# Bayesian Seismic Source Inversion With a 3-D Earth Model of the Japanese Islands



**Author(s):**
Simutė, Saulė; Boehm, Christian; Krischer, Lion; Gokhberg, Alexey; Vallée, Martin; Fichtner, Andreas

**Publication date:**
2023-01

**Permanent link:**
https://doi.org/10.3929/ethz-b-000594484



**Originally published in:**
Journal of Geophysical Research: Solid Earth 128(1), https://doi.org/10.1029/2022jb024231

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2022JB024231

# Bayesian Seismic Source Inversion With a 3-D Earth Model of the Japanese Islands

**Saulė Simutė[1,2], Christian Boehm[1], Lion Krischer[1], Alexey Gokhberg[1,3], Martin Vallée[2], and Andreas Fichtner[1]**

[1]Institute of Geophysics, ETH Zurich, Zurich, Switzerland, [2]CNRS, Institut de physique du globe de Paris, Université de Paris, Paris, France, [3]Fragata Computer Systems AG, Schwyz, Switzerland

**Key Points:**

- Full-waveform inversion models enable the use of shorter-period data in earthquake source inversion
- Using regional waveforms at 15 s period may reduce non-double-couple components by tens of percent
- Probability densities of source parameters below 20 s are non-Gaussian, thus demanding stochastic inversion approaches

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

S. Simutė,
saaule@gmail.com

**Citation:**

Simutė, S., Boehm, C., Krischer, L., Gokhberg, A., Vallée, M., & Fichtner, A. (2023). Bayesian seismic source inversion with a 3-D Earth model of the Japanese islands. *Journal of Geophysical Research: Solid Earth*, *128*, e2022JB024231. https://doi.org/10.1029/2022JB024231

Received 16 FEB 2022
Accepted 18 DEC 2022

**Author Contributions:**

**Conceptualization:** Saulė Simutė, Christian Boehm, Andreas Fichtner
**Formal analysis:** Saulė Simutė
**Funding acquisition:** Andreas Fichtner
**Methodology:** Saulė Simutė, Christian Boehm, Lion Krischer, Andreas Fichtner
**Software:** Saulė Simutė, Lion Krischer, Alexey Gokhberg
**Supervision:** Martin Vallée, Andreas Fichtner
**Validation:** Saulė Simutė
**Visualization:** Saulė Simutė
**Writing – original draft:** Saulė Simutė, Andreas Fichtner



**Abstract** We present probabilistic centroid-moment tensor solutions inferred from the combination of Hamiltonian Monte Carlo sampling and a 3-D full-waveform inversion Earth model of the Japanese islands. While the former provides complete posterior probability densities, the latter allows us to exploit waveform data with periods as low as 15 s. For the computation of Green's functions, we employ spectral-element simulations through the radially anisotropic and visco-elastic model, leading to substantial improvements of data fit compared to layered models. Focusing on 13 $M_w$ 4.8–$M_w$ 5.3 offshore earthquakes with a significant non-double-couple (non-DC) component, we simultaneously infer the centroid location, time and moment tensor without any a priori constraints on the faulting mechanism. Furthermore, we perform the inversions across several period bands, varying the minimum period between 15 and 50 s. Accounting for 3-D Earth structure at shorter periods can increase the double-couple (DC) component of an event, compared to the GCMT solution, by tens of percent. This suggests that non-DC events in the GCMT catalog may result from unmodeled Earth structure and the related limitation to longer-period data. We also observe that significant changes in source parameters, and the DC component in particular, may be related to only small waveform changes, thereby accentuating the importance of a reliable Earth model. Posterior probability density distributions become increasingly multimodal for shorter-period data that provide tighter constraints on source parameters. This implies, in our specific case, that stochastic approaches to the source inversion problem are required for periods below ∼20 s to avoid trapping in local minima.

**Plain Language Summary** In the majority of global earthquake catalogs, the earthquake solution, that is, centroid location, time and a rupture mechanism, is typically inferred assuming a 1-D Earth model. However, both earthquake source and Earth structure contribute to seismic recordings, meaning that unaccounted structure might map into and pollute the source solution. In this study we use a 3-D Earth structure of the Japanese islands to model the waveforms and infer earthquake parameters of 13 small-to-moderate magnitude offshore events. We do not put any a priori constraints on the faulting mechanism and let it be determined by the data. We perform stochastic inversions, which provide us with a collection of all plausible models ranked by their respective probability. When a 3-D Earth structure at shorter periods is taken into account, the earthquake mechanisms, investigated in this study, can be largely explained by a slip on the fault. We also observe that significant changes in source parameters may be related to tiny waveform changes, thereby accentuating the importance of a reliable Earth model.

## 1. Introduction

Earthquake source solutions are important in many fields of seismology, such as, but not limited to, seismic hazard, earthquake physics, seismotectonics, and seismic tomography. Although source inversion is an established discipline in seismology, obtaining a robust solution of moment tensor components together with a spatial and temporal location is still a challenging task. A largely unexplored potential lies in the adoption of more realistic Green's functions, which until recently have only accounted for radially symmetric Earth structure. In the wake of increasing computational power and growing number of full-waveform tomographic models, use of numerically computed Green's functions for complex regional Earth models has become possible.

**Writing – review & editing:** Saulė Simutė, Andreas Fichtner

### 1.1. Recent Developments in Source Inversion

Earthquake source mechanisms, in terms of the first-motion polarities and fault-plane solutions, have been studied since the beginning of the 20th century (Byerly, 1928; Galitzin, 1909; Omori, 1905), and the first computer programs, intended to aid the graphical analysis, were developed in the early 1960s (Kasahara, 1963; Knopoff, 1961). A significant development was accomplished by Backus and Mulcahy (1976a, 1976b), who derived a phenomenological representation for an indigenous source and showed that each seismic source can be described by a moment tensor, or a distribution thereof. Seismic source inversion has become routine since the end of the last century (e.g., Dziewoński et al., 1981; Kanamori & Given, 1981; Mendiguren, 1977). Since then, different approaches have been established to retrieve information on the source parameters, based on, for example, first-motion polarity (e.g., Hara et al., 2019; Kasahara, 1963; Knopoff, 1961; Lentas, 2017), body waveforms (e.g., Dreger & Helmberger, 1991; Dziewoński et al., 1981; Vallée et al., 2011), surface waves (e.g., Ferreira & Woodhouse, 2006; Kanamori & Given, 1981; Romanowicz, 1982), with a specific interest to the ultra-long-period W-phase (e.g., Duputel et al., 2012; Hayes et al., 2009; Kanamori & Rivera, 2008), or full waveforms, incorporating both body and surface wave signals (e.g., Dreger, 2003; Ekström et al., 2012; Hallo et al., 2017; Scognamiglio et al., 2016). While some methods might be more robust than others, all of them, to some degree, rely on how well one can predict the data for a given set of model parameters. The choice of the Earth model, hence, is of fundamental importance in earthquake source inversion, as unaccounted Earth structure might map into the source solution and potentially pollute it (e.g., Hjörleifsdóttir & Ekström, 2010; Smith & Ekström, 1996; Thurber, 1983; Woodhouse, 1983).

Until recently, radially symmetric Earth models have been predominantly used in source inversion studies, for they allow one to efficiently compute Green's functions. Lateral heterogeneities are then taken into account via empirical or theoretical corrections (Ferreira et al., 2011). For example, traces or even the different portions of Green's functions can be shifted independently to fit the data (e.g., Ford et al., 2009a, 2009b; Zhao & Helmberger, 1994; Zhu & Helmberger, 1996). However, such corrections might mask earthquake source effects. Theoretical surface wave corrections may be implemented in terms of mean phase slowness along the source-receiver great circle (e.g., Pondrelli et al., 2002; Woodhouse & Dziewoński, 1984), neglecting the amplitude effects. In addition, great circle approximations, relying on ray theory, do not account for finite-frequency effects of wave propagation, hence the corrections themselves might be erroneous. Another approach is to use multiple 1-D Earth models to account for differences in oceanic and continental crust (Lee et al., 2011), as is done for the National Research Institute for Earth Science and Disaster Prevention (NIED) earthquake catalog in Japan (Kubo et al., 2002). Alternatively, the dependence on structural models can be alleviated, focusing on those data which are less sensitive to crustal heterogeneities, such as the W-phase (Kanamori & Rivera, 2008) or the Pnl phase (Helmberger & Enge, 1980).

With increasing computational power, improving numerical methods (e.g., Afanasiev et al., 2018; Gokhberg & Fichtner, 2016; Komatitsch et al., 2010; Krischer et al., 2015; Nissen-Meyer et al., 2007; Wu et al., 2018) and theoretical developments (e.g., Fichtner, van Herwaarden et al., 2018; Thrastarson et al., 2020; Tromp et al., 2005; van Herwaarden et al., 2020), full-waveform tomographic models have been proliferating on both regional (e.g., Blom et al., 2020; Fichtner et al., 2009a; Krischer et al., 2018) and global scale (e.g., Bozdağ et al., 2016; Fichtner, van Herwaarden et al., 2018; French & Romanowicz, 2014). This has in turn enabled researches to start using numerically computed 3-D Green's functions for source inversion. Such type of studies have been performed for the Southern California region (Graves & Wald, 2001; Jia et al., 2020; Lee et al., 2011; Liu et al., 2004; X. Wang & Zhan, 2019, 2020; Zhao et al., 2006), the Australian region (Hejrani et al., 2017; Hingee et al., 2011), the Sichuan province in China (Zhu & Zhou, 2016) and more recently for offshore earthquakes along the Nankai trough in Japan (Takemura et al., 2018, 2020). Several studies have investigated the effect of the 3-D Green's functions on source resolvability using synthetic and observed data (e.g., Donner et al., 2020; Graves & Wald, 2001).

