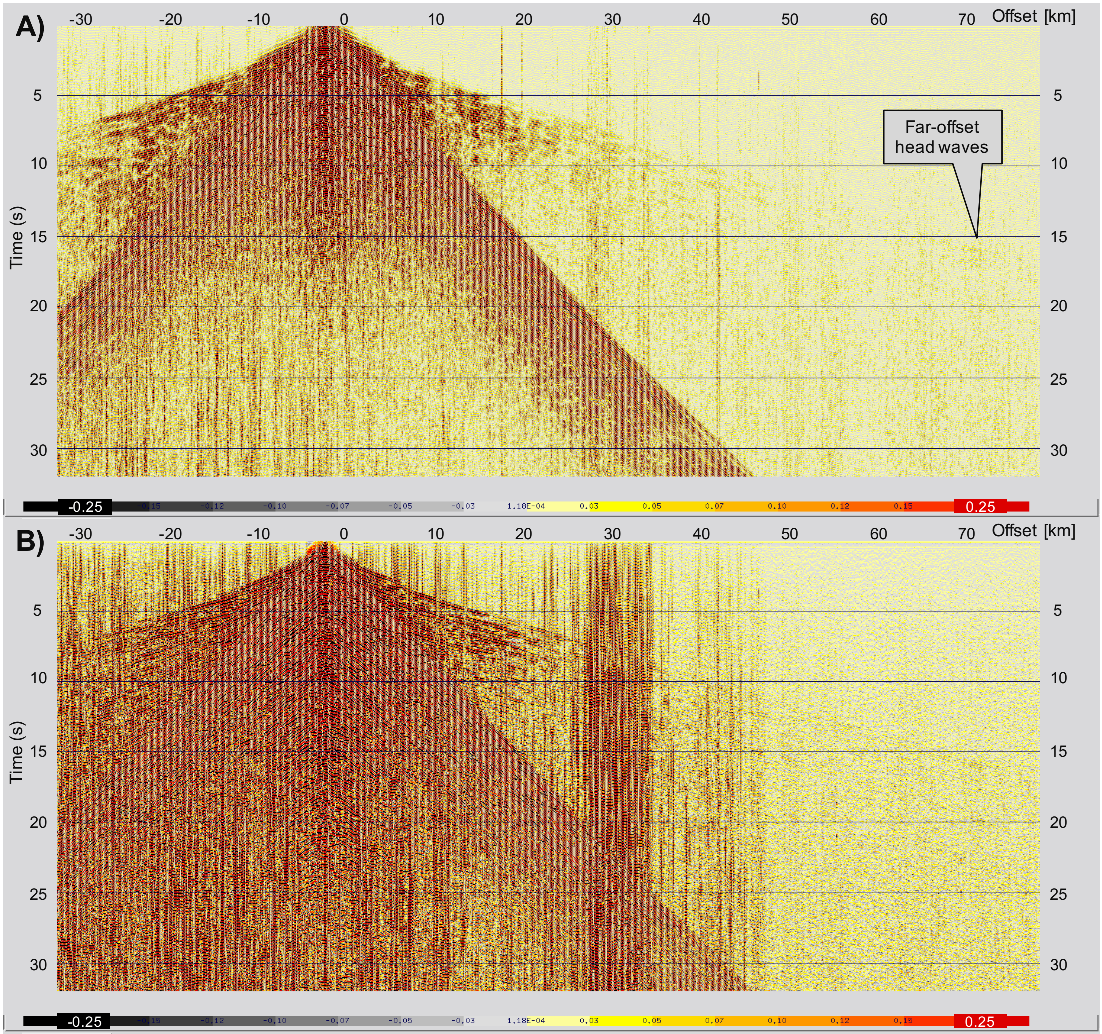

Figure A.15: OBS 5 is shown with two different Butterworth band-pass filters applied. A 1.0 Hz taper is added to each end of the band-pass filter in (A), that contains a frequency content between 6.0-8.0 Hz, and (B) shows a frequency content of 1.5-9.0 Hz.

at progressively larger maximum offsets, as documented in Table A.2. Seismic arrivals within the 2.0-4.0 Hz frequency band may be present at offsets upwards of 55 km, as shown in Figure A.14A. However, this interpretation is speculative. The presence of noise in the dataset becomes a significant problem between 4.0-6.0 Hz. Before 4.0 Hz, the data are relatively noise-free. This type of noise is persistent across many of the GSC instruments.

In Figure A.15B a 0.5/1.5-9.0/11.0 Hz Butterworth band-pass filter is applied to

OBS 17. This filter design encompasses all relevant seismic signal. It is unnecessary to apply this filter to the volume of OBS data processed for FWI as the extraction of monochromatic frequencies only occurs at frequencies within this pass band. Using this analysis, we decide to begin our FWI process from 2.0 Hz.

A.3.6 Semblance

We use a conventional coherency filter to further enhance low-amplitude, far-offset head waves. The SEMBSMOOTH module in Globe Claritas is used to apply a coherency filter to the volume of OBS data intended for first break picking only. Semblance values are computed for a defined window of seismic traces along angled straight lines using a slant-stack method. Coherencies are computed from these semblances and raised to a defined power to remove non-coherent data and then smoothed along the maximum semblance direction. The semblance filter is then applied by multiplying the smoothed data with computed coherency values and adding back a percentage of the original dataset.

This semblance processing workflow will distort AVO relations within the observed data. As shown in Figure A.5F, only the dataset intended for first break picking will have a coherency filter applied to it. Therefore, the volume of data intended for first break picking will have both a Butterworth and coherency filter applied to it, and the data volume for FWI will have neither applied. This subsection will show the effect of this coherency filter on the data following a description of the coherency filter's design.

We use a design window of 21 traces to compute semblances. Semblances within this 21 trace window are then raised to the power of 2 to form coherencies, and 201 different slowness dips are used to calculate semblances. Once coherencies are computed, we add back 40 percent of the original dataset to the coherency filtered data.

In Figure A.16A, OBS 15 is shown without the application of this coherency filter, and in Figure A.16B this coherency filter is applied. A colour-coded plot highlighting the normalized seismic amplitudes is shown with and without this coherency filter applied as well. The region of greatest change is highlighted in Figure A.16B between -40 to -30 km offset. It is easier to see the increased amplitude of these seismic arrivals in the normalized amplitude plots. This inconsistent amplitude increase with offset

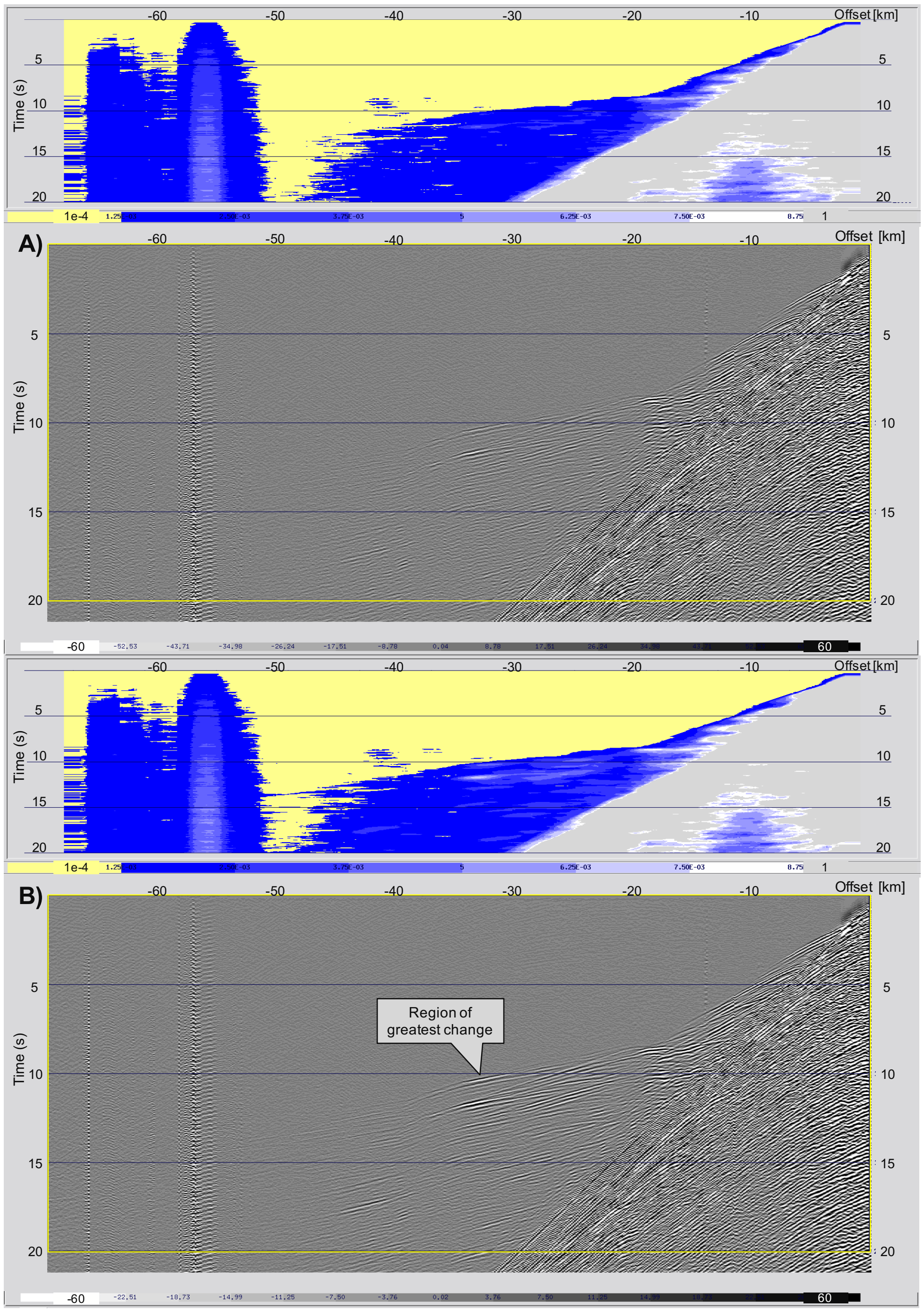

Figure A.16: (A) OBS 15 is without a semblance filter, and with a semblance filter in (B). The top panel for each image colour-codes the absolute value of the amplitudes shown in each OBS gather from the region shown by the yellow box.

is because the computed coherencies alter the AVO trends in the data, making this correction unacceptable for FWI. However, these far-offset arrivals are easier to pick for traveltime tomography following the application of this coherency filer.

A.3.7 RMS Amplitude Bulk Shift

It is beneficial to have comparable amplitude profiles for each OBS as the noisy trace attenuation in processing stage 2 (Subsection A.4.1) performs a statistical analysis of the noise root mean square (NRMS) profiles of each OBS. Normalizing the root mean square (RMS) amplitude of each OBS to one ensures that this statistical analysis is comparable across all OBS. This processing stage effectively removes the coupling bias between the OBS and seafloor from the data.

The SPHDIV module applies this amplitude bulk shift. All entries in SPHDIV are zero except the SFACTOR1, which multiplies all traces by a constant. The Globe Claritas XVIEW module computes the RMS amplitude for each OBS gather. The entire OBS dataset is processed to have an RMS amplitude of one for this study.

A.4 Data Processing Stage 2

The second stage of seismic data processing consists of seismic amplitude corrections. This stage of data processing uses an AMPLITUDE_PROCESSING Python module, which is a part of a suite of Python codes used in this study. There are two seismic data processing objectives for this stage of data processing to attenuate noisy seismic traces, and bulk-shift the amplitudes of the observed dataset to match those of the modelled dataset. The application of these corrections to the OBS data occurs during the FWI data generation stage (Figure A.1). Setting up the OBS data in Python requires a set of first break picks, which acts as a geometry file. Attenuating noisy traces requires first breaks to define the NRMS window as well. Applying the amplitude bulk shift requires a starting model to forward model waveforms and produce synthetic OBS gathers. First break picking and starting model generation are discussed in Appendix B.

A.4.1 Noisy Trace Attenuation

Much of this WARR acquisition is plagued by noise, as is expected in a real-data acquisition. Recall OBS 5, shown within the 4.0-6.0 Hz frequency band in Figure A.14B. Many of the traces in the OBS gather are dominated by high amplitude noise, especially within the 50 km offset range. Rather than removing these noisy traces, we scale down their amplitudes (Górszczyk et al., 2017). Analysis of the NRMS window occurs above the first breaks for each trace in an OBS gather. Traces with high NRMS values will have their amplitudes scaled down to a user defined level. What is problematic about this approach is the assumption of a correlation between high NRMS values and noisy traces in the OBS gather. We observe instances in this dataset where some high amplitude noise is present in only a portion of the seismic trace. This noise is present in the NRMS window, but not outside the NRMS window on some traces.

Figure A.17A shows an uninterpreted gather from OBS 20, and in Figure A.17B different noise types are interpreted for OBS 20. The circled green noise shown in Figure A.17B is high amplitude and confined to a single trace. It continues from the NRMS window down in time to the head waves in OBS 20. The circled yellow noise in Figure A.17B is of lower amplitude but spread across multiple traces in the OBS gather. This type of noise in the data does extend down to the head waves, but due to its low amplitude, it does not completely mask the signature of the head waves. The circled red noise in Figure A.17B is problematic for this noise removal approach. It is present in the NRMS window but dissipates before interfering with the head waves. This noise is of moderately high amplitude, and it may result in flagging the entire trace as noisy. If this is the case, the algorithm unnecessarily attenuates good signal underlying this trace. An interactive process that allows for better user interaction within the algorithm is not yet implemented.

The maximum and minimum offsets for this NRMS investigation must be defined. There are typically higher NRMS values at near offsets due to a smaller time window for the NRMS computation. The minimum offset for an NRMS trace analysis is 8 km. Noise at the furthest offsets becomes angled due to the shot wrap-around. Therefore, there is no vertical correlation in noise at this point, and this noise removal process is flawed. Based on the geometrical setup of this study, shot wrap-around noise occurs at approximately 90 km offsets. These offset constraints provide an 82 km window of offset to attenuate noisy traces, between 8 and 90 km.

Standard deviations are computed for the NRMS amplitude values within the offset range. Noisy traces are easily identifiable with high values of standard deviation that separates them from a background noise floor, which is a straight line fit to the NRMS amplitude values with standard deviations composing a covariance matrix. Standard deviations greater than a defined threshold will raise a flag to scale the amplitude of

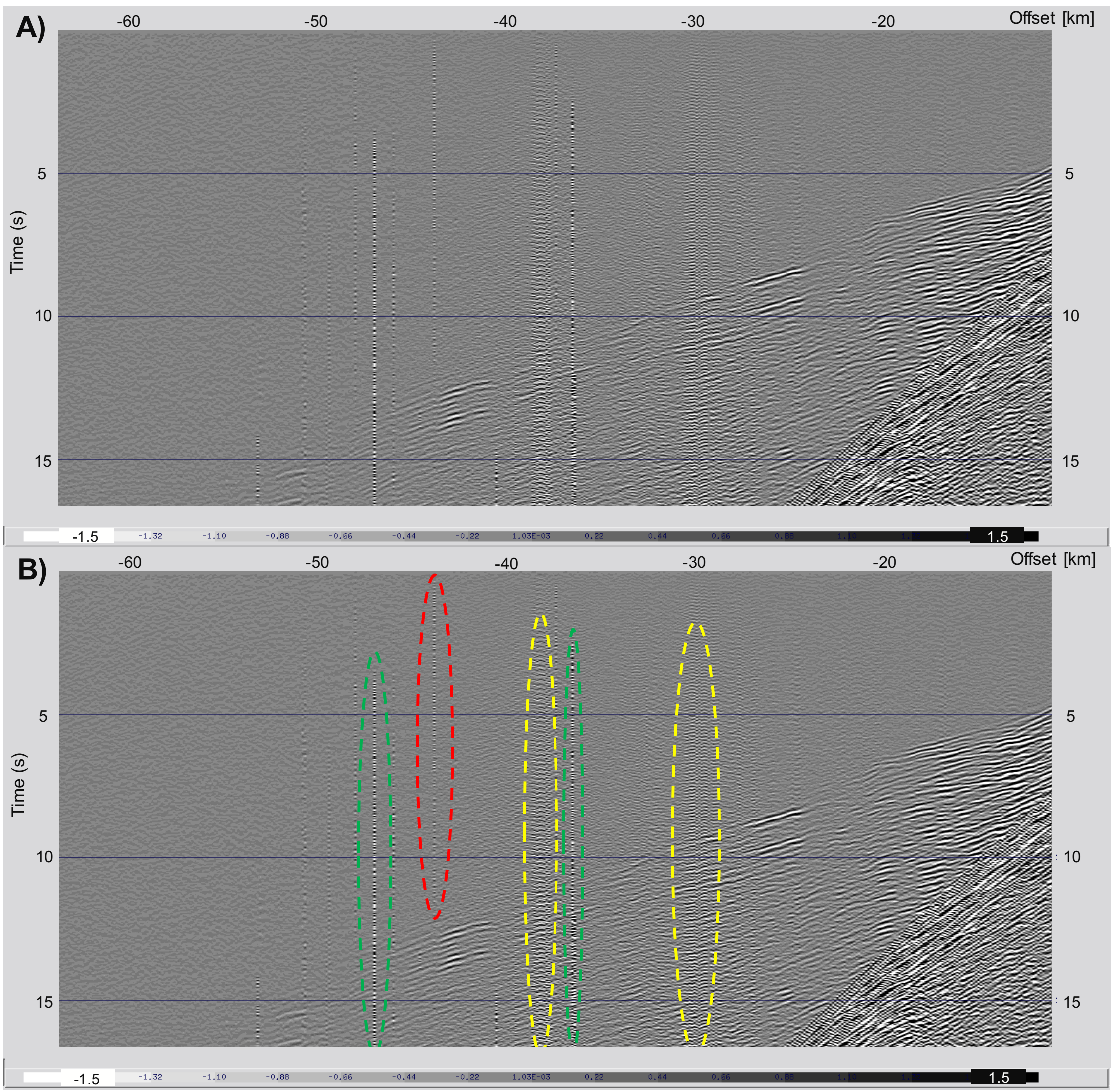

Figure A.17: (A) OBS 20 is shown, and shown with different noise types interpreted in (B). The green circles highlight a high-priority noise-type to attenuate, the yellow circles highlight a low-priority noise-type to attenuate, and the red circle highlights a problematic noise-type.

these traces down to a percentage of the background noise floor. Having a smaller standard deviation cutoff will flag more traces as noisy, resulting in the reduction of their amplitudes to a user-defined percentage of the background noise floor. This will ensure that no high-amplitude noisy traces are adversely affecting the objective function, especially at the sensitive far offsets.

OBS 17 is shown in Figure A.18A without trace amplitude attenuation. Standard deviations computed in the NRMS window are plotted in Figure A.19A. A cluster of high standard deviation values appears between offsets -22 to -12 km. A lower standard deviation cutoff of 0.003 is taken for OBS 17, as shown by the black line in Figure A.19A. More traces are flagged as noisy, as shown by the red points in Figures A.19A and A.19B. These flagged traces are scaled down to 80 % of the NRMS floor, where the dotted line shows the noise floor of the NRMS amplitudes in Figure A.19B. The green points in Figures A.19A and A.19B show the new standard deviations and NRMS amplitudes for the adjusted traces, respectfully. The correction that is applied to each trace to obtain the result shown in Figure A.19B is shown in Figure A.19C.

The result of this NRMS scaling applied to the data is shown in Figure A.18B. The amplitude scale in Figure A.18B is different than in Figure A.18A as a result of the amplitude bulk shift. The AMPLITUDE_PROCESS algorithm includes both the trace attenuation and a bulk-shift correction in a single amplitude correction file. However, the noisy trace attenuation process did an excellent job of reducing the amplitudes of the noisy traces in Figure A.18A. The most obvious region of improvement in Figure A.18B lies between -10 and -20 km offset. This region of improvement correlates with the region flagged with high standard deviations in Figure A.19A.

The AMPLITUDE_PROCESS module directly applies this noisy trace attenuation correction to the OBS traces. RMS trace amplitudes are now computed from these OBS traces with the trace attenuation correction applied. This RMS trace amplitude information is used in the next phase of the AMPLITUDE_PROCESS module, the amplitude bulk-shift.

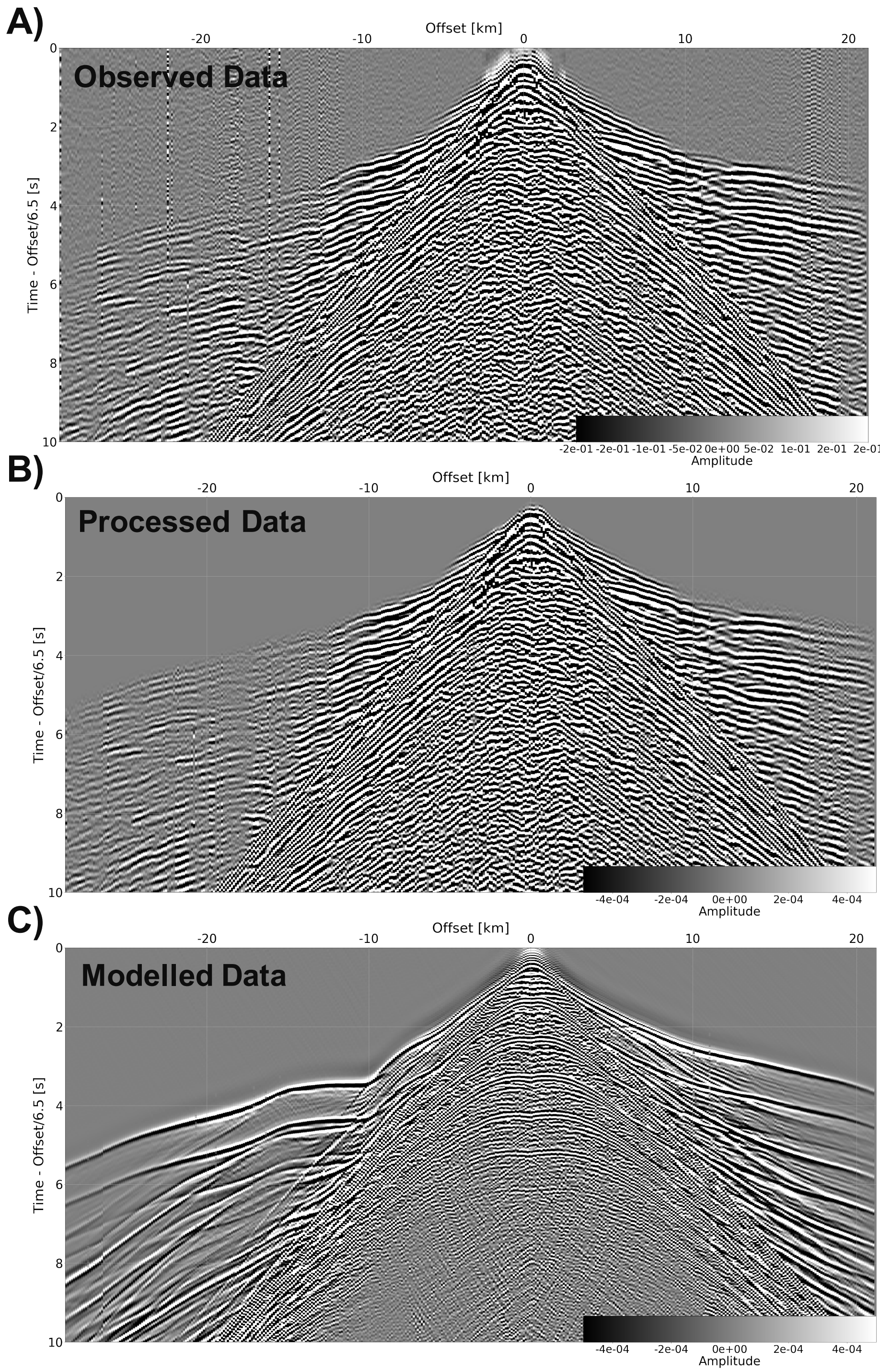

Figure A.18: (A) OBS 17 plotted in reduced time does not have the trace-attenuation workflow applied, and (B) shows with the trace-attenuation workflow applied. Additionally, we mute above the first break picks in (B) as described in section A.5. (C) A forward modelled shot gather.

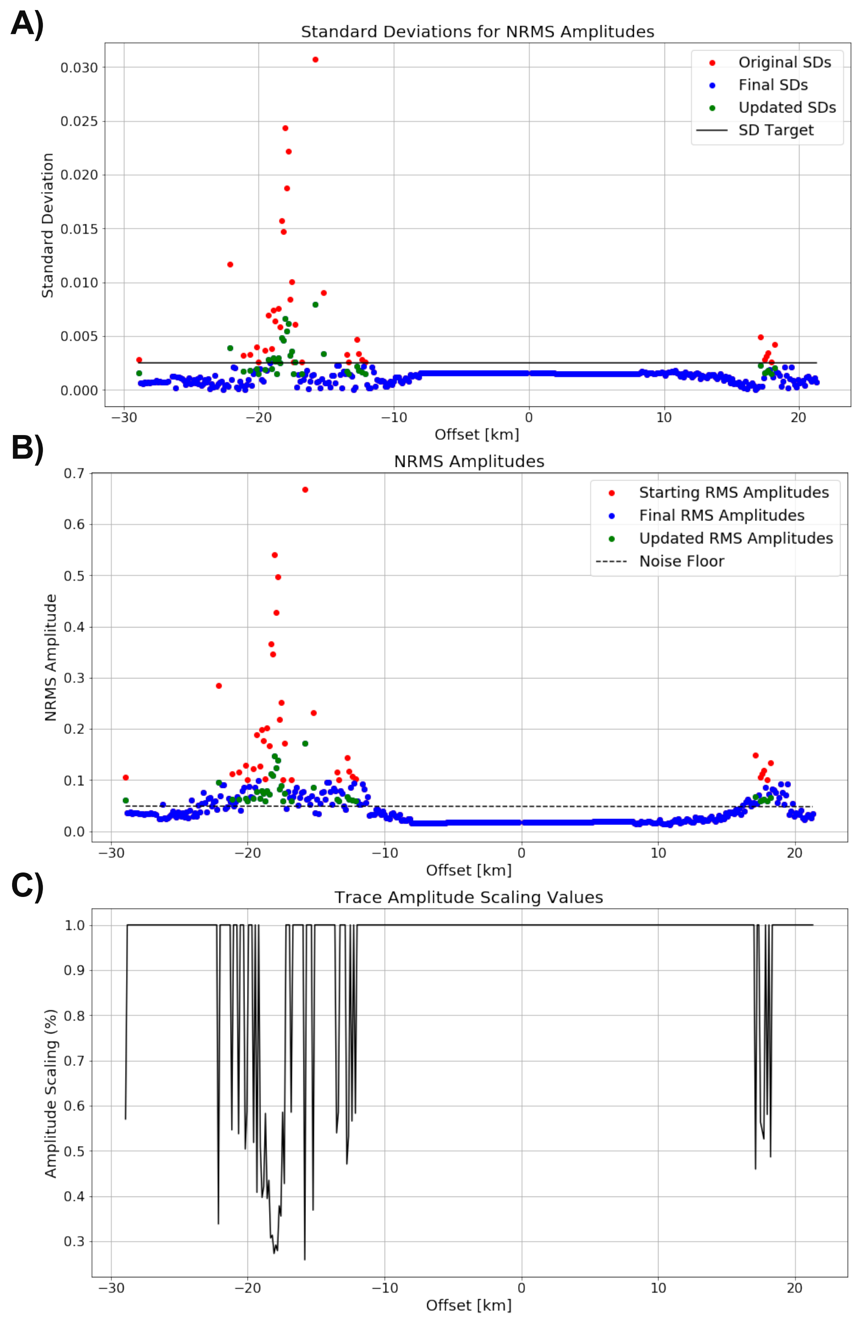

Figure A.19: Three plots illustrate the noisy trace attenuation process for OBS 17. (A) The standard deviations of the NRMS values, where the red points are traces that exceed the standard deviation cutoff, shown by the green line. The blue points are the standard deviations of the traces once scaled. (B) The amplitudes and scaled amplitudes, and (C) shows the amplitude-scaling trace coefficients.

A.4.2 Amplitude Bulk-Shift

In real-data OBS acquisitions, the coupling between the instrument and seafloor is likely not consistent for all instruments due to variations in near-surface geology. One method of fixing this instrument-seafloor coupling issue is to bulk shift all OBS data such that they have a consistent RMS amplitude. This technique I applied to the FWI data and discussed in subsection A.3.7. There is also a discrepancy between the observed data amplitudes, and the amplitudes predicted by mathematical modelling. Correcting for instrument coupling may not correct for this discrepancy between the predicted and observed data amplitudes.

The amplitude bulk-shift data processing step occurs immediately following the noisy trace attenuation step. A one-term linear optimization problem minimizes the differences between the RMS trace amplitudes of the observed and predicted data. In its current implementation, the algorithm computes RMS trace amplitudes. The RMS trace amplitudes of the observed data are higher than the modelled data due to a variety of factors, including noise. Merely taking the ratio of the RMS amplitudes for modelled and observed datasets and applying this value to the observed data as a correction is not sufficient. This is because the signal in which we wish to model through FWI does not contribute much to the trace RMS amplitude value. An additional constant is applied to the observed data on a trial and error basis to ensure its signal is of similar amplitude to that in the modelled data, not the RMS trace amplitude.

Figures A.18A and A.18B show OBS 17 with and without amplitude scaling applied to them for OBS 17. The modeled data for OBS 17 are shown in Figure A.18C. The amplitude range for the observed and modelled data following the application of the bulk-shift are aligned.

A.5 FWI Data Generation

The FWI data generation phase of OBS data processing takes the OBS data processed during stage one and converts the time-domain SEG-Y data into frequency-domain data formatted for TOY2DAC. The coded Python module TOY2DAC_FILES completes the FWI data generation processing step. This subsection will focus on the

SEG-Y data processing aspects of this code, namely the data-muting, data-damping, and Fourier transform processing steps.

A.5.1 Data-Muting

Data-muting, data-damping, and the amplitude correction (Section A.4) are all applied in a single function within the TOY2DAC_FILES module. Data mutes above the first breaks are applied first. The data are zero-phase, and to mute at the location of the first break pick would delete half of the first-arrival waveform. A buffer-zone of 100 times samples or 400 ms above the first break is not muted. A taper is applied to data in this buffer zone, improving the result of the Fourier transform. Recalling Figure A.18A in Section A.4, OBS 16 is shown without trace-mutes applied above the first breaks. In Figure A.18B, OBS 16 is shown with these trace-mutes applied and a tapered buffer-zone.

A.5.2 Data-Damping

Weighting the application of frequency-domain damping from the first breaks is effective for WARR investigations (Brossier et al., 2009; Górszczyk et al., 2017). TOY2DAC implements this by applying a weight to the observed and predicted data in the frequency domain. This weight is the exponential that accounts for the first break time in Equation 1.9. Therefore, we damp the observed data from $t = 0$ in the time-domain before a Fourier transform and frequency extraction. This procedure is undoubtedly a pit-fall, and it breaks the application of frequency-domain damping into two parts; damping from t=0, and applying a weight to replicate damping from the first breaks in the frequency domain. OBS 20 in Figure A.20 is shown with a maximum offset of 60 km and various damping values applied from its first breaks. In Figure A.20A, a high damping value of 5.0 is used to attenuate everything but the earliest head-wave arrivals. By progressively decreasing this damping value, more complex arrivals are progressively introduced to the dataset. For example, Figure A.20C uses a damping value of 0.5 which includes more complex arrivals, but only a small portion of the entire OBS dataset.

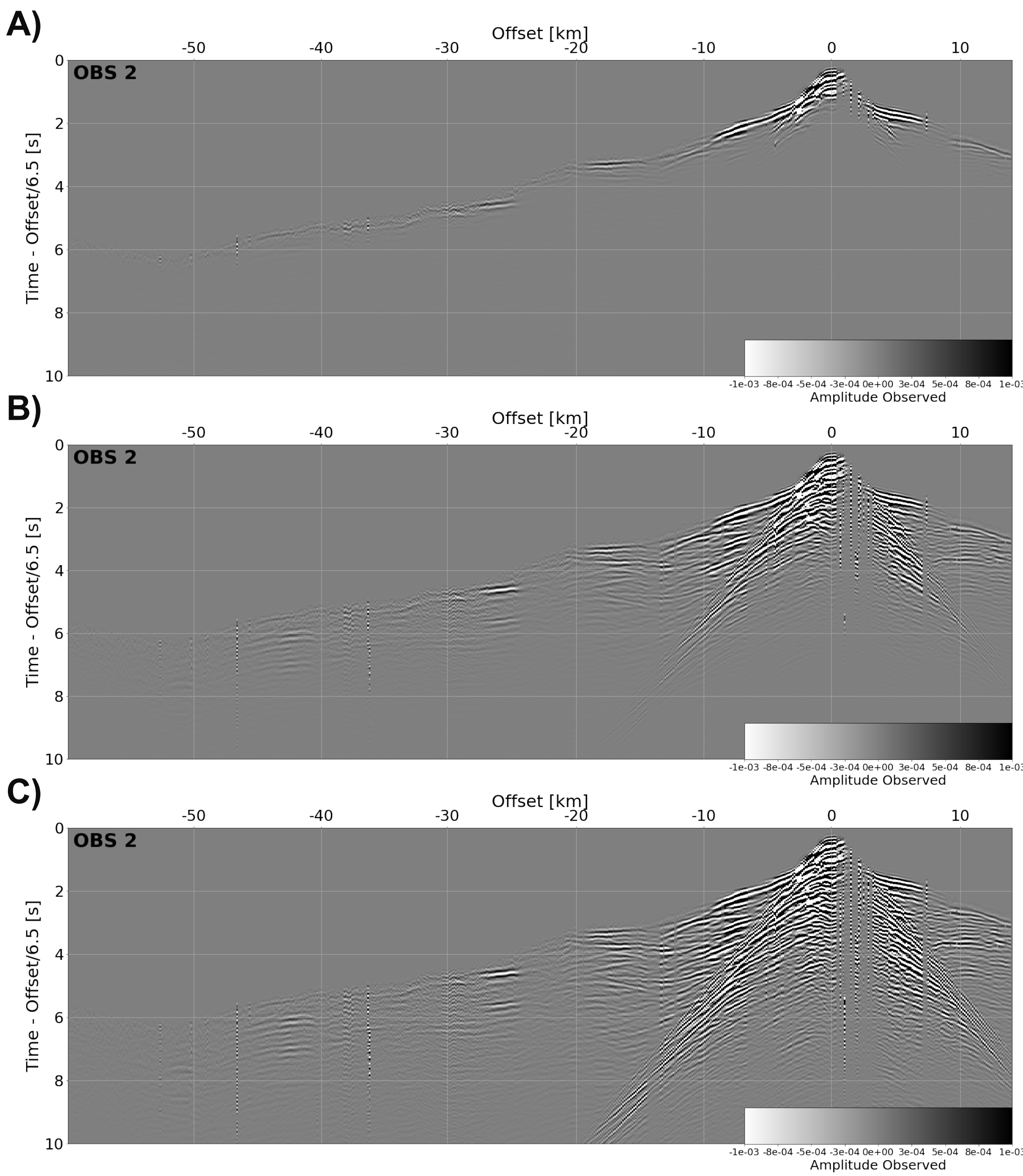

Figure A.20: OBS 20 with four different data damping values from the first break picks. (A) We use a damping value of 5.0, (B) shows a damping value of 1.0, and (C) shows a damping value of 0.5.

