Accounting for sub-wavelength heterogeneities in full waveform inversion based on wavefield gradient measurements

Kota Mukumoto, Yann Capdeville, Sneha Singh, Takeshi Tsuji

Geophys. J. Int. (2020) **000,** 000–000

Accounting for sub-wavelength heterogeneities in full waveform inversion based on wavefield gradient measurements

Kota Mukumoto, Yann Capdeville, Sneha Singh, Takeshi Tsuji

12 December 2023

keywords:

SUMMARY

Seismic gradient measurements from distributed acoustic sensors, rotational sensors and other instruments are becoming increasingly available for field surveys. Such measurements provide a wealth of information and are currently being considered for many applications. In this work, using a simple 2-D numerical approach, we tackle the implications of such wavefield gradient measurements on full waveform inversion (FWI) techniques using a simple two-dimensional numerical test. In particular, we study the impact of the wavefield gradient measurement sensitivity to heterogeneities that are much smaller than the minimum wavelength. Indeed, as shown previously through the homogenization method, small scale heterogeneities induce an unexpected coupling of the strain components to the wavefield gradient measurement. We further show that this coupling introduces a potential limitation to the FWI results if it is not taken into account. We demonstrate that a gradient measurement-based FWI can only reach the accuracy of a classical displacement field-based FWI if the coupling coefficients are also inverted. Furthermore, there appears to be no specific gain in using gradient measurements instead of conventional displacement (or velocity, acceleration) measurements to image structures. Nevertheless, the inverted correctors contain fine-scale heterogeneities information that could be exploited to reach an unprecedented resolution, particularly if an array of receivers is used.

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), seismic-wave 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.

2 FULL WAVEFORM INVERSION AND HOMOGENIZATION THEORY

2.1 Context

We consider a two-dimensional (2D) elastic domain $\mathbf{\Omega}$ with absorbing boundary conditions on its boundary $\partial\mathbf{\Omega}$. We use a Cartesian coordinate system in $\mathbf{\Omega}$, where the x-axis is the horisontal 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 equation:

$$\rho\partial_{tt}\mathbf{u} - \boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = \boldsymbol{f}_s\,, \tag{1}$$

$$\boldsymbol{\sigma} = \mathbf{c} : \boldsymbol{\epsilon}(\mathbf{u})\,, \tag{2}$$

where $\epsilon_{kl}(\mathbf{u}) = \frac{1}{2}(\partial_k u_l + \partial_l u_k)$ is the strain operator, $(k,l) \in \{x,z\}^2$, $\rho(\mathbf{x})$ 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 and $\mathbf{x} = (x,z)$ is the position vector in $\mathbf{\Omega}$. We use point sources, such that:

$$\boldsymbol{f}_s(\mathbf{x},t) = \mathbf{F}_s\,\delta(\mathbf{x}-\mathbf{x}_s)g(t) \tag{3}$$

for a point force, or:

$$\boldsymbol{f}_s(\mathbf{x},t) = -\mathbf{M}_s\cdot\boldsymbol{\nabla}\delta(\mathbf{x}-\mathbf{x}_s)g(t) \tag{4}$$

for a point moment tensor, where $\mathbf{F}_s$ is the force vector, $\mathbf{M}_s$ is the moment tensor, $\mathbf{x}_s$ is the source location and $g(t)$ is the source time wavelet.

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.

2.2 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 multi-scale heterogeneous media with no scale separation for wave propagation (Capdeville *et al.* 2010a; Capdeville *et al.* 2010b; Cupillard & Capdeville 2018).

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:

$$\varepsilon_0 = \frac{\lambda_0}{\lambda_{min}}. \tag{5}$$

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. 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).

The displacement $\mathbf{u}$, which is a solution of the wave equation (equations (1)-(2)), is a power-series approximation in ε_0:

$$\mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) = \mathbf{u}^*(\mathbf{x},\mathbf{y},t) + \varepsilon_0\mathbf{u}^1(\mathbf{x},\mathbf{y},t) + \varepsilon_0^2\mathbf{u}^2(\mathbf{x},\mathbf{y},t) + \dots , \tag{6}$$

where the homogenised coefficients $\mathbf{u}^i$ need to be determined. Similarly, the strain can also be expanded as a power-series approximation in ε:

$$\boldsymbol{\epsilon}^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) = \boldsymbol{\epsilon}^0(\mathbf{x},\mathbf{y},t) + \varepsilon_0\boldsymbol{\epsilon}^1(\mathbf{x},\mathbf{y},t) + \varepsilon_0^2\boldsymbol{\epsilon}^2(\mathbf{x},\mathbf{y},t) + \dots \tag{7}$$

where $\boldsymbol{\epsilon}^i$ are the strain homogenised expansion coefficients. Note that the exponent i is the power for ε_0 and an index for $\mathbf{u}^i$ and $\boldsymbol{\epsilon}^i$ in equations (6) and (7). We can also define a homogenised expansion for the rotational field $\omega = \frac{1}{2}\boldsymbol{\nabla}\times\mathbf{u}$:

$$\omega^{\varepsilon_0}(\mathbf{x},\mathbf{y},t) = \omega^0(\mathbf{x},\mathbf{y},t) + \varepsilon_0\omega^1(\mathbf{x},\mathbf{y},t) + \varepsilon_0^2\omega^2(\mathbf{x},\mathbf{y},t) + \dots . \tag{8}$$

The true $\mathbf{u}$, $\boldsymbol{\epsilon}$ and ω values are then found via:

$$\mathbf{u}(\mathbf{x},t) = \mathbf{u}^{\varepsilon_0}(\mathbf{x},\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0},t) , \tag{9}$$

$$\boldsymbol{\epsilon}(\mathbf{x},t) = \boldsymbol{\epsilon}^{\varepsilon_0}(\mathbf{x},\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0},t)\, and \tag{10}$$

$$\omega(\mathbf{x},t) = \omega^{\varepsilon_0}(\mathbf{x},\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0},t) . \tag{11}$$

The main results of the homogenisation development are the following:

- The displacement of the leading-order approximation (in ε_0) is independent of the small-scale variable $\mathbf{y}$:

$$\mathbf{u}^*(\mathbf{x},\mathbf{y},t) = \mathbf{u}^*(\mathbf{x},t) . \tag{12}$$

- $\mathbf{u}^*$ is the solution of the effective wave equation:

$$\rho^*\partial_{tt}\mathbf{u}^* - \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}^* = \boldsymbol{f}_s^* , \tag{13}$$

$$\boldsymbol{\sigma}^* = \mathbf{c}^* : \boldsymbol{\epsilon}^* , \tag{14}$$

where $\boldsymbol{\epsilon}^* = \boldsymbol{\epsilon}(\mathbf{u}^*)$ and $(\rho^*, \mathbf{c}^*)$ are the effective mechanical properties (see below), and $\boldsymbol{f}_s^*$ is the effective source (Burgos *et al.* 2016; Capdeville 2021).

- The first-order homogenised coefficient can be written as follows:

$$\mathbf{u}^1(\mathbf{x}, \mathbf{y}, t) = \boldsymbol{\chi}(\mathbf{x}, \mathbf{y}) : \boldsymbol{\epsilon}^*(\mathbf{x}, t) + \left\langle \mathbf{u}^1 \right\rangle (\mathbf{x}, t)\,, \tag{15}$$

where $\boldsymbol{\chi}$ is the first-order corrector (a third-order tensor) and $\langle \mathbf{u}^1 \rangle$ can generally be neglected.

- The strain of the leading-order approximation depends on the small-scale variable $\mathbf{y}$:

$$\boldsymbol{\epsilon}^0(\mathbf{x}, \mathbf{y}, t) = (\mathbf{I} + \boldsymbol{\epsilon}_{\mathbf{y}}(\boldsymbol{\chi})(\mathbf{x}, \mathbf{y})) : \boldsymbol{\epsilon}^*(\mathbf{x}, t)\,, \tag{16}$$

where $\mathbf{I}$ is the fourth-order identity tensor and $[\boldsymbol{\epsilon}_{\mathbf{y}}]_{ij}(\mathbf{v}) = \frac{1}{2}(\partial_{y_i} v_j + \partial_{y_j} v_i)$ for any $\mathbf{v}$.

- Similarly, the leading-order rotational coefficient can be written as:

$$\omega^0(\mathbf{x}, \mathbf{y}, t) = \omega^*(\mathbf{x}, t) + \frac{1}{2} \left(\boldsymbol{\nabla}_{\mathbf{y}} \times \boldsymbol{\chi}(\mathbf{x}, \mathbf{y}) \right) : \boldsymbol{\epsilon}^*(\mathbf{x}, t)\,, \tag{17}$$

where $\omega^* = \frac{1}{2} \boldsymbol{\nabla} \times \mathbf{u}^*$. Note that ω^* and ω^0 are generally different.

- The homogenised quantities of $\rho^*(\mathbf{x})$ and $\mathbf{c}^*(\mathbf{x})$, as well as the corrector $\boldsymbol{\chi}(\mathbf{x}, \mathbf{y})$, can be computed through the homogenised operator $\boldsymbol{\mathcal{H}}$:

$$(\rho^*, \mathbf{c}^*, \chi) = \boldsymbol{\mathcal{H}}(\rho, \mathbf{c})\,. \tag{18}$$

$\boldsymbol{\mathcal{H}}$ implies solving a set of equations known as the "cell problem," which is a set of static elasticity equations with a set of source terms (Capdeville *et al.* 2010b; Capdeville *et al.* 2015). The solution only depends on ε_0 and the elastic model; neither the time nor the seismic source incluence $\boldsymbol{\mathcal{H}}$. 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. $\boldsymbol{\mathcal{H}}$ 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. 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. $\varepsilon_0 \leq 0.5$ is considered a reasonable choice to obtain homogenised quantities and preserve the solution accuracy for most geological media (Capdeville *et al.* 2010b).

Finally, we can also find the relationship between the effective fields, and the true rotational and gradient of the wavefield. The leading-order approximation of the wavefield gradient is:

$$\begin{aligned} &\boldsymbol{\nabla}\mathbf{u}(\mathbf{x},t) = \boldsymbol{\nabla}\mathbf{u}^*(\mathbf{x},t) + \mathbf{G}(\mathbf{x}) : \boldsymbol{\epsilon}^*(\mathbf{x},t) + O(\varepsilon_0)\,, \\ &\text{where: } \mathbf{G}(\mathbf{x}) = \boldsymbol{\nabla}_{\mathbf{y}}\boldsymbol{\chi}(\mathbf{x},\mathbf{y})|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\,, \end{aligned} \tag{19}$$

and $\boldsymbol{\nabla}_{\mathbf{y}}$ is the gradient operator that is applied to the $\mathbf{y}$ variable.

The leading-order approximation of the rotational field is:

$$\begin{aligned} &\omega(\mathbf{x},t) = \omega^*(\mathbf{x},t) + \mathbf{J}(\mathbf{x}) : \boldsymbol{\epsilon}^*(\mathbf{x},t) + O(\varepsilon_0)\,, \\ &\text{where: } \mathbf{J}(\mathbf{x}) = \frac{1}{2}\boldsymbol{\nabla}_{\mathbf{y}} \times \boldsymbol{\chi}(\mathbf{x},\mathbf{y})|_{\mathbf{y}=\frac{\mathbf{x}}{\varepsilon_0}}\,. \end{aligned} \tag{20}$$

In practice, once the effective medium $(\rho^*, \mathbf{c}^*)$ is obtained, the effective wave equation (equations (13)–(14)), 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 (12), this is not the case for the strain and rotational wavefields respectively. Equations (19) and (20) show that an extra term is needed to obtain the correct leading-order approximations for the effective 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).

2.3 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 $\boldsymbol{\mathcal{M}}$, 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 $\boldsymbol{\mathcal{M}}$ is infinite-dimensional.

We can define the homogenised counterpart of each model $\mathbf{m} \in \boldsymbol{\mathcal{M}}$ at a given frequency band as $\mathbf{p}^* = \boldsymbol{\mathcal{H}}(\mathbf{m})$. Equation (18) 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 (20)). We therefore define the effective model space $\boldsymbol{\mathcal{M}}^*$, which contains all of the possible effective models $\mathbf{p}^*$. Symbolically, we have $\boldsymbol{\mathcal{M}}^* = \boldsymbol{\mathcal{H}}(\boldsymbol{\mathcal{M}})$, where $\boldsymbol{\mathcal{M}}^*$ is finite-dimensional.

2.4 Full waveform inversion based on displacement data

We consider an idealised FWI problem whose objective is to recover the true mechanical properties (ρ^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 $\boldsymbol{\mathcal{M}}$. 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 (1)–(2)). We define the least-squares misfit function and associated minimisation problem as:

$$\begin{aligned} E(\mathbf{m}) &= \sum_{r,s} \int_0^T (\mathbf{d}_s(\mathbf{x}_r, t) - \mathbf{u}(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}))^2 dt \,, \\ \bar{\mathbf{m}} &= \underset{\mathbf{m} \in \boldsymbol{\mathcal{M}}}{\operatorname{argmin}} \, E(\mathbf{m}) \,, \end{aligned} \tag{21}$$

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 (21) is generally an unsolvable

problem because $\boldsymbol{\mathcal{M}}$ 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:

$$E^{f_\mathrm{m}}(\mathbf{m}) = \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_\mathrm{m}}(\mathbf{x}_r, t) - \mathbf{u}^{f_\mathrm{m}}(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}))^2 dt \,, \tag{22}$$

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 $\boldsymbol{\mathcal{M}}^h$, whereby at least one model that minimises (22) can be found. Furthermore, $\boldsymbol{\mathcal{M}}^h$ can be designed in such a way that the wave equation (equations (1)–(2)) can be solved in a bounded time for every one of its models. This assumption and model design yield the following inverse problem:

$$\bar{\mathbf{m}}^h = \underset{\mathbf{m} \in \boldsymbol{\mathcal{M}}^h}{\operatorname{argmin}} E^{f_\mathrm{m}}(\mathbf{m}) \tag{23}$$

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 $\boldsymbol{\mathcal{M}}^h$ 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 $\boldsymbol{\mathcal{M}}^h$ 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 $\boldsymbol{\mathcal{M}}^*$, as defined in the previous section. We therefore define a new inverse problem for a fixed f_{m}:

$$
\begin{aligned}
E^*(\mathbf{m}^*) &= \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_{\mathrm{m}}}(\mathbf{x}_r, t) - \mathbf{u}^*(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}^*))^2 dt \,, \\
\bar{\mathbf{m}}^* &= \underset{\mathbf{m}^* \in \boldsymbol{\mathcal{M}}^*}{\operatorname{argmin}} E^*(\mathbf{m}^*) \,,
\end{aligned}
\tag{24}
$$

where $\mathbf{u}^*$ is the leading-order homogenised displacement, which is defined in equation (6). Here, we only use the displacement observations and zeroth-order homogenisation, which implies that no corrector is needed in the definition of $\boldsymbol{\mathcal{M}}^*$. 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 (12), $\mathbf{u}^*$ is insensitive to small-scale heterogeneities.

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

$$
E^{f_{\mathrm{m}}}(\mathbf{m}) = E^*(\mathbf{m}^*) + \mathcal{O}(\varepsilon_0) \,. \tag{25}
$$

The true $\mathbf{m}_t$ minimises E and also $E^{f_{\mathrm{m}}}$; therefore, $\mathbf{m}_t^* = \boldsymbol{\mathcal{H}}(\mathbf{m}_t)$ is a solution of the homogenisation FWI (HFWI) problem (24). 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:

$$
\bar{\mathbf{m}}^* = \boldsymbol{\mathcal{H}}(\mathbf{m}_t) \,. \tag{26}
$$

This important result shows we can access the effective true model by solving (24), which is a solvable inverse problem. In practice, we can only set up an explicit parametrisation of $\boldsymbol{\mathcal{M}}^*$ in the layered model case. We need to rely on an approximate space model $\boldsymbol{\mathcal{M}}^{*h}$ for the general case and use the homogenisation operator $\boldsymbol{\mathcal{H}}$ to project the solution into $\boldsymbol{\mathcal{M}}^*$. In practice, the construction of $\boldsymbol{\mathcal{M}}^{*h}$ is very similar to a classical construction of $\boldsymbol{\mathcal{M}}^h$. However, we need to have $\boldsymbol{\mathcal{M}}^* \subset \boldsymbol{\mathcal{H}}(\boldsymbol{\mathcal{M}}^{*h})$, which often implies that $\boldsymbol{\mathcal{M}}^{*h}$ 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.

