Exploring Full Waveform Inversion and Wavefield Gradient Measurements

椋本, 浩太

https://hdl.handle.net/2324/7182447

出版情報：Kyushu University, 2023, 博士（工学）, 課程博士
バージョン：
権利関係：

Exploring Full Waveform Inversion and Wavefield Gradient Measurements

Kota Mukumoto

Department of Earth Resources Engineering
Kyushu University

This dissertation is submitted for the degree of
Doctor of Philosophy

2024

Dedicated to my incredibly supportive parents and to my younger self,
who conquered the challenges of Ph.D. student life

Abstract

This thesis presents the application, theoretical investigations, and methodological developments of seismic imaging, in especially for full waveform inversion (FWI).

We start with the application of the FWI to image subsurface structures of Japan island. In this application, we aookued the FWI to the large volume of seismic data from the dense seismic network in Japan. The new seismic velocity model based on full computations of seismic wavefields is presented.

Subsequently, we focus on the theoretical investigations of FWI based on the wavefield gradients such as strain and rotational. Recent advancements in sensor technology have enabled the acquisition of wavefield gradient measurements. Although these measurements hold great potential for studying seismic waves, their unique sensitivity to much smaller-scale heterogeneities than the minimum wavelength poses challenges for FWI. We elucidate this distinctive challenge encountered in FWI when utilizing wavefield gradient measurements and propose approaches to mitigate the impact of small-scale heterogeneities.

Lastly, we explore the distinctive nature of wavefield gradient measurements to visualize small-scale structures with exceptionally high spatial resolution. Using distributed acoustic sensing (DAS), an emerging seismic sensor, we demonstrate through numerical simulations and real field experiments that strain measurements obtained by DAS can accurately delineate the location of embedded subsurface obstacles, showcasing unprecedented spatial resolution.

Acknowledgements

How lucky I am to have been surrounded by remarkable, supportive, and enthusiastic people in my Ph.D. journey. I would like to express my gratitude to everyone who has been involved.

Many thanks to supervisor Takeshi Tsuji. You definitely opened the doors to the world of science for me. I have enjoyed working with you. I have been impressed by your passion and hard works. I often wonder if there could be more than one of you, perhaps many clones, considering the vast amount of hard work you accomplish for science.

My great advisor Tatsunori Ikeda. Thanks for always having constructive discussions, funny chating, and enjoyable dinner. Among my favorite moments were leaving the laboratory together, engaging in enriching conversations while catching the same bus.

Special thanks to Yann Capdeville. Meeting you during my Ph.D. was undoubtedly the most important event. I am really happy to have the chance to work with you, as your papers significantly impressed me during my master degree. Our study together in Nantes provided me with an array of invaluable assets: theoretical knowledge, technical skills, a research philosophy, and also real life tips to enjoy the life.

To past and current loboratory's colleagues, thank you for your support, good talks, and eating together. It has reduced the pressure. A special thank you to Andri Hendriyana. Meeting you during my master's degree was instrumental in my decision to pursue a Ph.D. Your daily teachings on the fundamentals of seismology, often lasting until late at night, built my basement in seismology, particularly as I lacked a background in seismology. I recall our mountain climb together vividly, although I struggled to keep pace with you, impressed by your remarkable strength.

To Rodney, Shivangi, and Frank: Thank you for your great support during my stay in Nantes and always reminding me to enjoy staying in France, not only the research. The trip to La Rochelle, which you took me on, became one of my favorite places

To my fiance, Yako: Your cheerful spirit has been a tremendous help to me. Even though we were miles apart, chatting with you every day over the phone between Japan and France was a great support. I apologize for often saying, 'I have to go laboratory on that day'.

Finally, many thanks to my parents, Kimiko and Satoshi, for watching over me, supporting me, a being good family, and your love.

Jump across the fire to me! If you'll jump across the fire to me ...
Yukio Mishima (Shiosai)

Table of contents

List of figures

List of tables

Chapter 1

Introduction

The greatest part of human's questions arise from the surface level phanomena, which are visible and apparent. However, the most significant and impactful factors are hidden beneath the surface, eluding our perception and understanding.

This holds true for our planet Earth. On the surface, we witness the tangible results of Earth's activities—earthquakes, volcanic eruptions, and astonishing landscapes. Yet, the origins of these occurrences remain hidden beneath the surface, delving into the profound depths of the subsurface where mysteries unfold, unseen and enigmatic.

Seismic imaging is a powerful tool for visualizing subsurface structures using seismic waves, unraveling the enigmas of Earth's activities. This thesis presents the applications, theoretical foundations, and methodological advancements of seismic imaging. A short introduction is presented below.

1.1 Seismic wave

The seismic wave is propagating the subsurface, triggered by the natural phenomena such as earthquakes and volcanic activity, as well as by human-made events like nuclear tests and daily activities. The seismic wave is reflected and refracted at the boundaries of the subsurface as depicted in Fig.1.1. While traversing the subsurface, the wave encounters variations in speed, slowing down or accelerating, attributed to the different material densities of the Earth, ranging from soft to hard. In essence, the seismic wave contains a wealth of information about the Earth's subsurface structures.

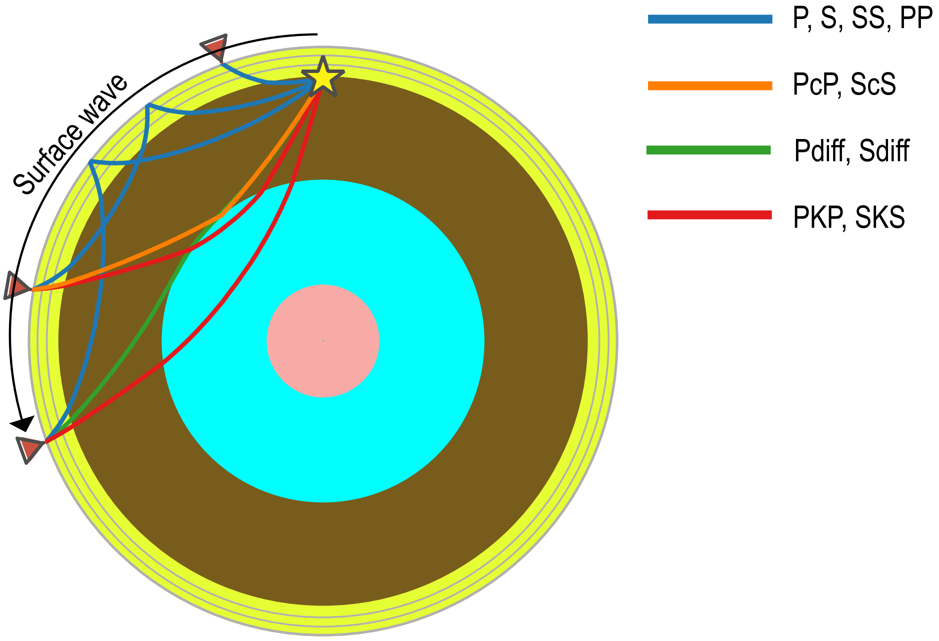


Fig. 1.1 The schematic illustration depicts ray paths of various seismic phases. Different seismic phases penetrate to different depths within the Earth. Starting from the inner part, the inner core (pink), outer core (blue), mantle (brown), and crust (yellow). Earthquake is denoted by stars, while seismic stations are represented by triangles. The ray paths were calculated and visualized using ObsPy modules (Beyreuther *et al.* 2010) with the velocity model proposed by (Kennett *et al.* 1995).

1.2 Seismic tomography

Seismic tomography is a powerful seismic imaging method for unveiling subsurface structures. By extracting valuable information embedded within seismic waves, seismic tomography offers a sophisticated lens to illuminate the hidden complexities of the Earth's subsurface. This method is similar to CT technology, which is used to image the internal structures of the human body. Both methods rely on the principles of wave physics to reconstruct detailed images of the subsurface, whether it's the Earth's structures or the human body.

The study, in many literatures, often referred as pioneering work of seismic tomography is the study by Aki & Lee (1976). They used first-arrival times of P-wave excited by earthquakes to determine hypocenter locations and subsurface medium parameters along straight ray path between hypocenters and receivers in the subsurface of Sanandreas fault zone. It's noteworthy that the development and application of seismic tomography dates back as early as 1971 in the study by Bois *et al.* (1971). Then, they applied their developed seismic tomography to cross-hole seismic data in oil fields (Bois *et al.* 1972). Their approach differed from Aki's study as they employed the shooting method, which can account for ray bending caused by heterogeneous structures. Around the same time, the first attempt to estimate a global Earth seismic velocity model was conducted (Dziewonski & Anderson 1981), leading to the establishment of the well-known Earth model called PREM. In their study, they utilized multiple seismic phases, including free oscillation, surface waves, and P- and S-body waves. The P- and S-body waves considered in their study were not restricted to just single P and S phases but also included other phases such as PcP and SKS. The integrated use of different seismic phases is crucial for resolving the entire Earth's structure because these various seismic phases penetrate to different depths even when observed from the same location (Fig.1.1). Since the pioneering study of global seismic tomography by Dziewonski & Anderson (1981), numerous studies of regional and global seismic tomography has been performed (Dziewonski 1984; Woodhouse & Dziewonski 1984; Kennett & Engdahl 1991; Fukao *et al.* 1992; Kennett *et al.* 1995; Mégnin & Romanowicz 2000; Tape *et al.* 2010; Fichtner *et al.* 2009b; Bozdağ *et al.* 2016).

In advance, we shortly introduce the inversion process of seimic tomography. First, the forward modeling of synthetic data $\mathbf{s}$, which corresponds to observed data $\mathbf{d}$, is defined as:

$$\mathbf{s} = \mathbf{G}(\mathbf{m})\,. \tag{1.1}$$

Here, $\mathbf{m}$ represents the model parameters, and $\mathbf{G}$ is the forward operator that relates the model $\mathbf{m}$ to the data $\mathbf{s}$. The associated inverse problem for estimating $\mathbf{m}$ is often formulated as a least-squares problem, which minimizes the misfit function between the observed $\mathbf{d}$ and synthetic $\mathbf{s}(\mathbf{m})$ data. The classical least-squares misfit function is defined as:

$$E(\mathbf{m}) = \frac{1}{2}||\mathbf{s}(\mathbf{m}) - \mathbf{d}||^2. \tag{1.2}$$

By expanding $\mathbf{E}(\mathbf{m})$ using a Taylor series and making a quadratic approximation, we can find the perturbation of model parameters $\delta\mathbf{m}$ that locally minimizes $\mathbf{E}(\mathbf{m})$, expressed as:

$$\delta\mathbf{m} = -\mathbf{H}^{-1}\mathbf{F}(\mathbf{s}(\mathbf{m}) - \mathbf{d})\,. \tag{1.3}$$

Here, $\mathbf{H}$ and $\mathbf{F} = \frac{\partial \mathbf{s}(\mathbf{m})}{\partial \mathbf{m}}$ represent the Hessian matrix and so called Frechet derivatives respected to model parameters. The model is iteratively updated by adding the estimated $\delta\mathbf{m}$ until convergence is achieved.

It is obvious that computation of Frechet derivatives requires computing forward modelling expressed by equation (1.1). Based on the methods used for computing seismic wave propagation and Frechet derivatives, in this study, we introduce three seismic tomography methods; seismic tomography based on geometrical ray theory, geometrical ray theory combined with Born approximation, and full waveform inversion (FWI).

Seismic tomography based on geometrical ray theory

The initial approach to seismic tomography relied on the high-frequency approximation of geometrical ray theory, where wave propagation is represented as rays carrying the energy of high-frequency waves along their paths. In this method, the synthetic data and Frechet derivatives are computed using forward modeling techniques such as eikonal solver and WKBJ seismograms, which are based on the geometrical ray theory.

The methods have a significant limitation: they ignored scattering effects, confining the Frechet derivatives, and consequently model estimations, to the line along the ray path between the source and receivers. However, it has been recognized that within the Fresnel volume (depicted in Figure 1.2), a tube-like area surrounding the central ray, multiple rays contribute to the observed wavefield at an observation point due to scattering (finite-frequency effects). In the presence of only the high-frequency approximation, the Fresnel volume converges to the central ray and the approximation is valid.

Real seismic wavefields are finite-frequency, and observed wavefields are influenced by heterogeneities outside the central ray. This contradiction between finite-frequency wave phenomena and the approximation underlying geometrical ray theory can distort seismic tomography results. For example, as shown in Figure 1.2, a velocity anomaly exists outside the ray path which the ray theory with high-frequency approximation expects. However, the observed seismic wave is affected by this velocity anomaly within the Fresnel volume. Consequently, seismic tomography based on geometrical ray theory may inaccurately attribute the misfit caused by the velocity anomaly to model parameters along the ray path, where the velocity anomaly is absent.

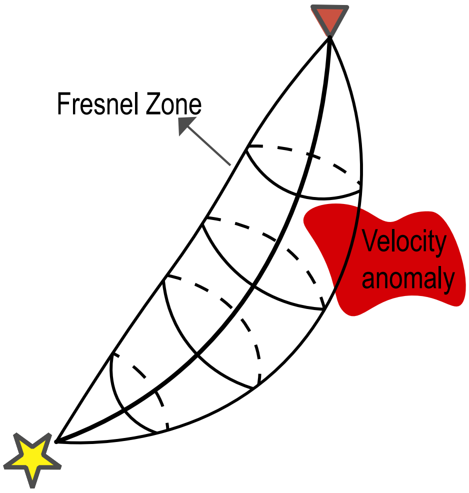


Fig. 1.2 The schematic image depicts the Fresnel volume, which influences the wavefield obtained at the seismic station. Traditional seismic ray tomography considers only the central ray path. However, in this schematic representation, the wavefield observed at the seismic station is affected by a velocity anomaly located outside the ray path. This discrepancy potentially leads to distortions in the resulting seismic tomographic images.

Geometrical ray theory + Born approximation

To address finite-frequency effects, significant advancements have been made, notably with the introduction of the Born approximation (Snieder 1988; Yomogida 1992; Dahlen *et al.* 2000). This approach accounts for first-order scattering effects into Frechet derivatives calculations. The incorporation of scattering effects led to substantial improvements in seismic tomography, as convincingly demonstrated by Montelli *et al.* (2004), which resulted in significant discoveries including the identification of new mantle plumes.

However, during the initial stages of methods based on the Born approximation, limited computational power necessitated the use of geometrical ray theory for computing synthetic data and Frechet derivatives. This choice was made to ensure computational resources remained within acceptable limits. It's worth noting that the high-frequency approximation within ray theory could potentially yield inaccurate model estimations that fail to fully capture certain wave propagation phenomena. For instance, numerical comparisons between forward modeling using ray theory and finite-difference methods solving directly wave equations, revealed limitations in ray theory's ability to model diffracted waves due to its high-frequency approximation (George *et al.* 1987).

From geometrical ray theory to full wavefield modeling

The rapid advancement in computational power has empowered us to simulate seismic wave propagation using sophisticated direct wave equation solvers, such as the finite-difference method (Boore 1972; Virieux 1986), direct solution method (Hara *et al.* 1993), and spectral element method (Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999). Computation of Frechet derivatives is efficiently achieved using these solvers through the adjoint method (Tarantola 1984), involving two distinct wavefield simulations: one originating from seismic sources and the other from seismic sensors. Frechet derivatives obtained through comprehensive wavefield simulations encompass all aspects of seismic wave propagations, leading to significantly enhanced accuracy in seismic tomographic images. Furthermore, this methodology naturally integrates multiple seismic phases, eliminating the need to confine the inversion to specific phases, thereby enhancing the overall reliability of the results.

In the context of this study, we term this approach based on full wavefield modeling as "full waveform inversion (FWI)". And the FWI is standing as the primary focus of this thesis.

1.3 Multi-scale problem in seismic tomography

The seismic data is always band-limited with minimum/maximum frequency, hence, minium/maximum wavelength present in its wavelength spectrum. It is well known that band-limited seismic wave is not affected by much smaller-scale subsurface structures than minimum wavelength, indeed, the seismic data lacks the information of smaller-scale structures. As a result, seismic tomography can only effectively estimate smooth structures. However, the concept of "smooth structures by seismic tomography" needs clarification because the Earth is, not smooth, heterogeneous at any scales, and estimated subsurface structures are different from true Earth structures.

The first success to clarify the "smooth structures by seismic tomography" has been done by Capdeville & Métivier (2018). Based on the two-scale homogenization theory (Capdeville *et al.* 2010a; Capdeville *et al.* 2010b), they showed that estimated subsurface structures by FWI using translational motions (displacement, velocity, and acceleration) are homogenized (effective) true Earth structures using fact that translational motions is dependent only on homogenized structures. This is the great leap forward to further understand about the seismic tomography.

On the other hand, the two-scale homogenization theory also raises a crucial question in FWI. While it reveals that translational motions depend solely on homogenized structures, it highlights that their spatial derivatives, namely seismic wavefield gradients, depend on both homogenized and small-scale structures (Capdeville 2021; Singh *et al.* 2020). This sparks the question: What happens if we use wavefield gradients for FWI instead of translational motions? Could it potentially enhance spatial resolution?

1.4 Outline

This thesis begins by providing the basics of FWI. Following this, we present the application of the FWI to image the subsurface structures in Japan island. We then proceed to theoretical investigation of multi-scale problems in the FWI based on the wavefield gradients. Finally, we investigate the simple method to image the location of small-scale subsurface structures based on wavefield gradients through numerical simulation and real field experiments. A detailed description of each chapter follows bellow.

Chapter 2: Basics of FWI

This chapter describes the some fundamentals of FWI, concentrating on the foundational principles that will be used in Chapters 3 and 4. The primary focus of this chapter is the derivation of Frechet derivatives calculation and exploration of inversion methods commonly employed in FWI studies.

Chapter3: FWI in Japan

Chapter 2 presents the application of FWI for imaging subsurface structures in Japan island. Most of previous seismic tomography study in Japan have used seismic tomography method based on geometrical ray theory. This work aims to estimate the new subsurface model in Japan by leveraging FWI in combination with the dense seismic station network available in the region. This study was published from Geophysical Journal International (Mukumoto & Tsuji 2023).

Chapter4: FWI based on wavefield gradients

This chapter is dedicated to the theoretical and numerical exploration of FWI using wavefield gradients (strain and rotational wavefields). We show the FWI based on wavefield gradients sufferred from the sensitivity of wavefield gradients to sub-wavelength subsurface structures. We discuss the difference between FWI using translational motions and wavefield gradients based on two-scale homogenization theory. Additionally, we propose a method to address the challenges posed by small-scale subsurface structures in FWI using wavefield gradients.

Chapter5: Detecting small-scale subsurface anomaly using DAS data

This last chapter explores the potential of imaging subsurface anomalies by leveraging the sensitivity of wavefield gradients to sub-wavelength heterogeneities. We show the arrays of sensor to obtain wavefield gradients possibly images the location of subsurface anomaly, even when the wavelength is much larger than the size of anomalies. As an illustrative case, we demonstrate the imaging capability of small-scale anomalies using a novel seismic sensor—distributed acoustic sensing (DAS) data—which records strain wavefields through both synthetic tests and real field experiments.

Chapter 2

Basics of full waveform inversion

2.1 Wave equation

We consider a elastic domain Ω (2D or 3D) enclosed by its boundary $\partial\Omega$, and cartesian coordinate denoted by $\mathbf{x} = \{x, y, z\}$ or $\mathbf{x} = \{x, z\}$ is adopted in Ω. The particle displacement $\mathbf{u}(\mathbf{x}, t)$ in Ω is manipulated by the elastic wave equation:

$$\rho\partial_{tt}\mathbf{u}(\mathbf{x}, t) - \nabla \cdot \boldsymbol{\sigma}(\mathbf{x}, t) = \boldsymbol{f}_s(\mathbf{x}_s, t)\,, \tag{2.1}$$

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \mathbf{c}(\mathbf{x}) : \boldsymbol{\epsilon}(\mathbf{u}(\mathbf{x}, t))\,, \tag{2.2}$$

where $\epsilon_{kl} = \frac{1}{2}(\partial_k u_l + \partial_l u_k)$ is the strain operator, $\rho(\mathbf{x})$ is the density, $\mathbf{c}(\mathbf{x})$ is the elastic tensor, $\boldsymbol{\sigma}(\mathbf{x}, t)$ is the stress tensor, $\boldsymbol{f}$ is the source term. The absorbing condition or free-surface boundary condition $\boldsymbol{\sigma} \cdot \mathbf{n} = 0$ are imposed on $\partial\Omega$. For the force, the source term can be:

$$\boldsymbol{f}_s(\mathbf{x}_s, t) = \mathbf{F}\delta(\mathbf{x} - \mathbf{x}_s)g(t)\,, \tag{2.3}$$

where $\mathbf{F}$ is the single force vector, $g(t)$ is the source term function, and $\mathbf{x}_s$ is the source location. For a moment tensor, source time can be described as:

$$\boldsymbol{f}_s(\mathbf{x}_s, t) = -\nabla\mathbf{M}\delta(\mathbf{x} - \mathbf{x}_s)g(t)\,, \tag{2.4}$$

where $\mathbf{M}$ is the moment tensor. Throughout this paper, we only consider point source.

For the ease of derivation of Frechet derivatives $\mathbf{F}$ in subsection 2.3, wave equation in frequency domain is used:

$$-\omega^2\rho(\mathbf{x})\mathbf{u}(\mathbf{x}, \omega) - \nabla \cdot \boldsymbol{\sigma}(\mathbf{x}, \omega) = \boldsymbol{f}_s(\mathbf{x}_s, \omega)\,. \tag{2.5}$$

2.2 Green tensor

For the simplicity, we use wave equation in frequency domain (equation (2.5)). The green tensor **G** is the earth response by impulsive source, namely

$$-\omega^2\rho(\mathbf{x})G_i^p(\mathbf{x},\omega;\mathbf{x}_s) - \partial_j c_{ijkl}\partial_k G_l^p(\mathbf{x},\omega;\mathbf{x}_s) = \delta(\mathbf{x}-\mathbf{x}_s)\delta_{ip} \tag{2.6}$$

where G_i^p is the i-th component of the response associated with p-th directional impulsive force. The green tensor is symmetric in the sense that

$$G_i^p(\mathbf{x},\omega;\mathbf{x}_s) = G_p^i(\mathbf{x}_s,\omega;\mathbf{x})\,. \tag{2.7}$$

Equation (2.7) shows the principle of seismic reciprocity, expressing source-receiver interchange. Based on the representation theorem (Dahlen & Tromp 1998) , the observation of i-th component displacement excited by p-th directional force can be expressed as:

$$u_i(\mathbf{x},\omega) = G_i^p(\mathbf{x},\omega;\mathbf{x}_s)F_p g(\omega)\,. \tag{2.8}$$

For the point moment tensor force, representation theorem reads:

$$u_i(\mathbf{x},\omega) = \partial_j^s G_i^p(\mathbf{x},\omega;\mathbf{x}_s)M_{pj} g(\omega)\,. \tag{2.9}$$

It is clear that impulsive response G_i^{pj} respected to pj-component of point moment tensor force is:

$$G_i^{pj}(\mathbf{x},\omega;\mathbf{x}_s) = \partial_j^s G_i^p(\mathbf{x},\omega;\mathbf{x}_s)\,. \tag{2.10}$$

In addition, the reciprocity of strain for point moment tensor can be written as (Ikelle & Amundsen 2018):

$$\partial_k G_i^{pj}(\mathbf{x},\omega;\mathbf{x}_s) = \partial_j^s G_p^{ik}(\mathbf{x}_s,\omega;\mathbf{x})\,. \tag{2.11}$$

2.3 Frechet derivatives

Here, we derive the Frechet derivative of the displacement **u** respected to model parameters **m**. The derived Frechet derivative is employed in the iterative algorithms to update model parameters, minimizing the misfit function. These algorithms will be explained in the next subsection.

Based on the wave equation, the born approximation reads (Tromp *et al.* 2005):

$$\delta \mathbf{u}_i(\mathbf{x}_r,\omega) = -\int_V \{\delta ln\rho(\mathbf{x})\dot{G}_i^j(\mathbf{x}_r,\omega;\mathbf{x})\dot{u}_j(\mathbf{x},\omega) - \delta lnc_{jklm}(\mathbf{x})\partial_k G_i^j(\mathbf{x}_r,\omega;\mathbf{x})\partial_l u_m(\mathbf{x},\omega)\}dV\,, \tag{2.12}$$

where dot subscript $\dot{}$ indicates time derivative, $\delta ln\rho = \delta\rho/\rho$, and $\delta lnc = \delta c/c$. Using reciprocity theorem (2.7), we can rewrite born approximation such as:

$$\delta \mathbf{u}_i(\mathbf{x}_r,\omega) = -\int_V \{\delta ln\rho(\mathbf{x})\dot{G}_j^i(\mathbf{x},\omega;\mathbf{x}_r)\dot{u}_j(\mathbf{x},\omega) - \delta lnc_{jklm}(\mathbf{x})\partial_k G_j^i(\mathbf{x},\omega;\mathbf{x}_r)\partial_l u_m(\mathbf{x},\omega)\}dV\,. \tag{2.13}$$

Based on equation (2.13), we deduce the frechet derivatives respected to density and elastic parameters respectively:

$$\frac{\partial u_i(\mathbf{x}_r,\omega)}{\partial\delta\rho} = \rho(\mathbf{x})\dot{G}_j^i(\mathbf{x},\omega;\mathbf{x}_r)\dot{u}_j(\mathbf{x},\omega)\,, \tag{2.14}$$

$$\frac{\partial u_i(\mathbf{x}_r,\omega)}{\partial\delta c_{jklm}} = -\partial_k G_j^i(\mathbf{x},\omega;\mathbf{x}_r)c_{jklm}\partial_k u_m(\mathbf{x},\omega)\,. \tag{2.15}$$

Therefore, Frechet derivatives respected to model parameters can be computed taking cross correlation of two wavefields: one is from seismic sources (forward wavefield) and another is from receiver (adjoint wavefield) excited by single force (adjoint source). The method to compute Frechet derivatives based on equations (2.14) and (2.15) is called adjoint method.

In the capter 4, we study about wavefield gradients based FWI. Therefore, we also derive Frechet derivatives for wavefield gradients. Taking spatial derivatives ∂_q^r both sides of equation (2.14) and (2.15), and using relation (2.10), we have:

$$\frac{\partial\partial_q^r u_i(\mathbf{x}_r,\omega)}{\partial\delta\rho} = \rho(\mathbf{x})\dot{G}_j^{iq}(\mathbf{x},\omega;\mathbf{x}_r)\dot{u}_j(\mathbf{x},\omega)\,, \tag{2.16}$$

$$\frac{\partial\partial_q^r u_i(\mathbf{x}_r,\omega)}{\partial\delta c_{jklm}} = -\partial_k G_j^{iq}(\mathbf{x},\omega;\mathbf{x}_r)c_{jklm}(\mathbf{x})\partial_k u_m(\mathbf{x},\omega)\,. \tag{2.17}$$

From the computational perspective, the Frechet derivastives of wavefield gradients can be computed same way as displacement. The difference is just that adjoint wavefields for wavefield gradients are excited by moment tensor force. For the Frechet derivatives of strain and rotational wavefields, we rewrite equations (2.16) and (2.17) using coupling tensor **J** and

$[\mathbf{K}^{qi}]_{pn} = \mathbf{J}_{qipn}$:

$$J_{qipn}\frac{\partial \partial_p^r u_n(\mathbf{x}_r,\omega)}{\partial \delta\rho} = \rho(\mathbf{x})\dot{G}_j^{\mathbf{K}^{qi}}(\mathbf{x},\omega;\mathbf{x}_r)\dot{u}_j(\mathbf{x},\omega)\,, \tag{2.18}$$

$$J_{qipn}\frac{\partial \partial_p^r u_n(\mathbf{x}_r,\omega)}{\partial \delta c_{jklm}} = -\partial_k G_j^{\mathbf{K}^{qi}}(\mathbf{x},\omega;\mathbf{x}_r)c_{jklm}(\mathbf{x})\partial_k u_m(\mathbf{x},\omega)\,. \tag{2.19}$$

We use coupling tensor $\mathbf{J}$ such as, for strain:

$$\mathbf{J}_{qipn} = (\delta_{qp}\delta_{in} + \delta_{qn}\delta_{ip})/2\,, \tag{2.20}$$

and for rotation:

$$\begin{aligned} \mathbf{J}_{qipn} &= \theta_{ipn}/2 \quad \text{(in 3D)}\,, \\ \mathbf{J}_{qipn} &= \theta_{pn}/2 \quad \text{(in 2D)}\,, \end{aligned} \tag{2.21}$$

where θ is Leve-Civita symbol. Note that the adjoint moment tensor source for rotational wavefields is nonsymmetric, a condition typically considered physically invalid. This is due to the purely mathematical nature of the adjoint method.

2.4 Descent based inversion method

In this section, we provide an overview of commonly used iterative inversion methods for minimizing the misfit function. Without delving into rigorous mathematical details, the Frechet derivative of the misfit function $E(\mathbf{m}) \in \mathcal{R}$ is defined as linear map $A_\mathbf{m}$ at $\mathbf{m}$ (Yamada 2009), which is described as:

$$\begin{gathered} A_\mathbf{m} : \delta\mathbf{m} \in \mathcal{M} \to A_\mathbf{m}(\delta\mathbf{m}) \in \mathcal{R} \quad \text{such that} \\ \lim_{||\delta\mathbf{m}||_\mathcal{M} \to 0} \frac{||E(\mathbf{m}+\delta\mathbf{m}) - E(\mathbf{m}) - A_\mathbf{m}(\delta\mathbf{m})||}{||\delta\mathbf{m}||_\mathcal{M}} = 0\,, \end{gathered} \tag{2.22}$$

where $||\cdot||_\mathcal{M}$ is norm defined in parameter space $\mathcal{M}$. We denote the Euclidean norm as $||\cdot||$ and its corresponding inner product as $\langle\rangle$. Assuming that we can represent the linear mapping as an inner product: $A_m(\delta\mathbf{m}) = \langle E'(\mathbf{m}), \delta\mathbf{m}\rangle$ and adopting Euclidean distance for parameter space, Frechet derivative of misfit function $E'(\mathbf{m})$ is the vector such that inner product with $\delta\mathbf{m}$ approximates the change in the misfit function $E(\mathbf{m}+\delta\mathbf{m}) - E(\mathbf{m})$. Moreover, if we assume $||\delta\mathbf{m}|| = ||E'(\mathbf{m})||$, we have:

$$-||E'(\mathbf{m})||^2 \leq \langle E'(\mathbf{m}), \delta\mathbf{m}\rangle \leq ||E'(\mathbf{m})||^2 \,. \tag{2.23}$$

Therefore, the direction of the Frechet derivative corresponds to the direction that maximizes the misfit function (ascent direction). Conversely, the negative direction (descent direction) minimizes the misfit function.

We have confirmed that the negative direction of the Frechet derivative of the misfit function represents the descent direction. However, the choice of the metric to measure distance in parameter space $\mathcal{M}$ influences the descent direction. Generally, there is no a priori reason to measure distance in parameter space using Euclidean distance. Consequently, we introduce a general symmetric metric matrix $\mathbf{P}$ in parameter space. Indeed, $||\delta\mathbf{m}||^2_{\mathcal{M}}$ is written as:

$$||\delta\mathbf{m}||^2_{\mathcal{M}} = \delta\mathbf{m}^t\mathbf{P}\delta\mathbf{m} \,. \tag{2.24}$$

This distance is equivalent to the Euclidean distance of the point where $\delta\mathbf{m}$ maps to another space by $\mathbf{P}^{\frac{1}{2}}$ (Xing *et al.* 2002). Assuming $||\mathbf{P}\delta\mathbf{m}|| = ||E'(\mathbf{m})||$, we have:

$$-||E'(\mathbf{m})||^2 \leq \langle \mathbf{P}^{\frac{1}{2}}E'(\mathbf{m}), \mathbf{P}^{\frac{1}{2}}\delta\mathbf{m}\rangle = \langle E'(\mathbf{m}), \mathbf{P}\delta\mathbf{m}\rangle \leq ||E'(\mathbf{m})||^2 \,. \tag{2.25}$$

Therefore, it is clear that $\delta\mathbf{m} = -\mathbf{P}^{-1}E'(\mathbf{m})$ is the descent direction that minimizes the misfit function. In practice, to ensure the positive definiteness of the metric matrix during the practical computation, a damping parameter λ is used. Therefore, the iterative algorithm for minimizing the misfit function is described as follows:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n(\mathbf{P}_n + \lambda_n\mathbf{I})^{-1}\mathbf{F}_n(\mathbf{s}(\mathbf{m}_n) - \mathbf{d}) \,, \tag{2.26}$$

where subscript n indicates n-th iteration, $\mathbf{I}$ is identity matrix, and α is step length.

Newton method

The Newton method is a second-order optimization technique. Referring to Equation 1.3, the iterative procedure can be written as:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - (\mathbf{H}_n + \lambda_n\mathbf{I})^{-1}\mathbf{F}_n(\mathbf{s}(\mathbf{m}_n) - \mathbf{d}) \,. \tag{2.27}$$

When $\mathbf{P}_n = \mathbf{H}_n$ and $\alpha_n = 1$, the Newton method is identical to equation 2.26. The Hessian matrix can be regarded as one type of metric matrix (CROCKETT & CHERNOFF 1955; Tarantola 2005).

Gauss-Newton method

The challenge with the Newton method lies in computing the inverse Hessian matrix. As the number of model parameters increases, the size of the Hessian matrix grows larger, making computations infeasible, particularly for large inverse problems such as seismic tomography. The Gauss-Newton method approximates the Hessian matrix (Tarantola 2005). To derive the approximated Hessian, we start with equation 1.2. Applying the partial derivative with respect to one component of model parameters m^i, we obtain:

$$\frac{\partial E(\mathbf{m})}{\partial m^i} = \frac{\partial s^j(\mathbf{m})}{\partial m^i}(s^j(\mathbf{m}) - d^j)\,. \tag{2.28}$$

Taking another partial derivative with respect to model parameter m^k, we get:

$$\frac{\partial^2 E(\mathbf{m})}{\partial m^k \partial m^i} = \frac{\partial s^j(\mathbf{m})}{\partial m^i}\frac{\partial s^j(\mathbf{m})}{\partial m^k} + \frac{\partial^2 s^j(\mathbf{m})}{\partial m^k \partial m^i}(s^j(\mathbf{m}) - d^j)\,. \tag{2.29}$$

The computation of the second term is complex, and usually, this term is assumed to be small enough. Therefore, with $F_{ij} = \frac{\partial s^j(\mathbf{m})}{\partial m^i}$, the approximated Hessian is given by:

$$\mathbf{H} = \frac{\partial^2 E(\mathbf{m})}{\partial \mathbf{m}^2} \approx \mathbf{F}^t\mathbf{F}\,. \tag{2.30}$$

Substituting the approximated Hessian in equation (2.27), Gauss-Newton inversion can be described as:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - (\mathbf{F}^t\mathbf{F} + \lambda_n\mathbf{I})_n^{-1}\mathbf{F}_n(\mathbf{s}(\mathbf{m}_n) - \mathbf{d})\,, \tag{2.31}$$

There are other methods, such as the quasi-Newton method, that also use approximated Hessians. L-BFGS and BFGS algorithms are often employed. These methods approximate the Hessian with successive gradients **F** from previous iterations (Nocedal 1980).

Steepest descent method

The steepest descent method simply use identity matrix **I** as the metric matrix in equation (2.26). Therefore, the steepest descent method can be written as:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n\mathbf{F}_n(\mathbf{s}(\mathbf{m}_n) - \mathbf{d})\,. \tag{2.32}$$

Chapter 3

Full waveform inversion in Japan

3.1 Introduction

The Japan Islands are located along the convergent boundary of four tectonic plates, whereby the Philippine Sea and Pacific plates are subducting beneath the Eurasia and Okhotsk plates. The interactions of these four plates are responsible for many of Japan's large-scale tectonic features (Fig.3.1). The Itoigawa–Shizuoka Tectonic Line forms the terrestrial boundary between the Pacific and Eurasia plates. Subduction of the Philippine Sea plate created the Izu–Bonin collision zone owing to the ongoing collision of the subducting Izu–Bonin arc and the overriding Honshu arc. The area also includes two prominent structural features, the Median Tectonic Line and the Niigata–Kobe Tectonic Zone. These major tectonic features, together with the many active volcanoes and sedimentary basins across central Japan, yield a regional seismic structure with substantial lateral heterogeneities. Previous geophysical studies of central Japan have mainly employed regional- and exploration-scale seismic tomography to constrain the seismic structure (Nakajima & Hasegawa 2007a; Nakajima & Hasegawa 2007b; Nishida *et al.* 2008; Nakajima *et al.* 2009; Arai & Iwasaki 2014; Miyoshi *et al.* 2017; Nimiya *et al.* 2020). For example, studies using first-arrival tomography (Nakajima & Hasegawa 2007a; Nakajima & Hasegawa 2007b; Nakajima *et al.* 2009) have revealed the slab geometry of the Philippine Sea Plate and investigated plausible relationships between arc magmatism and subducting oceanic plates. Nishida *et al.* (2008) and Nimiya *et al.* (2020) have analysed the ambient noise wavefield using seismic interferometry and clearly imaged underground features, including magmatic fluids and thick sedimentary successions.

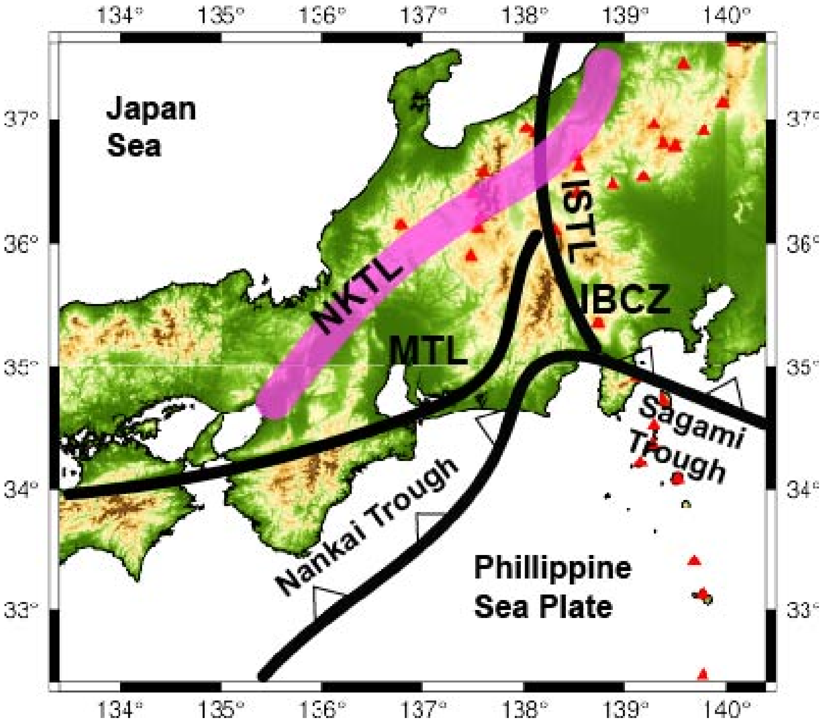


Fig. 3.1 Map of the study area. Black lines and the pink area indicate the major tectonic structures in the area: Itoigawa–Shizuoka Tectonic Line (ISTL), Median Tectonic Line (MTL), and Niigata–Kobe Tectonic Zone (NKTZ). Red triangles denote active volcanoes.

The recent development of full-waveform tomography techniques has improved our ability to resolve subsurface structures (Tarantola 1984; Mora 1987; Fichtner *et al.* 2009a; Fichtner 2010; Tape *et al.* 2010; Simutė *et al.* 2016; Bozdağ *et al.* 2016; Lei *et al.* 2020). This method enables the three-dimensional (3D) sensitivity distributions of the propagating seismic waves to be computed via full numerical simulations of seismic-wave propagation in heterogeneous media (Tromp *et al.* 2005; Fichtner 2010). In contrast to ray-based tomography, which approximates the seismic-wave propagation using high-frequency approximations of the seismic waves, full-waveform tomography utilises the complete waveform to extract all of the available information from the data. Miyoshi *et al.* (2017) have applied full-waveform tomography in the Kanto area using data from the F-net broadband seismic stations and obtained a more heterogeneous velocity model than that estimated using conventional ray-based tomography. Therefore, full-waveform tomography has the potential to better resolve the heterogeneous seismic velocity structures beneath central Japan.

The goals of this study are to resolve the crustal shear-wave (S-wave) velocity structures in central Japan using full-waveform tomography, and to evaluate its performance by comparing these results to those obtained via other tomographic methods. We apply full-waveform tomography to a large seismic waveform volume that was collected by the Hi-net and F-net networks (Okada *et al.* 2004). The resultant crustal S-wave velocity model was obtained after 17 iterations and effectively resolved the strong lateral velocity variations throughout the crust. The velocity anomalies in this model are in good agreement with the known geology of central Japan.

3.2 Method

3.2.1 3D seismic-wave simulation and initial model

Synthetic waveforms with a 240-s duration were calculated using the spectral element method, which is widely used in seismology owing to its accuracy and computational efficiency via parallel computing (Komatitsch & Tromp 1999; Komatitsch & Tromp 2002a). We used the SPECFEM3D Cartesian program for the 3D isotropic seismic-wave simulations (Peter *et al.* 2011). We calculated the seismic wavefields in our target model volume (133.5°–140.5°E × 32.5°–37.5°N × 0–90 km depth) using a mesh that is suitable for resolving seismic waves with 2.5-s periods and larger. The Clayton–Enquist–Stacey absorbing boundary condition was used (Clayton & Engquist 1977; Stacey 1988) in the simulations. We employed the laterally homogeneous JMA2001 seismic velocity model (Ueno 2002), which specifies the S- and compressional-wave (P-wave) velocity structures around Japan, as the initial velocity model. We confirmed that the JMA2001 model produced reasonable seismograms that were similar to the observations via a visual comparison of the simulations to the corresponding observations. We obtained the initial density structures using the empirical relationship between the P-wave velocity V_p and density ρ via the following relationship (Brocher 2005):

$$\rho = 1.6612V_p - 0.4721V_p^2 + 0.0671V_p^3 - 0.0043V_p^4 + 0.000106V_p^5, \tag{3.1}$$

where ρ is in g cm^{-3} and V_p is in km s^{-1}.

3.2.2 Computing the misfit function

High noise levels in the observational dataset will hamper the inversion, and increase the likelihood of obtaining incorrect solutions. Furthermore, cycle skipping can lead to a local minimum in the solution space of the waveform inversion that does not correspond to a best model. Cycle skipping can occur when the observed and synthetic waveforms are more than half a wavelength out of phase; therefore, careful data selection must be undertaken to prevent these incorrect representations of the seismic waveforms. We utilised the FLEXWIN program (Maggi *et al.* 2009) for the automatic determination of the time windows for the synthetic and observed waveform pairs. This time-window determination can be based on a number of parameters, including the time lag, cross-correlation coefficient between the observed and synthetic waveforms, and signal-to-noise ratio. We optimised the model parameters in two period ranges, 20–50 s and 10–30 s, and selected the time windows before the first iteration of each period range. We set the maximum time lag between the observed and synthetic waveforms to 12 s for the 20–50-s period range and 5 s for the 10–30-s period range during the time-window selection. We determined the time windows for 1,821 waveform pairs in the 20–50-s period range and 13,052 pairs in the 10–30-s period range for the full-waveform tomographic inversion.