A.5.3 Fourier Transform

Following data-muting, data-damping, and an amplitude correction (Section A.4), the SEG-Y data are transformed to the Fourier domain where the frequency-domain

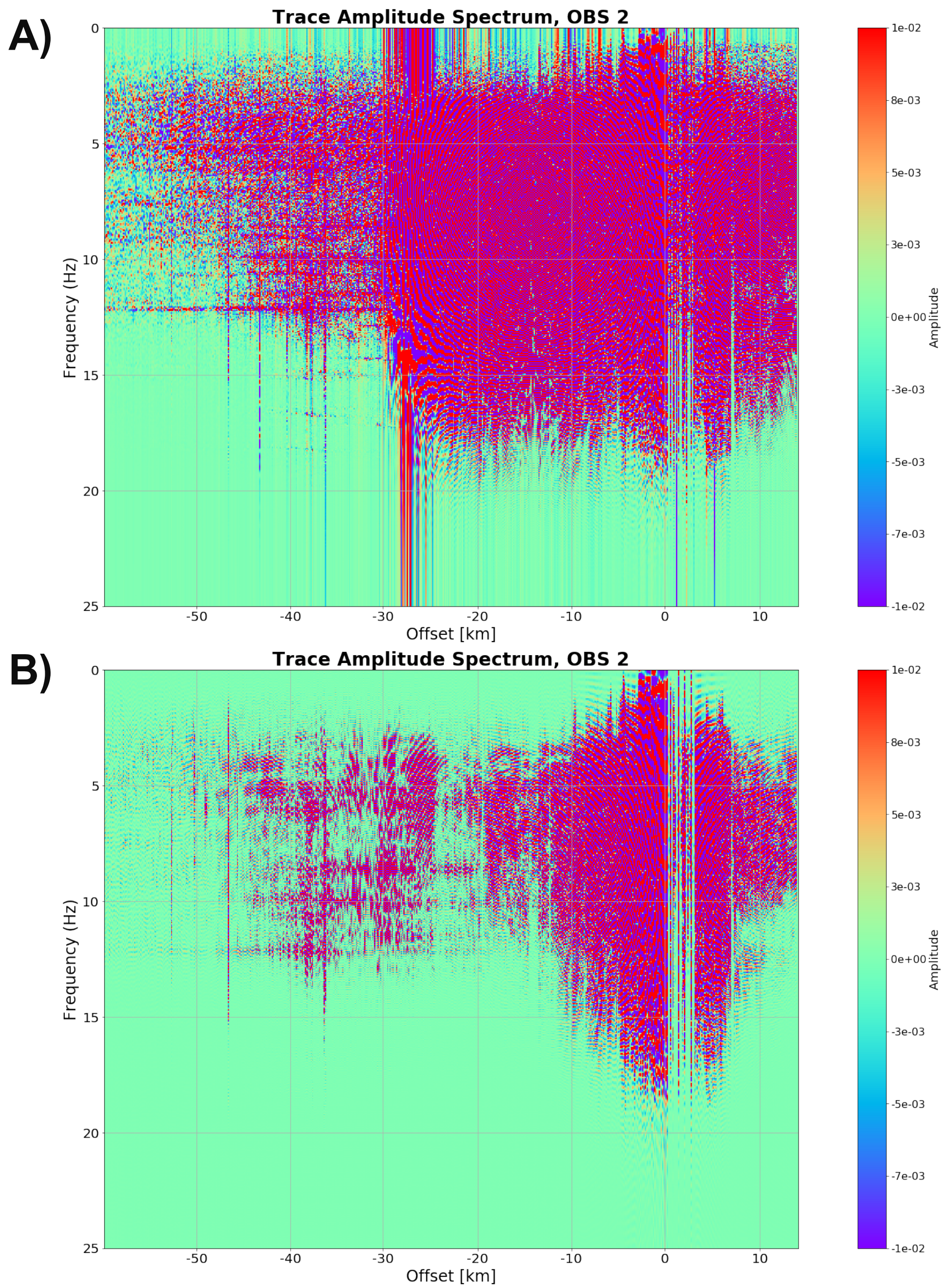

Figure A.21: Two different datasets for OBS 2 are Fourier transformed. (A) OBS 2 has a 60 km maximum offset constraint and no time-domain damping. (B) The data have a 60 km maximum offset with a damping value of 1.0 applied from the first breaks.

data are extracted. When completing the Fourier transform, the maximum frequency is governed by the sampling rate of the OBS data in time, $fmax = \frac{1}{sr_t}$. Recalling

Subsection A.3.4, the OBS data consistently have their sampling rate set to 4 ms, resulting in a maximum frequency of 250 Hz. A Frequency sampling is obtained from the maximum time of the OBS data, $sr_f = \frac{1}{tmax}$. The maximum time for the OBS data is 40 seconds. This maximum time results in a frequency sampling rate of 0.025 Hz.

In Figure A.21, two different trace amplitude spectra are shown for two different versions of the data from OBS 20. In Figure A.3A, the un-damped dataset is subject to a 1-D Fourier transform to produce the spectrum. In Figure A.20B, the dataset is moderately damped from its first breaks to produce the spectrum. Figure A.21 is shown to highlight what the frequency-domain spectrum will look like in the Laplace-Fourier domain.

To prepare the spectra shown in Figure A.3 for FWI using TOY2DAC, select frequencies are extracted from an appropriate spectrum and saved in another file. This "data_modeling" file is then converted to binary format and exported from the Python TOY2DAC_FILES module. For each combination of maximum offset and data-damping, a new trace amplitude spectrum is generated for all OBS.

Appendix B

Starting Model Generation

B.1 Introduction

Viscoacoustic FWI requires three starting models. The most important being the velocity model, which will prevent cycle-skipping if sufficiently accurate. Further requirements include starting density and attenuation models. This appendix discusses the generation of all three starting models with a focus on the starting velocity model. This introduction first provides some background information regarding TOMO2D and Plotsec. The following section of this appendix will explore the workflow required to produce the starting velocity model. The next section will discuss all three starting models for FWI. A suite of Python codes used for these steps are provided at `https://github.com/celw10/TOMO2D_2_TOY2DAC.git`.

B.1.1 First Arrival Traveltime Tomography using TOMO2D

Tomographic inversion requires a set of first break picks. This study uses Plotsec to pick first breaks for TOMO2D. Plotsec was developed by John Amor at the University of British Columbia. These first break picks must be put into a format required by TOMO2D. Using Plotsec, the written TXIN_2_TOMO2D_GEOM Python module converts the first break picks from Plotsec to TOMO2D format. The data are then inverted using TOMO2D (Korenaga et al., 2000) once this pick file is correctly formatted. A multiscale approach for FAT improves the inversion result. The Python script RUNTOMO2D automatically iterates through multiscale tomographic inversion. For each inversion group, RUNTOMO2D will run one inversion and forward model. All the relevant inversion and forward modelling parameters are set within RUNTOMO2D.

B.1.2 First Break Picking Using Plotsec

This subsection provides a discussion on how to use Plotsec due to an absence of known documentation. Plotsec first requires a plotsec_rsegy.com executable file that reads the SEGY seismic data, and then re-formats the data by stripping the SEGY headers from the data. The command plotsec_rsegy command will convert the SEGY data to Plotsec format. The plotsec_pick.com executable file then initiates an interactive window for first break picking, as shown in Figure B.1. The proper Plotsec header and data files are defined within plotsec_pick.com. Also, the input and output pick files must be defined here as well.

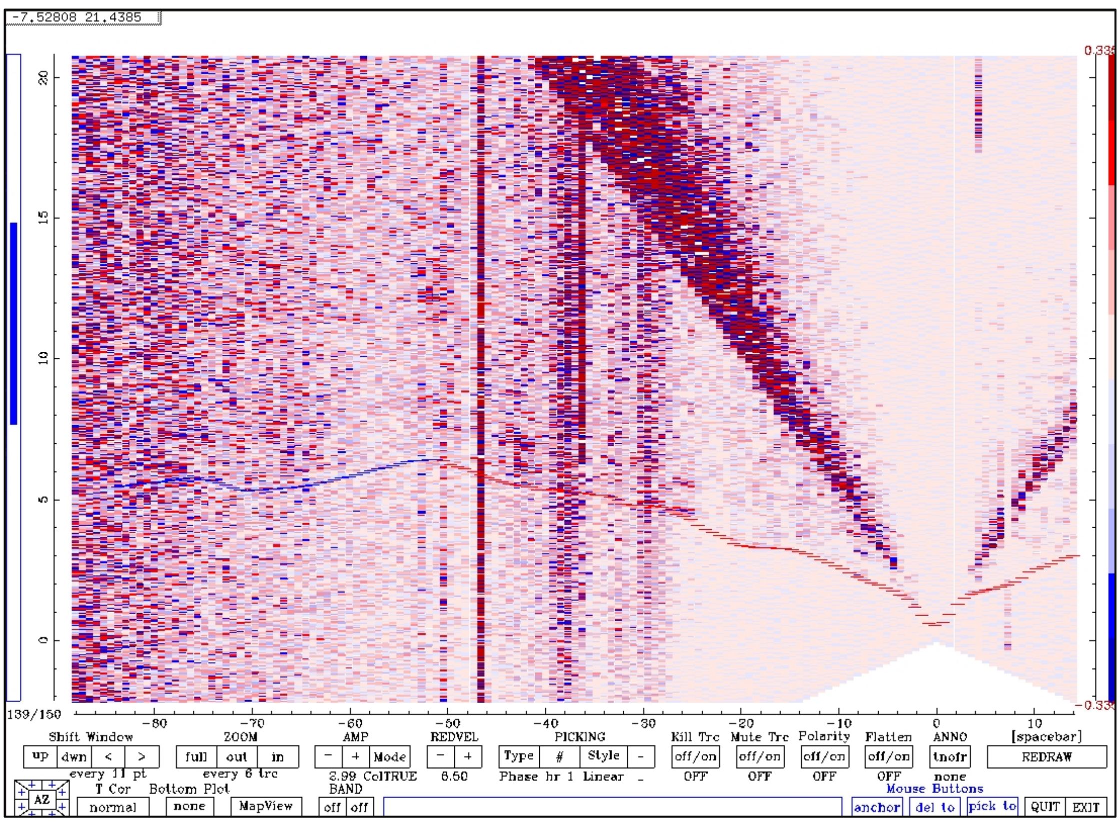

Figure B.1: An interactive window for first break picking within Plotsec.

The interactive window in Figure B.1 is opened by running the plotsec_pick command and is initially tough to navigate. The data window is controlled by shifting and zooming buttons found in the bottom left of the Plotsec interactive taskbar. A critical pitfall to cover is the "every 6 trc" text shown below the "ZOOM" mouse-buttons, meaning that the plot shows every sixth trace at the present window configuration. Adjusting any first breaks at this data zoom would only make these adjustments for every sixth trace. The "AMP" mouse-buttons control the amplitudes and colour bars to the right of "ZOOM", and the color bars are specifically controlled by the "Mode" mouse-button. Four plotting schemes are available, true-amplitude colour, true-amplitude wiggle, relative-amplitude colour, and relative-amplitude wiggle. The middle mouse-buttons beneath "PICKING" are essential for defining the type of first break pick (i.e. PmP, Pg, direct wave, etc.). The red first break picks in Figure B.1 is "phase 1", which for this example is the Pg arrival, and currently selected for manipulation. The mouse button above "hr 1" will select a different pick type, and there may be up to 10 different pick types.

When picking first breaks in this interactive window, a pick is made at one trace by left-clicking on that trace. Making a pick on a distant trace by right-clicking will interpolate between the first break picks when made from left to right. When interpolating the traces between these two picks, the style tab under "PICKING" may be set to minimum, maximum, zero-crossing, or linear. We find that interpolating any way but linearly results in a noisy set of first break picks. Pressing the space button will "redraw" the plot will then save these first break picks. The middle mouse button deletes first break picks. To delete all picks between two traces, moving left to right, one must use the left mouse button at one trace, and then the middle mouse button at the other. Extreme caution is advised during all stages of first break picking however as there is no "undo" button (that I'm aware of) within this application.

The "Fatten" tab near the right-hand side of the taskbar in Figure B.1 is a good QC tool. It is useful to look above a first break pick to ensure that it is roughly parallel with any multiples. Following this QC phase, the user must choose the "EXIT" button in the bottom right-hand corner, not the "QUIT" button. The "QUIT" button will not save the first break picks in the project, but the EXIT button will.

Exiting this interactive window will generate a file containing a new set of first break picks. The plotsec_amppk.com executable file now formats these new first break

Table B.1: Pick uncertainties are assigned based on a computed signal to noise ratio (SNR) within Plotsec.

Computed SNR	20	4	2	1.8	1.6	1.4	1.2
Pick Uncertainty (s)	0.030	0.045	0.060	0.080	0.105	0.130	0.160

picks. The Plotsec seismic data, headers, and pick files are all defined in the plotsec_amppk.com executable file. The computation of pick uncertainties requires the definition of some constants as well. Signal to noise ratios are computed 0.25 seconds above and below the first break picks in this study, and are assigned a pick uncertainty in time. The uncertainties for this summary are shown in Table B.1, and are the same as those used by Welford et al. (2015a). The plotsec_amppk command will then generate a first break pick file. Two different pick files are required, one with respect to offset and another with respect to the OBS model location. This shift is added or removed according to the "DORIG" parameter in the plotsec_amppk.com executable file, where the shift is just the position of the OBS in model space.

Each OBS will have a first break picking file generated through Plotsec. These first break pick files are placed in a shared folder, and sequentially labelled (i.e. 01, 02, 03...). To properly consolidate these first break picks into a single file, the Linux command "cat * > *output*" is used. These first break picks can now be put into proper TOMO2D format using the TXIN_2_TOMO2D_GEOM Python module.

B.2 Starting Model Generation

Following numerous failed attempts to obtain a starting velocity model, the workflow in Figure B.2 resulted in the generation of a sufficiently accurate starting velocity model. This workflow shows that any starting FWI model must be smooth and geologically reasonable, given the results of Welford et al. (2015a). Any recovered model should generally resemble their model, and if it does not, we need sound reasoning justifying the observed differences. TOMO2D will often juxtapose cells with a significant velocity contrast when there are issues with either the inversion parameters or first break picks. If the output model is geologically reasonable and smooth, an assessment of its accuracy takes place with a focus on cycle-skipping. If at any point the model is deemed to be not geologically accurate, not smooth enough, or does not adequately mitigate the risk of cycle-skipping, the TOMO2D parameters

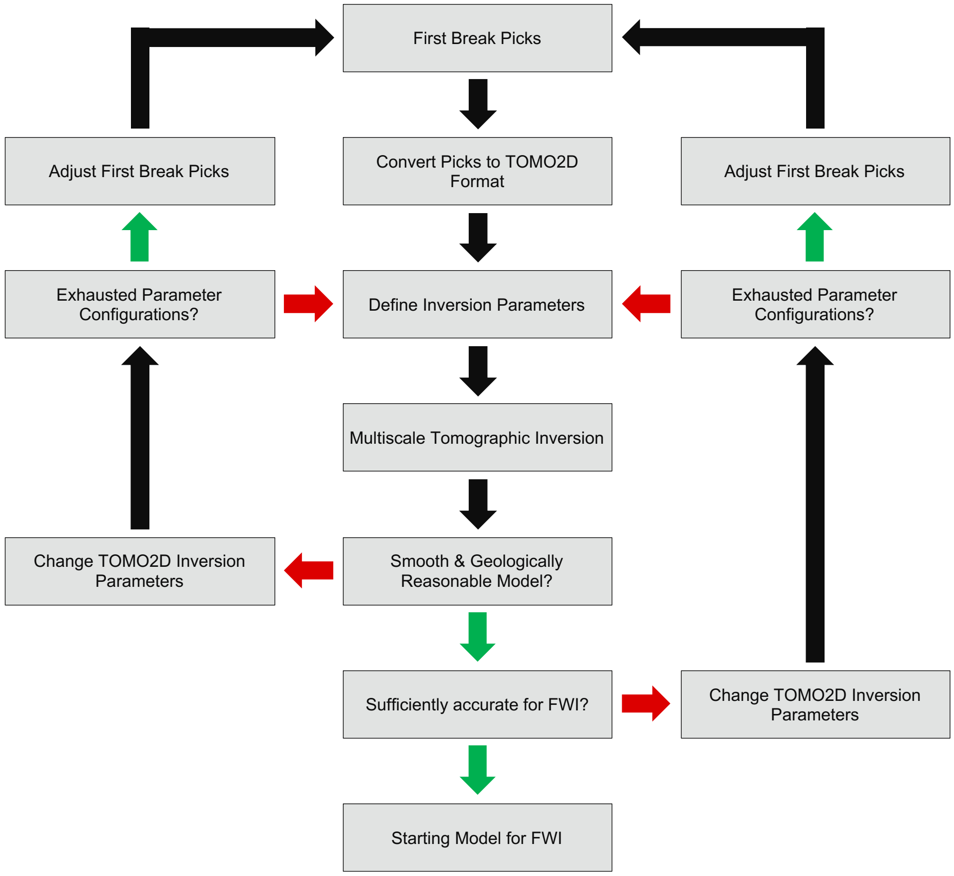

Figure B.2: A proposed workflow to build a starting velocity model for FWI that ensures it is geologically reasonable, smooth, and accurate. The workflow is an iterative process that aims at continually refining first-break picks and TOMO2D inversion parameters. Green arrows show the workflow for a "yes" answer to the decision points, and red arrows show the workflow for a "no" answer.

are first re-assessed. Following multiple inversion and forward modelling parameter re-assessments and continued inadequacy, problematic areas will have their first break picks adjusted. These adjustments will focus on OBS gathers, or regions within an OBS gather where the cycle-skipping criteria are not met.

B.2.1 Starting Model for First Arrival Traveltime Tomography

A 1-D velocity model that is the horizontal average of the final model from Welford et al. (2015a) is sufficient for a tomographic starting model. There are not any formal

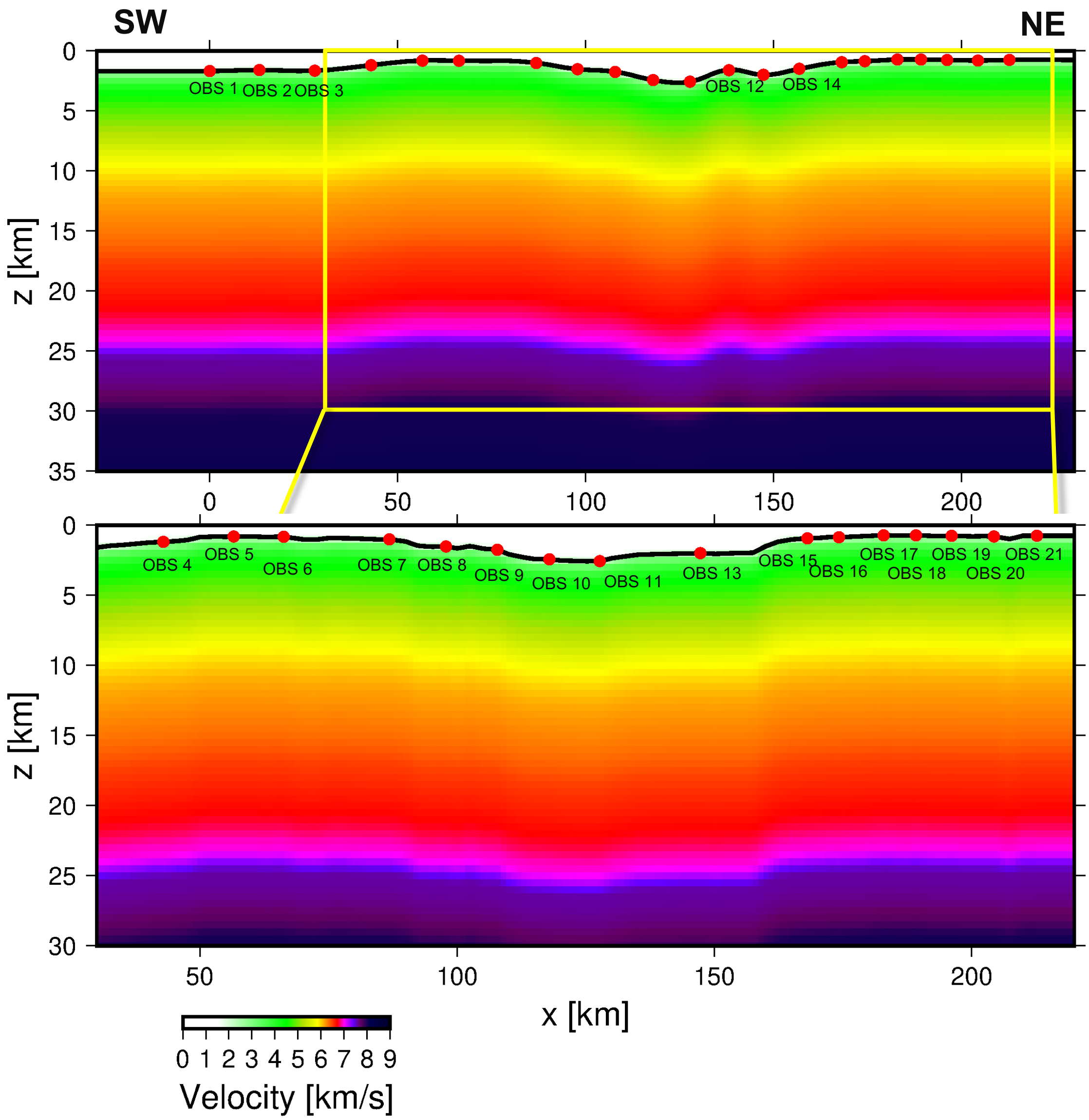

Figure B.3: A starting model encompassing the entire model space used by Welford et al. (2015a) is shown cropped and re-gridded. Five OBS are removed from the dataset as well.

structural restrictions imposed by TOMO2D on how the nodes for this model should be setup. However, these restrictions do exist within TOY2DAC. These starting FAT models are adjusted to be as close to a TOY2DAC model as possible, as TOY2DAC requires a constant node spacing in the x- and z-dimensions. FWI also requires a dense enough node spacing to avoid aliasing higher frequency waveforms.

In Figure B.3, we show the starting model with the model dimensions from Welford

Table B.2: A summary of the configurations for each of the three tomographic starting models. A variable x- and z-dimension mesh spacing is denoted by *.

Model Characteristics	Starting FAT Model One	Starting FAT Model Two	Starting FAT Model Three
Nodes (x, z)	301, 71	401, 141	951, 151
Mesh Spacing [m] (x, z)	866*, 500*	500, 250	200, 200
Model Size [km] (x)	-30-230	25-225	30-220
Model Size [km] (z)	0-35	0-35	0-30
Number of OBS	21	16	16
Interpolation Points for Bathymetric Profile	21 OBS	21 OBS	16 OBS and a 2.5 km spaced bathymetric profile

et al. (2015a). In this figure we also show a cropped version of this model, which is the final starting model for FAT. An important distinction between the two models is that the model dimensions from Welford et al. (2015a) include all 21 OBS, whereas the final starting model for FAT only includes 16 OBS. From Figure B.3, the starting model with the model dimensions from Welford et al. (2015a) will be referred to as *Starting FAT Model One*, and the final starting model for FAT will be referred to as *Starting FAT Model Three*. *Starting FAT Model Two* is an intermediate starting FAT model that exists between the two. Table B.2 summarizes the number of nodes, mesh spacing, model size, number of OBS, and bathymetric profile which distinguishes each of these starting FAT models. This table demonstrates that from starting FAT model one to three, the side of the model is decreasing, the node spacing is decreasing and becoming regular, the number of OBS included in the tomographic inversion decreases, and generating the bathymetric profile incorporates additional information.

The quality of the starting FAT model improved throughout this study. Starting FAT model one had a highly inconvenient node spacing as denoted in Table B.2, and it included all 21 OBS. Dramatic changes in the results follow once starting FAT model two is built. This improvement is largely due to the omission of the 5 OBS discussed in subsection A.2, and shown in Table B.2. In particular, the removal of OBS 12 and 14 resulted in the removal of some spurious velocity updates. Using a more accurate bathymetric profile in Starting FAT Model Three improves the FWI updates near-surface. The previous bathymetric profile for Starting FAT Model Two is interpolated between the 21 OBS, the removal of some OBS from the dataset. Taking time to configure the starting model for FAT properly would have saved time during the starting velocity model building procedure.

B.2.2 Iterative Tomographic Model Building

FAT initially uses the first break picks from Welford et al. (2015a) before iteratively updating the picks. These first break picks need to be carefully adjusted to meet the requirements for FWI. This means that gaps left in the picks from Welford et al. (2015a) must be filled, and the first break picks for FWI need to be closer to the true first break. In total, we preformed six iterations of first break pick adjustments using Plotsec. This is in conjunction with three different starting FAT models (Subsection B.2.1), and numerous TOMO2D inversion parameter adjustments. This iterative starting model building process is summarized in Table B.3.

In Table B.3, Pick 1 are the first break picks from Welford et al. (2015a), and Pick 7 are the final first break picks made in this study. Based on this study, we make some remarks about the TOMO2D parameters:

- The velocity correlation file is the most influential parameter.
- With this dataset, a smaller velocity correlation file generally outperformed a larger velocity correlation file. Other studies, however, are successful with a rather large velocity correlation size (Watremez et al., 2015).
- Adjusting the velocity correlation file and velocity smoothing is the best way to make meaningful impacts on model accuracy.
- Increasing the size of the forward star (for ray path estimation) or decreasing the segment length provides a more accurate result but is more computationally expensive.
- A larger number of interpolation points per segment provides a smoother result.
- A highly accurate inverted model can still be obtained with lower LSQR, Conjugate Gradient, and Brent Minimization tolerances. Lowering such tolerances results in a smoother model.
- Minimal changes occur when adjusting the depth weighting parameter and the depth smoothing parameter.
- A maximum velocity perturbation percentage was included during the earliest inversions to prevent large velocity jumps. It is unclear if this parameter is still required.

A summarizing figure for picks 1-7 is shown in the figures labeled in Table B.3. OBS 15 is shown in these images and is processed for first break picking, not FWI (see Section A.3). A few comments for each of these figures are listed below:

- Tomographic model 22 is produced using pick 1 and shown in Figure B.4. Pick 1 is the first break picks from Welford et al. (2015a). The goal of this stage was to figure out an initial set of TOMO2D parameters that could successfully invert the dataset and recover similar velocity anomalies as interpreted by Welford et al. (2015a).

- Tomographic model 37 is produced using pick 2 and shown in Figure B.5. Pick 2 is the first of six iterative first break pick updates. Pick 2 focuses on filling in the regions void of picks with poor SNR, and extending these first break picks to areas where Welford et al. (2015a) interpreted wide-angle reflections. The starting model assessment for model 37 shows a regression in model accuracy in comparison to model 22.

- Tomographic model 54 is produced using pick 3 and shown in Figure B.6. Now that gaps in the first break picks are filled, pick 3 attempts to move the picks in regions of low SNR governed by the results of forward modelling. If the forward modelling result suggests that a first break pick is too early, this pick is moved to a later time. The starting model assessment suggests that model 54 is more accurate than model 37, but the velocity model itself is quite geologically unreasonable.

- Tomographic model 56 is produced using pick 4 and shown in Figure B.7. Referring to Table B.2.2, there are actually no TOMO2D parameter adjustments from Model 54 to 56. Some poor first break picks are adjusted instead, resulting in a more geological and accurate starting model than model 54.

- Tomographic model 57 is produced using pick 5 and shown in Figure B.8. Again, no TOMO2D parameter adjustments have occurred since Model 54 (Table B.3). Some bad first break picks are removed for pick 5. An obvious example is tomographic station 19. There are subtle improvements noticeable in the smoothness of the model, but a prominent velocity feature at x=135 km is not sufficiently smooth.

- Tomographic model 66 is produced using pick 6 and shown in Figure B.9. Starting FAT Model Two is finally used for this tomographic inversion (Table B.2). The removal of OBS Binky and Kreig in the middle of the model seemed to correspond with a smoother model near x=135 km. The model, however, is considerably smoother throughout and is in general agreement with the velocity model from Welford et al. (2015a). Also, a starting model assessment suggests that the starting model is sufficiently accurate at 2.0 Hz to avoid cycle-skipping.

- Tomographic model 110 is produced using pick 7 and shown in Figure B.10. Starting FAT Model Three is used for this tomographic inversion (Table B.2). A few issues with pick 6 are cleaned up. Namely, the removal of some uncertain picks resulting in inaccuracies shown by the starting model assessments. Of more importance, first break picks are zero-phase as many of the picks were previously minimum phase picks (recall zero-phase processing in Subsection A.3.3). The tolerances used in this inversion are after Watremez et al. (2015).

Although it perhaps took a little longer than it should, FAT successfully provides a starting model suitable for FWI. The importance of the starting FAT model is much greater than anticipated, and it would have saved time to get this starting model right earlier. The iterative tomographic model building workflow used is successfully able to move away from the poor Model 22 (Figure B.4) toward the optimal Model 110 (Figure B.10). A combination of TOMO2D parameter and first break picking adjustments is required to achieve this. The FWI starting model, Model 110, is further analyzed in the following subsection.

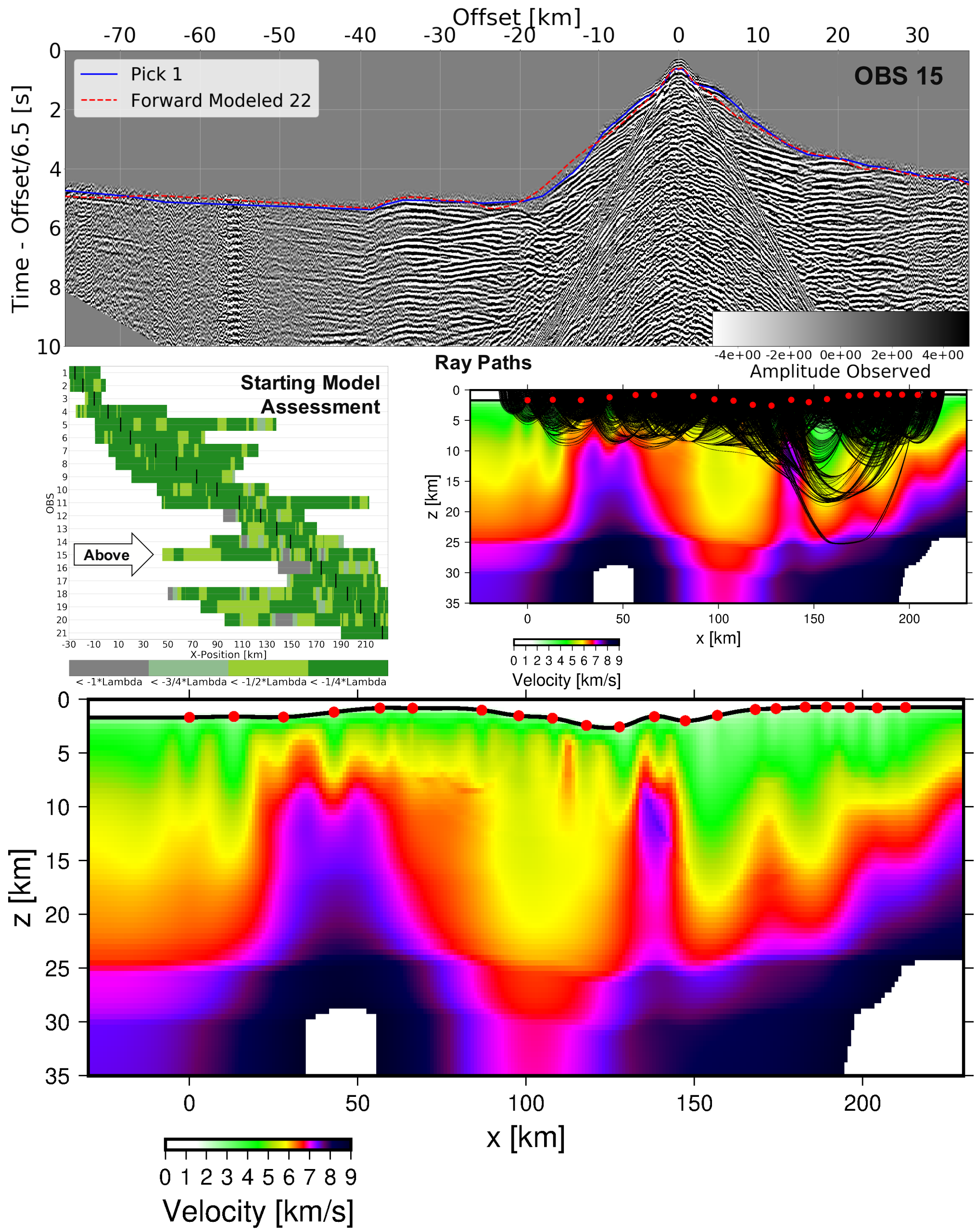

Figure B.4: The picked and forward modelled first breaks are shown for OBS 15, plotted in reduced time. A starting model assessment for a starting frequency of 2.0 Hz, and two velocity models, with and without ray paths, are shown for Tomographic model 22, which uses pick 1. Velocities of the features centred at x=50 and x=140 km are unreasonably high, and the model is not sufficiently smooth.

Table B.3: A summary of the inversion configurations for various iterations of first break picks. From left to right, the starting model for FAT, the iteration of first break picks, the figure showing this inversion, TOMO2D inversion parameters, TOMO2D forward modelling parameters, and the TOMO2D velocity correlation file for the x- and z-dimensions. The TOMO2D inversion parameters are shown as; iterations/depth weight/velocity smoothing/depth smoothing/maximum velocity perturbation percentage/LSQR tolerance. The TOMO2D forward model parameters are shown as: forward star x/forward star z/ray segment length/maximum number of interpolation points per segment/Conjugate Gradient tolerance/Brent Minimization tolerance. The parameters for the velocity correlation file in the x- and z-dimensions are both presented, in kilometres, as; range/minimum correlation length/maximum correlation length. For each new iteration of first break picks, any TOMO2D parameters that have changed from the previous are bold.