2.5 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:

$$\begin{aligned} E^{f_m}(\mathbf{m}) &= \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_m}(\mathbf{x}_r, t) - \omega^{f_m}(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}))^2 dt \,, \\ \bar{\mathbf{m}}^h &= \underset{\mathbf{m} \in \boldsymbol{\mathcal{M}}^h}{\operatorname{argmin}} E^{f_m}(\mathbf{m}) \,. \end{aligned} \tag{27}$$

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:

$$\begin{aligned} E^*(\mathbf{m}^*) &= \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_\mathrm{m}}(\mathbf{x}_r, t) - \omega^*(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{m}^*))^2 dt \,, \\ \bar{\mathbf{m}}^* &= \underset{\mathbf{m}^* \in \boldsymbol{\mathcal{M}}^*}{\operatorname{argmin}} E^*(\mathbf{m}^*) \,. \end{aligned} \tag{28}$$

However, $\omega^{f_\mathrm{m}} \neq \omega^*$ for the leading-order approximation, which means that $E^{f_\mathrm{m}}(\mathbf{m}) \neq E^*(\mathbf{m}^*)$. 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^*:

$$\begin{aligned} E^*(\mathbf{p}^*) &= \sum_{r,s} \int_0^T (\mathbf{d}_s^{f_\mathrm{m}}(\mathbf{x}_r, t) - \omega^0(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*))^2 dt \,, \\ \bar{\mathbf{p}}^* &= \underset{\mathbf{p}^* \in \boldsymbol{\mathcal{M}}^*}{\operatorname{argmin}} E^*(\mathbf{p}^*) \,, \end{aligned} \tag{29}$$

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 (20). Here $\boldsymbol{\mathcal{M}}^*$ 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 $E^{f_\mathrm{m}}(\mathbf{m}) = E^*(\mathbf{p}^*) + O(\varepsilon_0)$, which makes sure that we can relate the inverse problem solution $\bar{\mathbf{m}}^* = \bar{\mathbf{p}}^*|_\mathbf{m}$ to the effective true model $\mathbf{m}_t^*$.

3 INVERSION SCHEME

We use the same inversion framework as in Capdeville & Métivier (2018), whereby we spatially parametrise $\boldsymbol{\mathcal{M}}^{*h}$ by dividing the squared inversion domain $\boldsymbol{\mathcal{I}} \in \boldsymbol{\Omega}$ into non-overlapping $n \times n$ square elements: $\boldsymbol{\mathcal{I}} = \cup_{e=1}^{n^2} \boldsymbol{\mathcal{I}}_e^n$. An example of inversion domain $\boldsymbol{\mathcal{I}}$ with 22×22 elements ($\boldsymbol{\mathcal{I}}_e^{22}$) is shown in Fig. 1. A $N \times N$ degree polynomial basis is then used for each element $\boldsymbol{\mathcal{I}}_e^n$, and each inverted mechanical property is expanded within each element. Here we name such a parametrisation $\boldsymbol{\mathcal{P}}_n^N(\boldsymbol{\mathcal{I}})$. We use a fully anisotropic parametrisation for $\boldsymbol{\mathcal{M}}^{*h}$, 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 $\boldsymbol{\mathcal{M}}^{*h}$ if no correctors are inverted.

3.1 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:

$$\mathbf{m}^{i+1} = \mathbf{m}^i + ((\mathbf{F}^i)^T \cdot \mathbf{F}^i + \boldsymbol{\lambda}^i)^{-1}[(\mathbf{F}^i)^T \cdot (\mathbf{d} - \boldsymbol{\omega}(\mathbf{m}^i))]\,, \tag{30}$$

where $\mathbf{F}^i$ is the partial derivative matrix and $\boldsymbol{\lambda^i}$ is the damping parameter, which is used to stabilise the inversion. The damping parameter $\boldsymbol{\lambda^i}$ 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.

3.2 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:

$$\begin{aligned} E(\mathbf{J}_r) = &\sum_s \int_0^T (\mathbf{d}_s(\mathbf{x}_r, t) - \omega^0(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*))^2 dt \,, \\ &\text{where: } \omega^0(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*) = \omega^*(\mathbf{x}_r, t; \mathbf{x}_s, \mathbf{p}^*|_{\mathbf{m}}) + \mathbf{J}_r : \boldsymbol{\epsilon}^*(\mathbf{u}^*) \,, \end{aligned} \tag{31}$$

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:

$$\mathbf{J}_r = (\mathbf{F}_r^t \mathbf{F}_r)^{-1} \mathbf{F}_r^t \mathbf{b}, \tag{32}$$

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.2 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.

4 NUMERICAL EXPERIMENTS

4.1 Experiment design

4.1.1 Target model

The target model that the FWI tries to estimate is a 2D elastic medium, such as that shown in figure 1. The Ω domain is a 40×40 km region that consists of a homogeneous background with a 16 $\times$ 16 km heterogeneous area embedded at the centre. This heterogeneous area is made of 40 $\times$ 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.

4.1.2 Spectral element simulation setup

We use the spectral element method (Komatitsch & Vilotte 1998) to solve the wave equations (equations (1)-(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 $\times$ 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 $\times$ 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.

4.1.3 FWI acquisition geometry

The source and receiver geometries used in our numerical experiments are shown in Fig. 3. 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. 3): The receivers are located around the inverted area and are far from the heterogeneities.
- Configuration B (Fig. 3): 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).

4.1.4 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 $\mathbf{c}^{*,\text{iso}}$ following Browaeys & Chevrot (2004). The compressional- and shear-wave velocities are defined as follows:

$$V_P^*(\mathbf{x}) = \sqrt{(c_{1111}^{*,\text{iso}}(\mathbf{x})/\rho(\mathbf{x}))}\,, \tag{33}$$

and:

$$V_S^*(\mathbf{x}) = \sqrt{(c^{*,\text{iso}}_{1212}(\mathbf{x})/\rho(\mathbf{x}))}\,, \tag{34}$$

respectively, and the total anisotropy is defined as:

$$\text{aniso}(\mathbf{x}) = \frac{\sqrt{\sum_{ijkl}(c^{\text{iso}}_{ijkl}(\mathbf{x}) - c_{ijkl}(\mathbf{x}))^2}}{\sqrt{\sum_{ijkl}(c^{\text{iso}}_{ijkl}(\mathbf{x}))^2}}\,. \tag{35}$$

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:

$$E^2_m(\varepsilon_0, \mathbf{x}) = \frac{\sum_{n=1}^{n=N}\sum_{ij}(c^{*,n}_{ij}(\mathbf{x}) - c^{*,t}_{ij}(\mathbf{x}))^2}{N\sum_{ij}(c^{*,t}_{ij}(\mathbf{x}))^2}\,, \tag{36}$$

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:

$$\bar{E}^2_m(\varepsilon_0) = \frac{\int_{\boldsymbol{\mathcal{I}}}\sum_{n=1}^{n=N}\sum_{ij}(c^{*,n}_{ij}(\mathbf{x}) - c^{*,t}_{ij}(\mathbf{x}))^2 d\mathbf{x}}{N\int_{\boldsymbol{\mathcal{I}}}\sum_{ij}(c^{*,t}_{ij}(\mathbf{x}))^2 d\mathbf{x}}\,, \tag{37}$$

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 (33) and (34), 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:

$$\mathrm{Cov}^{\varepsilon_0}(\mathbf{x},\mathbf{y}) = \frac{\sum_{n=1}^{N}(p_n^*(\mathbf{x})-\mu_p(\mathbf{x}))(p_n^*(\mathbf{y})-\mu_p(\mathbf{y}))}{N}\,, \tag{38}$$

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:

$$\sigma^{\varepsilon_0}(\mathbf{x}) = \sqrt{\mathrm{Cov}^{\varepsilon_0}(\mathbf{x},\mathbf{x})}\,, \tag{39}$$

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

$$R^{\varepsilon_0}(\mathbf{x},\mathbf{y}) = \frac{\mathrm{Cov}^{\varepsilon_0}(\mathbf{x},\mathbf{y})}{\sigma^{\varepsilon_0}(\mathbf{x})\sigma^{\varepsilon_0}(\mathbf{y})}\,. \tag{40}$$

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.

4.2 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. 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 (20). Conversely, the rotational waveforms from the four receivers at location B, which is $\sim$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.

4.3 FWI results without correctors

4.3.1 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. 3).

The displacement-based inversion results, which are presented in Fig. 5, 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.

The rotational inversion results are shown in Fig. 6. 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 (20) and demonstrated in Fig.

4. 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.

Our assessment of the optimal ε_0 value, which is based on equations (37) and (37), is shown in Fig. 7. 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 $\varepsilon_0 = 1.0$ 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 $\varepsilon_0 = 1.0$ as the optimal choice that provides the best overall resolution and accuracy of the final model results.

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 $\varepsilon_0 = 1.0$ and shown in Fig. 8. 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 (40). 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.

4.3.2 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. 3). The forward modelling test presented in section 4.2 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. 9, 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 12).

The rotational-based inversion results are presented in Fig.10. 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.

The trade-off between the error of the mean model $\bar{E}_m$ and ε_0 are shown in Fig.11. 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.12 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 dis-

tributed. Furthermore, the spatial resolutions for the rotational case are no longer valid, as there are strong correlations at distant locations.

4.4 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 3). The correctors are obtained during the inversion based on equation (32). 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.13, 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.14 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.11). 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.15 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.

Table 1 provides a summary of the misfits, which are defined by equation (37) for a single model, in each case. Here we use single solution, as opposed to the mean model, to compute

Table 1. Model misfits, which are based on equation (37), 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

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.

Table 2. Model misfit based on equation (37) 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

5 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 from 52 sensors (26 rotational and 26 displacement sensors) for the inversion. We utilised a least-squares misfit function by combining equations (22) and (29) 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 (37)), which are based on $\varepsilon_0 = 1$, and presented in table 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.

6 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 $\times$ 1.8 km area (Fig. 16) for this numerical test; the absolute values of the inverted correctors at the final iteration for each component are illustrated in Fig. 16. 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.

7 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.17 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.

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 $\boldsymbol{\mathcal{M}}^{*h}$ that we have been using contain small-scale heterogeneities (continuity between elements is not enforced). Therefore, the correctors shown in Fig. 16 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 (6) 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.

DATA AVAILABILITY

This study used only synthetic data.

REFERENCES

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.

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.

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.

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.

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.

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

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

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

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

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).

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

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

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.

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

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.

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.

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

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

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

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

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.

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.

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.

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.

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.

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

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.

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.

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

Appendix A: Adjoint partial derivative calculations

A.1 Preliminary remarks

Here we use elastic wave equation in the frequency domain for simplicity:

$$-\omega^2\rho(\mathbf{x})u_i(\mathbf{x},\omega) - \partial_j c_{ijkl}(\mathbf{x})\partial_k u_l(\mathbf{x},\omega) = f_i(\mathbf{x},\omega)\,. \tag{A.1}$$

The Green's function for a single point force at $\mathbf{x}_s$ is described by $f_i^p(\mathbf{x},\omega) = \delta(\mathbf{x}-\mathbf{x}_s)\delta_{ip}$, which satisfies the equation:

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

The Green's function for a moment component point source $f_i^{pq}(\mathbf{x},\omega) = -\delta_{ip}\delta_{jq}\partial_j\delta(\mathbf{x}-\mathbf{x}_s)$ satisfies the equation:

$$-\omega^2\rho(\mathbf{x})G_i^{pq}(\mathbf{x},\omega;\mathbf{x}_s) - \partial_j c_{ijkl}(\mathbf{x})\partial_k G_l^{pq}(\mathbf{x},\omega;\mathbf{x}_s) = -\delta_{ip}\delta_{jq}\partial_j\delta(\mathbf{x}-\mathbf{x}_s)\,. \tag{A.3}$$

Finally, we note that G_i^M for the moment tensor force described in equation (4) can be written as:

$$-\omega^2\rho(\mathbf{x})G_i^{\mathbf{M}}(\mathbf{x},\omega;\mathbf{x}_s) - \partial_j c_{ijkl}(\mathbf{x})\partial_k G_l^{\mathbf{M}}(\mathbf{x},\omega;\mathbf{x}_s) = -M_{ij}\partial_j\delta(\mathbf{x}-\mathbf{x}_s)\,. \tag{A.4}$$

We then omit ω to denote the reciprocity theorem as follows:

$$G_i^j(\mathbf{x}_r;\mathbf{x}_s) = G_j^i(\mathbf{x}_s;\mathbf{x}_r) \tag{A.5}$$

and:

$$\partial_j^r G_i^{pq}(\mathbf{x}_r;\mathbf{x}_s) = \partial_p^s G_q^{ij}(\mathbf{x}_s;\mathbf{x}_r) \tag{A.6}$$

where $\partial_j^r = \partial/\partial_{x_{r,j}}$ $\partial_j^s = \partial/\partial_{x_{s,j}}$.

We also have:

$$G_i^{pq}(\mathbf{x}_r;\mathbf{x}_s) = \partial_p^s G_i^q(\mathbf{x}_r;\mathbf{x}_s)\,. \tag{A.7}$$

A.2 Partial derivatives for a single scattering point in $\mathbf{x}$

The Born approximation reads:

$$\delta u_i(\mathbf{x}_r) = -\omega^2 \delta\rho G_i^m(\mathbf{x}_r;\mathbf{x}) G_m^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s) - G_i^m(\mathbf{x}_r;\mathbf{x}) \partial_j \delta c_{mjkl} \partial_k G_l^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s)\,. \tag{A.8}$$

Equation (A.8) can be reformulated via integration by parts and the reciprocity principle as follows:

$$\delta u_i(\mathbf{x}_r) = -\omega^2 \delta\rho G_m^i(\mathbf{x};\mathbf{x}_r) G_m^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s) + \partial_j G_m^i(\mathbf{x};\mathbf{x}_r) \delta c_{mjkl} \partial_k G_l^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s)\,. \tag{A.9}$$

We can then deduce the equation for a gradient component:

$$\partial_p^r \delta u_i(\mathbf{x}_r) = -\omega^2 \delta\rho \partial_q^r G_m^i(\mathbf{x};\mathbf{x}_r) G_m^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s) + \partial_j \partial_q^r G_m^i(\mathbf{x};\mathbf{x}_r) \delta c_{mjkl} \partial_k G_l^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s)\,. \tag{A.10}$$

We substitute equation (A.7) into equation (A.11) to obtain:

$$\partial_q^r \delta u_i(\mathbf{x}_r) = -\omega^2 \delta\rho G_m^{iq}(\mathbf{x};\mathbf{x}_r) G_m^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s) + \partial_j G_m^{iq}(\mathbf{x};\mathbf{x}_r) \delta c_{mjkl} \partial_k G_l^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s)\,. \tag{A.11}$$

The coupling tensor $\mathbf{J}$, with $[\mathbf{K}^{qi}]_{pn} = J_{qipn}$, can be written as:

$$J_{qipn} \partial_p^r \delta u_n(\mathbf{x}_r) = -\omega^2 \delta\rho G_m^{\mathbf{K}^{qi}}(\mathbf{x};\mathbf{x}_r) G_m^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s) + \partial_j G_m^{\mathbf{K}^{qi}}(\mathbf{x};\mathbf{x}_r) \delta c_{mjkl} \partial_k G_l^{\mathbf{M}}(\mathbf{x};\mathbf{x}_s)\,. \tag{A.12}$$

We can utilise the coupling tensor $\mathbf{J}$ and Levi-Civita symbol θ to derive the adjoint source of the moment tensor for rotational wavefields as:

$$\begin{aligned} &\mathbf{J}_{qipn} = \theta_{ipn}/2 \quad \text{(in 3D)}\,, \\ &\mathbf{J}_{qipn} = \theta_{pn}/2 \quad \text{(in 2D)}\,. \end{aligned} \tag{A.13}$$

Furthermore, we can derive the adjoint source of the moment tensor for the strain wavefields as follows:

$$\mathbf{J}_{qipn} = (\delta_{qp}\delta_{in} + \delta_{qn}\delta_{ip})/2\,. \tag{A.14}$$

We stored the calculated full wavefields for both sources and adjoint sources on disk and computed the partial derivatives by convolving each source–receiver pair.

Appendix B: 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. A.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. A.2–A.4 to confirm the importance of including correctors to improve the FWI solution.

Appendix C: 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. A.5), and the other configuration involved placing all of the receivers within the heterogeneous area (”Receivers inside” in Fig. A.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.

C1. Receivers outside of the inversion area

Fig. A.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. A.7, with the residual values being comparable to those obtained via the displacement-based FWI.

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

Figures A.8 and A.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.

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. A.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.A.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 3 provides the misfits in the elastic tensors based on equation (37) 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.