The misfit between the synthetic and observed waveforms was quantified on the basis of the phase misfit using the time–frequency transform (Fichtner *et al.* 2008):

$$\mathrm{E}_{\mathrm{p}}^{2}(\mathrm{u}_{\mathrm{i}}, \mathrm{d}_{\mathrm{i}}) = \iint \mathrm{W}_{\mathrm{p}}^{2}(\mathrm{t}, \omega)\,[\phi_{\mathrm{i}}(\mathrm{t}, \omega; \mathbf{m}) - \phi_{\mathrm{i}}(\mathrm{t}, \omega)]^{2}\mathrm{d}\mathrm{t}\mathrm{d}\omega, \tag{3.2}$$

where d_{i} and u_{i} are the i-th component of the observed and synthetic velocity data, respectively, $\phi_{\mathrm{i}}(\mathrm{t}, \omega)$ and $\phi_{\mathrm{i}}(\mathrm{t}, \omega; \mathbf{m})$ are the time–frequency phases of the observed and synthetic velocity data, respectively, $\mathbf{m}$ is a model parameters, and W_p is the weighting function. The weighting function was chosen following Fichtner *et al.* (2008):

$$\mathrm{W}_{\mathrm{p}} = \log(1 + |\tilde{\mathrm{u}}_{\mathrm{i}}|) / \max_{\mathrm{t},\omega}\ \log(1 + |\tilde{\mathrm{u}}_{\mathrm{i}}|), \tag{3.3}$$

where $\tilde{\mathrm{u}}$ represents the velocity data on the time–frequency plane.

3.2.3 Updating the velocity model

We used a multiscale strategy that first recovered the smooth large-scale structures in the 20–50-s period range and then resolved the fine-scale structures in the 10–30-s period range. Only the S-wave velocities were updated after each iteration because the surface waves were the most energetic phase in our data.

We computed the gradient of the misfit function with respect to the model parameters using the adjoint method (Tromp *et al.* 2005; Fichtner 2010) and then updated the model using the steepest-descent method. Inverse Hessian kernels, which were approximated by their diagonal elements, were applied to the gradients. These approximate inverse Hessian kernels act as a preconditioner in above-mentioned gradient methods. The model update in the $i + 1$-th iteration is:

$$\mathbf{m}_{i+1} = \mathbf{m}_{i} - \alpha S[\tilde{\mathbf{H}}^{-1}(\mathbf{m}_{i})\mathbf{g}(\mathbf{m}_{i})]\,, \tag{3.4}$$

where $\mathbf{m}_i$ is the current model parameter; α is the step length, which scales with the descent direction; $\tilde{\mathbf{H}}(\mathbf{m})$ is the approximate Hessian; $\mathbf{g}(\mathbf{m})$ is the gradient of the misfit function; and S is the smoothing operator. Smoothing was performed via convolution with a 3D Gaussian. We set the horizontal and vertical standard deviations of the 3D Gaussian smoothing operator to 20 and 10 km for the 20–50-s period range, and 10 and 6 km for the 10–30-s period range, respectively. We used an iterative process to update the velocity model, starting with a step length that allowed us to update the model within a 2% maximum change. If the misfit reduction between the current and previous iteration was less than 1% over the past two iterations, then we doubled the maximum change in model to 4 %. We stopped the iteration when the change in the misfit between the current and previous iterations was less than 0.5%.

3.3 Data

The earthquake waveform data that were used in this study were collected by the nationwide Hi-net high-sensitivity seismograph network and F-net broadband seismograph network, which are both operated by the National Research Institute for Earth Science and Disaster Prevention (Okada *et al.* 2004). We used the seismogram records from 388 Hi-net and 28 F-net permanent stations (Fig. 3.2a), with all of these stations consisting of three-component velocity seismometers that are deployed in boreholes. We used the F-net stations for the inversions in both the 20–50-s and 10–30-s period ranges. We only used the Hi-net data for the inversions in the 10–30-s period range because the Hi-net seismometers have sensitivities of >1.0 Hz; we applied the sensitivity corrections proposed by Maeda *et al.* (2011) to enhance the long-period component of the Hi-net data for the inversions. We collected the seismogram records from 41 earthquakes that occurred during the 2004–2021 period, possessed moment magnitudes (Mw) in the $4.7 \leq Mw \leq 5.4$ range, and occurred at less than 60 km depth (Fig. 3.2b and c), thereby restricting the waveform data to seismogram records at source–receiver distances greater than 80 km. The number of available seismometers varied during 2004–2021 as new instruments were installed. We prepared an inversion dataset that consisted of 29 events for the velocity model inversion and an evaluation dataset that consisted of 12 events for the evaluation of the final velocity model. The earthquakes in the evaluation dataset were chosen to avoid locations that overlapped those in the inversion dataset (Figs. 3.2b and c). The earthquake parameters for the simulations were extracted from the global CMT catalogue (Ekström *et al.* 2012); these values were fixed while updating the velocity model to reduce the computational requirements of the inversion.

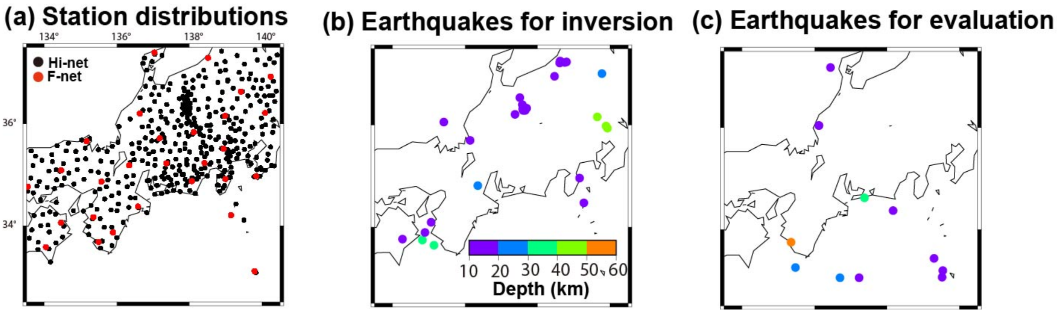


Fig. 3.2 Seismometer and earthquake distributions used in this study. (a) Hi-net (black circles) and F-net (red circles) station distributions. (b) Earthquakes used for the inversion. (c) Earthquakes used for the evaluation.

3.4 Results and Interpretation

The final S-wave velocity model was obtained after 11 iterations using the data in the 20–50-s period range, followed by 6 iterations using the data in the 10–30-s period range; four depth slices are shown in Fig. 3.3. We confirmed that the misfit converged to a minimum value after each set of iterations. The velocity model outside of the dashed lines in Fig. 3.3 denotes the model space in the inversion dataset that is insufficiently sensitive and cannot be accurately estimated (see section 5.1). Strong velocity variations are observed throughout the model: for example, the velocities range from <3000 m s^{-1} to ~4000 m s^{-1} at 5 km depth. The misfit was reduced by 16.4% during the 11 iterations using the data in the 20–50-s period range and a further 6.7% during the 6 iterations using the data in the 10–30-s period range (Fig. 3.4).

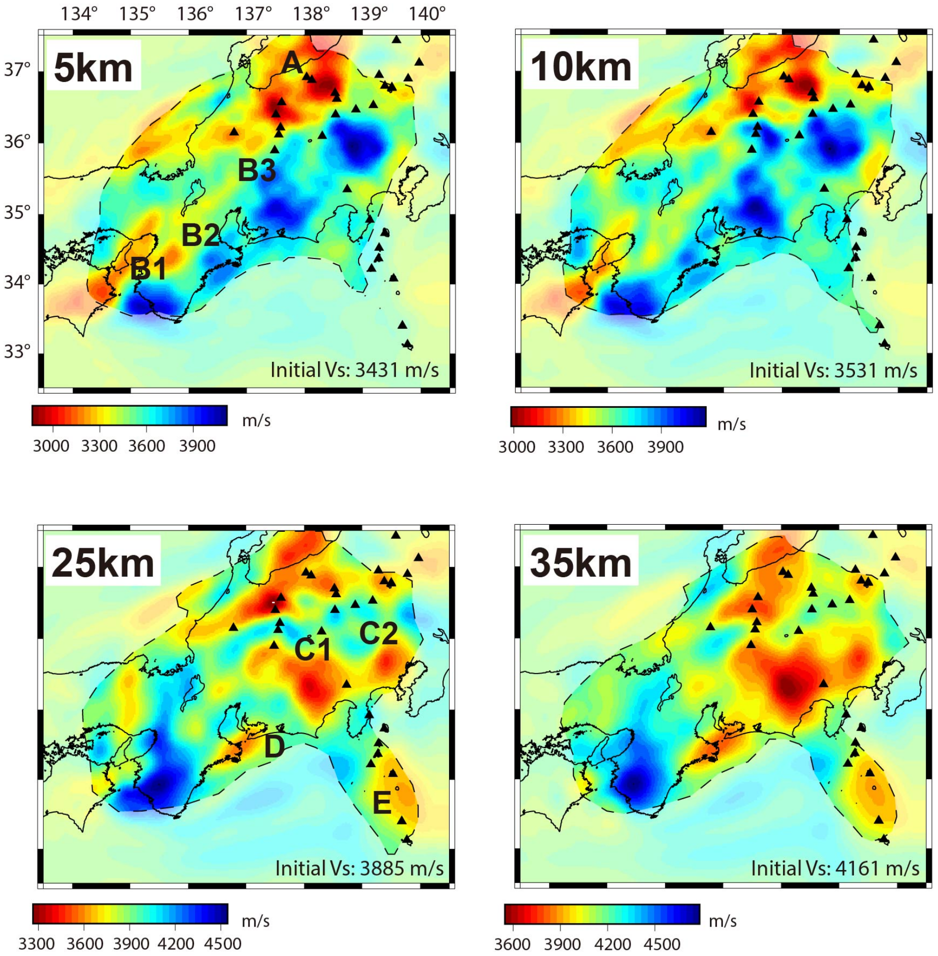


Fig. 3.3 Depth slices (5, 10, 25 and 35 km) through the resultant S-wave velocity model. The region enclosed by the black dashed line is used in the interpretations due to its enhanced sensitivity to model perturbations, as explained in section 5. The black triangles denote active volcanoes. The S-wave velocity of the initial model for a given depth slice is provided in the bottom-right corner of each panel.

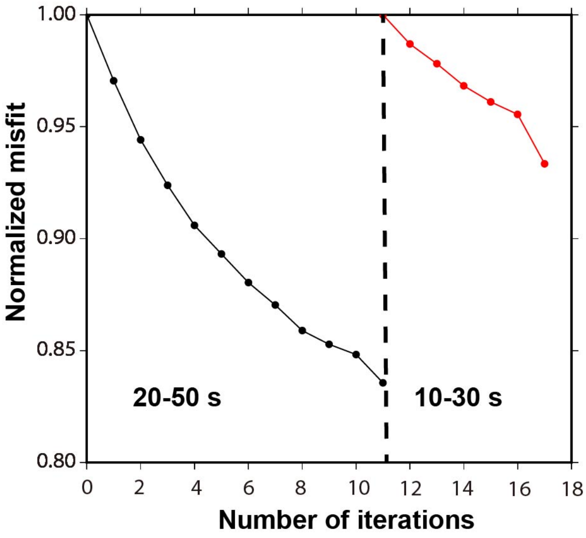


Fig. 3.4 Misfit reduction. Normalised misfit reductions in the 20–50-s (black) and 10–30-s (red) period ranges.

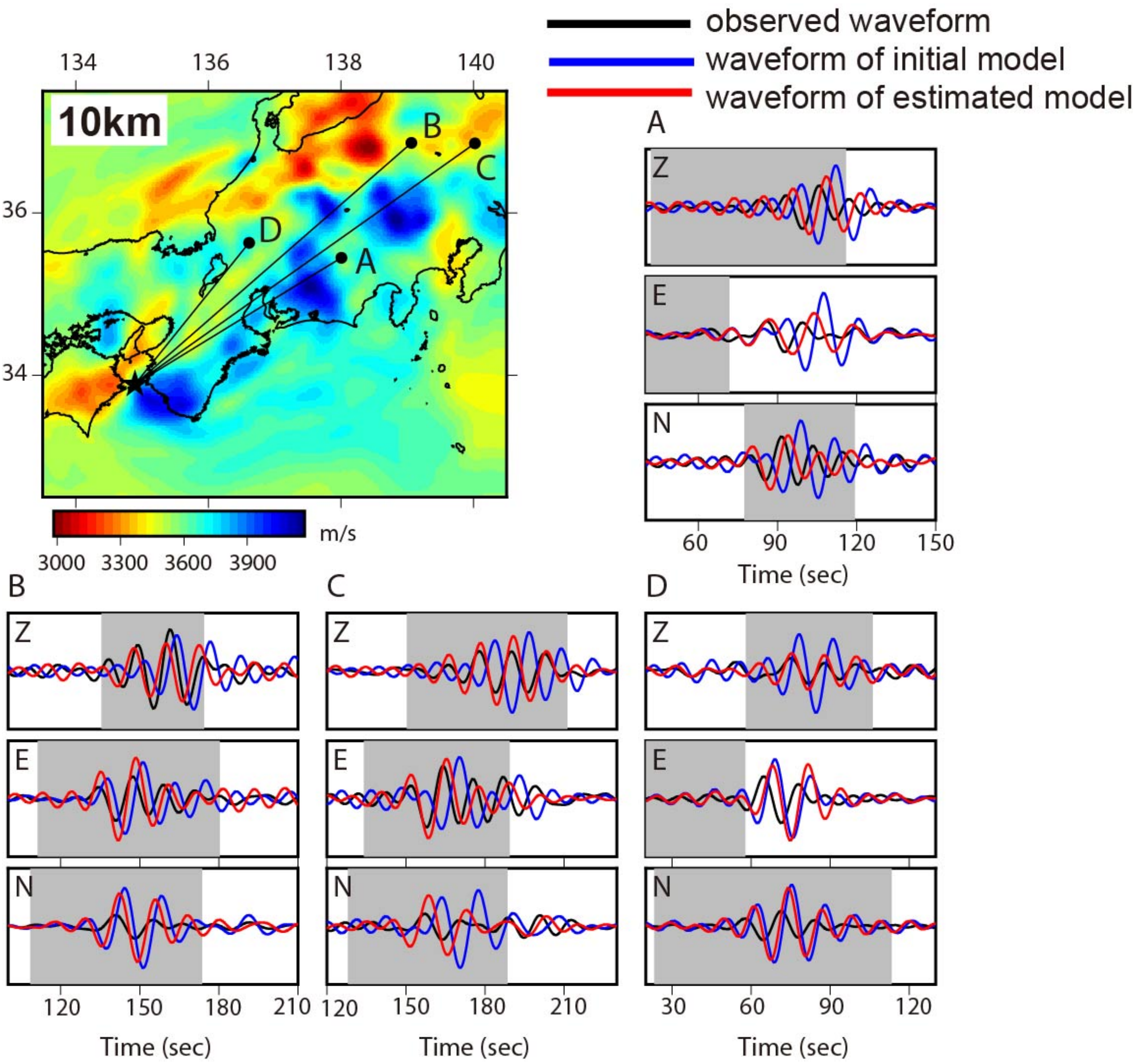


Fig. 3.5 Velocity model improvements to the waveform fits. The upper-left subfigure shows the final S-wave velocity model at 10 km depth, with the black star and circles marking a selected earthquake epicentre and four selected seismometers, respectively. Panels A–D show the observed (black lines), initial model (blue lines), and final model (red lines) waveforms for the corresponding seismometers in the upper-left subfigure. The vertical (Z), eastward (E), and northward (N) components are shown in each panel. Grey shaded areas denote the time windows that were selected by FLEXWIN.

Both the phases and amplitudes of the synthetic waveforms were improved relative to the observed waveforms after the inversion (17 iterations in total), as shown by the representative examples in Fig. 3.5. However, there were some instances where the synthetic waveforms were not altered during the inversion, as seen in some of the horizontal components in panel D of Fig. 3.5. This may result from either inappropriate earthquake source parameters or insufficient model parametrisation (elastic isotropic parametrisation was employed in our study).

The largest feature at shallow depths (≤ 10 km) in the S-wave velocity model is a low-velocity region in the northern part of the study area, where active volcanoes exist, whereas high velocities dominate in the southern part of the study area (Fig.3.3). This general finding is in agreement with previous studies (Nakajima & Hasegawa 2007a; Nishida *et al.* 2008). The low-velocity anomaly around region A corresponds to the Niigata sedimentary basin, which formed during the opening of the Japan Sea (Takano 2002), and active volcanoes such as Asama and Kusatsu-Shirane. Therefore, we attribute this low-velocity anomaly to the presence of both the sedimentary basin and magmatic fluids that are associated with back-arc volcanism. The high-velocity anomalies at these shallow depths are aligned southwest–northeast (B1–B3 in Figure 3), with this trend being in partial agreement with a previous velocity model that was constructed via ambient noise tomography (Nishida *et al.* 2008). The high-velocity anomaly beneath the southern end of the Kii Peninsula (B1 in Fig.3.3) could be related to the igneous complex that has been proposed in previous studies (Tsuji *et al.* 2015; Arnulf *et al.* 2022; Kimura *et al.* 2022). This high-density (and high-velocity) structure could be related to the slip behaviour in this region (Tsuji *et al.* 2017), as well as the fault segment boundary of interplate seismicity (Kimura *et al.* 2022). Figures 3.6(a) and (b) show the S-wave velocity distribution at 10 km depth, with the epicentres of the shallow (0–25 km depth; Yano et al., 2017) and deep (25–50 km depth; Yano et al., 2017) earthquakes overlain, respectively. The high-velocity regions at these shallow depths ($\leq$10 km) seem to overlie the areas of intense deep seismicity (Fig. 3.6b).

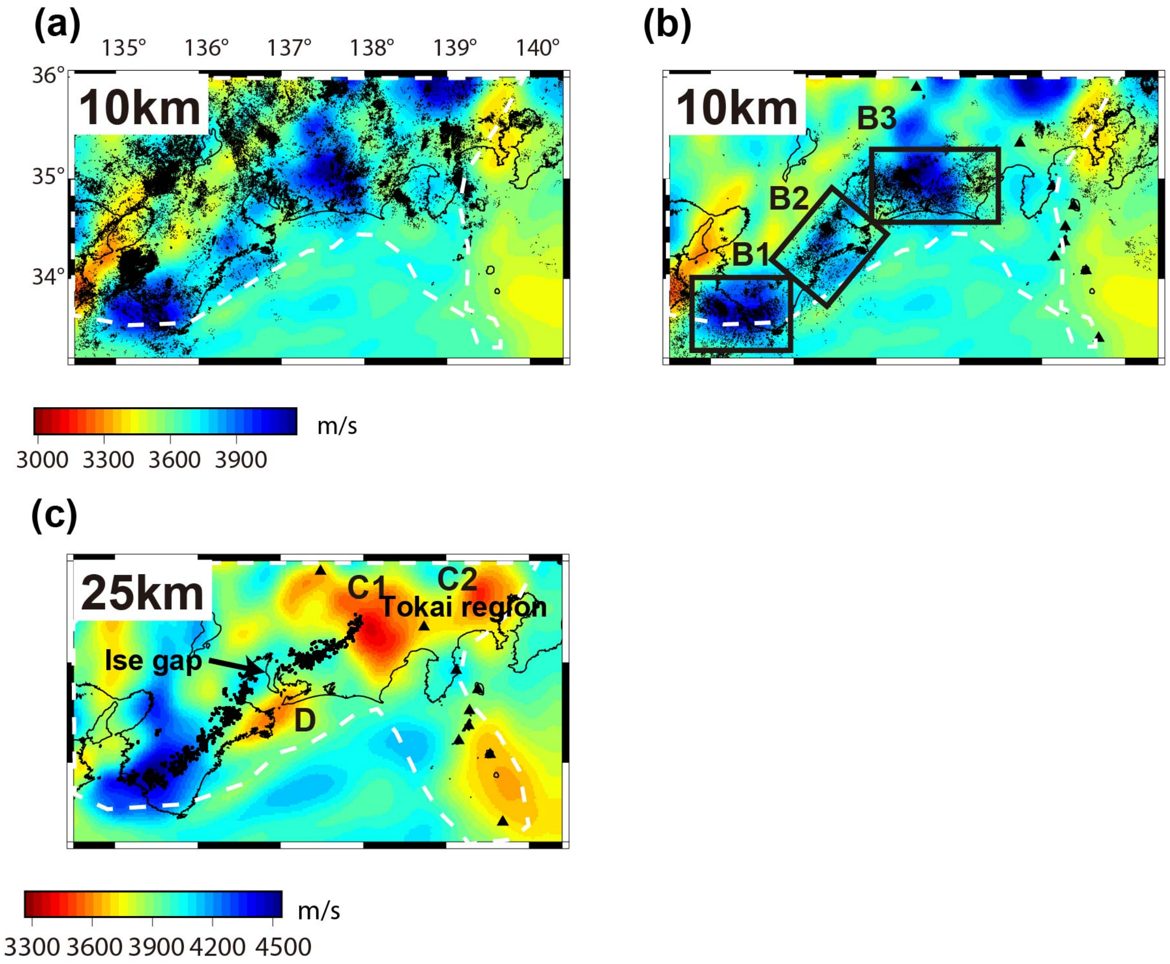


Fig. 3.6 Comparison of the S-wave velocity model and earthquake distributions. (a) S-wave velocity model at 10 km depth, with the epicentral distribution of <25 km depth earthquakes (Yano *et al.* 2017) overlain. (b) S-wave velocity model at 10 km depth, with the epicentral distribution of 25—50 km depth earthquakes (Yano *et al.* 2017) overlain. (c) S-wave velocity model at 25 km depth, with the epicentral distribution of low-frequency earthquakes (Kato & Nakagawa 2020) overlain. The dashed white lines are same as the dashed black lines in Figure 3. Black triangles denote active volcanoes.

The velocity model possesses four distinct low-velocity anomalies in the south at >25 km depth; these anomalies are labelled C1, C2, D, and E in Fig. 3.3. The boundary between the Philippine Sea and Eurasia plates is at ~30 km depth around anomalies C1, C2, and D (Hirose *et al.* 2008; Nakajima *et al.* 2009); these three anomalies may therefore be associated with the release of water via dehydration reactions in the subducting oceanic crust. The region between C1 and C2, which does not possess a significant low-velocity anomaly, corresponds to the Izu–Bonin collision zone, where Seno & Yamasaki (2003) have proposed that dehydration reactions are either absent or reduced in the subducting slab. Furthermore, Arai *et al.* (2014) have proposed that the thick crust of the Izu–Bonin arc on the oceanic plate inhibits the infiltration of seawater into the oceanic crust via an analysis of active-source seismic reflection data. The locations of these proposed dehydration reactions (and lack thereof) are consistent with the observed anomalies in our velocity model. Anomaly E lies beneath the chain of active volcanoes along the Izu–Bonin arc and may, therefore, be related to the magmatic activity associated with the arc. Figure 3.6(c) shows the S-wave velocity distribution at 25 km depth overlain by the distribution of low-frequency earthquakes (LFEs) estimated by Kato & Nakagawa (2020). Nakajima & Hasegawa (2016) have reported that LFE activity is significant above the subducted plate, where the P-wave velocity is low; however, there is low LFE activity beneath both Ise Bay and the Tokai region, where the P-wave velocity is not low. Therefore, the presence of low S-wave velocity anomalies in regions D, C1, and C2 is consistent with previous studies.

3.5 Discussion

3.5.1 Resolution analysis

We conducted a resolution analysis using point spread functions (PSFs) to assess the resolution of our tomographic inversions. PSFs are point-localised perturbations to a given background model, and the resolution analysis is based on the resulting blur and smear perturbations that is generated in a tomographic inversion (Fichtner & Trampert 2011). Here, we computed Hessians for the model perturbations using the following approximation:

$$\mathbf{H}(\mathbf{m})\delta\mathbf{m} \approx \mathbf{g}(\mathbf{m}+\delta\mathbf{m}), \tag{3.5}$$

where $\mathbf{H}(\mathbf{m})$ is the Hessian of the final model $\mathbf{m}$, $\delta\mathbf{m}$ is the model perturbation, and $\mathbf{g}(\mathbf{m}+\delta\mathbf{m})$ is the gradient of the perturbed model. We computed Hessian responses $\mathbf{H}\delta\mathbf{m}$ for the 20–50-s period range and employed smoothing using a 3D Gaussian with the same deviations in our velocity inversion. Figure 3.7 shows the results of individual PSF tests for velocity anomalies B1-B3, C1, C2, and D. We inserted perturbations consisting of 3D Gaussian functions with horizontal and vertical standard deviations of 15 and 6 km, respectively. The perturbations at 7 km depth, which corresponded to high-velocity anomalies B1—B3, possessed –15% peaks, and the perturbations at 25 km depth, which corresponded to low-velocity anomalies C1, C2, and D, possessed +15% peaks. The PSF test results for B2 and B3 (Figs 3.7b and c, respectively) show that the $H\delta m$ peaks coincide closely with the locations of the perturbations and therefore support the existence of high-velocity anomalies at these locations. However, the PSF test for B1 resulted in a peak with strong vertical smearing (Fig. 3.7a), which suggests that the termination of the high-velocity zone for B1, which extends to 35 km depth in our velocity model, is probably shallower. The PSF test results for C1, C2, and D demonstrate that the velocity perturbations were accurately mapped to $\mathbf{H}\delta\mathbf{m}$, with only slight shifts (Figs 3.7d–f). We conclude that low-velocity anomalies C1, C2, and D accurately reflect the presence of dehydration fluids around 25 km depth.

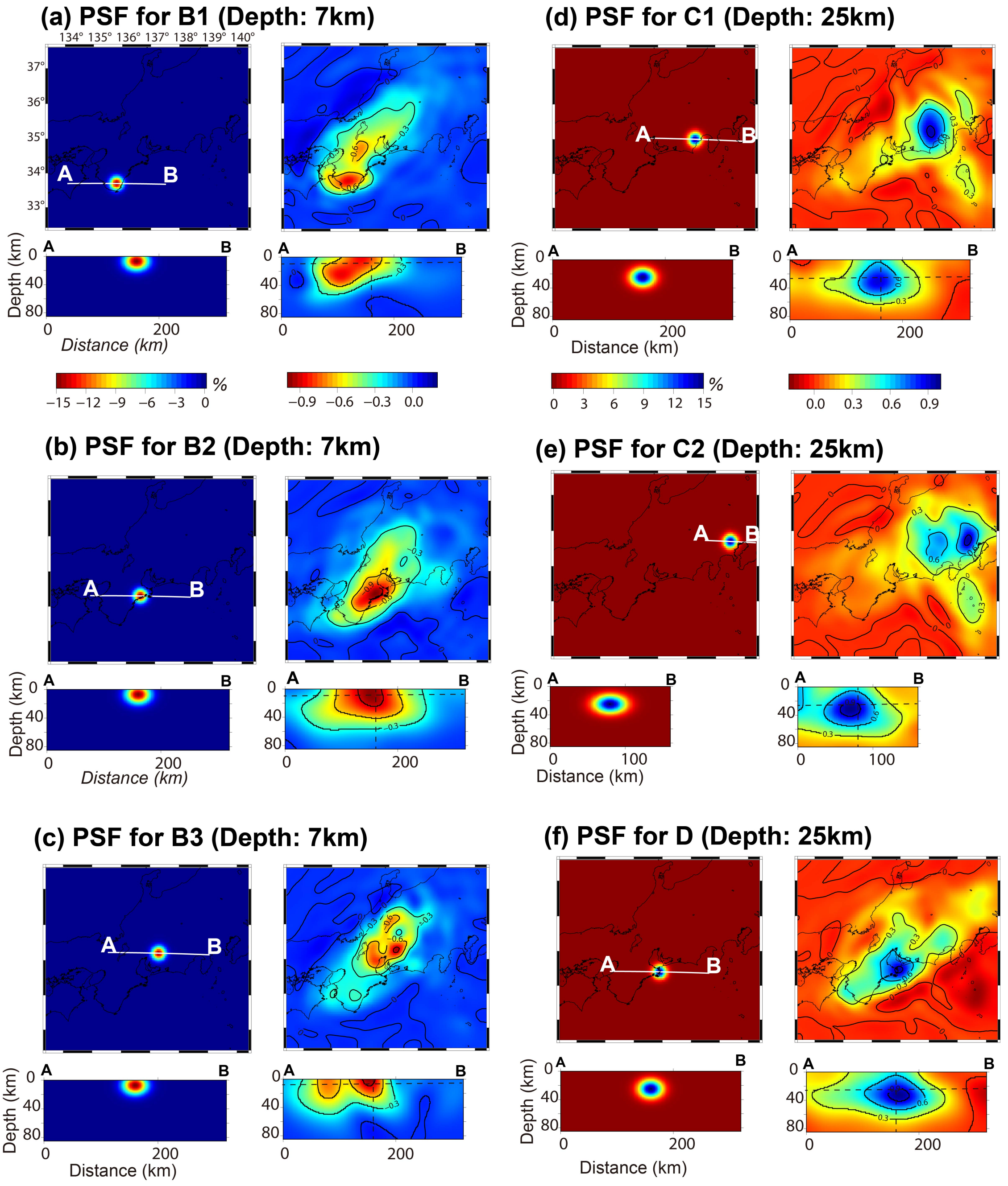


Fig. 3.7 PSF test results using 3D Gaussian functions. The upper and lower two figures in each panel are the horizontal model perturbations ($\delta \mathbf{m}$) and Hessian responses ($\mathbf{H}\delta \mathbf{m}$) for each depth slice and the corresponding vertical cross sections along profile AB, respectively. The intersection point of the dashed lines on the Hessian response cross sections indicates the location of the center of the corresponding perturbation. The perturbations from the final model are shown as percentages and the Hessian responses are normalised.

We investigated the sensitivity of our inversion specifications, such as the limited period ranges and source–station pairs, to the target model volume we chose for this study by computing $\mathbf{H}\delta\mathbf{m}$ using a constant volumetric perturbation of 50 m s^{-1} instead of the PSF test perturbations. The Hessian for this perturbation corresponds to the zero-wavenumber component of the Hessian in the frequency–wavenumber domain (Fichtner & Trampert 2011). Figure 3.8 shows the normalised response to this perturbation at our two period ranges. As expected, $\mathbf{H}\delta\mathbf{m}$ has large values at greater depths in the longer-period range than in the shorter-period range. Its value peaks at less than 40 km depth, with very low sensitivity at greater than 60 km depth; the 90 km depth of our model volume is therefore sufficient to compute the seismic waves included in our dataset. We set a threshold $\mathbf{H}\delta\mathbf{m}$ value of 0.2 to exclude the regions with insufficient sensitivity (dashed lines in Figs 3.3, 3.6, and 3.8).

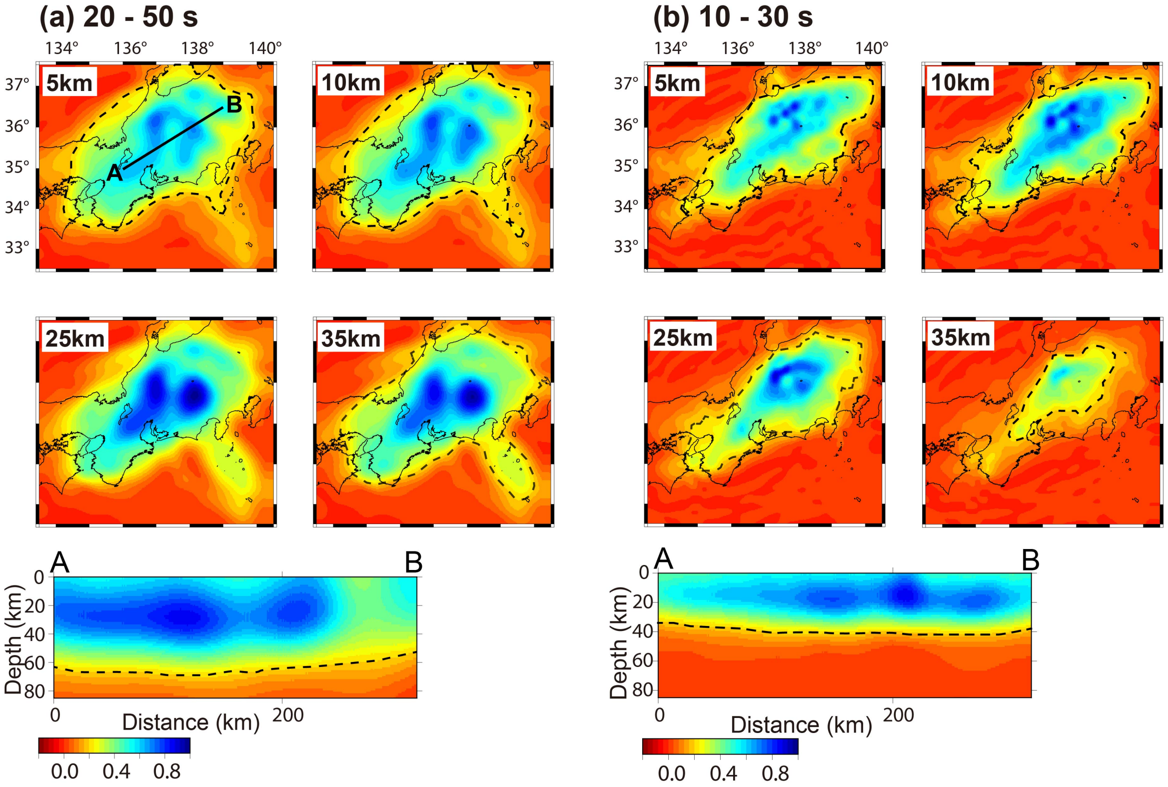


Fig. 3.8 PSF test results using a constant volumetric velocity perturbation. (a) Depth slices of the normalised $\mathbf{H}\delta\mathbf{m}$ for the 20–50-s period range. The vertical cross section for profile AB is shown in the bottom subfigure. The regions enclosed by the dashed black lines denote where $\mathbf{H}\delta\mathbf{m} > 0.2$. (b) The same as (a), but for the 10–30-s period range.

3.5.2 Model evaluation

The evaluation dataset (Fig. 3.2c), which was not included in the inversion, offers additional information about the inversion results. If the evaluation dataset causes only minimal to no reductions to the misfit, then the inversion solutions may be overestimated. We employed the normalised waveform difference misfit used in previous studies (e.g. Tape et al. 2010; Simutė et al. 2016):

$$\mathrm{E}^2(\mathrm{m}) = \frac{\int [\mathrm{d(t)} - \mathrm{u}(\mathrm{t};\mathbf{m})]^2 \mathrm{dt}}{\sqrt{\int \mathrm{d(t)}^2 \mathrm{dt} \int \mathrm{u}(\mathrm{t};\mathbf{m})^2 \mathrm{dt}}} \tag{3.6}$$

where $\mathrm{d(t)}$ and $\mathrm{u}(\mathrm{t};\mathbf{m})$ are the observed and synthetic data, respectively. We also calculated the travel-time misfits based on cross-correlations of the observed and synthetic waveforms. We computed the normalised waveform differences and travel-time misfits of both the inversion and evaluation datasets. These calculations were performed using the entire period range (10–50 s), and the dataset was selected by FLEXWIN using the observed data and synthetic waveforms from the initial model, with a maximum time lag of 12 s. We used the selected time windows for the travel-time misfit calculations, but not for the normalised waveform difference misfit calculations, to ensure that the entire seismograms were assessed. Histograms of the resultant normalised waveform difference and cross-correlation travel-time misfits of the initial and final models for the evaluation and inversion datasets (Fig. 3.9) suggest that our inversion has estimated reliable velocity structures.

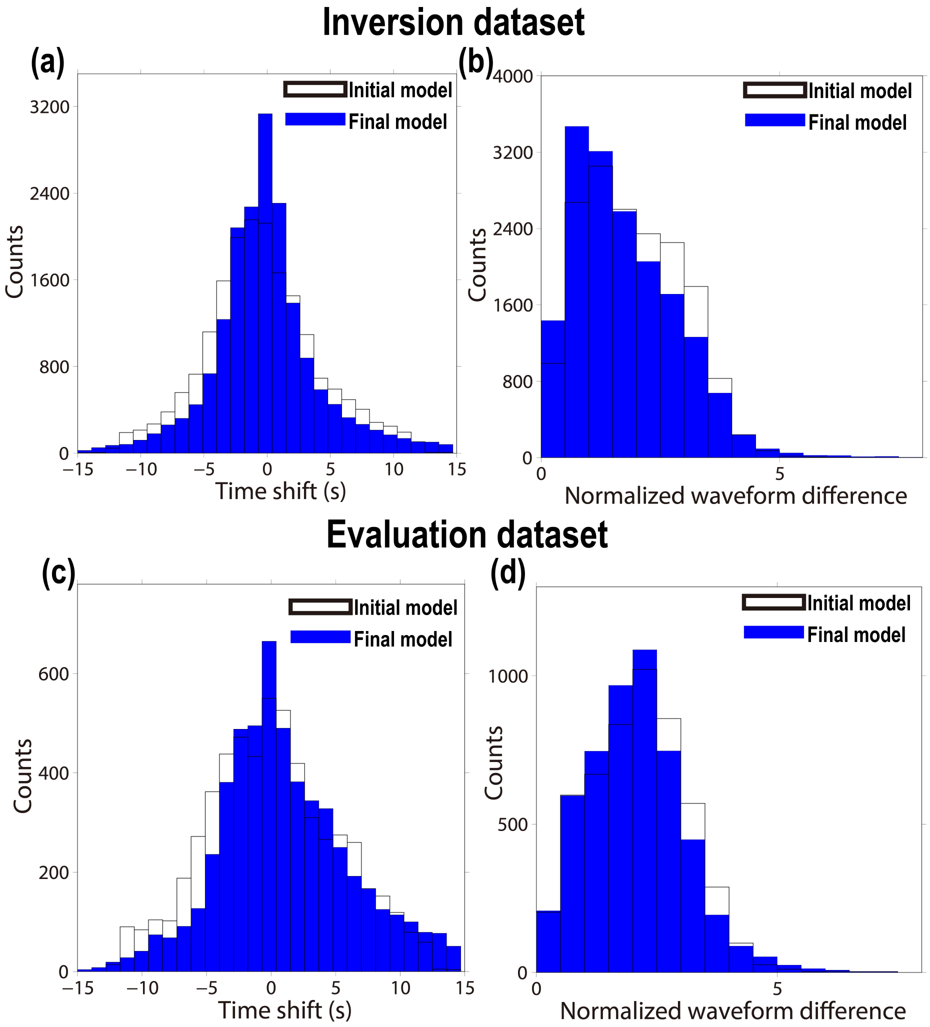


Fig. 3.9 Evaluations using the inversion (a,b) and evaluation (c,d) datasets based on the cross-correlation time shift (a,c) and normalised waveform difference (b,d). The data are filtered over the 10—50-s period range.

We assumed that the initial source parameters were accurate, thereby keeping the source parameters fixed during the inversion. However, both the model and source parameters contribute to waveform misfits. Numerical tests by Valentine & Woodhouse (2010) have suggested that the mislocation of the sources influences the recovered model parameters, such that a better result can be obtained by simultaneously inverting for both the source and model parameters. Some studies have inverted for the source parameters based on 3D wave propagation in the initial model before the structural inversion to reduce the uncertainties in the initial source parameters (Bozdağ *et al.* 2016; Lei *et al.* 2020). Source parameters can be also inverted at some iterations during the inversion of the structural parameters to take into account the changes in the structural parameters. Other studies minimised the uncertainties in the source parameters using the double-difference travel-time misfits (Yuan *et al.* 2016; Örsvuran *et al.* 2020).

Here, we recovered the S-wave velocity model assuming isotropic parametrisation. However, the assumption of isotropy in seismic tomography generally produces distortions in the recovered model because crustal rocks are commonly anisotropic. Even isotropic elastic heterogeneities can generate anisotropic effective medium for longer seismic waves (Backus 1962). Furthermore, Kashiwagi & Nakajima (2019) estimated the 3D P-wave attenuation structure of our study area and observed a high-attenuation zone. This can be also inferred from the low-velocity area in our model. Therefore, both anelastic dispersion and anisotropy may be contributing to the phase delays in our dataset. We further note that our presented S-wave velocity model can be improved by incorporating source parameters, anisotropy, and attenuation into the inversion.

3.5.3 Comparison with other velocity model

We compared our S-wave velocity model to that of Simutė *et al.* (2016) for the Japan Islands region at 20 km depth (Fig. 3.10). The Simutė *et al.* (2016) velocity model (hereafter CSEM_japan), which is part of the Collaborative Seismic Earth Model (CSEM) (Fichtner *et al.* 2018), consists of the vertical (SV) and horizontal (SH) S-wave velocity structures in the crust and upper mantle of the Japan Islands region. Simutė *et al.* (2016) resolved these SV and SH velocity structures via full-waveform tomography, which was based on the time–frequency phase misfits. Their optimisation scheme was based on conjugate-gradient optimisation, whereas our optimisation scheme is based on the steepest-descent method. They performed the inversion using earthquake waveforms that spanned the 20–80-s period range, but currently the updated model with a period range of 15–80-s is available. Their dataset contains waveforms from seismic networks in China, Taiwan, South Korea, and F-net in Japan, and their inverted region is much larger than ours. Although there are some differences in the parameters used in the CSEM_Japan model and our model (e.g., a shorter period range and a smaller number of elastic parameters in our model), similar features are observed in both model results. We calculated the isotropic S-wave velocity v_s for this comparison from the SV v_{sv} and SH v_{sh} velocities using the following equation: $v_s = \sqrt{\frac{2}{3}v_{sv}^2 + \frac{1}{3}v_{sh}^2}$ (e.g., Simutė et al., 2016). Our model possesses low-velocity regions in the northern and southern parts of the target volume, with the northern regions extending farther east–west than the southern region (Fig. 3.10a); this feature is also evident in the CSEM_Japan model (Fig. 3.10b). Our model contains a narrow band of high-velocity anomalies that separates the two low-velocity anomalies along the solid black line in Fig. 3.10(a). CSEM_Japan also possesses a relatively higher velocity around this line, although it does not split the low-velocity anomaly into separate parts. CSEM_Japan used a longer period range and data with longer source–receiver distances compared with our inversion, such that their data are likely dominated by seismic waves with longer periods than ours and are mainly sensitive to the greater than 20 km depth regions. This key difference in the inverted waveform datasets is probably the main reason for the observed differences between the CSEM_Japan model and our model.