Start Mdl.	Picks	Model and Figure	TOMO2D Inversion Params.	TOMO2D Fwd. Modeling Params.	VCorr. X [km]	VCorr. Z [km]
Model One	Pick 1	Model 22 B.4	5/0.005/50/50/10%/$1e^{-4}$	5/10/2/8/$1e^{-5}$/$1e^{-5}$	-30-230/2/8	0-25/2/8
Model One	Pick 2	Model 37 B.5	5/0.005/50/50/10%/$1e^{-4}$	**10**/10/**1**/**10**/$1e^{-5}$/$1e^{-5}$	-30-230/2/8	0-25/2/8
Model One	Pick 3	Model 54 B.6	5/0.005/**80**/50/10%/$1e^{-4}$	10/10/1/8/$1e^{-5}$/$1e^{-5}$	-30-230/2/**13**	0-25/2/**7**
Model One	Pick 4	Model 56 B.7	5/0.005/80/50/10%/$1e^{-4}$	10/10/1/8/$1e^{-4}$/$1e^{-4}$	-30-230/2/13	0-25/2/7
Model One	Pick 5	Model 57 B.8	5/0.005/80/50/10%/$1e^{-4}$	10/10/1/8/$1e^{-4}$/$1e^{-4}$	-30-230/2/13	0-25/2/7
Model Two	Pick 6	Model 66 B.9	**10**/0.005/80/50/10%/$1e^{-4}$	10/10/1/8/$1e^{-4}$/$1e^{-4}$	**25-225**/2/13	0-25/2/7
Model Three	Pick 7	Model 110 B.10	**5**/**0.1**/**70**/50/10%/$1e^{-3}$	10/10/1/**10**/$5e^{-4}$/$5e^{-4}$	**30-220**/**1**/**6**	0-25/**1**/**3**

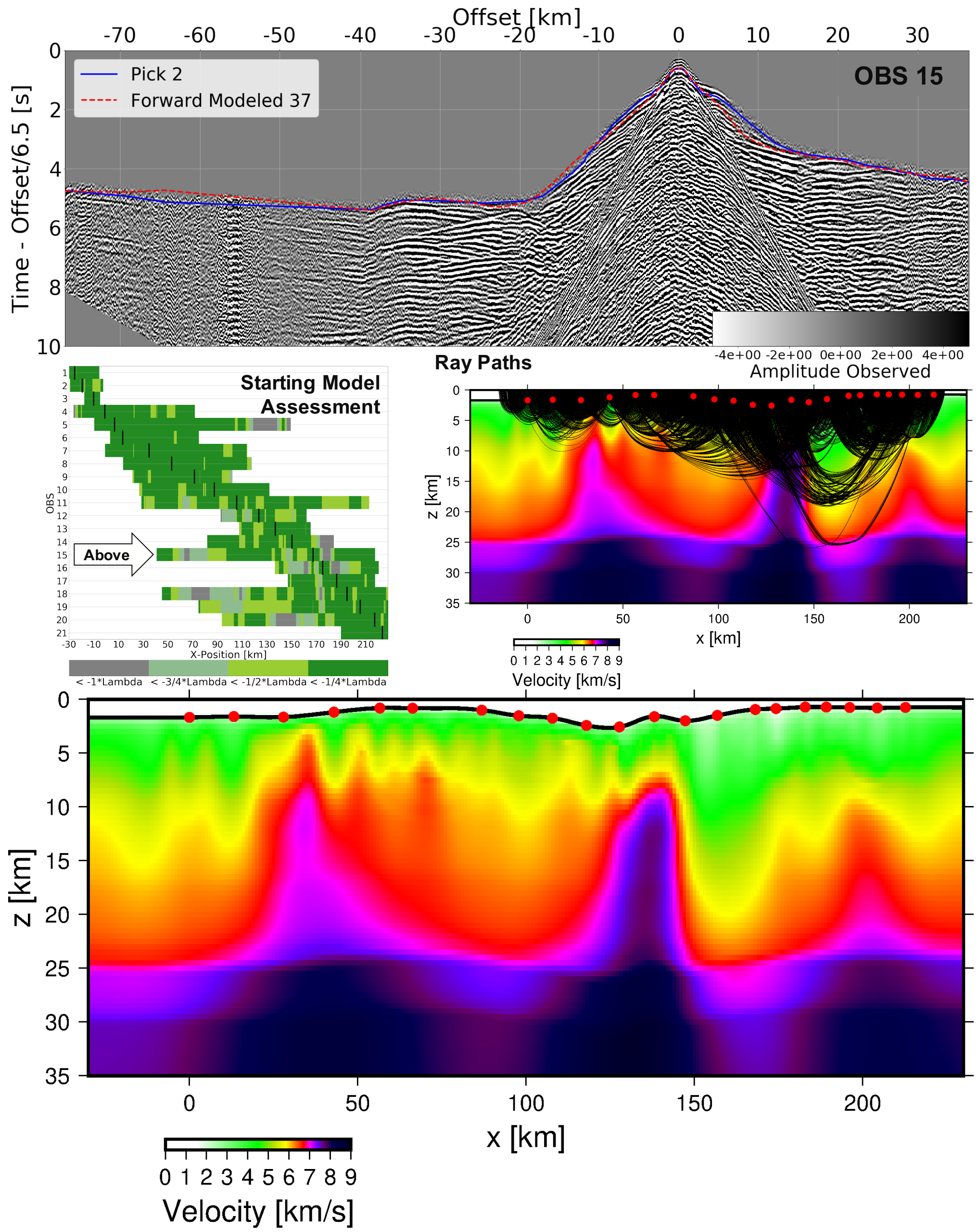

Figure B.5: The picked and forward modelled first breaks are shown for OBS 15, plotted in reduced time. A starting model assessment for a starting frequency of 2.0 Hz, and two velocity models, with and without ray paths, are shown for Tomographic model 37, which uses pick 2. This model is much smoother than the previous, but the high-velocity body at x=140 km is unreasonably high.

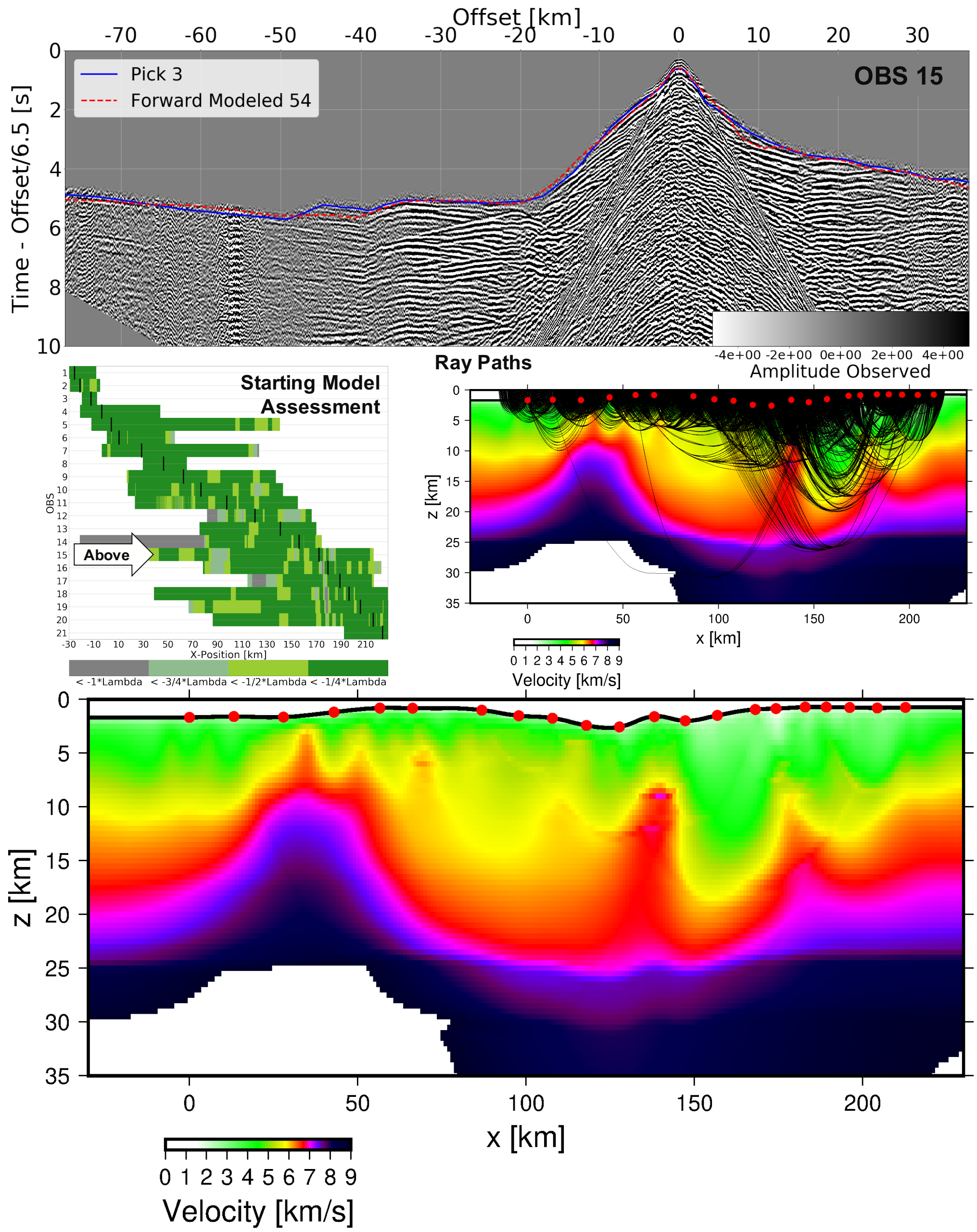

Figure B.6: The picked and forward modelled first breaks are shown for OBS 15, plotted in reduced time. A starting model assessment for a starting frequency of 2.0 Hz, and two velocity models, with and without ray paths, are shown for Tomographic model 54, which uses pick 3. This model has an unreasonably high-velocity feature centred at x=40 km, and it is not sufficiently smooth.

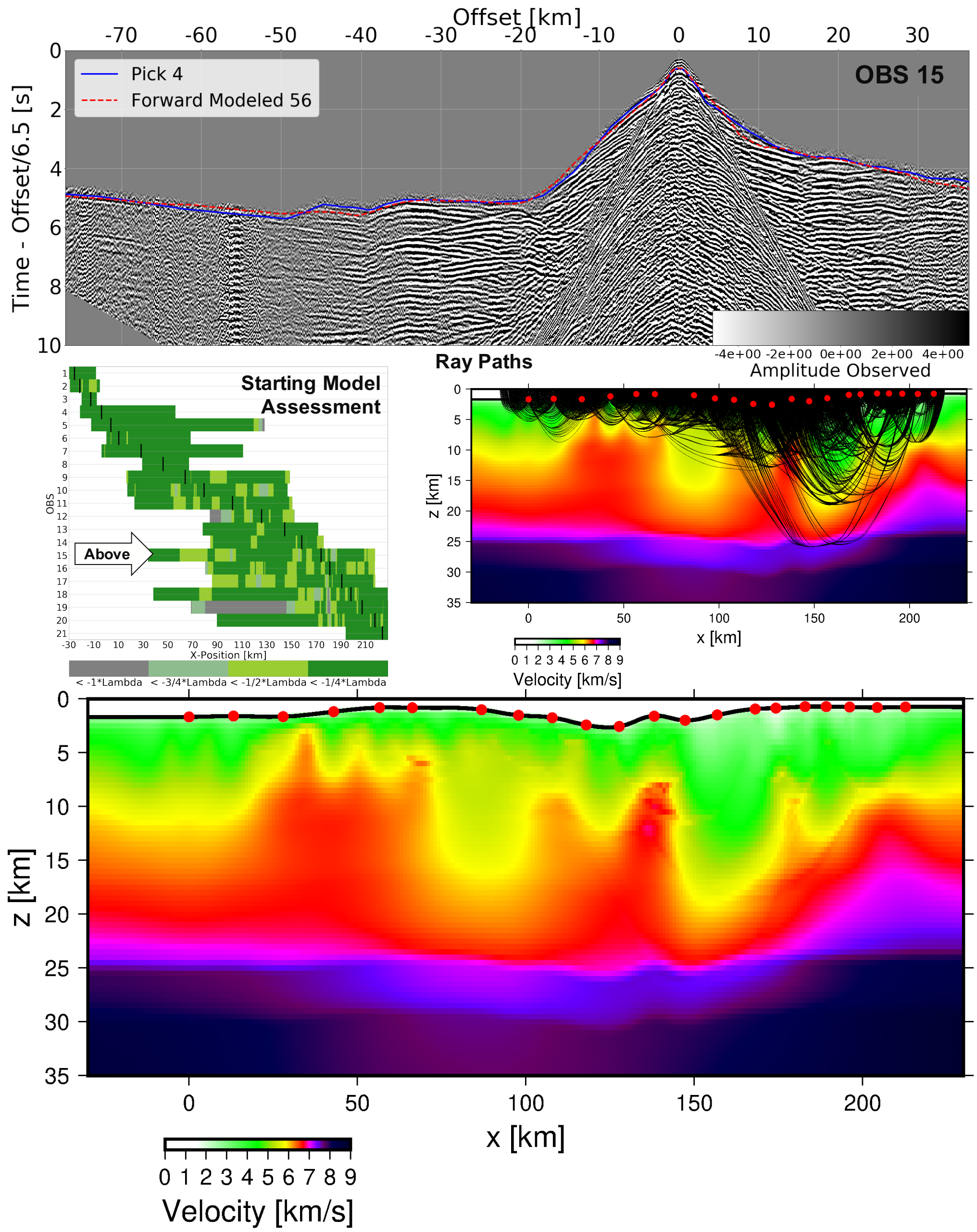

Figure B.7: The picked and forward modelled first breaks are shown for OBS 15, plotted in reduced time. A starting model assessment for a starting frequency of 2.0 Hz, and two velocity models, with and without ray paths, are shown for Tomographic model 56, which uses pick 4. Apart from one OBS, this model is reasonably accurate and geologically reasonable. However, the velocity updates are too sharp.

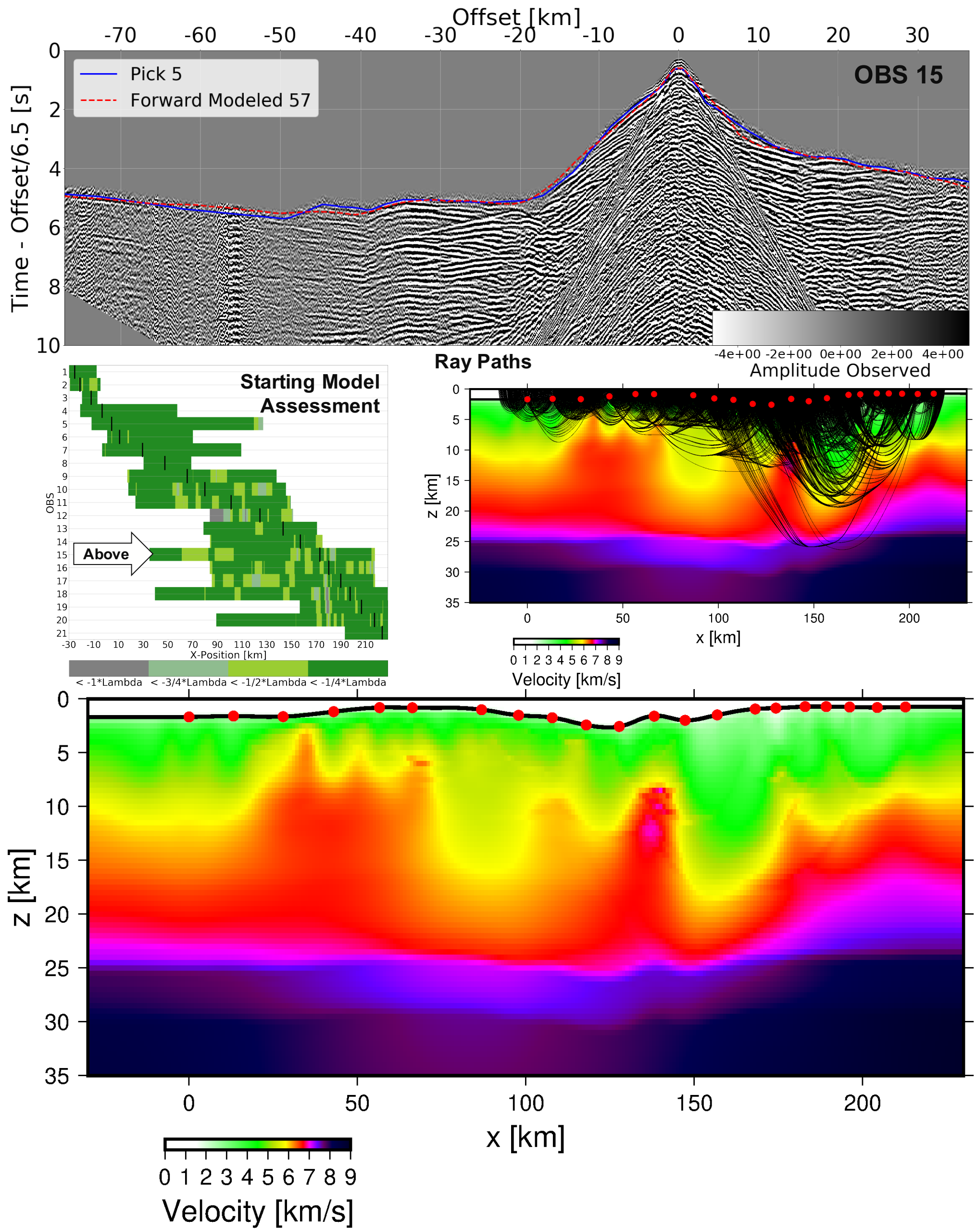

Figure B.8: The picked and forward modelled first breaks are shown for OBS 15, plotted in reduced time. A starting model assessment for a starting frequency of 2.0 Hz, and two velocity models, with and without ray paths, are shown for Tomographic model 57, which uses pick 5. This model is not smooth enough as sharp vertical contrasts are evident for some of the high-velocity updates.

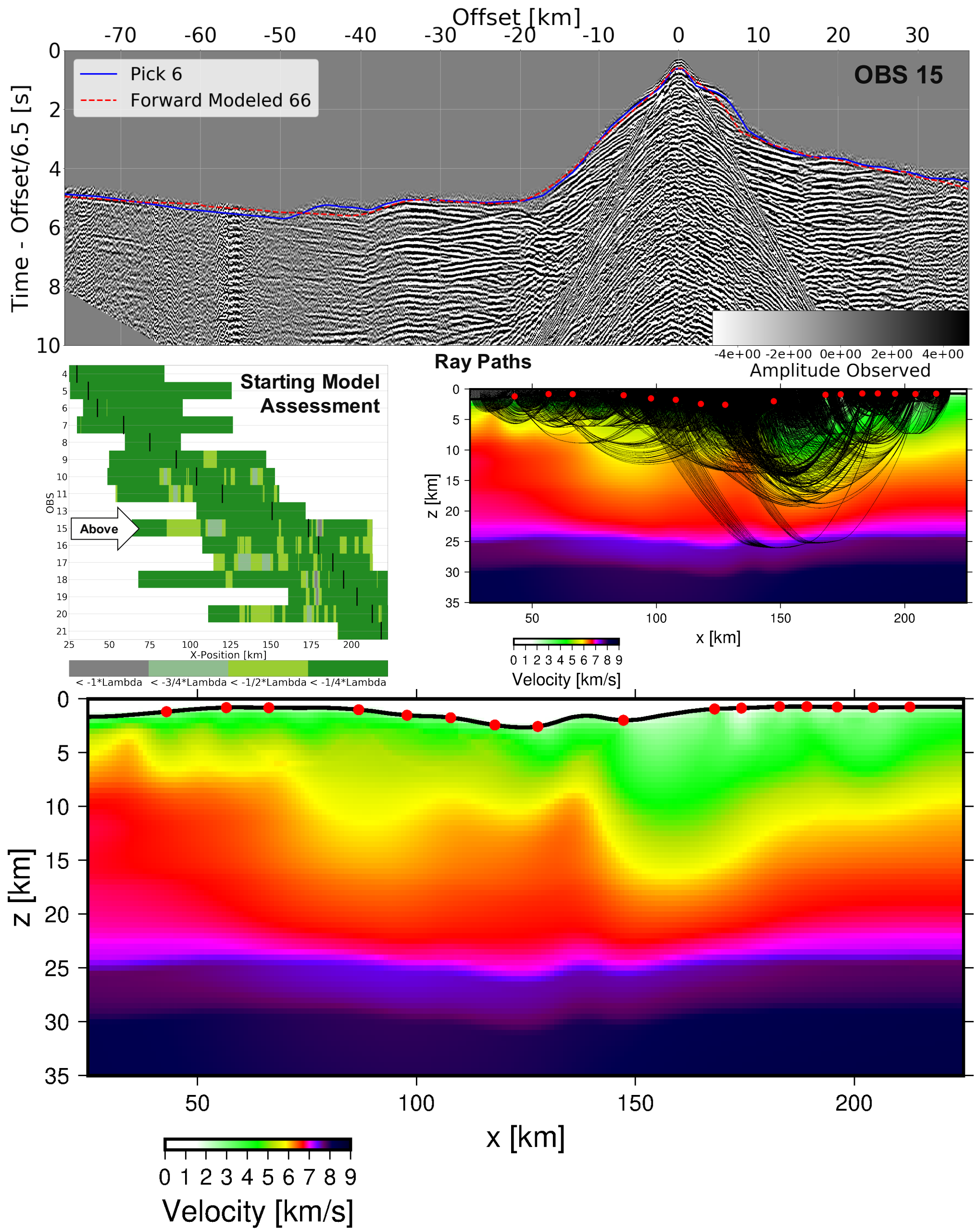

Figure B.9: The picked and forward modelled first breaks are shown for OBS 15, plotted in reduced time. A starting model assessment for a starting frequency of 2.0 Hz, and two velocity models, with and without ray paths, are shown for Tomographic model 66, which uses pick 6. Using starting FAT model two removes most of the sharp vertical velocity contrasts.

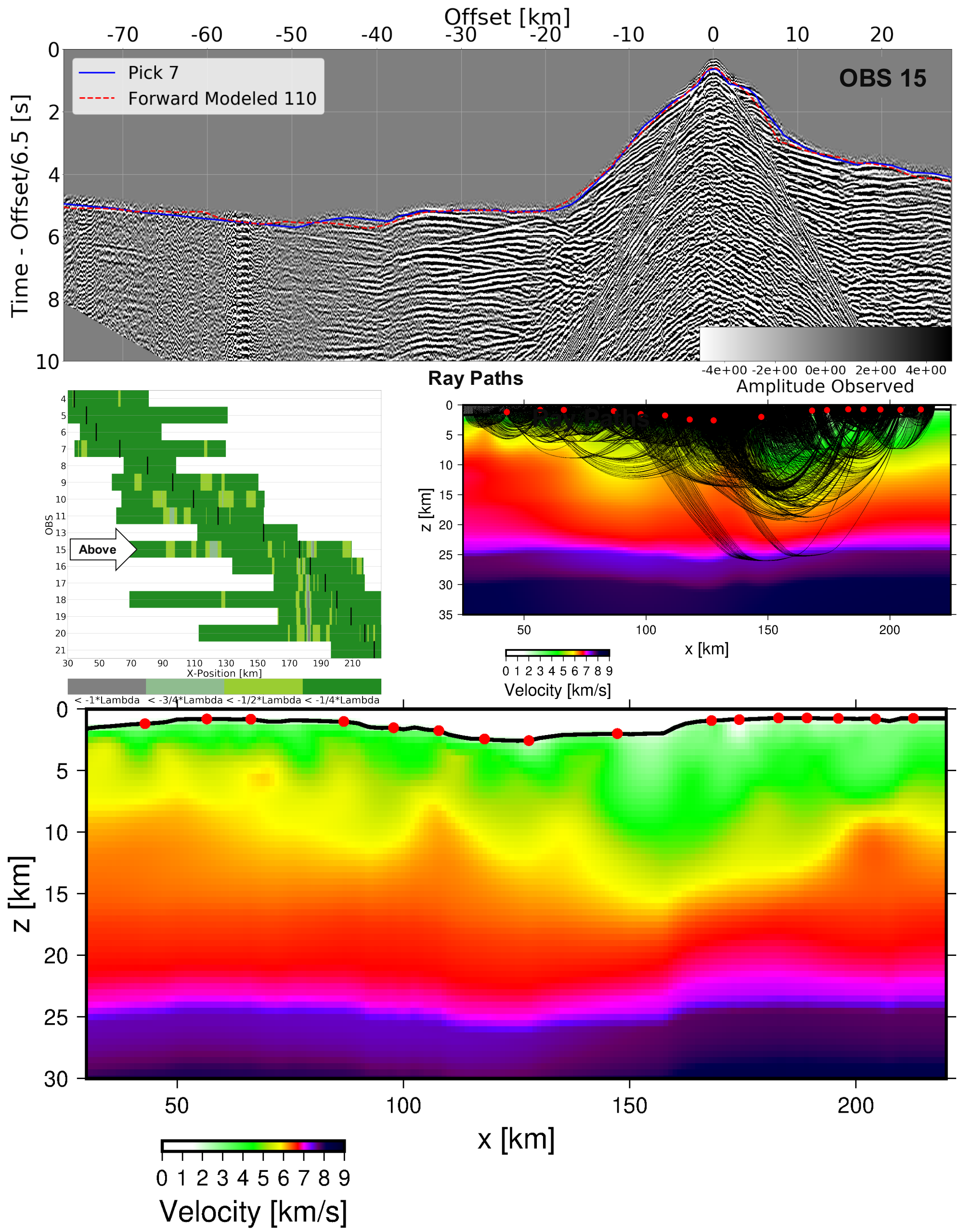

Figure B.10: The picked and forward modelled first breaks are shown for OBS 15, plotted in reduced time. A starting model assessment for a starting frequency of 2.0 Hz, and two velocity models, with and without ray paths, are shown for Tomographic model 110, which uses pick 7. This model is slightly more accurate than the previous, and exhibits more detail on the mid-crustal section while remaining smooth and geologically reasonable.

B.2.3 Final FWI Starting Model: Model 110

Tomographic Model 110 shown in Figure B.10 is selected to be the starting model for FWI. As shown by the cycle-skipping analysis in Figure B.10, it is highly accurate at 2.0 Hz. Assuming a 0.5 Hz frequency bandwidth for progressive frequency groups (Górszczyk et al., 2017), it must also be accurate at 2.5 Hz. The FWI starting model assessment image for Model 110 at 2.5 Hz is in Figure B.11A. This assessment is completed for 3.0 Hz as well in Figure B.11B. From this figure model 110 appears to be accurate enough that cycle-skipping should not be an issue, in theory. Based on Figure B.11, frequencies up to 3.0 Hz may be included in a starting frequency group with the risk of cycle-skipping still at a minimum.

There are particularities in the starting model assessment plots for 2.0 (Figure B.10), 2.5 and 3.0 Hz (Figure B.11). A prominent zone where model 110 is unable to reproduce the observed data is near x=180 km in Figure B.11. We probe this model for these poorly fit traces to ensure there are not any artifacts in the model which are responsible for producing these bad traces. Ray-paths corresponding to traces that are vulnerable to cycle-skipping and are not vulnerable to cycle-skipping at 2.5 Hz are shown in Figures B.12A and B.12B, respectfully. There are two main observations to

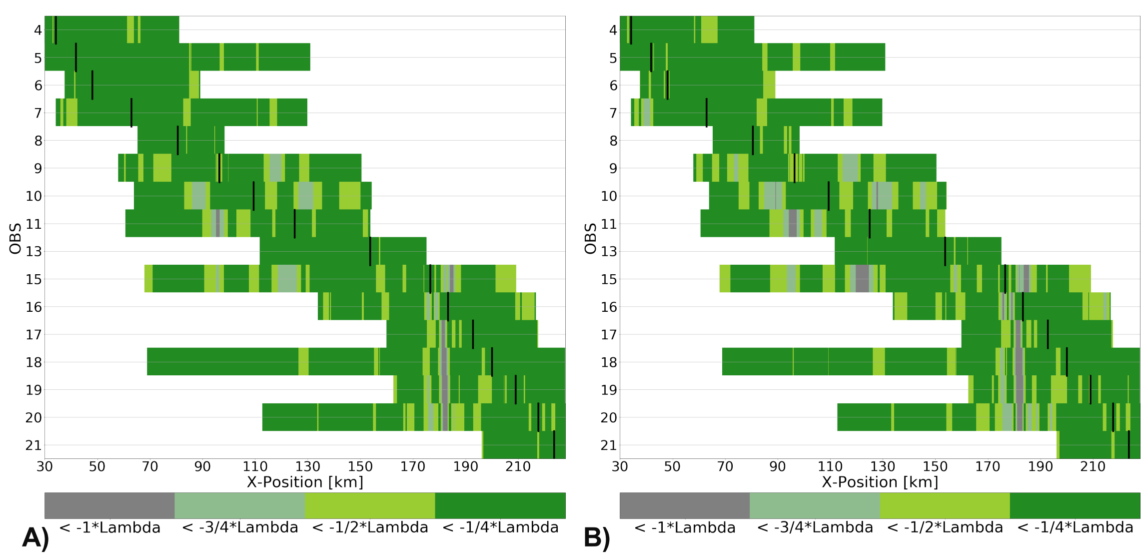

Figure B.11: (A) A starting model assessment for Model 110 is completed at 2.5 Hz, and at 3.0 Hz in (B). The grey regions correspond to areas where the traces are vulnerable to cycle-skipping, and in green, the traces are less vulnerable to cycle-skipping.

be made from this image. First, the number of traces for which the traveltime picks appear to be accurate enough to prevent convergence to a local minimum greatly outnumber the traces that are not. Secondly, the poor traces do not cluster in a single location in the subsurface that could cause their inaccuracies. There is one shallow region of potential concern near (x,z)=(175,3) km. This area is interpreted to correspond to a thin lens of evaporites toward the southern edge of the Hecataeus Rise using coincident seismic reflection data.

While Model 110 is sufficiently accurate for FWI, it is quite different from the final model from Welford et al. (2015a). The model from Welford et al. (2015a) is shown in Figure B.13A to be juxtaposed with Model 110 in Figure B.13B. Some reasons for the differences observed between the two figures are as follows:

- Welford et al. (2015a) do not use tomographic methods to produce their final model. The authors use tomographic methods as an outline for forward modelling.
- Interpreter bias is always challenging to eliminate, but forward modelling is more subjective than tomography.
- All 21 OBS are used by Welford et al. (2015a), whereas we deem only 16 of the 21 OBS suitable for FWI.
- The first break picks used to produce Model 110 are different than the first break picks used for forward modelling by Welford et al. (2015a). The requirements for FWI are much stricter than that for forward modelling, therefore the picks must be adjusted.
- No reflected phases are included during FAT performed in this study, whereas Welford et al. (2015a) models reflected phases.
- Model 110 that we develop here is better able to reproduce its first-break picks than the model of Welford et al. (2015a). This again, is a requirement for FWI. Welford et al. (2015a) cites an overall misfit of 143 ms for their model, whereas Model 110 has an overall misfit of 57.8 ms. Again, the model from Welford et al. (2015a) includes reflected phases in the computation of the RMS misfit whereas we do not.

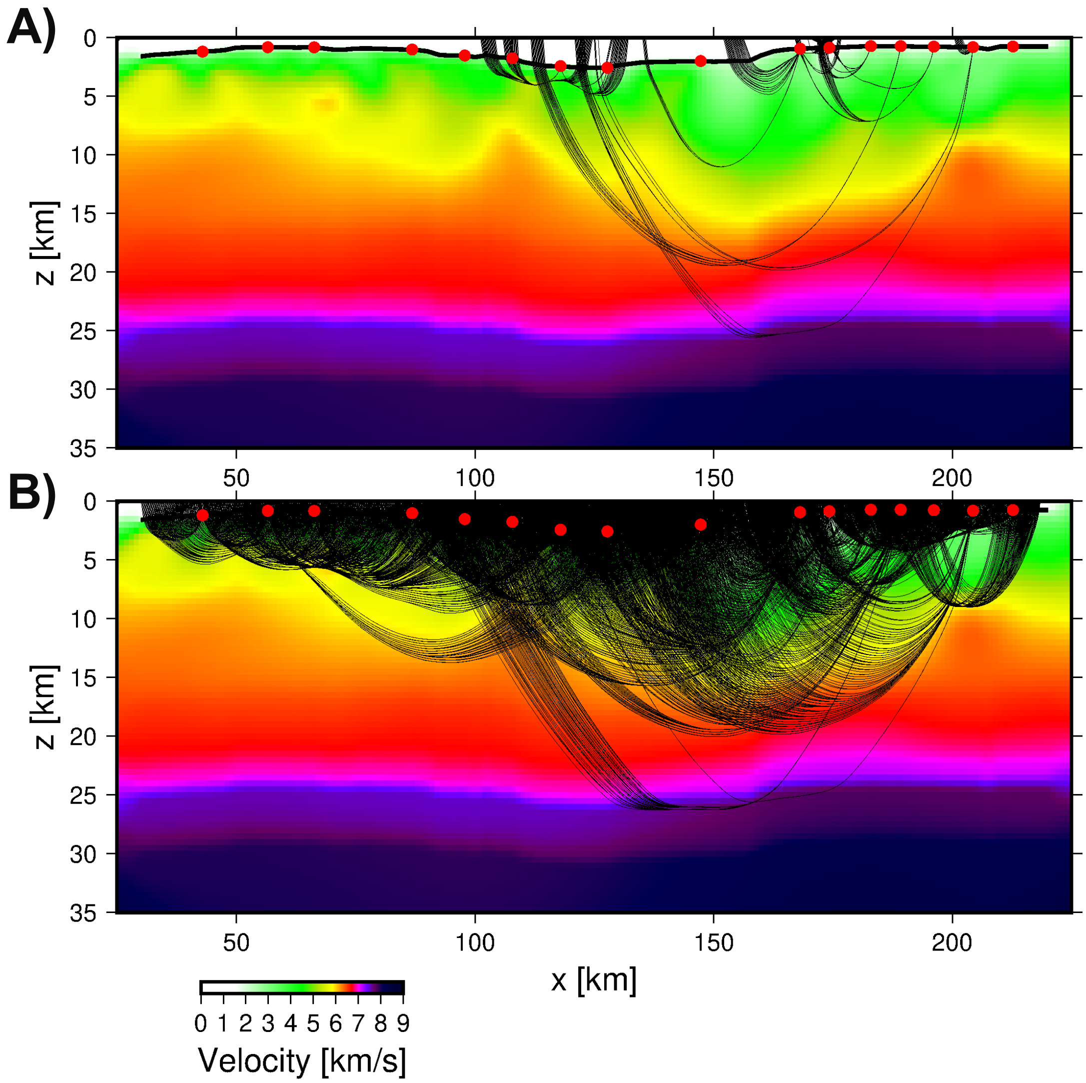

Figure B.12: Two sets of rays are traced through FWI starting Model 110. The cycle-skipping criteria is computed for a frequency of 2.5 Hz in both examples. (A) Rays that do not pass the cycle-skipping criteria, and (B) shows rays that pass the cycle-skipping criteria.

- The base of the model is unconstrained in Model 110, except for some refracted arrivals from 130-175 km in the x-dimension. Welford et al. (2015a) use gravity modelling to help constrain the position of the Moho.