Table 3. Model misfit based on equation (37) 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

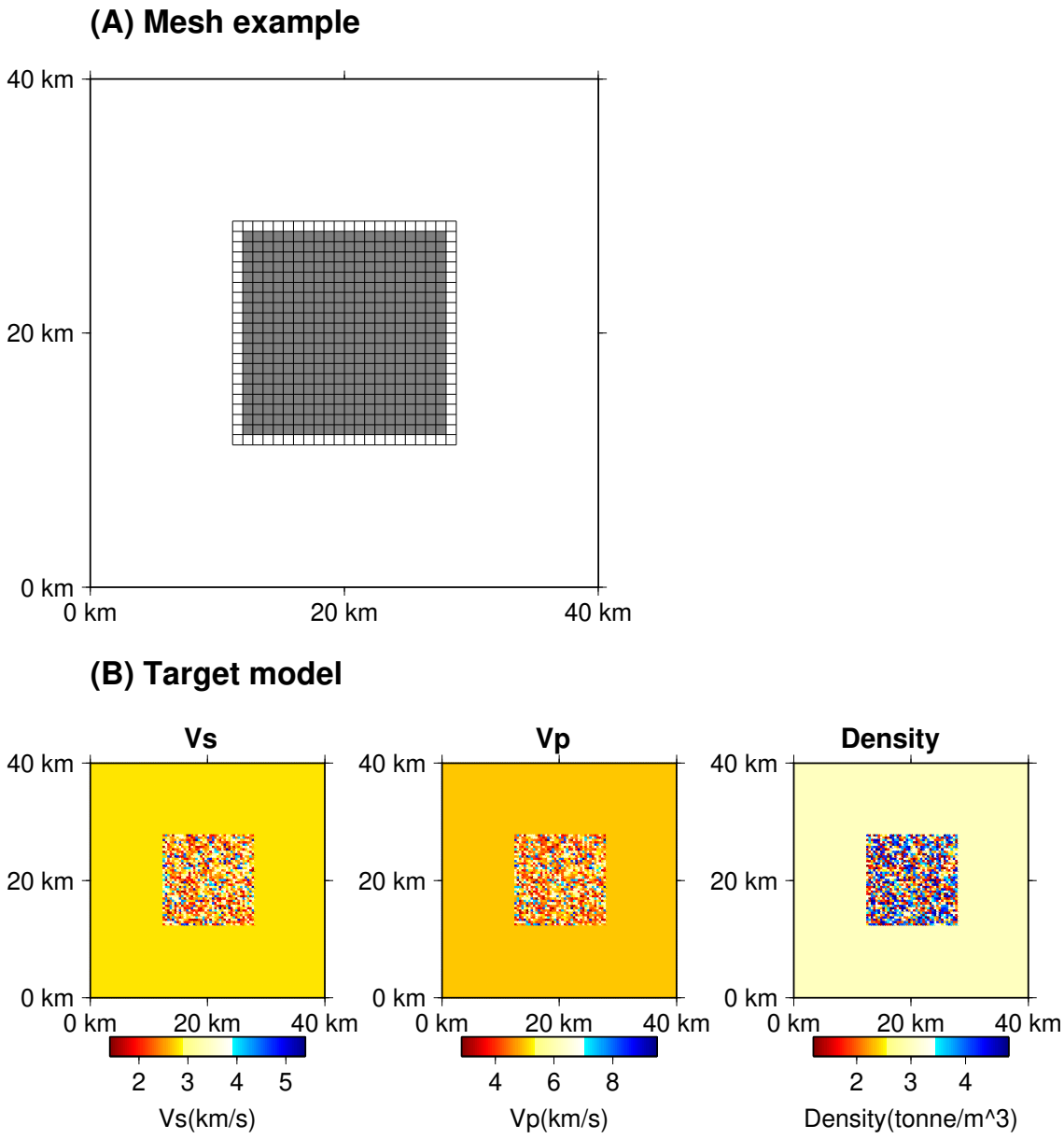

Figure 1. (A) Schematic diagram of the $\mathcal{I}_e^{22}$ mesh structure. The shaded grey area shows the target area to be inverted. Black lines define the element boundaries. (B) V_S (left), V_P (centre) and density (right) structures of the target model.

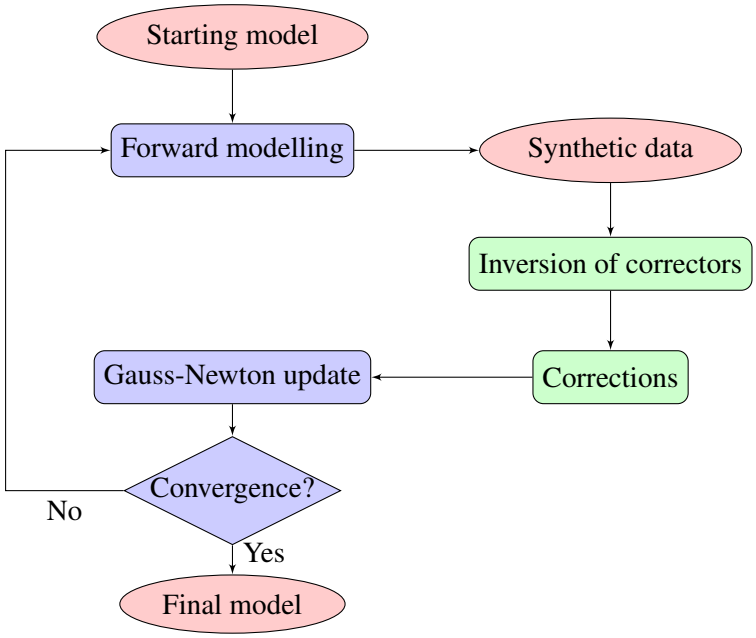

Figure 2. Flowchart of the FWI process, with correctors. The process begins with forwarding modelling to generate synthetic data. The correctors are inverted, and the synthetic data are corrected using the inverted correctors. Then, the inversion of the elastic mechanical properties is performed. If convergence is not obtained, the process is repeated.

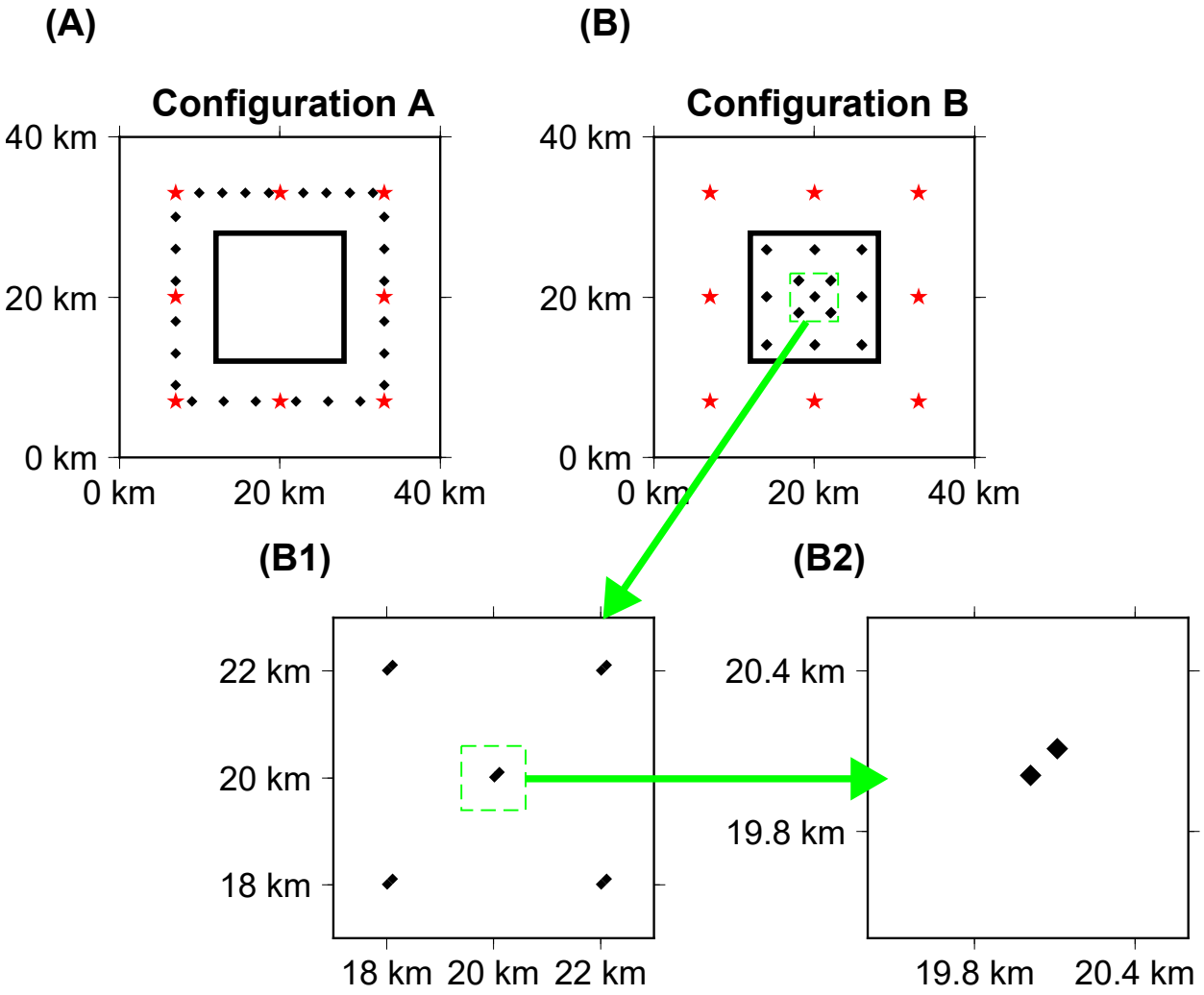

Figure 3. Source (red stars) and receiver (black diamonds) configurations employed for the synthetic tests. (A) Configuration A: All sources and receivers are positioned outside of the heterogeneous area. (B) Configuration B: All sources are positioned outside of the heterogeneous area, whereas all receivers 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.

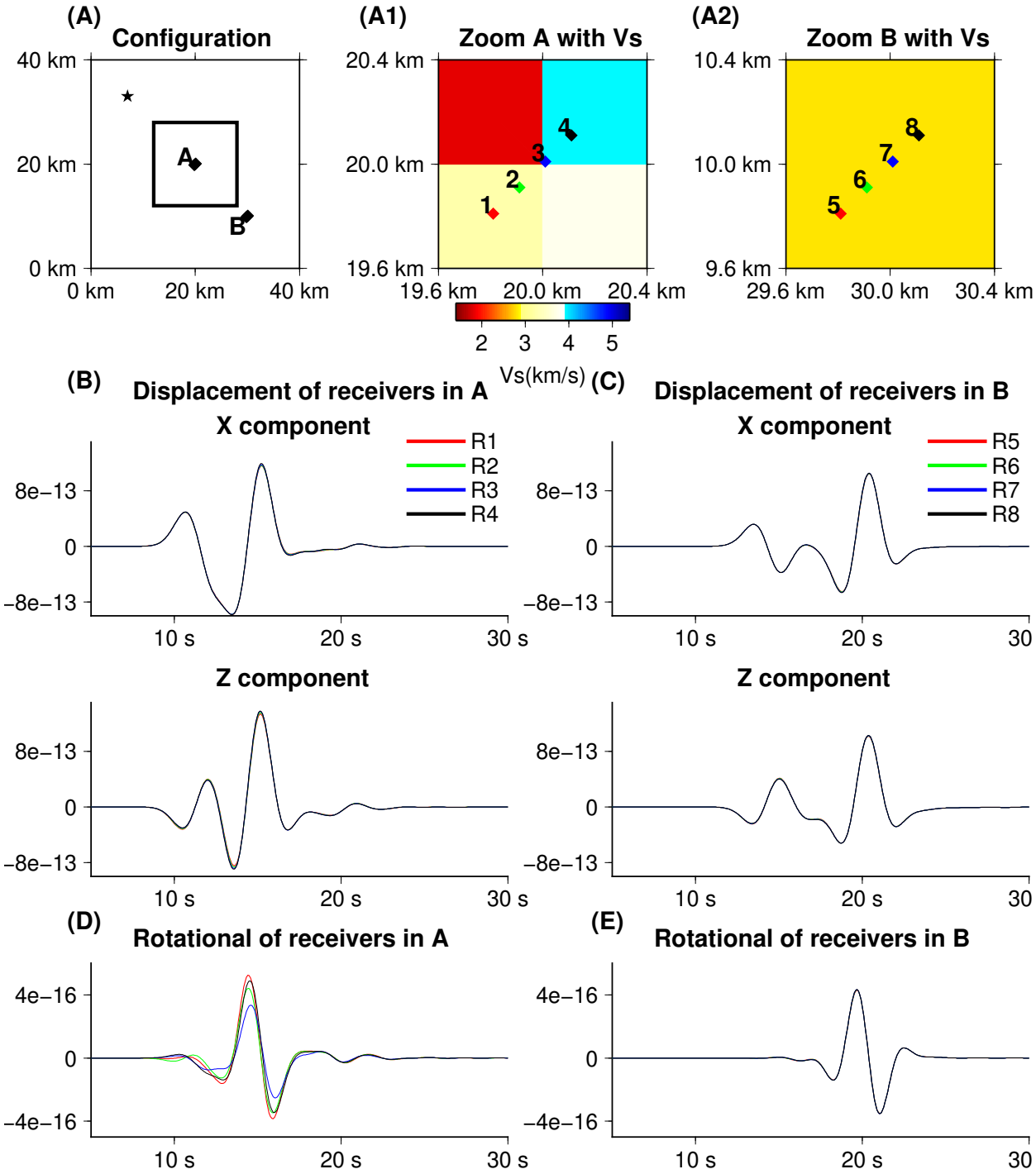

Figure 4. Numerically derived example of the impact of small-scale heterogeneities on displacement and rotational waveforms. (A) Source–receiver positions. Close-up views of the location (A1) A and (A2) B receiver positions, and their local V_S structures. Location A is within the heterogeneous area, 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.

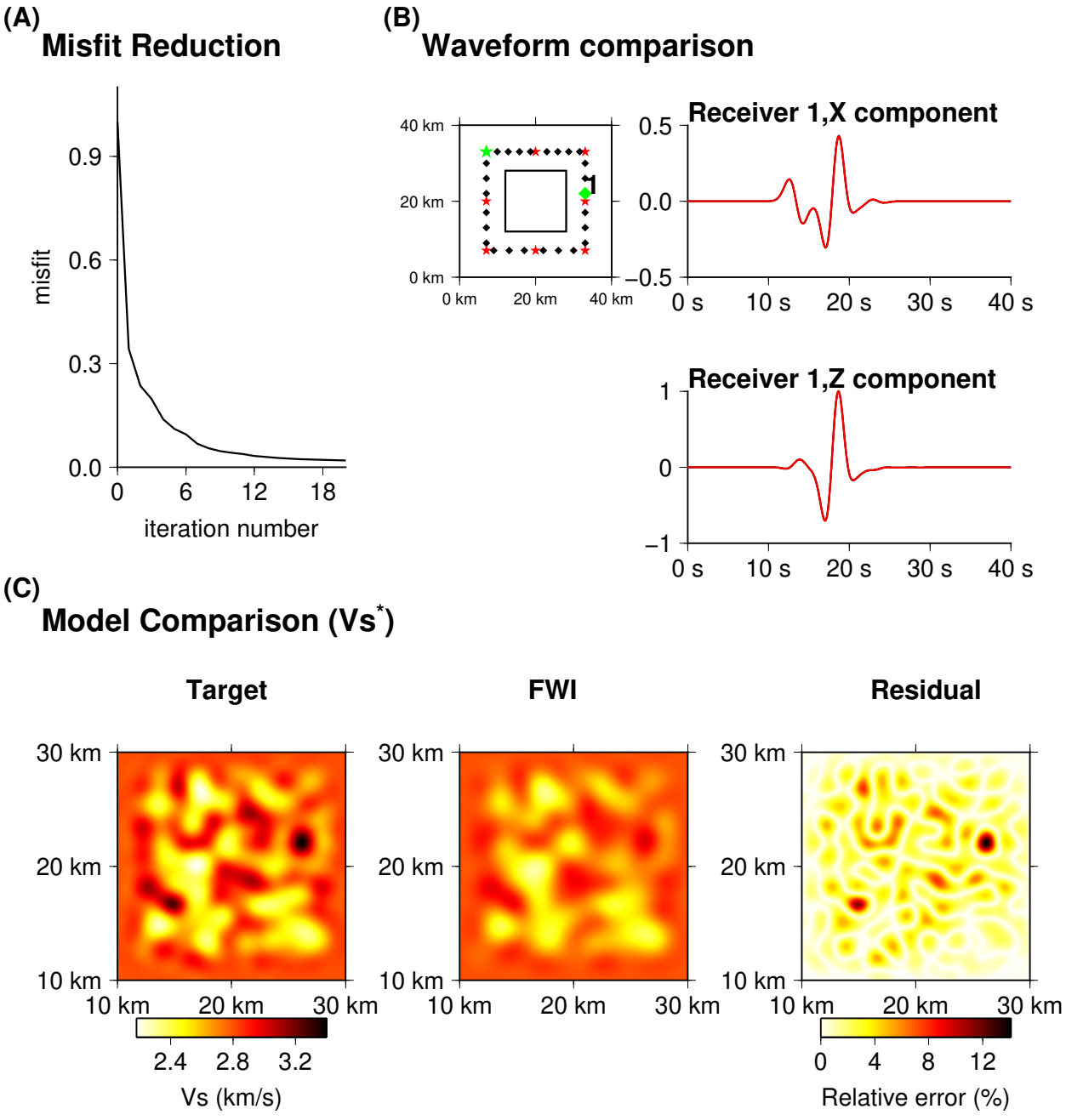

Figure 5. Displacement-based FWI results using receivers positioned outside of the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamond) positions 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.