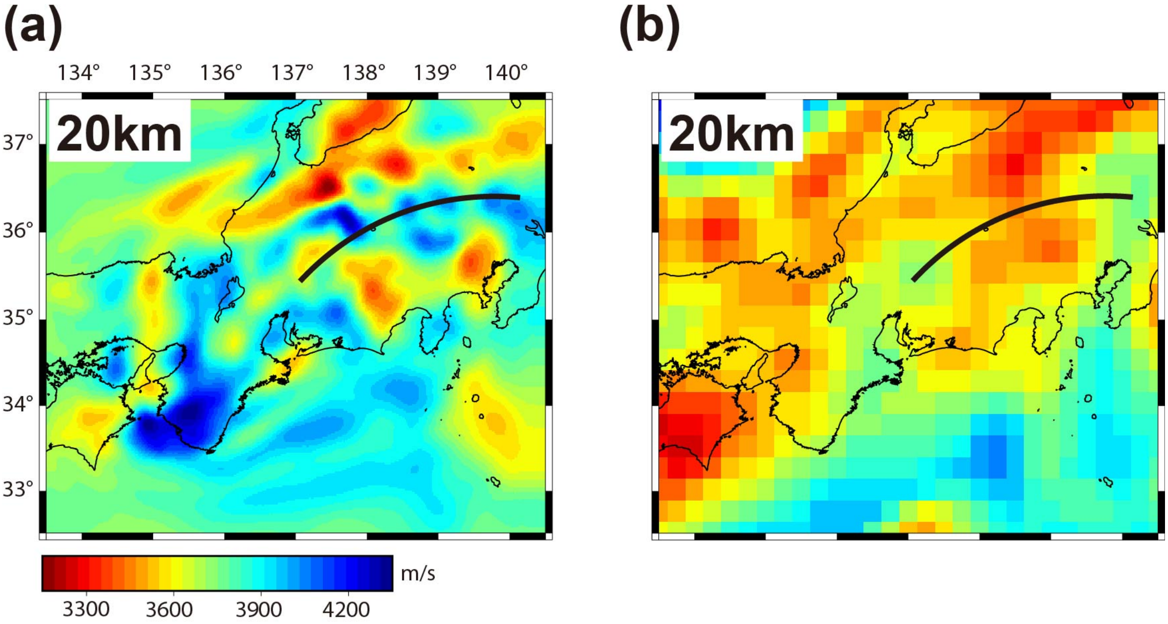


Fig. 3.10 Comparison of the obtained S-wave velocity structure in this study and the CSEM for the Japan Islands region (Simutė et al. 2016) at 20 km depth. Note that a relatively high velocity can be seen around solid black line in both velocity models. (a) S-wave velocity structure from this study. (b) CSEM for the Japan Islands region (Simutė et al. 2016).

3.6 Conclusions

We constructed a 3D S-wave velocity model of the crust beneath central Honshu Island, Japan, via full-waveform tomography of the seismic waveforms of 29 earthquakes recorded by 388 Hi-net and 28 F-net seismic stations. Our inversion was designed to minimise the time–frequency phase misfit between the observed and synthetic seismic waveforms in the 50–20-s and 30–10-s period ranges. The final model resolved strong horizontal heterogeneities, with lateral velocity variations of up to 1200 m s^{-1} (2800–4000 m s^{-1} range) imaged. The low-velocity anomalies we resolved appear to correspond to a thick sedimentary basin, volcanic fluids, and fluid migration owing to dehydration of the subducting plate.

We also imaged three high-velocity regions at shallow depths ($\leq$ 10 km) that overlie areas of intense seismicity. These high-velocity anomalies also agree with the igneous complex that has been inferred to influence the regional slip behaviour. PSF tests confirmed that our dataset possesses the necessary sensitivity to reliably delineate these low- and high-velocity anomalies. The similar performance of our model with an evaluation dataset supports the accuracy of the inversion results.

This study confirms that full-waveform tomography and the mixed waveform dataset from F-net and Hi-net stations can resolve an S-wave velocity structure that is consistent with the known geology of Japan. The seismic structures in our velocity model are comparable to those obtained in previous tomographic studies. Unlike previous S-wave velocity models of Japan, which have been largely obtained via first-arrival tomographic inversions, our model is based on the isotropic elastic seismic-wave equation and is therefore more suitable for modelling waveforms in the 10–50-s period range. The combination of full-waveform tomography, F-net data, and Hi-net data will lead to more accurate seismic velocity models throughout the Japan Islands region, although it has not yet been confirmed whether the earthquake data recorded by the F-net and Hi-net stations have the necessary resolution to resolve the fine-scale velocity structures in other geologically complex regions, such as Kyushu and Hokkaido.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science (JSPS) through a Grant-in-Aid for JSPS Research Fellow (JP21J21871) and KAKENHI grants (JP20H01997 and JP21H05202). We used the open-source software SPECFEM3D Cartesian for seismic wave simulation, and we thank the developers of SPECFEM3D. We also thank the National Research Institute for Earth Science and Disaster Prevention for providing us with Hi-net and F-net data. Seismic wave simulations were conducted on the ITO supercomputer system at the Research Institute for Information Technology, Kyushu University. We used the open-source software GMT (Wessel et al. 2013) to construct the figures.

Chapter 4

Full waveform inversion based on wavefield gradients

4.1 Introduction

Recent advances in seismic instrumentation have enabled the use of wavefield gradient measurements to more comprehensively characterise ground motions and further constrain various seismic phenomena, including earthquake mechanisms, the directionality of seismic waves and anisotropic structures. However, it is important to carefully examine the feasibility of directly applying seismic methods that have been used for translational measurements to wavefield gradient measurements.

The ability to acquire seismic wavefield gradient measurements has existed for decades. Tiltmeters, which measure the tilt of the ground surface (spatial derivatives of displacement), have traditionally been used to investigate tidal tilt (Nishimura 1950; Melchior 1966). Seismic observations have also been made using tilt meters (Sassa & Nishimura 1951); however, measurements of seismic rotational motion have been a long-standing challenge due to their much smaller amplitudes compared with the translational motions that are excited by earthquakes. Recent advances in sensor technologies, such as ring laser gyroscopes (McLeod *et al.* 1998), have enabled rotational motion measurements, and subsequent studies have shown the benefits of using rotational sensors in various geological settings. For example, collocated measurements of rotation and translational motions have been analysed to retrieve the local phase velocity (Mikumo & Aki 1964; Igel *et al.* 2005; Fichtner & Igel 2009). Recently, the theory to measure the dispersion relation using both translational and wavefield gradient measurements at single observation point in weakly anisotropic media has been introduced (Tang *et al.* 2023). Detailed strain measurements from distributed acoustic sensing (DAS) arrays have led to a significant increase in the number of geophysical studies that implement dynamic subsurface strains (Willis *et al.* 2016; Zhan 2020). These recent advances in sensor technology and the increased use of rotational sensors and DAS arrays have enabled a transition towards actively observing and utilising measurements of seismic wavefield gradients.

Many seismic analyses based on rotational sensor and DAS observations have been conducted. Phase-velocity and polarisation analyses using rotational sensors (Wassermann *et al.* 2016; Keil *et al.* 2021), and structural imaging (Parker *et al.* 2018), near-surface velocity monitoring (Dou *et al.* 2017) and earthquake studies (Ajo-Franklin *et al.* 2019) using DAS cables have made important advances in seismology. These studies have demonstrated the potential of using wavefield gradient measurements to address fundamental seismological questions; however, it is important to pay attention to the distinct nature of the wavefield gradients compared with translational motions when analysing wavefield gradient measurements.

Wavefield gradients can be significantly influenced by local small-scale heterogeneities in the subsurface that distort both the phase and amplitudes of the wavefield gradients (King & Bilham 1973; Harrison 1976; van Driel *et al.* 2012; Singh *et al.* 2020; Muir & Zhan 2022). This issue has been acknowledged as a critical problem in tilt measurements (Melchior 1966; Lennon & Baker 1973; Goulty 1976). A previous study reported sudden changes in tilt values due to local subsurface heterogeneities at the sensor, such as cavities, discontinuities and elastic contrasts, and also the local topography, and noted discrepancies between the theoretical and observed tilt values (Meertens *et al.* 1989). Similar effects have been observed in rotational sensor and DAS measurements (Singh *et al.* 2020; Muir & Zhan 2022). Singh *et al.* (2020) obtained a strong fit between the observed and synthetic displacement waveforms using a smooth Earth model whereas the observed rotational waveforms from the same locations did not fit the synthetic waveforms well; they attributed these findings to the effects of small-scale heterogeneities on measurements of the wavefield gradients. Two-scale homogenisation theory, which has recently been introduced in seismology, can mathematically explain the effects of small-scale heterogeneities on wavefield gradient measurements (Capdeville *et al.* 2010a; Capdeville *et al.* 2010b; Cupillard & Capdeville 2018). This theory shows that coupling of the effective strain to gradient measurements can result in the effects of small-scale heterogeneities. Singh *et al.* (2020) proposed a method for inverting the coupling-term coefficients to address the discrepancies between the observed and simulated rotational data. These coefficients can then be used to correct the simulated data and explain the observed wavefield gradients (Singh *et al.* 2020; Muir & Zhan 2022). Therefore, both the coupling-term coefficients (hereafter "correctors") and a smooth tomographic model are required to accurately simulate wavefield gradient observations, as the ability to obtain high-resolution tomographic models with small-scale structures is generally challenging.

Methods have been developed to account for the effects of small-scale heterogeneities when forward modelling wavefield gradients. However, few studies have investigated the effects of small-scale heterogeneities on wavefield gradients in seismic inverse problems, such as full waveform inversion (FWI). Although it may be straightforward to assume that a FWI using wavefield gradients can recover the small-scale structures around the receivers, thereby generating higher-resolution images than those obtained using translational motions, the impacts of small-scale heterogeneities on wavefield gradient-based FWI have not been thoroughly investigated. This study aims to investigate the effects of wavefield gradient measurements on FWI results and address the impact of small-scale heterogeneities. Our findings demonstrate that small-scale heterogeneities can strongly affect the FWI inversion results. Given the sensitivity of FWI frameworks to these small-scale effects, we propose a new FWI framework that can account for small-scale heterogeneities and improve the accuracy of FWI results.

4.2 Two-scale homogenisation

Two-scale homogenisation is a mathematical framework that can deal with multiscale problems. It was originally developed for periodic and stochastic elastostatic and elastodynamic problems (Sanchez-Palencia 1980; Bensoussan *et al.* 1978) before being extended to deterministic multiscale heterogeneous media with no scale separation for wave propagation (Capdeville *et al.* 2010a; Capdeville *et al.* 2010b; Cupillard & Capdeville 2018). In seismology, the two-scale homogenization was first introduced in 1D heterogeneous media by Capdeville *et al.* (2010a), then extended to 2D (Capdeville *et al.* 2010b; Guillot *et al.* 2010) and 3D (Cupillard & Capdeville 2018) cases. The homogenization can be used to estimate the homogenized (effective or equivalent) model while keeping the wave propagation in original model. Because the homogenized model can be sampled by corser mesh than one which needs for original model, computational costs can be significantly reduced. Furthermore, we can investigate the effects of small-scale structures on the seismic source inversion using homogenization theory (Burgos *et al.* 2016). Recently, the concept of homogenization for heterogeneous media was extended to seismic sources (Capdeville 2021), and (Renat *et al.* 2022) reveals that the image of seismic source by time-reversal method is equivalent to homogenized point source.

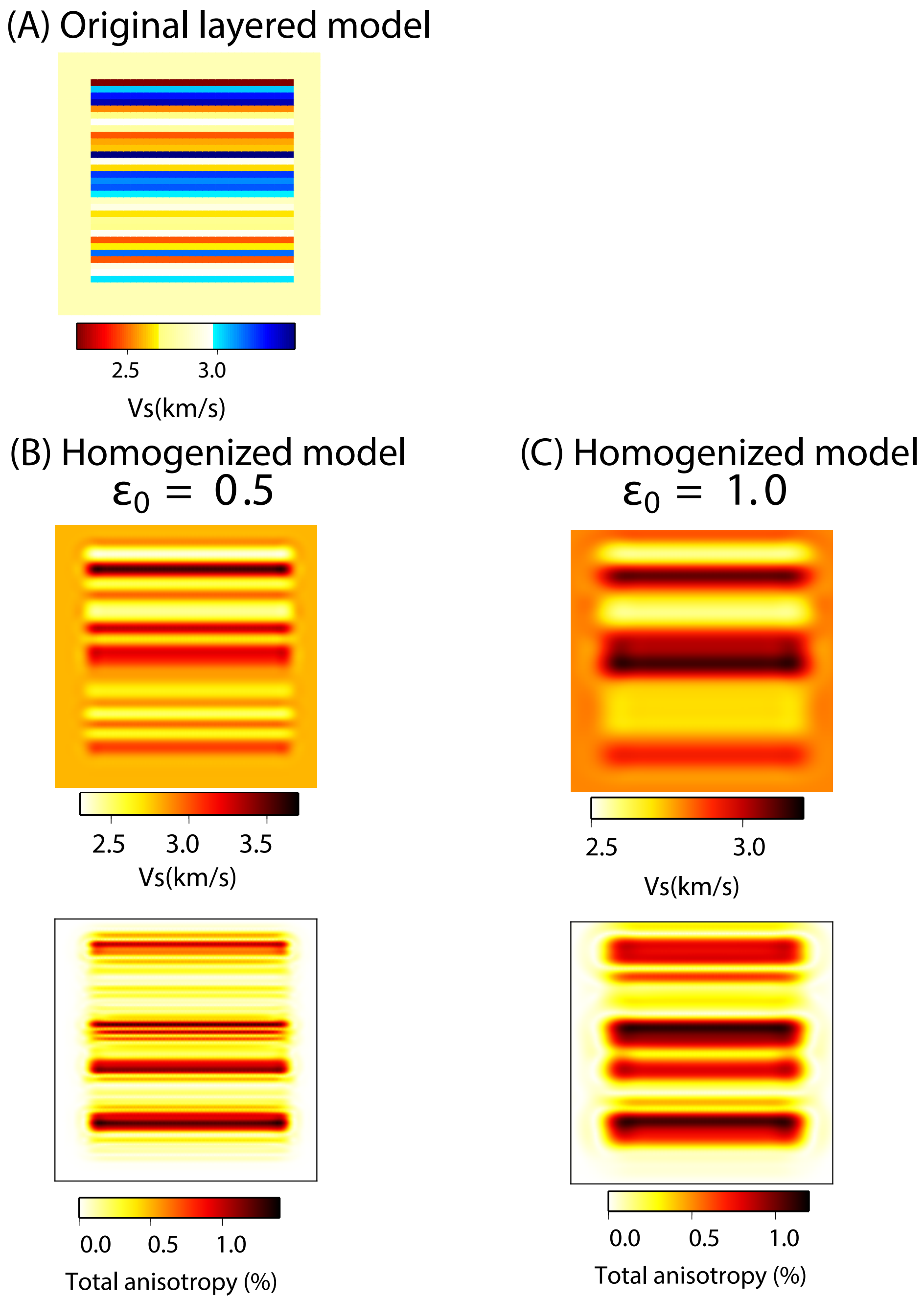
4.2.1 Scale separation

Homogenisation theory considers scales of heterogeneities as either small-scale (microscopic scale) or large-scale (macroscopic scale) features. Small-scale heterogeneities are meant to be homogenised and replaced by effective properties. For periodic heterogeneities, the small scale is the periodic structure, and the large scale is the constant background. In seismology, there is no periodicity in the media and we rely on an arbitrary scale λ_0 to separate small for large scales. Homogenisation is an asymptotic method based on a small parameter ε_0, which is defined as follows:

$$\varepsilon_0 = \frac{\lambda_0}{\lambda_{min}}. \tag{4.1}$$

ε_0 measures the position of the separation between the small and large scales with respect to the minimum wavelength λ_{min}. Note that any value can be chosen for ε_0. However, $\varepsilon_0 = 0.5$ (meaning that heterogeneities half the size of the minimum wavelength and smaller are considered small-scale heterogeneities) is a good approximation that has generally been employed for most cases. Fig.4.1 shows the example of homogenized model with different scale separation: $\varepsilon_0 = 0.5$ and $\varepsilon_0 = 1.0$. When ε_0 is smaller, the detailed structures of original model is preserved which leads to higher accuracy of wave propagation. In this example, the original model is simple layered model, therefore, homogenization technique developed by Backus (1962) also can be used.

The two-scale homogenisation method relies on two space variables: the regular space variable $\mathbf{x}$ (also called the large- or macroscopic-scale variable) and a new space variable $\mathbf{y} = \frac{\mathbf{x}}{\varepsilon_0}$ (also called the small- or microscopic-scale variable). A new mathematical problem is derived, whereby all of the physical quantities depend on both the macro-scale spatial variable $\mathbf{x}$ and the micro-scale spatial variable $\mathbf{y}$. The details of the derivation of this homogenisation problem can be found in Capdeville *et al.* (2020).



(A) Original layered model

(B) Homogenized model $\varepsilon_0 = 0.5$

(C) Homogenized model $\varepsilon_0 = 1.0$

Fig. 4.1 (A) V_S of original layered model. (B) V_S and total anisotropy of homogenized model with $\varepsilon_0 = 0.5$. (C) V_S and total anisotropy of homogenized model with $\varepsilon_0 = 1.0$.

We define the cell average $\langle\cdot\rangle$ for a periodic function $h(\mathbf{x},\mathbf{y})$ over the cell (grid interval) $\mathbf{Y}$:

$$\langle h\rangle(\mathbf{x}) = \frac{1}{\mathbf{Y}}\int_{\mathbf{Y}} h(\mathbf{x},\mathbf{y})d\mathbf{y}\,. \tag{4.2}$$

Therefore, after averaging over the cell, the function $h(\mathbf{x}, \mathbf{y})$ is independent on small-scale variable.

Further, we write the solution of the wave equation (equations (2.1)–(2.2)) in the power-series of ε_0:

$$\mathbf{u}^{\varepsilon_0}(\mathbf{x}, \mathbf{y}, t) = \mathbf{u}^0(\mathbf{x}, \mathbf{y}, t) + \varepsilon_0 \mathbf{u}^1(\mathbf{x}, \mathbf{y}, t) + \varepsilon_0^2 \mathbf{u}^2(\mathbf{x}, \mathbf{y}, t) + \ldots, \tag{4.3}$$

and the stress is:

$$\boldsymbol{\sigma}^{\varepsilon_0}(\mathbf{x}, \mathbf{y}, t) = \varepsilon_0^{-1} \boldsymbol{\sigma}^{-1}(\mathbf{x}, \mathbf{y}, t) + \boldsymbol{\sigma}^0(\mathbf{x}, \mathbf{y}, t) + \varepsilon_0 \boldsymbol{\sigma}^1(\mathbf{x}, \mathbf{y}, t) + \varepsilon_0^2 \boldsymbol{\sigma}^2(\mathbf{x}, \mathbf{y}, t) + \ldots. \tag{4.4}$$

Note that the exponent i is the power for ε_0 and an index i is for $\mathbf{u}^i$ and $\boldsymbol{\sigma}^i$. All coefficients are periodic in small-scale variable $\mathbf{y}$. Using small- and large-scale variables, we assume that we are able to define $\rho^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x})$ and $\mathbf{c}^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \mathbf{c}(\mathbf{x})$. Both quantities are also periodic in $\mathbf{y}$. Detailed definitions can be found in (Capdeville *et al.* 2010b). Inserting above two equations (4.3) and (4.4) into the wave equation (equations (2.1)–(2.2)), using $(\rho_0^{\varepsilon}, \mathbf{c}_0^{\varepsilon})$, and identifying the same term in ε_0^i, we obtain:

$$\begin{aligned} \rho^{\varepsilon_0} \partial_{tt} \mathbf{u}^i - \nabla_{\mathbf{x}} \cdot \boldsymbol{\sigma}^i - \nabla_{\mathbf{y}} \boldsymbol{\sigma}^{i+1} &= \boldsymbol{f} \delta_{i,0}\,, \\ \boldsymbol{\sigma}^i = \mathbf{c}^{\varepsilon_0} : (\boldsymbol{\epsilon}_{\mathbf{x}}(\mathbf{u}^i) &+ \boldsymbol{\epsilon}_{\mathbf{y}}(\mathbf{u}^{i+1})\,. \end{aligned} \tag{4.5}$$

For the above equation (4.5), we used the relation $\nabla_{\mathbf{x}} = \nabla_{\mathbf{x}} + \frac{1}{\varepsilon_0} \nabla_{\mathbf{y}}$.

4.2.2 Asymptotic solution of homogenization problem

The equations (4.5) should be solved for each index i. $i = -2$ for first equation and $i = -1$ for second equation in equations (4.5), we can show:

$$\mathbf{u}^0(\mathbf{x}, \mathbf{y}, t) = \langle \mathbf{u}^0 \rangle (\mathbf{x})\,. \tag{4.6}$$

This means that the first term of solution of wave equation $\mathbf{u}^0$ is independent on small-scale variable $\mathbf{y}$. Indeed, by leading-order approximation of solution of wave equation, the displacement is dependent on only large-scale variable $\mathbf{x}$. In this study, we use symbol $*$ to denote homogenized quantities which does not depend on small-scale. Therefore, we write $\mathbf{u}^0(\mathbf{x}, \mathbf{y}, t) = \mathbf{u}^*(\mathbf{x}, t)$.

$i = -1$ for first equation and $i = 0$ for second equation in equations (4.5), we have:

$$\nabla_{\mathbf{y}} \cdot \boldsymbol{\sigma}^0 = 0\,, \tag{4.7}$$

$$\boldsymbol{\sigma}^0 = \mathbf{c}^{\varepsilon_0} : (\boldsymbol{\epsilon}_{\mathbf{x}}(\mathbf{u}^*) + \boldsymbol{\epsilon}_{\mathbf{y}}(\mathbf{u}^1))\,. \tag{4.8}$$

Combining these two equations, the linear differential equation can be obtained:

$$\nabla_{\mathbf{y}} \cdot (\mathbf{c}^{\varepsilon_0} : \boldsymbol{\epsilon}(\mathbf{u}^*)) = -\nabla_{\mathbf{y}} \cdot (\mathbf{c} : \boldsymbol{\epsilon}_{\mathbf{y}}(\mathbf{u}^1))\,. \tag{4.9}$$

Supposing the solution of this differential equation as:

$$\mathbf{u}^1(\mathbf{x}, \mathbf{y}, t) = \boldsymbol{\chi}(\mathbf{x}, \mathbf{y}) : \boldsymbol{\epsilon}(\mathbf{u}^*)(\mathbf{x}) + \langle \mathbf{u}^1 \rangle(\mathbf{x})\,, \tag{4.10}$$

where $\boldsymbol{\chi}$ is the so-called first-order corrector, we obtain the cell problem:

$$\nabla_{\mathbf{y}} \cdot (\mathbf{c}^{\varepsilon_0} : \boldsymbol{\epsilon}(\mathbf{u}^*)) = -\nabla_{\mathbf{y}} \cdot (\mathbf{c} : \boldsymbol{\epsilon}(\boldsymbol{\chi}^{\varepsilon_0} : \boldsymbol{\epsilon}(\mathbf{u}^*))\,. \tag{4.11}$$

The cell problem (as defined in Equation (4.11)) can be solved by numerical methods like the finite-element method (Capdeville *et al.* 2020) and Fourier-based solvers (Capdeville *et al.* 2015).

Next, taking cell average of equation (4.8), we have:

$$\langle \boldsymbol{\sigma}^0 \rangle(\mathbf{x}) = \mathbf{c}^*(\mathbf{x}) : \boldsymbol{\epsilon}_{\mathbf{x}}(\mathbf{u}^*)\,, \tag{4.12}$$

where homogenized elastic tensor $\mathbf{c}^*$ is:

$$\mathbf{c}^*(\mathbf{x}) = \langle \mathbf{c}^{\varepsilon_0}(\mathbf{I} + \frac{1}{2}(\nabla_{\mathbf{y}}\boldsymbol{\chi} +^t \nabla_{\mathbf{y}}\boldsymbol{\chi})) \rangle\,. \tag{4.13}$$

The equation (4.5) for $i = 0$ gives:

$$\rho^{\varepsilon_0}\partial_{tt}\mathbf{u}^* - \nabla_{\mathbf{x}} \cdot \boldsymbol{\sigma}^0 - \nabla_{\mathbf{y}} \cdot \boldsymbol{\sigma}^1 = \boldsymbol{f}\,. \tag{4.14}$$

Taking cell average and using the fact that $\langle \partial_y h(\mathbf{x}, \mathbf{y}) \rangle = 0$, we obtain homogenized wave equation:

$$\rho^*\partial_{tt}\mathbf{u}^* - \nabla_{\mathbf{x}} \cdot (\mathbf{c}^* : \boldsymbol{\epsilon}_{\mathbf{x}}(\mathbf{u}^*)) = \boldsymbol{f}\,, \tag{4.15}$$

where $\rho^*(\mathbf{x}) = \langle \rho_0^{\varepsilon} \rangle$. Hence, $\mathbf{u}^*$ solely depends on the homogenized medium $(\rho^*, \mathbf{c}^*)$. Consequently, in computing the displacement wavefield, we can solve the wave equation using the homogenized model (e.g. (B) in Fig.4.1) instead of the original model (e.g. (A) in

Fig.4.1). We have assumed that $(\rho^{\varepsilon_0}(\mathbf{x},\mathbf{y}),\mathbf{c}^{\varepsilon_0}(\mathbf{x},\mathbf{y}))$ is periodic in $\mathbf{y}$. However, such set up of $\rho^{\varepsilon_0}(\mathbf{x},\mathbf{y})$ and $\mathbf{c}^{\varepsilon_0}(\mathbf{x},\mathbf{y})$ is not straightforward due to the elusive scale separation in general heterogeneous media. To overcome this challenge, Capdeville *et al.* (2010a) introduced a technique using a low-pass filter to separate scales. For a detailed explanation of this construction, readers are referred to Capdeville *et al.* (2020).

For the discussions of follwing section, we define the homogenization operator $\mathcal{H}$ to compute the homogenized quantities of $(\rho^*,\mathbf{c}^*,\chi)$:

$$(\rho^*,\mathbf{c}^*,\chi)=\mathcal{H}(\rho,\mathbf{c})\,. \tag{4.16}$$

$\mathcal{H}$ implies solving a cell problem as defined by equation (4.11). The solution only depends on ε_0 and the elastic model; neither the time nor the seismic source incluence $\mathcal{H}$. The presence of normal stress-free boundary conditions requires the use of effective boundary conditions (Capdeville & Marigo 2008; Capdeville *et al.* 2013) to homogenise the small-scale variations in both the local topography and structures just beneath the subsurface. However, we omit the effective boundary condition in this study for simplicity. $\mathcal{H}$ is a non-linear operator that often yields non-intuitive homogenisation results. For example, the effective properties are almost always anisotropic, even if the true model is isotropic (e.g Fig.4.1). Solving the cell problem and then finding the effective mechanical properties generally involve a numerical solver, with the layered-medium case being the only case that leads to an analytical solution. Homogenisation theory yields the well-known results of Backus (1962) for the layered-medium case. We do not expand further on this aspect because it is not required for our discussion; see Capdeville *et al.* (2020) for a complete discussion.

4.2.3 Wavefield gradients dependent on small-scale structures

With the fact that $\mathbf{u}^0(\mathbf{x},\mathbf{y},t)=\mathbf{u}^*(\mathbf{x})$, we can rewrite the equation (4.3) as:

$$\mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t)=\mathbf{u}^*(\mathbf{x},t)+\varepsilon_0\mathbf{u}^1(\mathbf{x},\mathbf{y},t)+\varepsilon_0^2\mathbf{u}^2(\mathbf{x},\mathbf{y},t)+\dots\,. \tag{4.17}$$

Again, this equation means that the displacement wavefield can be well approximated with $\mathbf{u}^*(\mathbf{x},t)$ assuming $\varepsilon_0\to 0$. Likewise the displacement, we write strain $\boldsymbol{\epsilon}^{\varepsilon_0}$ and rotation $\boldsymbol{\omega}^{\varepsilon_0}$ as:

$$\boldsymbol{\epsilon}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t)=\boldsymbol{\epsilon}^0(\mathbf{x},\mathbf{y},t)+\varepsilon_0\boldsymbol{\epsilon}^1(\mathbf{x},\mathbf{y},t)+\varepsilon_0^2\boldsymbol{\epsilon}^2(\mathbf{x},\mathbf{y},t)+\dots \tag{4.18}$$

$$\omega^{\varepsilon_0}(\mathbf{x},\mathbf{y},t)=\omega^0(\mathbf{x},\mathbf{y},t)+\varepsilon_0\omega^1(\mathbf{x},\mathbf{y},t)+\varepsilon_0^2\omega^2(\mathbf{x},\mathbf{y},t)+\dots\,. \tag{4.19}$$

We need to look for the coefficients $\boldsymbol{\epsilon}^0(\mathbf{x},\mathbf{y},t)$ and $\omega^0(\mathbf{x},\mathbf{y},t)$. To find these coefficients, we take the gradient of equation (4.17):

$$\begin{aligned}
\nabla\mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) =& \nabla_{\mathbf{x}}\mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) + \frac{1}{\varepsilon_0}\nabla_{\mathbf{y}}\mathbf{u}^{\varepsilon_0} \\
=& \nabla_{\mathbf{x}}\{\mathbf{u}^*(\mathbf{x},t) + \varepsilon_0\mathbf{u}^1(\mathbf{x},\mathbf{y},t) + \varepsilon_0^2\mathbf{u}^2(\mathbf{x},\mathbf{y},t) + ...\} \\
&+ \nabla_{\mathbf{y}}\{\mathbf{u}^1(\mathbf{x},\mathbf{y},t) + \varepsilon_0\mathbf{u}^2(\mathbf{x},\mathbf{y},t) + \varepsilon_0^2\mathbf{u}^3(\mathbf{x},\mathbf{y},t) + ...\} \\
=& \nabla_{\mathbf{x}}\mathbf{u}^*(\mathbf{x},t) + \nabla_{\mathbf{y}}\mathbf{u}^1(\mathbf{x},\mathbf{y},t) + \varepsilon_0\{\nabla_{\mathbf{x}}\mathbf{u}^1(\mathbf{x},\mathbf{y},t) + \nabla_{\mathbf{y}}\mathbf{u}^2(\mathbf{x},\mathbf{y},t)\} + ... \\
=& \nabla_{\mathbf{x}}\mathbf{u}^*(\mathbf{x},t) + \nabla_{\mathbf{y}}\boldsymbol{\chi}(\mathbf{x},\mathbf{y}) : \boldsymbol{\epsilon}(\mathbf{u}^*) + \varepsilon_0\{\nabla_{\mathbf{x}}\mathbf{u}^1(\mathbf{x},\mathbf{y},t) + \nabla_{\mathbf{y}}\mathbf{u}^2(\mathbf{x},\mathbf{y},t)\} + ...
\end{aligned} \tag{4.20}$$

Therefore, the strain expantion can be written as:

$$\boldsymbol{\epsilon}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) = \boldsymbol{\epsilon}^*(\mathbf{x},t) + \boldsymbol{\epsilon}_{\mathbf{y}}(\boldsymbol{\chi})(\mathbf{x},\mathbf{y}) : \boldsymbol{\epsilon}^*(\mathbf{x},t) + \varepsilon_0(\boldsymbol{\epsilon}_{\mathbf{x}}(\mathbf{u}^1) + \boldsymbol{\epsilon}_{\mathbf{y}}(\mathbf{u}^2)) + ..., \tag{4.21}$$

where we write simply $\boldsymbol{\epsilon}^*(\mathbf{x},t) = \boldsymbol{\epsilon}(\mathbf{u}^*)(\mathbf{x},t)$. Simiraly, take the curl operator of equation (4.17):

$$\begin{aligned}
\nabla\times\mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) =& \nabla_{\mathbf{x}}\mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) + \frac{1}{\varepsilon_0}\nabla_{\mathbf{y}}\times\mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) \\
=& \nabla_{\mathbf{x}}\times\mathbf{u}^*(\mathbf{x},t) + (\nabla_{\mathbf{y}}\times\boldsymbol{\chi}(\mathbf{x},\mathbf{y})) : \boldsymbol{\epsilon}(\mathbf{u}^*)(\mathbf{x},t) \\
&+ \varepsilon_0\{\nabla_{\mathbf{x}}\times\mathbf{u}^1(\mathbf{x},\mathbf{y},t) + \nabla_{\mathbf{y}}\times\mathbf{u}^2(\mathbf{x},\mathbf{y},t)\} + ...
\end{aligned} \tag{4.22}$$

Then, the rotation expation can be written as:

$$\omega^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) = \omega^*(\mathbf{x},t) + \frac{1}{2}(\nabla_{\mathbf{y}}\times\boldsymbol{\chi}(\mathbf{x},\mathbf{y})) : \boldsymbol{\epsilon}^*(\mathbf{x},t) + \varepsilon_0\{\omega^1_{\mathbf{x}}(\mathbf{x},\mathbf{y},t) + \omega^2_{\mathbf{y}}(\mathbf{x},\mathbf{y},t)\} + ..., \tag{4.23}$$

where $\omega^*(\mathbf{x},t) = \frac{1}{2}\nabla_{\mathbf{x}}\times\mathbf{u}^*(\mathbf{x},t)$, $\omega^1_{\mathbf{x}}(\mathbf{x},t) = \frac{1}{2}(\nabla_{\mathbf{x}}\times\mathbf{u}^1(\mathbf{x},\mathbf{y},t))$, and $\omega^2_{\mathbf{y}}(\mathbf{x},\mathbf{y},t) = \frac{1}{2}(\nabla_{\mathbf{y}}\times\mathbf{u}^2)(\mathbf{x},\mathbf{y},t)$. Finally, comparing the equations (4.18) and (4.19) with the equations (4.21) and (4.23) respectively, we can obtain zeroth order terms of strain and rotation:

$$\boldsymbol{\epsilon}^0(\mathbf{x},\mathbf{y},t) = \boldsymbol{\epsilon}^*(\mathbf{x},t) + \boldsymbol{\epsilon}_{\mathbf{y}}(\boldsymbol{\chi})(\mathbf{x},\mathbf{y}) : \boldsymbol{\epsilon}^*(\mathbf{x},t) \tag{4.24}$$

$$\omega^0(\mathbf{x},\mathbf{y},t) = \omega^*(\mathbf{x},t) + \frac{1}{2}(\nabla_{\mathbf{y}}\times\boldsymbol{\chi}(\mathbf{x},\mathbf{y})) : \boldsymbol{\epsilon}^*(\mathbf{x},t)\,. \tag{4.25}$$

Therefore, we can deduce significant results about displacement, strain and rotation with the leading order approximation. For the displacement:

$$\mathbf{u}(\mathbf{x},t) = \mathbf{u}^*(\mathbf{x},t) + O(\varepsilon_0)\,. \tag{4.26}$$

For the strain:

$$\begin{aligned}&\boldsymbol{\epsilon}(\mathbf{x},t) = \boldsymbol{\epsilon}^*(\mathbf{x},t) + \boldsymbol{\epsilon}(\boldsymbol{\chi})(\mathbf{x}) : \boldsymbol{\epsilon}^*(\mathbf{x},t) + O(\varepsilon_0)\,,\\ &\text{where: } \boldsymbol{\epsilon}(\boldsymbol{\chi})(\mathbf{x}) = \boldsymbol{\epsilon}(\boldsymbol{\chi})(\mathbf{x},\mathbf{y})|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\end{aligned} \tag{4.27}$$

For the rotation:

$$\begin{aligned}&\omega(\mathbf{x},t) = \omega^*(\mathbf{x},t) + \mathbf{J}(\mathbf{x}) : \boldsymbol{\epsilon}^*(\mathbf{x},t) + O(\varepsilon_0)\,,\\ &\text{where: } \mathbf{J}(\mathbf{x}) = \frac{1}{2}\boldsymbol{\nabla}_{\mathbf{y}} \times \boldsymbol{\chi}(\mathbf{x},\mathbf{y})|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\,.\end{aligned} \tag{4.28}$$

In addition, from the equation (4.20), the gradient can be written as:

$$\begin{aligned}&\boldsymbol{\nabla}\mathbf{u}(\mathbf{x},t) = \boldsymbol{\nabla}\mathbf{u}^*(\mathbf{x},t) + \mathbf{G}(\mathbf{x}) : \boldsymbol{\epsilon}^*(\mathbf{x},t) + O(\varepsilon_0)\,,\\ &\text{where: } \mathbf{G}(\mathbf{x}) = \boldsymbol{\nabla}_{\mathbf{y}}\boldsymbol{\chi}(\mathbf{x},\mathbf{y})|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\,.\end{aligned} \tag{4.29}$$

In practice, once the homogenized medium ($\rho^*, \mathbf{c}^*$) is obtained, the homogenized wave equation (equation (4.15)), which is the standard wave equation, can be then solved to obtain the effective displacement $\mathbf{u}^*$ and its derivative $\boldsymbol{\epsilon}^*$, $\boldsymbol{\nabla}\mathbf{u}^*$ and ω^*. Nevertheless, if using $\mathbf{u}^*$ we obtain directly the true displacement with a good approximation via equation (4.26), this is not the case for the strain and rotational wavefields respectively. Equations (4.27)–(4.29) show that an extra term is needed to obtain the correct leading-order approximations for the effective gradient, strain and rotational fields. This extra term depends on $\mathbf{y}$, thereby suggesting that the waveforms of the wavefield gradients are affected by the small-scale structures of the medium. It also implies that the gradient and rotational field are sensitive to small-scale structures, whereas the displacement is not. In practice, this small-scale heterogeneity causes the strain components to couple to the expected gradient of the rotational components. This effect has already been observed in both synthetic and real data (van Driel *et al.* 2012; Singh *et al.* 2020; Muir & Zhan 2022).

4.3 Full waveform inversion based on homogenization theory

4.3.1 Context

We use a Cartesian coordinate system in two-dimensional (2D elastic model domain in $\boldsymbol{\Omega}$ with absorbing boundary conditions on its boundary $\partial\boldsymbol{\Omega}$, where the x-axis is the horizontal axis and the z-axis is the vertical axis. The particle displacement vector $\mathbf{u}(\mathbf{x}, t)$ in $\mathbf{\Omega}$ is driven by the elastic wave equations (2.1)–(2.2).

The mechanical properties $(\rho(\mathbf{x}), \mathbf{c}(\mathbf{x}))$ are considered heterogeneous at all spatial scales, as expected in geological media. The two-scale homogenisation mathematical framework is therefore appropriate to deal with such heterogeneities.

4.3.2 Earth models and associated spaces

We first define the space of possible Earth models prior to introducing the inverse problem. An Earth model $\mathbf{m}$ is defined by its mechanical properties $(\rho(\mathbf{x}), \mathbf{c}(\mathbf{x}))$ at any location $\mathbf{x}$ in the domain $\mathbf{\Omega}$. We define the admissible model space $\mathcal{M}$, which contains all of the physically admissible Earth models. If we assume that $\mathbf{\Omega}$ is heterogeneous at all scales (similar to the real Earth), then $\mathcal{M}$ is infinite-dimensional.

We can define the homogenised counterpart of each model $\mathbf{m} \in \mathcal{M}$ at a given frequency band as $\mathbf{p}^* = \mathcal{H}(\mathbf{m})$. Equation (4.16) shows that $\mathbf{p}^*$ is more than just the effective mechanical properties $\mathbf{m}^* = (\rho^*, \mathbf{c}^*)$, as it generally contains the corrector χ. $\mathbf{m}^*$ is the restriction of $\mathbf{p}^*$ to the mechanical properties: $\mathbf{m}^* = \mathbf{p}^*|_{\mathbf{m}}$. Note that the actual definition of the effective model generally depends on the receiver geometry, type of observation and homogenisation order. If there are only displacement observations and the homogenisation order is 0, then there is no need for correctors. However, if the observations are rational, even for zeroth-order homogenisation, then the corrector curl is needed at the receiver locations (see equation (4.28)). We therefore define the effective model space $\mathcal{M}^*$, which contains all of the possible effective models $\mathbf{p}^*$. Symbolically, we have $\mathcal{M}^* = \mathcal{H}(\mathcal{M})$, where $\mathcal{M}^*$ is finite-dimensional.

4.3.3 Full waveform inversion based on displacement data

We consider an idealised FWI problem whose objective is to recover the true mechanical properties $(\rho^t, \mathbf{c}^t)$ of the elastic domain $\mathbf{\Omega}$ using seismic waveforms. We assume that the true model $\mathbf{m}_t = (\rho^t, \mathbf{c}^t)$ belongs to $\mathcal{M}$. We consider $\mathbf{\Omega}$ with N_r receivers located at $\mathbf{x}_r$, $r \in \{1, ..., N_r\}$ and N_s sources located at $\mathbf{x}_s$, $s \in \{1, ..., N_s\}$. The signal from each source is recorded for a time duration T at each receiver. We assume that the displacement data set is accurately modelled by solving the wave equation (equations (2.1)–(2.2)). We define the least-squares misfit function and associated minimisation problem as:

$$\begin{aligned} E(\mathbf{m}) &= \sum_{r,s} \int_0^T (\mathbf{d}_s(\mathbf{x}_r, t) - \mathbf{u}(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}))^2 dt\,, \\ \bar{\mathbf{m}} &= \underset{\mathbf{m}\in\mathcal{M}}{\operatorname{argmin}}\, E(\mathbf{m})\,, \end{aligned} \tag{4.30}$$

where $\mathbf{d}_s(\mathbf{x}_r, t)$ represents the displacement data generated by source number s and recorded by receiver number r. Under perfect conditions, the inverse problem has a unique solution and $\bar{\mathbf{m}} = \mathbf{m}_t$ (Nachman 1988; Nakamura & Uhlmann 1994). However, equation (4.30) is generally an unsolvable problem because $\mathcal{M}$ is an infinite-dimensional space and the numerical cost of solving the wave equation is not bounded.

The classical strategy to obtain a solvable inverse problem is to limit the data frequency band. Therefore, a low-pass filter $\mathcal{F}^{f_\mathrm{m}}$ is introduced, such that the filtered data $\mathbf{d}_s^{f_\mathrm{m}} = \mathcal{F}^{f_\mathrm{m}}(\mathbf{d}_s)$ have no signal beyond the maximum frequency f_m in the frequency domain. A new misfit is then defined from this filtered data set as follows:

$$E^{f_\mathrm{m}}(\mathbf{m}) = \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_\mathrm{m}}(\mathbf{x}_r, t) - \mathbf{u}^{f_\mathrm{m}}(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}))^2 dt\,, \tag{4.31}$$

where $\mathbf{u}^{f_\mathrm{m}}$ are the synthetic data, which are computed with a low-pass-filtered source time function $g_s^{f_\mathrm{m}} = \mathcal{F}^{f_\mathrm{m}}(g_s)$. The maximum frequency f_m ensures that the wavefield has a minimum wavelength λ_min for most media (see Zhao *et al.* (2016) and Capdeville *et al.* (2020) for exceptions). Based on the common assumption that displacement wavefields are insensitive to spatial heterogeneities that are smaller than λ_min, it is assumed that it is possible to introduce a finite-dimensional model space $\mathcal{M}^h$, whereby at least one model that minimises (4.31) can be found. Furthermore, $\mathcal{M}^h$ can be designed in such a way that the wave equation (equations (2.1)–(2.2)) can be solved in a bounded time for every one of its models. This assumption and model design yield the following inverse problem:

$$\bar{\mathbf{m}}^h = \underset{\mathbf{m}\in\mathcal{M}^h}{\operatorname{argmin}}\, E^{f_\mathrm{m}}(\mathbf{m}) \tag{4.32}$$

that can be solved in practice, where h is a ‘resolution’ parameter, which is a number that characterises the discretisation made in the finite-dimensional approximation. For example, if the elastic model is spatially represented by constant velocity blocks, then h can be the size of these blocks. h is generally directly related to λ_min; however, other information, such as the illumination angles, offset ranges or data coverage, can influence the choice of h.