The non-linear relationship between data and model parameters, such as centroid location and centroid time of an earthquake, make it difficult to tackle the source inversion with deterministic approaches. The least-squares method, for example, provides a single solution and does not account for non-uniqueness, which can arise due to insufficient data coverage and modeling inaccuracies. Furthermore, uncertainty information, derived by linearization methods, is only representative if the objective functional is indeed quadratic or otherwise have little meaning at all (Sambridge & Mosegaard, 2002). To tackle these challenges, we resort to probabilistic inference, which provides a collection of all plausible models ranked by their respective probability. Statistical inferences

can be made from the ensemble to assess the uncertainty, and covariance matrices can be recovered to study the inter-parameter trade-offs. Such an approach respects the non-uniqueness, avoids the subjective regularization required by the deterministic inversion, and delivers uncertainty measures as part of the solution. However, a more vigorous exploration of the model space typically comes with higher computational costs.

Stochastic approaches were used in the inversions of microseismic events (e.g., Pugh et al., 2016; Shang & Tkalčić, 2020), of events with anomalously high non-double-couple (non-DC) component (e.g., Mustać & Tkalčić, 2016), finite-fault inversions (e.g., Dettmer et al., 2014; Duputel et al., 2014; Minson et al., 2014), and for earthquake early warning purposes (Cua & Heaton, 2007). Only a few probabilistic studies have been performed using fully heterogeneous Earth models (e.g., Lee et al., 2011).

### 1.2. Hamiltonian Monte Carlo

The performance of traditional stochastic random walk methods, such as Metropolis-Hastings (Hastings, 1970; Metropolis et al., 1953), tends to scale poorly with increasing dimension (Betancourt, 2017). One way to guarantee the efficacy of sampling is through informed proposals, a strategy to ensure that the transitions largely follow the contours of high-probability mass (Betancourt, 2017; Neal, 1996, 2011). Informed algorithms, such as Hamiltonian Monte Carlo (HMC) are designed to make use of the information outside of a simple target distribution evaluation at a given point (e.g., Khoshkholgh et al., 2020; Zanella, 2020). HMC relies on the gradient information of the misfit in order to guide the sampler toward the areas of high-probability mass. It can be regarded as a hybrid approach encompassing the virtues of both gradient-based optimization and derivative-free Markov chain Monte Carlo methods (Fichtner, Zunino, & Gebraad, 2018).

HMC is particularly useful for multi-dimensional problems with high-quality data or weakly constrained priors, which, in traditional, derivative-free sampling algorithms would result in a low acceptance rate and a slow convergence. Although introduced in the 1980s (Duane et al., 1987), HMC has only recently gained popularity in geophysics. Maiti and Tiwari (2009) implemented HMC-based neural networks to analyze well log data, Muir and Tkalčić (2020) applied HMC for a lowermost mantle study, Sen and Biswas (2017) and Biswas and Sen (2017) used HMC in 1-D and 2-D seismic inversions, respectively, while Fichtner, Zunino, and Gebraad (2018) and Gebraad et al. (2020) further proved the potential of HMC for non-linear seismic tomography problems. Very recently Aleardi et al. (2020) used HMC in the context of dispersion curves inversion, while Koch et al. (2020) implemented adjoint HMC in the context of engineering. A variant of HMC that tunes itself while sampling was presented by Fichtner et al. (2021).

The potential of HMC in earthquake source inversion was demonstrated by Fichtner and Simutė (2018), where HMC was adapted for efficient source studies in complex media, with synthetic examples and a real-data illustration. A study of induced earthquake parameters, which also takes advantage of linearized HMC, has been performed by Masfara et al. (2022). In this work, we largely rely on the methodology presented in Fichtner and Simutė (2018) and perform multiple source inversions with an expanded and improved data set.

### 1.3. Motivation and Outline

Green's functions computed for laterally averaged structure are not adequate for tectonically complex areas, especially subduction zones, which require a proper incorporation of 3-D Earth structure (e.g., Engdahl et al., 1977; Igel et al., 2002). Simplified Earth models affect the estimation of the centroid location and time (e.g., Dziewoński & Woodhouse, 1983; Ferreira & Woodhouse, 2006; Hjörleifsdóttir & Ekström, 2010; Morales-Yáñez et al., 2020; Smith & Ekström, 1996; Thurber, 1983), the seismic moment (e.g., Patton & Randall, 2002), as well as the moment tensor itself (e.g., Ferreira & Woodhouse, 2006; Hejrani et al., 2017; Newrkla et al., 2019; Scognamiglio et al., 2016; Woodhouse, 1983), which often manifest as spurious non-double couple components (Zahradník et al., 2015). However, radially symmetric Earth models, which allow for a computationally efficient way to obtain Green's functions, are still commonly used in source inversion studies on the grounds that a suitable data selection might isolate data pertaining principally to the source (e.g., Ford et al., 2009a; Mustać & Tkalčić, 2016; Staehler & Sigloch, 2014; Woodhouse, 1983).

Motivated by the effects that unaccounted Earth structure potentially has on earthquake source solutions and endorsing the need for uncertainty information, we propose a stochastic earthquake source inversion, based on

the HMC sampling algorithm. We start by introducing a heterogeneous, viscoelastic and radially anisotropic Earth model of the crust and upper mantle beneath the Japanese islands region, which is constructed for this study (Section 2). We then introduce the formulation and practical aspects of forward and inverse problems (Sections 3–5). Finally, we present multi-period centroid moment tensor inversion results of earthquakes at the Izu-Bonin trench (Section 6). Owing to the Bayesian framework, we retrieve the uncertainty information as well as the inter-parameter trade-offs. We discuss the implications as well as the limitations of the study in Section 7 and provide concluding remarks in Section 8.

This study should be viewed as a step toward establishing the methodology of source inversion with a 3-D Earth model and HMC framework. Having presented the basic concept in Fichtner and Simutė (2018), here we proceed with (a) applying the approach to several earthquakes in a realistic setup in Japan, (b) establishing technicalities of the algorithm, such as computation and storage of receiver-side Green's strains, and (c) studying the effect of 3-D Earth model and HMC framework on a smaller number of source solutions. Hence, we admittedly handle several aspects, such as data and modeling errors or selection of measurement window, pragmatically. For a production code, some workflows could be automated and more consideration could be given to a more elaborate implementation of data errors.

## 2. Velocity Model for the Japanese Islands

To reduce the effect of 3-D Earth structure on estimated source parameters, we construct a full-waveform inversion model for the Japanese islands region, building on the velocity model previously constructed by Simutė et al. (2016) on the basis of waveform data in the 20–80 s period range. The model is viscoelastic, radially anisotropic and 3-D heterogeneous. For forward and adjoint modeling, we employ the GPU-accelerated spectral-element wave equation solver SES3D (Fichtner et al., 2009b; Gokhberg & Fichtner, 2016). Earthquakes source parameters are obtained from the regional NIED earthquake catalog (Kubo et al., 2002). We use time-frequency phase misfits (Fichtner & Igel, 2008) to quantify differences between observed and synthetic waveforms within automatically selected measurement time windows where waveform similarity is sufficient to avoid cycle skips (Krischer et al., 2015). The final model is the result of an iterative conjugate-gradient minimization of the misfit, with gradients computed by adjoint techniques (Fichtner et al., 2006; Tarantola, 1988; Tromp et al., 2005). We invert for isotropic P velocity $v_p$, SV velocity $v_{sv}$, SH velocity $v_{sh}$, and density $\rho$. Furthermore, we implement viscoelastic attenuation by using the QL6 attenuation model of Durek and Ekström (1996), which is, however, kept constant throughout the inversion. Since the focus of this work is on source inversion, we refer to Simutė et al. (2016) for a more detailed and technical description of the well-established full-waveform inversion method.

Starting with the model presented in Simutė et al. (2016), we performed 14 additional iterations using waveform data with a slightly broadened period range of 15–80 s. First, we completed seven iterations for the larger model domain shown in Figure 1, which we also used previously (Simutė et al., 2016). Subsequently, we performed the remaining seven iterations for a smaller domain and with additional regional data, as indicated in Figure 1. This was intended to specifically improve that part of the model which we later use for the computation of Green's functions, needed for the Bayesian source inversion. With this concrete application in mind, we primarily focus on waveform fit, limiting the presentation of the structural model to a short paragraph at the end of this section.

The overall waveform misfit decreased by 24% after the first seven iterations in the larger initial domain, and by another 21% during the subsequent seven iterations in the smaller domain. More details on the misfit evolution are shown in the Figure S1 in Supporting Information S1.

In Figure 2, we present a small but representative collection of waveform comparisons across the model domain for four $M_w$ 5.0–$M_w$ 5.8 shallow- to intermediate-depth events, situated (a) at the Izu-Bonin trench, (b) off Kyushu, (c) in the Sea of Japan, and (d) off the east coast of Honshu. Together with the selected stations they represent model parts which are relatively well covered by the data. Still, the waveform fit is not uniform across the model domain. Observed waveforms for some paths, such as between event 19 and station BO.KSK, or event 30 and the stations in central Japan, are well explained in terms of both phase and amplitude. Other paths, in contrast, are characterized by a good match in phase but show discrepancies in amplitude; for example, the path between event 16 and station BO.ABU. The latter is a general feature observed across the majority of the traces, suggesting that the source mechanism may need improvement.