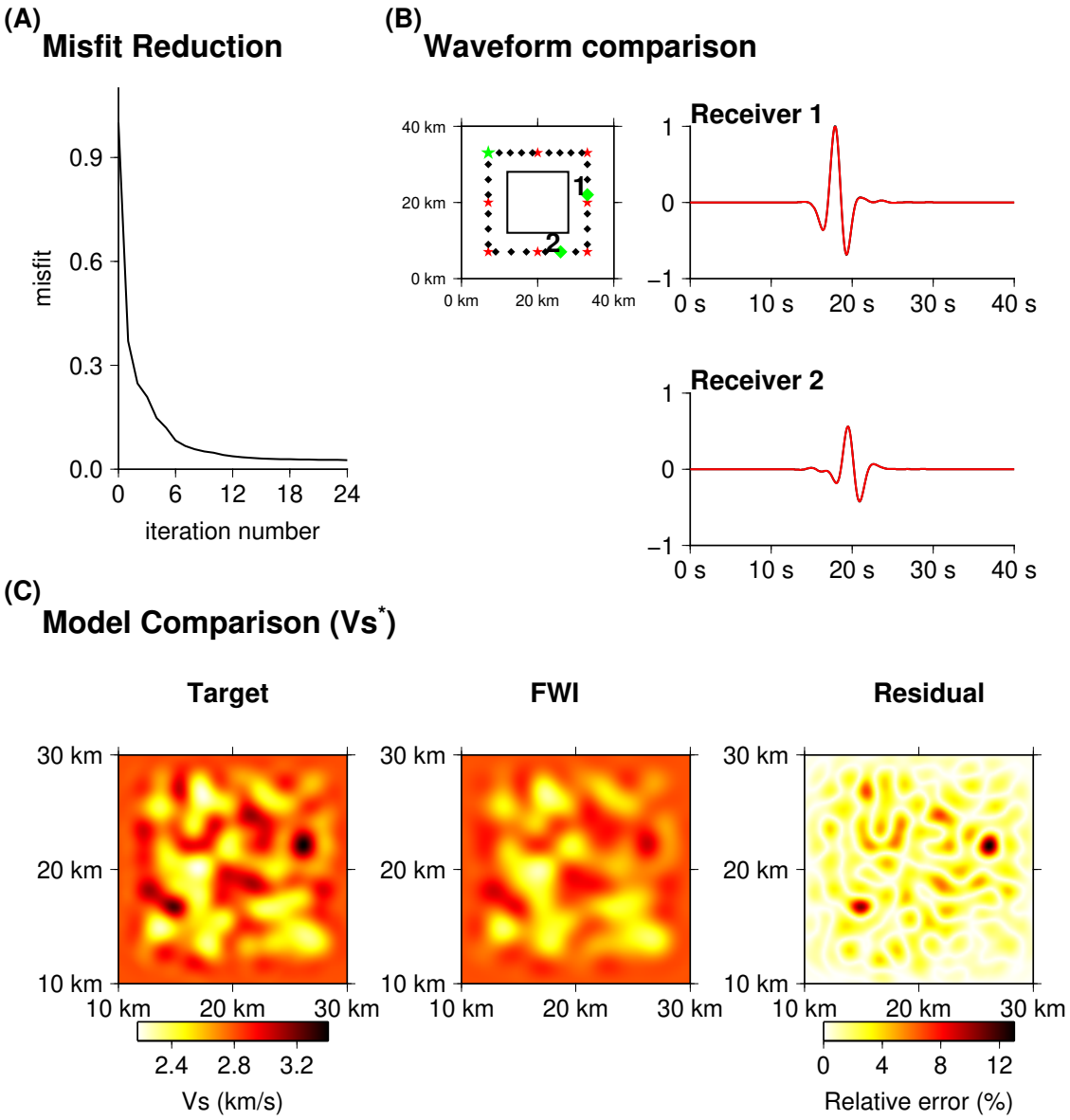

Figure 6. Rotational-based FWI results using receivers positioned outside of the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions 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.

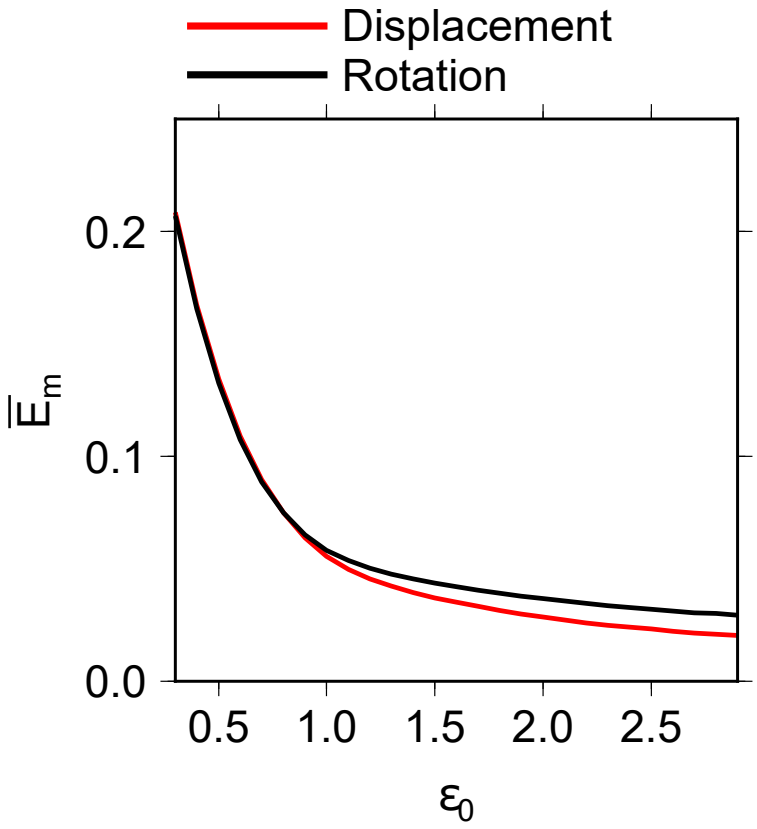

Figure 7. Error $\bar{E}_m$ as a function of ε_0 for V_S^*, which is derived using displacement (red) and rotational (black) receivers positioned outside of the heterogeneous area.

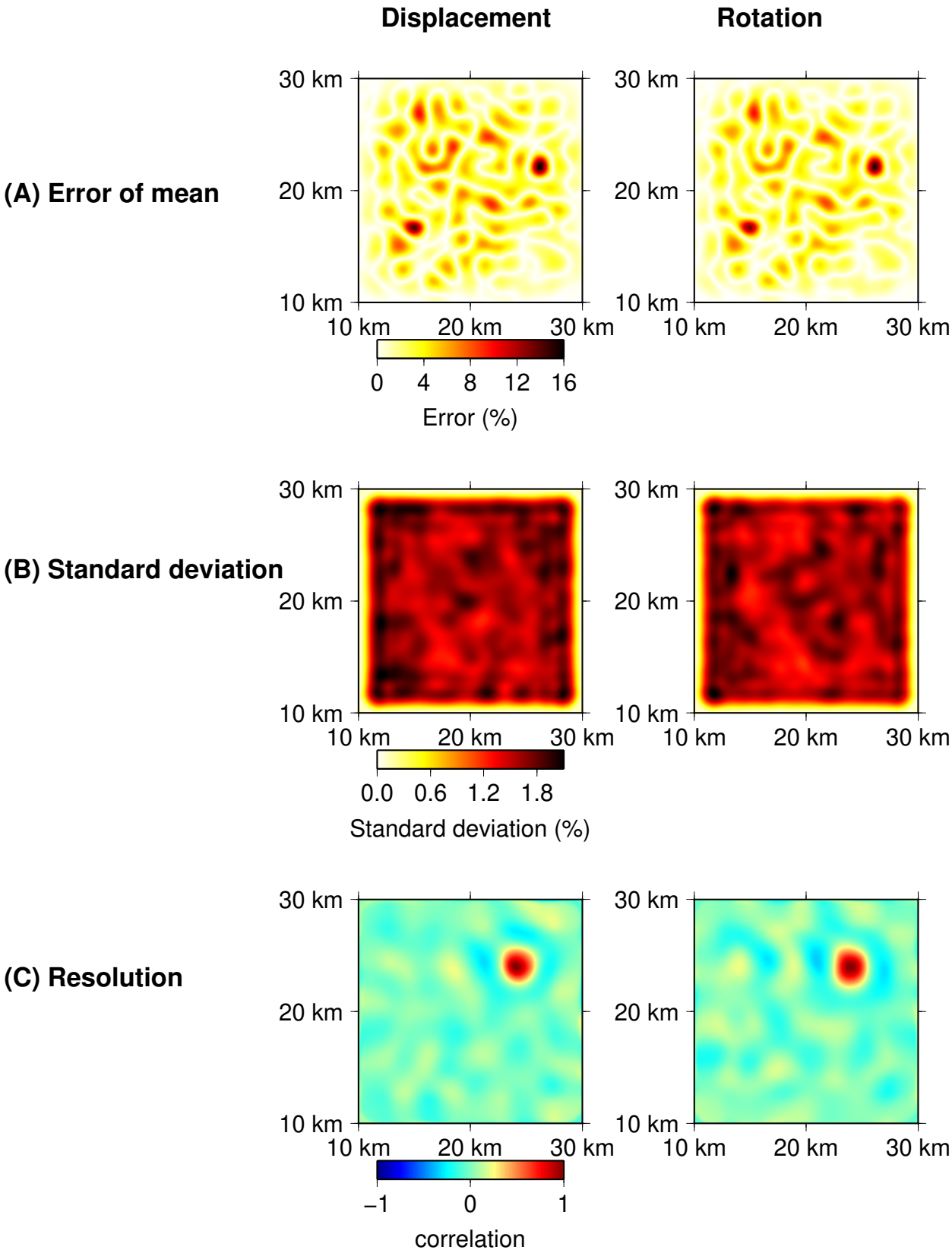

Figure 8. Statistics of the displacement- and rotational-based FWI results using receivers positioned outside of the heterogeneous area. (A) Error of the mean models of V_S^*. (B) Standard deviation of V_S^*. (C) Resolution of V_S^*. The error and standard deviation values are presented as fluctuations from a background value of 2.8 km/s.

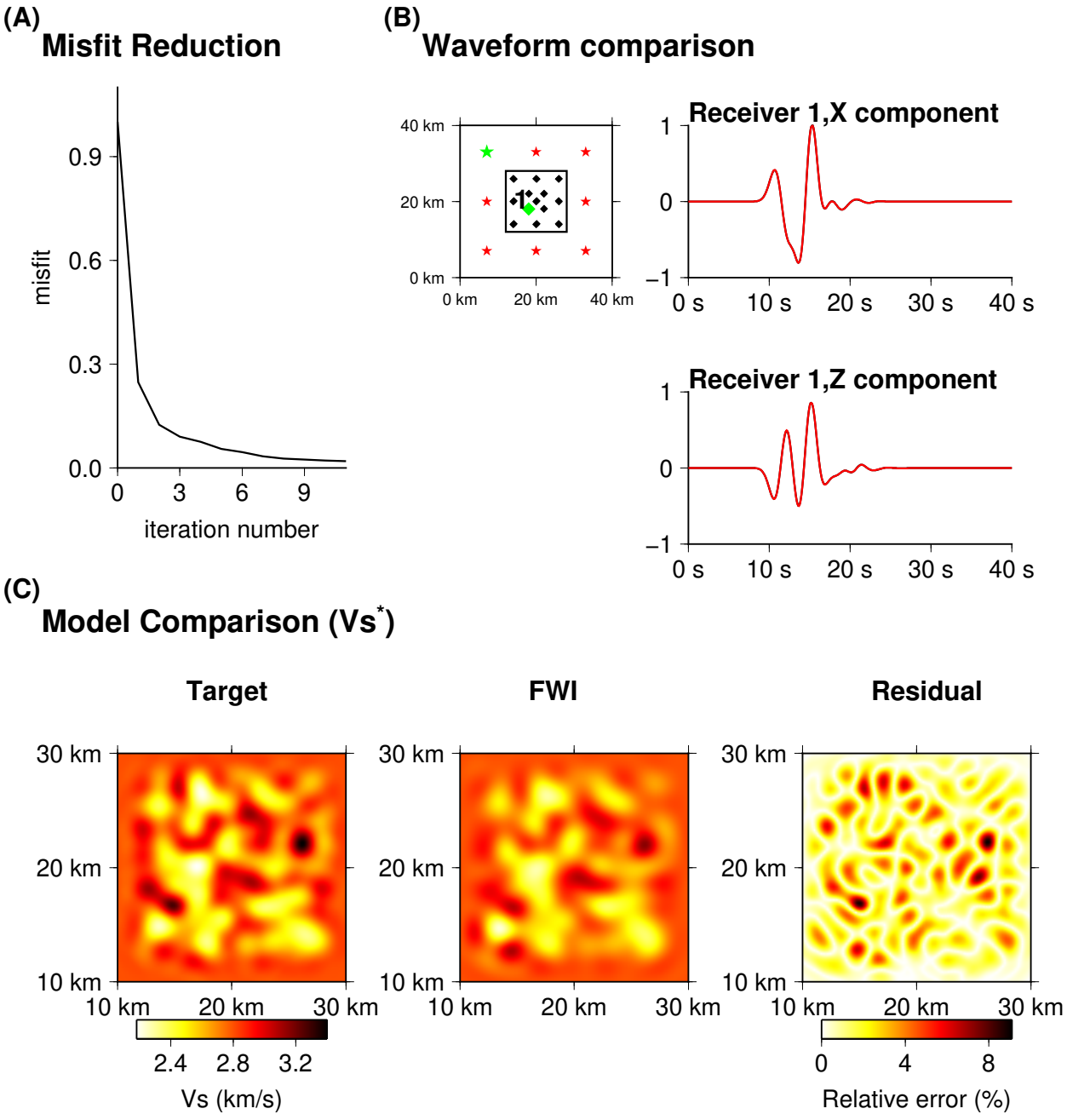

Figure 9. Displacement-based FWI results using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamond) positions 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.