One serious drawback arises when the introduction of frequency-limited data and $\mathcal{M}^h$ make the inverse problem solvable: the relation between the true model $\mathbf{m}_t$ and the inverse problem solution $\bar{\mathbf{m}}^h$ is unknown. Furthermore, there are many different ways to design $\mathcal{M}^h$ for a given experiment geometry and λ_min, with a different $\bar{\mathbf{m}}^h$ for each of them.

Capdeville & Métivier (2018) outlined an approach to solve this problem as follows. For a given signal maximum frequency f_m, we search for a solution in $\mathcal{M}^*$, as defined in the previous section. We therefore define a new inverse problem for a fixed f_{m}:

$$\begin{aligned}
E^*(\mathbf{m}^*) &= \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_{\mathrm{m}}}(\mathbf{x}_r, t) - \mathbf{u}^*(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}^*))^2 dt \,, \\
\bar{\mathbf{m}}^* &= \operatorname*{argmin}_{\mathbf{m}^* \in \mathcal{M}^*} E^*(\mathbf{m}^*) \,,
\end{aligned} \tag{4.33}$$

where $\mathbf{u}^*$ is the leading-order homogenised displacement, which is defined in equation (4.3). Here, we only use the displacement observations and zeroth-order homogenisation, which implies that no corrector is needed in the definition of $\mathcal{M}^*$. This simplifies $\mathbf{u}^*$ to the displacement obtained in the effective media $\mathbf{m}^*$, such that $\mathbf{u}^* = \mathbf{u}^{f_{\mathrm{m}}}(\mathbf{m}^*)$. As shown in equation (4.26), $\mathbf{u}^*$ is insensitive to small-scale heterogeneities.

For any model $\mathbf{m}$ and its homogenised version $\mathbf{m}^* = \mathcal{H}(\mathbf{m})$, it can be shown that:

$$E^{f_{\mathrm{m}}}(\mathbf{m}) = E^*(\mathbf{m}^*) + O(\varepsilon_0) \,. \tag{4.34}$$

The true $\mathbf{m}_t$ minimises E and also $E^{f_{\mathrm{m}}}$; therefore, $\mathbf{m}_t^* = \mathcal{H}(\mathbf{m}_t)$ is a solution of the homogenisation FWI (HFWI) problem (4.33). If we assume this solution is unique (this a strong assumption; see the discussion in Capdeville & Métivier (2018)), then the solution of the HFWI problem is related to the true model as follows:

$$\bar{\mathbf{m}}^* = \mathcal{H}(\mathbf{m}_t) \,. \tag{4.35}$$

This important result shows we can access the effective true model by solving (4.33), which is a solvable inverse problem. In practice, we can only set up an explicit parametrisation of $\mathcal{M}^*$ in the layered model case. We need to rely on an approximate space model $\mathcal{M}^{*h}$ for the general case and use the homogenisation operator $\mathcal{H}$ to project the solution into $\mathcal{M}^*$. In practice, the construction of $\mathcal{M}^{*h}$ is very similar to a classical construction of $\mathcal{M}^h$. However, we need to have $\mathcal{M}^* \subset \mathcal{H}(\mathcal{M}^{*h})$, which often implies that $\mathcal{M}^{*h}$ is fully anisotropic. The numerical tests shown in Capdeville & Métivier (2018) show that the solution of the HFWI problem is indeed the homogenised true model.

4.3.4 Full waveform inversion based on rotational waveforms

We now consider the case where the observations $\mathbf{d}$ are rotational waveforms. The FWI inverse problem is very similar to the displacement case, and it also needs to be based on limited frequency band data to be solvable:

$$
\begin{aligned}
&E^{f_m}(\mathbf{m}) = \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_m}(\mathbf{x}_r, t) - \omega^{f_m}(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}))^2 dt\,, \\
&\bar{\mathbf{m}}^h = \underset{\mathbf{m}\in\mathcal{M}^h}{\operatorname{argmin}} E^{f_m}(\mathbf{m})\,.
\end{aligned}
\tag{4.36}
$$

Although this problem is solvable, the relationship between the solution $\bar{\mathbf{m}}^h$ and true model $\mathbf{m}_t$ is lost.

Following what was done for the displacement case in the previous section, we can define a homogenised inverse problem based on rotational data as follows:

$$
\begin{aligned}
&E^*(\mathbf{m}^*) = \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_m}(\mathbf{x}_r, t) - \omega^*(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}^*))^2 dt\,, \\
&\bar{\mathbf{m}}^* = \underset{\mathbf{m}^*\in\mathcal{M}^*}{\operatorname{argmin}} E^*(\mathbf{m}^*)\,.
\end{aligned}
\tag{4.37}
$$

However, $\omega^{f_m} \neq \omega^*$ for the leading-order approximation, which means that $E^{f_m}(\mathbf{m}) \neq E^*(\mathbf{m}^*)$. This implies that we cannot relate the solution $\bar{\mathbf{m}}^*$ of the inverse problem to the true model $\mathbf{m}_t$. In practice, a solution to the above problem can be found, but it contains errors and/or differences compared with the homogenised true model $\mathbf{m}_t^*$ that cannot be controlled. We can correct this problem by relying on ω^0 instead of ω^*:

$$
\begin{aligned}
&E^*(\mathbf{p}^*) = \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_m}(\mathbf{x}_r, t) - \omega^0(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*))^2 dt\,, \\
&\bar{\mathbf{p}}^* = \underset{\mathbf{p}^*\in\mathcal{M}^*}{\operatorname{argmin}} E^*(\mathbf{p}^*)\,,
\end{aligned}
\tag{4.38}
$$

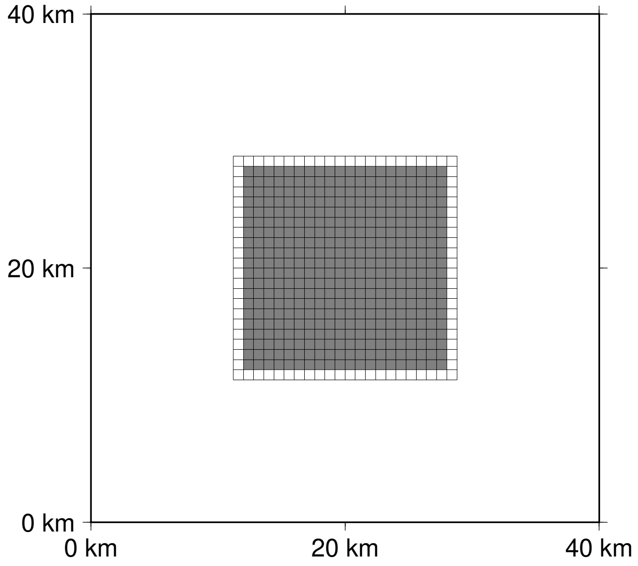
where $\omega^0(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*) = \omega^*(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*|_\mathbf{m}) + \mathbf{J}_r : \boldsymbol{\epsilon}^*(\mathbf{u}^*)$, with $\mathbf{J}_r = \mathbf{J}(\mathbf{x}_r)$, from equation (4.28). Here $\mathcal{M}^*$ contains all of the possible models $\mathbf{p}^* = ((\rho^*, \mathbf{c}^*), \mathbf{J}_r, r \in \{1, ..., N_r\})$. The last inverse problem definition leads to $E^{f_m}(\mathbf{m}) = E^*(\mathbf{p}^*) + O(\varepsilon_0)$, which makes sure that we can relate the inverse problem solution $\bar{\mathbf{m}}^* = \bar{\mathbf{p}}^*|_\mathbf{m}$ to the effective true model $\mathbf{m}_t^*$.

4.4 Inversion scheme

We use the same inversion framework as in Capdeville & Métivier (2018), whereby we spatially parametrise $\mathcal{M}^{*h}$ by dividing the squared inversion domain $\mathcal{I} \in \boldsymbol{\Omega}$ into non-overlapping $n \times n$ square elements: $\mathcal{I} = \cup_{e=1}^{n^2} \mathcal{I}_e^n$. An example of inversion domain $\mathcal{I}$ with 22

$\times$ 22 elements ($\mathcal{I}_e^{22}$) is shown in Fig. 4.2. A $N \times N$ degree polynomial basis is then used for each element $\mathcal{I}_e^n$, and each inverted mechanical property is expanded within each element. Here we name such a parametrisation $\mathcal{P}_n^N(\mathcal{I})$. We use a fully anisotropic parametrisation for $\mathcal{M}^{*h}$, thereby implying that each c_{ij} and ρ is inverted. Therefore, we have $7 \times n^2 \times (N+1)^2$ free parameters to characterise $\mathcal{M}^{*h}$ if no correctors are inverted.

(A) Mesh example



(B) Target model

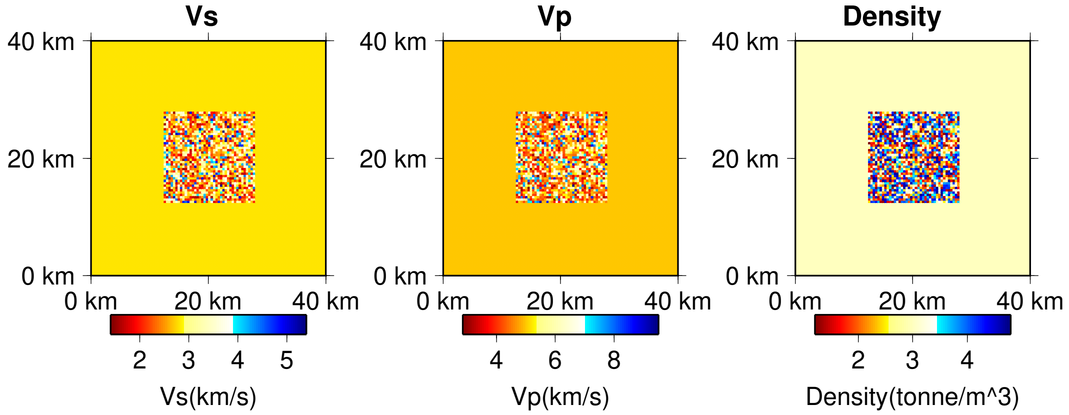


Fig. 4.2 (A) Schematic diagram of the $\mathcal{I}_e^{22}$ mesh structure. The shaded grey area shows the target area to be inverted. Black lines define the element boundaries. (B) V_S (left), V_P (centre) and density (right) structures of the target model.

4.4.1 Gauss–Newton iterative inversion

We iteratively update the model parameters using a Gauss–Newton inversion scheme to minimise the misfit function. The i-th iteration model $\mathbf{m}^i$ for the rotational wavefields is updated as follows:

$$\mathbf{m}^{i+1} = \mathbf{m}^i + ((\mathbf{F}^i)^T \cdot \mathbf{F}^i + \lambda^i)^{-1}[(\mathbf{F}^i)^T \cdot (\mathbf{d} - \omega(\mathbf{m}^i))] , \tag{4.39}$$

where $\mathbf{F}^i$ is the partial derivative matrix and $\boldsymbol{\lambda}^{\boldsymbol{i}}$ is the damping parameter, which is used to stabilise the inversion. The damping parameter $\boldsymbol{\lambda}^{\boldsymbol{i}}$ is optimised in each numerical test. The partial derivative matrix $\mathbf{F}^i$ with respect to the model parameters were evaluated using the adjoint technique (Tarantola 1984; Tarantola 1988; Pratt *et al.* 1998). The detailed calculations of the partial derivatives for both the displacement and rotational wavefields are provided in appendix A.

We set up two criteria for convergence:

- A 98% decrease in the misfit between the observed and synthetic data, where the initial model in the inversion is a homogeneous model.
- A <1% decrease in the misfit from the previous iteration.

4.4.2 FWI with correctors

The model space for a rotational-based FWI includes both the effective mechanical properties and correctors. Adding the corrector parameters to the inverted parameters can be done via the same framework that was used for the elastic parameters alone. The Gauss–Newton scheme remains the same for this case; the main modifications are an increase in the size of the partial Hessian matrix to be inverted and slightly more complex programming to keep track of the inverted parameters. The adjoint sources for the rotational data are moment tensors, as explained in appendix A. Therefore, moment tensors of the adjoint sources for the rotational data should be corrected with inverted correctors (Burgos *et al.* 2016; Capdeville 2021).

Here we rely on a two-stage approach to simplify the computations. During each Gauss–Newton iteration, the mechanical parameters are inverted to obtain the current model $\mathbf{m}^i$ in the first stage, and the correctors are then independently inverted following Singh *et al.* (2020) in the second stage. We minimise the squared error by summing over all of the sources for a given rotational receiver location $\mathbf{x}_r$ as follows:

$$\begin{aligned} E(\mathbf{J}_r) &= \sum_s \int_0^T (\mathbf{d}_s(\mathbf{x}_r, t) - \omega^0(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*))^2 dt\,, \\ &\text{where: } \omega^0(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*) = \omega^*(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*|_{\mathbf{m}}) + \mathbf{J}_r : \boldsymbol{\epsilon}^*(\mathbf{u}^*)\,, \end{aligned} \tag{4.40}$$

and ω^* and $\boldsymbol{\epsilon}^*$ are the synthetic data and strain computed in the current mechanical model $\mathbf{m}^i$, respectively. The misfit is simple quadratic with respect to $\mathbf{J}_r$ and can be inverted via a simple least-squares approach:

$$\mathbf{J}_r = (\mathbf{F}_r^t \mathbf{F}_r)^{-1} \mathbf{F}_r^t \mathbf{b}, \tag{4.41}$$

where $\mathbf{F}_r$ is the partial derivative matrix that is limited to specific receiver, and $\mathbf{b}$ is the residual between the target and synthetic data. This calculation is a small computation (three and six free parameters for 2D and three-dimensional (3D) problems, respectively) and under-parameterised, which implies that damping is not needed to obtain a stable result. Fig.4.3 shows the overview of the FWI with correctors employed in this study.

This two-step inversion process is expected to slow down the Gauss–Newton inversion convergence compared with a single-step inversion scheme. However, we generally do not observe a significant difference in the number of iterations compared with the cases where correctors are not needed.

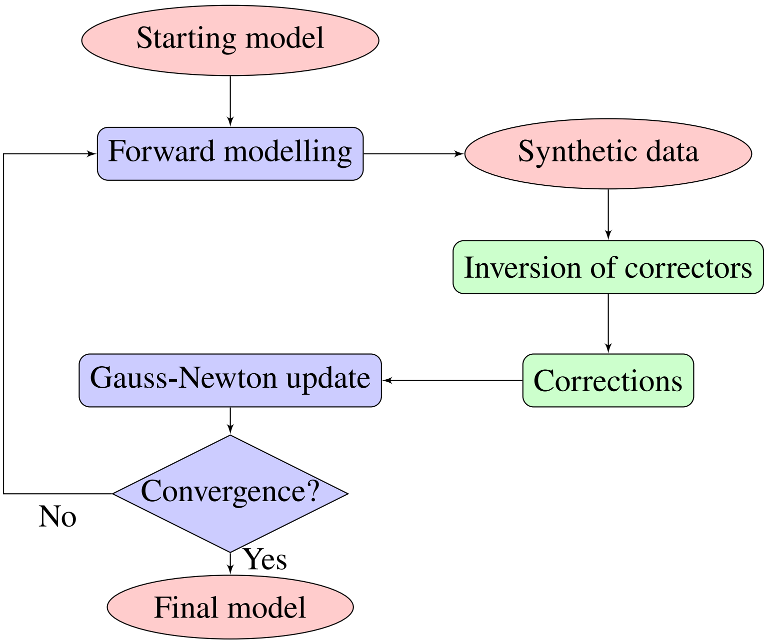


Fig. 4.3 Flowchart of the FWI process, with correctors. The process begins with forwarding modelling to generate synthetic data. The correctors are inverted, and the synthetic data are corrected using the inverted correctors. Then, the inversion of the elastic mechanical properties is performed. If convergence is not obtained, the process is repeated.

4.5 Numerical experiments

4.5.1 Experiment design

Target model

The target model that the FWI tries to estimate is a 2D elastic medium, such as that shown in figure 4.2. The $\mathbf{\Omega}$ domain is a 40×40 km region that consists of a homogeneous background with a 16 × 16 km heterogeneous area embedded at the centre. This heterogeneous area is made of 40 × 40 square anomalies, with each square possessing constant mechanical properties that are randomly generated. The background model is isotropic with shear and compressional velocities (V_S and V_P, respectively), and a density of 2.8 km/s, 5.0 km/s and 3.0 t.m^{-3}, respectively. The random mechanical properties of the heterogeneous region are generated using a flat probability distribution that is within 60% of the elastic tensor and density of the homogeneous background values.

Spectral element simulation setup

We use the spectral element method (Komatitsch & Vilotte 1998) to solve the wave equations (equations (2.1)-(2.2)) for both generating the data to be inverted and solving the inverse problem. The absorbing boundaries for each of our experimental cases are based on perfectly matched layers (Festa & Vilotte 2005) that are used all edges of the domain. Two different spectral element meshes are used:

- We rely on a square 100 × 100 element mesh to generate the data to be inverted. We use a polynomial degree of 10 in each tensorial direction of each element. The mesh is designed such that all of the material discontinuities are honoured by an element boundary, which is important to obtain good accuracy.
- A different spectral element mesh than the previous one is used when computing the full wavefields from the sources and adjoint sources to compute the partial derivatives. This mesh consists of 50 × 50 elements with a polynomial degree of 4.

The different spatial meshes for generating the target and synthetic data sets during the inversion process allows us to avoid an "inverse crime" because the numerical errors due to spatial discretisation are different. Furthermore, we do not introduce any a priori knowledge of the heterogeneity geometry in the inversion modelling mesh.

FWI acquisition geometry

The source and receiver geometries used in our numerical experiments are shown in Fig. 4.4. The sources are located all around the inverted area, with each consisting of a single vector force that produces a Ricker wavelet. This wavelet possesses a 0.25-Hz central frequency and 0.7-Hz maximum source frequency. The associated minimum wavelength in the background medium is 4 km, which is 10 times larger than the lengths of the small anomalies in the target model.

We use two different geometries for the receivers:

- Configuration A (Fig. 4.4): The receivers are located around the inverted area and are far from the heterogeneities.
- Configuration B (Fig. 4.4): The receivers are located within the inverted area and are in contact with the heterogeneities.

This setup is designed to investigate the impact of small-scale heterogeneities on FWI using rotational wavefields recorded using two receiver configurations: closely located receivers (configuration B) or not (configuration A).

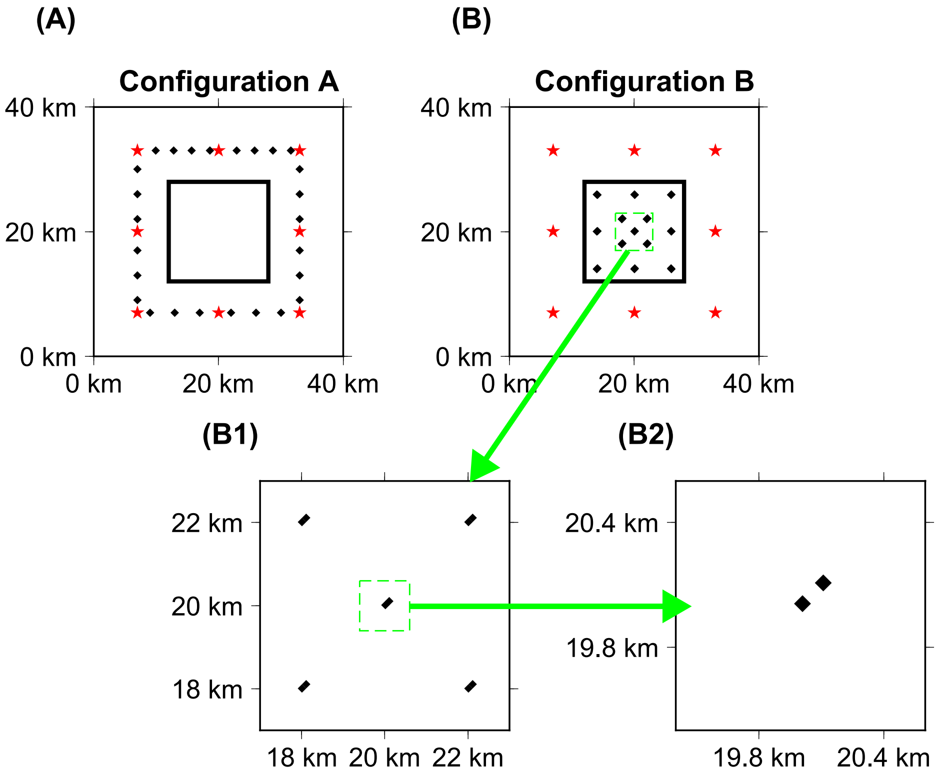


Fig. 4.4 Source (red stars) and receiver (black diamonds) configurations employed for the synthetic tests. (A) Configuration A: All sources and receivers are positioned outside of the heterogeneous area. (B) Configuration B: All sources are positioned outside of the heterogeneous area, whereas all receivers are positioned within the heterogeneous area. (B1,B2) Close-up views of the receiver arrangement in Configuration B for the areas indicated by the dashed green squares.

Representation and evaluation of the FWI results

We first need to define what will be presented and how the results will be assessed prior to conducting the numerical experiments.

The FWI results are always a full elastic tensor $\mathbf{c}$ after homogenization. We present the results by first projecting the anisotropic elastic tensor $\mathbf{c}^*$ to the nearest isotropic tensor $\mathbf{c}^{*,\mathrm{iso}}$ following Browaeys & Chevrot (2004). The compressional- and shear-wave velocities are defined as follows:

$$V_P^*(\mathbf{x}) = \sqrt{(c_{1111}^{*,\mathrm{iso}}(\mathbf{x})/\rho(\mathbf{x}))}\,, \tag{4.42}$$

and:

$$V_S^*(\mathbf{x}) = \sqrt{(c_{1212}^{*,\mathrm{iso}}(\mathbf{x})/\rho(\mathbf{x}))}\,, \tag{4.43}$$

respectively, and the total anisotropy is defined as:

$$\operatorname{aniso}(\mathbf{x}) = \frac{\sqrt{\sum_{ijkl}(c^{\mathrm{iso}}_{ijkl}(\mathbf{x}) - c_{ijkl}(\mathbf{x}))^2}}{\sqrt{\sum_{ijkl}(c^{\mathrm{iso}}_{ijkl}(\mathbf{x}))^2}} . \tag{4.44}$$

One objective of our work is to compare the rotational-based FWI results to the displacement-based FWI results, particularly to assess the accuracy and resolution of the two approaches. A classical way to perform this analysis is to analyse the diagonal of the partial Hessian and compute the point spread functions (Fichtner & Trampert 2011). However, our use of a homogenisation step in the FWI, which is a non-linear step, makes it impossible for us to work directly on the partial Hessian. We overcome this difficulty by performing a brute force statistical analysis. We generate 200 different data sets for both the displacement and rotational receiver configuration cases by adding random noise to the data. This added noise possesses a standard deviation of 10% of the original data set and spans the same frequency band as the original data set. Two hundred inversions are then performed, one for each data set. We define the inversion error as follows:

$$E^2_m(\varepsilon_0, \mathbf{x}) = \frac{\sum_{n=1}^{n=N} \sum_{ij}(c^{*,n}_{ij}(\mathbf{x}) - c^{*,t}_{ij}(\mathbf{x}))^2}{N \sum_{ij}(c^{*,t}_{ij}(\mathbf{x}))^2} , \tag{4.45}$$

where $\mathbf{c}^{*,n}$ and $\mathbf{c}^{*,t}$ are the homogenised FWI model results for noise number n and the true homogenised elastic tensor, respectively, for a given ε_0. We finally define the inverted area-averaged inversion error as:

$$\bar{E}^2_m(\varepsilon_0) = \frac{\int_{\mathcal{I}} \sum_{n=1}^{n=N} \sum_{ij}(c^{*,n}_{ij}(\mathbf{x}) - c^{*,t}_{ij}(\mathbf{x}))^2 d\mathbf{x}}{N \int_{\mathcal{I}} \sum_{ij}(c^{*,t}_{ij}(\mathbf{x}))^2 d\mathbf{x}} , \tag{4.46}$$

Analysing $\bar{E}^2_m(\varepsilon_0)$ as a function of ε_0 allows us to find an estimate of the smallest ε_0 that can be used. This smallest ε_0 can be seen as the overall resolution limit of the inversion. Constraining ε_0 in this manner highlights the compromise between the resolution and error in the inversion: using a smaller ε_0 would generate higher-resolution images with larger error, whereas using a larger ε_0 instils good confidence in the accuracy of the imaged structures at a lower spatial resolution.

We also assess the inversion results using three statistical measures: the standard deviations, mean values and resolutions (correlations) of V^*_S and V^*_P (obtained from equations (4.42) and (4.43), respectively), and density. The mean values are the same as the noise-free inversion results and are used to compare the inversion results and target model for a given ε_0. We note that the mean values and noise-free inversion results are not exactly the same, but it does not change the point of the paper and we ignore this differences. For a given ε_0, the covariance of mechanical property p^* (V_P^*, V_S^* or the effective density) is written as:

$$\mathrm{Cov}^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \frac{\sum_{n=1}^{N} (p_n^*(\mathbf{x}) - \mu_p(\mathbf{x}))(p_n^*(\mathbf{y}) - \mu_p(\mathbf{y}))}{N}, \tag{4.47}$$

where μ_p is the mean value of p, and N is the number of samples. The deviation σ^{ε_0} can be derived in terms of the covariance as follows:

$$\sigma^{\varepsilon_0}(\mathbf{x}) = \sqrt{\mathrm{Cov}^{\varepsilon_0}(\mathbf{x}, \mathbf{x})}, \tag{4.48}$$

such that the resolutions R^{ε_0} are then written as follows:

$$R^{\varepsilon_0}(\mathbf{x}, \mathbf{y}) = \frac{\mathrm{Cov}^{\varepsilon_0}(\mathbf{x}, \mathbf{y})}{\sigma^{\varepsilon_0}(\mathbf{x})\sigma^{\varepsilon_0}(\mathbf{y})}. \tag{4.49}$$

Here we provide the V_S^* results for our presented inversion and statistical analyses. Comprehensive results, which include the V_P^* and effective density results, are provided in the appendix. Note that both σ^{ε_0} and R^{ε_0} depend on ε_0. This co-dependence introduces a trade-off between σ^{ε_0} and R^{ε_0}. Therefore, the resolution limit and overall error cannot be measured independently for this inversion process.

4.5.2 A simple forward modelling example of a small-scale heterogeneity effect on the data

We first illustrate the different impacts of small-scale heterogeneities on displacement and rotational measurements prior to introducing the FWI results. Example of rotational and displacement waveforms for the two source—receiver configurations are shown in Fig.4.5. Here, four receivers are located in close proximity to each other at two distant locations, A and B, whereby they are spaced at 100-m horizontal and vertical intervals from each other. We acquired both the displacement and rotational waveforms using a source that generates a Ricker wavelet with a 0.7-Hz maximum frequency and 4-km minimum wavelength. It should be noted that station intervals and lengths of the anomalies are much smaller than the 4-km minimum wavelength. The displacement waveforms from the four receivers at location A, which is within the heterogeneous area, are the same, as expected, whereas there are observable differences among the rotational waveforms. These differences in the rotational waveforms are due to the corrector in equation (4.28). Conversely, the rotational waveforms from the four receivers at location B, which is ~2 km from the heterogeneous area, are the same. This is because the effect of the corrector is limited to the heterogeneous area, with no small-scale effects propagating outside of the heterogeneous area.

This example clearly highlights that a smooth tomographic model would have a hard time explaining data that are contaminated by small-scale heterogeneities. It also underscores why it is necessary introduce correctors to account for such effects.

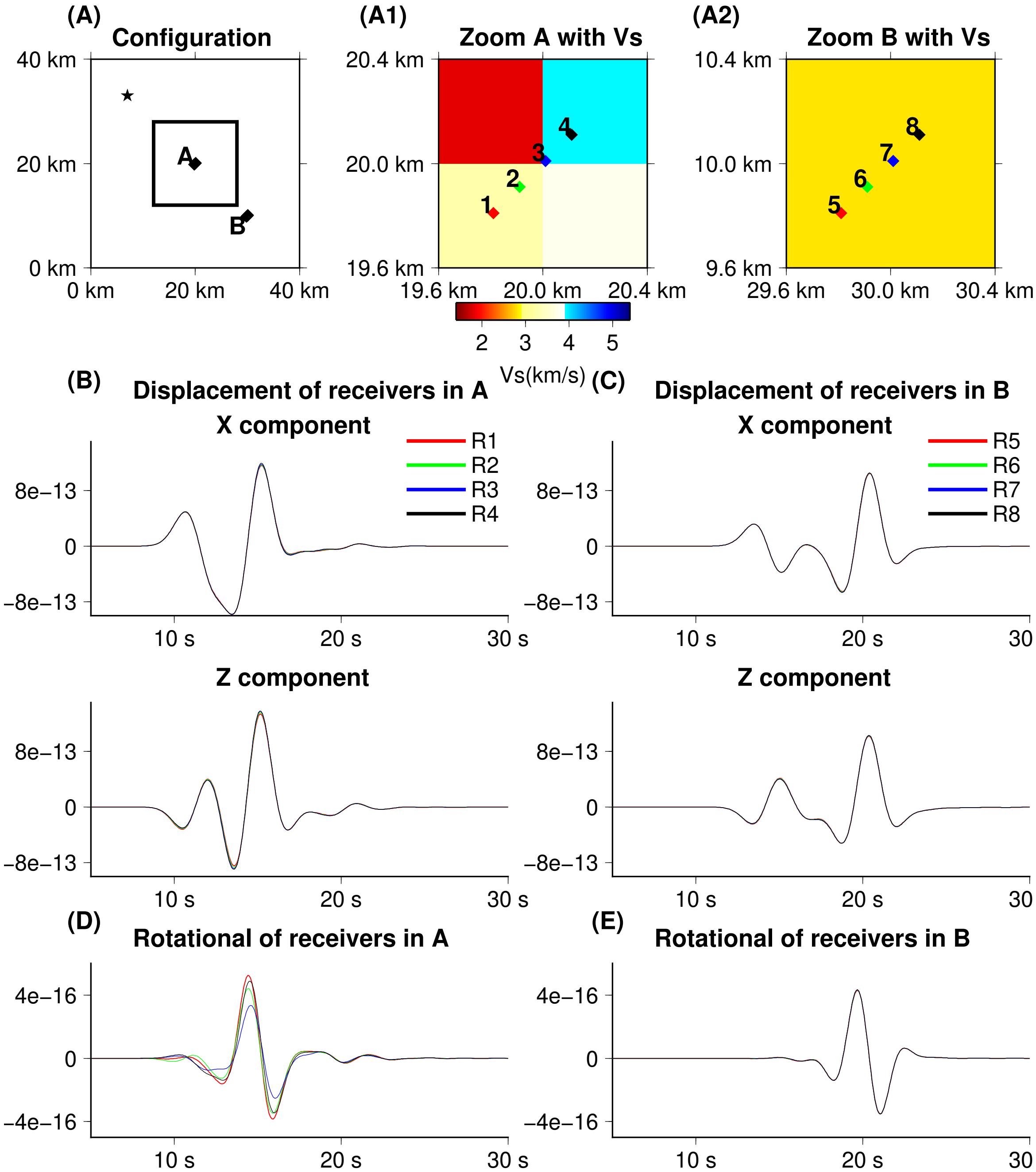


Fig. 4.5 Numerically derived example of the impact of small-scale heterogeneities on displacement and rotational waveforms. (A) Source–receiver positions. Close-up views of the location (A1) A and (A2) B receiver positions, and their local V_S structures. Location A is within the heterogeneous area, whereas location B is outside of the heterogeneous area. Displacement waveforms for the (B) location A and (C) location B receivers. Rotational waveforms for the (D) location A and (E) location B receivers.

4.5.3 FWI results without correctors

Receivers outside of the heterogeneous area

Here we present displacement and rotational inversion results for receivers positioned outside the heterogeneous area ("configuration A" in Fig. 4.4).

The displacement-based inversion results, which are presented in Fig. 4.6, show that the misfit decreases steadily throughout all of the Gauss–Newtown iterations and satisfies the convergence criteria at iteration 19. The strong fit between the target and synthetic data indicates that the final model can adequately explain the target data. We have used $\varepsilon_0 = 1.0$ to compute the homogenised models, which is a good compromise between resolution and accuracy, as we will discuss later. The residual is calculated as the deviation from a background V_S value of 2.8 km/s. A comparison of the modelled V_S^* structures shows that the final model exhibits a broadly similar pattern to that in the target model, with the exception of some high-velocity anomalies in the target model that are not well-recovered in our final inverted model.

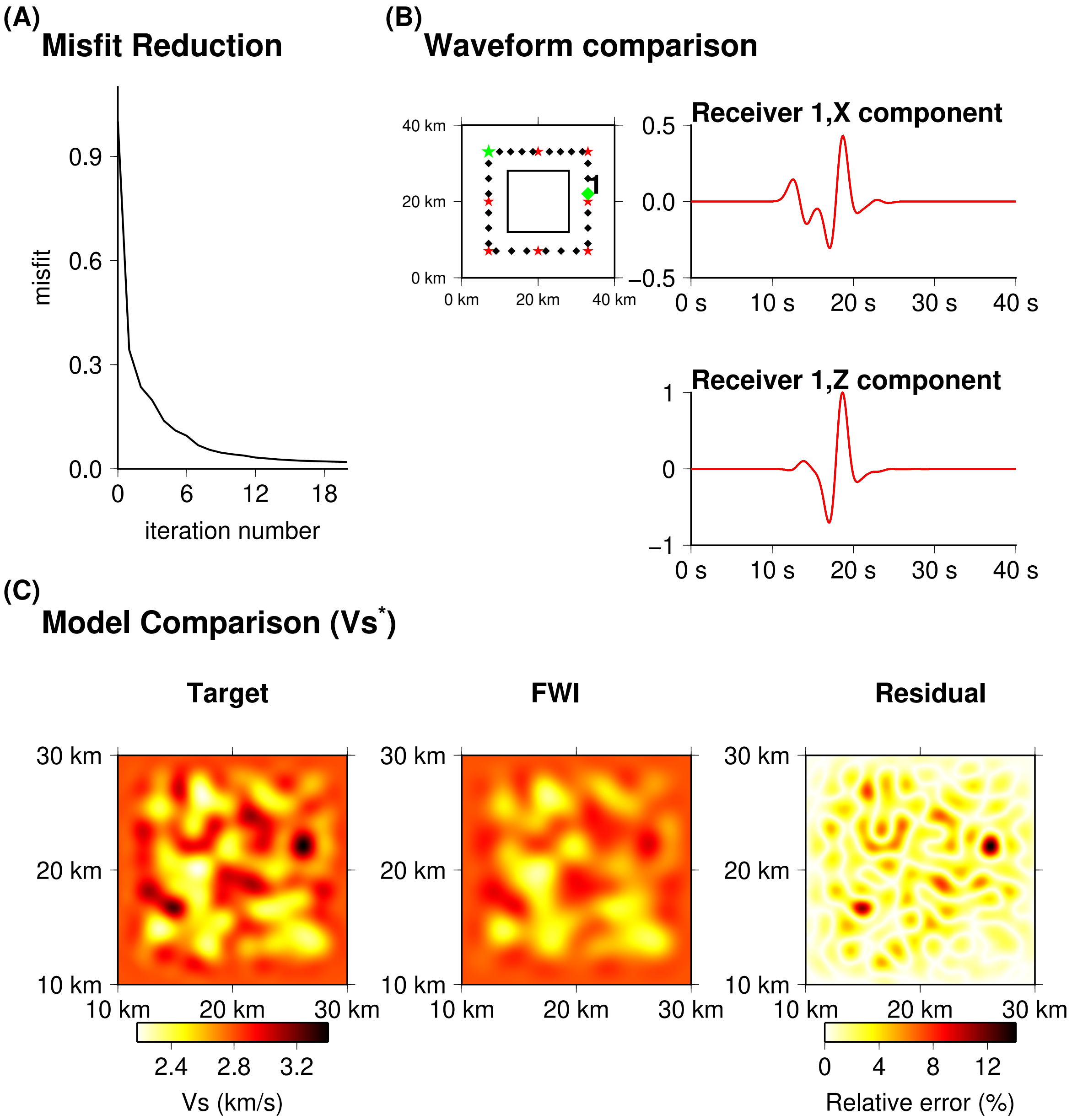


Fig. 4.6 Displacement-based FWI results using receivers positioned outside of the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamond) positions indicating the source–receiver pair that was used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers. (C) Target and final V_S^* models and their corresponding residuals. Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from a background value of 2.8 km/s.

The rotational inversion results are shown in Fig. 4.7. The chosen receiver configuration makes it possible to conduct a rotational-based FWI without any expected small-scale heterogeneity effects because the rotational waveforms are only influenced by small-scale heterogeneities that are in close proximity to the receiver locations, as explained in equation (4.28) and demonstrated in Fig. 4.5. The presented waveform comparisons show a strong fit between the waveforms for the target and final models, which indicates that the rotational- and displacement-based FWI models are comparable in their abilities to explain the target data. The V_S residuals between the target and final models for the rotational waveform case are also comparable to those for the displacement case.

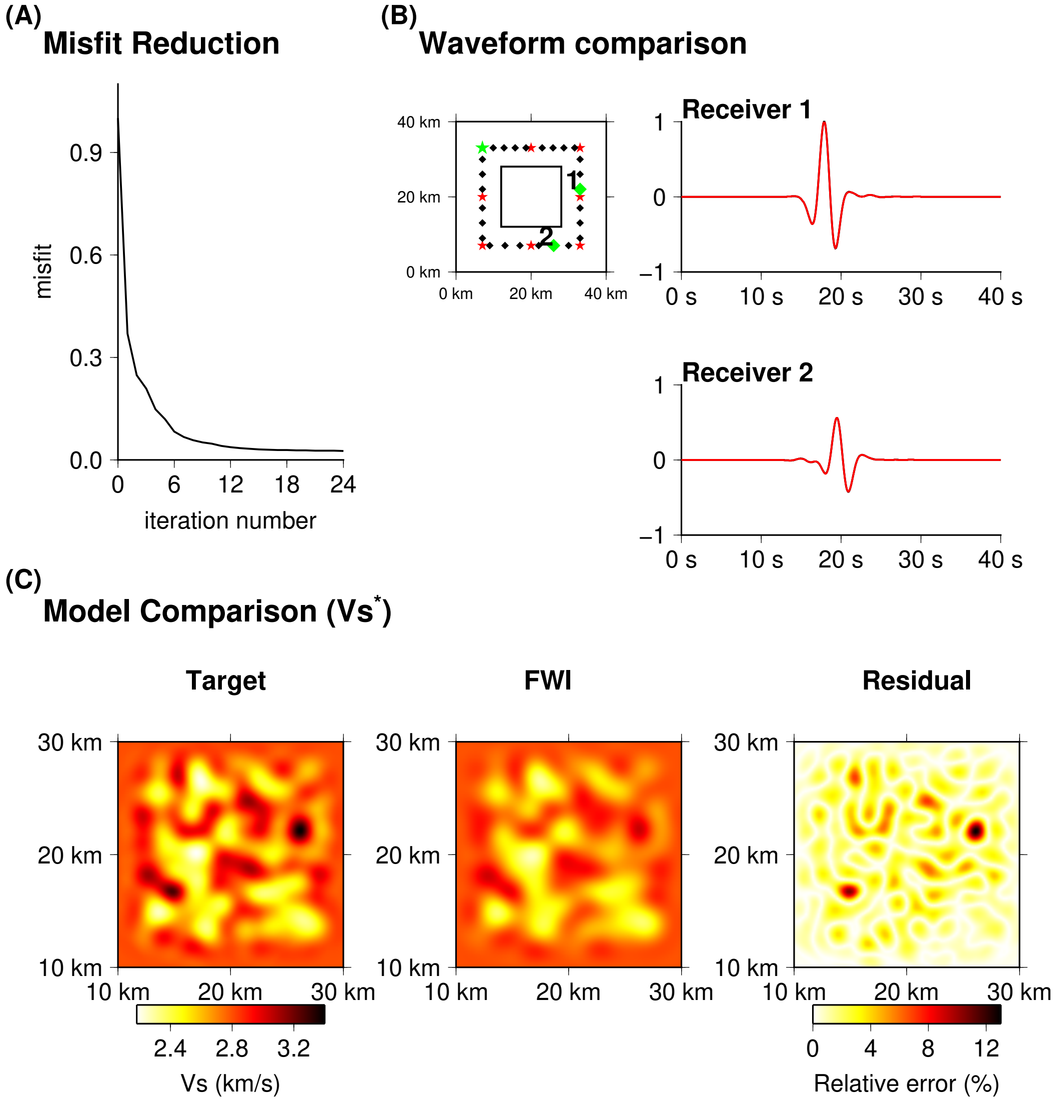


Fig. 4.7 Rotational-based FWI results using receivers positioned outside of the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions indicating the source–receiver pairs that were used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers. (C) Target and final V_S^* models and their corresponding residuals. Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from a background value of 2.8 km/s.

Our assessment of the optimal ε_0 value, which is based on equations (4.46) and (4.46), is shown in Fig. 4.8. We used the mean model from 200 samples, as opposed to a single solution, for the model error $\bar{E}_m$, which was computed for models that used ε_0 in the 0.3—2.9 range. The model error $\bar{E}_m$ increases exponentially as ε_0 decreases, with the transition to a rapid increase in $\bar{E}_m$ occurring around ε_0 = 1.0 for both the displacement- and rotation-based FWI. This result suggests that the displacement- and rotation-based FWI have almost the same resolution limit. We therefore select ε_0 = 1.0 as the optimal choice that provides the best overall resolution and accuracy of the final model results.