There are still some similarities between this starting model and that of Welford et al. (2015a). A large sedimentary basin from x=145 km to x=200 km is present in both models. Other features are blurry in Model 110, where they are sharp in the

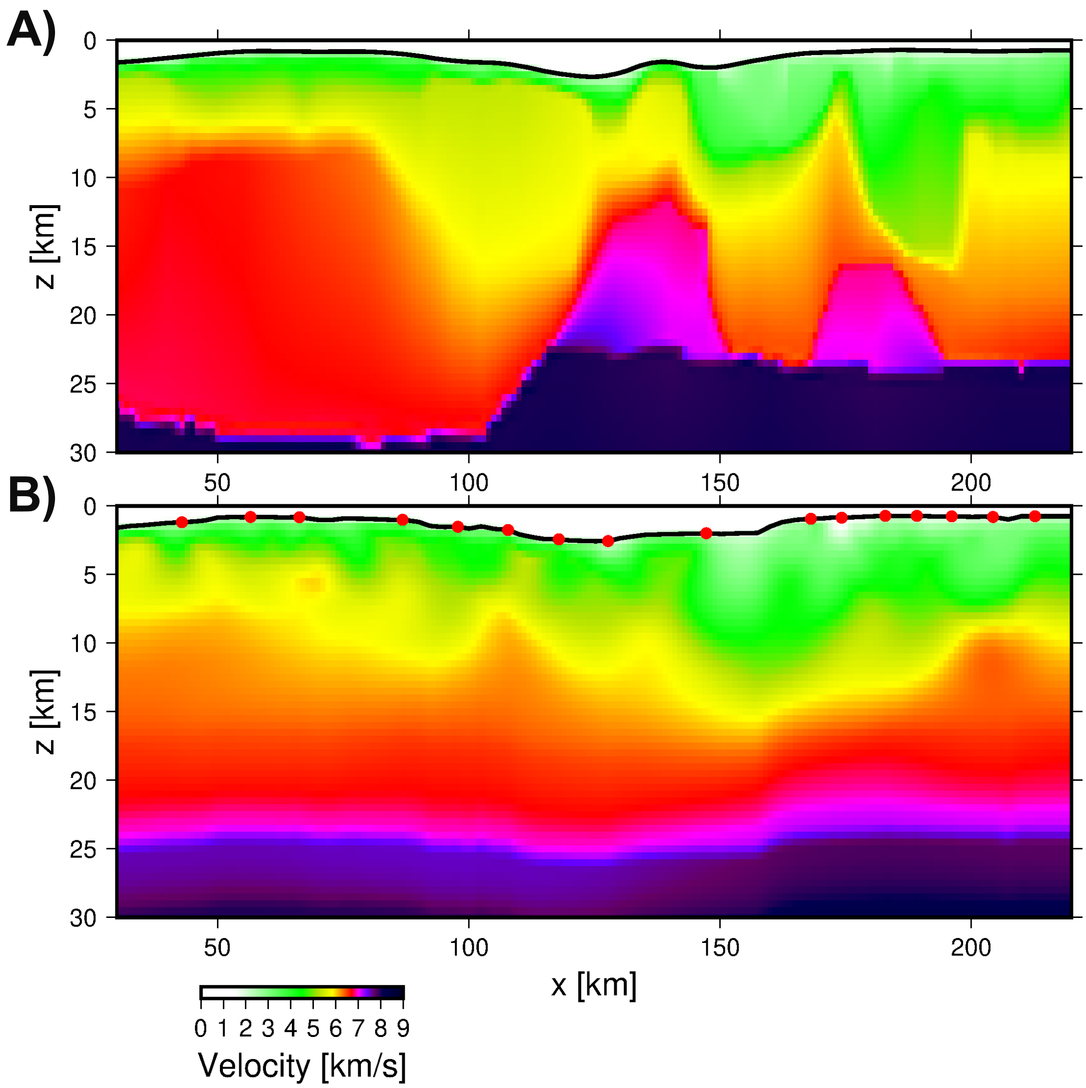

Figure B.13: (A) The final model from Welford et al. (2015a) is cropped to the dimensions of Model 110. (B) Model 110 is shown again in to allow for a direct comparison between the two models.

forward model. The high-velocity region in Model 110 from x=100 to x=140 km may correspond to the first high-velocity body interpreted by Welford et al. (2015a). A second high-velocity body interpreted at x=130 km by Welford et al. (2015a) may be present in Model 110. Velocities approach 7.0 km/s in the same location at Model 110, but a potential anomaly would not stand out from the background velocity due to an absence of ray coverage to the right. Welford et al. (2015a) model a high-velocity shallow body in Model 110 at x=135 km and z=4 km. This feature is interpreted to be an autochthonous salt wedge (Reiche et al., 2016). The lateral variability of the ECB from x=30 km to x=100km is relatively consistent in both models as well.

B.3 FWI Starting Models

An initial velocity model, discussed in the previous section, is the most difficult of the three starting models to obtain. Density and attenuation models are the other two starting models required for visco-acoustic FWI. Unlike the velocity model, the density and attenuation models remain constant during FWI. This procedure is consistent with that of Górszczyk et al. (2017). This section briefly discusses necessary alterations to the starting velocity model (Model 110 in Section B.2) for FWI, as well as the generation of density and attenuation models.

B.3.1 Velocity Model Alteration

Two significant problems are addressed with the final tomographic velocity model before its use as the starting velocity model for FWI. The first problem is related to the format of the velocity model in TOMO2D in comparison to what is required by TOY2DAC. The second problem is that the sampling interval must be increased for FWI. At least three grid points per wavelength must be available at the highest frequency and lowest velocity to be able to resolve the waveform without spatial aliasing (Hicks, 2002). A requirement of at least four grid points per wavelength is set for this study, to be on the safe side.

The MAKE_MODELS Python module addresses both problems with this velocity mesh output from TOMO2D. The velocity model output from TOMO2D is hanging from the bathymetry, meaning that the mesh excludes water velocities, only incorporating a bathymetric profile. TOY2DAC requires the velocity model to be in a matrix format with water nodes included. The TOMO2D starting velocity model must be altered to the TOY2DAC matrix format, where it can then be up-scaled to a finer node-spacing. The nearest neighbour interpolation scheme is required to ensure the retention of the starting model structures during up-scaling.

A rough calculation may suggest what node spacing should be appropriate given a maximum frequency for FWI. Assuming a minimal water velocity of 1500 m/s, and a maximum frequency of 7.5 Hz, a node spacing of 50 m is required to have four nodes per wavelength. This requires a model upscaling factor of 4, given that the spacing of the TOMO2D model is 200 m (Table B.3). An alternate option would be to upscale the mesh by a factor of 5, which would allow for data at frequencies up to 9.375 Hz to be inverted with a 40 m node spacing. Computational time greatly increases with a smaller node spacing, and an F-K analysis of the data reveals that spatial aliasing is present at approximately 8 Hz (Figure A.12). We decided to proceed with a mesh spacing of 50 m for the starting model. This starting velocity model for FWI

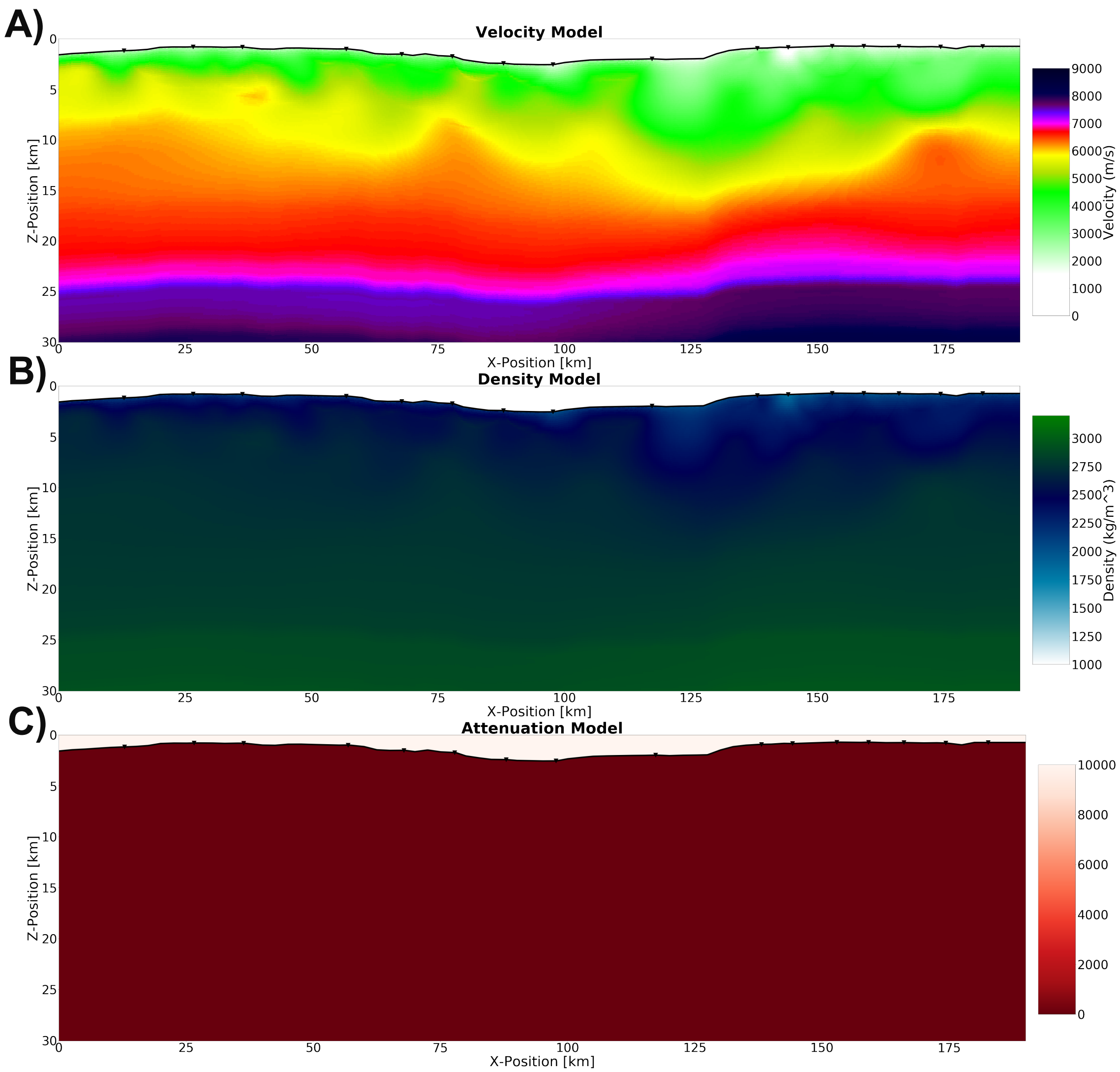

Figure B.14: (A) The final up-scaled velocity model, the final density model (B), and the final attenuation model (C).

is shown in Figure B.14A. It has 3801 nodes in the x-dimension, and 601 nodes in the z-dimension. The density and attenuation models have the same structure as this starting velocity model

B.3.2 Density Model Generation

Many different methods have been used to define a density model for FWI applications in the past. Operto et al. (2006) used a constant density model of 1.0 throughout FWI. Davy et al. (2017) and Górszczyk et al. (2017) both use Gardner's relation to formulate a relationship between the seismic velocities and subsurface rock-densities (Gardner et al., 1974). This relationship between velocity and density is,

$$\rho = 0.23v^{0.25}, \tag{B.1}$$

where ρ is the density and v is the velocity. We obtain a density model in this study using Gardner's relation to transform the seismic velocities into densities. This density model is shown in Figure B.14B. It is output using the MAKE_MODELS Python module. In this module, densities above the bathymetry are $1.0kg/m^3$. The density model is held constant throughout FWI. However, the decision to use Gardner's relation is not without controversy. Gardner's relation is derived for sedimentary rocks, and may not be completely valid for our crustal-scale of investigation. This issue is discussed in subsection 4.3.2.

B.3.3 Attenuation Model Generation

Like the density model, the attenuation model is held constant throughout FWI. We model seismic attenuation through the use of complex-valued velocities. The attenuation model should match the AVO trend of the modelled data with the observed data. Unlike density, there is no straightforward relationship between inelastic attenuation and rock properties (Winkler et al., 1979).

We generate a simplistic attenuation model following the technique of Górszczyk et al. (2017). This attenuation model consists of two values, one for the water column and one for the subsurface. We use a low attenuation value of 10,000 for the water

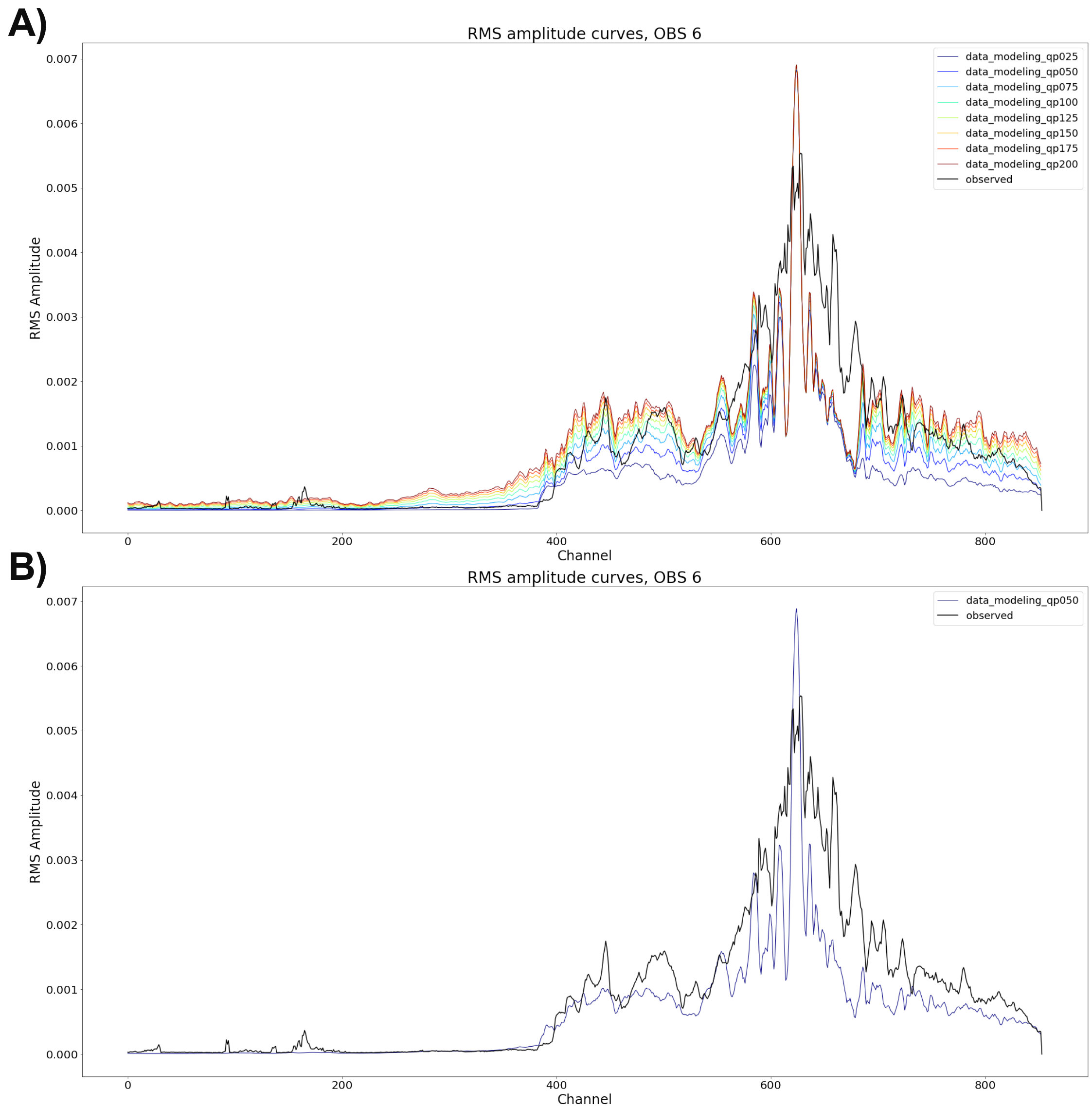

Figure B.15: (A) The RMS amplitude curves for OBS 15 for the observed dataset, and for 8 forward modeled datasets with attenuation values in the subsurface ranging from 25 to 200. (B) The optimal subsurface attenuation value, 50, is plotted along with the observed data.

column, and a high attenuation value of 50 characterizes the subsurface. In comparison, Górszczyk et al. (2017) uses a slightly higher subsurface attenuation value of 200. Even with a rough attenuation model, Kurzmann et al. (2013) are able to improve synthetic visco-acoustic FWI results. Their conclusion, however, was that a smoothed attenuation model is more effective than a homogeneous one. The tools to estimate this smoothed attenuation model are beyond the scope of this study.

We generate synthetic OBS gathers with 8 different subsurface attenuation values to select a subsurface attenuation value that best matches the observed data. The final velocity and density models shown in Figures B.14A and B.14B are used to generate these data, but the subsurface attenuation is varied. In increments of 25, the subsurface attenuation is increased from 25 to 200 in each of the eight forward models to compute RMS trace amplitudes. The RMS trace amplitudes for the observed data are multiplied by a constant to have similar RMS trace amplitudes as the modelled data. This step occurs before the amplitude correction discussed in Section A.4.

Figure B.15A shows nine RMS trace amplitude curves for observed and modelled data for OBS 15. The observed data are plotted as a thick black line, and the modelled data are plotted with hotter colours corresponding to lower attenuation values. The amplitude curves corresponding to modelled datasets with lower Q values have lower RMS amplitude trace values than those with higher Q values, as anticipated. The amplitude curve for the observed data appears to be in agreement with higher Q values at channels > 400 and in agreement with lower Q values for channels < 400. This dramatic change is likely due to the loss of the direct wave at t=20s. Also, the RMS values in the observed data are higher than the modelled data because the modelled data have no noise. Once the direct wave is lost, it is evident that an attenuation value of 50 best matches the amplitudes between the observed and predicted data. Figure B.15B shows the amplitude curve for the observed data and the amplitude curve produced with our chosen subsurface attenuation value of 50.

Appendix C

FWI Synthetic Testing

C.1 Introduction

We have five goals for synthetic testing, the first of which is to ensure the robustness of the TOY2DAC algorithm. The second is to compare the multiscale inversion strategy of Górszczyk et al. (2017) to traditional FWI techniques, and a third is to explore the effect of data sparsity. We investigate these first three goals with the synthetic Marmousi2 model (Martin et al., 2006). The fourth and fifth goals are to prove that FWI is possible at the scale and sparsity of the Eastern Mediterranean dataset and to find some optimal inversion parameters to begin a parametric study test for the real data inversion. A full-scale synthetic Eastern Mediterranean model is constructed to prove the viability of this FWI experiment and provide an initial set of inversion parameters for the real-data. This synthetic model is a cropped version of the final model from Welford et al. (2015a).

While Chapter 3 investigates data sparsity, a complete investigation is presented in this Appendix using the Marmousi model. Furthermore, the source geometry has been changed from being positioned at the seafloor to being positioned near the surface. We break up the multiscale inversion strategy used in this study, after Górszczyk et al. (2017), into four segments and progressively apply it to three different Marmousi acquisitions with 144, 18, and 2 sources. Our baseline inversion only progresses through frequencies using an overlapping frequency approach. We then introduce a progressive offset strategy along with the preconditioned l-BFGS algorithm. We then change the to a quasi-progressive structure, and lastly we apply FWI will take place in the Laplace-Fourier domain with linear data weighting.

C.1.1 Marmousi Synthetic Setup

The true Marmousi Model is shown in Figure C.1A. It is obtained through the installation of TOY2DAC and is specifically the Marmousi2 version of the model (Martin et al., 2006). The starting synthetic model is a smoothed version of the true model. Because the model comes directly from TOY2DAC, we are not able to determine the precise smoothing parameters. The density model is constant at 1000 kg/m^3, and the attenuation model is constant and small with a value of 1000. All models are composed of 681 and 141 nodes in the x- and z-dimensions spaced at 25 m, and the Marmousi models span 17 km and 3.5 km in the x- and z-dimensions, respectfully.

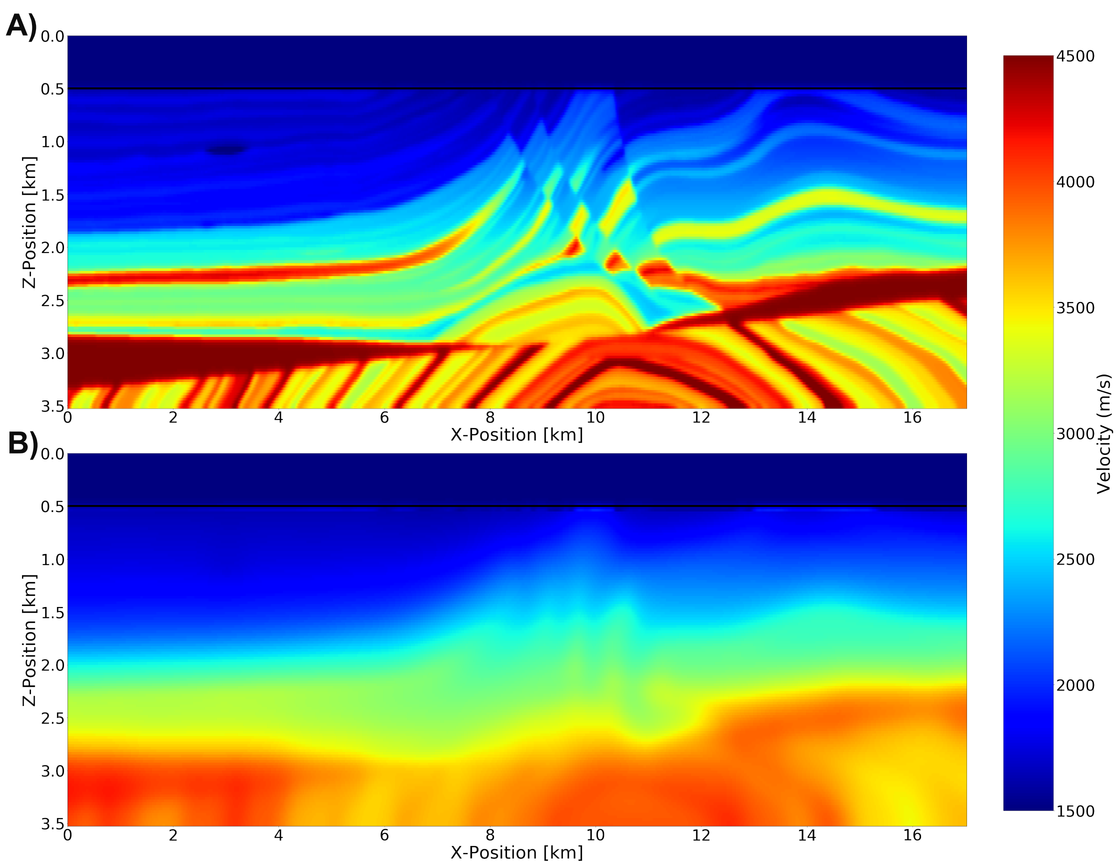

Figure C.1: (A) The true Marmousi2 velocity model, and (B) shows a starting Marmousi velocity model for FWI.

It is crucial to assess the starting model for cycle-skipping to ensure that FWI starts at an appropriate frequency. A starting model assessment at 2.5 Hz for a subset of the 144 sources is in Figure C.2A. Forward modelled arrival times through the starting and true Marmousi models yield a traveltime difference, which is then computed as a percentage of a seismic cycle, for a given frequency. The majority of the traces satisfy cycle skipping at 2.5 Hz. The first frequency group will then span from 2.0 to 2.5 Hz.

The 18 sources used to produce Figure C.2B represent an eighth of the available data. This dataset is considered to be sparse dataset for synthetic testing. Furthermore an ultra-sparse dataset is generated using only two of the 144 available sources. This ultra-sparse dataset is significant because the two sources are spaced 10 km apart, which resembles the average OBS spacing for the real Eastern Mediterranean data acquisition. For each shot in the Marmousi model, there are 660 receivers.

C.1.2 Synthetic Eastern Mediterranean Model

Chapter 3 explains how the synthetic Eastern Mediterranean model is setup. In essence, it is geometrically a replication of the real-data Eastern Mediterranean inversion with the true model being the final model from Welford et al. (2015a), and the starting model being a smoothed version of the truth.

The full-scale synthetic inversion begins at 0.25 Hz. To further justify this starting frequency, a starting model assessment for FWI is shown for 0.75 Hz in Figure C.2B. There is a risk of cycle skipping even at this low frequency due to the high degree of smoothing we use to generate the starting velocity model. As the synthetic Eastern Mediterranean inversion represents an idealized scenario, low frequencies between 0.25 and 0.75 Hz will compose the starting frequency group in this inversion.

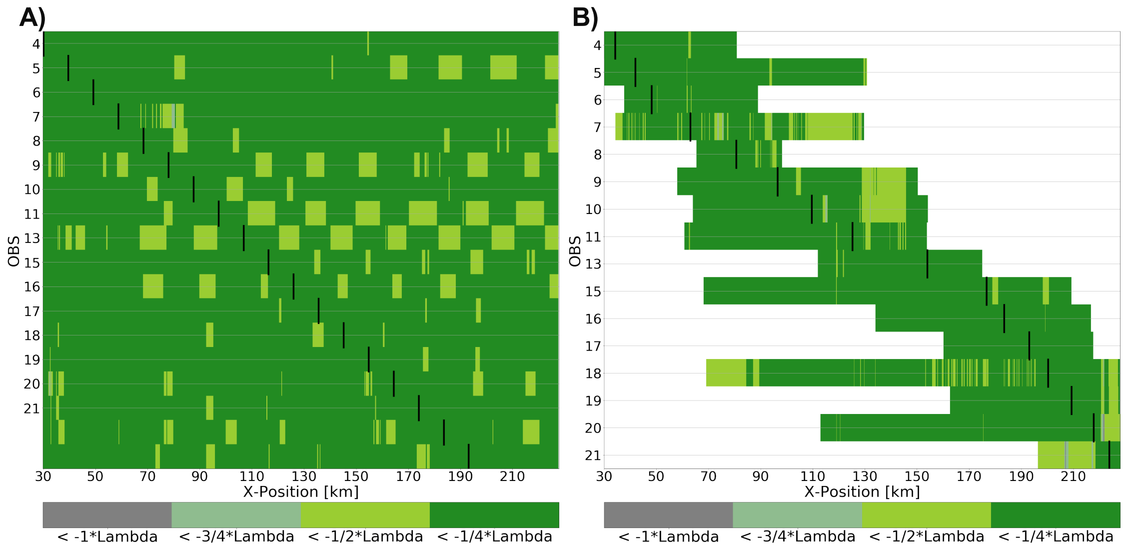

Figure C.2: FWI starting model assessments for both synthetic models. (A) Assesses the Marmousi starting model at 2.5 Hz assuming a sparse geometry of 18 OBS. The synthetic Eastern Mediterranean model is assessed at 0.75 Hz in (B).

C.1.3 FWI Computational Setup

External libraries are required by TOY2DAC to perform FWI. The reader is encouraged to read the TOY2DAC documentation for a further discussion on TOY2DAC installation and software dependencies. The TOY2DAC documentation provides URLs for various software requirements, and one may find further information there. For this study, TOY2DAC was compiled on both a UNIX personal workstation and a high-performance computing (HPC) cluster. For use on a personal UNIX workstation, TOY2DAC is compiled in sequential, but in parallel for use on the HPC cluster. Sequential and parallel installations of TOY2DAC are then different, mainly due to the software packages they require. Both C and FORTRAN compilers are required to compile versions 2.5 and 2.6 of TOY2DAC.

A MUMPS solver version 5.1.2 is used by this study to compile TOY2DAC. The MUMPS solver performs lower-upper (LU) decomposition on large sparse matrices. It requires the installation of at least METIS and SCOTCH as ordering libraries. METIS and SCOTCH versions 5.1.0 and 6.0.5 are used for the UNIX installation in sequential. For the parallel installation on TORNGAT, only PARMETIS is used, which is the minimum, as stated in TOY2DAC documentation. PARMETIS is the parallel version of METIS, and is installed on TORNGAT as version 4.0.3. The MUMPS solver also requires linear algebra packages. For the UNIX installation in sequential, BLAS and LAPACK version 3.8.0 is used. For the parallel installation on TORNGAT, MKL (INTEL) is used and is pre-installed on select TORNGAT cores. The final library required by TOY2DAC is the SEISCOPE Optimization Toolbox, which contains a suite of numerical optimization routines used by TOY2DAC (Métivier & Brossier, 2016). This study downloaded version 1.0 of the Optimization Toolbox from the same site as TOY2DAC, `https://seiscope2.osug.fr/SEISCOPE-OPTIMIZATION-TOOLBOX`.

FWI is a computationally expensive inverse problem (Virieux & Operto, 2009). Though computational time is saved by casting FWI as a frequency-domain inversion problem, large multiscale inversions are run on TORNGAT in parallel for this study. The number of model nodes has the largest impact on the computational cost, which is an issue as higher frequency inversions require a finer mesh spacing to avoid aliasing. Inversions on TORNGAT use two different nodes, one with 64 Gb RAM and 16 cores, and another with 256 Gb RAM and 24 cores. For Marmousi synthetic inversions, four separate inversions may be run on one 64 Gb node, and a maximum of two Eastern Mediterranean (synthetic and real as they have the same model sizes) may be run on one 64 Gb node. However, this is entirely dependent on the number of frequencies included in the inversion, as memory requirements increase with the number of frequencies in each group (Brossier et al., 2009). Cores are evenly distributed for each inversion to run in parallel. Marmousi inversions typically run on the order of hours, while Eastern Mediterranean inversions may take multiple days to run.

Transforming the frequency-domain data to the time-domain for visualization requires an inverse Fourier transform. The frequency-domain data are tapered for such an inverse transform, meaning that frequencies higher than the principle frequencies of interest are forward modelled for the purpose of being tapered. A mesh with a dense node spacing is then required. The 256 Gb node is required to forward model the data, which takes multiple days to complete. Only one forward model may be run on this node at once due to a dense mesh of 20 m or 25 m node spacing for a 190*30 km mesh.

Table C.1: A summary of the constant TOY2DAC inputs for the Marmousi inversions.

FWI Input	
Source Estimation	Mean source over all shots
Deadzone	None, 0 m
l-BFGS Memory Parameter	10
Convergence Criteria Model Update	$1e^{-4}$
Velocity Bounds	1000 m/s & 5000 m/s
Data Weighting	All weighted 1.0
FDFD Input	
PML Coefficients and Number of PML	90 & 10
Hicks Interpolation	On
Free Surface	On
Source Type	Explosive
Receiver Type	Pressure
MUMPS Input	
Pivot Order (ICNTL7)	METIS
Memory Relaxation (ICNTL14)	50 (%)
Memory (ICNTL23)	10 Gb

C.2 Marmousi Synthetic Tests

TOY2DAC parameters for this synthetic inversion that do not change are shown in Table C.1 for frequency-domain finite-difference (FDFD), FWI, and MUMPS files. The inputs for the Frequency management and TOY2DAC input files are varied depending on the inversion performed. We provide these parameters in a table for each section when a new set of inversion parameters are defined. All inversions model isotropic materials, specified in the TOY2DAC input file, and P-wave velocity is the only recovered parameter for all inversions. All TOY2DAC parameters listed in Table C.1 are defined in TOY2DAC, Optimization Toolbox, and MUMPS documentations.

RUNTOY2DAC is an algorithm written to run multiscale TOY2DAC inversions automatically. TOY2DAC must be called for each inversion group to run. Manually adjusting the inversion parameters after each inversion group is cumbersome when the essence of the multiscale inversion strategy is a three-loop inversion. RUNTOY2DAC, written in Python, runs through three loops defined by frequency, damping, and offset where inversion parameters are changed. RUNTOY2DAC is a part of the Python software package developed for this study found at `https://github.com/celw10/TOMO2D_2_TOY2DAC.git`.

C.2.1 Configuration One: Baseline

The first synthetic FWI test increases frequency groups from low to high within the 2-12 Hz bandwidth for Marmousi inversions. Shown in Table C.2, an overlapping frequency group strategy is first used (Brossier et al., 2009). The overlapping frequency group strategy retains no lower frequencies, so the long-wavelength structure recovered by the inversion may be overprinted by short-wavelength structure (Górszczyk et al., 2017). This study uses three frequencies per group after Brossier et al. (2009), and the first frequency of each group overlaps with the last frequency in the previous group. The spacing between all frequencies is shown to be consistent at 0.5 Hz.

Table C.3 defines the varied inversion parameters for this Marmousi inversion study. At this point, only the first loop of the three-loop multiscale inversion strategy is defined, which results in a faster inversion. A non-preconditioned l-BFGS algorithm performs numerical optimization for this first synthetic study. We use five iterations of FWI per frequency group, resulting in the simplest multiscale FWI strategy we will study.

Figure C.3 shows three different inverted models with three different geometries reflecting different degrees of data sparsity. In addition, the residual velocity field for each inversion is shown. Subtracting the true velocities from the inverted model yields the residual velocity field. Therefore it is a measure of how good the inversion is, a perfect inversion would show no velocity residuals. The observations from this inversion, for three different synthetic geometries, are summarized as follows:

- More sources produce a superior inversion.

Table C.2: An overlapping frequency strategy for FWI. All frequencies are shown in Hz.

Overlapping Frequencies					
Group 1	[2.0,2.5,3.0]	Group 7	[5.0,5.5,6.0]	Group 13	[8.0,8.5,9.0]
Group 2	[2.5,3.0,3.5]	Group 8	[5.5,6.0,6.5]	Group 14	[8.5,9.0,9.5]
Group 3	[3.0,3.5,4.0]	Group 9	[6.0,6.5,7.0]	Group 15	[9.0,9.5,10.0]
Group 4	[3.5,4.0,4.5]	Group 10	[6.5,7.0,7.5]	Group 16	[9.5,10.0,10.5]
Group 5	[4.0,4.5,5.0]	Group 11	[7.0,7.5,8.0]	Group 17	[10.0,10.5,11.0]
Group 6	[4.5,5.0,5.5]	Group 12	[7.5,8.0,8.5]	Group 18	[10.5,11.0,11.5]
				Group 19	[11.0,11.5,12.0]

Table C.3: A description of the varied parameters for the Marmousi synthetic study. *The number of iterations is typically varied in loop two, but they remain constant for this synthetic example. **A non-preconditioned l-BFGS algorithm performs numerical optimization.

Inversion Parameter	Multiscale Inversion Loop	Definition
Frequency Group	Loop One	Overlapping w/o low frequencies retained
Laplace Constant	Loop Two	N/A
Iterations	Loop Two	Five iterations per frequency group*
Offset	Loop Three	N/A
Gradient Preconditioner	Loop Three	N/A**
Data Weighting	Loop Three	N/A

- The inversion with 144 sources adequately recovers the upper section of the Marmousi model with minimal velocity residuals outside of the complex structural region.

- There is a poor recovery of the lower section of the model with all datasets.

- Even with a simple overlapping frequency grouping strategy, the sparse-data inversion can adequately resolve the upper section of the Marmousi model.

- The sparse dataset appears to suffer from an acquisition footprint, vertically oriented artifacts visible in the residual velocities that align vertically with the sources. This footprint is much weaker than the example from subsection 3.5.1 where the sources are placed on the seafloor. Arrows in Figure C.3B highlight the orientation of this footprint, which appears as a positive residual velocity.

- Unlike inverting the full synthetic dataset, the sparse dataset is plagued by noise in the updated velocities that produce a "dimpled" texture in the updated and residual velocity field.

- Inverting the ultra-sparse dataset is mostly unsuccessful. Only the uppermost velocities are well-recovered, and one cannot interpret the presence of the Marmousi structure without prior knowledge of its existence.