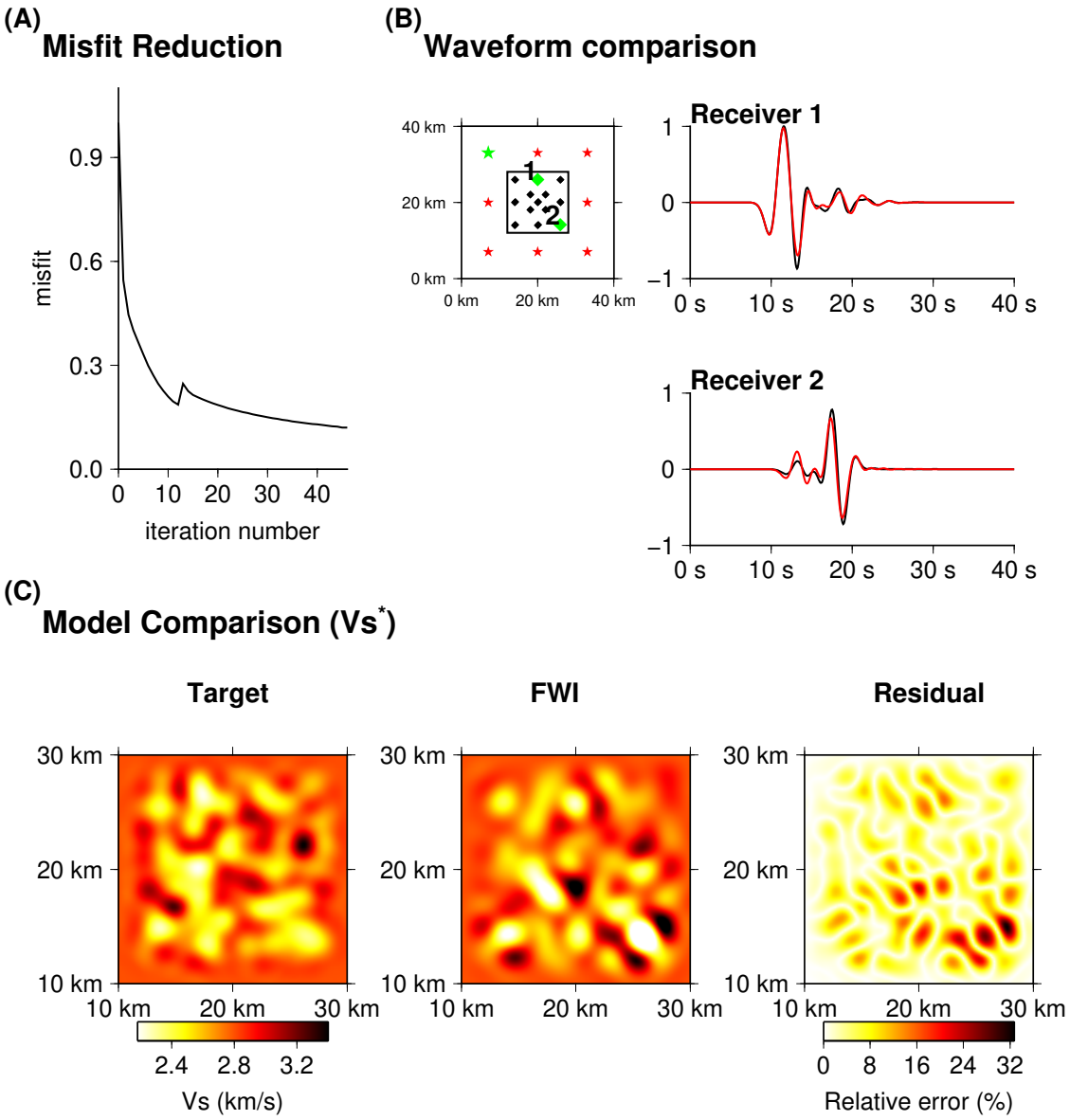

Figure 10. Rotational-based FWI results, without correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (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.

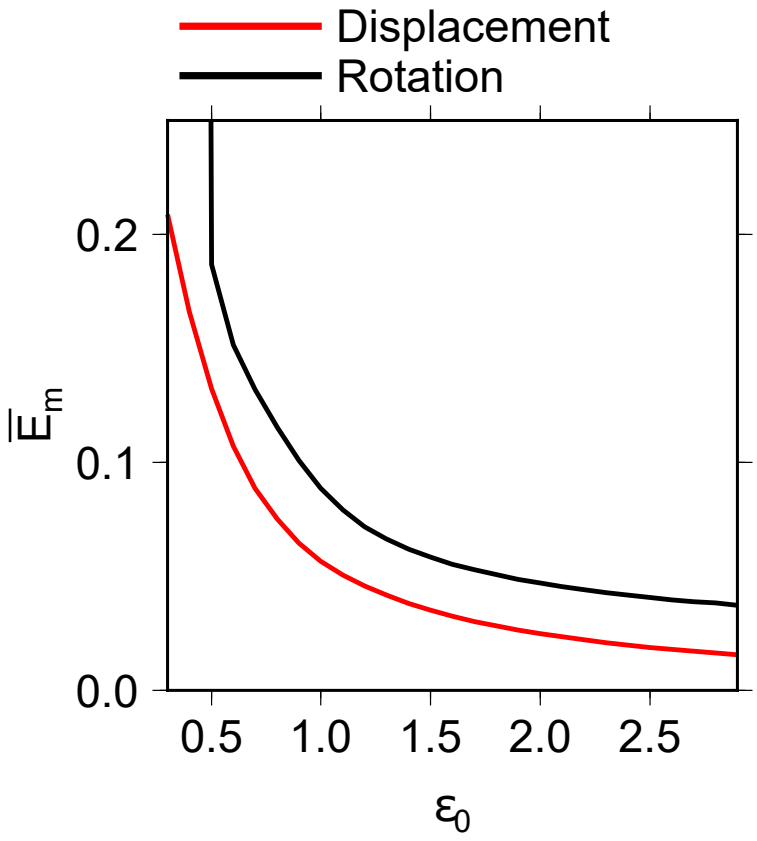

Figure 11. Error $\bar{E}_m$ as a function of ε_0 for V_S^*, which is derived from the FWI results, without correctors, using displacement (red) and rotational (black) receivers positioned within the heterogeneous area.

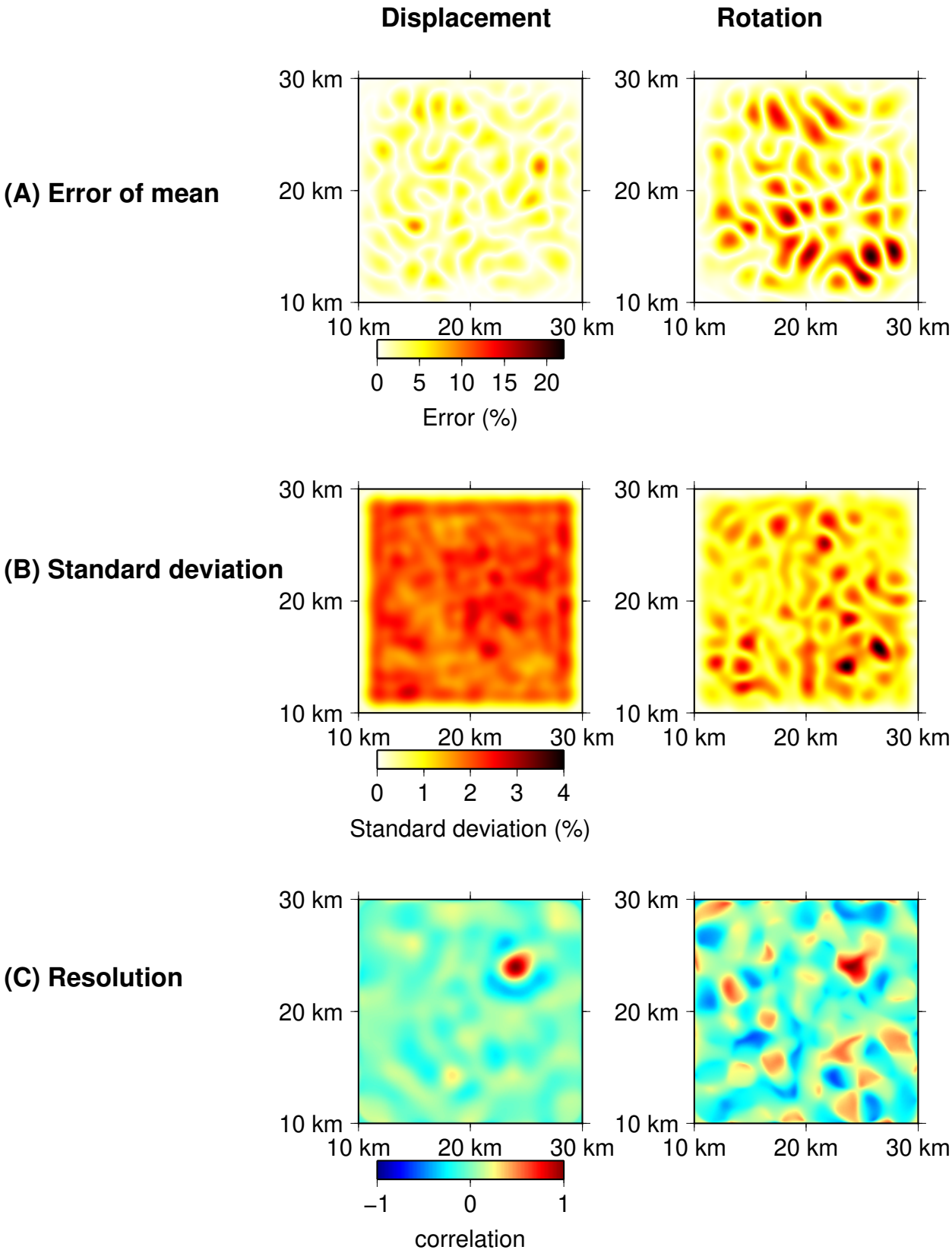

Figure 12. Statistics of the displacement- and rotational-based FWI results, without correctors, using receivers positioned within the heterogeneous area. (A) Error of the mean models of V_S^*. (B) Standard deviation of V_S^*. (C) Resolution of V_S^*. The error and standard deviation values are presented as fluctuations from a background value of 2.8 km/s.

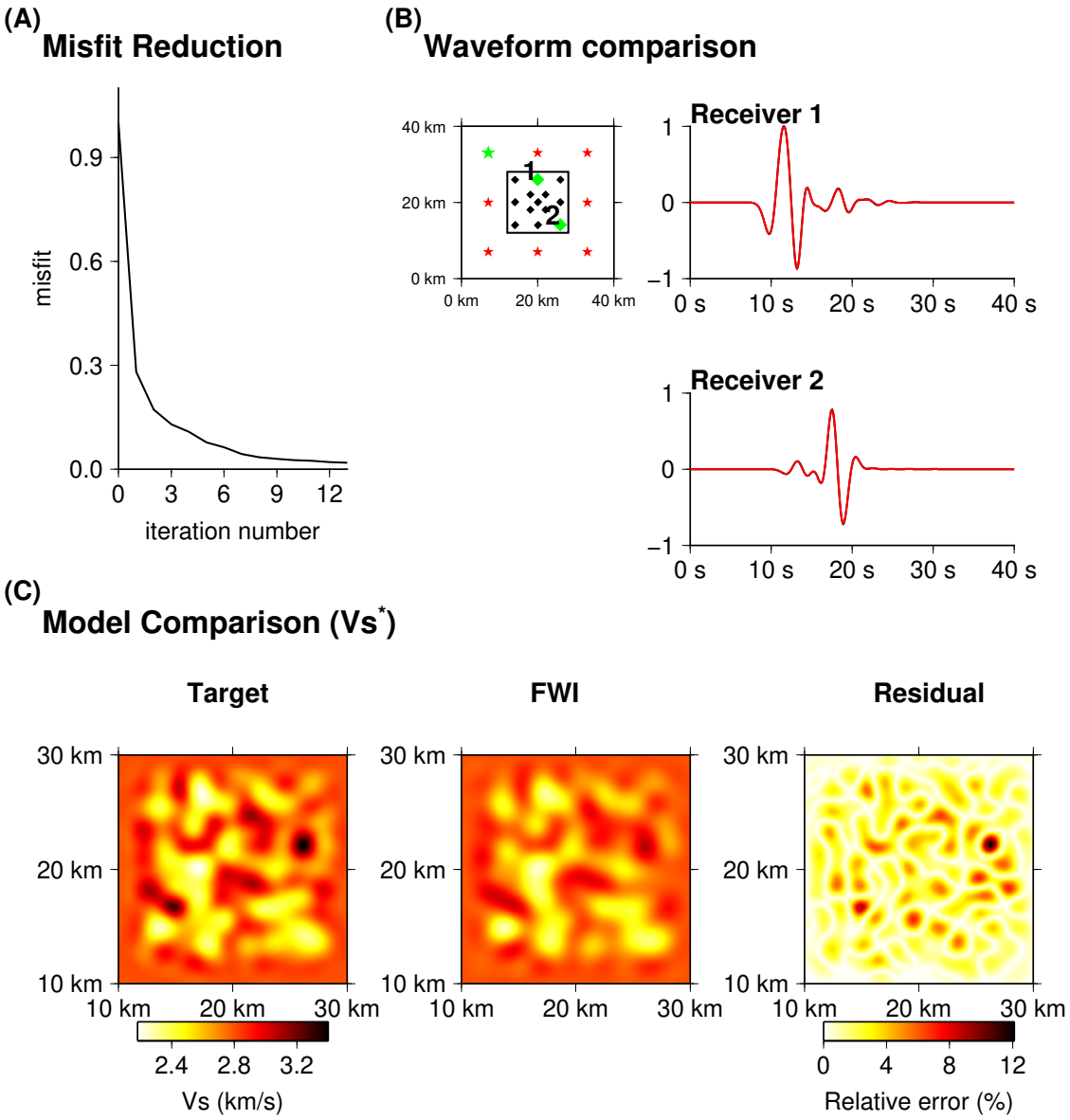

Figure 13. Rotational-based FWI results, with correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) 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.

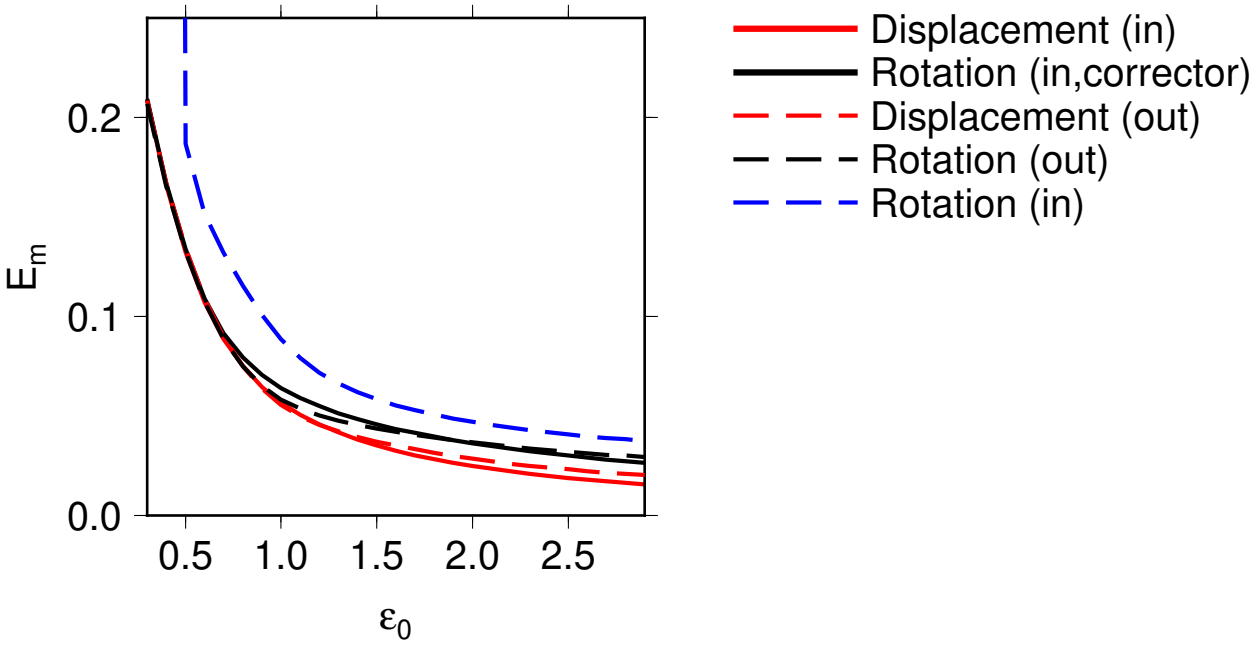

Figure 14. Error $\bar{E}_m$ as a function of ε_0 for V_S^*, which is derived from the FWI results, with correctors, using displacement (solid red) and rotational (solid black) receivers positioned within the heterogeneous area. Results from the other cases (dashed curves) are also shown for comparison.