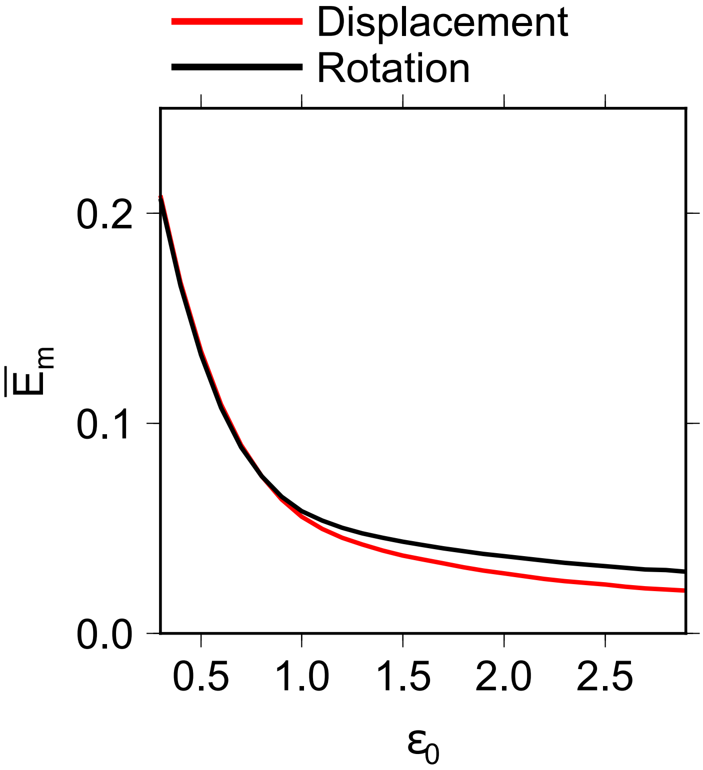


Fig. 4.8 Error $\bar{E}_m$ as a function of ε_0 for V_S^*, which is derived using displacement (red) and rotational (black) receivers positioned outside of the heterogeneous area.

Our statistical comparisons between the displacement and rotational cases, which consist of the error of mean values, standard deviation, and resolution between the target and final models, are based on ε_0 = 1.0 and shown in Fig. 4.9. The resolutions were computed by selecting a single grid point $\mathbf{y}_i$ and computing the correlations against all of the $\mathbf{x}$ points based on equation (4.49). Strong correlations (approaching unity) indicate that these points cannot be determined separately; therefore, the areas with strong correlations correspond to the spatial resolution of the final model. All of the statistical quantities exhibit comparable V_S^* trends between the displacement and rotational cases.

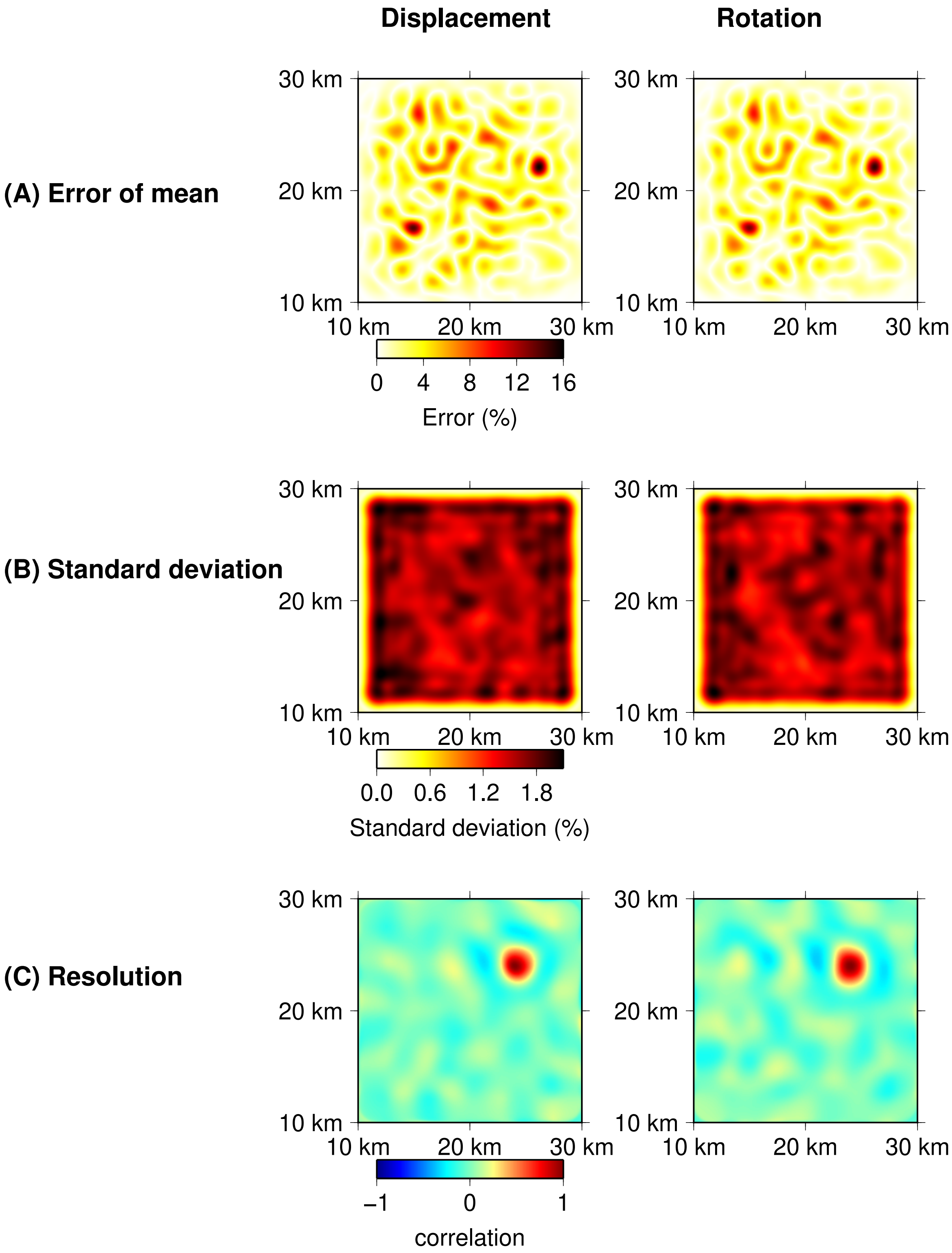


Fig. 4.9 Statistics of the displacement- and rotational-based FWI results using receivers positioned outside of the heterogeneous area. (A) Error of the mean models of V_S^*. (B) Standard deviation of V_S^*. (C) Resolution of V_S^*. The error and standard deviation values are presented as fluctuations from a background value of 2.8 km/s.

Receivers within the heterogeneous area

Here we present the displacement- and rotational-based inversion results using receivers positioned within the heterogeneous area ("configuration B" in Fig. 4.4). The forward modelling test presented in section 4.5.2 indicates that it is necessary to invert for the corrector to obtain good results. Nevertheless, we ignore this fact here and perform the displacement- and rotational-based FWI without correctors. The displacement-based inversion results, which are presented in Fig. 4.10, show that the misfit steadily decreased throughout all of the iterations and satisfied the convergence criteria at iteration 11. A strong fit between the target and synthetic data is confirmed, thereby indicating that the final model can explain the target data. Our model comparisons reveal that the errors between the target and final models for the receivers within the heterogeneous area are slightly better than the model comparisons when all of the receivers are positioned outside of the heterogeneous area. The nearby heterogeneities have a small effect on the displacement measurements, and subsequently the FWI process, as theoretically expected (equation 4.26).

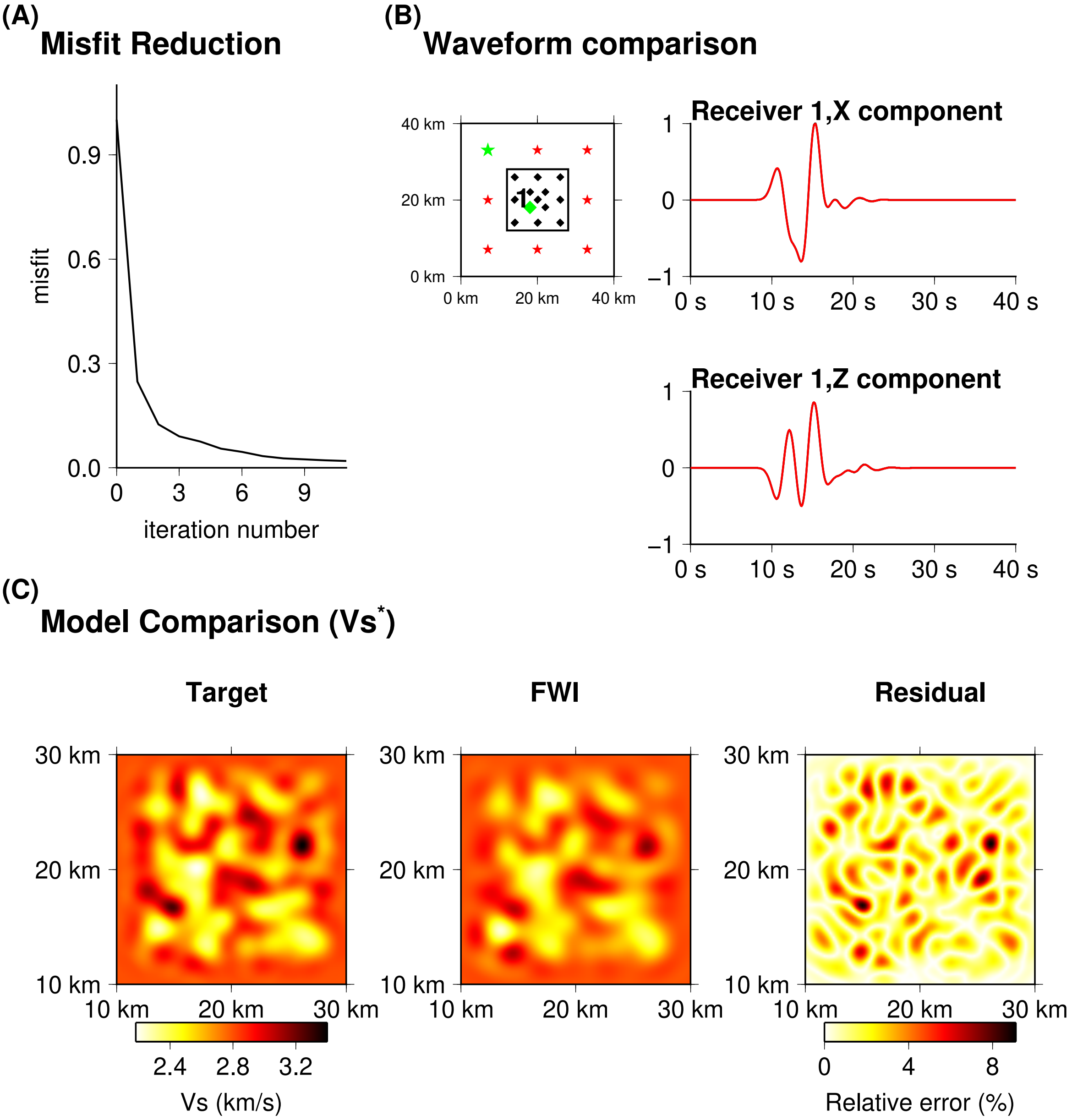


Fig. 4.10 Displacement-based FWI results using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamond) positions indicating the source–receiver pair that was used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers. (C) Target and final V_S^* models and their corresponding residuals. Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from a background value of 2.8 km/s.

The rotational-based inversion results are presented in Fig.4.11. The misfit reductions satisfied the convergence criteria at iteration 46, despite a temporary increase in the misfit at iteration 13. The comparison between the target and synthetic models suggests that the final model is unable to fully explain the target model, which indicates that the inversion did not converge to the best solution in this case. Furthermore, the V_S residuals in the rotational case are larger than those in the displacement case, despite having the same source and receiver configurations. We did not observe significant misfits in the waveforms and model parameters when the receivers were positioned outside of the heterogeneous area. The larger residuals between the target and final models and larger misfits between the waveforms are likely associated with the presence of small-scale heterogeneities in close proximity to the receiver locations, as theoretically expected.

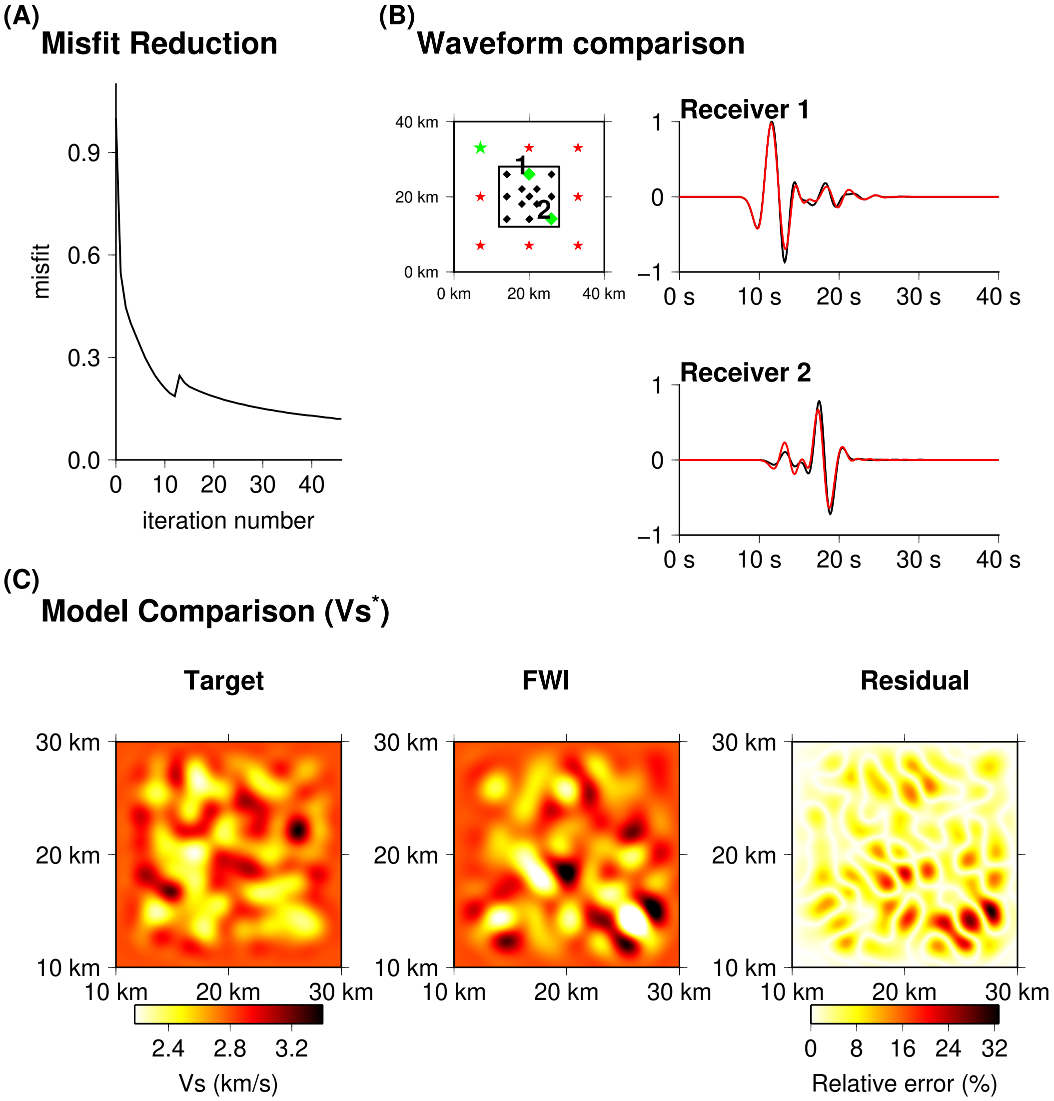


Fig. 4.11 Rotational-based FWI results, without correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions indicating the source–receiver pairs that were used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers. (C) Target and final V_S^* models, and their corresponding residuals. Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from a background value of 2.8 km/s.

The trade-off between the error of the mean model $\bar{E}_m$ and ε_0 are shown in Fig.4.12. The overall trend of $\bar{E}_m$ is similar to that when the receivers are placed within the heterogeneous area. However, the $\bar{E}_m$ results show that errors from the rotational-based FWI are consistently larger than those from the displacement-based FWI, regardless of the ε_0 value. The effects of small-scale heterogeneities can be clearly seen, even if the scale is limited to very large scales.

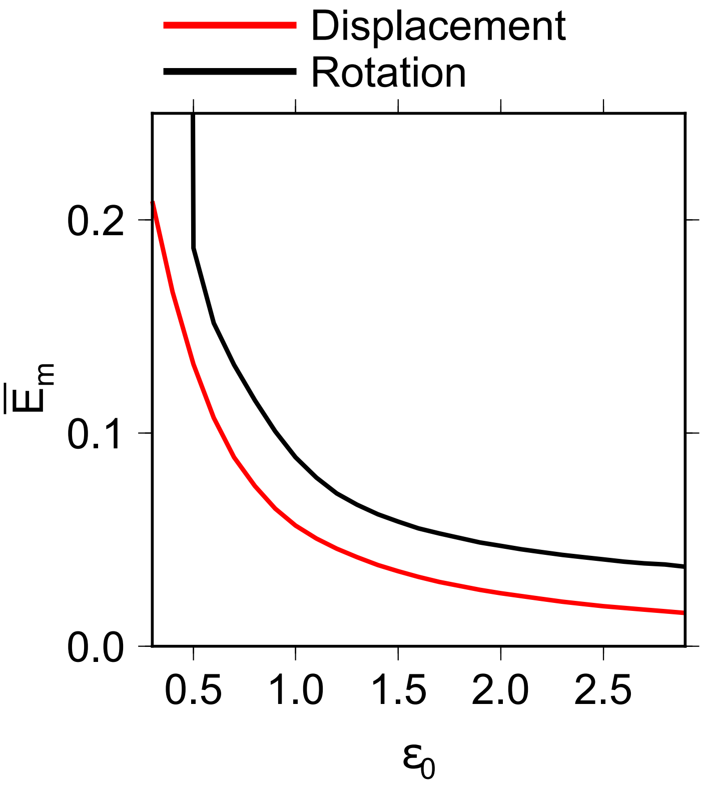


Fig. 4.12 Error $\bar{E}_m$ as a function of ε_0 for V_S^*, which is derived from the FWI results, without correctors, using displacement (red) and rotational (black) receivers positioned within the heterogeneous area.

Fig.4.13 presents our statistical comparison of the displacement- and rotational-based FWI results for $\varepsilon_0 = 1.0$. The error of the mean value of V_S for the rotational case indicates that the inversion could not determine the model parameters as accurately as in the other cases. The deviations and resolutions also show peculiar results for the rotational case. Unlike the displacement case, the deviations for the rotational case are not uniform, and the largest anomalies are distributed. Furthermore, the spatial resolutions for the rotational case are no longer valid, as there are strong correlations at distant locations.

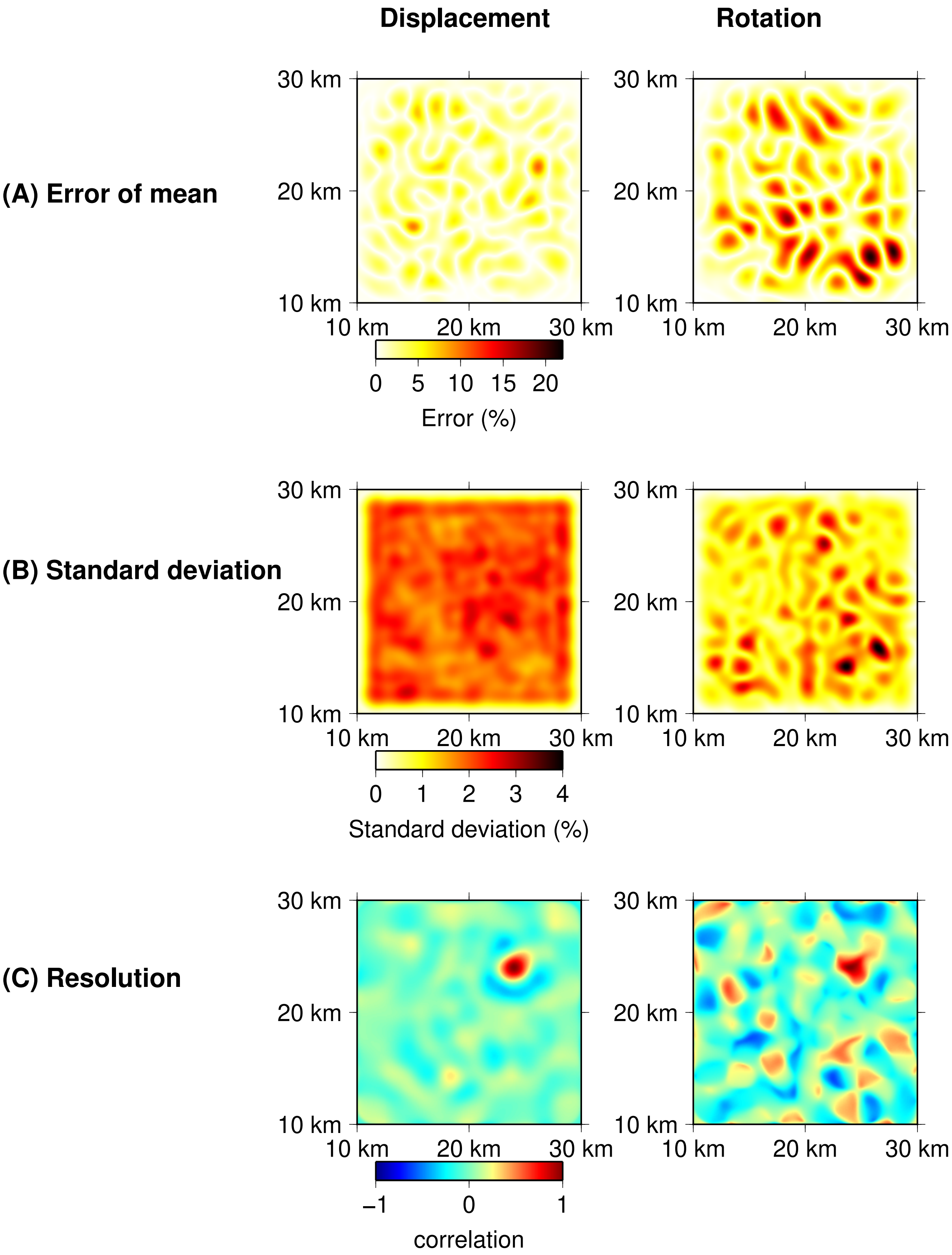


Fig. 4.13 Statistics of the displacement- and rotational-based FWI results, without correctors, using receivers positioned within the heterogeneous area. (A) Error of the mean models of V_S^*. (B) Standard deviation of V_S^*. (C) Resolution of V_S^*. The error and standard deviation values are presented as fluctuations from a background value of 2.8 km/s.

4.5.4 FWI results with correctors

Here we present the rotational-based FWI results, with correctors, for receivers located within the heterogeneous area ("configuration B" in Figure 4.4). The correctors are obtained during the inversion based on equation (4.41). Incorporating correctors in the rotational-based FWI procedure is theoretically expected to enhance the results compared with the FWI procedure without correctors.

The inversion results, which are presented in Fig.4.14, show that the misfit decreased in all of the iterations and met the convergence criteria at iteration 12. The strong fit between the waveforms for the target and synthetic models indicates that the final model with correctors can adequately explain the target data, whereas the final model without correctors is unable to achieve such a fit. Our model comparisons yield smaller residuals between the models with correctors than those without, thereby suggesting that including correctors led to improvements in the final model.

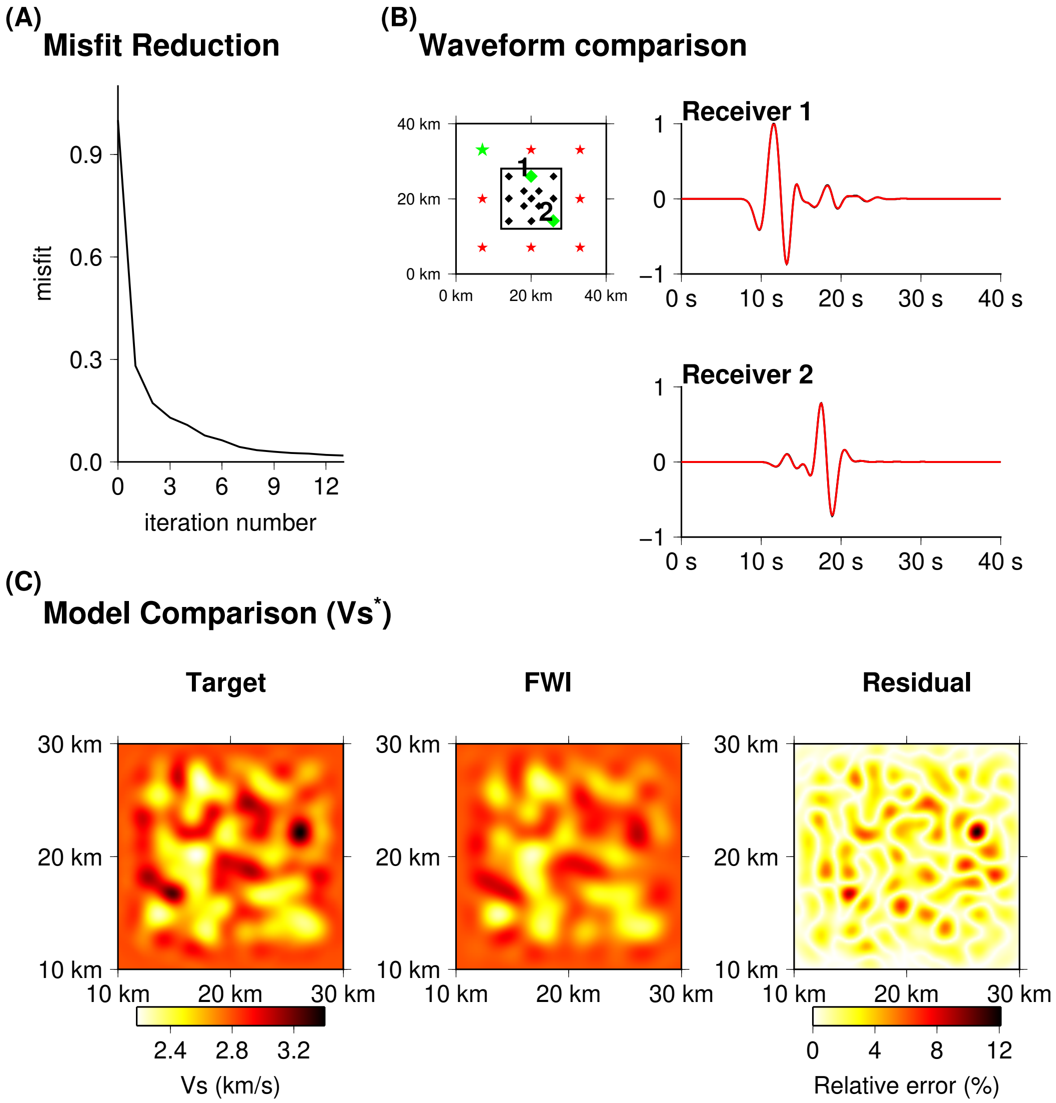


Fig. 4.14 Rotational-based FWI results, with correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions indicating the source–receiver pairs that were used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers. (C) Target and final V_S^* models, and their corresponding residuals. Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from a background value of 2.8 km/s.

Fig.4.15 shows relationship between the error of the mean model $\bar{E}_m$ and ε_0. A larger error $\bar{E}_m$ than that for the displacement case was observed when correctors were not included in the rotational-based FWI (Fig.4.12). However, the rotational-based FWI with correctors possesses an error that is almost the same as that for the displacement case. The resolution limits for both the rotational and displacement cases can be chosen to be 1.0 based on the $\bar{E}_m$ results, which indicates that there is no difference in the resolution limit between the rotational and displacement cases.

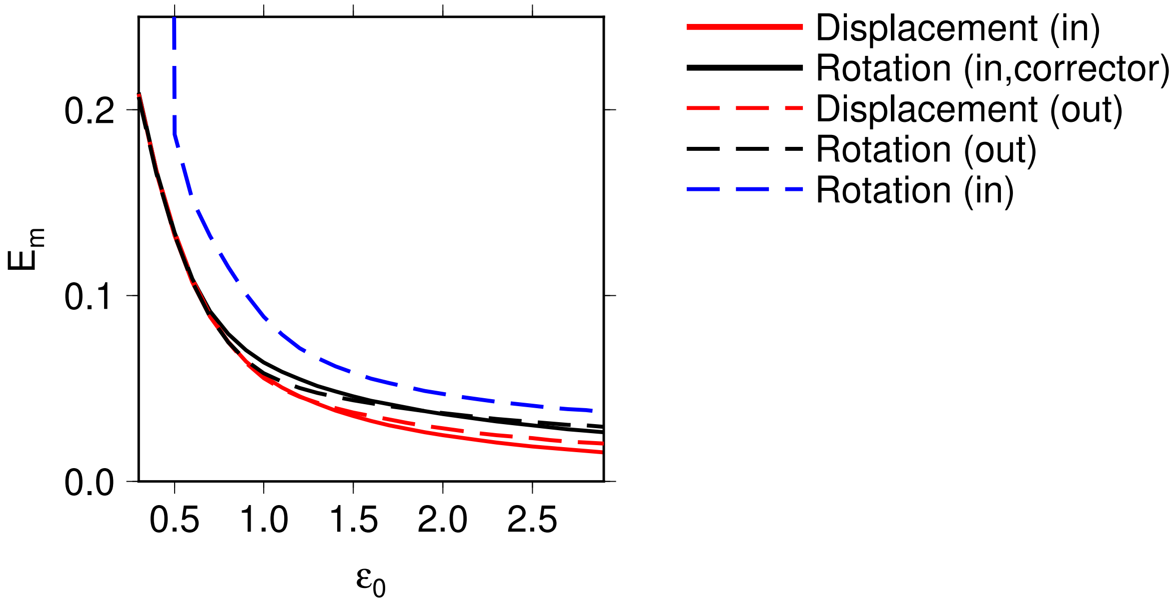


Fig. 4.15 Error $\bar{E}_m$ as a function of ε_0 for V_S^*, which is derived from the FWI results, with correctors, using displacement (solid red) and rotational (solid black) receivers positioned within the heterogeneous area. Results from the other cases (dashed curves) are also shown for comparison.

Fig.4.16 presents the statistical results based on $\varepsilon_0 = 1$. As discussed in the previous section, the FWI without correctors resulted in large errors in the mean values and peculiar outcomes for the deviations and resolutions. However, we obtained similar statistical results to those in the case for receivers positioned outside of the heterogeneous area when correctors were included in the inversion process. This indicates that the inclusion of correctors in the FWI procedure mitigates the effects of small-scale heterogeneities.

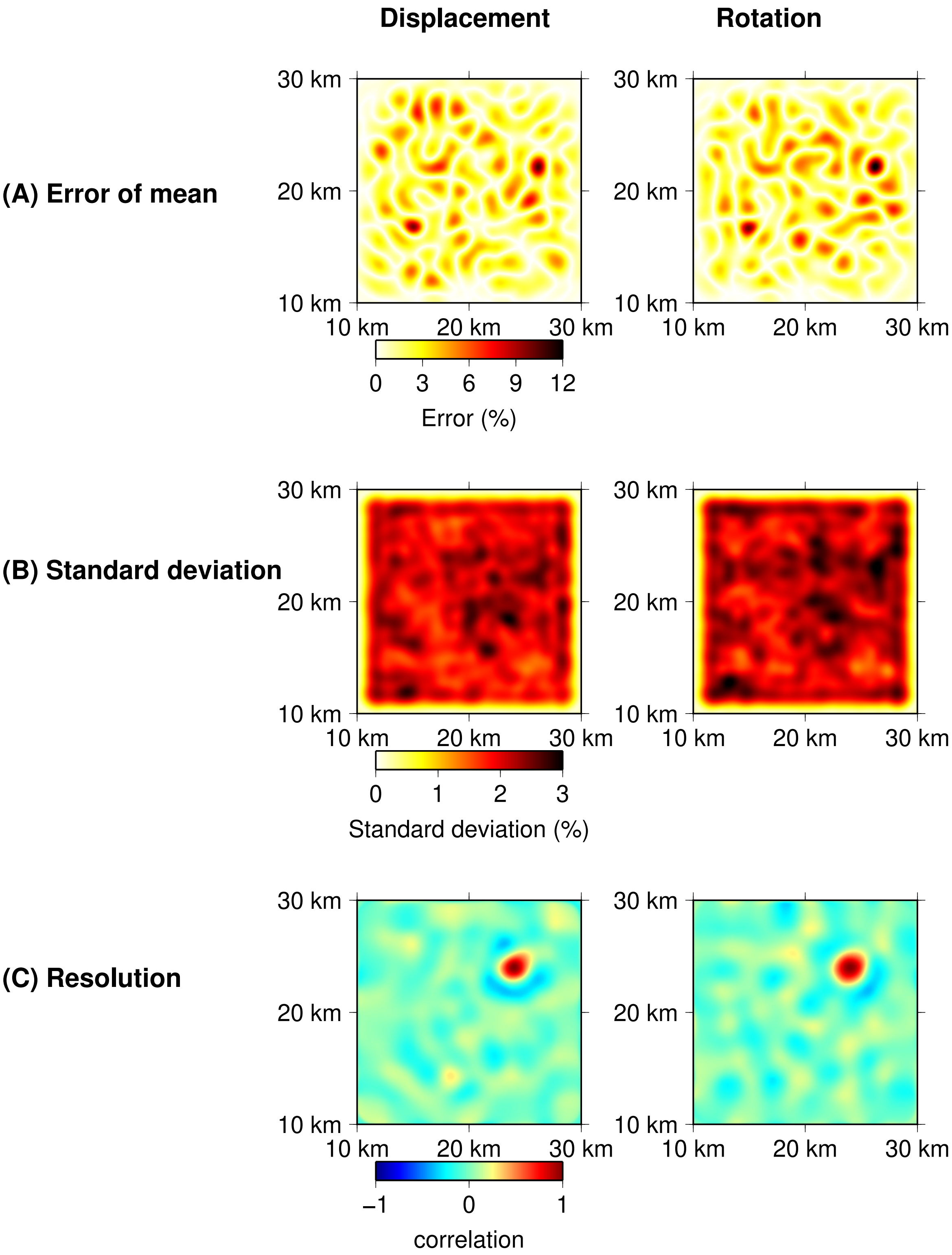


Fig. 4.16 Statistics of the displacement- and rotational-based FWI results, with correctors, using receivers positioned within the heterogeneous area. (A) Error of the mean models of V_S^*. (B) Standard deviation of V_S^*. (C) Resolution of V_S^*. The error and standard deviation values are presented as fluctuations from a background value of 2.8 km/s.

Table 4.1 Model misfits, which are based on equation (4.46), for the numerical FWI results using displacement and rotational data.

Case	$\bar{E}_m(\varepsilon_0 = 1)$
FWI without correctors	
Displacement and receivers outside of the heterogeneous area	0.041
Rotational and receivers outside of the heterogeneous area	0.045
Displacement and receivers within the heterogeneous area	0.042
Rotational and receivers within the heterogeneous area	0.095
FWI with correctors	
Rotational and receivers within the heterogeneous area	0.049

Table 4.1 provides a summary of the misfits, which are defined by equation (4.46) for a single model, in each case. Here we use single solution, as opposed to the mean model, to compute the model error $\bar{E}_m$. The misfits for the FWI without correctors indicate that the misfit is largest when rotational data are used and all of the receivers are located within the heterogeneous area. However, incorporating correctors in the inversion process significantly reduces the misfit, with a reduction from 0.095 (without correctors) to 0.049 (with correctors) observed. The misfit of 4.86 for the rotational-based FWI with correctors is comparable to the FWI results when receivers are placed outside of the heterogeneous area.

We present comprehensive inversion and statistical results, including the V_P^*, density* and anisotropy* results for three cases (FWI using displacement receivers positioned within the heterogeneous area; FWI, without correctors, using rotational receivers positioned within the heterogeneous area; and FWI, with correctors, using rotational receivers positioned within the heterogeneous area), in the appendix.

A short investigation of additional FWI results using strain receivers is also presented in appendix. The conclusion of the strain-based FWI results is same as that for the rotational case: the FWI does not converge to the best model when the effects of small-scale heterogeneities are present. It is therefore necessary to include correctors in strain-based FWI to effectively mitigate any small-scale effects.

4.6 Joint displacement- and rotational-based FWI

Gradient sensors are often collocated with seismometers for more comprehensive field studies. Here we perform a numerical experiment where both displacement and rotational sensors are used. We used the same receiver and source configurations as in the previous section, whereby the receivers are positioned within the heterogeneous area, and assume that both displacement and rotational data are recorded at each receiver location. We therefore incorporated data

Table 4.2 Model misfit based on equation (4.46) for the FWI results of the joint displacement and rotational data experiments.

Case	$\bar{E}_m(\varepsilon_0 = 1)$
FWI without correctors	
Joint displacement and rotational data	0.066
FWI with correctors	
Joint displacement and rotational data	0.038

from 52 sensors (26 rotational and 26 displacement sensors) for the inversion. We utilised a least-squares misfit function by combining equations (4.31) and (4.38) in this numerical experiment. A data covariance matrix that is different from the identity matrix is used to balance the significant amplitude differences between the displacement and rotational data to ensure that both data types are equally weighted in the inversion (Tarantola & Valette 1982). The results are evaluated using the model misfit (equation (4.46)), which are based on $\varepsilon_0 = 1$, and presented in table 4.2. The results show that the misfits in the model parameters are minimised when correctors are incorporated. These misfits are even smaller than those in the case when only displacement data are used; this is likely related to the inclusion of additional data in the joint displacement and rotational data case. This numerical experiment indicates that the inclusion of correctors in the inversion process is necessary, even when translational motions are used with wavefield gradients.

4.7 Using correctors to observe sub-wavelength heterogeneities

We have demonstrated that the rotational-based FWI is hampered by the sensitivity of the rotational wavefield to small-scale structures rather than improving the spatial resolution. Although incorporating correctors in the FWI process can help mitigate the effects of small-scale heterogeneities, it can only produce tomographic images that are at the same resolution as displacement-based FWI results, as shown in the statistical evaluations. Therefore, rotational-based FWI can only estimate the effective mechanical properties at best, and are therefore unable to capture fine-scale heterogeneities in the tomographic images, despite the sensitivity of the rotational wavefield to small-scale heterogeneities. However, the inverted correctors in our new FWI framework may contain information about any small-scale structures that may be present. Here we demonstrate how small-scale structures can be observed in the inverted correctors.

We randomly place 500 rotational receivers within a 1.8 × 1.8 km area (Fig. 4.17) for this numerical test; the absolute values of the inverted correctors at the final iteration for each component are illustrated in Fig. 4.17. The diagonal components of the correctors appear to only possess high values in the vicinity of the vertices of the small-scale anomalies, whereas the off-diagonal components of the correctors more clearly show the defined boundaries of the small-scale anomalies. Therefore, the inverted correctors in our new FWI framework can provide indications of the presence of strong small-scale discontinuities in the subsurface that are difficult to image using conventional seismic tomography methods.

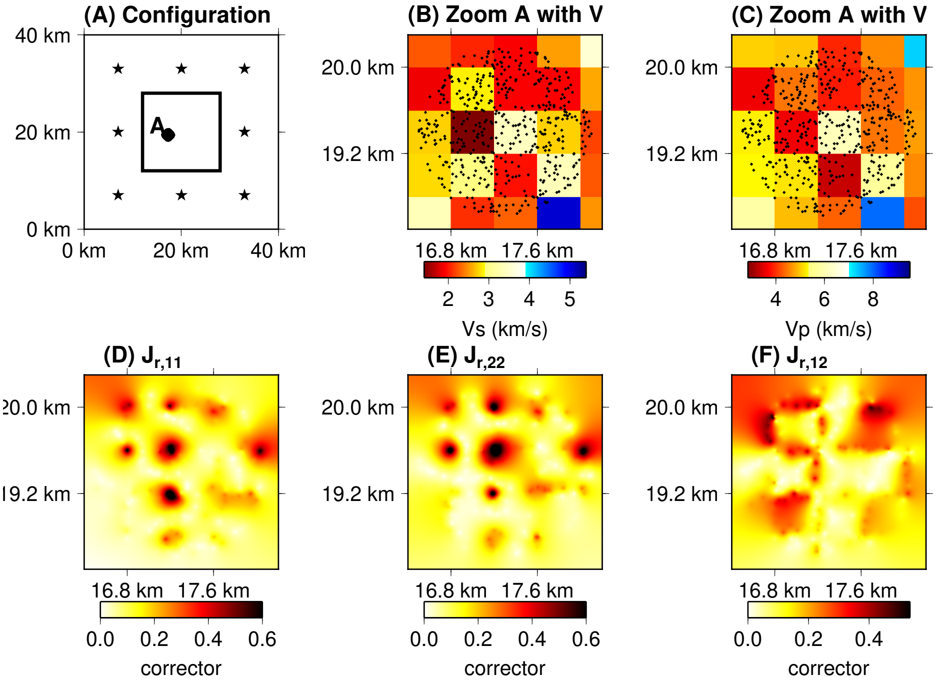


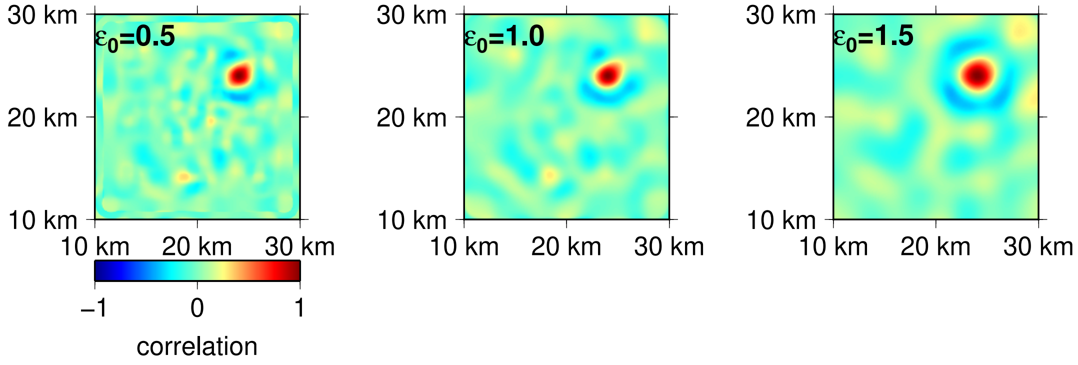
Fig. 4.17 Numerical demonstration to show indications of small-scale structures via the inverted correctors. (A) Source (black stars) and receiver (black diamonds) configuration for the numerical experiment. Five-hundred rotational receivers are positioned at location A. (B) V_S and (C) V_P heterogeneities around location A. Black dots represent the receiver positions. Inverted diagonal components, (D) $J_{r,11}$ and (E) $J_{r,22}$. (F) Inverted off-diagonal component, $J_{r,12}$.