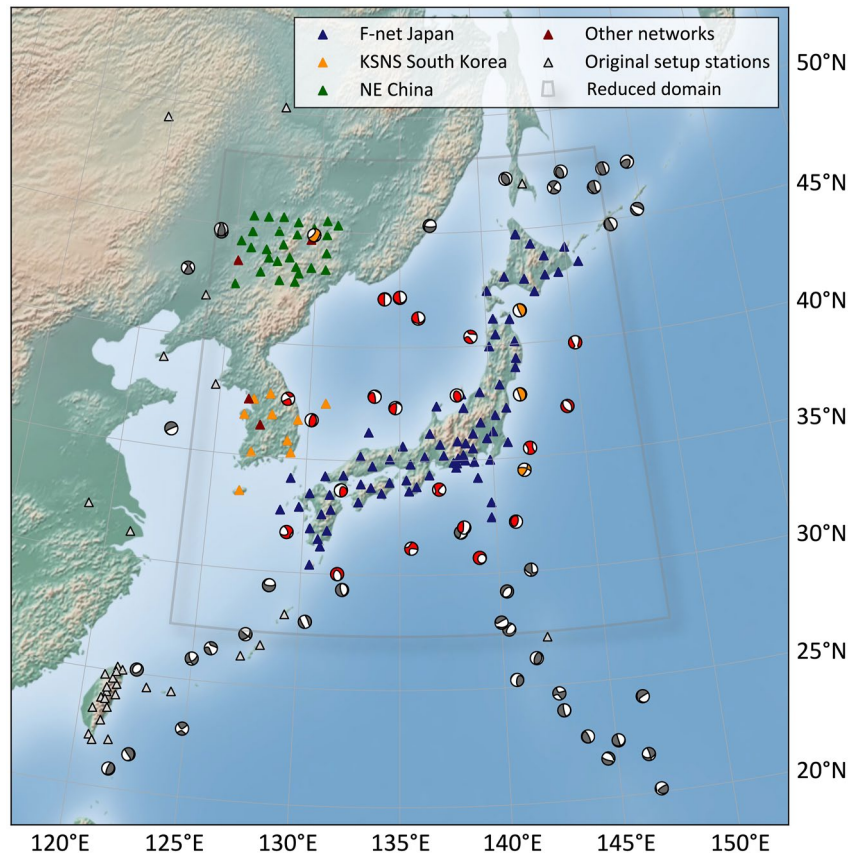


**Figure 1.** Source-receiver setups for tomographic inversions in the initial large domain and the smaller focused domain. Within the large domain, we used 58 earthquakes, depicted as red and gray focal mechanisms, and all the stations except for the NE China array, shown as triangles. The smaller domain comprises 20 events from the original setup shown in red and four new events in orange together with all seismic stations depicted in non-gray color.

To assess the importance of using a 3-D model for source inversion in our study region, we compare computed waveforms for our 3-D model (red solid waveforms in Figure 2) and its laterally averaged 1-D version (red dashed waveforms in Figure 2). The 3-D full-waveform inversion model produces a substantially better waveform fit than the 1-D model, for which time shifts can be on the order of tens of seconds. For the whole-Earth 1-D model AK135 (Kennett et al., 1995) results are similar, thus corroborating that the lateral heterogeneities in our velocity model are indeed required to fit waveform data at periods between 15 and 80 s, as studied here.

In Figure 3 we compare the whole-seismogram waveform fit at short and long periods. The root-mean square error is computed as:

$$\chi_i = \frac{\sqrt{\int_0^T \left[u_i^0(t) - u_i(t)\right]^2 dt}}{\sqrt{\int_0^T u_i^0(t)^2 dt}}, \tag{1}$$

where $u_i^0(t)$ denotes the $i$-component of the observations, $u_i(t)$ the $i$-component of the synthetic seismograms, and $T$ is the duration of the time series. Misfit at long-periods (50–80 s) is low throughout the domain, with remaining discrepancies close to the expected noise level. Misfits at short-period (15–80 s), on the other hand, have more variability, largely correlating with the geological complexity along the source-receiver path. The implications for the source inversion will be further discussed in Section 5.4.

A collection of depth and cross-sectional slices through the tomographic model in terms of deviation of the isotropic S velocity $v_s$ from the lateral average $\bar{v}_s$ is shown in Figure 4. We compute isotropic S velocity as

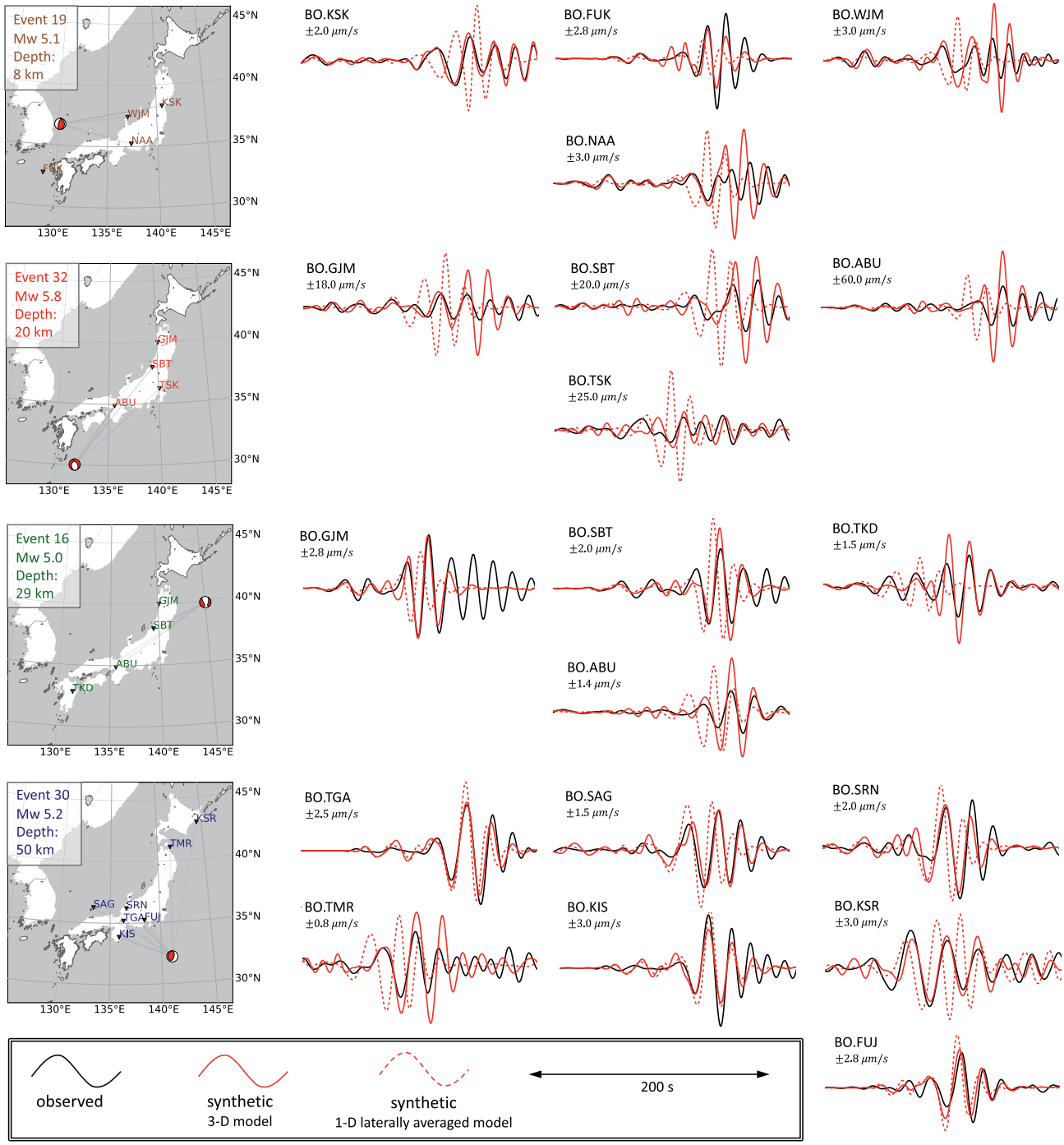


**Figure 2.** Representative collection of observed waveforms (black), synthetic waveforms computed for the final 3-D model (solid red) and synthetic waveforms computed for the 1-D laterally averaged model (dashed red). The waveforms are filtered between 15 and 80 s. We show the vertical component of the waveforms for four events and selected stations, with the source-receiver configuration specified in a separate map for each earthquake. Event information, shown in the top left corner of each map, are National Research Institute for Earth Science and Disaster Prevention Centroid-Moment-Tensor solutions (Fukuyama et al., 2001).

$v_s = \sqrt{\frac{2}{3}v_{sv}^2 + \frac{1}{3}v_{sh}^2}$ (e.g., Babuška & Cara, 1991; Panning & Romanowicz, 2006). The lateral average $\bar{v}_s$ (Figure S2 in Supporting Information S1), more depth slices (Figure S3 in Supporting Information S1), and depth profiles (Figure S4 in Supporting Information S1) as well as anisotropy (Figure S5 in Supporting Information S1), are presented as supplementary information.