If a dense dataset is available, a complex inversion strategy may not be required. This conclusion, however, is based upon a synthetic example, so real-data of similar scale likely require a more robust inversion strategy. The biggest surprise of this experiment was the ability of the inversion to still recover most of the Marmousi

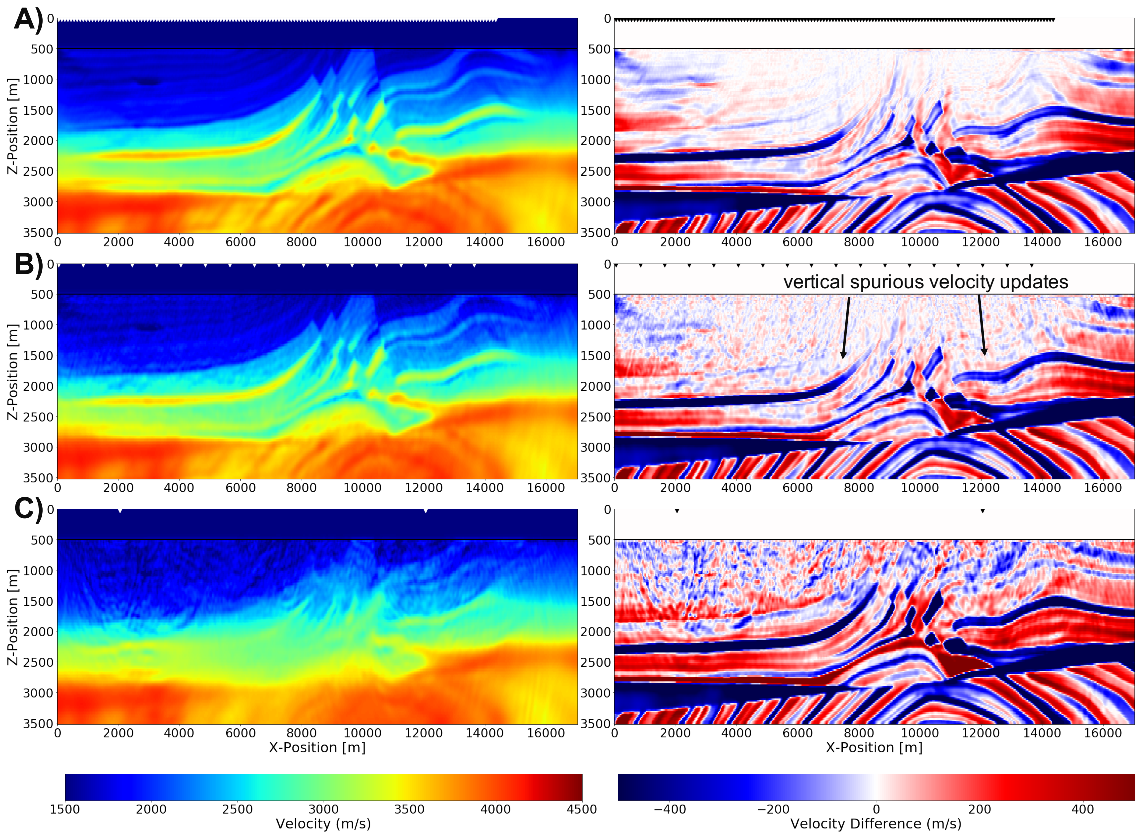

Figure C.3: Synthetic datasets with 144, 18, and 2 sources are inverted using the Marmousi2 model. The recovered velocity model is shown to the left, and the residual velocity field is shown to the right. The residual field is obtained by subtracting the true model from the inverted model. A baseline inversion is completed with 144 sources in (A), with 18 sources in (B), and with 2 sources in (C).

structure with just an eighth of the available data. Of course, using 144 sources produces a better result than using 18, but the difference is not too significant. The difference between 2 and 18 sources is significant to the point where two sources will not allow for a suitable convergence of the inverse problem.

A comparison of the results in this subsection to those shown in Figure 3.2 returns interesting observations. The dense and sparse datasets from this subsection have 144 and 18 sources, which is roughly comparable to the 170 and 17 sources used in subsection 3.5.2, respectfully. The sparse dataset in Figure 3.2A reveals a strong acquisition footprint. This observation yields two interpretations, one being that the placement of sources at the seafloor from Figure 3.2 or the surface in this example will generate different artifacts for sparse acquisitions. Additionally, running 50 iterations per frequency group has likely over-fit the data as a ”dimpled” texture is visible in the velocity residuals for Figure 3.2A. This texture is much more pronounced than in Figure C.3A when we run only 5 iterations per frequency group. However, the dense dataset in Figure 3.2C does not suffer from this over-fitting artifact. This observation leads to an important interpretation that the number of FWI iterations must be controlled for sparse datasets to avoid over-fitting.

C.2.2 Configuration Two: Maximum Offset Constraints

We now employ a strategy that progressively increases the maximum offset inverted for within the outermost frequency loop. Within this innermost offset loop, we use a preconditioned l-BFGS algorithm, where smaller preconditioner values favour deeper velocity updates from FWI (Ravaut et al., 2004; Górszczyk et al., 2017). Deeper portions of the velocity model are recovered by FWI by aligning a progressively decreasing damping values of the gradient preconditioner with progressively larger maximum offsets. The FWI inversion parameters are adjusted accordingly in table C.4. This approach should improve the poor recovery of deep structures observed in subsection C.2.1. The results for each geometry are shown in Figure C.4.

Some observations from this multiscale inversion test can be summarized as follows:

- The result of the full-data inversion has dramatically improved. The entirety of the main Marmousi model is imaged with minimal residual velocities.
- The sparse-data inversion has improved. The ”dimpled texture from the sparse-data inversion with only an overlapping frequency group strategy approach is

Table C.4: A description of the varied parameters for the Marmousi synthetic study. *The number of iterations is typically varied in loop two, however they remain constant for this synthetic example.

Inversion Parameter	Multiscale Inversion Loop	Definition
Frequency Group	Loop One	Overlapping w/o low frequencies retained
Laplace Constant	Loop Two	N/A
Iterations	Loop Two	Five iterations per maximum offset*
Offset	Loop Three	[4, 8, 12, 17] (km)
Gradient Preconditioner	Loop Three	[$1e^{-2}$, $1e^{-3}$, $1e^{-4}$, $1e^{-5}$]
Data Weighting	Loop Three	N/A

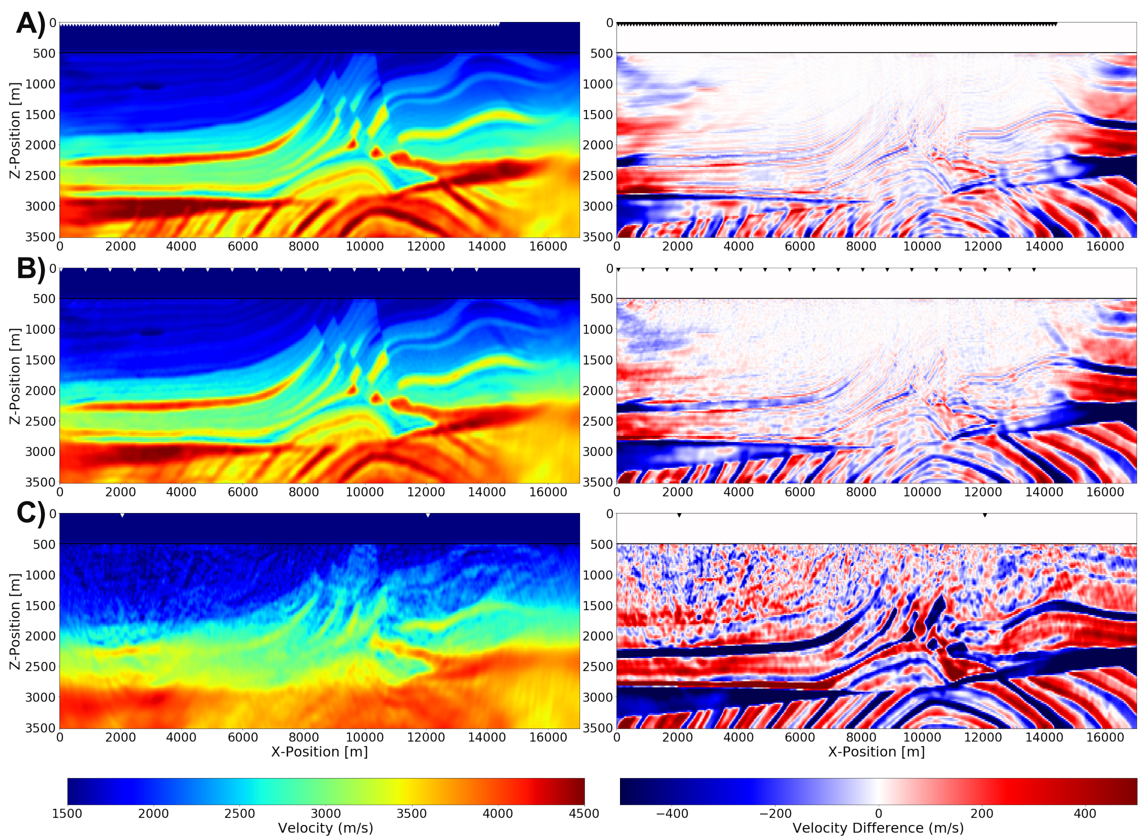

Figure C.4: Synthetic datasets with 144, 18, and 2 sources are inverted using the Marmousi2 model. The recovered velocity model is shown to the left, and the residual velocity field is shown to the right. The residual field is obtained by subtracting the true model from the inverted model. A multiscale offset inversion is completed with 144 sources in (A), with 18 sources in (B), and with 2 sources in (C).

absent with the multiscale offset strategy. The acquisition footprint shown in the residual velocities is reduced as well.

- Both the sparse and full datasets are successfully recovering deeper structure from the Marmousi model.

- The ultra-sparse-data inversion has improved as well, but the Marmousi structure is still challenging to resolve. This inversion is plagued by noisy velocity updates.

- One may begin to make out the main aspects of the Marmousi structure from the ultra-sparse-data inversion.

Performing five iterations of FWI per offset group amounts to a maximum of 20 total FWI iterations per frequency group. For the 19 frequency groups in Table C.2, a maximum of 380 iterations of FWI are available following the introduction of the offset loop. In comparison, running a maximum of five iterations per frequency group without the offset loop amounts to a maximum of 95 total iterations. The number of total iterations will grow again once the Laplace-Fourier approach is incorporated. We attribute this success to: using a superior preconditioned l-BFGS algorithm, decimating the data by maximum offset, and increasing the maximum number of FWI iterations. The inversion may not complete each defined iteration due to the implementation of a model update convergence criteria, defined in Equation 3.1, and shown for this example in Table C.1. Doing so ensures that each iteration generates a significant decrease in the objective function, and we refrain from overfitting the data.

Incorporating the inner loop of this multiscale FWI strategy yields dramatic improvements over those in subsection C.2.1. The full- and sparse-data inversions successfully recover the subsurface velocity structure at depth. The ultra-sparse data inversion has improved also, but this result is still not satisfactory.

C.2.3 Configuration Three: Quasi-Progressive Frequencies

The progressive increase of maximum offset retains lower offsets in previously inverted offset bands, but until now, the frequency strategy does not retain previous inverted-for frequencies (Table C.2). Retaining most inverted frequencies in a group and progressively adding high frequencies is known as the quasi-progressive frequency strategy (Górszczyk et al., 2017). It is more computationally expensive but should improve the inversion results as it will retain the low-wavelength structure. This new quasi-progressive frequency strategy for the 2-12 Hz Marmousi frequency bandwidth is in Table C.5 for the synthetic Marmousi inversion. The increase in computational cost with respect to the overlapping strategy is directly attributed to a progressive increase in the number of frequencies simultaneously inverted for.

The current FWI parameter configuration is shown in Table C.6 and the results are shown in Figure C.5. Observations made from these inversions can be summarized as follows:

Table C.5: A quasi-progressive frequency strategy for FWI. All frequencies are shown in Hz.

Quasi-Progressive Frequencies	
Group 1	[2.0, 2.25, 2.5]
Group 2	[2.0, 2.5, 3.0]
Group 3	[2.0, 2.5, 3.0, 3.5]
Group 4	[2.0, 3.0, 3.5, 4.0]
Group 5	[2.0, 3.0, 3.5, 4.0, 4.5]
Group 6	[2.0, 3.0, 4.0, 4.5, 5.0]
Group 7	[2.0, 3.0, 4.0, 4.5, 5.0, 5.5]
Group 8	[2.0, 3.0, 4.0, 5.0, 5.5, 6.0]
Group 9	[2.0, 3.0, 4.0, 5.0, 5.5, 6.0, 6.5]
Group 10	[2.0, 3.0, 4.0, 5.0, 6.0, 6.5, 7.0]
Group 11	[2.0, 3.0, 4.0, 5.0, 6.0, 6.5, 7.0, 7.5]
Group 12	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 7.5, 8.0]
Group 13	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 7.5, 8.0, 8.5]
Group 14	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 8.5, 9.0]
Group 15	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 8.5, 9.0, 9.5]
Group 16	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 7.5, 8.0, 9.0, 9.5, 10.0]
Group 17	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 7.5, 8.0, 9.0, 9.5, 10.0, 10.5]
Group 18	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 7.5, 8.0, 9.0, 10.0, 10.5, 11.0]
Group 19	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 7.5, 8.0, 9.0, 10.0, 10.5, 11.0, 11.5]
Group 20	[2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 7.5, 8.0, 9.0, 10.0, 11.0, 11.5, 12.0]

- Both the full and sparse data inversions improve toward the edges and bottom of the model.

- The sparse data inversion experiences a further removal of the "dimpled" texture and faded artifacts related to acquisition footprint.

- The improvements for the full and sparse datasets are not as significant as incorporating the multiscale offset approach.

- There is a further improvement noted in the inverted model produced by the

Table C.6: A description of the varied parameters for the Marmousi synthetic study. *The number of iterations is typically varied in loop two, however they remain constant for this synthetic example.

Inversion Parameter	Multiscale Inversion Loop	Definition
Frequency Group	Loop One	Quasi-progressive frequencies
Laplace Constant	Loop Two	N/A
Iterations	Constant	Five iterations per maximum offset*
Offset	Loop Three	[4, 8, 12, 17] (km)
Gradient Preconditioner	Loop Three	[$1e^{-2}$, $1e^{-3}$, $1e^{-4}$, $1e^{-5}$]
Data Weighting	Loop Three	N/A

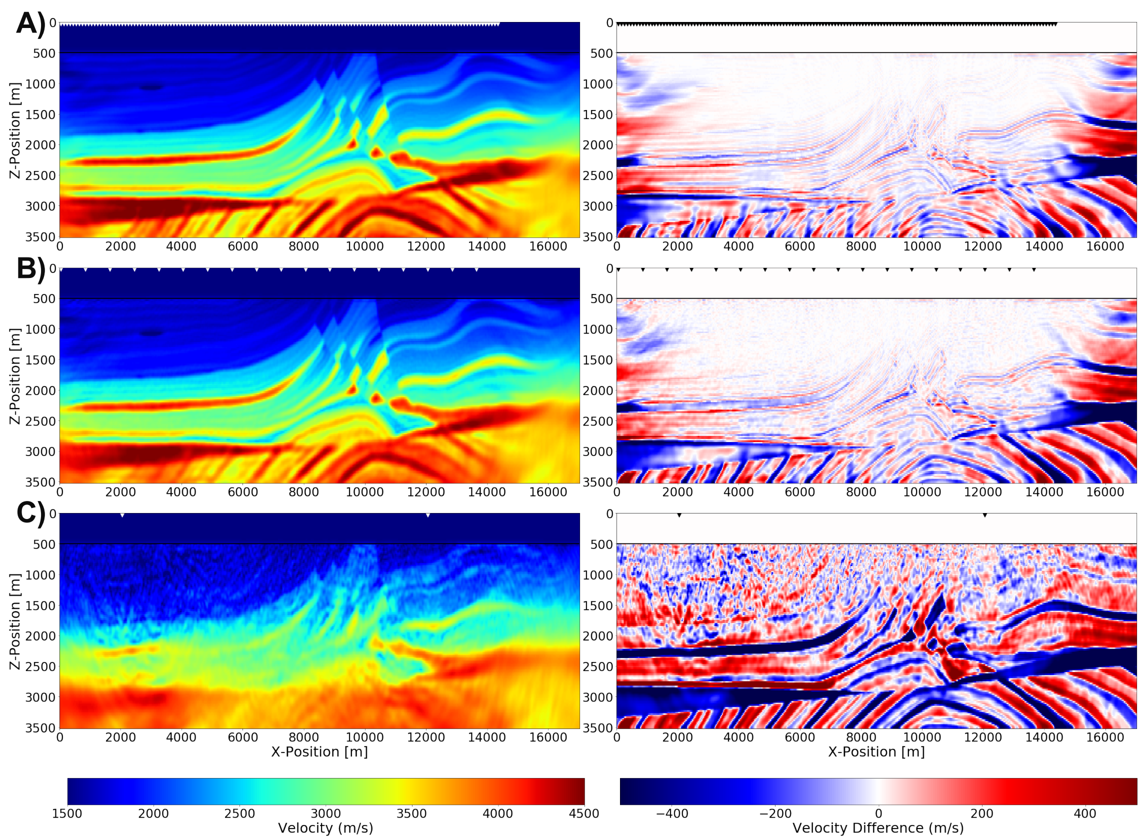

Figure C.5: Synthetic datasets with 144, 18, and 2 sources are inverted using the Marmousi2 model. The recovered velocity model is shown to the left, and the residual velocity field is shown to the right. The residual field is obtained by subtracting the true model from the inverted model. FWI using a quasi-progressive frequency strategy is completed with 144 sources in (A), with 18 sources in (B), and with 2 sources in (C).

ultra-sparse dataset. The high velocity layers of the Marmousi structure are better defined.

This frequency strategy, which retains low frequency information throughout the inversion, improves the inverted model with all three datasets. Incorporating more frequencies into an inversion increases the redundancy of signal in the data, which is unnecessary when synthetic data are modelled, but important for real-data inversions. With regards to computational time, there is approximately a four times increase to add the multiscale offset strategy to the baseline inversions from subsection C.2.1. For the quasi-progressive frequency strategy, it takes approximately nine times as long to invert the data with respect to the baseline inversions. It is unclear that incorporating this quasi-progressive frequency strategy is beneficial for the dense dataset once computational time is considered. It took approximately five times as long to perform the dense data inversion than the sparse data inversions. Comparing Figures C.5B and C.4A, the sparse data result with a quasi-progressive frequency strategy strongly resembles the full data inversion with only the multiscale offset approach. This sparse data inversion is about twice as fast as the full data inversion. When considering computational time, the most efficient strategy is to use the sparse dataset with this multiscale inversion strategy. The notion that an eighth of the available data may outperform the full dataset warrants further investigation in future studies.

C.2.4 Multiscale Configuration Four: Data-Damping and Regularization

We now extend our work to incorporate the full multiscale strategy of Górszczyk et al. (2017) by performing FWI in the Laplace-Fourier domain. It resembles the approach used in subsection 3.5.1, but with slightly different damping values. Subsection 3.5.1 also describes a new iteration structure from Górszczyk et al. (2017) as well as a linear data weight applied to the data. Other than the different damping values, the main differences between the multiscale test discussed here and that shown in Figure 3.2B are the placement of sources and the frequency strategy employed. This example uses a quasi-progressive frequency strategy (Table C.5), but inversions in subsection 3.5.1 uses a progressive frequency strategy. A progressive frequency strategy includes every previously inverted-for frequency in each frequency group, whereas the quasi-progressive strategy systematically omits frequencies previoulsy inverted-for in the interest of saving computational time.

Table C.7: A description of the varied parameters for the Marmousi synthetic study.

Inversion Parameter	Multiscale Inversion Loop	Definition
Frequency Group	Loop One	Increasing frequencies
Laplace Constant	Loop Two	[4.0, 2.0, 1.0, 0.5]
Iterations	Loop Two	[[5],[5,5],[5,5,5],[5,5,5,5]]
Offset	Loop Three	[4, 8, 12, 17] (km)
Gradient Preconditioner	Loop Three	[1e^{-2}, 1e^{-3}, 1e^{-4}, 1e^{-5}]
Data Weighting	Loop Three	0.5 min offset, 1.0 max offset

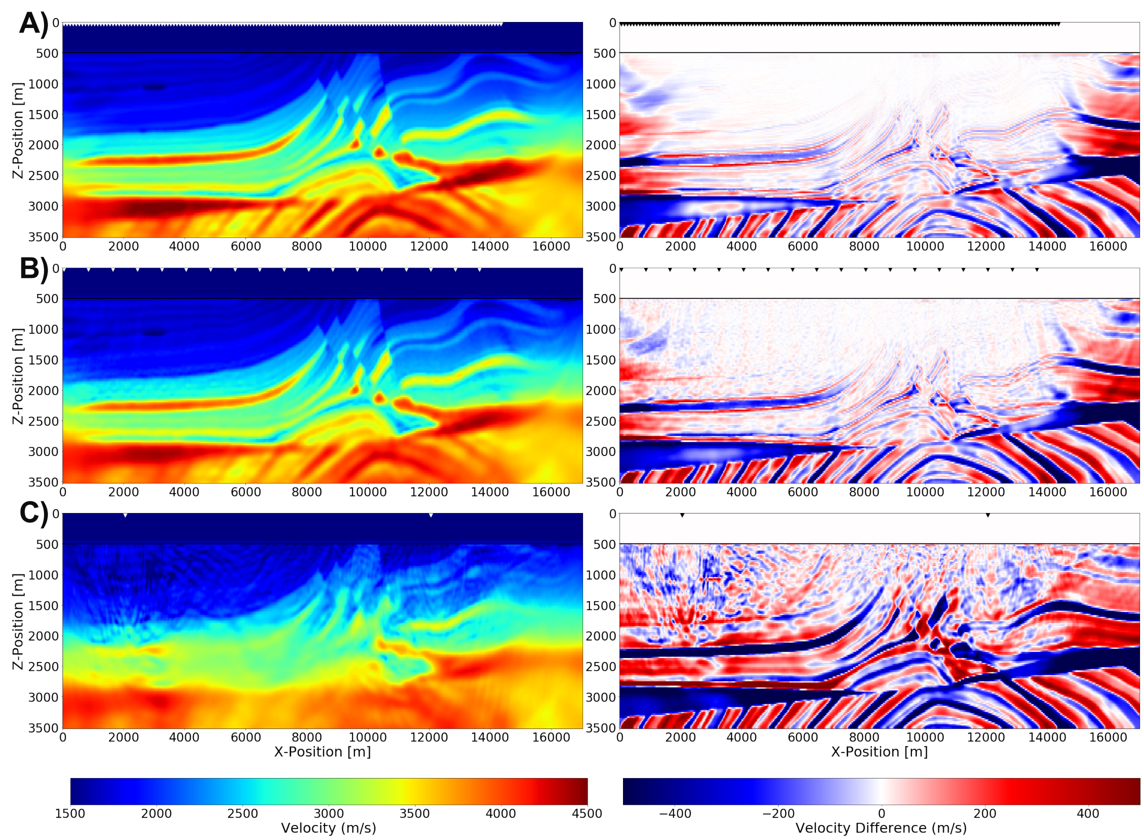

Figure C.6: Synthetic datasets with 144, 18, and 2 sources are inverted using the Marmousi2 model. The recovered velocity model is shown to the left, and the residual velocity field is shown to the right. The residual field is obtained by subtracting the true model from the inverted model. A full multiscale FWI strategy is now implemented with 144 sources in (A), with 18 sources in (B), and with 2 sources in (C).

The summary of the Marmousi multiscale inversion parameters for the data-damping and regularization test is in Table C.7. The results for this inversion are shown in Figure C.6. Observations from these results may be summarized as follows:

- Both inversions using the full and sparse acquisitions have regressed in deep portions of the model. Only above 1.5 km near the center of the model does there appear to be no change.
- The acquisition footprint is now visible in the velocity residual model for the sparse acquisition once again.
- There is improvement in the upper section for the ultra-sparse inversion. The

Marmousi structure is much more defined here.

- The deeper sections in the ultra-sparse inversion are not as well-recovered in comparison to the quasi-progressive inversion.

The regression in the recovered FWI model for the dense and sparse datasets is puzzling. An interpretation as to why these inversions have gotten worse is provided in subsection 4.2.1 as it summarizes these results.

C.3 Eastern Mediterranean Synthetic Tests

This section is intended to be a prelude to the full-scale synthetic Eastern Mediterranean inversion shown in Figure 3.4. There are two tests performed in this subsection, the first of which highlights the importance of a weighted Laplace-Fourier inversion at the crustal scale. The second investigates the various preconditioned optimization techniques provided by the SEISCOPE Optimization Toolbox (Métivier & Brossier, 2016). Other than the aforementioned parameter tests, the FWI parameters for the synthetic inversions are identical to those used in subsection 3.5.2. Table C.8 shows the constant inversion parameters for all synthetic Eastern Mediterranean inversions.

Parameter definitions for offset, Laplace constant, iterations, and the gradient preconditioner are provided in Table 3.2. A genuine progressive frequency approach is now used instead of the quasi-progressive frequency approach used in subsection C.2.3. The progressive approach retains every frequency previously inverted, while consistently growing each frequency group. Table C.9 defines the progressive frequency approach for the 0.25-5.25 Hz bandwidth used to invert the synthetic Eastern Mediterranean model.

Table C.8: A summary of the TOY2DAC inputs for the synthetic Eastern Mediterranean inversions that are not varied.

FWI Input	
Source Estimation	Mean source over all shots
Data Weighting Function	0 min offset, 1 max offset
l-BFGS Memory Parameter	10
Inner CG iterations for TRN/TGN	10
Convergence Criteria Model Update	$1e^{-5}$
Smoothing [x,z]	[0.4,0.2]
Velocity Bounds	1450 m/s & 8000 m/s
FDFD Input	
PML Coefficients and Number of PML	90 & 10
Hicks Interpolation	On
Free Surface	On
Source Type	Explosive
Receiver Type	Pressure
MUMPS Input	
Pivot Order (ICNTL7)	METIS
Memory Relaxation (ICNTL14)	50 (%)
Memory (ICNTL23)	24 Gb
Frequency Management	
Frequency Strategy	True progressive (Table C.9)

Table C.9: A true progressive frequency strategy for FWI. All frequencies are shown in Hz.

Progressive Frequencies	
Group 1	[0.25, 0.5, 0.75]
Group 2	[0.25, 0.5, 0.75, 1.25]
Group 3	[0.25, 0.5, 0.75, 1.25, 1.75]
Group 4	[0.25, 0.5, 0.75, 1.25, 1.75, 2.25]
Group 5	[0.25, 0.5, 0.75, 1.25, 1.75, 2.25, 2.75]
Group 6	[0.25, 0.5, 0.75, 1.25, 1.75, 2.25, 2.75, 3.25]
Group 7	[0.25, 0.5, 0.75, 1.25, 1.75, 2.25, 2.75, 3.25, 3.75]
Group 8	[0.25, 0.5, 0.75, 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25]
Group 9	[0.25, 0.5, 0.75, 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25, 4.75]
Group 10	[0.25, 0.5, 0.75, 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25, 4.75, 5.25]

C.3.1 Importance of Weighting from the First Breaks

Much time was spent trying to figure out why FWI was not able to recover the deep-crustal structure during full-scale synthetic inversions. This section emphasizes the importance of using the Laplace-Fourier approach correctly for crustal-scale investigations. Figure C.7 shows the recovered FWI model if damping is not weighted from the first breaks. The updated velocity field in Figure C.7B shows that structure down to approximately 10 km depth is recovered. The residual velocity field in Figure C.7C shows an acquisition footprint on the velocity updates as well. Figure C.7D shows the difference in the FWI model when damping is weighted from the first breaks (Figure 3.4). Evidently the application of damping from the first breaks is quite important for the recovery of deep velocity structure.

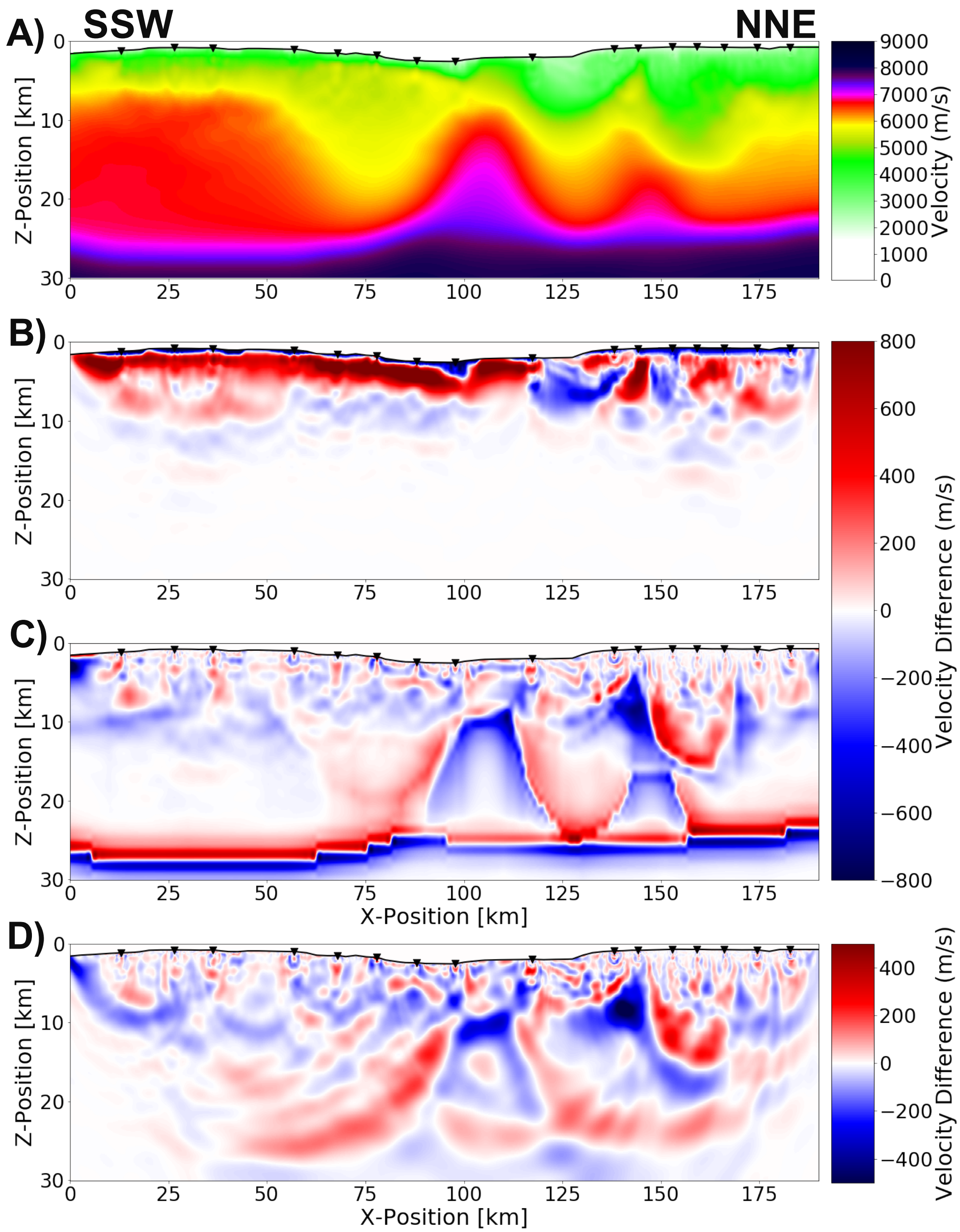

Figure C.7: **Inversion damping from t=0:** The synthetic Eastern Mediterranean model inverted without the weighting of Laplace-Fourier domain waveform inversion from the first breaks. The inverted synthetic Eastern Mediterranean is in (A), the updated velocity field is in (B), (C) shows the residual velocity field, and (D) shows the difference between this recovered model and that in Figure 3.4.

C.3.2 Optimization Technique Assessments

FWI is performed for the synthetic Eastern Mediterranean model using a preconditioned steepest descent algorithm and a preconditioned nonlinear conjugate gradient (NLCG). The result using steepest descent is shown in Figure C.8, and the result using NLCG is shown in Figure C.9. These results are compared to the preconditioned l-BFGS algorithm presented in subsection 3.5.2. Unfortunately, the truncated Newton and Gauss-Newton methods are unable to acceptably converge. The number of iterations and computational time required for each algorithm are as follows.

- Preconditioned l-BFGS: 271 iterations and 47057 seconds.
- Preconditioned NLCG: 349 iterations and 64314 seconds.
- Preconditioned Steepest Descent: 303 iterations and 69969 seconds.

The NLCG algorithm provides the next best result following l-BFGS as shown in Figures C.9C and C.9D. Preconditioned steepest descent provides the poorest result as shown in Figures C.8C and C.8D.

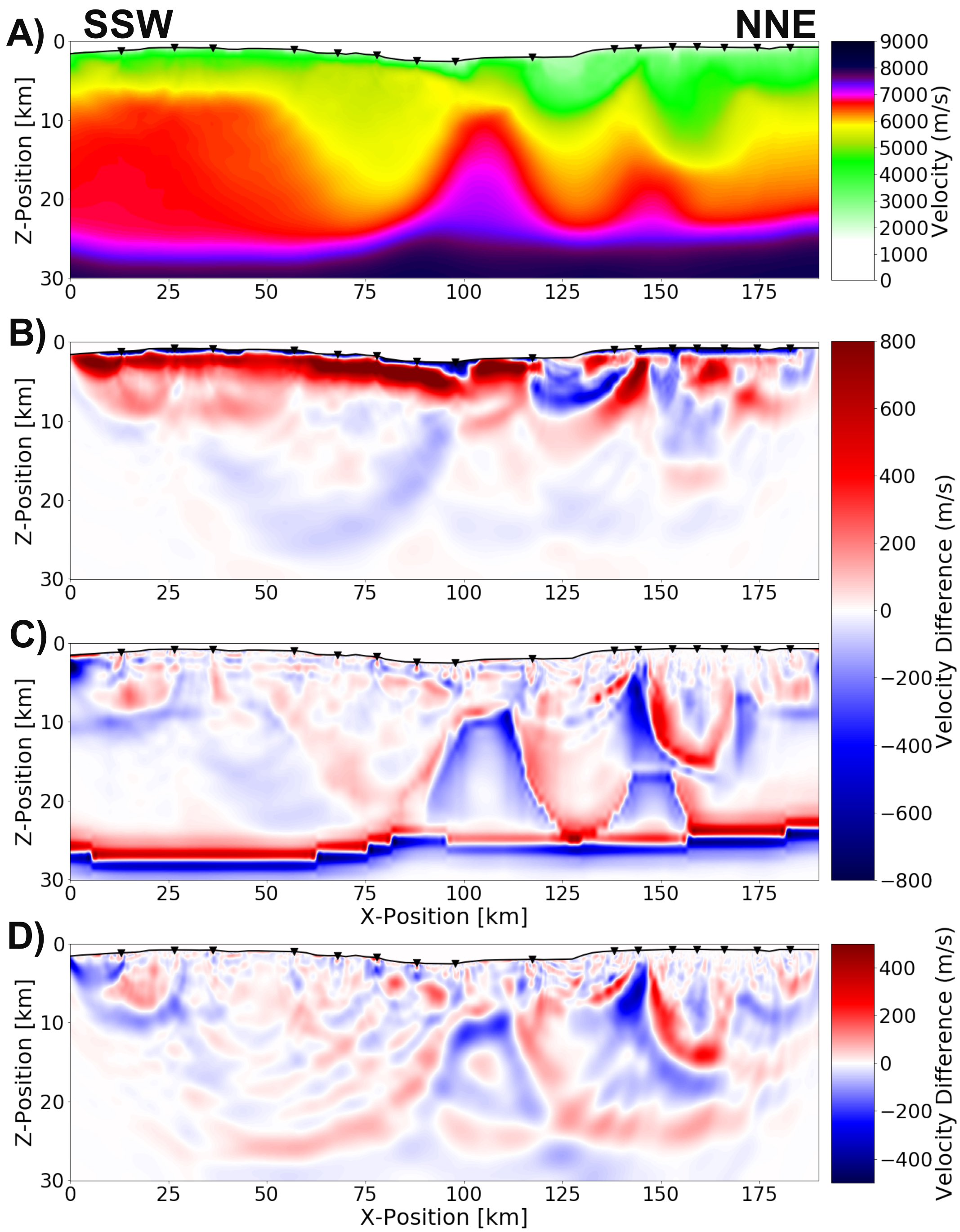

Figure C.8: **Preconditioned steepest descent:** The synthetic Eastern Mediterranean model inverted using a preconditioned steepest descent algorithm. The inverted synthetic Eastern Mediterranean is in (A), the updated velocity field is in (B), (C) shows the residual velocity field, and (D) shows the difference between this recovered model and that in Figure 3.4.