4.8 Discussions and conclusions

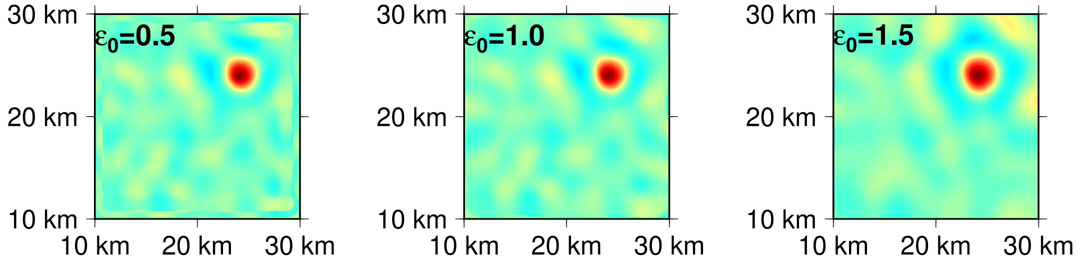
We have first shown the necessary difference in model parametrisation between displacement- and wavefield gradient-based FWI. The leading-order approximation in the two-scale homogenisation mathematical framework highlights that the displacement-based FWI only requires effective elastic mechanical properties to explain the data, whereas the wavefield gradient-based FWI also requires the corrector in the model space to explain the data. Current FWI applications implicitly rely on the FWI solution space being within the $\mathcal{M}^*$ subspace, such that the solution space is only represented by the effective elastic mechanical properties. However, FWI applications that are based on wavefield gradients, such as rotational and strain wavefields, require the solution space to extend to the entire $\mathcal{M}^*$, thereby including the effective elastic mechanical properties and correctors, as theoretically expected. Our numerical experiments have confirmed that this extension of the solution space is necessary, as the effective true model was only retrieved in the wavefield gradient-based FWI when correctors were included.

Another important result is that the resolution limit is close to $\varepsilon_0 = 1$ for both the displacement- and rotational-based FWI, such that there is no difference in terms of resolution. The model error exponentially increases as ε_0 decreases, with a rapid increase observed when $\varepsilon_0 < 1$. The resolution images that were obtained for different ε_0 values are also helpful for assessing the resolution limit. Fig.4.18 shows the resolution images for all of the cases and $\varepsilon_0 = 0.5\,, 1\,, 1.5$. The size of the area with a strong correlation is almost same between $\varepsilon_0 = 0.5$ and $\varepsilon_0 = 1$ for all of the cases, and then increases when $\varepsilon_0 = 1.5$. We therefore suggest that the resolution limit is around $\varepsilon_0 = 1$ for both the displacement- and rotational-based FWI. Finally, we note that $\varepsilon_0 = 1$ is generally not a valid value. Other configurations, such as either more receivers or more scattering, may allow this value to be lowered to around $\varepsilon_0 = 0.5$. $\varepsilon_0 = 0.5$ is sufficient to model data in most geological media, such that it is unnecessary to lower ε_0 below this limit.

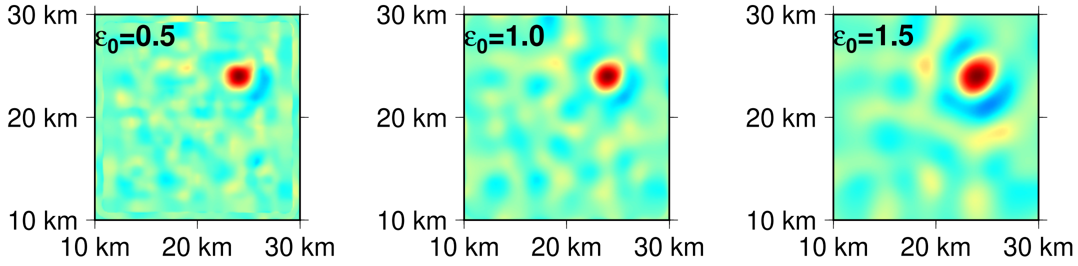
(A) Displacement: receivers inside heterogeneous area



(B) Displacement: receivers outside heterogeneous area



(C) Rotation: receivers inside heterogeneous area (corrector)



(D) Rotation: receivers outside heterogeneous area

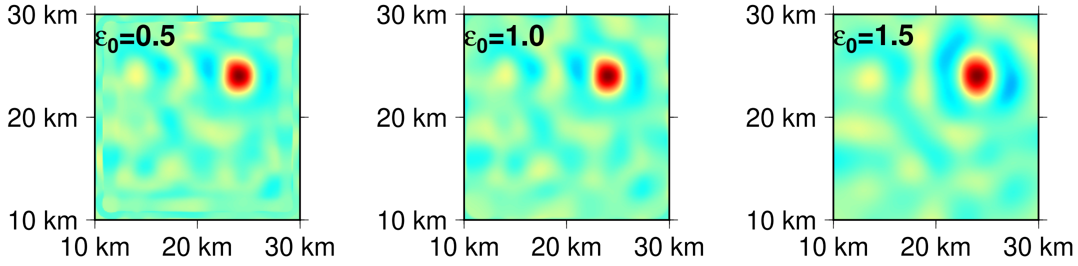


Fig. 4.18 Resolution images of V_S^* for different ε_0. The resolutions for $\varepsilon_0 = 0.5$, $\varepsilon_0 = 1.0$ and $\varepsilon_0 = 1.5$ are shown. The presented results are for displacement receivers positioned (A) within and (B) outside of the heterogeneous area, and rotational receivers positioned (C) within and (D) outside of the heterogeneous area. We used correctors in the FWI process for (C).

We have found no advantage of using wavefield gradients instead of displacement data during the FWI process in terms of accuracy and resolution. We note that this conclusion is only true for excellent data coverage. We have not investigated cases with partial data coverage; therefore, it is still possible that gradient data could make a significant difference in such contexts.

We have shown that the inverted correctors can serve as indicators of small-scale heterogeneities, which can lead to high-resolution mapping of complex geological structures, as long as such features are in the vicinity of the receivers. Therefore, we may be able to use arrays of many receivers to locate abrupt subsurface discontinuities, such as faults, if they are located near the sensors. DAS cables may be a viable option since they can easily obtain a significant number of very short interval strain data, which can assist in imaging small, abrupt, subsurface discontinuities via the corrector.

It is important to note that we have only inverted part of the correctors in this study. Indeed, the parametrisations of $\mathcal{M}^{*h}$ that we have been using contain small-scale heterogeneities (continuity between elements is not enforced). Therefore, the correctors shown in Fig. 4.17 cannot be directly compared with the corrector that would be obtained from the direct homogenisation problem. Such a comparison requires the correctors that are derived from the inversion process to be combined with those obtained from applying the homogenisation operator $\mathcal{H}$ to the inverse solution.

Although the FWI with correctors can mitigate the effects of small-scale heterogeneities, it cannot completely eliminate these effects. Our approach assumes that evaluating up to the first-order term in equation (4.3) is sufficient, such that the effect of the higher-order terms are negligible and can be ignored. However, it should be noted that the effects of these higher-order terms become more pronounced as the elastic contrasts increase. Therefore, our method is still considered to be affected by small-scale heterogeneities to some extent. Nevertheless, all of the synthetic tests in this study show that a better solution is obtained using correctors.

This study demonstrates that there is no intrinsic gain in using gradient sensors instead of conventional seismometers for FWI studies, as rotational-based FWI results can only recover information of the same quality as translation-based FWI results if receiver-coupling correctors are included in the inversion process. Nevertheless, these correctors contain sub-wavelength heterogeneities that can potentially be exploited to obtain a resolution at the sensor array scale.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science (JSPS) through a Grant-in-Aid for JSPS Research Fellowship (22KJ2397) and JSPS Overseas Challenge Program for Young Researchers. This work was also supported by JSPS KAKENHI (Grant Numbers JP20H01997). This work was completed in part with resources provided by the GLiCID (www.glicid.fr) Computing Center. Yann Capdeville's work was partly supported by the ANR ASIRISplus-SDS grant ANR-19-CE22-0015. This work has benefited from discussions within the SPIN-ITN European project.

Chapter 5

Detecting small-scale subsurface anomaly using DAS data

5.1 Introduction

Seismic tomography stands as a crucial tool for delving into subsurface structures across various scales, ranging from local to global. This method has been instrumental in furnishing substantial insights into Earth's structures (Dziewonski & Anderson 1981; Tape *et al.* 2010; Kamei & Pratt 2013). In its early stages, seismic tomography primarily enabled the estimation of large-scale subsurface models, often resulting in homogeneous models (Aki & Lee 1976; Dziewonski 1984). Over time, significant efforts have been undertaken to enhance the spatial resolution of subsurface structure imaging. These endeavors involve the expansion of permanent seismic station networks (Romanowic & Giardini 2001), advancements in theoretical and methodological approaches (Tarantola 1984; Tarantola 2005; Fichtner *et al.* 2009a), and the rapid growth of computational power. However, the spatial resolution of seismic tomography inherently constrained by around half the minimum wavelength in seismic data (Capdeville & Métivier 2018). Due to the finite-frequency of seismic sources and computational constraints, the seismic data is band-limited, consequently, the limitations of spatial resolution in seismic tomography are beyond our immediate control. In this study, we explore the potential of imaging subsurface obstacles whose size is much smaller than the half minimum wavelength using new emerging seismic sensors, distributed acoustic sensing (DAS).

DAS is a novel technique that allows the measurement of strain using only a single cable (Grattan & Sun 2000). The interrogator, positioned at one end of the DAS cable, directs a laser pulse into the cable and measures the backscattered photons reflected return to the interrogator.

Subsequently, the interrogator measures the phase shift between the backscattered photons reflected at two points. Because this phase shift is quasi-linearly proportional to the change in strain (Grattan & Sun 2000), one DAS cable works as multiple strain sensors. The straight line distance between two points of origins of backscattered photons is called gauge length which determines the spatial resolution of DAS measurements. The gauge length depends on the application ranging from a few tens of centimeters to a few tens of meters. The strain of DAS measurement can be well approximated by average strain rate (Wang *et al.* 2018; Daley *et al.* 2016; Hubbard *et al.* 2022), as described by:

$$\epsilon_{\mathrm{DAS}}(\mathbf{x}, t) = \frac{1}{g} \int_{-g/2}^{g/2} \boldsymbol{A}[\ \dot{\boldsymbol{\epsilon}}(\mathbf{u}(\mathbf{x}, t))]\ dl\,, \tag{5.1}$$

where ϵ_{DAS} is the strain rate of DAS, g is gauge length, and $\boldsymbol{A}$ is the operater to obtain the strain alog the DAS cable. The DAS strain rate is also often approximated to array derived strain rate (Wang *et al.* 2018):

$$\epsilon_{\mathrm{AD}}(\mathbf{x}, t) = \frac{v(\mathbf{x} + g\mathbf{l}, t) - v(\mathbf{x}, t)}{g}\,, \tag{5.2}$$

where $\mathbf{l}$ is the unit vector parallel to the DAS cable, and v is the particle velocity parallel to the DAS cable. The equation (5.1) shows the DAS measurement is equivalent to the averaged line integral of strain rate along the DAS cable, while the equation 5.2 is derived from the equation (5.1) and useful when the DAS measurement is compared with the geophone-array derived strain rate (Wang *et al.* 2018). The measurement points can be spatially dense, even a few tens of centimeters over a DAS cable of a few tens of kilometers, although spatial resolution is controlled by the gauge length.

In the geophysical study, the application of DAS was started from oil and gas industry for reservoir monitoring by employing DAS in the borehole (Daley *et al.* 2013). Soon, the potential of DAS in regional and global seismological studies has been highlited in recent years (Zhan 2020). Because of the dense spatial observation of strain wavefields by DAS, it simply enables us to conduct array analysis such as surface wave tomography (Yuan *et al.* 2020) and characterizing fault zone (Atterholt *et al.* 2022). Although DAS is definitely promissing tools to further understanding of seismological targets, we have to pay attention to the distinct nature of the strain rate from traditional seismic observations of translational motions such as displacement, velocity, and acceleration.

It has been recognized that wavefield gradients such as rotation and strain are distorted by localized small-scale structures (King & Bilham 1973; Harrison 1976; van Driel *et al.* 2012; Singh *et al.* 2020; Muir & Zhan 2022). Capdeville *et al.* (2020) numerically showed that localized jump in strain occurs at small-scale velocity anomaly but not in the displacement. Singh *et al.* (2020) and Muir & Zhan (2022) showed effects of small-scale structures on the real field data of rotational and DAS measurements. While this unique characteristic of wavefield gradients can present challenges for forward and inverse problems based on wavefield gradients, necessitating correction (Singh et al. 2020; Chapter 4), it also opens the door to the potential imaging of localized small-scale structures in the subsurface with unpreceedented spatial resolution.

In this study, we demonstrate the capacity to image small-scale anomalies using DAS observations through synthetic tests and real-field data, leveraging the principles of two-scale homogenization theory (Capdeville *et al.* 2020).

5.2 Theoretical insight of small-scale effects on DAS measurement

The localized small-scale effects on the DAS measurement can be explained using two-scale homogenization theory (Capdeville *et al.* 2010a; Capdeville *et al.* 2020). Based on the equations (4.27) and (5.1), we obtain:

$$\epsilon_{\mathrm{DAS}}(\mathbf{x},t) = \frac{1}{g}\int_{-g/2}^{g/2} \boldsymbol{A}[\,\dot{\boldsymbol{\epsilon}}(\mathbf{u}^{*}(\mathbf{x},t))]\,dl + \frac{1}{g}\int_{-g/2}^{g/2} \boldsymbol{A}[\,\boldsymbol{\epsilon}(\chi)(\mathbf{x},\mathbf{y}) : \dot{\boldsymbol{\epsilon}}(\mathbf{u}^{*})(\mathbf{x},t)]\,dl + O(\varepsilon_0)\,. \tag{5.3}$$

Knowing that only the corrector χ depends on small-scale structures, the second term represents the small-scale effects on the DAS measurements. For the sake of simplicity, we orientate the fiber optic cable parallel to the x-direction in cartesian coordinate system throughout this study. Furthermore, assuming that homogenized strain $\dot{\boldsymbol{\epsilon}}(\mathbf{u}^{*})$ is constant along the gauge length, we can simplify the equation (5.3) as:

$$\begin{aligned} \epsilon_{\mathrm{DAS}}(\mathbf{x},t) =& \dot{\epsilon}(\mathbf{u}^{*})(\mathbf{x},t)_{11} + G^{\mathrm{DAS}}(\mathbf{x})_{11kl}\dot{\epsilon}(\mathbf{u}^{*})(\mathbf{x},t)_{kl} + O(\varepsilon_0) \\ \text{where}\quad & G^{\mathrm{DAS}}(\mathbf{x})_{11kl} = \frac{1}{g}\int_{-g/2}^{g/2} \epsilon(\chi)(\mathbf{x},\mathbf{y})_{11kl}\big|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\,dx\,. \end{aligned} \tag{5.4}$$

Based on this equation, the small-scale effects can be seen as the coupling of averaged correctors over the gauge length and homogenized strain. The assumption that homogenized strain is constant over the gauge length would be valid when the minimum wavelength $\lambda_{\min}$ is $\lambda_{\min} \gg g$.

If we use the equation (5.2) instead of (5.1), it provides the different small-scale effects from what we see in the equation (5.4). Based on the equations (4.10) and (4.17), the particle velocity can be written as:

$$\mathbf{v}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) = \mathbf{v}^*(\mathbf{x},t) + \epsilon_0 \boldsymbol{\chi}(\mathbf{x},\mathbf{y}) : \dot{\boldsymbol{\epsilon}}(\mathbf{u}^*(\mathbf{x},t))\,. \tag{5.5}$$

Then, using first-order Taylor's approximation, the particle velocity at the location $(\mathbf{x}+g\mathbf{j},\mathbf{y})$ can be written as:

$$\mathbf{v}^{\varepsilon_0}(\mathbf{x}+g\mathbf{j},\mathbf{y},t) = \mathbf{v}^*(\mathbf{x},t) + g\nabla_{\mathbf{j}}\mathbf{v}^*(\mathbf{x},t) + \varepsilon_0 \boldsymbol{\chi}(\mathbf{x}+g\mathbf{j},\mathbf{y}) : \boldsymbol{\epsilon}(\mathbf{u}^*)(\mathbf{x},t)\,. \tag{5.6}$$

Subtracting (5.5) from (5.6), we obtain:

$$\begin{aligned}\frac{\mathbf{v}^{\varepsilon_0}(\mathbf{x}+g\mathbf{j},\mathbf{y},t) - \mathbf{v}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t)}{g} =& \nabla_{\mathbf{j}}\mathbf{v}^*(\mathbf{x},t) + \mathbf{G}^{AD}(\mathbf{x}) : \dot{\boldsymbol{\epsilon}}(\mathbf{u}^*)(\mathbf{x},t) + \varepsilon_0 \nabla_{\mathbf{j}}\boldsymbol{\chi}(\mathbf{x},\mathbf{y}) : \dot{\boldsymbol{\epsilon}}(\mathbf{u}^*)(\mathbf{x},t)\,,\\ &\text{where} \quad \mathbf{G}^{AD}(\mathbf{x}) = \frac{1}{g}(\boldsymbol{\chi}(\mathbf{x},\mathbf{y}+g\mathbf{j}/\varepsilon_0) - \boldsymbol{\chi}(\mathbf{x},\mathbf{y}))|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\,.\end{aligned} \tag{5.7}$$

Finally, assuming $\mathbf{j}$ is parallel to the x-direction, we obtain:

$$\begin{aligned}\epsilon_{\mathrm{AD}}(\mathbf{x},t) &= \nabla_1 v_1^*(\mathbf{x},t) + G^{\mathrm{AD}}(\mathbf{x})_{1kl}\dot{\boldsymbol{\epsilon}}(\mathbf{u}^*)(\mathbf{x},t)_{kl} + O(\varepsilon_0)\\ \text{where} \quad G^{\mathrm{AD}}(\mathbf{x})_{1kl} &= \frac{1}{g}(\chi(\mathbf{x},\mathbf{y}+g\mathbf{j}/\varepsilon_0)_{1kl} - \chi(\mathbf{x},\mathbf{y})_{1kl})|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\,.\end{aligned} \tag{5.8}$$

Therefore, the small-scale effects between (5.4) and (5.8) are different. Hence, we have to note this difference when we compare the DAS measurement with the array derived strain using geophones or when computing the DAS measurements by numerical simulation.

Since the corrector represents localized quantities at observation points (see chapter 4), it tends to approach zero in regions distant from specific small-scale structures. Consequently, the second term induces localized and abrupt distortions in strain waveforms at small-scale structures. In essence, the observational data from DAS contain valuable information about these small-scale structures, offering an opportunity to image them through the inversion of correctors.

5.3 Inversion of correctors

Here, we consider the inversion of correctors based on the misfit function:

$$E(\mathbf{G}_r^{\mathrm{DAS}}) = \sum_s \int_0^T (\epsilon_{s,\mathrm{DAS}}(\mathbf{x}_r,t) - \dot{\epsilon}_s(\mathbf{u}^*)(\mathbf{x}_r,t)_{11})^2 dt\,, \tag{5.9}$$

where T is the time length of the observation, r is the receier index, and s is the source index.

The corrector can be easily inverted if we have homogenized strain $\dot{\boldsymbol{\epsilon}}_s(\mathbf{u}^*)$. One strategy to invert correctors is computing homogenized strain by forward modeling, which is the same approach used in Singh et al. (2020) and Chapter 4. However, this approach is possible only if we have good priori knowledge about seismic sources and large-scale subsurface models. In this study, we propose a different approach to invert correctors.

We consider the separation of the DAS measurement $\dot{\epsilon}_{\text{DAS}}$ into two parts; homogenized strain rate and small-scale term including correctors without knowing seismic sources and subsurface models. The seismic data has a finite-frequency band, hence, the minimum wavelength λ_{min} exists in most cases. The minimum wavelength can be roughly estimated from the maximum frequency of the data and minimum seismic velocity. Knowing minimum seismic velocity may not be difficult, which could be done by simple analysis of DAS data such as surface wave analysis and analyzing travel times from the shot gather. If we assume the fiber optic cable is placed parallel to the propagation direction of seismic wave, the homogenized strain rate would consist of a lower wavenumber than $k_{\text{max}} = 1/\lambda_{\text{min}}$. While the small-scale effects on DAS measurements would consist of higher wavenumber than k_{max} because the corrector causes the abrupt distortions in DAS measurement. Hence, we may be able to distinguish between the first- and second-term of the equation (5.4) using the spatial filtering:

$$\begin{aligned}
&\mathcal{F}^{k_{\text{max}}}(\epsilon_{\text{DAS}}) = \dot{\epsilon}(\mathbf{u}^*)_{11} \\
&G^{\text{DAS}}_{11kl}\dot{\epsilon}(\mathbf{u}^*)_{kl} = \epsilon_{\text{DAS}} - \mathcal{F}^{k_{\text{max}}}(\epsilon_{\text{DAS}}) \,,
\end{aligned} \tag{5.10}$$

where $\mathcal{F}^{k_{\text{max}}}$ is the spatial low-pass filter with the cut-off frequency $k_{\text{max}} = 1/\lambda_{\text{min}}$. Upon on this separation, we can only obtain the one-component of homogenized strain along the DAS cable (xx-component in this study). Therefore, inversion of the corrector is also limited to only one component of corrector G_{1111}. The inversion of corrector can be done based on the equation:

$$G^{\text{DAS}}(\mathbf{x}_r)_{1111} = \frac{\sum_s \int_0^T (\epsilon_{s,\text{DAS}}(\mathbf{x}_r, t) - \mathcal{F}^{k_{\text{max}}}(\epsilon_{s,\text{DAS}}(\mathbf{x}_r, t)))\mathcal{F}^{k_{\text{max}}}(\epsilon_{s,\text{DAS}}(\mathbf{x}_r, t))dt}{\sum_s \int_0^T \mathcal{F}^{k_{\text{max}}}(\epsilon_{s,\text{DAS}}(\mathbf{x}_r, t))^2 dt} \tag{5.11}$$

5.4 Synthetic test

In this section, we present examples illustrating the impact of small-scale effects on DAS data and the effectiveness of inverted correctors to image small-scale obstacles in the subsurface.

5.4.1 Numerical setting

For the synthetic test simulating the strain rate of DAS measurements, we utilize the model and source-receiver configuration depicted in Fig. 5.1. The model dimensions are 40 m (x-direction) × 40 m (y-direction) × 100 m (z-direction). The background medium parameters consist of V_P of 300 m/s, V_S of 120 m/s, and a density of 1.4 t.m^{-3}. The center point of the small obstacle, whose size is 0.4 m (x-direction) × 0.4 m (y-direction) × 0.5 m (z-direction), is located at the point (16.6, 20.2, 0.25) in meters. This obstacle has V_P of 90 m/s, V_S of 40 m/s, and density of 0.42 t.m^{-3}. The DAS cable is placed in a straight line from (6.0, 20.6, 0.5) to (26.0, 20.6, 0.6), parallel to the surface and x-axis. The gauge length and channel interval of this DAS cable are set to 0.4 m and 0.2 m. The seismic source is characterized by Ricker-type wavelet.

We use SPECFEM3D software (Komatitsch & Tromp 2002a; Komatitsch & Tromp 2002b) to compute the strain rate. The absorbing boundary condition is employed at edges of model except for the surface.

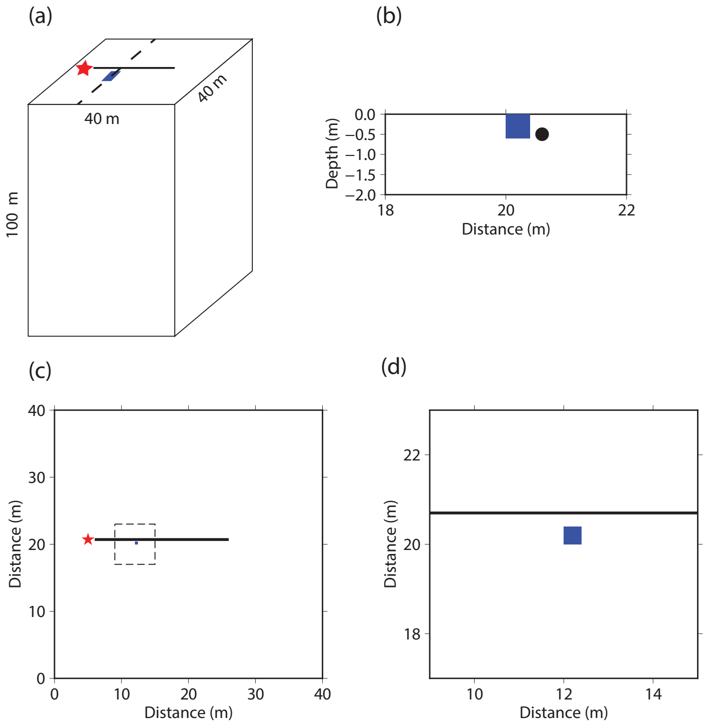


Fig. 5.1 Configuration of synthetic test. (a) Overview of the model for synthetic test with source (red star), DAS cable (black line), and small obstacle (blue). The length of DAS cable is 20 m and the size of obstacle is 0.4 m (x-direction) × 0.4 m (y-direction) × 0.5 m (z-direction). (b) Vertical section along the dashed line in panel (a). The obstacle (blue square) and the location of DAS cable (black circle) are shown. The obstacle is embedded from surface to 0.5 m depth. At the same depth of 0.5 m, the DAS cable is placed 20 cm far from DAS cable horizontally. (c) Plane view of panel (a). (d) Zoom of dashed sauare in panel (c).

5.4.2 Small-scale effects

In the initial synthetic experiment, we set the dominant frequency of the seismic source to 20 Hz. After computing the waveforms of strain rate, the bandpass filter is applied from 20 to 40 Hz. Based on the maximum frequency of 40 Hz and V_S of the background model, we can roughly estimate the minimum wavelength (λ_{min}) of 3.0 m. Therefore, the size of the anomaly is 6 times smaller vertically and 7.5 times smaller horizontally than the minimum wavelength.

First, we compute the strain rates without the small anomaly, hence, the model is a homogeneous model. The small-scale effects manifest as localized distortions of strain waveforms. Therefore, in this homogeneous case, higher spatial wavenumber components than $1/\lambda_{\mathrm{min}} = 0.33$ could be nearly zero. Fig.5.2 shows the waveform along the DAS cable for the homogeneous case. For the comparison, we also plot the waveform after applying the spatial lowpass filter with a maximum wavenumber ($k_{\max}$) of 0.33 1/m. The spatial filtered waveform well fit to the original waveform except at the edge of waveform (this is due to the short length of DAS cable in our synthetic test).

Fig.5.3 displays the waveforms of the DAS cable for the heterogeneous case with the anomaly. In contrast to the homogeneous case, a discrepancy between spatial filtered data and unfiltered data is clearly visible at the position of the anomaly (16.6 m offset). However, no such discrepancy is visible in the velocity seismogram, as shown in Fig.5.4. The localized discrepancy, which can be seen as a higher spatial component than the expected $k_{\max}$ (shown in Fig.5.3), is due to the small anomaly.

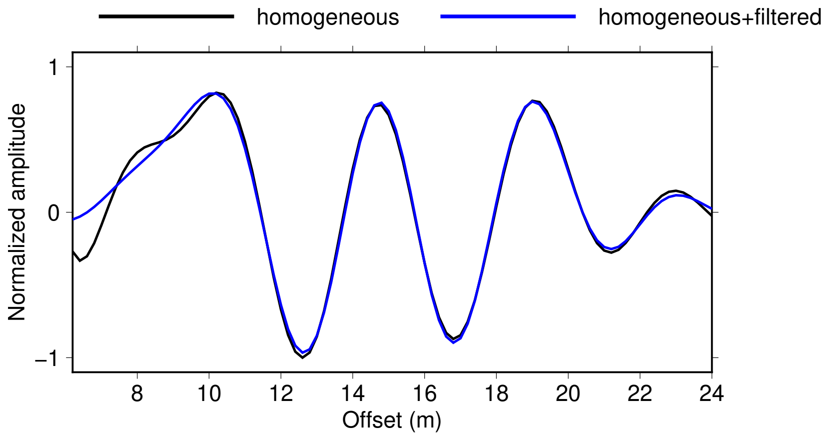


Fig. 5.2 The strain rates in the homogeneous model without small anomaly. The strain rates after applying a bandpass filter between 20–40 Hz (black line) and subsequently after applying a spatial lowpass filter with a cutoff maximum wavenumber of 0.33 1/m (blue line) are presented. The amplitudes are normalized by maximum amplitude of two waveforms. The offset is along the x-direction from the point (0, 20.6, 0.5). The red square at the bottom of the figure indicates the position of the anomaly.

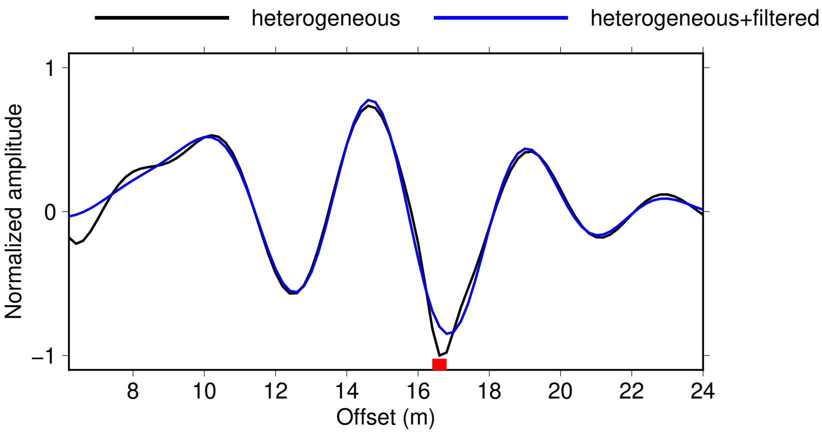


Fig. 5.3 The strain rates in the model with a small obstacle. The strain rates after applying a bandpass filter between 20–40 Hz (black line) and subsequently after applying a spatial lowpass filter with a cutoff maximum wavenumber of 0.33 1/m (blue line) are presented. The amplitudes are normalized by maximum amplitude of two waveforms. The offset is along the x-direction from the point (0, 20.6, 0.5). The red square at the bottom of the figure indicates the position of the anomaly.

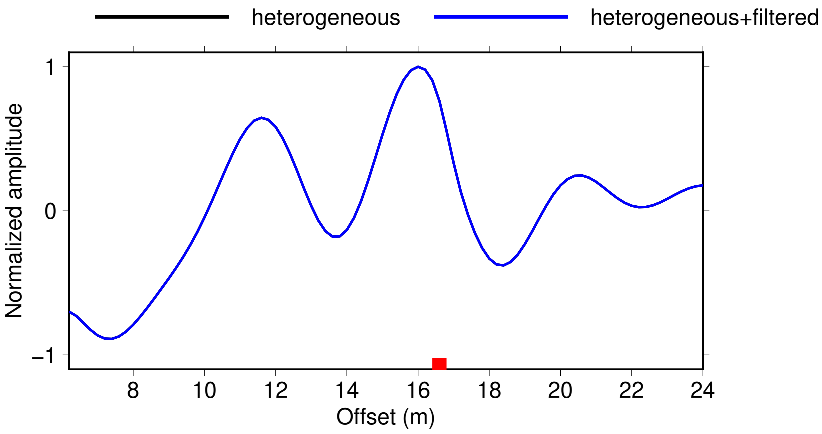


Fig. 5.4 The velocity waveform in the model with a small obstacle. The strain rates after applying a bandpass filter between 20–40 Hz (black line) and subsequently after applying a spatial lowpass filter with a cutoff maximum wavenumber of 0.33 1/m (blue line) are presented. The amplitudes are normalized by maximum amplitude of two waveforms. The offset is along the x-direction from the point (0, 20.6, 0.5). The red square at the bottom of the figure indicates the position of the anomaly.

5.4.3 Imaging the small anomaly by corrector

In Fig.5.5, we present the absolute values of inverted correctors based on the equation 5.11, comparing scenarios with and without the anomaly. Notably, a significant amplitude of correctors is located precisely at the position of the anomaly. In the absence of the anomaly, the expected nearly zero values for correctors exhibit some fluctuations, possibly due to filtering effects. These fluctuations may arise from an incomplete setting of the cut-off frequency or discontinuities in the input signal at both edges for $\mathcal{F}^{k_{\max}}$. However, despite these variations, pronounced changes in correctors are distinctly visible at the location of the anomaly.

The ability to image small-scale anomaly persists even when utilizing a lower frequency range. In the synthetic test conducted at lower frequencies, the center frequency of the source Ricker wavelet is set to 5 Hz and the bandpass filter in the 1–10 Hz range is applied for the strain rate. To handle with the longer minimum wavelength when applying spatial lowpass filtering, we extended the offset of DAS cable to 36 m. With the maximum frequency of 10 Hz, the minimum wavelength could be roughly estimated as $\lambda_{\min} = 12m$. The cut-off wavenumber for lowpass filtering to obtain homogenized strain rate is set at $k_{\max} = 0.0833$. Fig.5.6 illustrates the inverted correctors for both scenarios, with and without an anomaly. The clear alignment of the peak amplitude with the anomaly location is evident, although the spatial filtering effects due to the short length of DAS cable compared to the wavelength are significant. It's worth noting that the size of the obstacle is 24 times smaller vertically and 30 times smaller horizontally than the minimum wavelength. Any seismic tomography methods struggle to precisely locate such small-scale obstacles. However, the peak amplitude of the corrector provides a precise indicator of the small anomaly's location.

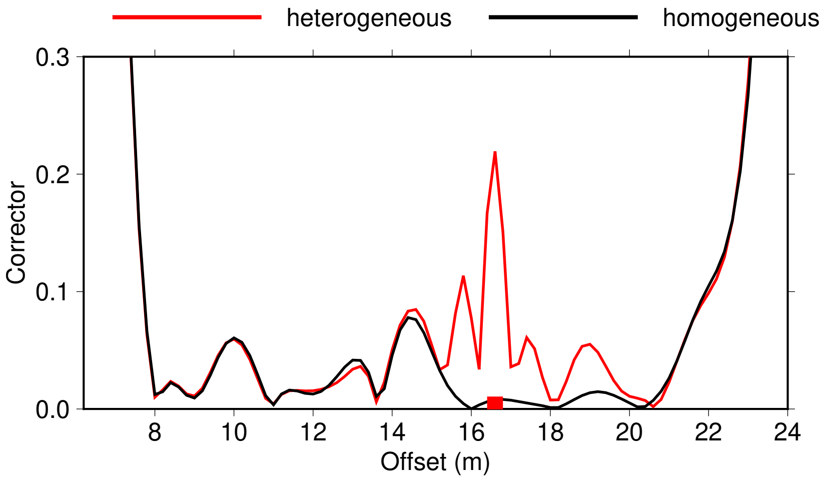


Fig. 5.5 Inverted correctors with (red line) and without (black line) the anomaly are shown. For the inversion of correctors, the data was initially filtered using a bandpass filter in the 20–40 Hz range, followed by the application of a spatial lowpass filter with a cut-off wavenumber of 0.33 1/m to obtain homogenized strain rate.

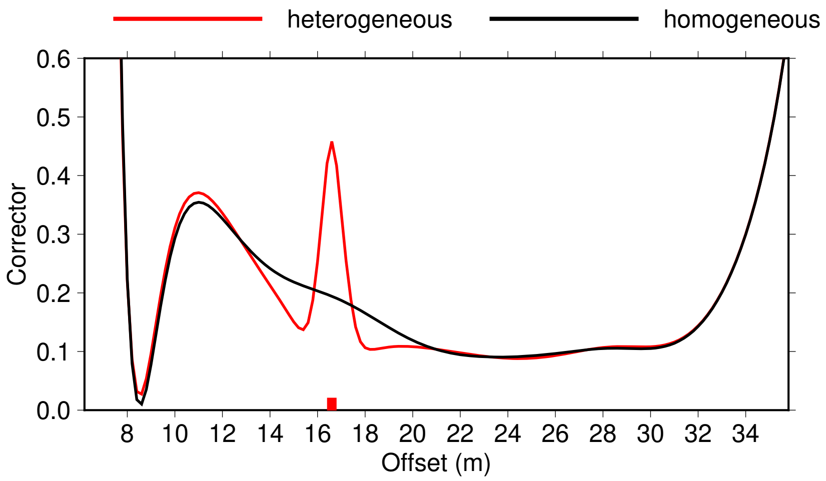


Fig. 5.6 Inverted correctors with (red line) and without (black line) the anomaly are shown. For the inversion of correctors, the data was initially filtered using a bandpass filter in the 1–10 Hz range, followed by the application of a spatial lowpass filter with a cut-off wavenumber of 0.0833 1/m to obtain homogenized strain rate.

5.5 Field experiments

5.5.1 DAS experiment in Kyushu University

In October 2023, the DAS cable was deployed at the field within Kyushu University. The primary objective of the project was to assess the viability of the DAS cable for subsurface monitoring at the CCS site. The DAS cable configuration consists of three main straight lines, as illustrated in Fig. 5.7: two lines were deployed at just beneath the surface, while one line is deployed within a borehole. The DAS cable was deployed at approximately 0.5 m depth. For the specific objectives of this study, focused on investigating small-scale effects on DAS data and exploring the potential for imaging small-scale structures, only a segment of the DAS cable (red rectangle in Fig.5.7) was used to prevent interference with the primary goal of the project.

We configured the observation parameters of the DAS system with a 0.4 m gauge length, a sampling frequency of 5000 Hz, and a spatial sampling interval of 0.2 m. Seismic waves were generated using a hummer shot on an iron plate (Fig. 5.8). This hummer shot produced seismic waves with a dominant frequency between 20–40 Hz, as illustrated in Fig. 5.9. Therefore, we applied a bandpass filter between 20–40 Hz before the analysis. The shot gather was obtained by stacking 20 shots. Based on prior surface wave analysis and shot gather, we roughly estimated the V_S in the field to be approximately 120 m/s, resulting in $\lambda_{\min}$

of approximately 3.0 m. Therefore, when performing an inversion of correctors, we use a spatial lowpass filter with a cut-off frequency of 0.33 1/m to obtain homogenized strain rate.

We conducted two field experiments. In the first experiment, an empty box served as a small-scale anomaly in the subsurface, while the second experiment employed perlite (Fig. 5.10). Perlite, being a lightweight material, is commonly used to enhance soil drainage, classifying both the empty box and perlite as low-velocity anomalies in the subsurface. Following the first experiment with the box, we removed the box from the ground and placed perlite at another location. The configurations of the obstacles, along with their sizes and relative positions to the DAS cable, are depicted in Figs. 5.11 and 5.12. The size of the box is less than 0.5 m, which is 6 times smaller than λ_{min}, while the size of the perlite is 0.4 m, which is 7.5 times smaller than λ_{min}. Both obstacles were placed in the vicinity of the DAS cable, 0.2 m away from it. In addition, note that the shot points are different between two experiments.

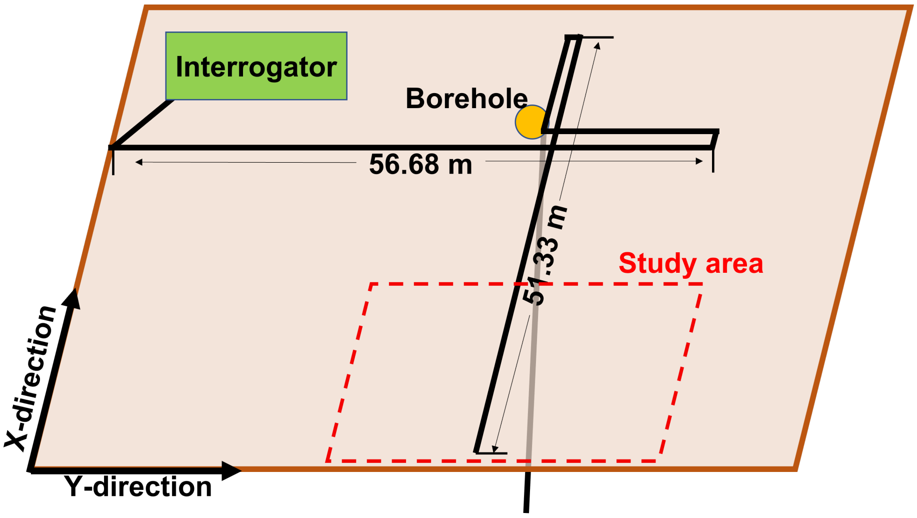


Fig. 5.7 The overview of DAS configuration of the field experiment. The DAS cable (black lines), borehole (yellow circle), and interrogator (green square) are shown. For the experiment in this study, only the DAS cable inside the study area (red dashed square) is used.

Fig. 5.8 The seismic wave is generated by a shot on the iron plate using a hummer.

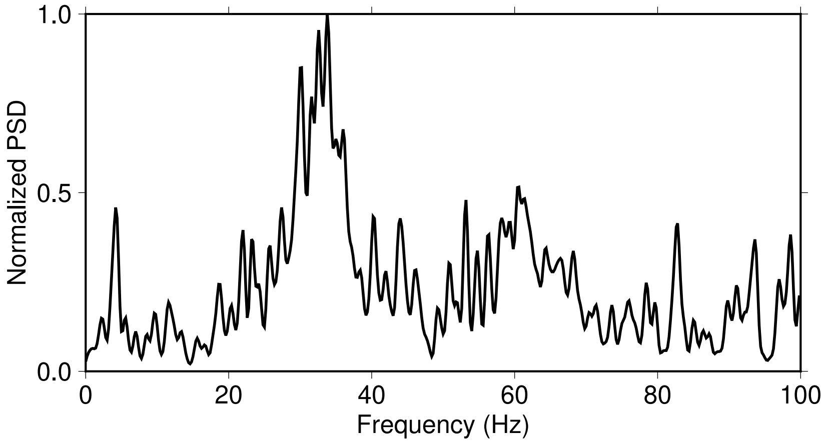


Fig. 5.9 The example of power spectral density of the hummer generated seismic wave.

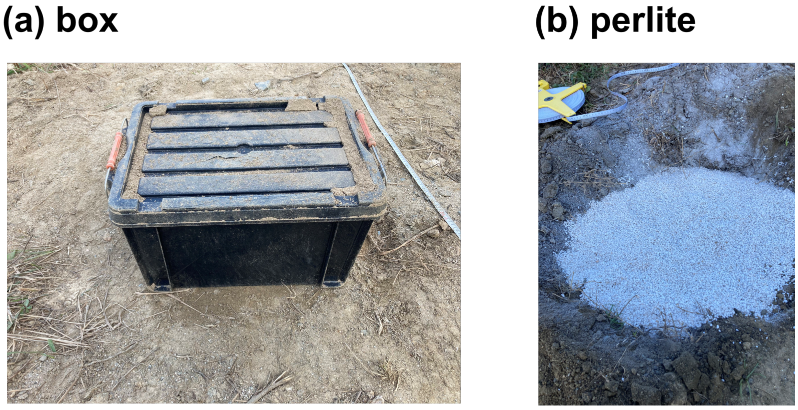


Fig. 5.10 The pictures of obstacles for the field experiment. (a) The empty box for the first experiment. (b) The perlite for the second experiment.