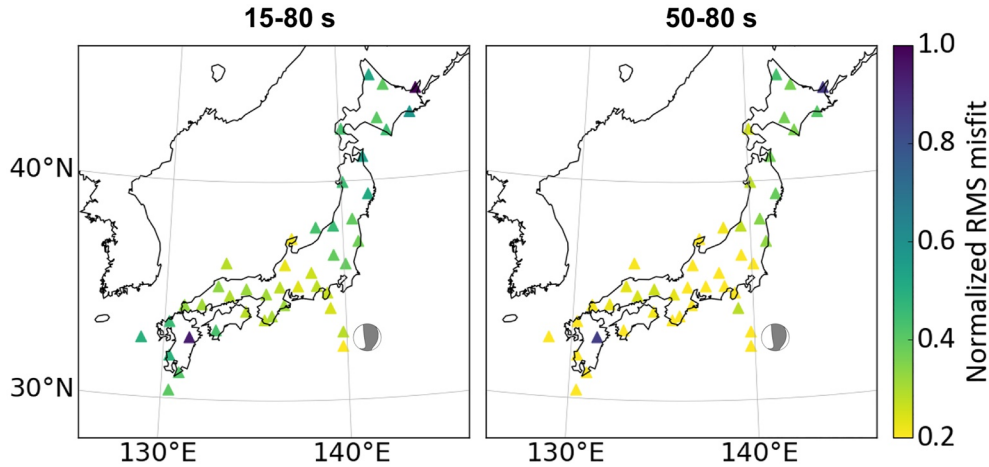


**Figure 3.** Root-mean square (RMS) misfit between the vertical component of the observed data and the synthetic seismograms calculated for the Global Centroid-Moment-Tensor solution for the $M_w$ 5.2 event at 50 km depth depicted as a gray focal mechanism. The waveforms are filtered between 15 and 80 s (left) and 50–80 s (right). Misfits are normalized to the largest value of both scenarios. While longer-period data are well explained at the majority of the stations, more variation in misfit is present at shorter periods. Generally, stations in central Japan exhibit a very good fit, while those further away from the event, such as in Hokkaido or Kyushu, are characterized by slightly elevated misfits.

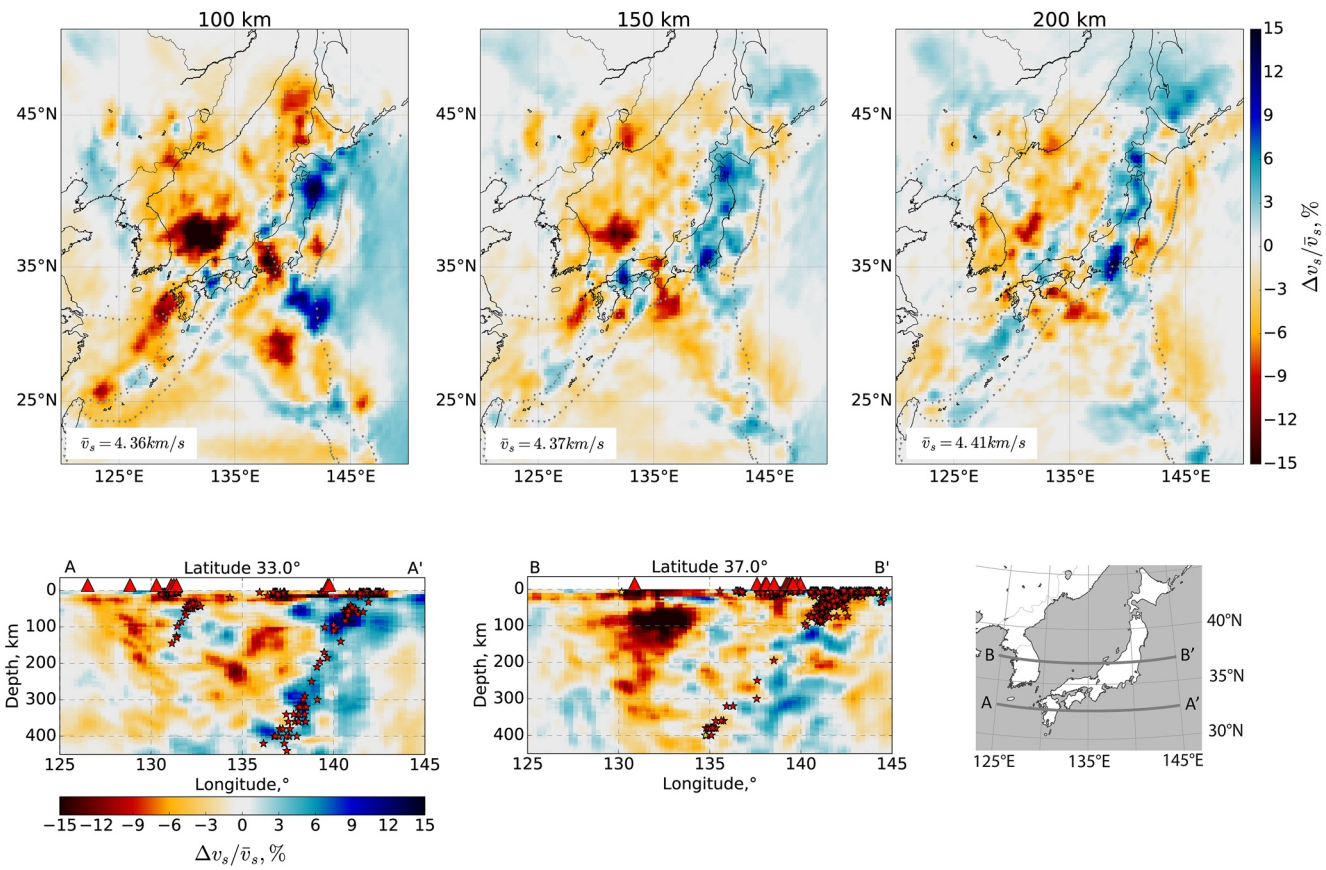


**Figure 4.** Percentage perturbations of the isotropic S velocity, computed as $\frac{v_s - \bar{v}_s}{\bar{v}_s} \times 100\%$, where $\bar{v}_s$ is the lateral average of $v_s$ for each depth. (Top panels) Horizontal slices, with dashed gray lines representing plate boundaries. (Bottom panel) Vertical cross-sections. Red and yellow stars represent earthquakes since 1997 and earthquakes used in the tomographic study, respectively, within 1° of the slice. Red triangles represent Holocene volcanoes (Siebert et al., 2010).

## 3. Forward Problem

### 3.1. Representation of the Displacement Field

In a point-source configuration the $i$-component of the displacement field $\boldsymbol{u}$ can be expressed as a convolution of a time-dependent moment tensor $M_{nq}(t)$ and Green's strains $G_{in,q}(\boldsymbol{x}, t; \boldsymbol{\xi}, \tau)$ (Aki & Richards, 2002):

$$u_i(\boldsymbol{x}, t) = M_{nq}(t) * G_{in,q}(\boldsymbol{x}, t; \boldsymbol{\xi}, \tau), \tag{2}$$

where Green's function $G_{in}(\boldsymbol{x}, t; \boldsymbol{\xi}, \tau)$ is the $i$-component of the displacement field recorded at location $\boldsymbol{x}$ and time $t$ due to an impulse in $n$-direction at location $\boldsymbol{\xi}$ and time $\tau$, and $G_{in,q}(\boldsymbol{x}, t; \boldsymbol{\xi}, \tau)$ is a spatial gradient of Green's function with respect to the $q$-coordinate of the source location, with, $q$ denoting a derivative with respect to $\xi_q$. Einstein notation is implied.

In practice, it is often assumed that separate moment tensor components have the same time dependence, or the same source time function $s(t)$ (e.g., Dziewoński et al., 1981; Ekström et al., 2012; Mustać & Tkalčić, 2016; Takemura et al., 2020; Zhu & Zhou, 2016), in which case a component of the displacement field can be expressed as:

$$u_i(\boldsymbol{x}, t) = M_{nq}\, s(t) * G_{in,q}(\boldsymbol{x}, t; \boldsymbol{\xi}, \tau). \tag{3}$$

The source duration of $\sim M_w$ 5 events, which we consider in this study, is usually a few seconds long (e.g., Vallée & Douet, 2016). Hence, we assume an instantaneous source time function, which is a sufficient approximation for the shortest periods we work with, that is, 15 s, which was also shown in the pilot study by Fichtner and Simutė (2018). Assuming the same source time function for all events and all moment tensor components, we can convolve $s(t)$ with the Green's strains at the time of computation, in which case the displacement field becomes a linear combination of convolved Green's strains scaled by the moment tensor elements.

To ensure a rapid forward problem for probabilistic inference, we pre-compute and store the Green's strains in a database, taking advantage of spatial reciprocity. The merits of reciprocity for the computations of Green's functions have also been exploited in previous studies (e.g., Eisner & Clayton, 2001; Hejrani et al., 2017; Lee et al., 2011; Okamoto et al., 2018; Takemura et al., 2020; Zhao et al., 2006). The reciprocal formulation of Equation 3 can be found in Fichtner and Simutė (2018).

### 3.2. Database of Green's Strains

We compute Green's strains numerically with the spectral-element solver SES3D (Fichtner et al., 2009b; Gokhberg & Fichtner, 2016). Enabled by reciprocity, we treat seismic stations as virtual sources and save the wavefield across the actual source area of interest, that is, the Izu-Bonin trench. To ensure a continuous representation of the wavefield within the domain, we store the wavefield on Gauss-Lobatto-Legendre (GLL) points of the fourth-order spectral-element method (SEM) grid and use the built-in polynomial interpolation of SEM to extract the wavefield for any spatial coordinate. This contrasts with the common practice of storing the Green's functions on a pre-defined grid (e.g., Hejrani et al., 2017; Lee et al., 2011; Takemura et al., 2020; Vackár et al., 2017), where one has to implement an interpolation routine or deal with a finite number of discrete locations and possibly limit the spatial resolution of the earthquake location. Storing the wavefield itself allows us to extract the strains for any potential source location and be exempt from any additional parametrization effects. We compute the database for over 50 selected F-net broadband stations (Figure 5) uniformly distributed across the network (National Research Institute for Earth Science and Disaster Resilience, 2021).