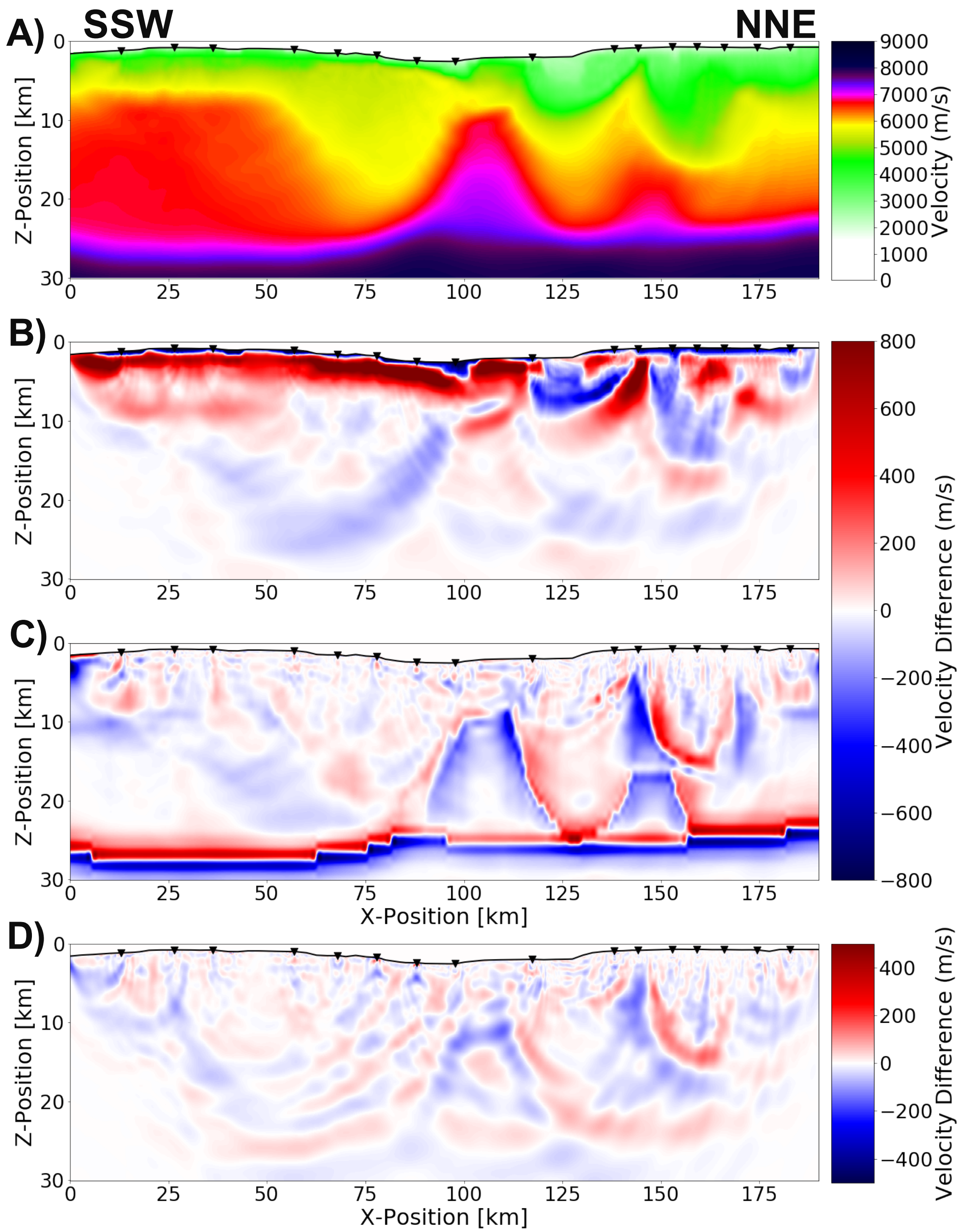

Figure C.9: **Preconditioned NLCG:** The synthetic Eastern Mediterranean model inverted using a preconditioned NLCG algorithm. The inverted synthetic Eastern Mediterranean is in (A), the updated velocity field is in (B), (C) shows the residual velocity field, and (D) shows the difference between this recovered model and that in Figure 3.4.

Appendix D

FWI Real Data Testing and Analysis

D.1 Introduction

This appendix is intended to provided supplementary QC materials for the final FWI model shown in Chapter 3. As FWI is performed on a sparse OBS dataset, the recovered velocity model is highly uncertain. In a general sense, it fits the data better than the starting velocity model, which is positive, but much of the data are poorly recovered. As opposed to the shot-frequency analysis of the data fit in Figure 3.8, a conventional plot of the total cost function, the relative cost function, and number of computed gradients is shown in Figure D.1. It is evident in this figure that the total cost function is progressively decreased to a minimum. The jumps in the cost function and relative cost function is related to the introduction and removal of data in the observed dataset as a function of the multiscale FWI strategy. Generally, far offsets and lower Laplace constants have higher cost functions.

This appendix will first show frequency-domain traces at 5.0 Hz for each OBS. Then, the time-domain predicted shot gathers are compared to the starting and observed time-domain data. The plots are meant to provide full transparency toward the FWI real-data results shown in Chapter 3. Discussion of the results in this section may be found in Chapter 4

The frequency domain traces are obtained by forward modelling the final frequency

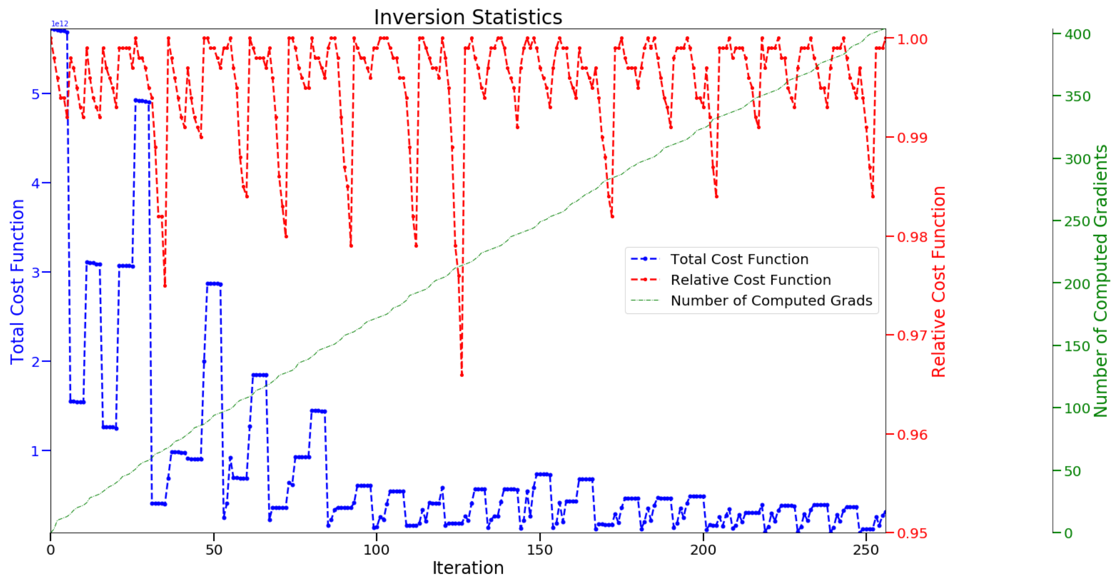

Figure D.1: A plot showing the total and relative cost functions and number of computed gradients for the recovered FWI model.

group through the final model, generating the final predicted dataset. This is done with the application of the final Laplace constant, which is not weighted from the first breaks. This frequency domain trace is then weighted from the first breaks once it is generated and the linear data-weight is applied with the first five km of offset muted.

An inverse Fourier transform is required to obtain the time-domain wavefields. The starting models and final FWI velocity model are up-scaled to a 20 m node spacing from a 50 m node spacing. This allows for frequencies up to 18.125 Hz to be modeled while avoiding temporal aliasing, assuming four nodes per wavelength and a minimal velocity of 1450 m/s. In practice, 3600 discrete frequencies are modeled with a frequency spacing of 0.005 Hz. These frequencies are then padded with zeros to form a matrix with frequencies up to 250 Hz. This allows for a time-domain sampling rate of 4 ms following the inverse Fourier transform, which is equivalent to the observed data (see subsection A.3.4).

D.1.1 Inversion Parameters

The TOY2DAC FWI parameters used to generate the real-data results in Chapter 3 are presented in Table D.1. Those presented here omit the inversion parameters presented in subsection 3.6.3.

Table D.1: A summary of the constant TOY2DAC inputs for the Marmousi inversions.

FWI Input	
Source Estimation	Mean source over all shots
Deadzone	None, 0 m
l-BFGS Memory Parameter	3
Velocity Bounds	1450 m/s & 8250 m/s
FDFD Input	
PML Coefficients and Number of PML	90 & 10
Hicks Interpolation	On
Free Surface	On
Source Type	Vertical force
Receiver Type	Pressure
MUMPS Input	
Pivot Order (ICNTL7)	METIS
Memory Relaxation (ICNTL14)	50 (%)
Memory (ICNTL23)	48 Gb
Computational Aspects	
Cores	8
Computational Time	30 hours

D.2 Frequency-Domain QC

The summary plots in Figure 3 provide a good understanding of the shot-frequency combinations that better match the observed dataset compared to the starting model. However, the match against the observed data is harder to quantify. It is impractical to show each shot-frequency combination, therefore the observed and predicted frequency domain traces are shown for 5.0 Hz for each OBS. Discussion pertaining to the analysis of these results is provided in Chapter 4.

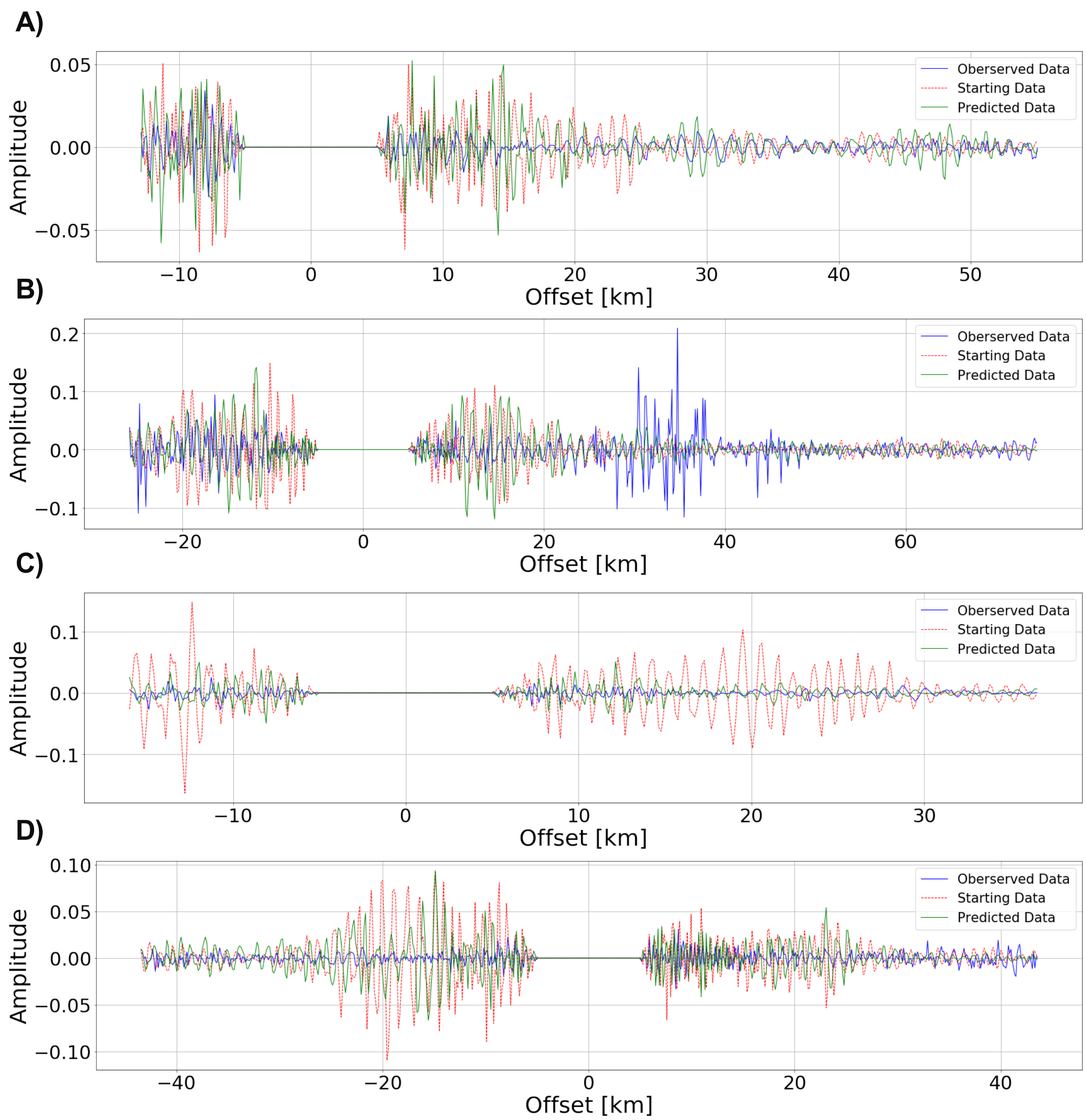

Figure D.2: The observed frequency domain trace is plotted with the predicted frequency domain trace from FWI at 5.0 Hz for four OBS; (A) OBS 4, (B) OBS 5, (C) OBS 6, and (D) OBS 7.

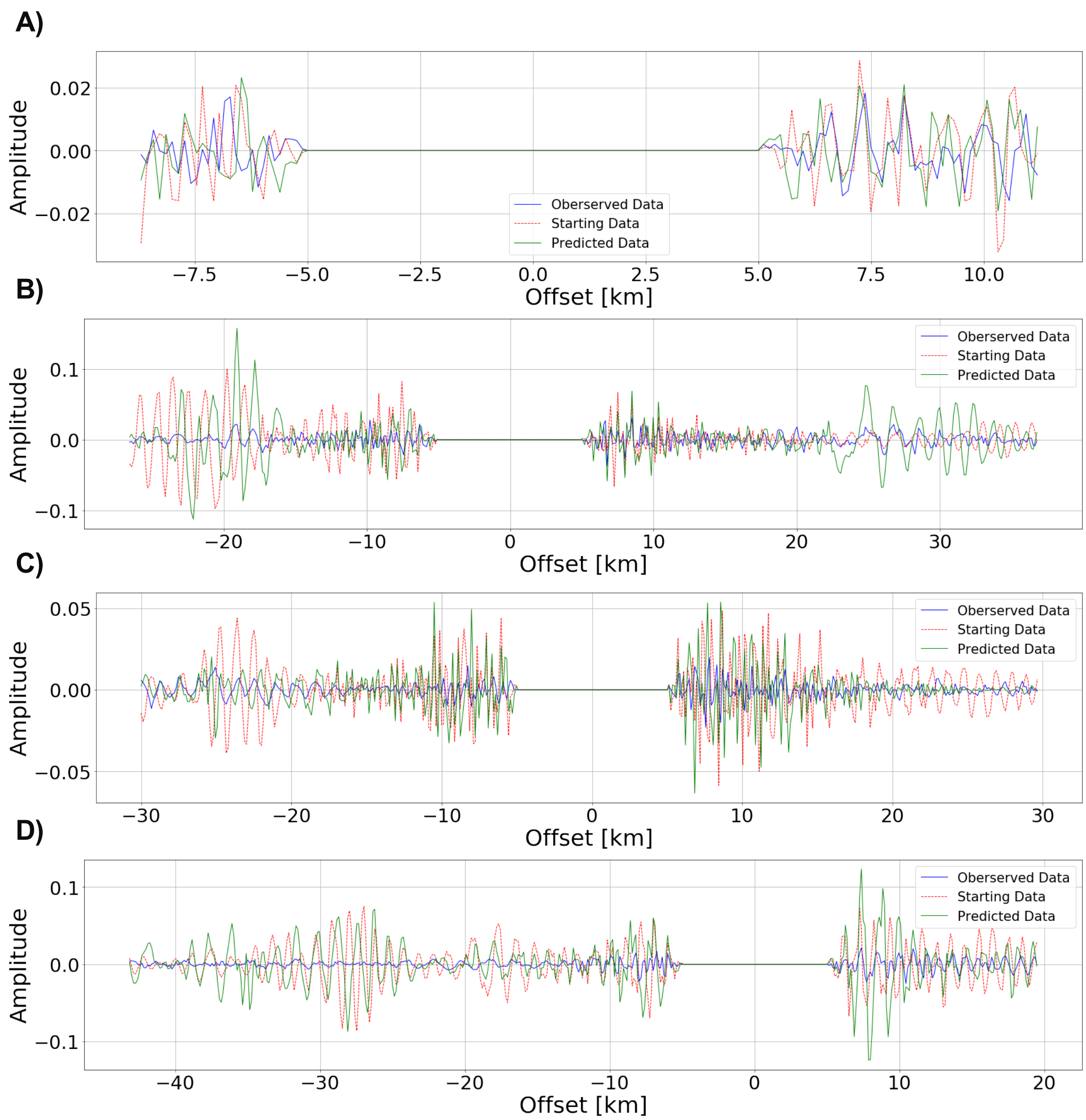

Figure D.3: The observed frequency domain trace is plotted with the predicted frequency domain trace from FWI at 5.0 Hz for four OBS; (A) OBS 8, (B) OBS 9, (C) OBS 10, and (D) OBS 11.

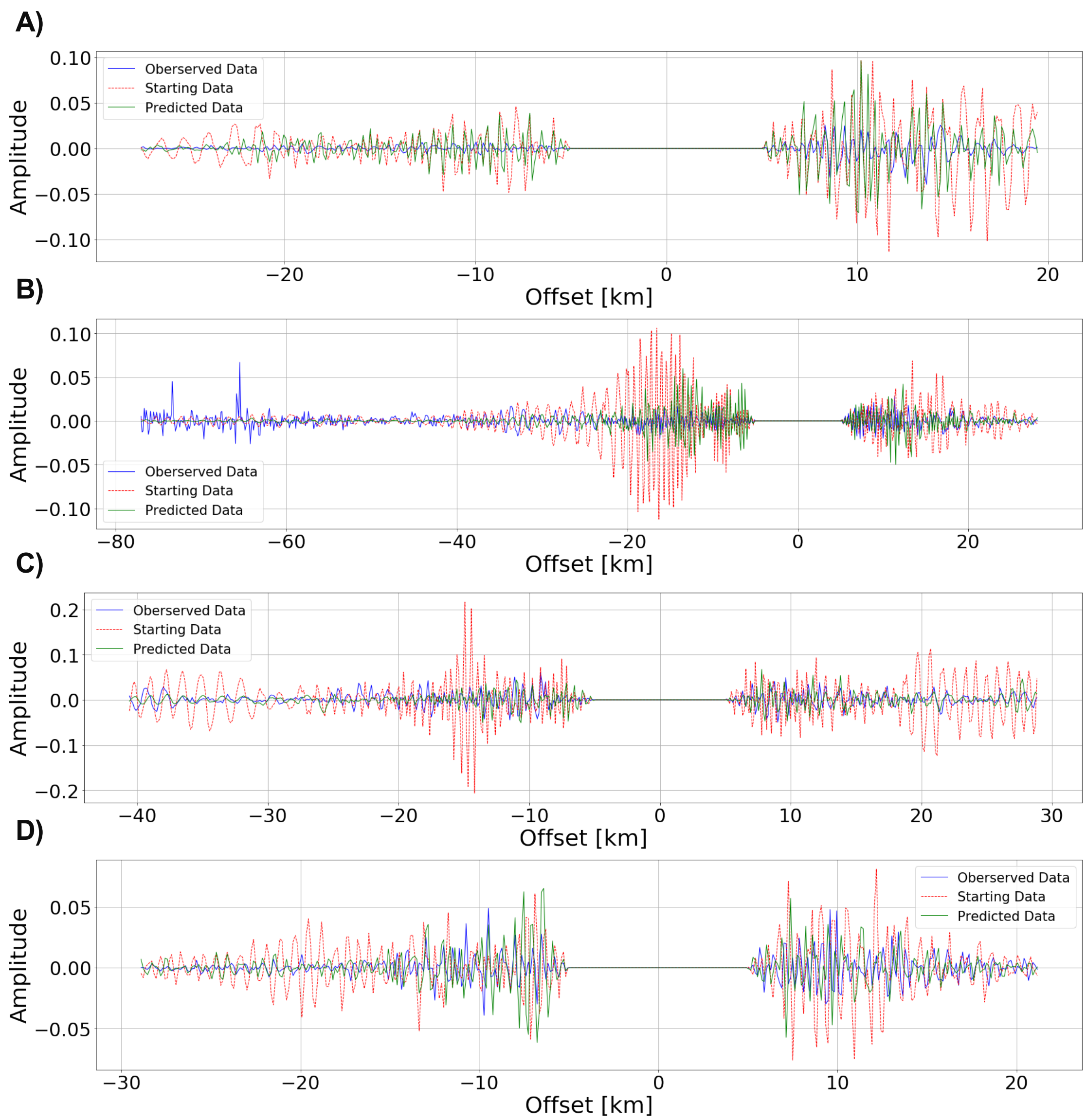

Figure D.4: The observed frequency domain trace is plotted with the predicted frequency domain trace from FWI at 5.0 Hz for four OBS; (A) OBS 13, (B) OBS 15, (C) OBS 16, and (D) OBS 17.

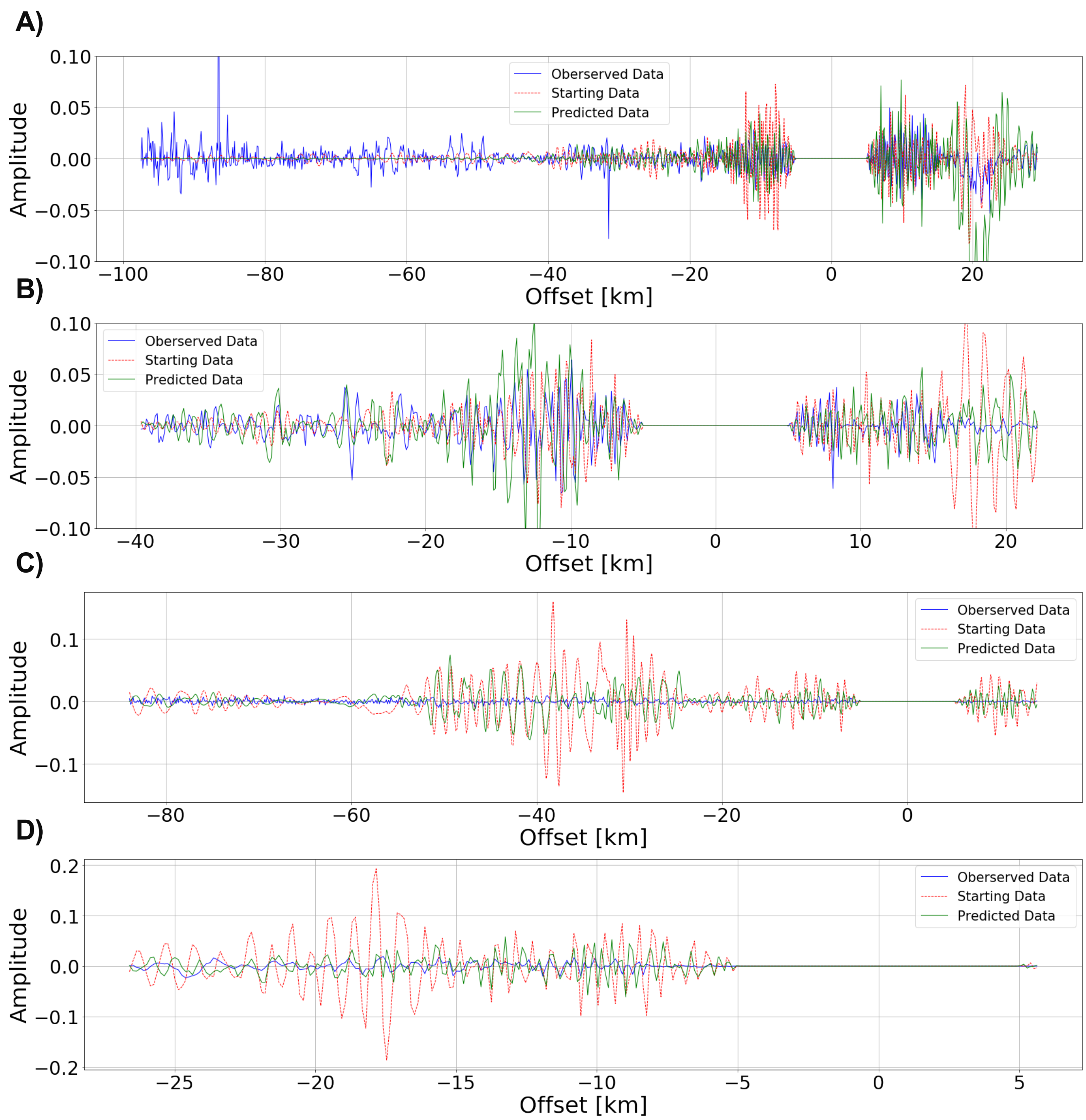

Figure D.5: The observed frequency domain trace is plotted with the predicted frequency domain trace from FWI at for four OBS; (A) OBS 18, (B) OBS 19, (C) OBS 20, and (D) OBS 21. For OBS 13, 14, and 16, the data shown are at 5.0 Hz, but for OBS 20 the data are shown at 3.0 Hz.

D.3 Time-domain QC

For each OBS, three shot gathers are shown. The data are forward modelled through the starting model, and shown on top (A). In the middle, the data are forward modelled through the final FWI model (B). The results for OBS 8 are omitted here as they are shown in Figure 3.10. The observed data are shown at the bottom and obey the data processing flow for the FWI dataset discussed in Appendix A. However, a Butterworth band-pass filter of 0.5/2.0-7.25/10.0 Hz is also applied. The purpose of this Butterworth filter is to highlight the frequency band where FWI is performed, 2.0-7.25 Hz. Tapers for the frequency domain data are designed to pass signal in this bandwidth as well, however the taper for the frequency domain data are longer than the time domain, 7.25-18.0 Hz, to avoid artifacts produced by the inverse Fourier transform. Like the previous section, Chapter 4 discusses the results from this appendix.