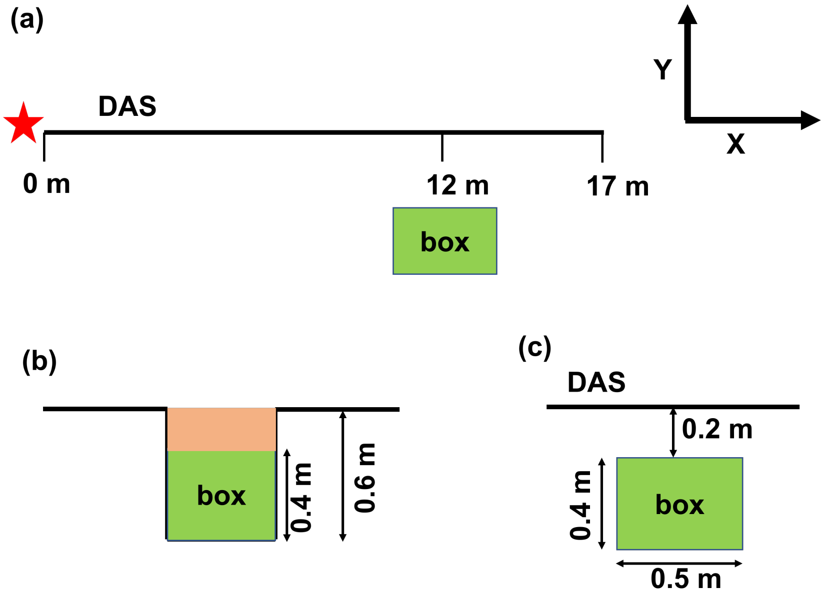


Fig. 5.11 The diagram illustrates the configurations of the DAS cable (depicted by the black line), seismic source (indicated by the red star), and the embedded box (highlighted by the green square). (a) Depicts the DAS cable and box configurations used in the analysis, with the box positioned at a 12 m offset. (b) Presents the depth section of the embedded box configuration, situated at a depth of 0.6 m. (c) Offers a top-down view of the DAS and box configurations, showcasing the box positioned 0.2 m away from the DAS cable.

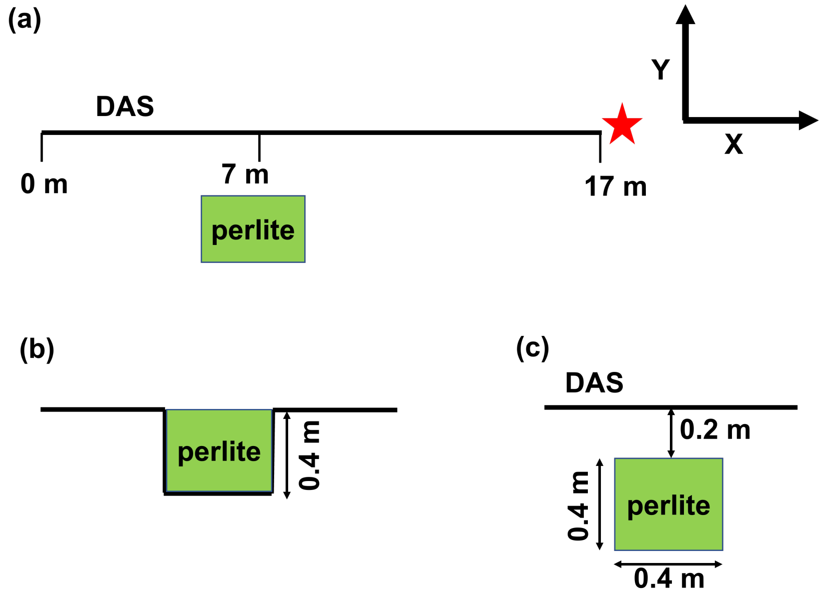


Fig. 5.12 The diagram illustrates the configurations of the DAS cable (depicted by the black line), seismic source (indicated by the red star), and the embedded perlite (highlighted by the green square). (a) Depicts the DAS cable and box configurations used in the analysis, with the perlite positioned at a 7 m offset. (b) Presents the depth section of the embedded perlite configuration, situated from the surface to 0.4 m depth. (c) Offers a top-down view of the DAS and perlite configurations, showcasing the perlite positioned 0.2 m away from the DAS cable.

5.5.2 First experiment: box

Fig. 5.13 displays unfiltered and spatial low-pass filtered DAS data waveforms (cut-off frequency: 0.33 1/m) before the box was embedded. The widespread discrepancies between the two waveforms indicate substantial subsurface heterogeneity at small scales.

Fig. 5.14 presents two waveforms (unfiltered and filtered) after the box was embedded. Noticeable fluctuations between two waveforms are clearly visible around the box location. However, these fluctuations are not localized at the box location. Additionally, the localized abrupt change observed at a 6.2 m offset in the waveform before embedding the box is not present after the box is embedded. Although we currently lack a clear explanation for this observation, one plausible reason is that the shot point of this first experiment is close to the endpoint of the DAS cable where the cable is exposed above the surface, which could make the DAS measurement unstable around the shot point.

Fig. 5.15 displays the inverted correctors. It is clear that the correctors are amplified after embedding the box around its position. However, as also inferred from the waveforms, the amplified correctors are not localized at the box position. In addition, significant changes in corrector values are present close to the shot point. Based on this result of inverted correctors, it would be challenging to interpret the precise location of the embedded box, but a rough estimation of the location is still possible.

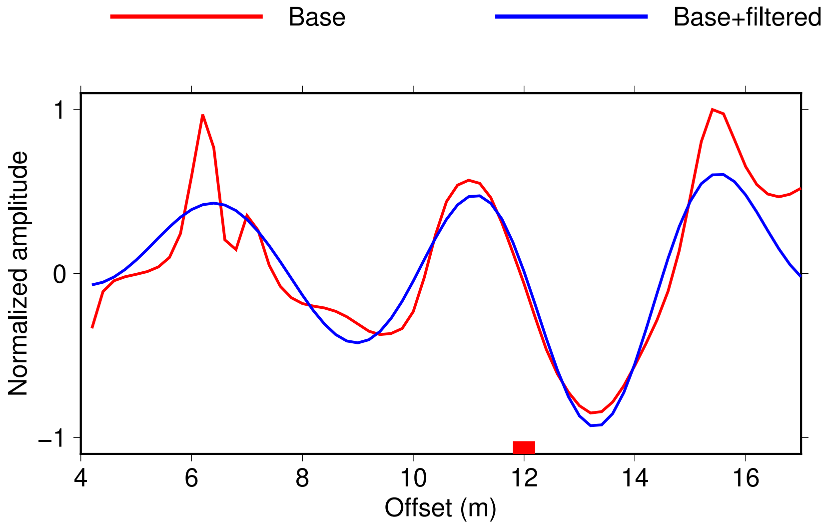


Fig. 5.13 Waveforms along the DAS cable before embedding the box (base data). Both the unfiltered (blue line) and spatially filtered (red line) waveforms are presented. The red square, positioned at a 12 m offset beneath the figure, indicates the location of the box.

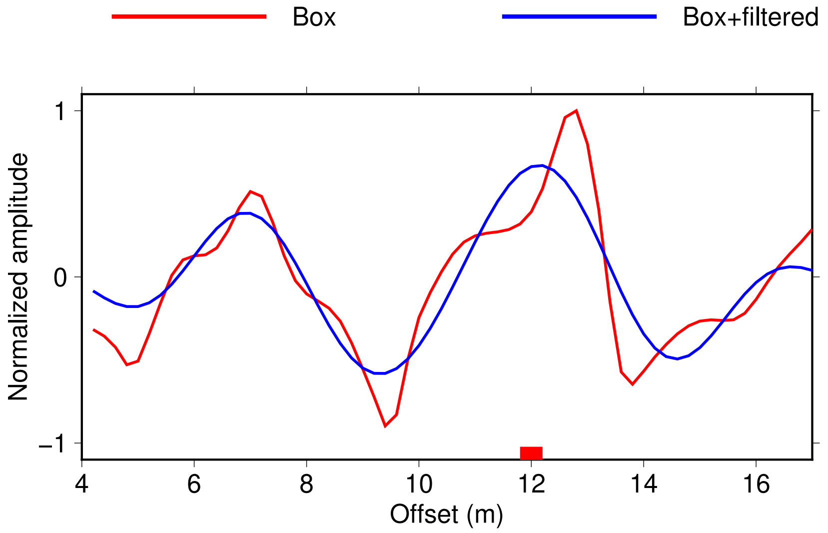


Fig. 5.14 Waveforms along the DAS cable after embedding the box. Both the unfiltered (blue line) and spatially filtered (red line) waveforms are presented. The red square, positioned at a 12 m offset beneath the figure, indicates the location of the box.

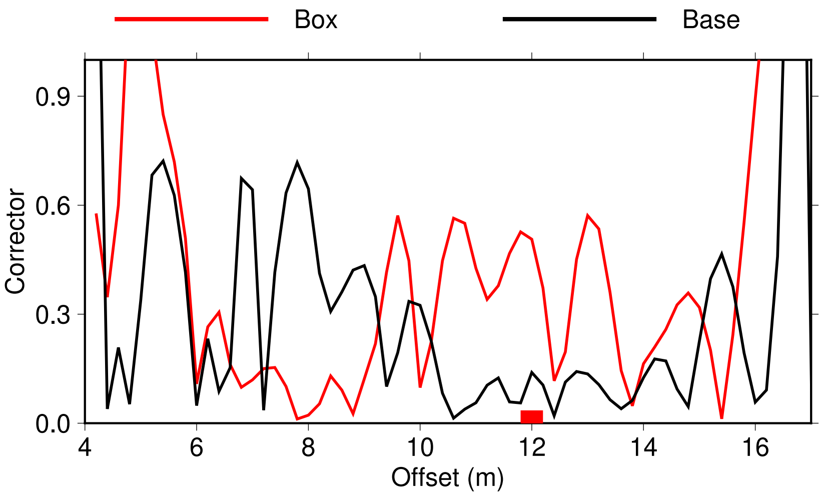


Fig. 5.15 Inverted correctors. Inverted correctors with (red line) and without (black line) box are shown. The red square, positioned at a 12 m offset beneath the figure, indicates the location of the box.

5.5.3 Second experiment: perlite

Figs. 5.16 and 5.17 display unfiltered and spatial low-pass filtered (cut-off frequency: 0.33 1/m) waveforms before (Fig.5.16) and after (Fig.5.17) embedding the perlite. Similar to the previous experiment, small-scale effects are present everywhere even before embedding perlite. Despite this, a strong spatial abrupt change in strain rate is pronounced at the perlite location after embedding.

Fig.5.18 shows the inverted correctors. The strong peak in corrector values at a 7 m offset corresponds to the location of the perlite. However, the corrector values change everywhere when comparing before and after embedding perlite, similar to the previous experiment. Despite these peculiar behaviors of correctors before and after embedding perlite, the inverted correctors after embedding precisely exhibit a localized peak at the perlite location, which means that the DAS measurements have the potential to obtain the location of small-scale anomalies in the subsurface through correctors.

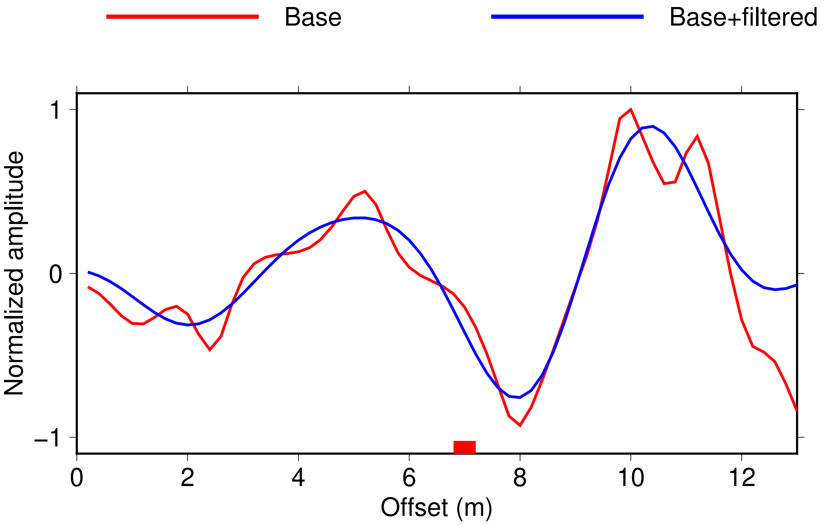


Fig. 5.16 Waveforms along the DAS cable before embedding the perlite (base data). Both the unfiltered (blue line) and spatially filtered (red line) waveforms are presented. The red square, positioned at a 12 m offset beneath the figure, indicates the location of the perlite.

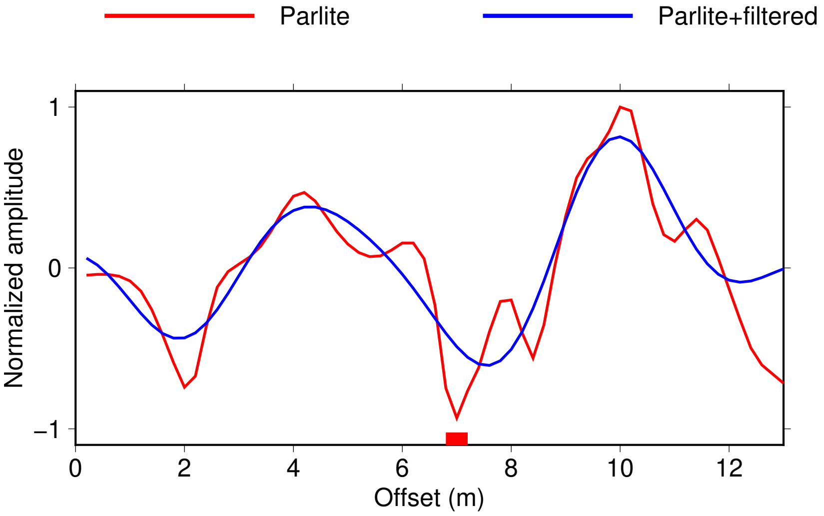


Fig. 5.17 Waveforms along the DAS cable after embedding the perlite. Both the unfiltered (blue line) and spatially filtered (red line) waveforms are presented. The red square, positioned at a 12 m offset beneath the figure, indicates the location of the perlite.

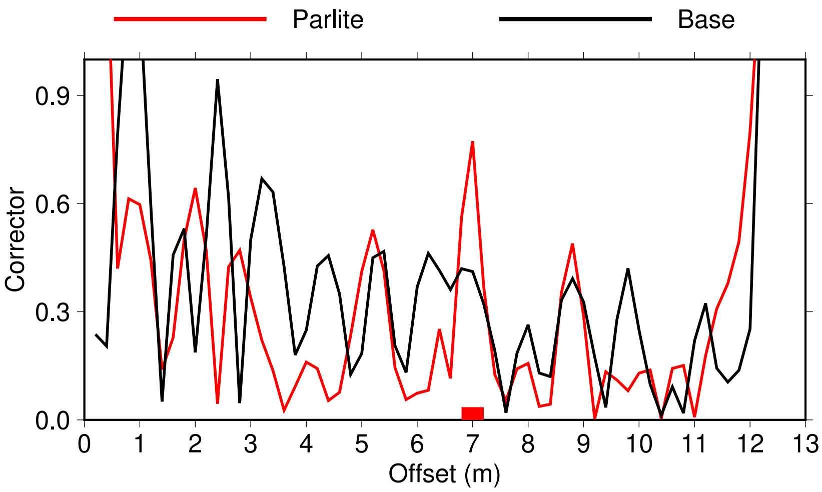


Fig. 5.18 Inverted correctors. Inverted correctors with (red line) and without (black line) perlite are shown. The red square, positioned at a 12 m offset beneath the figure, indicates the location of the perlite.

5.6 Discussion and Conclusion

We have first provided a theoretical foundation demonstrating that sub-wavelength anomalies can significantly influence DAS measurements, as elucidated by the two-scale homogenization theory. These small-scale effects can be shown as a coupling between homogenized strain and correctors. Subsequently, we proposed a method to invert correctors from DAS strain rate data without requiring knowledge of subsurface models or seismic source mechanisms. The method relies on spatial lowpass filtering to obtain homogenized strain rate. It's worth noting that this method has a limitation. We can only use seismic waves propagating parallel to the DAS cable, which imposes strong constraints on available seismic sources.

We employed numerical simulations to demonstrate how small-scale anomalies can be imaged through the inverted correctors derived from DAS measurements. The synthetic DAS measurements clearly exhibited small-scale effects on waveforms, absent in velocity seismograms. Furthermore, the inverted correctors precisely indicated the location of the small-scale anomaly. Despite the difficulties for any seismic imaging methods to image such small-scale anomaly, the inverted correctors from DAS measurements showcased the promising potential for imaging such small-scale subsurface structures.

Furthermore, we conducted a real field experiment using DAS and small obstacles with a size at least 6 times smaller than the minimum wavelength. While we successfully observed small-scale effects on the DAS waveforms, including clear peak amplitudes of inverted correctors at the locations of embedded small obstacles, the results were intricate. Based on comparing inverted correctors before and after embedding obstacles, we observed changes in correctors everywhere, including locations without any embedded obstacles. The reasons for these intricate results should be clarified by further real field experiments. Probably, the lower frequency signals with high S/N ratio such earthquakes ciyld yield better results due to its smooth and high quality data. Nevertheless, our field experiment results clearly demonstrated the impacts of small-scale structures on DAS measurements and the potential to image small-scale structures using correctors.

In both our synthetic tests and real field experiments, we deliberately prepared small-scale anomalies with sizes similar to the gauge length. If the anomaly is significantly smaller than the gauge length, based on the equation (5.4), the small-scale effects would be smoothed out from the DAS measurement and disappeared. Indeed, the spatial resolution limit for imaging small-scale structures using correctors would be comparable to the gauge length.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science (JSPS) through a Grant-in-Aid for JSPS Research Fellowship (22KJ2397) and JSPS Overseas Challenge Program for Young Researchers. The data acquisition by DAS was supported by INPEX CORPORATION.

Chapter 6

Conclusion

In Chapter 3, we initiated our exploration with the practical application of FWI by focusing on estimating subsurface structures of Japan Island. Using extensive seismic data from Japan's dense permanent seismic network, our goal was to achieve the accurate estimation of subsurface structures possible. The estimated seismic velocity model was derived from comprehensive computations of seismic waves, enabling the utilization of entire waveform data rather than specific portions, such as first-arrival times. This estimated seismic velocity model holds promise for various seismic applications, including moment tensor inversion for earthquakes.

Chapter 4 was dedicated to the theoretical and numerical exploration of FWI using wavefield gradient measurements. In contrast to conventional translational motions, wavefield gradient measurements exhibit sensitivity to small-scale structures, even when the wavelength significantly exceeds the size of these structures. This small-scale effect, explained by homogenization theory as the coupling between homogenized strain and correctors, presents challenges for FWI based on wavefield gradients. Despite the wavefield gradient measurement's sensitivity to small-scale structures, they fail to enhance spatial resolution in estimated subsurface images. Consequently, we proposed a method to mitigate these small-scale effects. Our approach involved inverting correctors alongside elastic mechanical properties, revealing its efficiency in eliminating small-scale effects and achieving subsurface model imaging with equivalent accuracy to methods using translational motions. Furthermore, this study demonstrated the potential of inverted correctors to visualize small-scale structures with unprecedented high spatial resolution.

Chapter 5 demonstrated that correctors inverted from wavefield gradient measurements can effectively image small-scale structures through synthetic and real field experiments. Our method enabled corrector inversion without requiring prior knowledge of subsurface models or seismic source information. Employing straightforward spatial filtering, this method proved remarkably simple yet effective. Real field experiments using DAS showed the sensitivity of DAS measurements to small-scale structures, even smaller than the minimum wavelength, indicating its potential for imaging local small-scale structures at the vicinity of DAS cables. Given DAS's significant emergence in seismology, our field experiments underscore a timely and novel application of DAS measurements. Because the field experiments were completed just two weeks before the submission of this Ph.D. thesis, certain tasks remain. Notably, exploring the possibility of inverting medium parameters or elastic contrasts through correctors. Additionally, investigating the effects of topography is crucial. The localized distortions of wavefield gradient measurements might be associated not only with small-scale structures but also with local topography variations.

Bibliography

Ajo-Franklin, J. B., Dou, S., Lindsey, N. J., Monga, I., Tracy, C., Robertson, M., Rodriguez Tribaldos, V., Ulrich, C., Freifeld, B., Daley, T. et al., 2019. Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection. *Scientific reports 9*(1), 1328.

Aki, K. & Lee, W., 1976. Determination of three-dimensional velocity anomalies under a seismic array using first p arrival times from local earthquakes: 1. a homogeneous initial model. *Journal of Geophysical research 81*(23), 4381–4399.

Arai, R. & Iwasaki, T., 2014. Crustal structure in the northwestern part of the izu collision zone in central japan. *Earth, Planets and Space*, **66**, 1–12.

Arai, R., Iwasaki, T., Sato, H., Abe, S. & Hirata, N., 2014. Contrasting subduction structures within the philippine sea plate: Hydrous oceanic crust and anhydrous volcanic arc crust. *Geochemistry, Geophysics, Geosystems 15*(5), 1977–1990.

Arnulf, A. F., Bassett, D., Harding, A. J., Kodaira, S., Nakanishi, A. & Moore, G., 2022. Upper-plate controls on subduction zone geometry, hydration and earthquake behaviour. *Nature Geoscience 15*(2), 143–148.

Atterholt, J., Zhan, Z. & Yang, Y., 2022. Fault zone imaging with distributed acoustic sensing: Body-to-surface wave scattering. *Journal of Geophysical Research: Solid Earth 127*(11), e2022JB025052.

Backus, G., 1962. Long-wave elastic anisotropy produced by horizontal layering. *J. Geophys. Res. 67*(11), 4427–4440.

Bensoussan, A., Lions, J.-L. & Papanicolaou, G., 1978. *Asymptotic Analysis of Periodic Structures*. North Holland.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J., 2010. Obspy: A python toolbox for seismology. *Seismological Research Letters 81*(3), 530–533.

Bois, P., La Porte, M., Lavergne, M. & Thomas, G., 1971. Essai de determination automatique des vitesses sismiques par mesures entre puits. *Geophysical Prospecting 19*(1), 42–83.

Bois, P., La Porte, M., Lavergne, M. & Thomas, G., 1972. Well-to-well seismic measurements. *Geophysics 37*(3), 471–480.

Boore, D. M., 1972. *Finite-difference method for seismic wave propagation in heterogeneous materials*, Chapter 11. Academic Press, Inc.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D., 2016. Global adjoint tomography: first-generation model. *Geophysical Journal International 207*(3), 1739–1766.

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.

Browaeys, J. T. & Chevrot, S., 2004. Decomposition of the elastic tensor and geophysical applications. *Geophys. J. Int.*, **159**, 667–678.

Burgos, G., Capdeville, Y. & Guillot, L., 2016. Homogenized moment tensor and the effect of near-field heterogeneities on nonisotropic radiation in nuclear explosion. *J. Geophys. Res. 121*(6), 4366–4389.

Capdeville, Y., 2021. Homogenization of seismic point and extended sources. *Geophysical Journal International 226*(2), 1390–1416.

Capdeville, Y., Cupillard, P. & Singh, S., 2020. In B. Moseley & L. Krischer (Eds.), *Machine Learning in Geosciences*, Volume 61 of *Advances in Geophysics*, pp. 217 – 306. Elsevier.

Capdeville, Y., Guillot, L. & Marigo, J. J., 2010a. 1-D non periodic homogenization for the wave equation. *Geophys. J. Int.*, **181**, 897–910.

Capdeville, Y., Guillot, L. & Marigo, J. J., 2010b. 2D nonperiodic homogenization to upscale elastic media for P-SV waves. *Geophys. J. Int.*, **182**, 903–922.

Capdeville, Y. & Marigo, J. J., 2008. Shallow layer correction for spectral element like methods. *Geophys. J. Int.*, **172**, 1135–1150.

Capdeville, Y. & Métivier, L., 2018. Elastic full waveform inversion based on the homogenization method: theoretical framework and 2-d numerical illustrations. *Geophysical Journal International 213*(2), 1093–1112.

Capdeville, Y., Stutzmann, E., Wang, N. & Montagner, J.-P., 2013. Residual homogenization for seismic forward and inverse problems in layered media. *Geophys. J. Int. 194*(1), 470–487.

Capdeville, Y., Zhao, M. & Cupillard, P., 2015. Fast fourier homogenization for elastic wave propagation in complex media. *Wave Motion*, **54**, 170–186.

Clayton, R. & Engquist, B., 1977. Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the seismological society of America 67*(6), 1529–1540.

CROCKETT, J. B. & CHERNOFF, H., 1955. Gradient methods of maximization. *Pacific J. Math*, **5**, 33–50.

Cupillard, P. & Capdeville, Y., 2018. Non-periodic homogenization of 3-d elastic media for the seismic wave equation. *Geophysical Journal International 213*(2), 983–1001.

Dahlen, F., Hung, S.-H. & Nolet, G., 2000. Frechet kernels for finite-frequency traveltimes – I. theory. *Geophys. J. Int.*, **141**, 157–174.

Dahlen, F. A. & Tromp, J., 1998. *Theoretical Global Seismology*. Princeton University Press. NJ.

Daley, T., Miller, D., Dodds, K., Cook, P. & Freifeld, B., 2016. Field testing of modular borehole monitoring with simultaneous distributed acoustic sensing and geophone vertical seismic profiles at citronelle, alabama. *Geophysical Prospecting 64*(5), 1318–1334.

Daley, T. M., Freifeld, B. M., Ajo-Franklin, J., Dou, S., Pevzner, R., Shulakova, V., Kashikar, S., Miller, D. E., Goetz, J., Henninges, J. et al., 2013. Field testing of fiber-optic distributed acoustic sensing (das) for subsurface seismic monitoring. *The Leading Edge 32*(6), 699–706.

Dou, S., Lindsey, N., Wagner, A. M., Daley, T. M., Freifeld, B., Robertson, M., Peterson, J., Ulrich, C., Martin, E. R. & Ajo-Franklin, J. B., 2017. Distributed acoustic sensing for seismic monitoring of the near surface: A traffic-noise interferometry case study. *Scientific reports 7*(1), 11620.

Dziewonski, A., 1984. Mapping the lower mantle: determination of lateral heterogeneity in p velocity up to degree and order 6. *J. Geophys. Res. 89*(B7), 5929–5952.

Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference Earth model. *Phys. Earth Planet. Inter.*, **25**, 297–356.

Ekström, G., Nettles, M. & Dziewoński, A., 2012. The global cmt project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, **200**, 1–9.

Festa, G. & Vilotte, J.-P., 2005. The newmark scheme as velocity-stress time-staggering: an efficient implementation for spectral element simulations of elastodynamics. *Geophys. J. Int.*, **161**, 789–812.

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

Fichtner, A. & Igel, H., 2009. Sensitivity densities for rotational ground-motion measurements. *Bulletin of the Seismological Society of America 99*(2B), 1302–1314.

Fichtner, A., Kennett, B. L., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental-and global-scale full-waveform inversion in the time–frequency domain. *Geophysical Journal International 175*(2), 665–685.

Fichtner, A., Kennett, B. L., Igel, H. & Bunge, H.-P., 2009a. Full seismic waveform tomography for upper-mantle structure in the australasian region using adjoint methods. *Geophysical Journal International 179*(3), 1703–1725.

Fichtner, A., Kennett, B. L. N., Igel, H. & Bunge, H. P., 2009b. Full waveform tomography for upper-mantle structure in the australasian region using adjoint methods. *Geophys. J. Int.*, **179**, 1703–1725.

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

Fichtner, A., van Herwaarden, D.-P., Afanasiev, M., Simutė, S., Krischer, L., Çubuk-Sabuncu, Y., Taymaz, T., Colli, L., Saygin, E., Villaseñor, A. et al., 2018. The collaborative seismic earth model: generation 1. *Geophysical research letters 45*(9), 4007–4016.

Fukao, Y., Obayashi, M., Inoue, H. & Nenbai, M., 1992. Subducting slabs stagnant in the mantle transition zone. *Journal of Geophysical Research: Solid Earth 97*(B4), 4809–4822.

George, T., Virieux, J. & Madariaga, R., 1987. Seismic wave synthesis by gaussian beam summation; a comparison with finite differences. *Geophysics 52*(8), 1065–1073.

Goulty, N. R., 1976. Strainmeters and tiltmeters in geophysics. *Tectonophysics 34*(3-4), 245–256.

Grattan, K. & Sun, T., 2000. Fiber optic sensor technology: an overview. *Sensors and Actuators A: Physical 82*(1-3), 40–61.

Guillot, L., Capdeville, Y. & Marigo, J. J., 2010. 2-D non periodic homogenization for the SH wave equation. *Geophys. J. Int.*, **182**, 1438–1454.

Hara, T., Tsuboi, S. & Geller, R. J., 1993. Inversion for laterally heterogeneous upper mantle s-wave velocity structure using iterative waveform inversion. *Geophysical Journal International 115*(3), 667–698.

Harrison, J., 1976. Cavity and topographic effects in tilt and strain measurement. *Journal of Geophysical Research 81*(2), 319–328.

Hirose, F., Nakajima, J. & Hasegawa, A., 2008. Three-dimensional seismic velocity structure and configuration of the philippine sea slab in southwestern japan estimated by double-difference tomography. *Journal of Geophysical Research: Solid Earth 113*(B9).

Hubbard, P. G., Vantassel, J. P., Cox, B. R., Rector, J. W., Yust, M. B. & Soga, K., 2022. Quantifying the surface strain field induced by active sources with distributed acoustic sensing: Theory and practice. *Sensors 22*(12), 4589.

Igel, H., Schreiber, U., Flaws, A., Schuberth, B., Velikoseltsev, A. & Cochard, A., 2005. Rotational motions induced by the m8. 1 tokachi-oki earthquake, september 25, 2003. *Geophysical research letters 32*(8).

Ikelle, L. T. & Amundsen, L., 2018. *Introduction to petroleum seismology*. Society of Exploration Geophysicists.

Kamei, R. & Pratt, R., 2013. Inversion strategies for visco-acoustic waveform inversion. *Geophys. J. Int. 194*(2), 859–884.

Kashiwagi, H. & Nakajima, J., 2019. Three-dimensional seismic attenuation structure of central japan and deep sources of arc magmatism. *Geophysical Research Letters 46*(23), 13746–13755.

Kato, A. & Nakagawa, S., 2020. Detection of deep low-frequency earthquakes in the nankai subduction zone over 11 years using a matched filter technique. *Earth, Planets and Space 72*(1), 1–9.

Keil, S., Wassermann, J. & Igel, H., 2021. Single-station seismic microzonation using 6c measurements. *Journal of Seismology*, **25**, 103–114.

Kennett, B. L., Engdahl, E. & Buland, R., 1995. Constraints on seismic velocities in the earth from traveltimes. *Geophysical Journal International 122*(1), 108–124.

Kennett, B. L. N. & Engdahl, E. R., 1991. Traveltimes for Global Earthquake Location and Phase Identification. *Geophys. J. Int.*, **105**, 429–465.

Kimura, G., Nakamura, Y., Shiraishi, K., Fujie, G., Kodaira, S., Tsuji, T., Fukuchi, R. & Yamaguchi, A., 2022. Nankai forearc structural and seismogenic segmentation caused by a magmatic intrusion off the kii peninsula. *Geochemistry, Geophysics, Geosystems 23*(8), e2022GC010331.

King, G. & Bilham, R., 1973. Tidal tilt measurement in europe. *Nature 243*(5402), 74–75.

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

Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation—i. validation. *Geophysical Journal International 149*(2), 390–412.

Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—ii. three-dimensional models, oceans, rotation and self-gravitation. *Geophysical Journal International 150*(1), 303–318.

Komatitsch, D. & Vilotte, J. P., 1998. The spectral element method: an effective tool to simulate the seismic response of 2D and 3D geological structures. *Bull. Seism. Soc. Am.*, **88**, 368–392.

Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D., 2020. Global adjoint tomography—model glad-m25. *Geophysical Journal International 223*(1), 1–21.

Lennon, G. & Baker, T., 1973. The earth tide signal and its coherency. *Quarterly Journal of the Royal Astronomical Society*, **14**, 161.

Maeda, T., Obara, K., Furumura, T. & Saito, T., 2011. Interference of long-period seismic wavefield observed by the dense hi-net array in japan. *Journal of Geophysical Research: Solid Earth 116*(B10).

Maggi, A., Tape, C., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography. *Geophysical Journal International 178*(1), 257–281.

McLeod, D., Stedman, G., Webb, T. & Schreiber, U., 1998. Comparison of standard and ring laser rotational seismograms. *Bulletin of the Seismological Society of America 88*(6), 1495–1503.

Meertens, C., Levine, J. & Busby, R., 1989. Tilt observations using borehole tiltmeters: 2. analysis of data from yellowstone national park. *Journal of Geophysical Research: Solid Earth 94*(B1), 587–601.

Mégnin, C. & Romanowicz, B., 2000. The 3D shear velocity structure of the mantle from the inversion of body, surface and higher modes wave forms. *Geophys. J. Int.*, **143**, 709–728.

Melchior, P. J., 1966. *The earth tides*. Oxford.

Mikumo, T. & Aki, K., 1964. Determination of local phase velocity by intercomparison of seismograms from strain and pendulum instruments. *Journal of Geophysical Research 69*(4), 721–731.

Miyoshi, T., Obayashi, M., Peter, D., Tono, Y. & Tsuboi, S., 2017. Adjoint tomography of the crust and upper mantle structure beneath the kanto region using broadband seismograms. *Progress in Earth and Planetary Science 4*(1), 1–20.

Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E. & Hung, S., 2004. Finite-frequency tomography reveals a variety of mantle plumes,. *Science*, **303**, 338–343.

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

Muir, J. B. & Zhan, Z., 2022. Wavefield-based evaluation of das instrument response and array design. *Geophysical Journal International 229*(1), 21–34.

Mukumoto, K. & Tsuji, T., 2023. 3-d crustal shear wave velocity model derived from full-waveform tomography for central honshu island, japan. *Geophysical Journal International 235*(1), 366–376.

Nachman, A. I., 1988. Reconstructions from boundary measurements. *Annals of Mathematics 128*(3), 531–576.

Nakajima, J. & Hasegawa, A., 2007a. Deep crustal structure along the niigata-kobe tectonic zone, japan: Its origin and segmentation. *Earth, planets and space*, **59**, e5–e8.

Nakajima, J. & Hasegawa, A., 2007b. Subduction of the philippine sea plate beneath southwestern japan: Slab geometry and its relationship to arc magmatism. *Journal of Geophysical Research: Solid Earth 112*(B8).

Nakajima, J. & Hasegawa, A., 2016. Tremor activity inhibited by well-drained conditions above a megathrust. *Nature communications 7*(1), 13863.

Nakajima, J., Hirose, F. & Hasegawa, A., 2009. Seismotectonics beneath the tokyo metropolitan area, japan: Effect of slab-slab contact and overlap on seismicity. *Journal of Geophysical Research: Solid Earth 114*(B8).

Nakamura, G. & Uhlmann, G., 1994. Global uniqueness for an inverse boundary problem arising in elasticity. *Inventiones mathematicae 118*(1), 457–474.

Nimiya, H., Ikeda, T. & Tsuji, T., 2020. Three-dimensional s wave velocity structure of central japan estimated by surface-wave tomography using ambient noise. *Journal of Geophysical Research: Solid Earth 125*(4), e2019JB019043.

Nishida, K., Kawakatsu, H. & Obara, K., 2008. Three-dimensional crustal s wave velocity structure in japan using microseismic data recorded by hi-net tiltmeters. *Journal of Geophysical Research: Solid Earth 113*(B10).

Nishimura, E., 1950. On earth tides. *Eos, Transactions American Geophysical Union 31*(3), 357–376.

Nocedal, J., 1980. Updating Quasi-Newton Matrices With Limited Storage. *Mathematics of Computation 35*(151), 773–782.

Okada, Y., Kasahara, K., Hori, S., Obara, K., Sekiguchi, S., Fujiwara, H. & Yamamoto, A., 2004. Recent progress of seismic observation networks in japan—hi-net, f-net, k-net and kik-net—. *Earth, Planets and Space 56*(8), xv–xxviii.

Örsvuran, R., Bozdağ, E., Modrak, R., Lei, W. & Ruan, Y., 2020. Double-difference measurements in global full-waveform inversions. *Geophysical Journal International 220*(1), 661–680.

Parker, L., Thurber, C., Zeng, X., Li, P., Lord, N., Fratta, D., Wang, H., Robertson, M., Thomas, A., Karplus, M. et al., 2018. Active-source seismic tomography at the brady geothermal field, nevada, with dense nodal and fiber-optic seismic arrays. *Seismological Research Letters 89*(5), 1629–1640.

Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casarotti, E., Le Loher, P., Magnoni, F., Liu, Q., Blitz, C. et al., 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes. *Geophysical Journal International 186*(2), 721–739.

Pratt, R., Shin, C. & Hicks, G., 1998. Gauss-newton and full newton methods in frequency domain seismic waveform inversion. *Geophys. J. Int.*, **133**, 341–362.

Renat, Z., Cupillard, P. & Capdeville, Y., 2022. Interpretation of time reversal focal spots based on point-source homogenization. *Geophysical Journal International 231*(3), 1653–1671.

Romanowic, B. & Giardini, D., 2001. The future of permanent seismic networks.

Sanchez-Palencia, E., 1980. *Non homogeneous media and vibration theory*. Number 127 in Lecture Notes in Physics. Berlin: Springer.

Sassa, K. & Nishimura, E., 1951. On phenomena forerunning earthquakes. *Eos, Transactions American Geophysical Union 32*(1), 1–6.

Seno, T. & Yamasaki, T., 2003. Low-frequency tremors, intraslab and interplate earthquakes in southwest japan—from a viewpoint of slab dehydration. *Geophysical Research Letters 30*(22).

Simutė, S., Steptoe, H., Cobden, L., Gokhberg, A. & Fichtner, A., 2016. Full-waveform inversion of the japanese islands region. *Journal of Geophysical Research: Solid Earth 121*(5), 3722–3741.

Singh, S., Capdeville, Y. & Igel, H., 2020, February. Correcting wavefield gradients for the effects of local small-scale heterogeneities. *Geophysical Journal International 220*(2), 996–1011.

Snieder, R., 1988. Large-scale waveform inversions of surface waves for lateral heterogeneity: 1. theory and numerical examples. *Journal of Geophysical Research: Solid Earth 93*(B10), 12055–12065.

Stacey, R., 1988. Improved transparent boundary formulations for the elastic-wave equation. *Bulletin of the Seismological Society of America 78*(6), 2089–2097.

Takano, O., 2002. Changes in depositional systems and sequences in response to basin evolution in a rifted and inverted basin: an example from the neogene niigata-shin'etsu basin, northern fossa magna, central japan. *Sedimentary Geology 152*(1-2), 79–97.

Tang, L., Igel, H. & Montagner, J.-P., 2023. Single-point dispersion measurement of surface waves combining translation, rotation and strain in weakly anisotropic media: theory. *Geophysical Journal International 235*(1), 24–47.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern california crust based on spectral-element and adjoint methods. *Geophys. J. Int. 180*(1), 433–462.

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

Tarantola, A., 1988. Theoritical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure Appl. Geophys. 128*(1/2), 365–399.

Tarantola, A., 2005. *Inverse problem theory and methods for model parameter estimation*. SIAM.

Tarantola, A. & Valette, B., 1982. Inverse problems= quest for information. *J. Geophys 50*(3), 150–170.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys. J. Int.*, **160**, 195–216.

Tsuji, T., Ashi, J., Strasser, M. & Kimura, G., 2015. Identification of the static backstop and its influence on the evolution of the accretionary prism in the nankai trough. *Earth and Planetary Science Letters*, **431**, 15–25.

Tsuji, T., Minato, S., Kamei, R., Tsuru, T. & Kimura, G., 2017. 3d geometry of a plate boundary fault related to the 2016 off-mie earthquake in the nankai subduction zone, japan. *Earth and Planetary Science Letters*, **478**, 234–244.

Ueno, H., 2002. Improvement of hypocenter determination procedures in the japan meteorological agency. *QJ Seismol.*, **65**, 123–134.

Valentine, A. P. & Woodhouse, J. H., 2010. Reducing errors in seismic tomography: combined inversion for sources and structure. *Geophysical Journal International 180*(2), 847–857.

van Driel, M., Wassermann, J., Nader, M. F., Schuberth, B. S. & Igel, H., 2012. Strain rotation coupling and its implications on the measurement of rotational ground motions. *Journal of Seismology*, **16**, 657–668.

Virieux, J., 1986. P-SV wave propagation in heterogeneous media: velocity-stress finite difference method. *Geophysics*, **51**, 889–901.

Wang, H. F., Zeng, X., Miller, D. E., Fratta, D., Feigl, K. L., Thurber, C. H. & Mellors, R. J., 2018. Ground motion response to an ml 4.3 earthquake using co-located distributed acoustic sensing and seismometer arrays. *Geophysical Journal International 213*(3), 2020–2036.

Wassermann, J., Wietek, A., Hadziioannou, C. & Igel, H., 2016. Toward a single-station approach for microzonation: Using vertical rotation rate to estimate love-wave dispersion curves and direction finding. *Bulletin of the Seismological Society of America 106*(3), 1316–1330.

Willis, M. E., Barfoot, D., Ellmauthaler, A., Wu, X., Barrios, O., Erdemir, C., Shaw, S. & Quinn, D., 2016. Quantitative quality of distributed acoustic sensing vertical seismic profile data. *The Leading Edge 35*(7), 605–609.

Woodhouse, J. & Dziewonski, A., 1984. Mapping the upper mantle: three-dimensional modeling of earth structure by inversion of seismic waveforms. *J. Geophys. Res. 89*(B7), 5953–5986.

Xing, E., Jordan, M., Russell, S. J. & Ng, A., 2002. Distance metric learning with application to clustering with side-information. *Advances in neural information processing systems*, **15**.

Yamada, I., 2009. *Functional analysis for engineering*. SUURIKOUGAKUSHA.

Yano, T. E., Takeda, T., Matsubara, M. & Shiomi, K., 2017. Japan unified high-resolution relocated catalog for earthquakes (juice): crustal seismicity beneath the japanese islands. *Tectonophysics*, **702**, 19–28.

Yomogida, K., 1992. Fresnel zone inversion for lateral heterogeneities in the earth. *Pure and Applied Geophysics*, **138**, 391–406.

Yuan, S., Lellouch, A., Clapp, R. G. & Biondi, B., 2020. Near-surface characterization using a roadside distributed acoustic sensing array. *The Leading Edge 39*(9), 646–653.

Yuan, Y. O., Simons, F. J. & Tromp, J., 2016. Double-difference adjoint seismic tomography. *Geophysical Journal International 206*(3), 1599–1618.