The source area of interest extends between 140°E and 143°E, 30°N–35°N, and down to 110 km depth. The downsampled wavefield with a time increment of 2 s takes 27 Gb of space for a single virtual source, and the total storage requirements are 4.2 Tb. The database is stored on Piz Daint supercomputer in the Swiss National Supercomputing Center, which we use to rapidly perform the inversions (Swiss National Supercomputing Center, 2021).

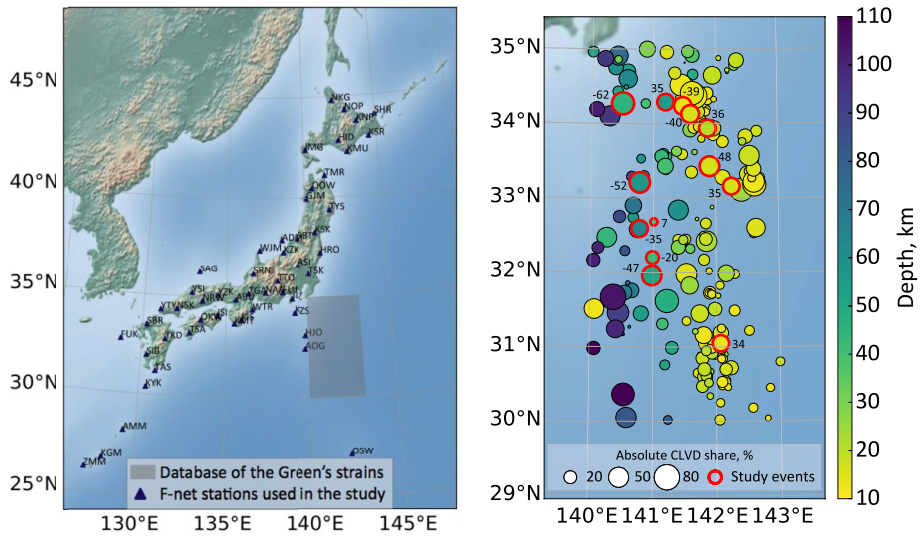


**Figure 5.** (Left) Setup of the stations and the source area used in the source inversion. Receiver-side Green's strains were computed from each seismic station, acting as a virtual source, and stored within the shaded source area. The strain database extends from the surface to 110 km depth. (Right) Distribution of earthquakes within the horizontal extent of the source area as given in the Global Centroid-Moment-Tensor catalog between 1997 and 2020 (The Global CMT Project, 2021). The color of the circles corresponds to the depth of an earthquake, and the size to the absolute share of the compensated linear vector dipole (CLVD) component. Events used in this study are outlined in red. Depth scale saturates at a maximum depth.

## 4. Earthquake Selection

### 4.1. Moment Tensor Decomposition

Moment tensor decomposition is used for physical interpretation and classification of the seismic sources. Typically, a moment tensor is divided into isotropic (ISO) and deviatoric components. The latter can be further decomposed into a number of equivalent force combinations, such as double-couple (DC) and compensated linear vector dipole (CLVD) parts (Jost & Herrmann, 1989; Vavryčuk, 2015). In this study, we compute CLVD component of deviatoric events, that is, those from the GCMT catalog as $2 \times \frac{|m_3|}{|m_1|} \times 100\%$, where $m_{1,2,3}$ are the eigenvalues and $|m_1| > |m_2| > |m_3|$ (Dziewoński et al., 1981). The largest positive eigenvalue corresponds to the tension axis, the largest negative (in terms of the absolute value) to the compression axis, and the smallest absolute eigenvalue, $m_3$, to the null axis. A perfect CLVD event has a value of 100%, while a pure DC event 0%, as $m_3 = 0$ for DC events (Dziewoński et al., 1981). The moment tensor decomposition of non-deviatoric events to ISO, DC and CLVD components is done after Vavryčuk (2015), which has been adopted in previous studies (e.g., Shang & Tkalčić, 2020; Vackár et al., 2017; Yu et al., 2019). We note that the deviatoric part of the moment tensor can be decomposed into a number of equivalent force combinations, and hence the choice is subjective (e.g., Jost & Herrmann, 1989).

### 4.2. Earthquakes in the Area

The choice of the Izu-Bonin trench as a source area is primarily motivated by its high seismicity extending from the shallow crust and along the subduction interface, which gives us a range of earthquake mechanisms and depths to study. Mostly regional source-receiver distances allow for a safe point-source approximation, while the choice in general represents a realistic setup with imperfect station coverage.

The Izu-Bonin trench marks the boundary between the subducting Pacific plate and the Philippine Sea plate. The trench is situated nearly linearly from north to south. It is a steeply dipping subduction zone, with the angles of the Wadati-Benioff zone between 50° and 70° (Faccenna et al., 2018). Along the Izu-Bonin slab, seismicity extends from the shallow surface down to the transition zone in the south and ~410 km depth in the north (Dziewoński et al., 1981; Ekström et al., 2012; Hayes, 2018; Hayes et al., 2012; Seno & Eguchi, 1983). Following the global trend, the majority of events are located in the upper ~60 km (Hasegawa, 2011; Kong et al., 2018) (Figure 5).

At this depth, seismicity primarily occurs as a low-angle interplate thrust faulting, reflecting the relative motion of the convergent plates (Hasegawa, 1990, 2011). Deeper down, earthquakes mostly take place within the slab (Hasegawa, 1990, 2011). In the Izu-Bonin arc—an old plate subduction zone—these intraslab events have the compressional axis predominantly oriented in the dip direction (Hasegawa, 2011). In the overriding plate, the compression is accommodated by intraplate thrust fault or strike-slip fault earthquakes, with compressional axis oriented in the direction of plate convergence (Hasegawa, 1990).

Given the complex nature of the subduction zone, earthquake mechanisms are diverse. Notably, there are numerous strongly non-DC events, with CLVD component reaching up to 80% of the total moment (Figure 5). There are physical explanations for CLVD mechanisms, such as simultaneous faulting of two non-parallel planes (Kuge & Kawakatsu, 1993), or complex faulting with any deviation from unidirectionality in terms of a strike, dip or rake, with volcanic caldera collapse being a perfect example how many nearly simultaneous slips on a curved fault result in an effective vertical-CLVD earthquake (Fichtner & Tkalcic, 2010; Nettles & Ekström, 1998; Shuler et al., 2013). However, very often an apparent CLVD component is an artifact caused by a modeling error. Incorrect Earth structure, especially around the hypocenter of the earthquake, has a significant influence on moment tensor estimation (Burgos et al., 2016; Shuler et al., 2013). For an intuitive understanding, one can think in terms of a first-polarity inversion and a simple double-couple earthquake. The take-off angle depends on the velocity structure in which the earthquake is embedded. When the take-off angle is incorrect, the inferred pressure (P) and tension (T) axes, which correspond to the middle of dilatational and compressional quadrants, respectively, are also incorrect (e.g., Newrkla et al., 2019). The inconsistencies can go a long way, such that the mechanism can no longer be explained by a double-couple, and the errors in modeling are then compensated by introducing a CLVD component. Inability to clearly distinguish between the physical versus apparent CLVD component inhibits our understanding of earthquake physics, while possibly incorrect focal mechanisms hinders the accurate delineation of the local tectonic setting. Hence, by incorporating complex Earth structure in our study we expect to see whether the CLVD component is a physical feature of the earthquake or an artifact due to modeling errors.

### 4.3. Study Events

We study events of moderate magnitudes, which fall within the area of our strain database (Figure 5) and have a significant CLVD component. Preference is given to more recent earthquakes away from the database boundaries. We selected 13 events from the Global Centroid-Moment-Tensor (GCMT) catalog (Dziewoński et al., 1981; Ekström et al., 2012; The Global CMT Project, 2021). The earthquakes have moment magnitudes between $M_w$ 4.8–5.3, are distributed within 13 and 64 km depth with CLVD component ranging between 7% and 62% (in absolute sense), with a median value of 36% (Figures 5 and 6, Table S1 in Supporting Information S1).

## 5. Inverse Problem

### 5.1. Bayesian Inference

We work in the Bayesian framework, where according to Bayes' theorem (Bayes & Price, 1763) the posterior probability density $\pi(\boldsymbol{q}|\boldsymbol{d})$ of the model vector $\boldsymbol{q}$ given the data $\boldsymbol{d}$ is:

$$\pi(\boldsymbol{q}|\boldsymbol{d}) = k\pi(\boldsymbol{d}|\boldsymbol{q})\pi(\boldsymbol{q}). \tag{4}$$

Bayes' theorem provides a framework to enhance the existing knowledge, or the prior probability density $\pi(\boldsymbol{q})$, with the new information from the data, that is, the likelihood $\pi(\boldsymbol{d}|\boldsymbol{q})$ (Fichtner, 2021; Mosegaard & Sambridge, 2002; Mosegaard & Tarantola, 2002; Sambridge & Gallagher, 2011). The likelihood term contains information on the data fit, that is, how well the current model can explain the data. A constant $k$ ensures the integral of the posterior probability density over the model space is equal to one (e.g., Mustać & Tkalčić, 2016; Sambridge & Mosegaard, 2002; Staehler & Sigloch, 2014).