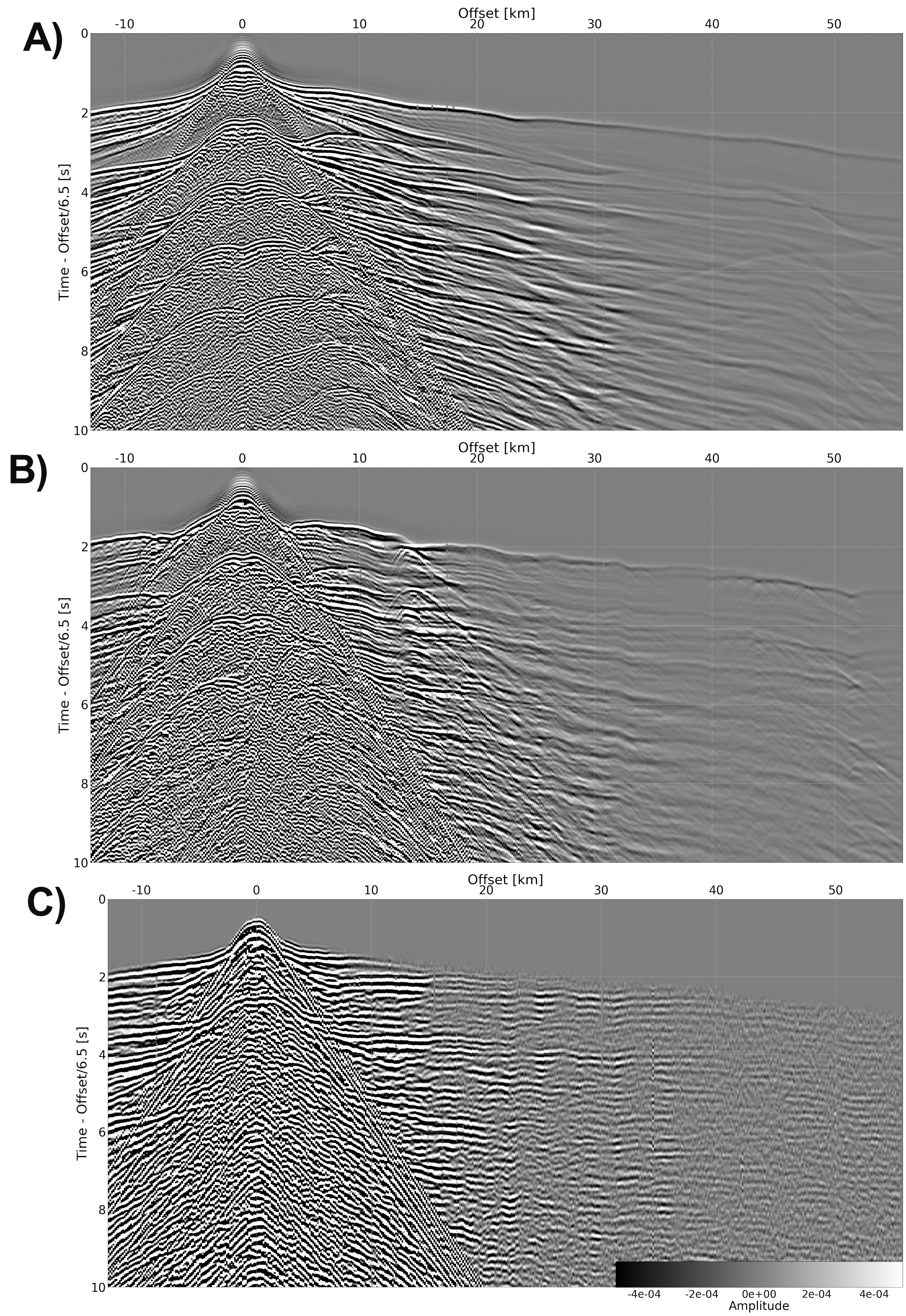

Figure D.6: For OBS 4, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

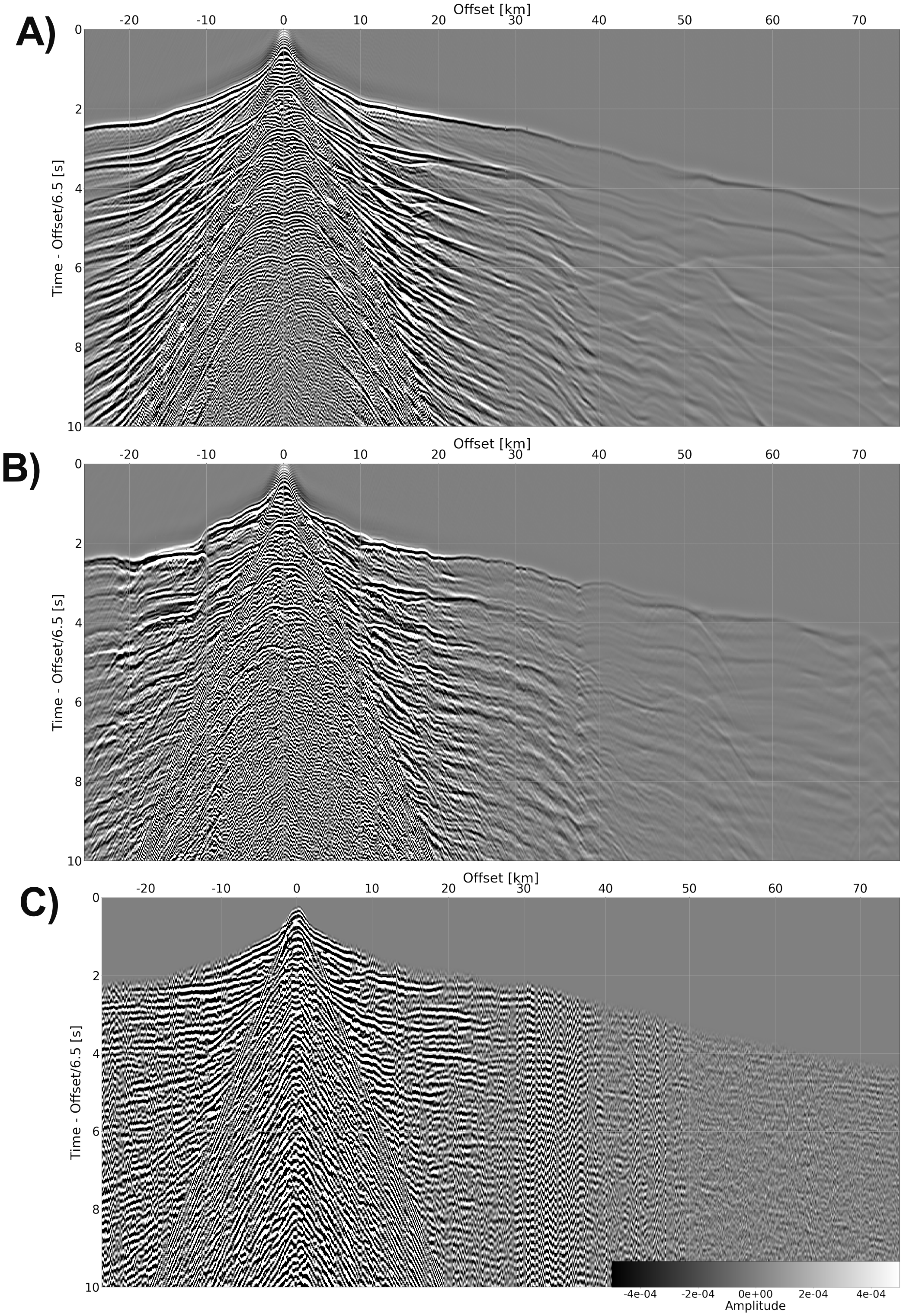

Figure D.7: For OBS 5, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

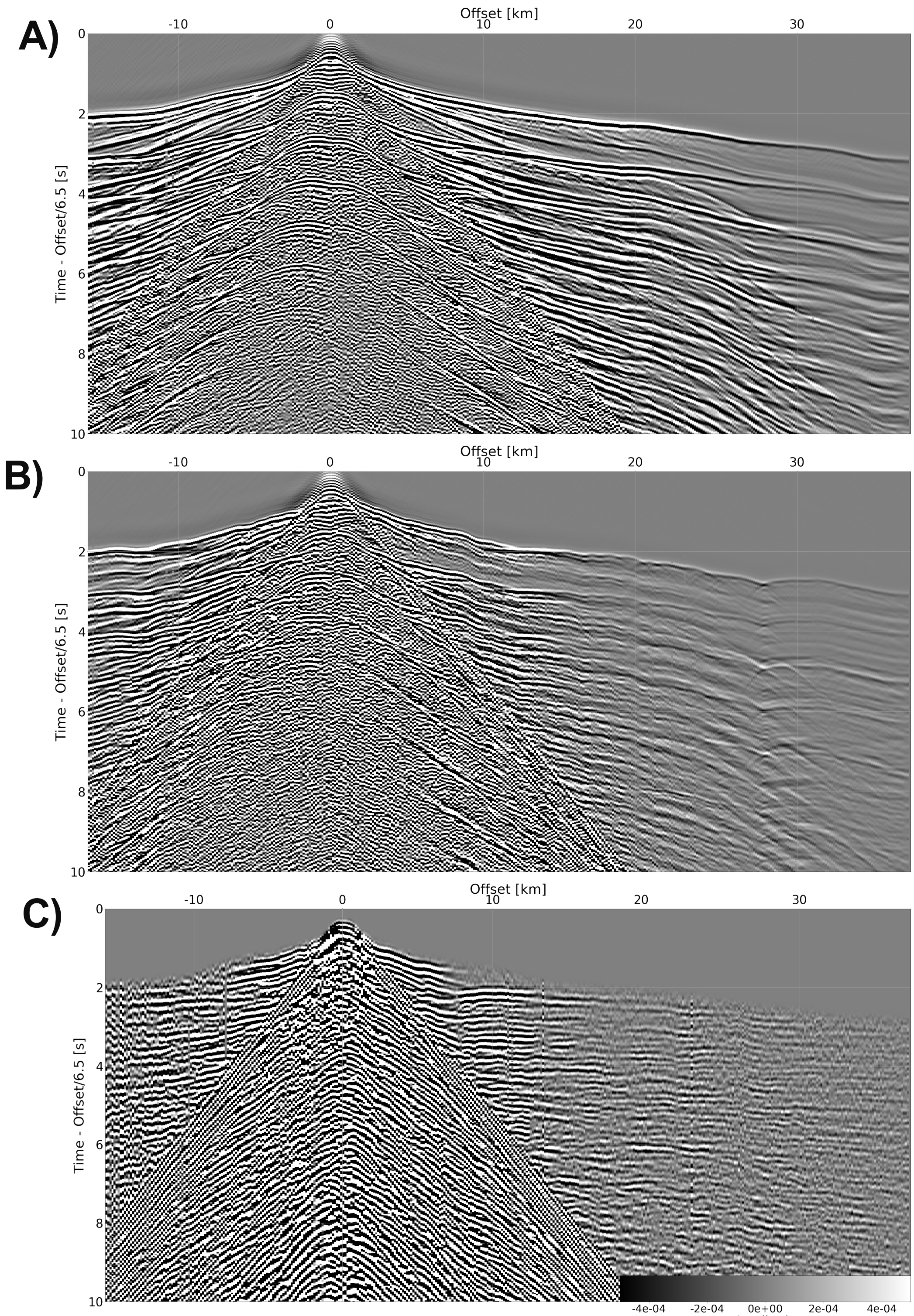

Figure D.8: For OBS 6, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

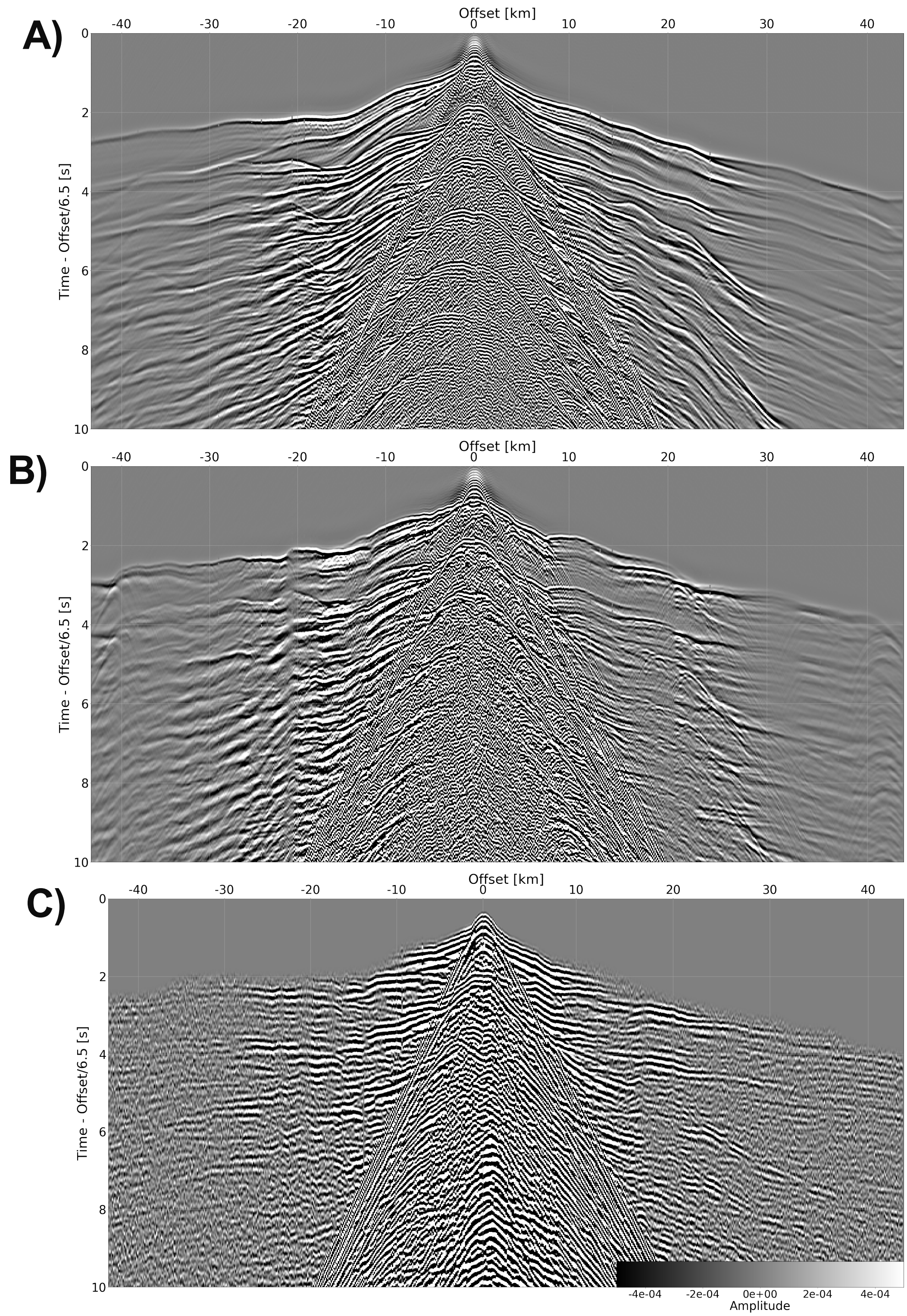

Figure D.9: For OBS 7, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

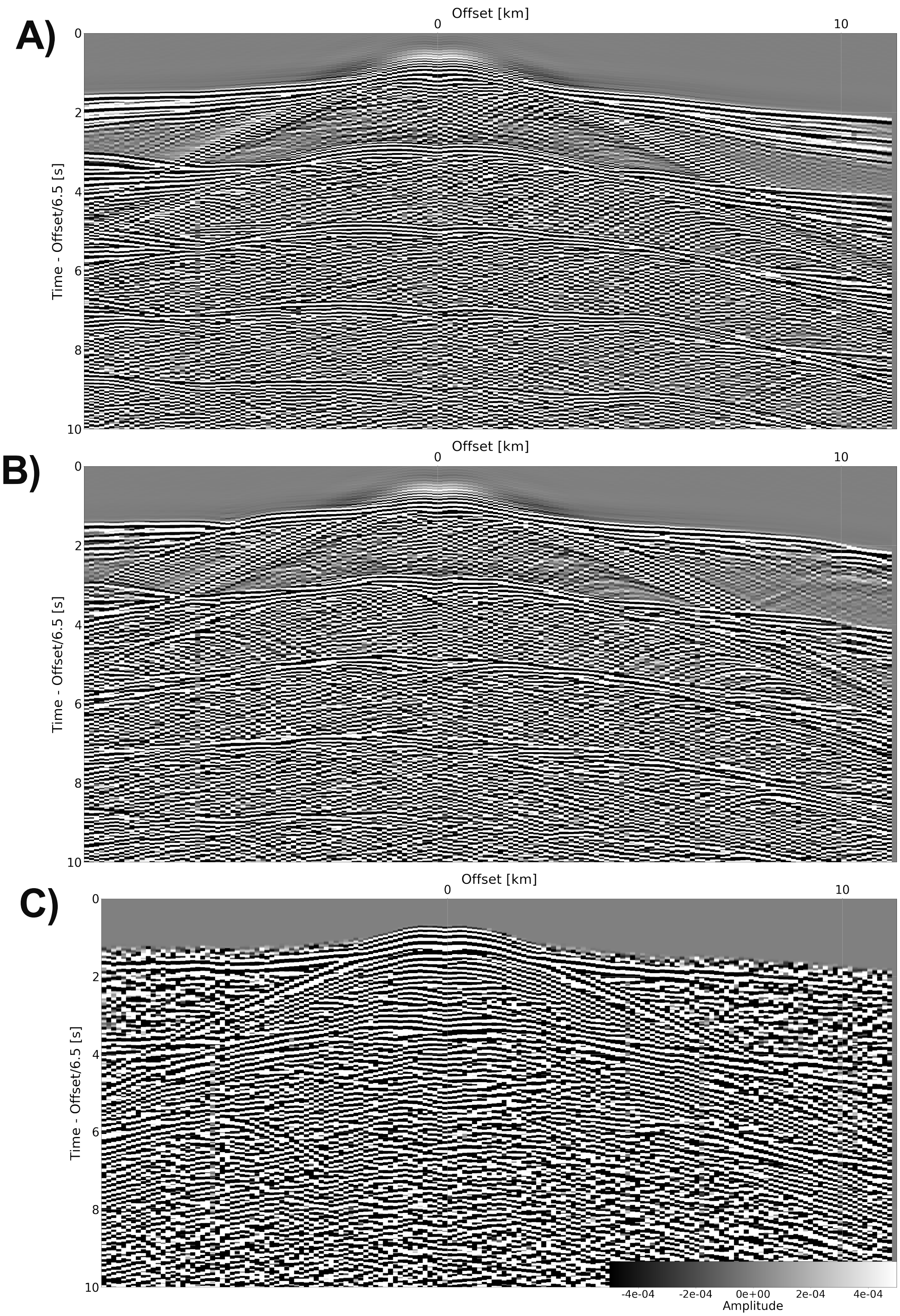

Figure D.10: For OBS 8, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

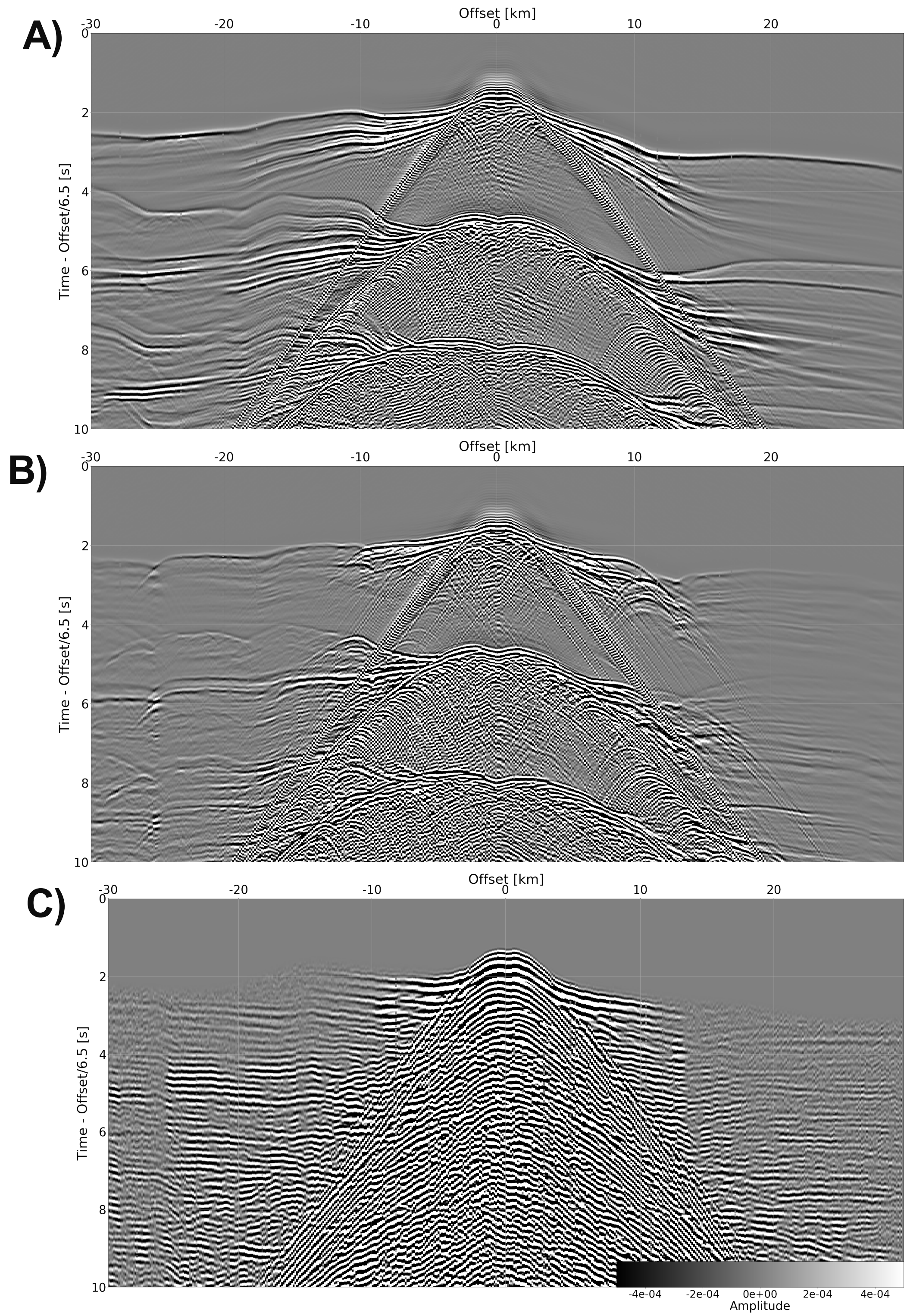

Figure D.11: For OBS 10, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

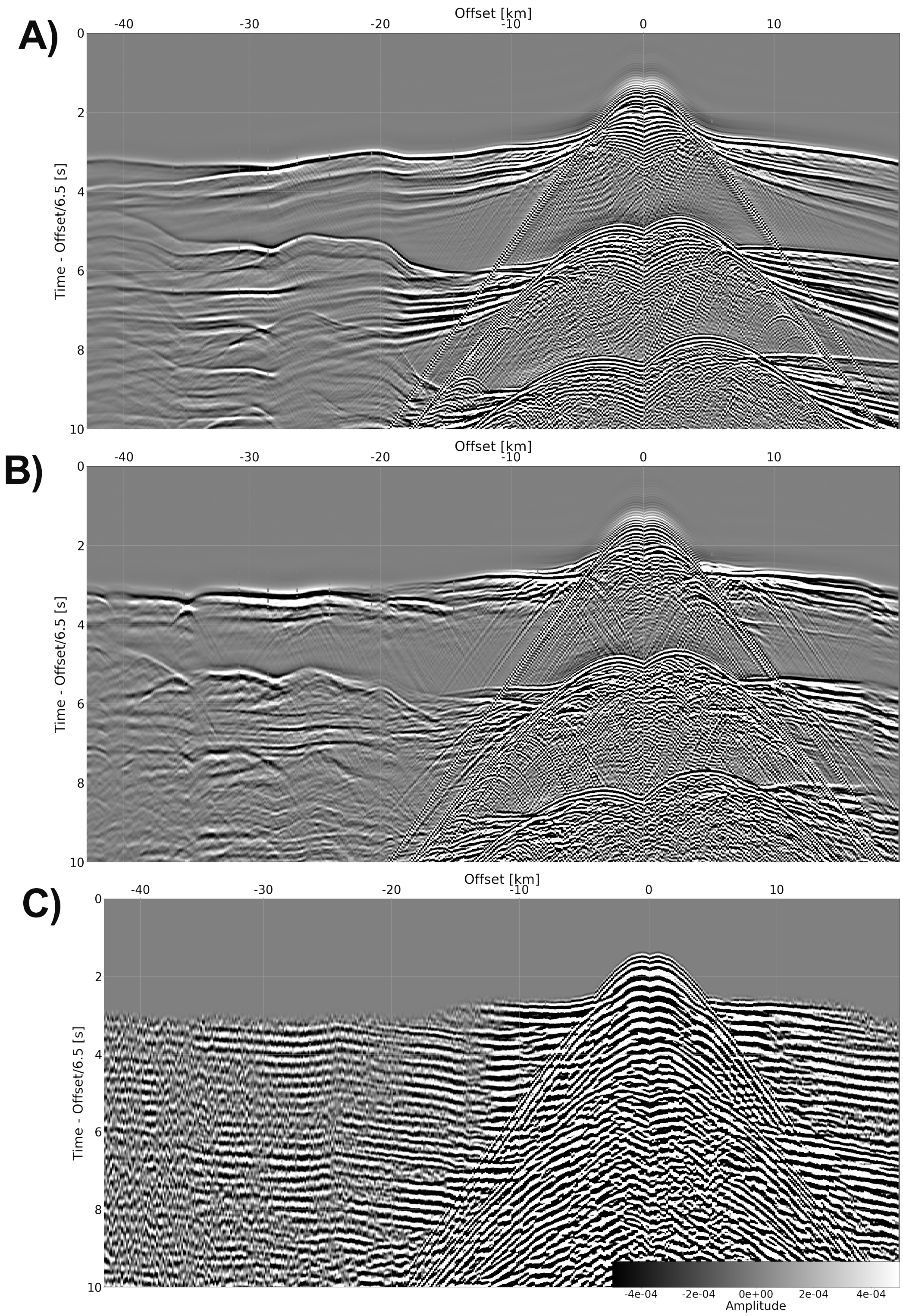

Figure D.12: For OBS 11, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

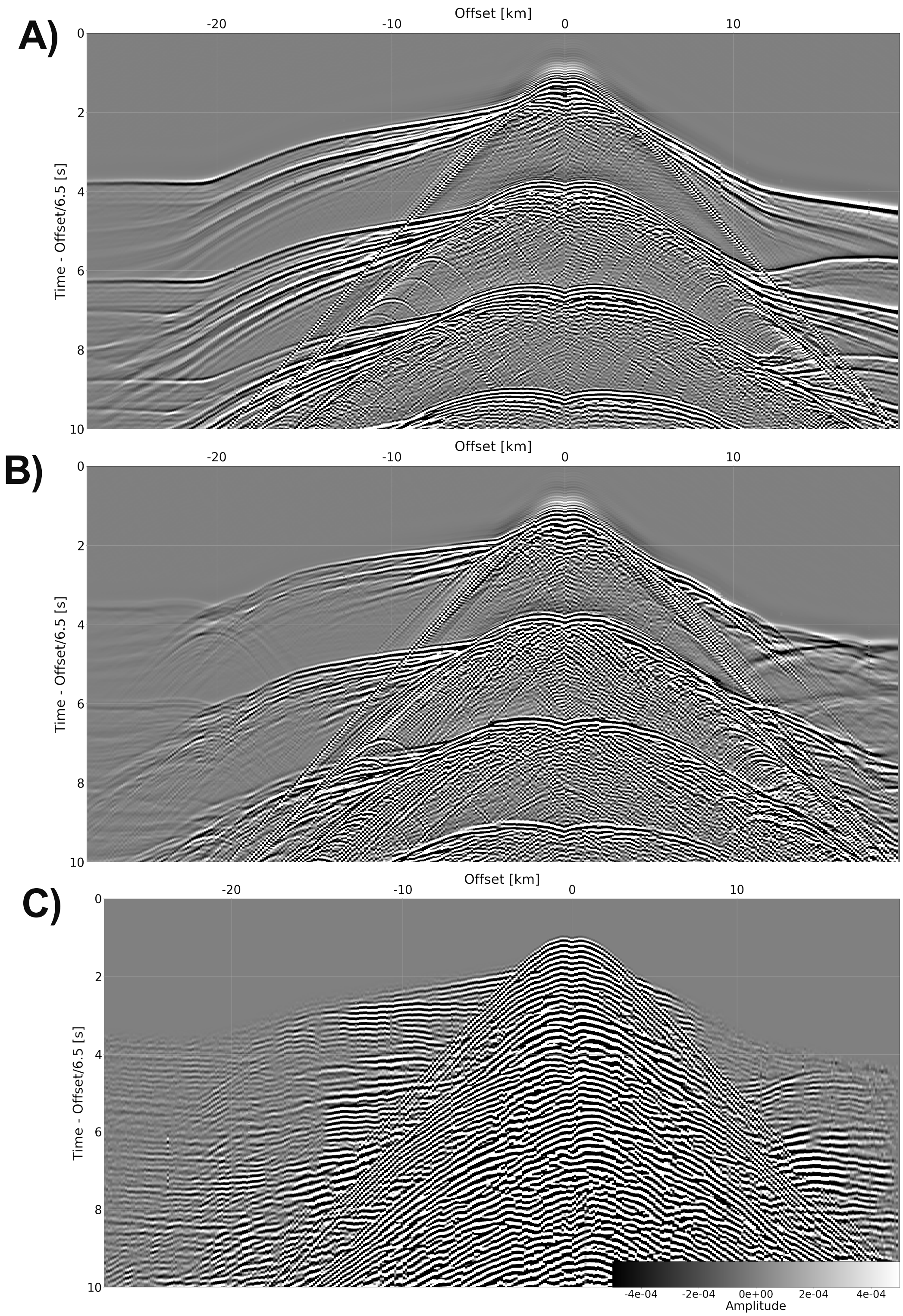

Figure D.13: For OBS 13, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

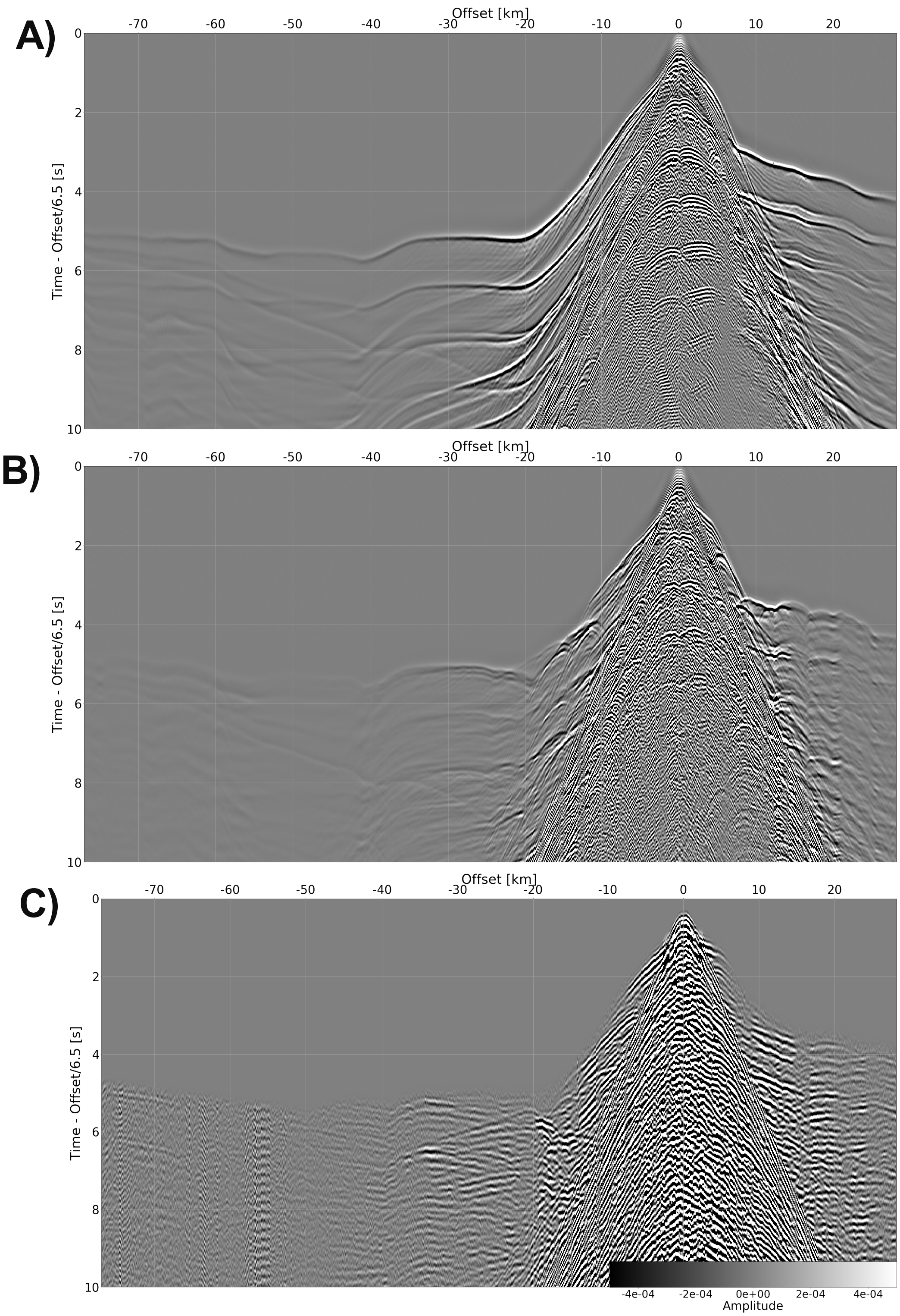

Figure D.14: For OBS 15, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

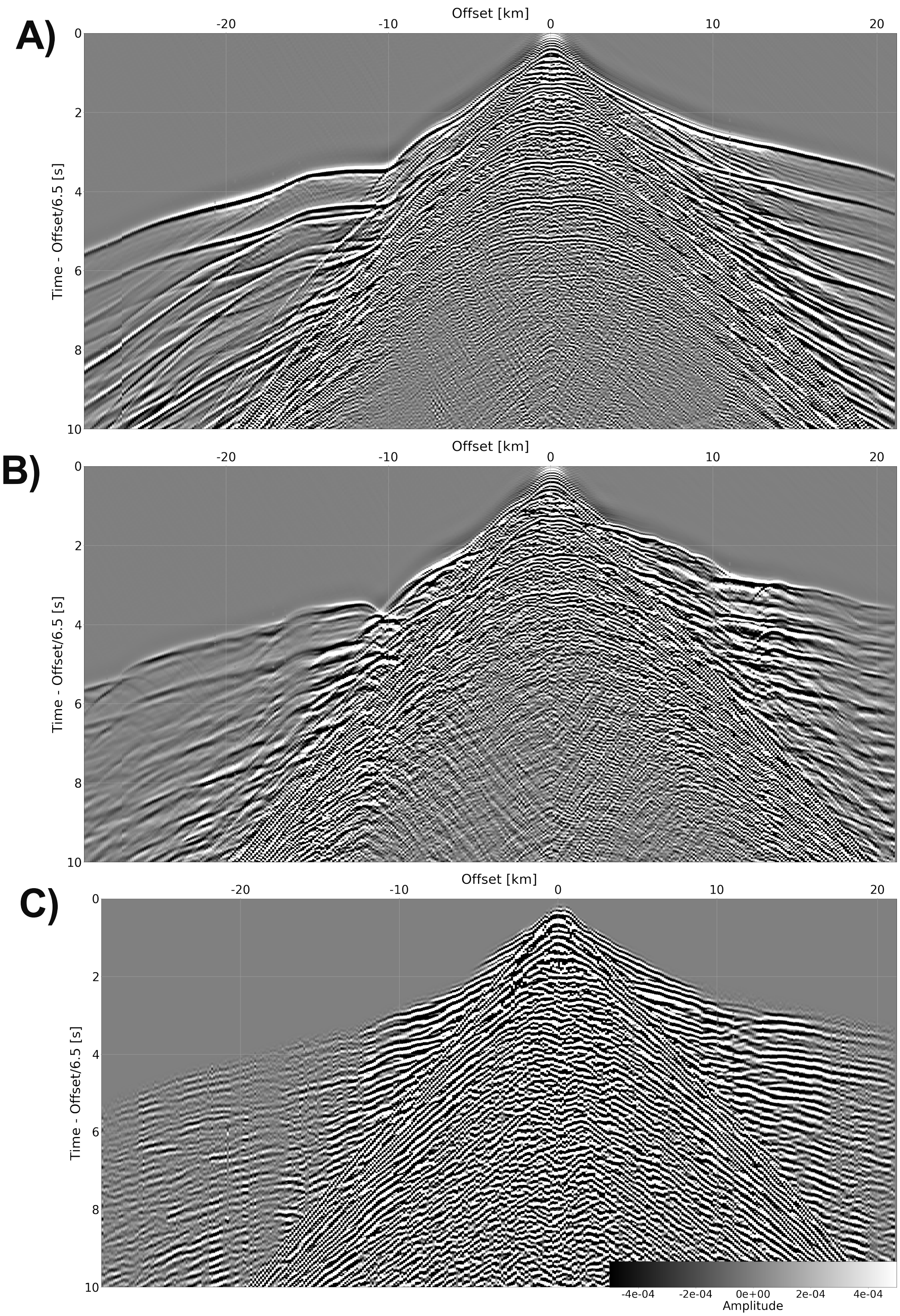

Figure D.15: For OBS 17, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

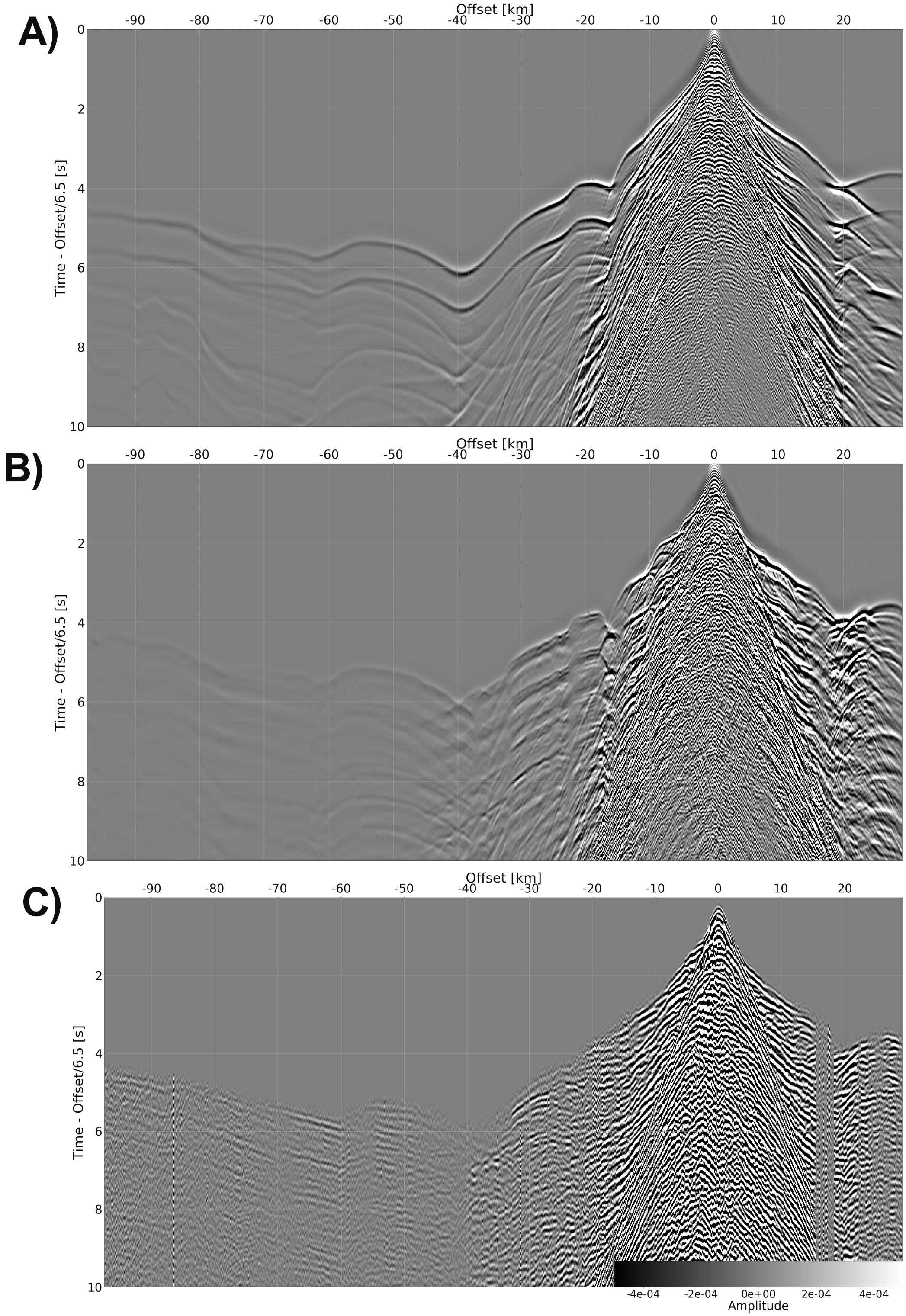

Figure D.16: For OBS 18, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

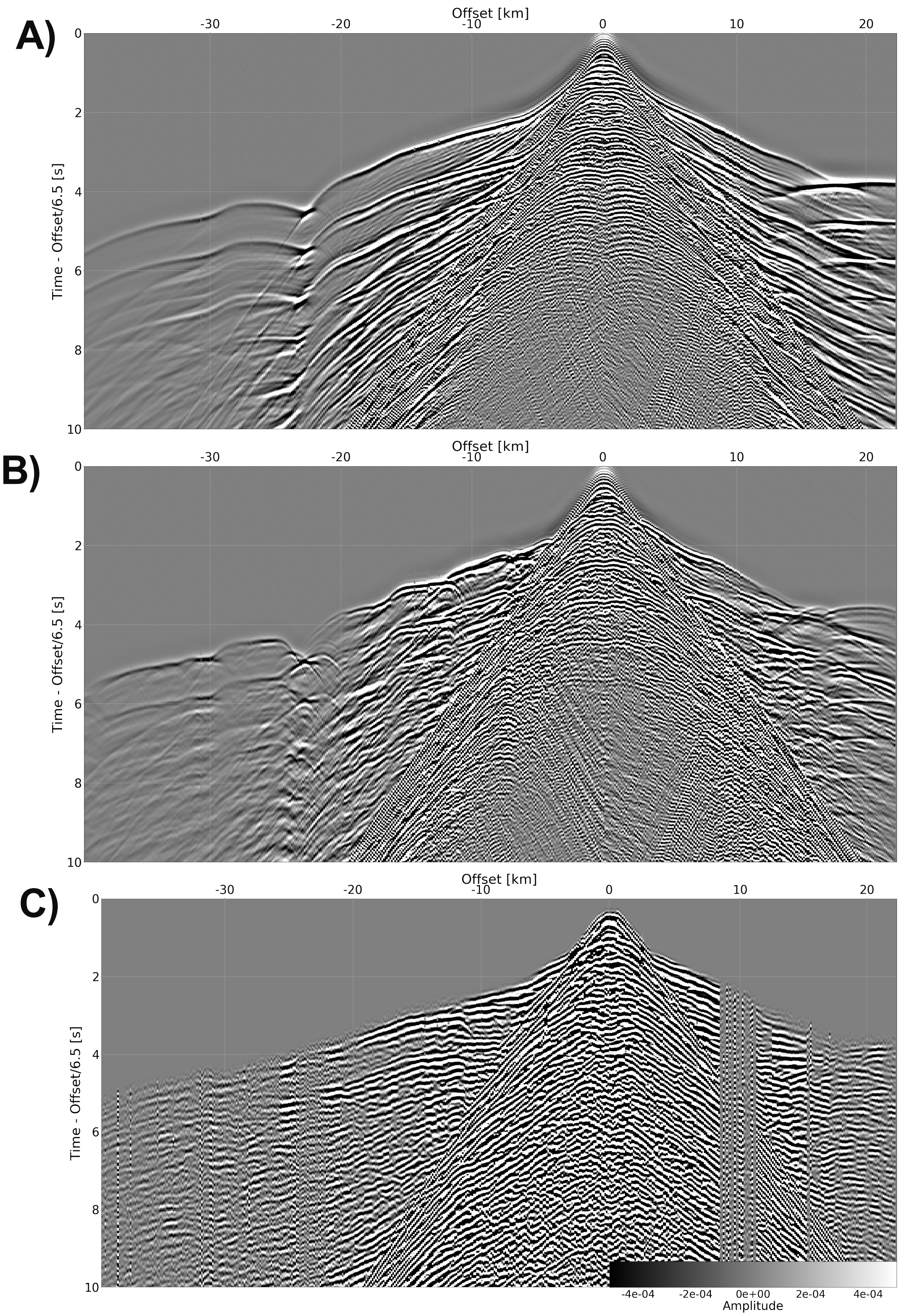

Figure D.17: For OBS 19, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

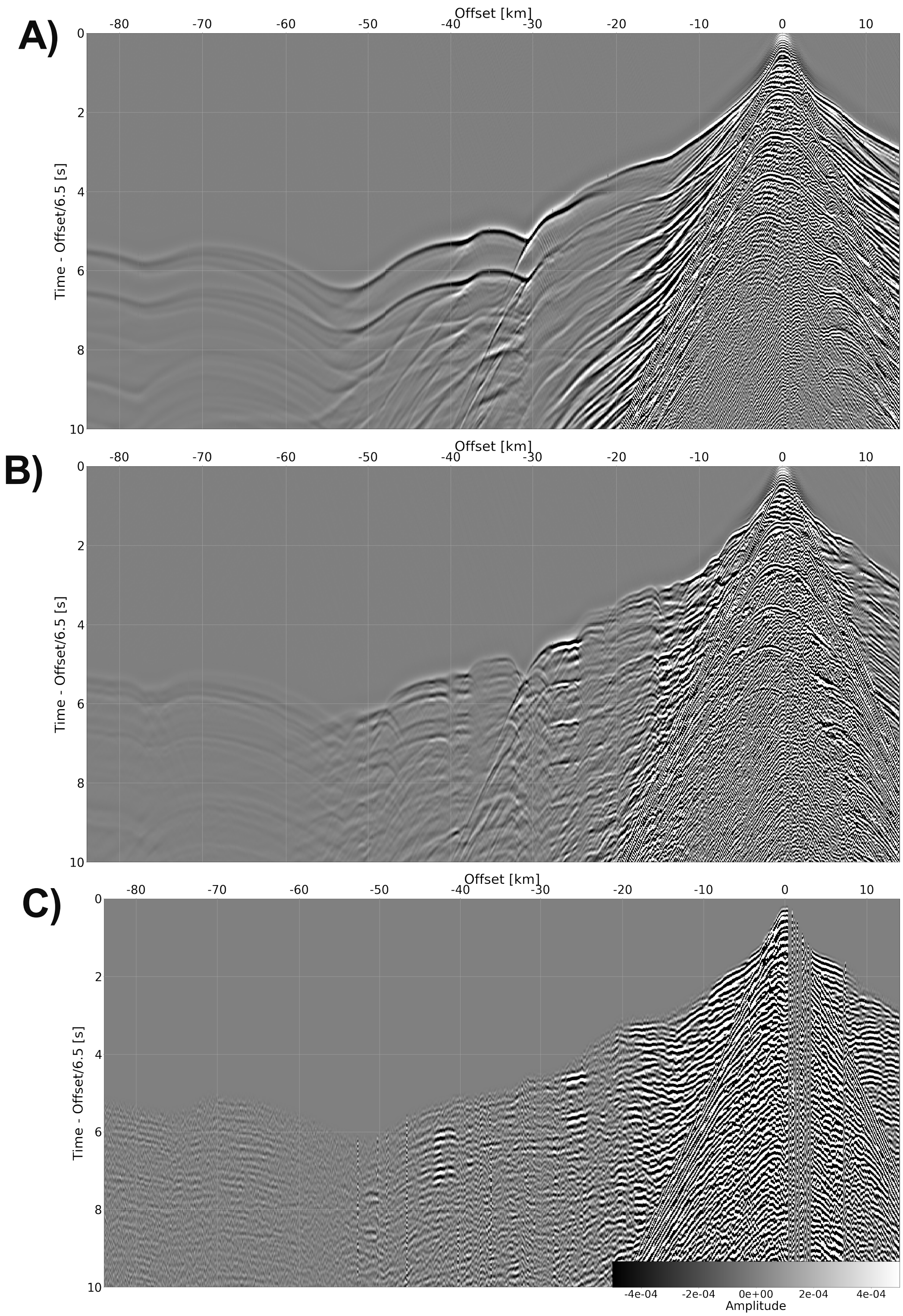

Figure D.18: For OBS 20, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

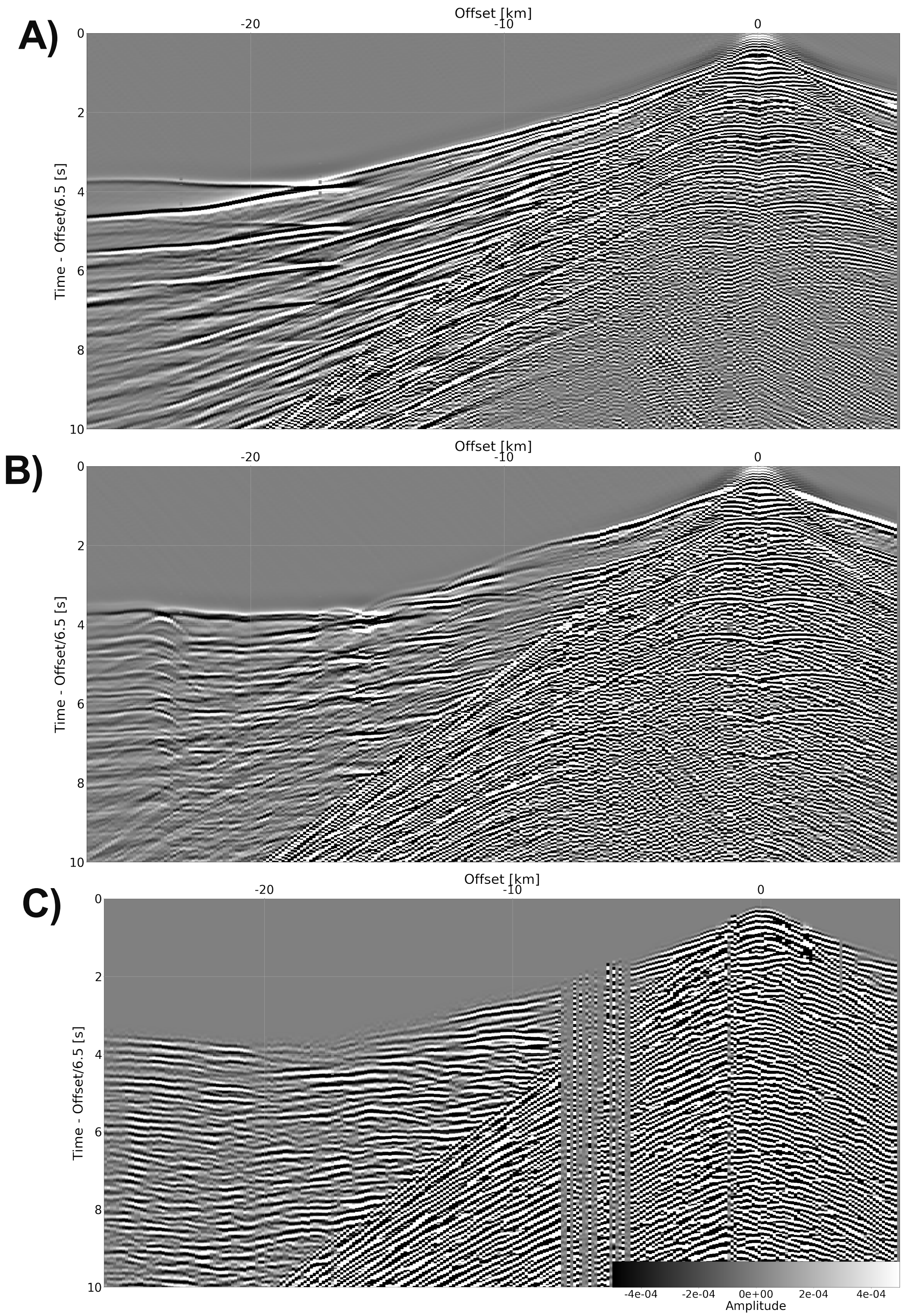

Figure D.19: For OBS 21, the starting data are in (A), (B) shows the predicted data, and (C) shows the observed seismic data.

Bibliography

Aki, K. & Richards, P. G., 1980. *Quantitative Seismology*, 2nd edn.

Amestoy, P. R., Duff, I. S., & L'Excellent, J. Y., 2000. Multifrontal parallel distributed symmetric and unsymmetric solvers, *Comput. Methods Appl. Mech. Engrg.*, **184**, 501–520.

Bailey, W. R., Holdsworth, R. E., & Swarbrick, R. E., 2000. Kinematic history of a reactivated oceanic suture: The Mamonia Complex Suture Zone, SW Cyprus, *Journal of the Geological Society, London*, **157**, 1107–1126.

Ben-Avraham, Z., Ginzburg, A., Makris, J., & Eppelbaum, L., 2002. Crustal structure of the Levant Basin, eastern Mediterranean, *Tectonophysics*, **346**, 23–43.

Ben-Avraham, Z., Shoham, Y., & Ginzburg, A., 1976. Magnetic Anomalies in the Eastern Mediterranean and the Tectonic Setting of the Eratosthenes Seamount, *Geophys. J. R. astr. Soc.*, **45**, 105–123.

Bird, P., 2003. An updated digital model of plate boundaries, *Geochemistry, Geophysics, Geosystems*, **4**(3).

Brenders, A. J. & Pratt, R. G., 2007. Full waveform tomography for lithospheric imaging: Results from a blind test in a realistic crustal model, *Geophys. J. Int.*, **168**, 133–151.

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.

Brossier, R., Virieux, J., & Operto, S., 2008. Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling, *Geophys. J. Int.*, **175**, 541–559.

Brossier, R., Operto, S., & Virieux, J., 2009. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics*, **74**(6), 105–118.

Bunks, C., Saleck, F., Zaleski, S., & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics*, **60**(5), 1457–1473.

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C., 1995. A Limited Memory Algorithm for Bound Constrained Optimization, *SIAM J. Sci. Comput.*, **16**(5), 1190–1208.

Calon, T. J., Aksu, A. E., & Hall, J., 2005a. The Neogene evolution of the Outer Latakia Basin and its extension into the Eastern Mesaoria Basin (Cyprus), Eastern Mediterranean, *Marine Geology*, **221**, 61–94.

Calon, T. J., Aksu, A. E., & Hall, J., 2005b. The Oligocene-Recent evolution of the Mesaoria Basin (Cyprus) and its western marine extension, Eastern Mediterranean, *Marine Geology*, **221**, 95–120.

Claerbout, J. F., 1971. Toward a unified theory of reflector mapping, *Geophysics*, **36**(3), 467–481.

Dai, Y. H. & Yuan, Y., 1999. A nonlinear conjugate gradient method with a strong global convergence property, *SIAM Journal on Optimization*, **10**(1).

Davy, R. G., Morgan, J. V., Minshull, T. A., Bayrakci, G., Bull, J. M., Klaeschen, D., Reston, T. J., Sawyer, D. S., Lymer, G., & Cresswell, D., 2017. Resolving the fine - scale velocity structure of continental hyperextension at the Deep Galicia Margin using full-waveform inversion, *Geophys. J. Int.*, **212**, 244–263.

Eakin, D. H., McIntosh, K. D., Van Avendonk, H. J. A., Lavier, L., Lester, R., Liu, C. S., & Lee, C. S., 2014. Crustal-scale seismic profiles across the Manila subduction zone: The transition from intraoceanic subduction to incipient collision, *Journal of Geophysical Research: Solid Earth*, **119**, 1–17.

Faccenna, C., Bellier, O., Martinod, J., Piromallo, C., & Regard, V., 2006. Slab detachment beneath eastern Anatolia: A possible cause for the formation of the North Anatolian fault, *Earth and Planetary Science Letters*, **242**, 85–97.

Feld, C., Mechie, J., Hübscher, C., Hall, J., Nicolaides, S., Gurbuz, C., Bauer, K., Louden, K., & Weber, M., 2017. Crustal structure of the Eratosthenes Seamount, Cyprus and S. Turkey from an amphibian wide-angle seismic profile, *Tectonophysics*, **700-701**, 32–59.

Frizon de Lamotte, D., Raulin, C., Mouchot, N., Wrobel-Daveau, J.-C., Blanpied, C., & Ringenbach, J.-C., 2011. The southernmost margin of the Tethys realm during the Mesozoic and Cenozoic: Initial geometry and timing of the inversion processes, *Tectonics*, **30**(TC3002).

Gardner, G. H. F., Gardner, L. W., & Gregory, A. R., 1974. Formation Velocity and Density- The Diagnostic Basics for Stratigraphic Traps, *Geophysics*, **39**(6), 770–780.

Gardosh, M. A., Garfunkel, Z., Druckman, Y., & Buchbinder, B., 2010. Tethyan rifting in the Levant Region and its role in Early Mesozoic crustal evolution, *Geological Society Special Publication*, **341**, 9–36.

Garfunkel, Z., 1998. Constrains on the origin and history of the Eastern Mediterranean basin, *Tectonophysics*, **298**, 5–35.

Golonka, J., 2007. Late Triassic and Early Jurassic palaeogeography of the world, *Palaeogeography, Palaeoclimatology, Palaeoecology*, **244**, 297–307.

Górszczyk, A., Operto, S., & Malinowski, M., 2017. Toward a robust workflow for deep crustal imaging by FWI of OBS data: The eastern Nankai Trough revisited, *Journal of Geophysical Research: Solid Earth*, **122**, 4061–4630.

Granot, R., 2016. Palaeozoic oceanic crust preserved beneath the eastern Mediterranean, *Nature Geoscience*, **9**, 701–705.

Hall, J., Aksu, A. E., Calon, T. J., & Yaşar, D., 2005a. Varying tectonic control on basin development at an active microplate margin: Latakia Basin, Eastern Mediterranean, *Marine Geology*, **221**, 15–60.

Hall, J., Calon, T. J., Aksu, A. E., & Meade, S. R., 2005b. Structural evolution of the Latakia Ridge and Cyprus Basin at the front of the Cyprus Arc, Eastern Mediterranean Sea, *Marine Geology*, **221**, 261–297.

Hicks, G. J., 2002. Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions, *Geophysics*, **67**(1), 156–166.

Hustedt, B., Operto, S., & Virieux, J., 2004. Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling, *Geophys. J. Int.*, **157**, 1269–1296.

Kamei, R. & Pratt, R. G., 2013. Inversion strategies for visco-acoustic waveform inversion, *Geophys. J. Int.*, **194**, 859–884.

Kamei, R., Pratt, R. G., & Tsuji T, 2012. Waveform tomography imaging of a megasplay fault system in the seismogenic Nankai subduction zone, *Earth and Planetary Science Letters*, **317-318**, 343–353.

Kamei, R., Pratt, R. G., & Tsuji, T., 2013. On acoustic waveform tomography of wide-angle OBS data—strategies for pre-conditioning and inversion, *Geophys. J. Int.*, **194**, 1250–1280.

Kamei, R., Pratt, R. G., & Tsuji, T., 2014. Misfit functionals in Laplace-Fourier domain waveform inversion, with application to wide-angle ocean bottom seismograph data, *Geophysical Prospecting*, **62**, 1054–1074.

Kempler, D., 1998. Eratosthenes Seamount: The Possible Spearhead of Incipient Continental Collision in the Eastern Mediterranean, *Proceedings of the Ocean Drilling Program, Scientific Results*, **160**, 709–721.

Klimke, J. & Ehrhardt, A., 2014. Impact and implications of the Afro-Eurasian collision south of Cyprus from reflection seismic data, *Tectonophysics*, **626**, 105–119.

Klingelhoefer, F., Labails, C., Cosquer, E., Rouzo, S., Géli, L., Aslanian, D., Olivet, J. L., Sahabi, M., Nouzé, H., & Unternehr, P., 2009. Crustal structure of the SW-Moroccan margin from wide-angle and reflection seismic data (the DAKHLA experiment) Part A: Wide-angle seismic models, *Tectonophysics*, **468**, 63–82.

Korenaga, J., Holbrook, W. S., Kent, G. M., Kelemen, P. B., Detrick, R. S., Larsen, H. C., Hopper, J. R., & Dahl-Jensen, T., 2000. Crustal structure of the southeast Greenland margin from joint refraction and reflection seismic tomography, *Journal of Geophysical Research*, **105**(B9), 21591–21614.

Korenaga, J., Holbrook, W. S., Detrick, R. S., & Kelemen, P. B., 2001. Gravity anomalies and crustal structure at the southeast Greenland margin, *Journal of Geophysical Research*, **106**(B5), 8853–8870.

Kurzmann, A., Przebindowska, A., Köhn, D., & Bohlen, T., 2013. Acoustic full waveform tomography in the presence of attenuation: A sensitivity analysis, *Geophys. J. Int.*, **195**, 985–1000.

Lailly, P., 1983. The seismic inverse problem as a sequency of before stack migrations, *Conference on Inverse Scattering*, pp. 206–220.

Lambare, G., Operto, S., Podvin, P., & Thierry, P., 2003. 3D ray+ Born migration/inversion–Part 1: Theory, *Geophysics*, **68**(4), 1348–1356.

Le Pichon, X. & Kreemer, C., 2010. The Miocene-to-Present Kinematic Evolution of the Eastern Mediterranean and Middle East and Its Implications for Dynamics, *Annu. Rev. Earth Planet. Sci.*, **38**, 323–351.

Ludwig, J. F., Nafe, J. E., & Drake, C. L., 1970. Seismic Refraction, in *The Sea*, edited by A. E. Maxwell, pp. 53–84, Wiley-Interscience, New York, 4th edn.

Makris, J., Abraham, Z. B., Behle, A., Ginzburg, A., Giese, P., Steinmetz, L., Whitmarsh, R. B., & Eleftheriou, S., 1983. Seismic refraction profiles between Cyprus and Israel and their interpretation, *Geophys. J. R. astr. Soc.*, **75**, 575–591.

Malinowski, M., Operto, S., & Ribodetti, A., 2011. High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full-waveform inversion, *Geophys. J. Int.*, **186**, 1179–1204.

Martin, G. S., Wiley, R., & Marfurt, K. J., 2006. Marmousi2: An elastic upgrade for Marmousi, *The Leading Edge*, **25**, 156–166.

McKenzie, D., 1972. Active Tectonics of the Mediterranean Region, *Geophys. J. R. astr. Soc.*, **30**, 109–185.

McPhee, P. J. & van Hinsbergen, D. J., 2019. Tectonic reconstruction of Cyprus reveals Late Miocene continental collision of Africa and Anatolia, *Gondwana Research*, **68**, 158–173.

Mereu, R. F., Epili, D., & Green, A. G., 1990. Pg shingles: preliminary results from the onshore GLIMPCE refraction experiment, *Tectonophysics*, **173**, 617–626.

Métivier, L. & Brossier, R., 2016. The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication, *Geophysics*, **81**(2), F1–F15.

Métivier, L., Brossier, R., Operto, S., & Virieux, J., 2013. Full Waveform Inversion and the Truncated Newton Method, *SIAM J. Sci. Comput.*, **59**(2), B401–B437.

Métivier, L., Bretaudeau, F., Brossier, R., Operto, S., & Virieux, J., 2014. Full waveform inversion and the truncated Newton method: Quantitative imaging of complex subsurface structures, *Geophysical Prospecting*, **62**, 1353–1375.

Montadert, L., Nicolaides, S., Semb, P. H., & Lie, Ø., 2014. Petroleum Systems Offshore Cyprus, in *Petroleum Systems of the Tethyan Region*, edited by L. Marlow, C. Kendall, & L. Yose, AAPG Memoir, 106th edn.

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

Mora, P., 1988. Elastic wave-field inversion of reflection and transmission data, *Geophysics*, **53**(6), 750–759.

Netzeband, G. L., Gohl, K., Hübscher, C. P., Ben-Avraham, Z., Dehghani, G. A., Gajewski, D., & Liersch, P., 2006. The Levantine Basin-crustal structure and origin, *Tectonophysics*, **418**, 167–188.

Nocedal, J. & Wright, S., 2006. *Numerical optimization, series in operations research and financial engineering*, 2nd edn.