Zhan, Z., 2020. Distributed acoustic sensing turns fiber-optic cables into sensitive seismic antennas. *Seismological Research Letters 91*(1), 1–15.

Zhao, M., Capdeville, Y. & Zhang, H., 2016. Direct numerical modeling of time-reversal acoustic subwavelength focusing. *Wave Motion*, **67**, 102–115.

Appendix: Comprehensive results of FWI based on wavefield gradients

Appendix A: Supplementary inversion results

We present comprehensive inversion and statistical results for three different cases: displacement-based FWI using receivers positioned within the heterogeneous area; rotational-based FWI, without correctors, using receivers positioned within the heterogeneous area; and rotational-based FWI, with correctors, using receivers positioned within the heterogeneous area. We show the target model and FWI solutions for each case, with $\varepsilon_0 = 1$, in Fig. 1. The inversion fails to converge to the homogenised target model when correctors are not included in the FWI process. We present the statistical results in Figs. 2–4 to confirm the importance of including correctors to improve the FWI solution.

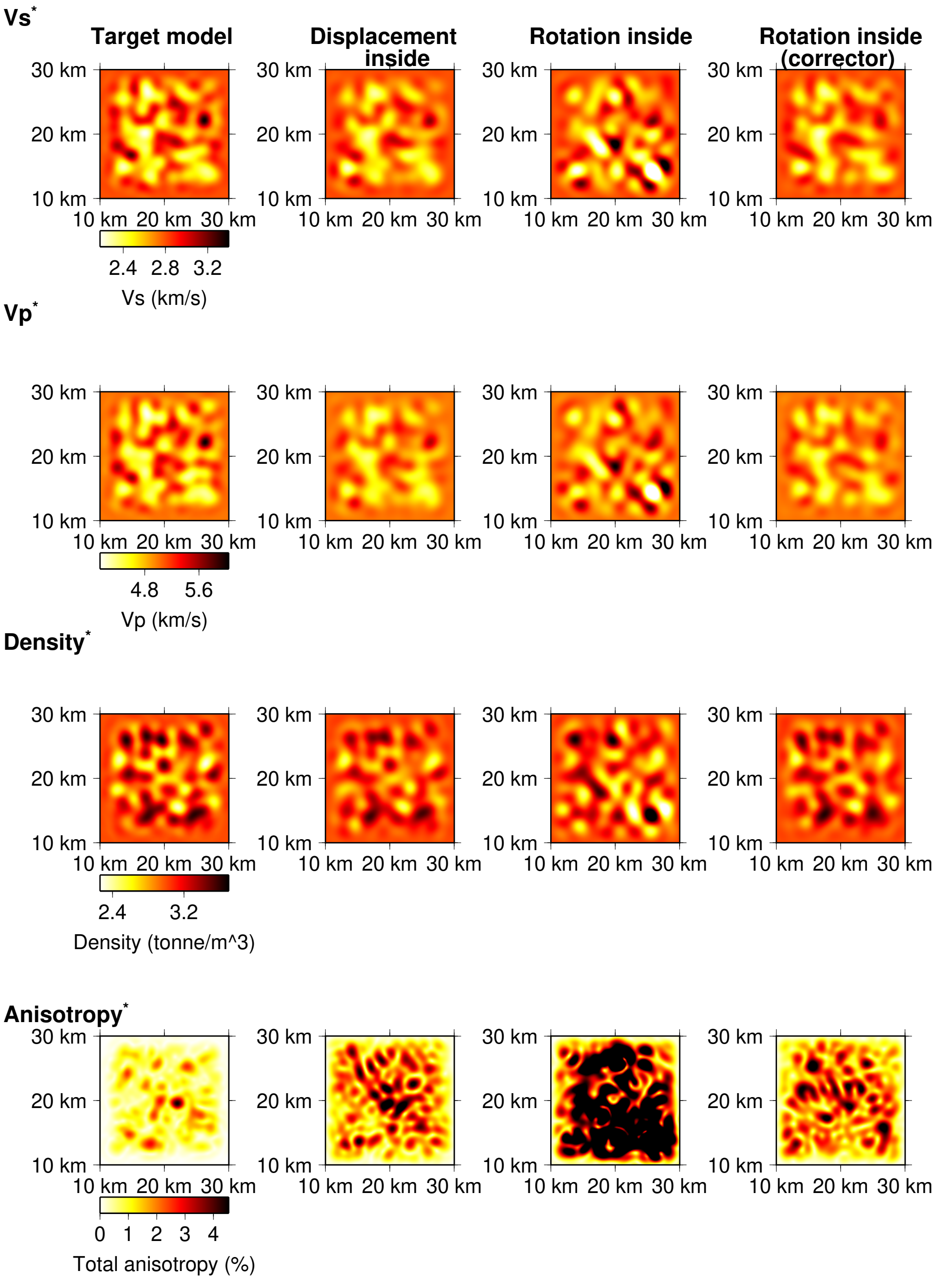


Fig. 1 Comparison of the FWI results. The V_S^*, V_P^*, density* and anisotropy* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Target model ("Target model"). Displacement-based FWI results using receivers positioned within the heterogeneous area ("Displacement inside"). Rotational-based FWI results, without correctors, using receivers positioned within the heterogeneous area ("Rotation inside"). Rotational-based FWI results, with correctors, using receivers positioned within the heterogeneous area ("Rotation inside (corrector)").

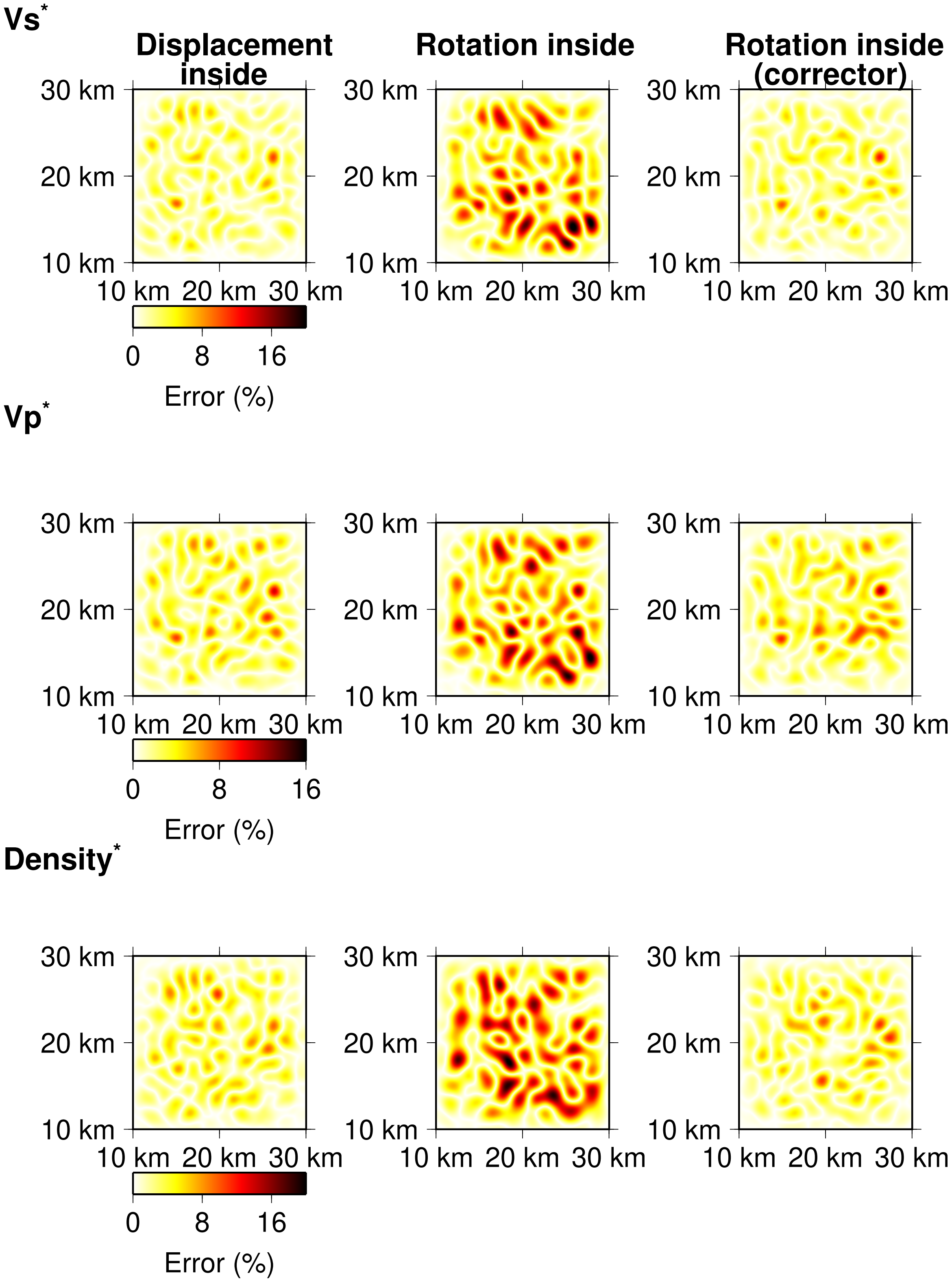


Fig. 2 Comparison of the error of mean models of the FWI results: The V_S^*, V_P^* and density* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Mean model for displacement receivers positioned within the heterogeneous area ("Displacement inside"). Mean model, without correctors, for rotational receivers positioned within the heterogeneous area ("Rotation inside"). Mean model, with correctors, for rotational receivers positioned within the heterogeneous area ("Rotation inside (corrector)").

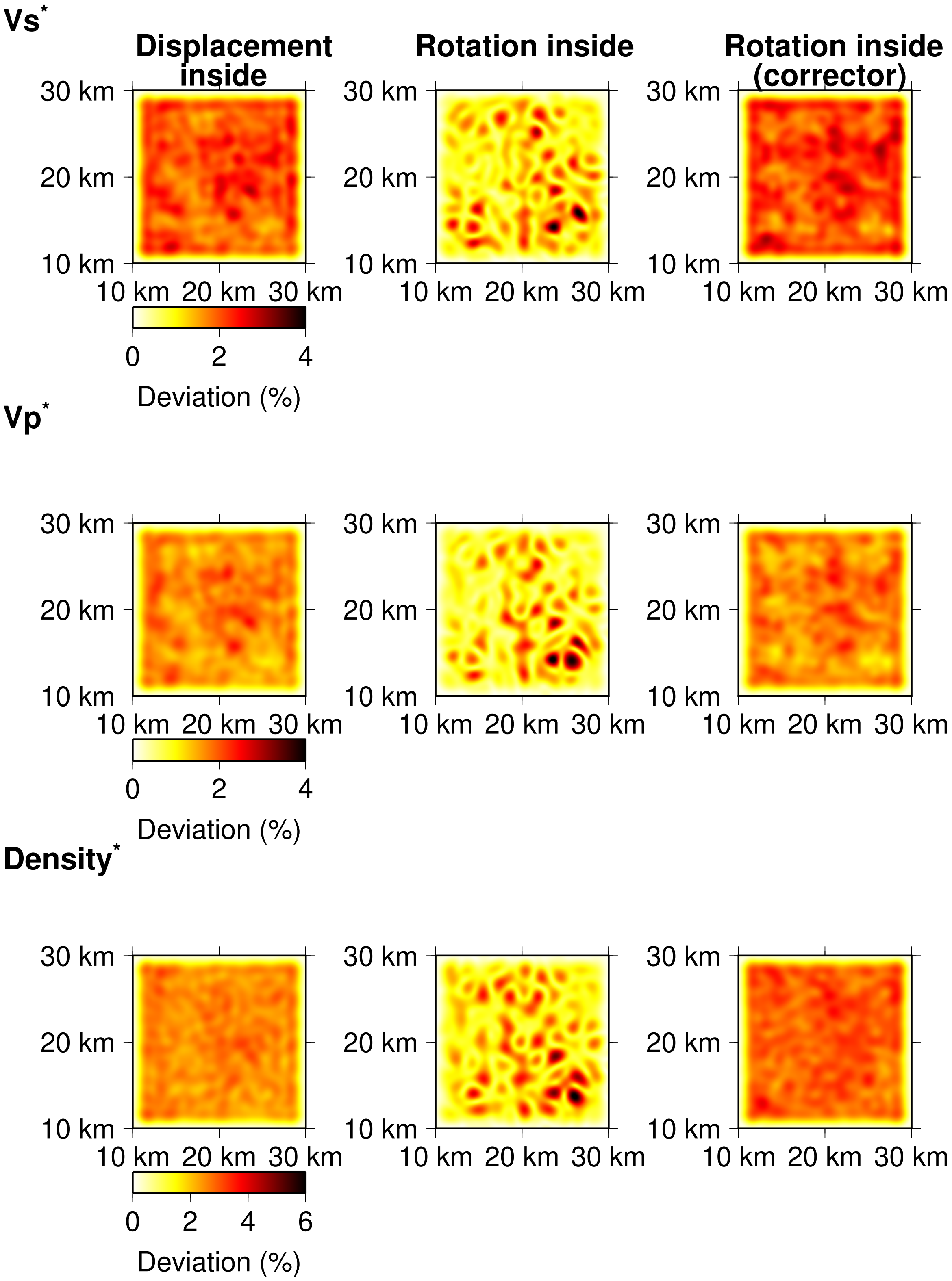


Fig. 3 Comparison of the standard deviations in the FWI results from the target model. The V_S^*, V_P^* and density* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Deviation for displacement receivers positioned within the heterogeneous area ("Displacement inside"). Deviation, without correctors, for rotational receivers positioned within the heterogeneous area ("Rotation inside"). Deviation, with correctors, for rotational receivers positioned within the heterogeneous area ("Rotation inside (corrector)").

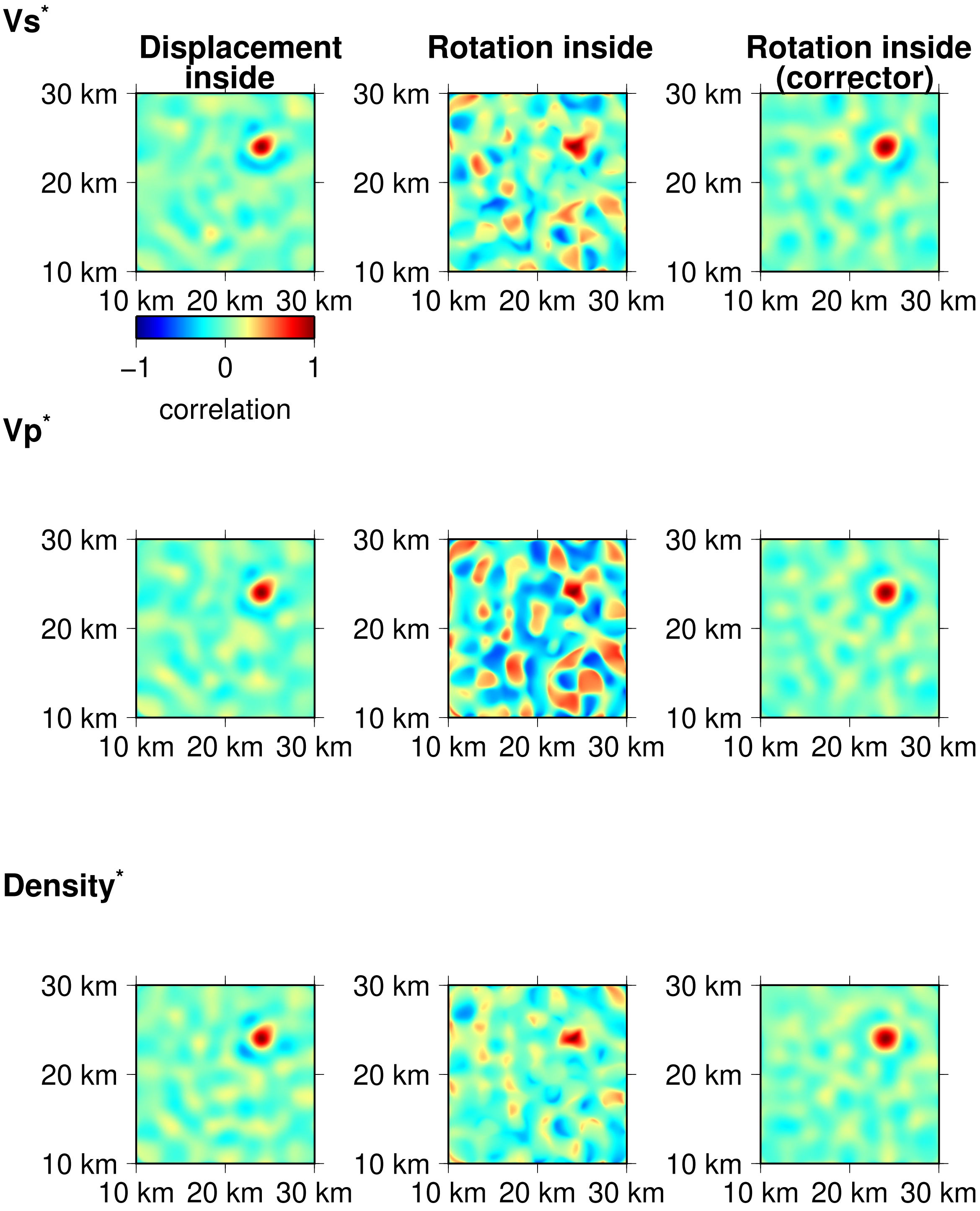


Fig. 4 Comparison of the resolutions of the FWI results. The V_S^*, V_P^* and density* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Resolution for displacement receivers positioned within the heterogeneous area ("Displacement inside"). Resolution, without correctors, for rotational receivers positioned within the heterogeneous area ("Rotation inside"). Resolution, with correctors, for rotational receivers positioned within the heterogeneous area ("Rotation inside (corrector)").

Appendix B: Strain-based FWI

Here we present numerical FWI experiments using strain data. We followed the inversion scheme that is explained in section 3 for this numerical experiment. We performed numerical tests with two different receiver configurations. One configuration involved placing all of the receivers outside of the heterogeneous area ("Receivers outside" in Fig. 5), and the other configuration involved placing all of the receivers within the heterogeneous area ("Receivers inside" in Fig. 5). We only used the vertical component along the vertical axis of the strain data at each receiver location in these numerical experiments. This setup was designed to simulate a DAS-based FWI; however, we did not consider the gauge length and strain rate of DAS measurements in our numerical experiments.

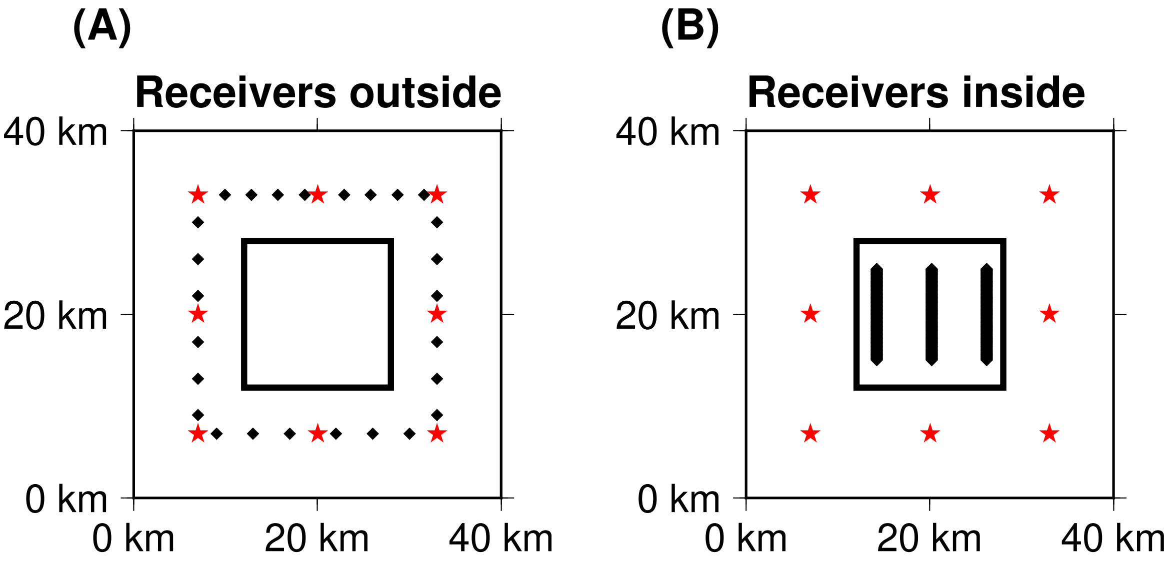


Fig. 5 Source (red stars) and receiver (black diamonds) configurations for the synthetic tests of the strain-based FWI. (A) Configuration with 8 sources and 26 receivers positioned outside of the heterogeneous area. (B) Configuration with 8 sources positioned outside of the heterogeneous area and 150 receivers positioned within the heterogeneous area.

C1. Receivers outside of the inversion area

Fig. 6 presents the misfit reduction results and waveform comparisons using receivers positioned outside of the heterogeneous area. A steady misfit reduction is observed throughout the iterative process, with the final misfit satisfying the convergence criteria at iteration number 41. There is a strong agreement between the synthetic and target waveforms, which demonstrates that the final model can adequately explain the target data. The final V_S^*, V_P^*, and density* models, and the corresponding residuals between the target and final models are shown in Fig. 7, with the residual values being comparable to those obtained via the displacement-based FWI.

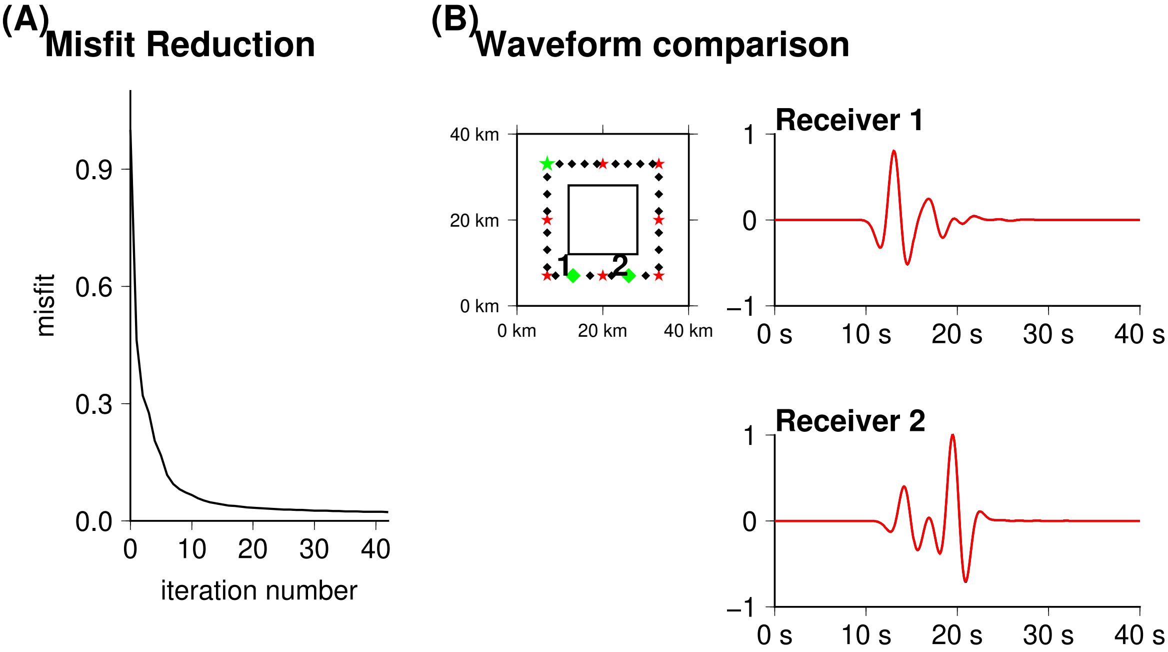


Fig. 6 Strain-based FWI results using receivers positioned outside of the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions indicating the source–receiver pairs used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers.

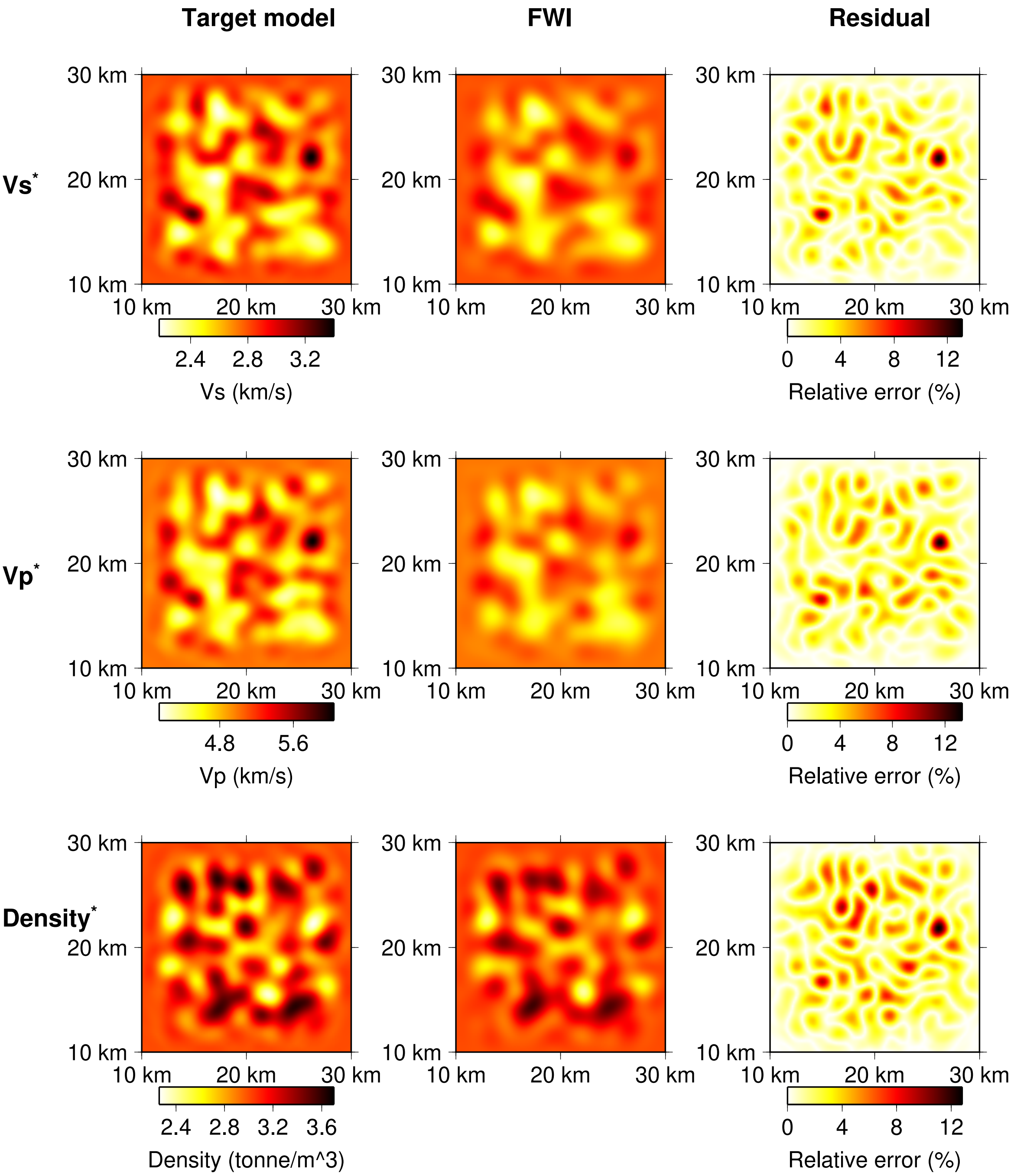


Fig. 7 Strain-based FWI results using receivers positioned outside of the heterogeneous area. The V_S^*, V_P^* and density* structures are shown for the target ("Target model") and final models ("FWI"), and their corresponding residuals ("Residuals"). Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from background V_S (2.8 km/s), V_P (5.0 km/s) and density (3.0 t.m^{-3}) values.

C2. Receivers within the inversion area for the FWI without correctors

Figures 8 and 9 illustrate the inversion results when the receivers are positioned within the heterogeneous area and correctors are not included. The misfit decreased until iteration 2, but then increased in subsequent iterations. Furthermore, there is a large degree of misfit between the synthetic and target waveforms, as well as large residuals between the target and final models. Similar to the rotational-based FWI case, the strain-based FWI without correctors fails to converge to the best model and is therefore incapable of explaining the target data when the effects of small-scale heterogeneities are present in the target data.

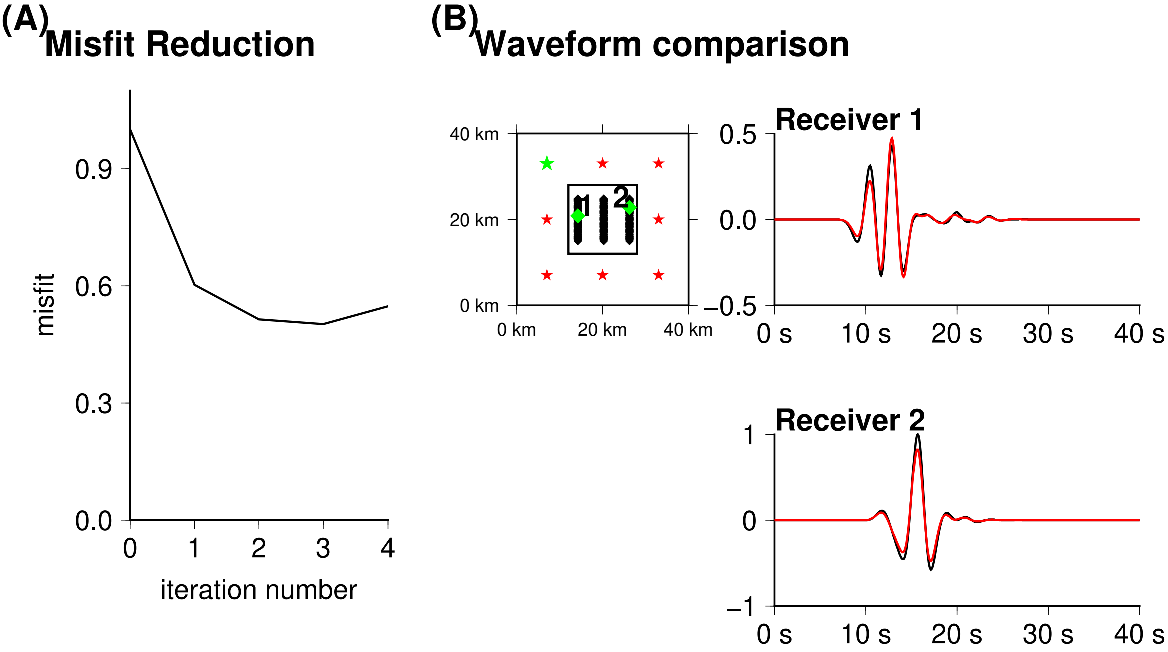


Fig. 8 Strain-based FWI results, without correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions indicating the source–receiver pairs used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers.

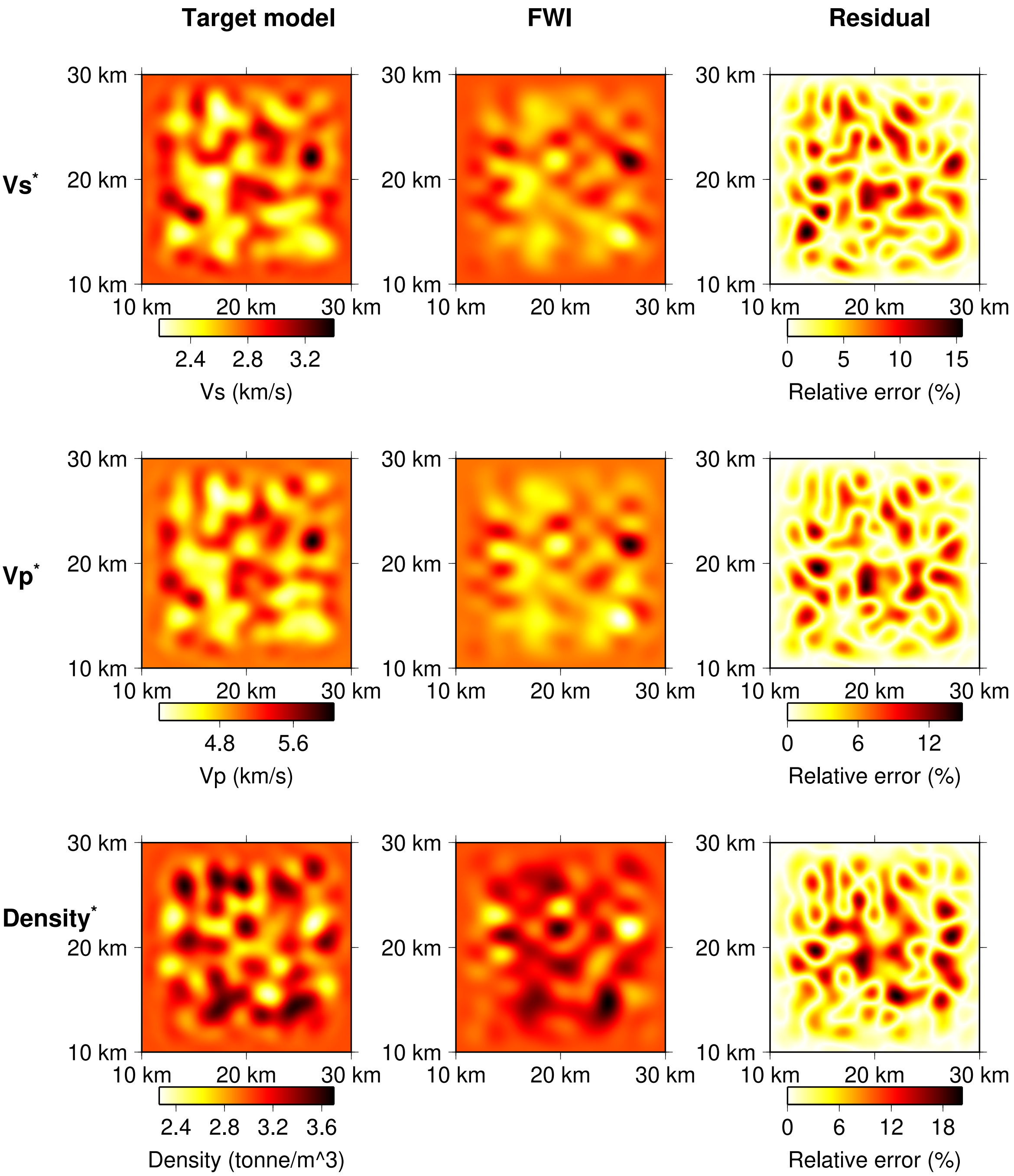


Fig. 9 Strain-based FWI results, without correctors, using receivers positioned within the heterogeneous area. The V_S^*, V_P^* and density* structures are shown for the target ("Target model") and final models ("FWI"), and their corresponding residuals ("Residual"). Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from background V_S (2.8 km/s), V_P (5.0 km/s) and density (3.0 t.m^{-3}) values.

C3. Receivers within inversion area for the FWI with correctors

We then performed the FWI with correctors using receivers that were positioned within the heterogeneous area. Fig. 10 shows the misfit reduction and waveform comparisons. The misfit steadily decreased throughout the iterative process, and a strong fit between the synthetic and target waveforms was observed, thereby indicating that FWI with correctors can adequately explain the target data. Fig.11 presents the model comparisons and residuals for each parameter, whereby the residuals show improvements in all of the parameters compared to the results without correctors. Table 1 provides the misfits in the elastic tensors based on equation (4.46) for a single model in each case, with the incorporation of correctors leading to an improvement in the accuracy of the inverted elastic parameters. Therefore, it is necessary to incorporate correctors in the FWI process to obtain accurate results when the strain observations are used.

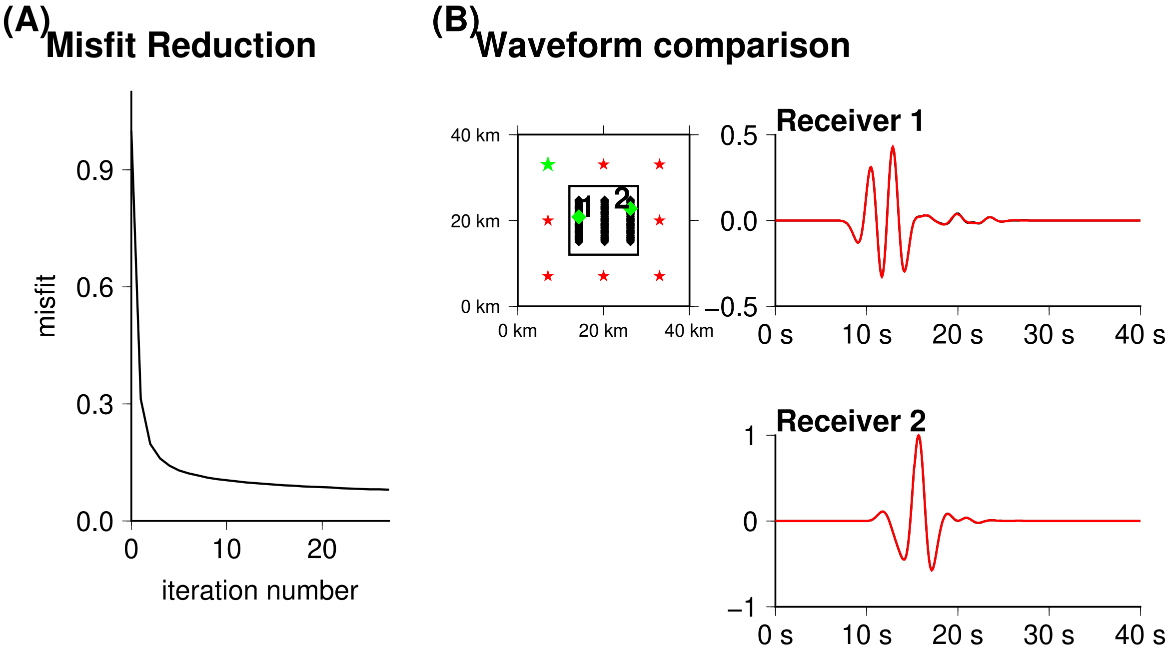


Fig. 10 Strain-based FWI results, with correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions indicating the source–receiver pairs that were used to generate the presented target (black) and synthetic (red) waveforms in the right panel. The waveforms are normalised by the maximum amplitude of the selected receivers.

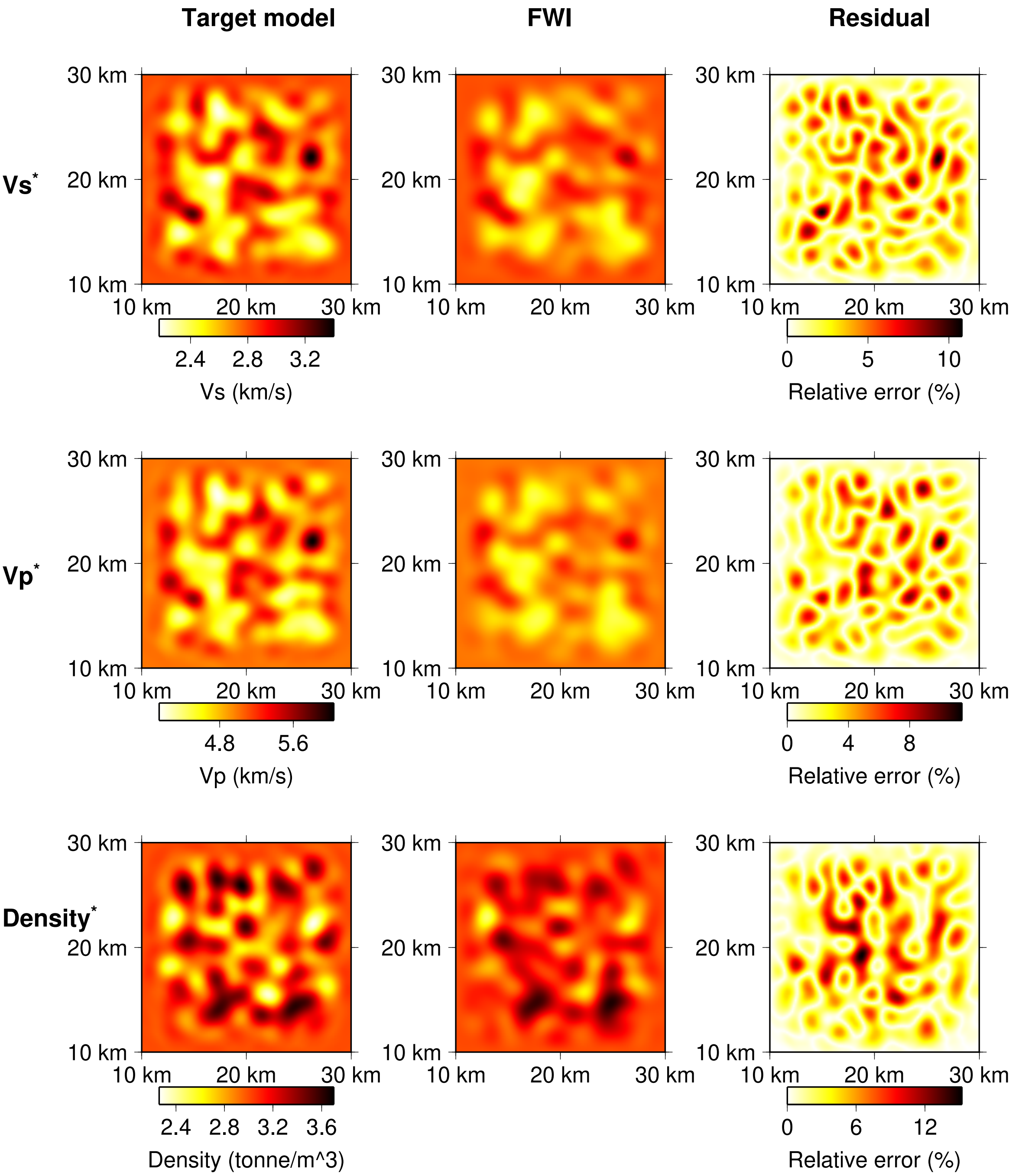


Fig. 11 Strain-based FWI results, with correctors, using receivers positioned within the heterogeneous area. The V_S^*, V_P^* and density* structures are shown for the target ("Target model") and final models ("FWI"), and their corresponding residuals ("Residual"). Both the target and final models are homogenised using $\varepsilon_0 = 1.0$. The residuals between the target and final models are represented as fluctuations from background V_S (2.8 km/s), V_P (5.0 km/s) and density (3.0 t.m^{-3}) values.

Table 1 Model misfit based on equation (4.46) for results of numerical experiment using strain data.

Case	$\bar{E}_m(\varepsilon_0 = 1)$
FWI without correctors	
Strain data and receivers outside heterogeneous area	0.041
Strain data and receivers inside heterogeneous area	0.078
FWI with correctors	
Strain data and receivers inside heterogeneous area	0.057