We express the likelihood $\pi(\boldsymbol{d}|\boldsymbol{q})$ as the exponential function of the negative $L_2$ misfit between the synthetic and the observed waveforms, $\mathbf{s}$ and $\mathbf{d}$, respectively:

$$\pi(\boldsymbol{d}|\boldsymbol{q}) \propto \exp(-\chi), \tag{5}$$

$$\chi = \frac{1}{2}(\mathbf{s}(\mathbf{q}) - \mathbf{d})^T \boldsymbol{C}_D{}^{-1}(\mathbf{s}(\mathbf{q}) - \mathbf{d}), \tag{6}$$

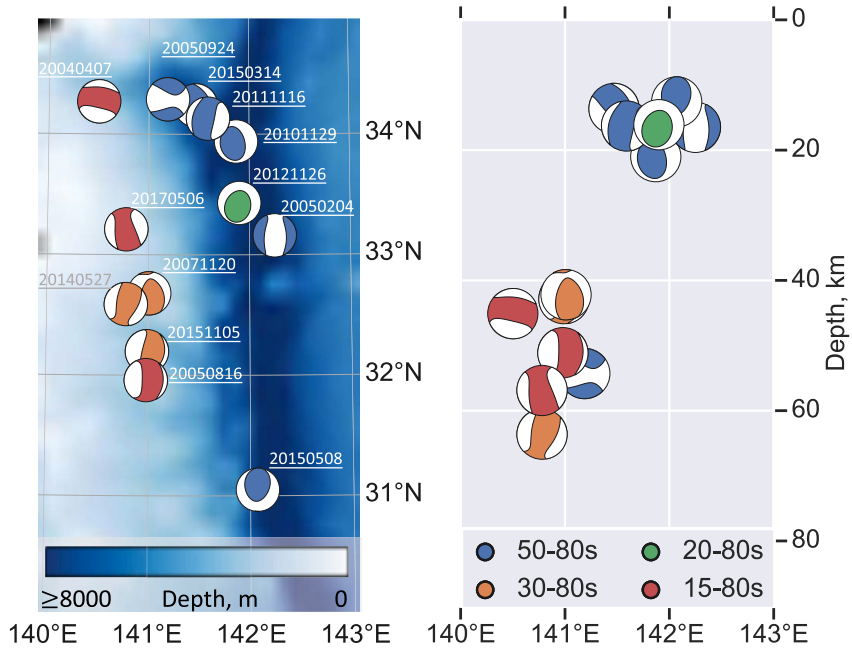


**Figure 6.** Distribution of earthquakes selected for this study. Earthquakes are plotted in terms of their focal mechanisms, with colors representing the shortest acceptable inversion periods (see Section 5.4 for more details). Left: horizontal distribution of the events and their IDs plotted on the bathymetric map. (Right) Depth distribution of the study events. Deeper events can be modeled over a wider frequency range compared to the shallow ones.

where $\boldsymbol{C_D}$ denotes the data covariance matrix.

### 5.2. Hamiltonian Monte Carlo

We pursue Bayesian inversion with the HMC method. Originally introduced as *hybrid* Monte Carlo (Duane et al., 1987), the method derives from molecular dynamics simulation, used to study the properties of many-body systems by solving Newton's equations of motion (e.g., Alder & Wainwright, 1959; Neal, 1993; Xu & Li, 2008). In short, HMC can be regarded as an efficient proposal mechanism, which relies on exploiting gradient information of the model parameters. The main idea of the algorithm is to follow a contour of high probability, which is achieved by balancing the gradient, or the *force*, by an artificially introduced *momentum*.

To set the stage for HMC, we first expand our model parameter space, described by the $N_q$-dimensional position vector $\boldsymbol{q}$, with auxiliary momentum parameters $\boldsymbol{p}$. For physical intuition of the sampling process one could imagine a mechanical particle in phase space. To propose a new sample, a particle is set into motion by randomly assigning momentum to each model parameter. The particle then travels along the trajectory for some artificial time $\tau$. The end of the trajectory serves as a new proposal. By marginalizing over the artificially introduced momentum, we can retain only the position variables, that is, the physical part of the phase space, which we are actually interested in. Mathematically, the trajectory is governed by Hamiltonian dynamics:

$$\frac{dq_i}{d\tau} = \frac{\partial K}{\partial p_i}, \qquad \frac{dp_i}{d\tau} = \frac{\partial U}{\partial q_i}, \qquad i = 1, \dots, N_q, \tag{7}$$

where potential energy $U$ is expressed as:

$$U(\boldsymbol{q}) = -\ln \pi(\boldsymbol{q}|\boldsymbol{d}), \tag{8}$$

and kinetic energy must be defined by the implementation. In this study we use:

$$K(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T \mathbf{M}^{-1} \boldsymbol{p}, \tag{9}$$

where the positive-definite mass matrix $\mathbf{M}$ is a tuning parameter, which, generally speaking, acts as a scaling parameter to ensure that momentum is tailored to the sensitivity of each model parameter. This, in turn, allows us to explore the space equally well for each parameter. The solution of Hamilton's equations throughout the artificial time $\tau$ represents the evolution of the model in phase space. The discretization of artificial time and the total length of the trajectory, $L$, are the tuning parameters of HMC. The dynamics conserve the total energy $H$ throughout the Hamiltonian trajectory:

$$H(\boldsymbol{q},\boldsymbol{p}) = U(\boldsymbol{q}) + K(\boldsymbol{p}). \tag{10}$$

To solve the differential equations (Equation 7), we numerically integrate using a leapfrog algorithm, which preserves the volumes of regions of phase space, meaning, that by moving from one region in phase space to another, the points retain the same volume and reversibility in time, two properties important for HMC (Neal, 1993, 2011). The total energy, on the other hand, is not conserved by the leapfrog algorithm, and this affects the acceptance rate of the proposed samples.

The algorithm is performed in steps, starting with some model $\mathbf{q}$:

1. Draw momentum values from the multivariate normal distribution $\exp\left(-\frac{1}{2}\boldsymbol{p}^T\mathbf{M}^{-1}\boldsymbol{p}\right)$.
2. With $\mathbf{q}$ and $\mathbf{p}$ specified, solve Hamilton's equations (Equation 7).
3. The end of the trajectory marks a newly proposed sample in terms of $\mathbf{q}(L)$ and $\mathbf{p}(L)$. After evaluating the total energy of the new sample, $H$, the model is accepted with probability:

$$\Pi_{accept} = \min\left[1, \frac{\exp[-H(\boldsymbol{p}(L),\boldsymbol{q}(L))]}{\exp[-H(\boldsymbol{p},\boldsymbol{q})]}\right]. \tag{11}$$

4. Repeat the procedure from step (1). If the sample is accepted, use $\mathbf{q}(L)$ as a new starting point, otherwise, return to the beginning of the trajectory and reuse model $\mathbf{q}$.

For technical aspects and choice of tuning parameters of HMC we refer the reader to Fichtner and Simutė (2018) and supplementary information (Text S3 in Supporting Information S1).

### 5.3. Inversion Parameters

In this study we seek a centroid-moment tensor solution, which means we simultaneously infer a centroid location, centroid time and a moment tensor of an earthquake (Dziewoński & Woodhouse, 1983). We invert for a full moment tensor, that is, six independent components $M_{ij}$. By not imposing any constraints on the faulting mechanism, we allow the mechanism to be determined freely by the data. If a parameter, or a combination of parameters, cannot be constrained by the data, that is, it lies in the null space of the model space, this shall be seen in the uncertainties provided by the probabilistic inference. For the comparison purposes, we also run a separate inversion for each case imposing a zero-trace constraint on the moment tensor, which denotes a source without the isotropic component.

Working with the moment tensor components $M_{ij}$ is a subjective choice, and various alternatives exist (Tape & Tape, 2013). An advantage of the probabilistic approach used in this work is that the subjective component is explicit, and that it can be modified via a simple re-parameterization of the involved probability densities. Hence, if needed, results can easily be presented in any different parameterization, without suffering from subjective regularization bias.

We work with three-component displacement velocity waveforms, which are low- and highpass filtered with a zero-phase, three-corner filter from the ObsPy library (Beyreuther et al., 2010). For the misfit computation we select measurement windows from the full waveforms manually. We normalize the amplitudes of each window to the largest of the corresponding event and frequency band in order to preserve information carried by the body waves, which would otherwise be suppressed by the larger-amplitude surface waves. Our measurements include body waves and fundamental- and higher-mode surface waves. No arbitrary waveform shifting is required, as all relevant time shifts are taken care of by the previous full-waveform inversion.

Assuming uncorrelated Gaussian data noise, our data covariance matrix $\boldsymbol{C_D}$ is a diagonal matrix, entries of which we conservatively estimate from the pre-signal noise. We express prior probability density on model parameters

as Gaussian distributions with standard deviations around 10 times larger than the parameter mean for the moment tensor elements ($1 \times 10^{17}$ N m), and 2°, 20 km, and 2 s for horizontal location, depth, and centroid time, respectively. As a prior mean we commonly use the solution provided by the GCMT (Dziewoński et al., 1981; Ekström et al., 2012). However, for some events (IDs 20050816, 20040407) we update the location first by running a preliminary inversion with a reduced number of samples, and then use the posterior maximum-likelihood model from this inversion as a prior mean for the main inversion. The choice of our prior is rather an attempt to encode ignorance than to represent actual prior information, which is not really available. Our goal is to ensure that the posterior distribution is primarily controlled by the data and not biased too much by the prior. We refer to the posterior maximum-likelihood model or maximum-likelihood model in short, as the one having the minimum potential energy $U$ (Equation 8).