Nocquet, J. M., 2012. Present-day kinematics of the Mediterranean: A comprehensive overview of GPS results, *Tectonophysics*, **579**, 220–242.

Operto, S., Virieux, J., Dessa, J. X., & Pascal, G., 2006. Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography: Application to the eastern Nankai trough, *Journal of Geophysical Research*, **111**(B09306).

Operto, S., Virieux, J., Ribodetti, A., & Anderson, J. E., 2009. Finite-difference frequency-domain modeling of viscoacoustic wave propagation in 2D tilted transversely isotropic (TTI) media, *Geophysics*, **74**(5), T75–T95.

Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Metiver, L., & Virieux, J., 2013. A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to pratice, *The Leading Edge*, **32**(9), 1040–1052.

Paige, C. C. & Saunders, M. A., 1982. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares, *ACM Transactions on Mathematical Software*, **8**(1), 43–71.

Papazachos, B. C. & Papaioannou, C. A., 1999. Lithospheric boundaries and plate motions in the Cyprus area, *Tectonophysics*, **308**, 193–204.

Pilidou, S., Priestley, K., Jackson, J., & Maggi, A., 2004. The 1996 Cyprus earthquake: a large, deep event in the Cyprean Arc, *Geophys. J. Int.*, **158**, 85–97.

Piromallo, C. & Morelli, A., 2003. P wave tomography of the mantle under the Alpine-Mediterranean area, *Journal of Geophysical Research: Solid Earth*, **108**(B2).

Plessix, R. E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.*, **167**, 495–503.

Plessix, R. E., Milcik, P., Rynja, H., Stopin, A., Matson, K., & Abri, S., 2013. Multiparameter full-waveform inversion: Marine and land examples, *The Leading Edge*, **32**(9), 1030–1038.

Portner, D. E., Delph, J. R., Berk Biryol, C., Beck, S. L., Zandt, G., Özacar, A. A., Sandvol, E., & Türkelli, N., 2018. Subduction termination through progressive slab deformation across Eastern Mediterranean subduction zones from updated P-wave tomography beneath Anatolia, *Geosphere*, **14**(3), 907–925.

Pratt, G. R. & Worthington, M. H., 1990. Inverse Theory Applied to Multi-Source Cross-Hole Tomography. Part 1: Acoustic Wave-Equation Method, *Geophysical Prospecting*, **38**, 287–310.

Pratt, R. G., 1990. Frequency-domain elastic wave modeling by finite differences: a tool for crosshole seismic imaging, *Geophysics*, **55**(5), 626–632.

Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model, *Geophysics*, **64**(3), 888–901.

Pratt, R. G., Song, Z. M., Williamson, P., & Warner, M., 1996. Two-dimensional velocity models from wide-angle seismic data by wavefield inversion, *Geophys. J. Int.*, **124**, 323–340.

Pratt, R. G., Shin, C., & Hicks, G. J., 1998. Gaussˆ Newton and full Newton methods in frequencyˆ space seismic waveform inversion, *Geophys. J. Int.*, **133**, 341–362.

Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O. I., & Kommedal, J. H., 2011. On the footprint of anisotropy on isotropic full waveform inversion: The Valhall case study, *Geophys. J. Int.*, **187**, 1495–1515.

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A., & Dell'Aversana, P., 2004. Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: Application to a thrust belt, *Geophys. J. Int.*, **159**, 1032–1056.

Reiche, S. & Hübscher, C., 2015. The Hecataeus Rise, easternmost Mediterranean: A structural record of Miocene-Quaternary convergence and incipient continent-continent-collision at the African-Anatolian plate boundary, *Marine and Petroleum Geology*, **67**, 368–388.

Reiche, S., Hübscher, C., & Ehrhardt, A., 2016. The impact of salt on the late Messinian to recent tectonostratigraphic evolution of the Cyprus subduction zone, *Basin Research*, **28**(5), 569–597.

Reilinger, R., McClusky, S., Vernant, P., Lawrence, S., Ergintav, S., Cakmak, R., Ozener, H., Kadirov, F., Guliev, I., Stepanyan, R., Nadariya, M., Hahubia, G., Mahmoud, S., Sakr, K., ArRajehi, A., Paradissis, D., Al-Aydrus, A., Prilepin, M., Guseva, T., Evren, E., Dmitrotsa, A., Filikov, S. V., Gomez, F., Al-Ghazzi, R., & Karam, G., 2006. GPS constraints on continental deformation in the Africa-Arabia-Eurasia continental collision zone and implications for the dynamics of plate interactions, *Journal of Geophysical Research: Solid Earth*, **111**(B05411), 1–26.

Robertson, A., Eaton, S., Follows, E., & McCallum, J., 1991. The Role of Local Tectonics Versus Global Sea-Level Change in the Neogene Evolution of the Cyprus Active Margin, in *Sedimentation, Tectonics and Eustasy: Sea-Level Changes at Active Margins*, edited by M. D. I. M, pp. 331–369, Blackwell Scientific.

Robertson, A. H., 1998. Tectonic significance of the Eratosthenes Seamount: A continental fragment in the process of collision with a subduction zone in the eastern Mediterranean (Ocean Drilling Program Leg 160), *Tectonophysics*, **298**, 63–82.

Robertson, A. H., Emeis, K.-c., Richter, C., Blanc-valleron, M.-m., Bouloubassi, I., Brumsack, H.-j., Cramp, A., Di Stefano, G. J., Flecker, R., Frankel, E., Howell, M. W., Janecek, T. R., Jurado, M. J., S Kemp, A. E., Koizumi, I., Kopf, A., Major, C. O., Mart, Y., C Pribnow, D. F., Rabaute, A., Roberts, A. E., Rullkotter, J., Sakamoto, T., Spezzaferri, S., Staerker, T. S., Stoner, J. S., Whiting, B. M., & Woodside, J. M., 1998. Collision-related break-up of a carbonate platform (Eratosthenes Seamount) and mud volcanism on the Mediterranean Ridge: preliminary synthesis and implications of tectonic results of ODP Leg 160 in the Eastern Mediterranean Sea, *Geological Society, London, Special Publications*, **131**, 242–271.

Robertson, A. H. F. & Shipboard Scientific Party, 1996. Tectonic Introduction, *Proceedings of the Ocean Drilling Program, Initial Reports*, **160**, 5–20.

Robertson, A. H. F., Parlak, O., & Ustaomer, T., 2012. Overview of the Palaeozoic-Neogene evolution of Neotethys in the Eastern Mediterranean region (southern Turkey, Cyprus, Syria), *Petroleum Geoscience*, **18**, 381–404.

Schattner, U., 2010. What triggered the early-to-mid Pleistocene tectonic transition across the entire eastern Mediterranean?, *Earth and Planetary Science Letters*, **289**, 539–548.

Schildgen, T. F., Yildirim, C., Cosentino, D., & Strecker, M. R., 2014. Linking slab break-off, Hellenic trench retreat, and uplift of the Central and Eastern Anatolian plateaus, *Earth-Science Reviews*, **128**, 147–168.

Sears, T. J., Barton, P. J., & Singh, S. C., 2010. Elastic full waveform inversion of multicomponent ocean-bottom cable seismic data: Application to Alba Field, U. K. North Sea, *Geophysics*, **75**(6), R109–R119.

Segev, A., Sass, E., & Schattner, U., 2018. Age and structure of the Levant basin, Eastern Mediterranean, *Earth-Science Reviews*, **182**, 233–250.

Shin, C. & Cha, Y. H., 2009. Waveform inversion in the Laplace-Fourier domain, *Geophys. J. Int.*, **177**, 1067–1079.

Shin, C. & Ha, W., 2008. A comparison between the behavior of objective functions for waveform inversion in the frequency and Laplace domains, *Geophysics*, **73**(5), VE119–VE133.

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

Shin, C., Min, D. J., Marfurt, K. J., Lim, H. Y., Yang, D., Cha, Y., Ko, S., Yoon, K., Ha, T., & Hong, S., 2002. Traveltime and amplitude calculations using the damped wave solution, *Geophysics*, **67**(5), 1637–1647.

Shipp, R. M. & Singh, S. C., 2002. Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data, *Geophys. J. Int.*, **151**, 325–344.

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

Stampfli, G. M. & Borel, G., 2002. A plate tectonic model for the Paleozoic and Mesozoic constrained by dynamic plate boundaries and restored syntetic ocean isochrons, *Earth and Planetary Science Letters*, **196**, 17–33.

Stampfli, G. M., Hochard, C., Vérard, C., Wilhem, C., & VonRaumer, J., 2013. The formation of Pangea, *Tectonophysics*, **593**, 1–19.

Symeou, V., Homberg, C., Nader, F. H., Darnault, R., Lecomte, J. C., & Papadimitriou, N., 2018. Longitudinal and Temporal Evolution of the Tectonic Style Along the Cyprus Arc System, Assessed Through 2-D Reflection Seismic Interpretation, *Tectonics*, **37**, 30–47.

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

Tarantola, A., 1984. Linearized Inversion of Seismic Reflection Data, *Geophysical Prospecting*, **32**, 998–1015.

ten Brink, U. S., Al-Zoubi, A. S., Flores, C. H., Rotstein, Y., Qabbani, I., Harder, S. H., & Keller, G. R., 2006. Seismic imaging of deep low-velocity zone beneath the Dead Sea basin and transform fault: Implications for strain localization and crustal rigidity, *Geophysical Research Letters*, **33**(L24314), 1–6.

Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics*, **51**(10), 1954–1966.

Toksöz, N. M. & Johnston, D. H., 1981. *Seismic Wave Attenuation*, Socoiety of Exploration Geophysicists, Tulsa, OK.

Tonn, R., 1991. the Determination of the Seismic Quality Factor Q From Vsp Data: a Comparison of Different Computational Methods, *Geophysical Prospecting*, **39**, 1–27.

van Hinsbergen, D. J., Torsvik, T. H., Schmid, S. M., Maţenco, L. C., Maffione, M., Vissers, R. L., Gürer, D., & Spakman, W., 2020. Orogenic architecture of the Mediterranean region and kinematic reconstruction of its tectonic evolution since the Triassic, *Gondwana Research*, **81**, 79–229.

Vidal, N., Alvarez-Marrón, J., & Klaeschen, D., 2000. Internal configuration of the Levantine Basin from seismic reflection data (eastern Mediterranean), *Earth and Planetary Science Letters*, **180**, 77–89.

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

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C., Conroy, G., & Bertrand, A., 2013. Anisotropic 3D full-waveform inversion, *Geophysics*, **78**(2), R59–R80.

Watremez, L., Helen Lau, K. W., Nedimović, M. R., & Louden, K. E., 2015. Traveltime tomography of a dense wide-angle profile across Orphan Basin, *Geophysics*, **80**(3), B69–B82.

Weber, M., Abu-Ayyash, K., Abueladas, A., Agnon, A., Al-Amoush, H., Babeyko, A., Bartov, Y., Baumann, M., Ben-Avraham, Z., Bock, G., Bribach, J., El-Kelani, R., Förster, A., Förster, H. J., Frieslander, U., Garfunkel, Z., Grunewald, S., Götze, H. J., Haak, V., Haberland, C., Hassouneh, M., Helwig, S., Hofstetter, A., Jäckel, K. H., Kesten, D., Kind, R., Maercklin, N., Mechie, J., Mohsen, A., Neubauer, F. M., Oberhänsli, R., Qabbani, I., Ritter, O., Rümpker, G., Rybakov, M., Ryberg, T., Scherbaum, F., Schmidt, J., Schulze, A., Sobolev, S., Stiller, M., Thoss, H., Weckmann, U., & Wylegalla, K., 2004. The crustal structure of the Dead Sea Transform, *Geophys. J. Int.*, **156**, 655–681.

Welford, J. K., Hall, J., Hübscher, C., Reiche, S., & Louden, K., 2015a. Crustal seismic velocity structure from Eratosthenes Seamount to Hecataeus Rise across the Cyprus Arc, eastern Mediterranean, *Geophys. J. Int.*, **200**, 933–951.

Welford, J. K., Hall, J., Rahimi, A., Reiche, S., Hübscher, C., & Louden, K., 2015b. Crustal structure from the hecataeus rise to the levantine basin, easternmediterranean, from seismic refraction and gravity modelling, *Geophys. J. Int.*, **203**, 2005–2069.

Williams, C., 2020. *Wide-Angle, Full Waveform Inversion with a Sparse Ocean-Bottom Seismometer Dataset, Imaging the Cyprus Arc from the Eratosthenes Seamount to the Hecataeus Rise*, Master of science, Memorial University of Newfoundland.

Williams, C., Malcolm, A., & Welford, J. K., 2019a. A multiscale full-waveform invserion strategy for sparse wide-angle datasets, in *SEG Techanical Program Expanded Abstracts*, San Antonio.

Williams, C., Malcolm, A., & Welford, J. K., 2019b. Structure of the Eratosthes Seamount through Wide-Angle, Full Waveform Imaging with a Sparse Dataset, in *AGU Fall Meeting*, pp. T23B–07, San Francisco.

Williamson, P. R., 1991. A guide to the limits of resolution imposed by scattering in ray tomography, *Geophysics*, **56**(2), 202–207.

Winkler, K., Nur, A., & Galdwin, M., 1979. Friction and seismic attenuation in rocks, *Geophysical Research Letters*, **227**, 528–531.

Wu, Y., Louden, K. E., Funck, T., Jackson, H. R., & Dehler, S. A., 2006. Crustal structure of the central Nova Scotia margin off Eastern Canada, *Geophysical Journal International*, **166**, 978–906.

Yilmaz, Ö., 2001. *Seismic Data Analysis: Processing, Inversion and Interpretation of Seismic Data*, Socoiety of Exploration Geophysicists, 2nd edn.

Zelt, C. A. & Smith, R. B., 1992. Seismic traveltime inversion for 2-D crustal velocity structure, *Geophys. J. Int.*, **108**, 16–34.