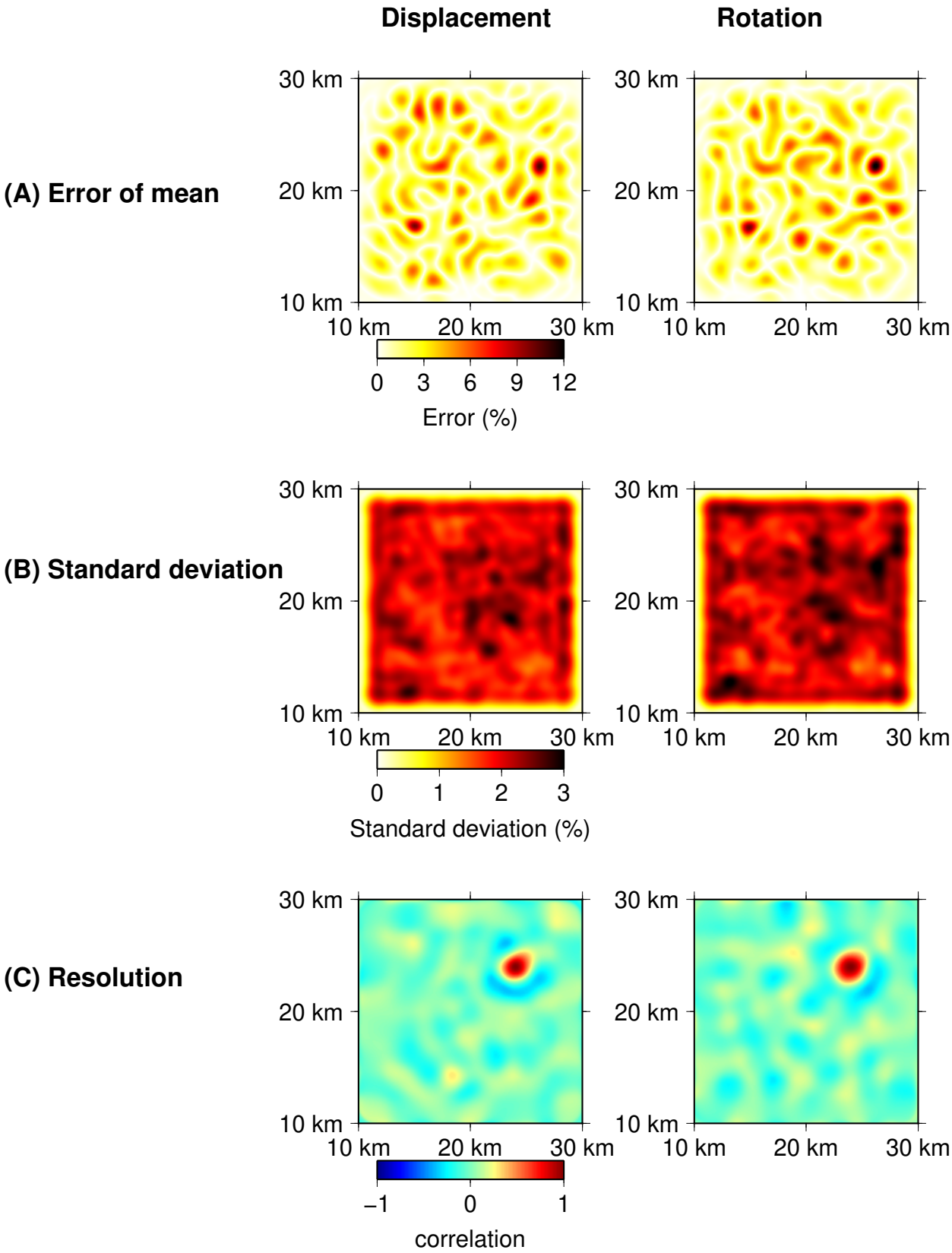

Figure 15. Statistics of the displacement- and rotational-based FWI results, with correctors, using receivers positioned within the heterogeneous area. (A) Error of the mean models of V_S^*. (B) Standard deviation of V_S^*. (C) Resolution of V_S^*. The error and standard deviation values are presented as fluctuations from a background value of 2.8 km/s.

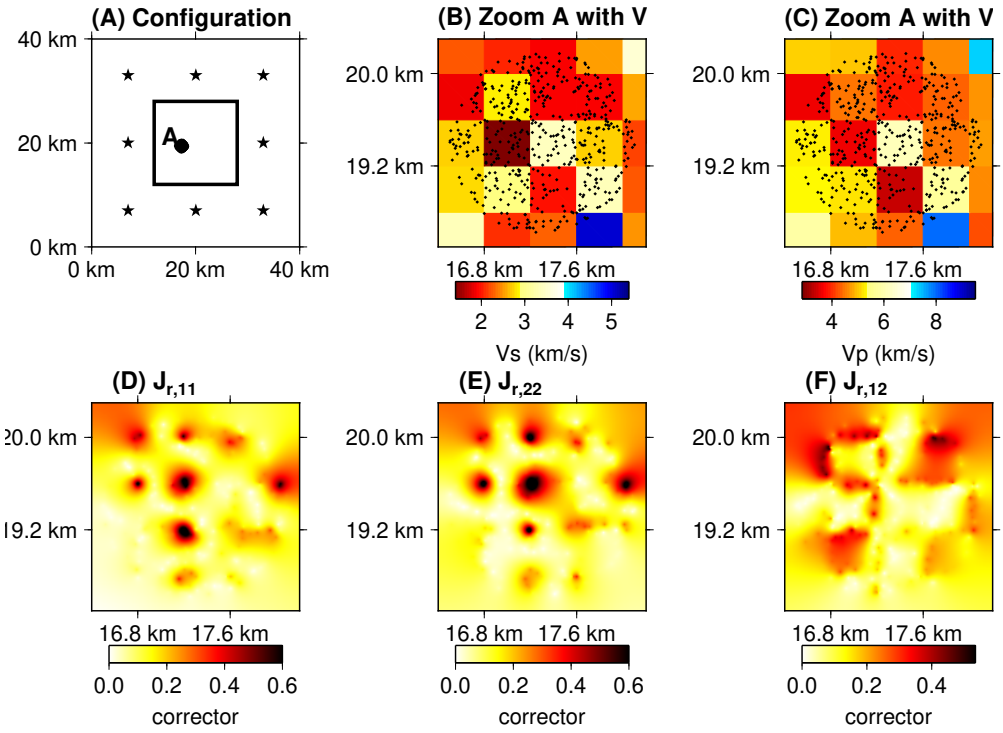

Figure 16. Numerical demonstration to show indications of small-scale structures via the inverted correctors. (A) Source (black stars) and receiver (black diamonds) configuration for the numerical experiment. Five-hundred rotational receivers are positioned at location A. (B) V_S and (C) V_P heterogeneities 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, $J_{r,12}$.

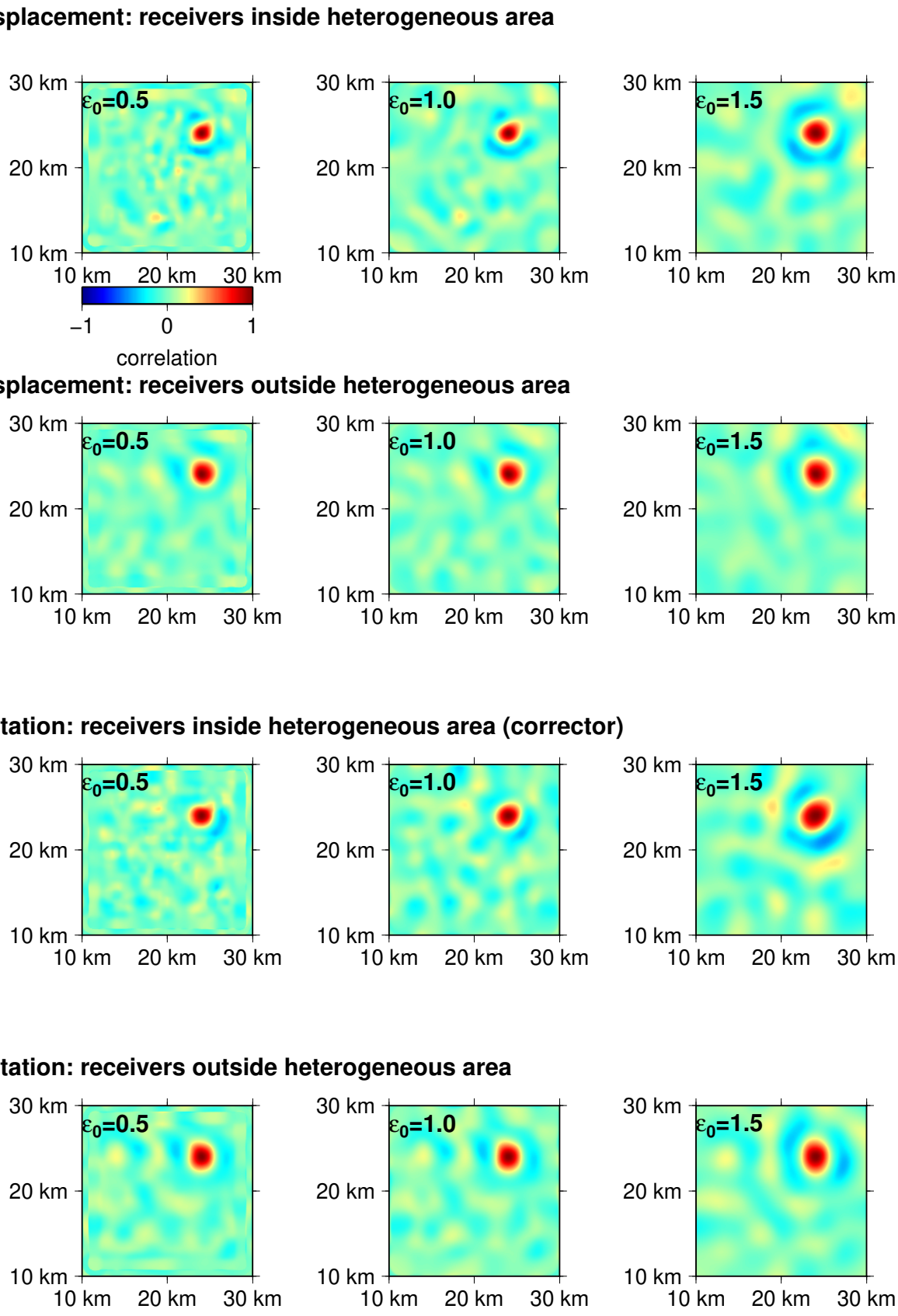

Figure 17. Resolution images of V_S^* for different ε_0. The resolutions for $\varepsilon_0 = 0.5$, $\varepsilon_0 = 1.0$ and $\varepsilon_0 = 1.5$ are shown. The presented results are for displacement receivers positioned (A) within and (B) outside of the heterogeneous area, and rotational receivers positioned (C) within and (D) outside of the heterogeneous area. We used correctors in the FWI process for (C).

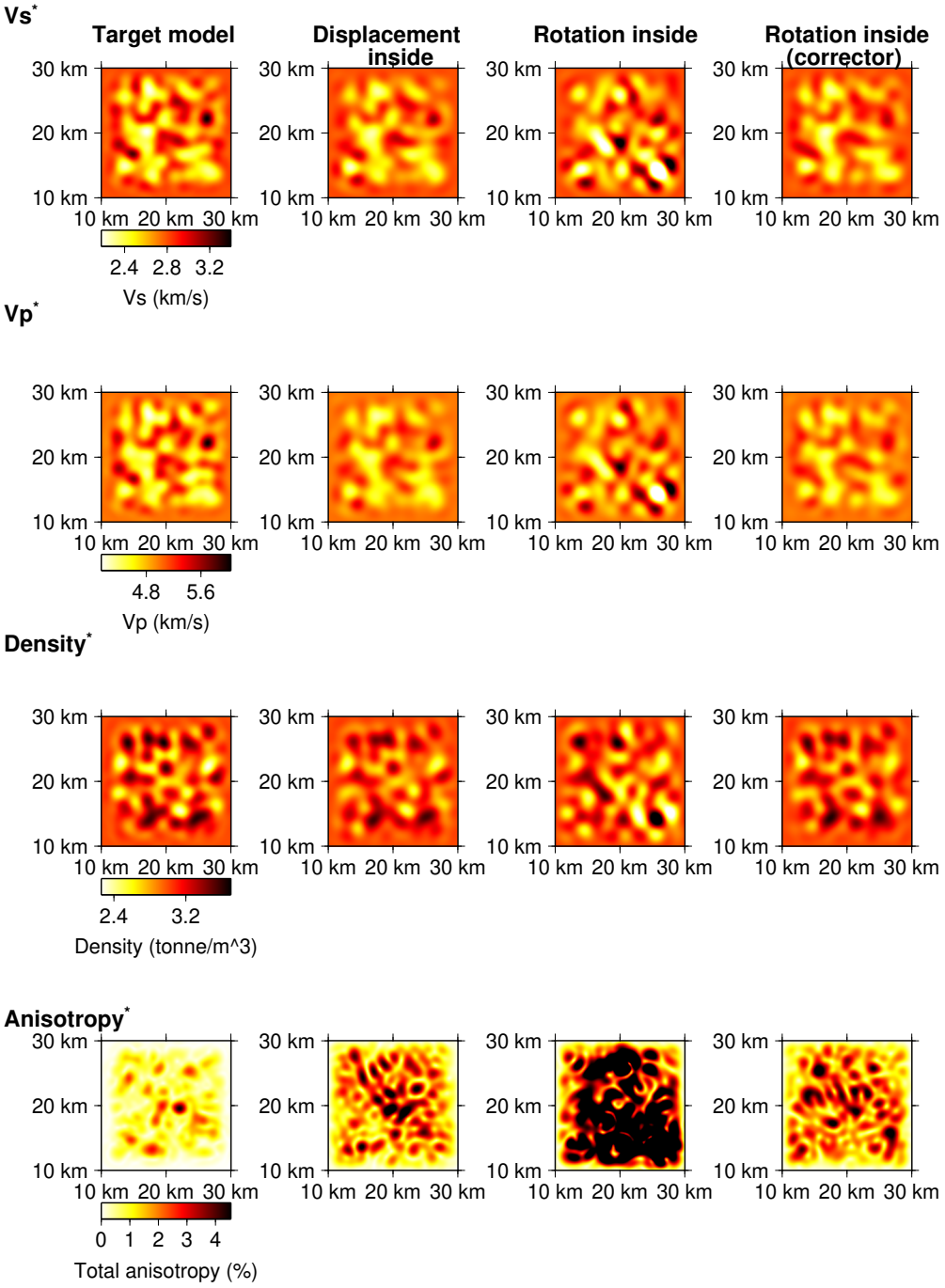

Figure A.1. Comparison of the FWI results. The V_S^*, V_P^*, density* and anisotropy* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Target model (”Target model”). Displacement-based FWI results using receivers positioned within the heterogeneous area (”Displacement inside”). 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)”).

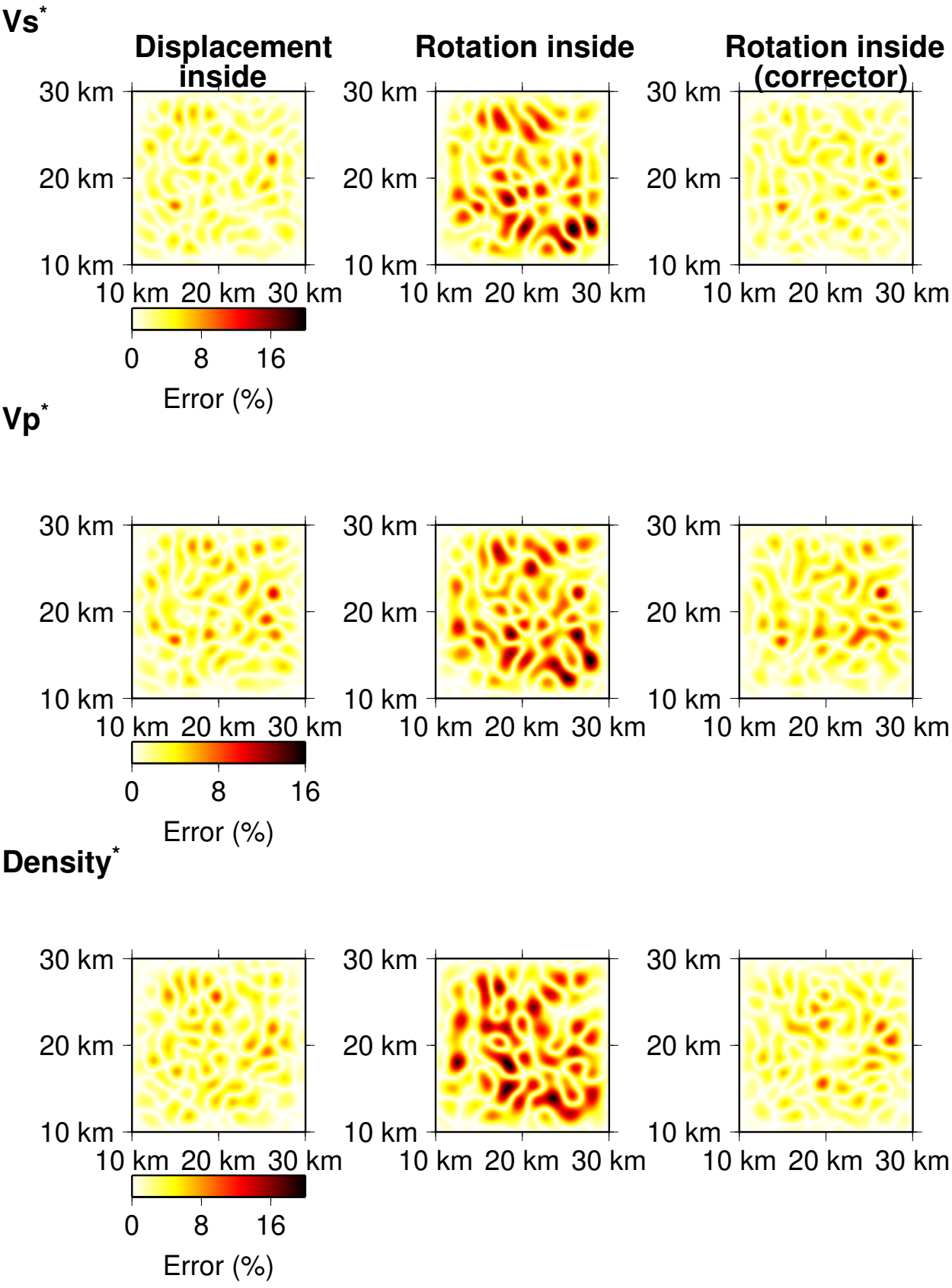

Figure A.2. Comparison of the error of mean models of the FWI results: The V_S^*, V_P^* and density* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Mean model for displacement receivers positioned within the heterogeneous area ("Displacement inside"). Mean model, without correctors,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