The choice of the parameterization and a corresponding prior is somewhat philosophical because it is inherently subjective. The only exceptions are when the prior is determined by fundamental physics, which is not the case here. In a simpler example, one may perform a seismic tomography in terms of velocity and a homogeneous prior in velocity, or with slowness and a homogeneous prior in slowness. Being Jeffreys parameters, one is inherently as good as the other. What is important, is that these choices are made explicit and that the results are interpreted while taking these choices into account.

### 5.4. Multi-Period Band Inversion

Seismic waves traveling through complex geology for many wavelengths accumulate complicated path effects, such as frequency-dependent scattering or focusing in the presence of seismic velocity heterogeneities, which become more pronounced at shorter periods (e.g., Ferreira & Woodhouse, 2007; Igel & Gudmundsson, 1997; Igel et al., 2002). While our tomographic model can explain the majority of the waveforms in the 15–80 s period band, some complexities remain unaccounted for. This is primarily because strongly heterogeneous geology (e.g., accretionary prisms in the subduction zones) cannot be fully resolved by our limited data, especially due to non-uniform source-receiver distribution.

As expected, our model explains longer-period data better (Figure 3), but omitting shorter periods reduces the information content carried by the waveforms. Hence, we are faced with a trade-off between a very good long-wavelength Earth model and the available short-period information, which is necessary to constrain a full moment tensor, including its isotropic component.

In our approach we perform multi-period band inversions, which means that we invert the same event using different period data, that is, 15–80, 20–80, 30–80, and 50–80 s. We start with the short-period data inversions, that is, 15–80 s, and gradually expand the period band until an adequate waveform fit between the synthetic data for the maximum-likelihood model and the observed data is achieved. In the analysis we only consider those event—period-band configurations for which a plausible source model exists. This step is needed to account for uneven Earth model resolution. We determine whether the waveform fit between the observed and the synthetic data for the maximum-likelihood model is adequate using the $L_2$ misfit (Equation 6) and visual evaluation. The $L_2$ misfit is suitable for the inversion, where each event is considered separately, however, the absolute comparison across events is problematic, because the misfit depends on the subjectively assigned data error, that is, the estimated data noise and its assumed distribution. Therefore, in addition to the $L_2$ misfit, we also evaluate misfits visually.

At shorter inversion periods, the misfits across events vary more significantly than at longer periods (Figure 7). Waveforms of some events, which tend to be deeper, are explained better than those from other events, which tend to be shallower. This is illustrated in the right plot of Figure 7, where the waveforms for event 20150314 (GMCT depth 13.6 km) are not well explained at short periods, but the fit becomes adequate at long periods. Event 20050816 (GMCT depth 51.1 km), on the other hand, has an adequate waveform fit throughout all the period bands. In our approach by varying the period, we seek an Earth model, which could largely explain the observed data and could therefore, be used for the source inversion.

In the following section we will investigate the results of three events inverted with 15–80, 30–80, and 50–80 s data and discuss the differences in solutions as seen by different period data. We will then group the events according to their shortest acceptable inversion period and provide a general overview of events from each period

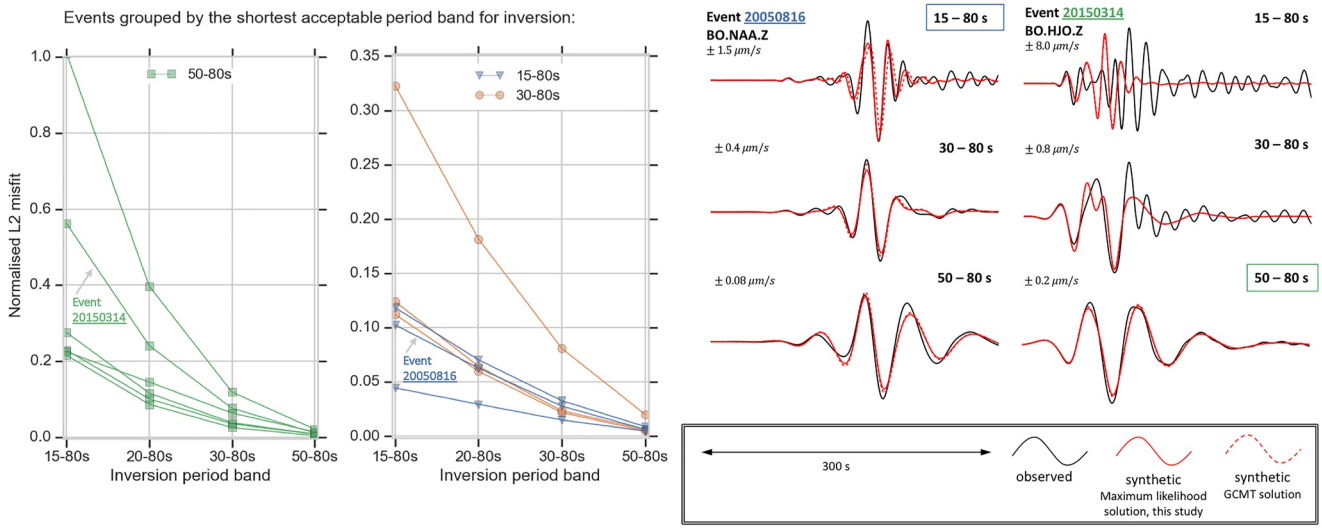


**Figure 7.** (Left) $L_2$ misfit (Equation 6) between the observed and the synthetic data for the maximum-likelihood model from the inversion at each of the four period bands. Each line corresponds to an event and color to the shortest acceptable period band at which a plausible source model exists. Data are plotted on two subplots for clarity; note a scale difference. The misfits are computed for the vertical component and are normalized to the largest value for visualization. Misfits have a high variability at short periods (0.049–1), but converge to a similar value for long-period inversions (0.004–0.05), indicating that data are not equally well explained for events at short periods. (Right) Waveform fit for two selected earthquakes in different period bands. The acceptable inversion period band of event 20050816 (in blue) is 15–80 s. It has a good waveform fit at short, as well as long periods. The acceptable inversion period band of event 20150314 (in green) is 50–80 s. For all events, waveforms are better explained with increasing periods.

band: three aforementioned events inverted with 15–80 s period data, one event with 20–80 s, three events with 30–80 s, and six events inverted with 50–80 s period data.

## 6. Inversion Results

### 6.1. Multi-Period Band Inversions of Selected Events

We present three events, each inverted with 15–80, 30–80, and 50–80 s period data, and discuss how data content affects the inversion results. The comparison is only possible for a small collection of events, for which the minimum misfit in all frequency bands is acceptable. Results, in terms of the maximum-likelihood solutions, are presented in Table S2 in Supporting Information S1.

#### 6.1.1. Moment Tensor Solutions

We present ensembles of effective moment tensor solutions in Figure 8. While the inversion results from 30 to 80 and 50 to 80 s period data are similar to each other, the short-period inversion (15–80 s) constrains a more distinctive source mechanism. Results from the short-period inversion reveal a significantly larger DC component compared to the longer periods. Our 3-D Earth model enables us to use short-period data, which, in turn, minimizes, what appears to be, an apparent CLVD component. At long periods, despite the 3-D Earth model, the inherent trade-offs in the source parameters still persist, and an increase in DC component is smaller (Dufumier & Rivera, 1997; Fitch et al., 1981).

Using the laterally averaged 1-D model, as in Figure 2, only allows us to work in the 50–80 s period band because data fit is too poor at shorter periods. As a consequence, the available waveform information is insufficient to modify the prior significantly. The maximum-likelihood solution for the 1-D model is therefore close to the GCMT solution. This confirms that the quality of source inversions in densely instrumented regions is mostly limited by the quality of the Earth model and less by coverage. Though the comparison is made with respect to a 1-D Earth model, one-dimensionality is by itself unlikely to result in a small DC component. Instead, it is the resulting limitation to long period-data and the large null-space that it entails.