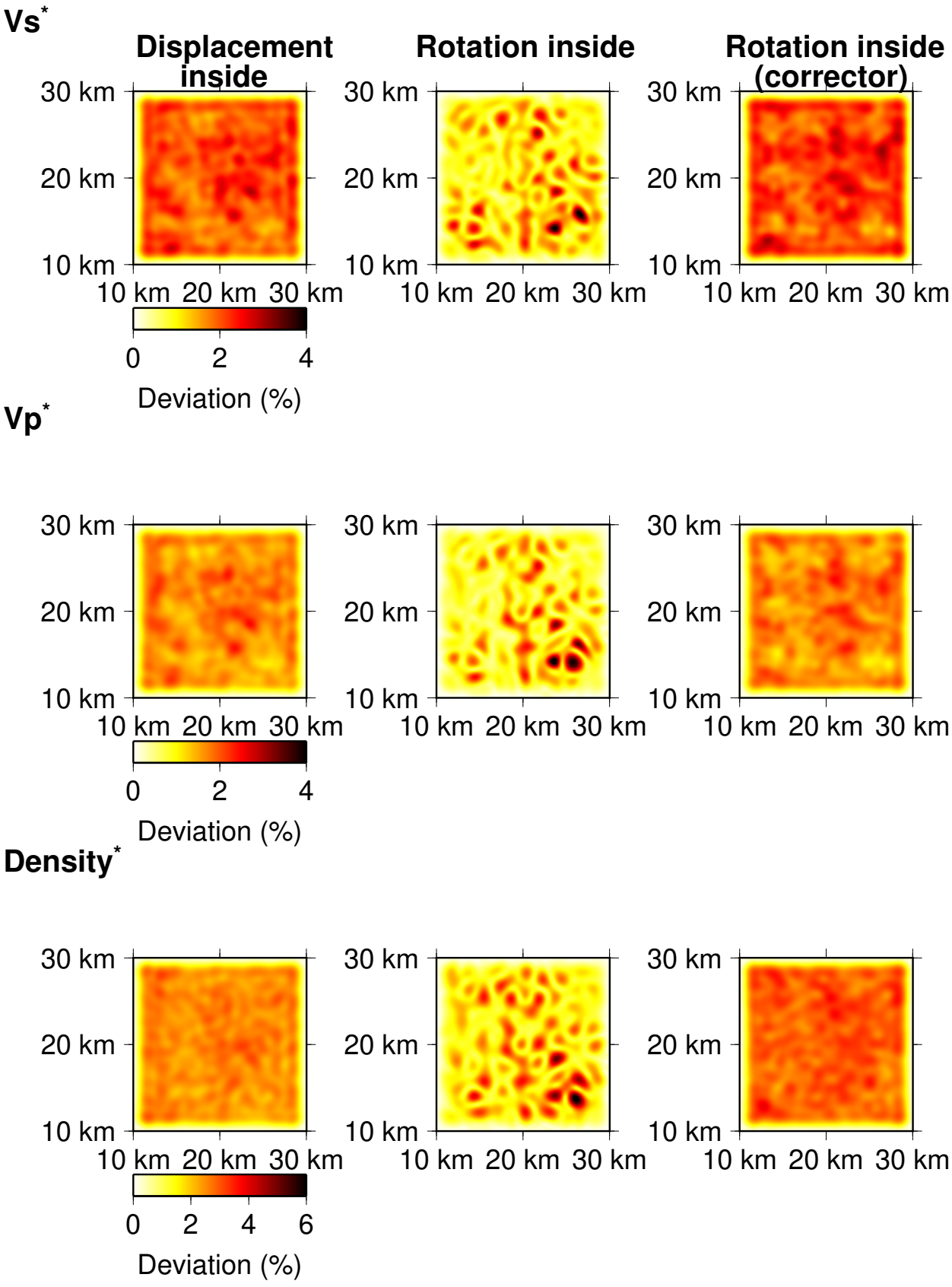

Figure A.3. Comparison of the standard deviations in the FWI results from the target model. The V_S^*, V_P^* and density* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Deviation for displacement receivers positioned within the heterogeneous area ("Displacement inside"). Deviation, 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)").

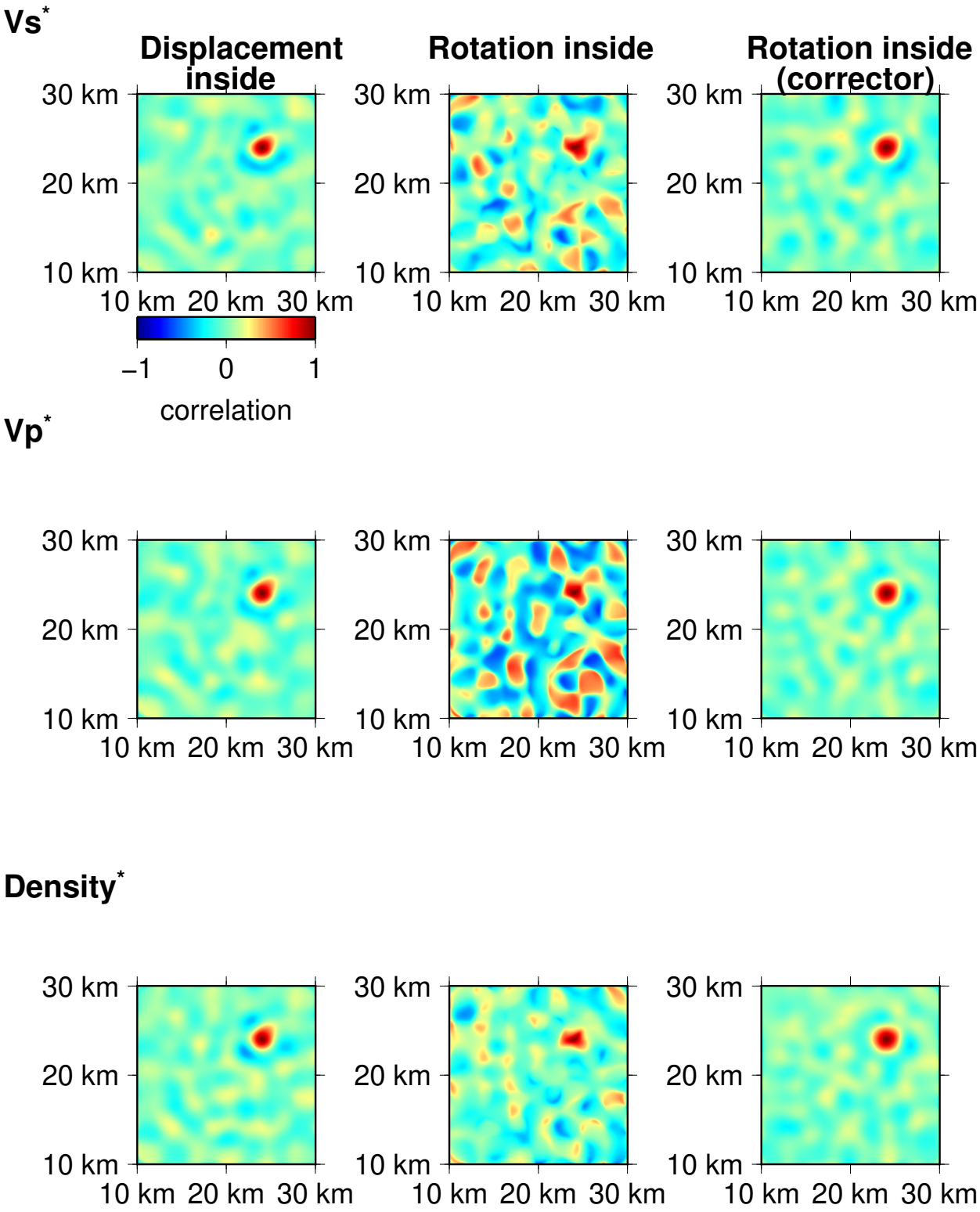

Figure A.4. Comparison of the resolutions of the FWI results. The V_S^*, V_P^* and density* structures, which are homogenised using $\varepsilon_0 = 1.0$, are shown. Resolution for displacement receivers positioned within the heterogeneous area (”Displacement inside”). Resolution, without correctors,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

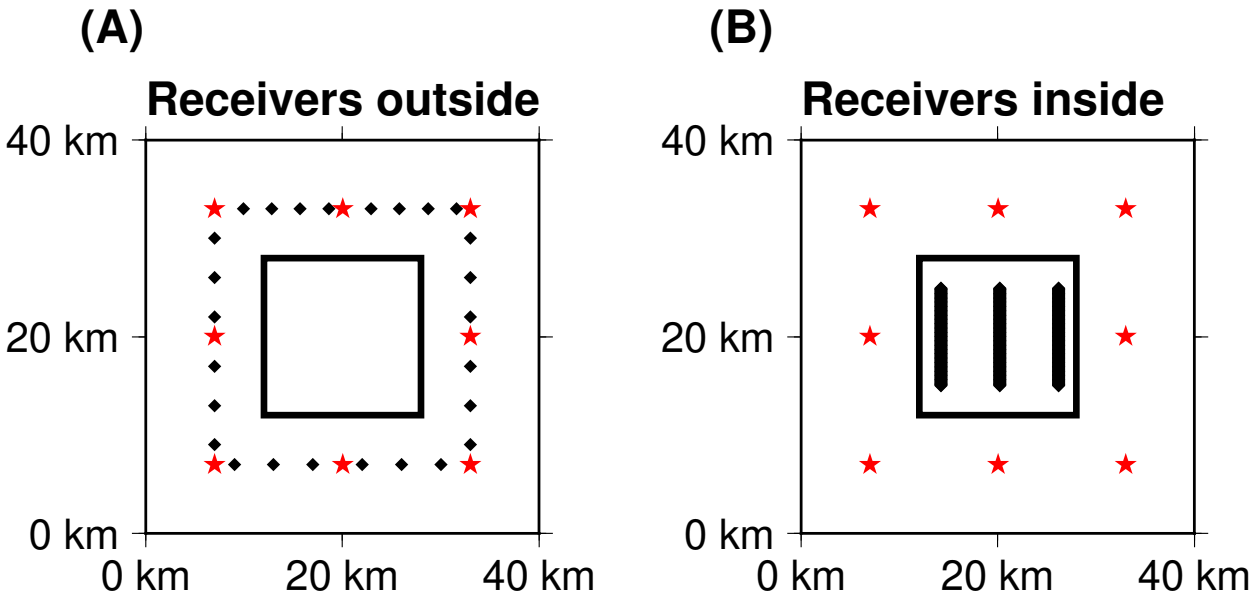

Figure A.5. Source (red stars) and receiver (black diamonds) configurations for the synthetic tests of the strain-based FWI. (A) Configuration with 8 sources and 26 receivers positioned outside of the heterogeneous area. (B) Configuration with 8 sources positioned outside of the heterogeneous area and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

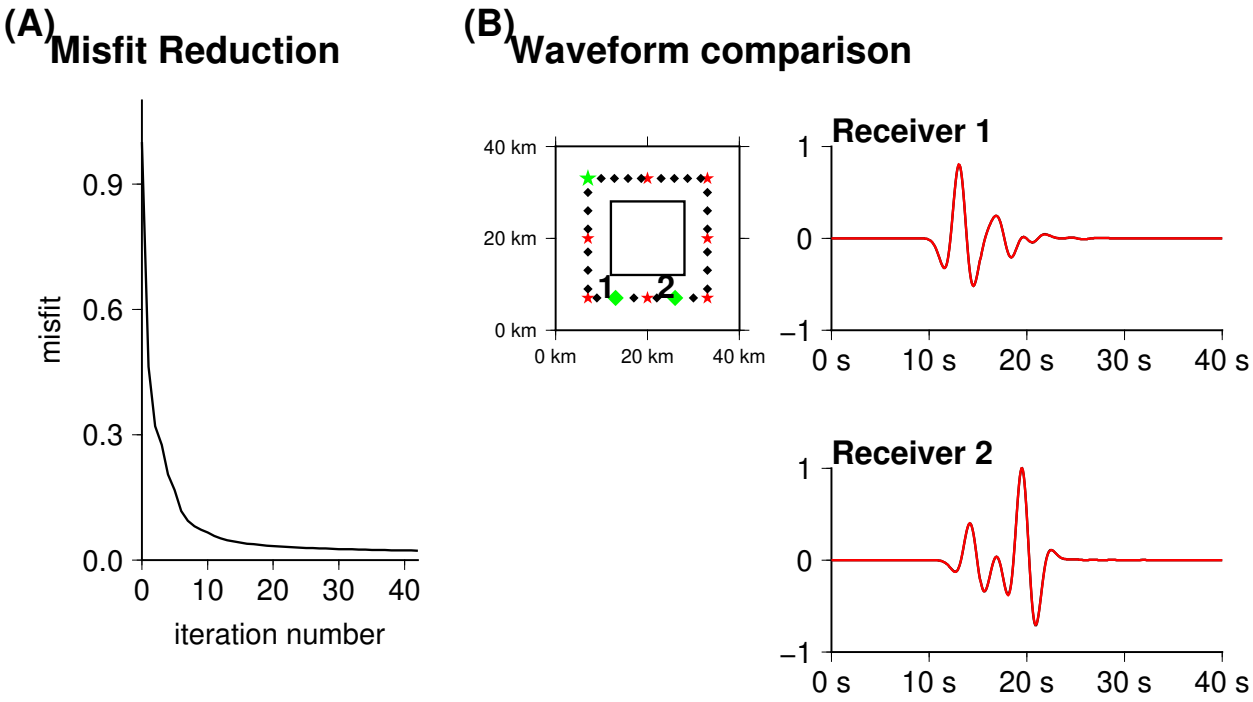

Figure A.6. Strain-based FWI results using receivers positioned outside of the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) positions 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.

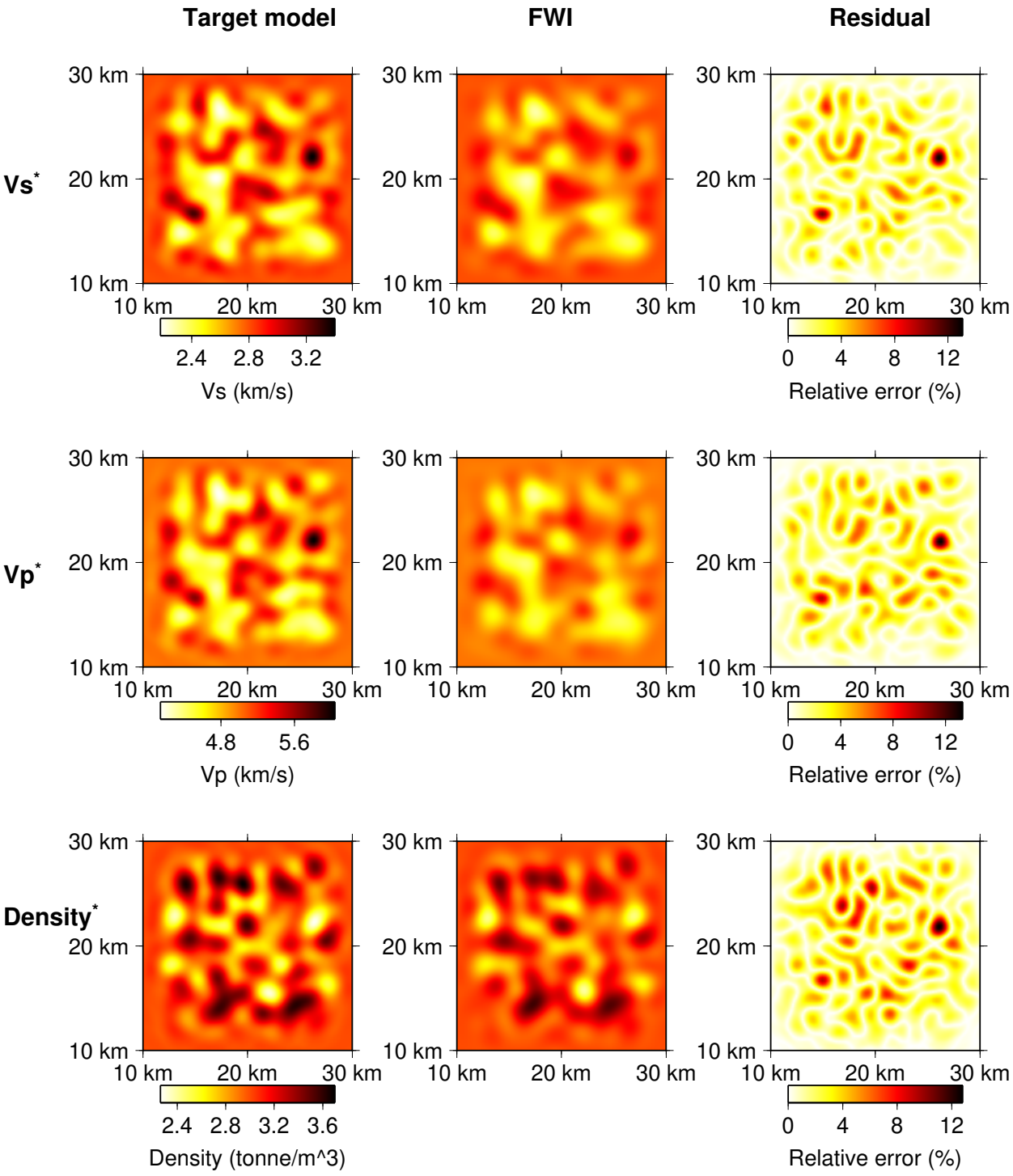

Figure A.7. Strain-based FWI results using receivers positioned outside of the heterogeneous area. The V_S^*, V_P^* and density* structures are shown for the target ("Target model") and final models ("FWI"), and their corresponding residuals ("Residuals"). Both the target and final models are homogenised 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.

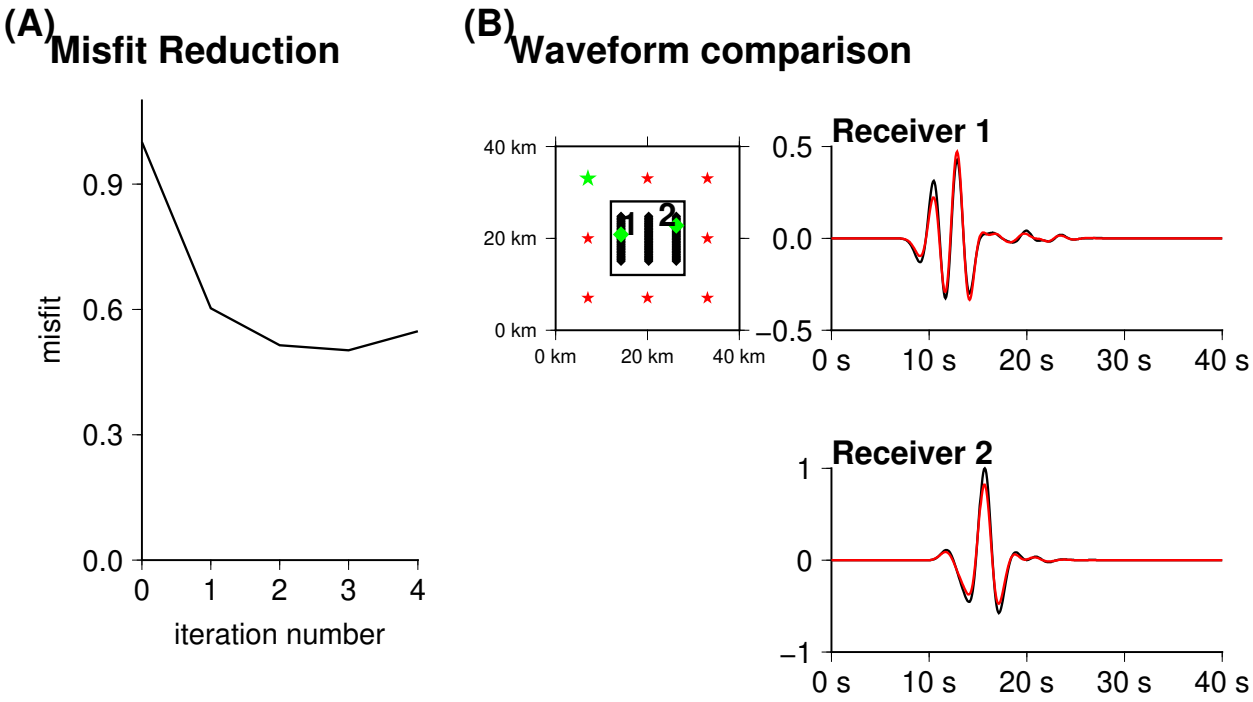

Figure A.8. Strain-based FWI results, without correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) 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.

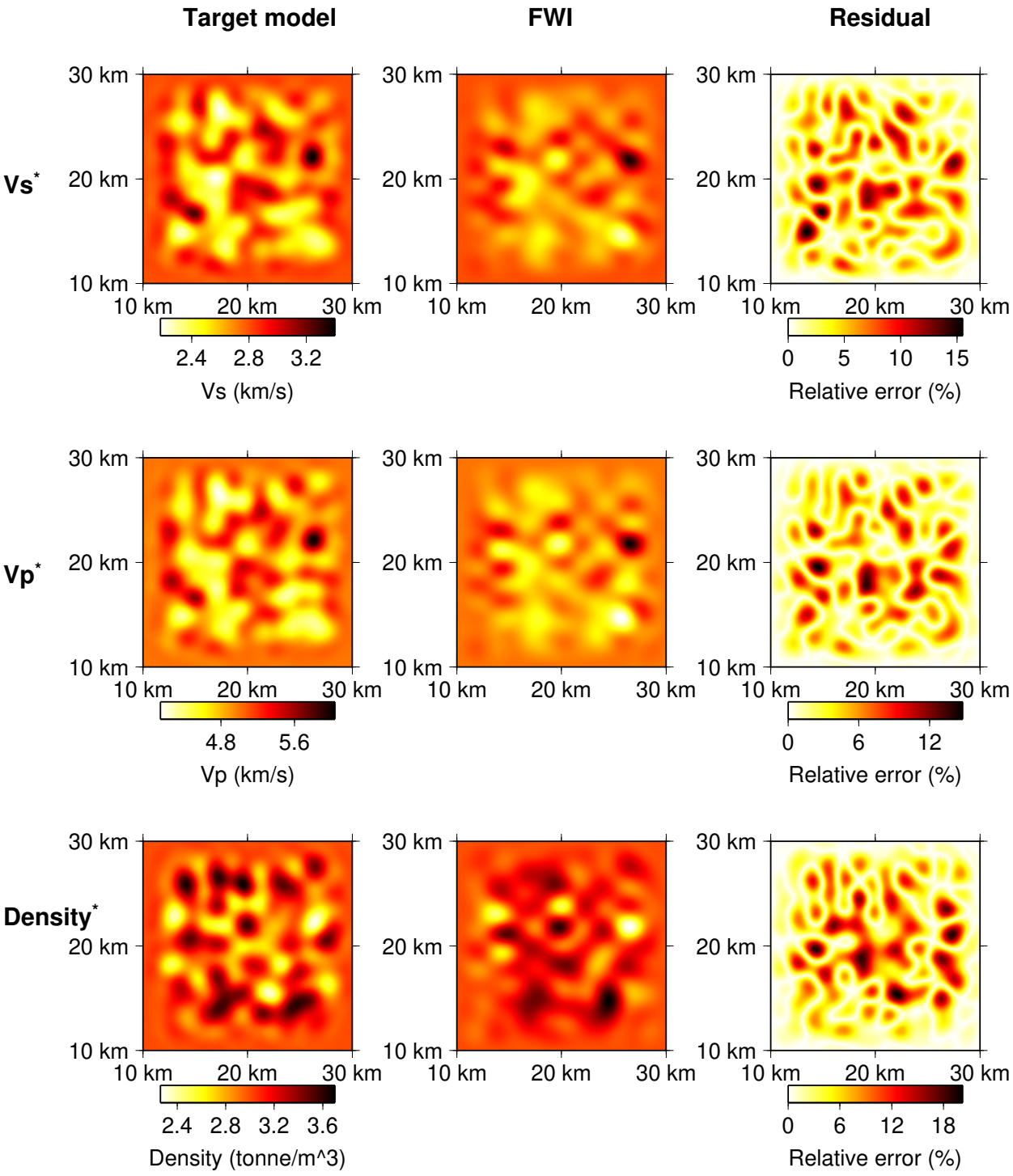

Figure A.9. Strain-based FWI results, without correctors, using receivers positioned within the heterogeneous area. The V_S^*, V_P^* and density* structures are shown for the target ("Target model") and final models ("FWI"), and their corresponding residuals ("Residual"). Both the target and final models 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.

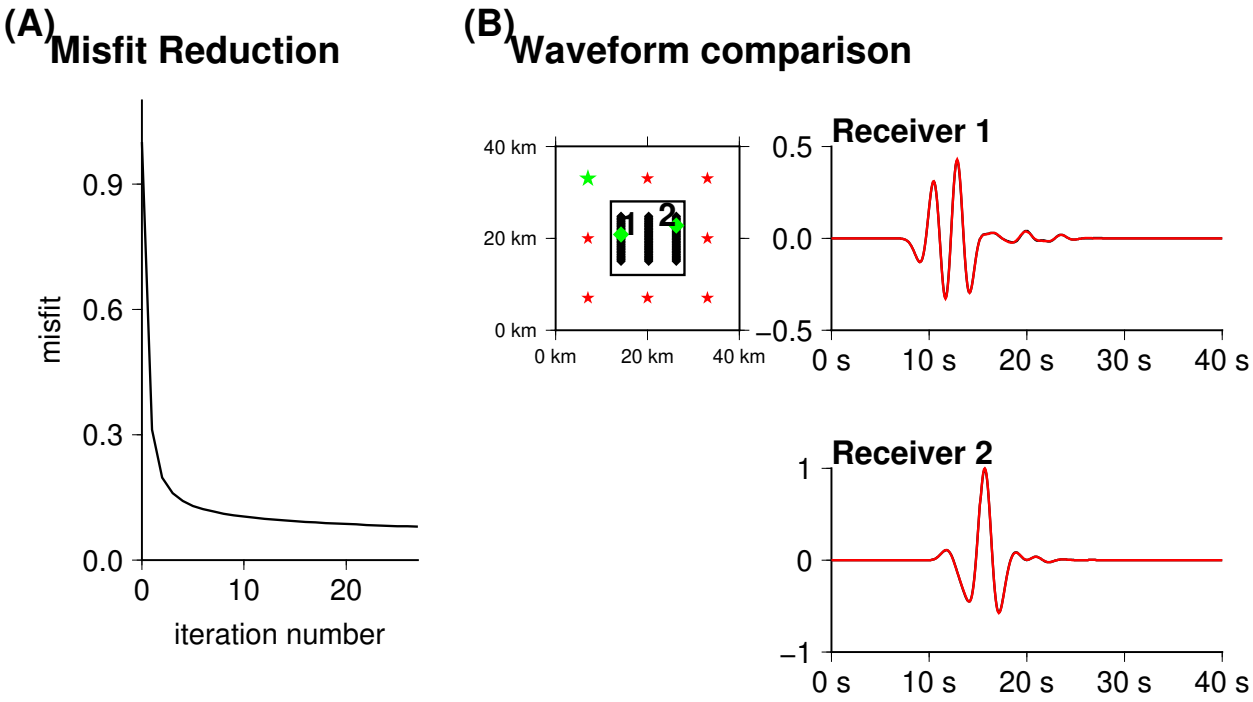

Figure A.10. Strain-based FWI results, with correctors, using receivers positioned within the heterogeneous area. (A) Misfit reduction. (B) Waveform comparisons between the target and synthetic data. Source–receiver positions are shown in the left panel, with the green source (star) and receiver (diamonds) 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.

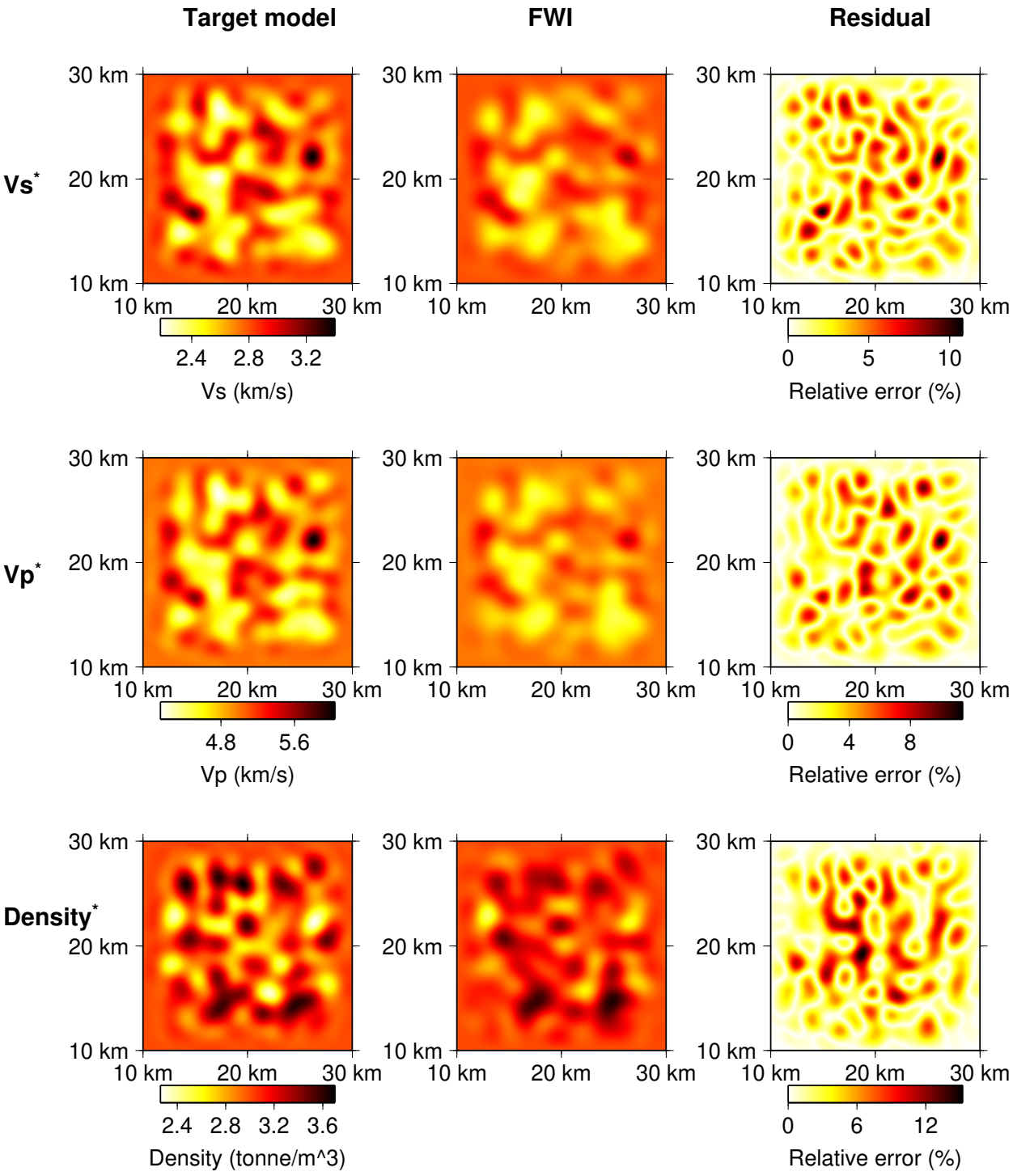

Figure A.11. Strain-based FWI results, with correctors, using receivers positioned within the heterogeneous area. The V_S^*, V_P^* and density* structures are shown for the target (”Target model”) and final models (”FWI”), and their corresponding residuals (”Residual”). Both the target and final models 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.