We note that using a 3-D Earth model and fixing the isotropic component to zero allows for a greater DC component compared to the GCMT solution across all period band inversions, with the largest DC share typically

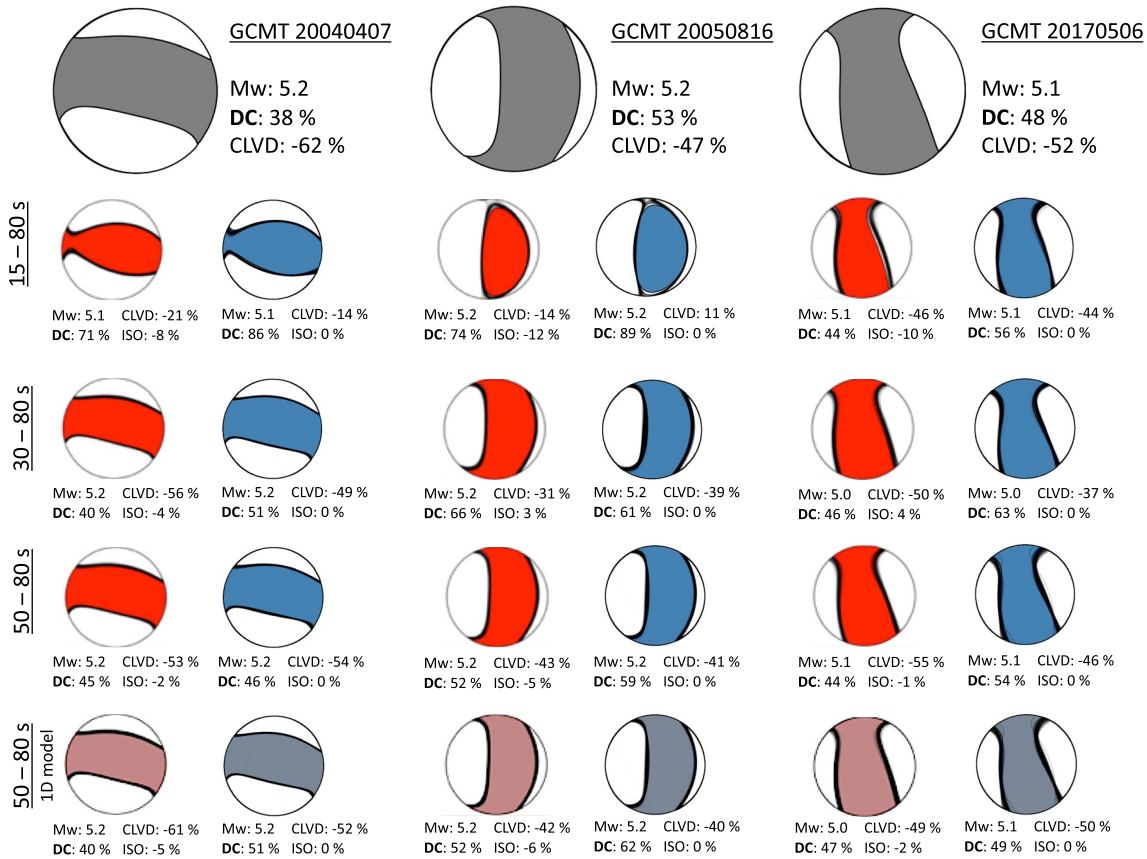


**Figure 8.** Ensembles of the focal mechanisms for three events inverted with 15–80, 30–80, and 50–80 s period data. Gray lines within the beachballs represent every 100th accepted model and the colored mechanisms correspond to the maximum-likelihood model. Red beachballs represent unconstrained inversions, and the blue ones inversions where the isotropic component is fixed to zero. Gray beachballs at the top represent the Global Centroid-Moment-Tensor (GCMT) solution. The double-couple component of the constrained inversion exceeds that of the GCMT for all events in all period bands. Generally, the double-couple component reaches the highest value for the shortest-period data inversion and decreases with increasing period band. The bottom row with beachballs in pale colors shows solutions in the 50–80 s period band based on Green's functions for the 1-D laterally averaged model, used before in Figure 2.

reached at the shortest periods. Double-couple components in constrained inversions tends to be higher than in the corresponding unconstrained inversions.

### 6.1.2. Event Locations

The ensembles of depth and horizontal location are presented in Figure 9. There is variability of centroid locations across different period bands. Similarly to the moment tensor solutions, the centroid locations from the short-period data inversions deviate more significantly from the longer-period ones as well as from the GCMT solution. We note that event relocations are accompanied by a change in the moment tensor.

The most significant difference between the short- and the long-period depth solutions is seen for event 20050816. Long-period data prefer locations near the GCMT solution, while the location mean from the short period inversion (15–80 s) is around 9 km shallower and around 10 km northward than the corresponding mean from 50 to 80 s data inversion.

### 6.1.3. Posterior Probability Densities

Generally, shorter-period data put tighter constraints on the source parameters compared to the longer-period data. This is reflected in the variances of individual model parameters for 15–80 and 50–80 s period data inversions (Figure 10 and Figure S8 in Supporting Information S1). Variances of the moment tensor components for short-period data inversion are 17% smaller than those at long period for event 20040407 and 35% for event 20050816. The largest differences are for the diagonal elements ($M_{xx}$, $M_{yy}$, $M_{zz}$) of the moment tensor, indicating that including short-period data improves the resolvability of the isotropic component. In the case of event

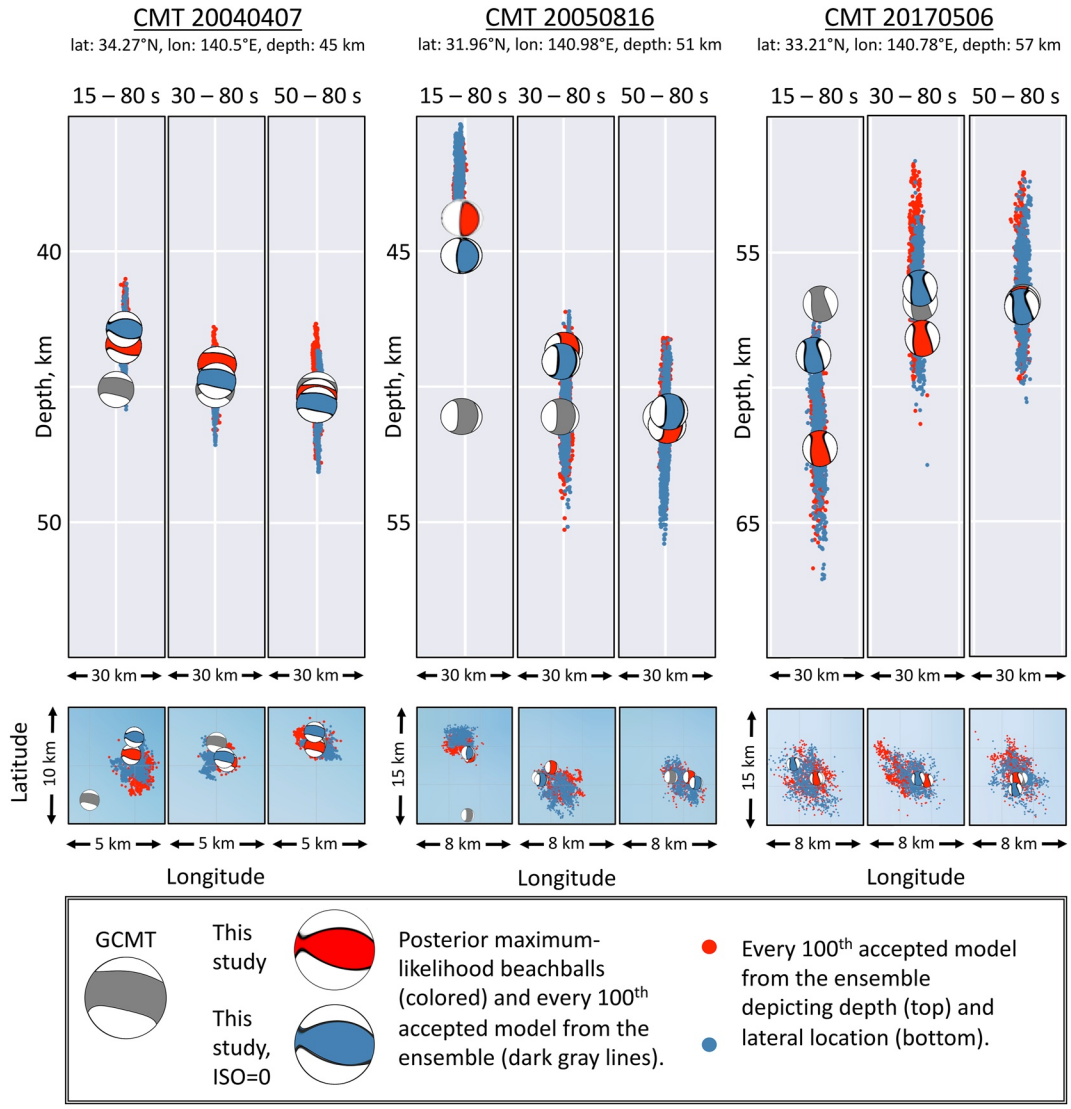


**Figure 9.** Ensembles of the location for three study events inverted with data of three different period bands, 15–80, 30–80, and 50–80 s. Red and blue dots represent every 100th accepted model from the unconstrained and constrained inversions, respectively. The maximum-likelihood model from each inversion is represented as a beachball following the same color code. Gray beachballs correspond to the Global Centroid-Moment-Tensor (GCMT) solution. The model space around the GCMT location of event 20050816 has been explored at short-periods during the preliminary inversion, which was performed before the main one in order to update the event location used as a prior mean (Figure S6 in Supporting Information S1).

20040407, location parameters have similar constraints from short- and long-period data inversions, while depth is much better constrained with short-period data for event 20050816 (Figure 10).

To demonstrate the inter-parameter trade-offs, we plot 2-D probability density marginals between the depth and all the other model parameters for 15–80 and 50–80 s period inversions (Figure 11). When shorter-period data are used, the posterior probability density distributions are more complicated (e.g., depth vs. $M_{yz}$, $M_{xx}$) and appear multimodal. Because the multi-parameter distributions are not exactly Gaussian, maximum-likelihood models do not correspond to the mean values and might lie outside of the main probability density volumes (Figure 11).

Depth trades offs with the majority of the moment tensor components. For most of the parameters the correlations are consistent for short- and long-period data inversions (e.g., depth vs. $M_{xx}$, $M_{yy}$). However, a positive trade-off between depth and $M_{zz}$ at short periods becomes a negative one at long periods. When compared across different events, we see that the trade-offs are not consistent, suggesting that it might be the earthquake mecha-

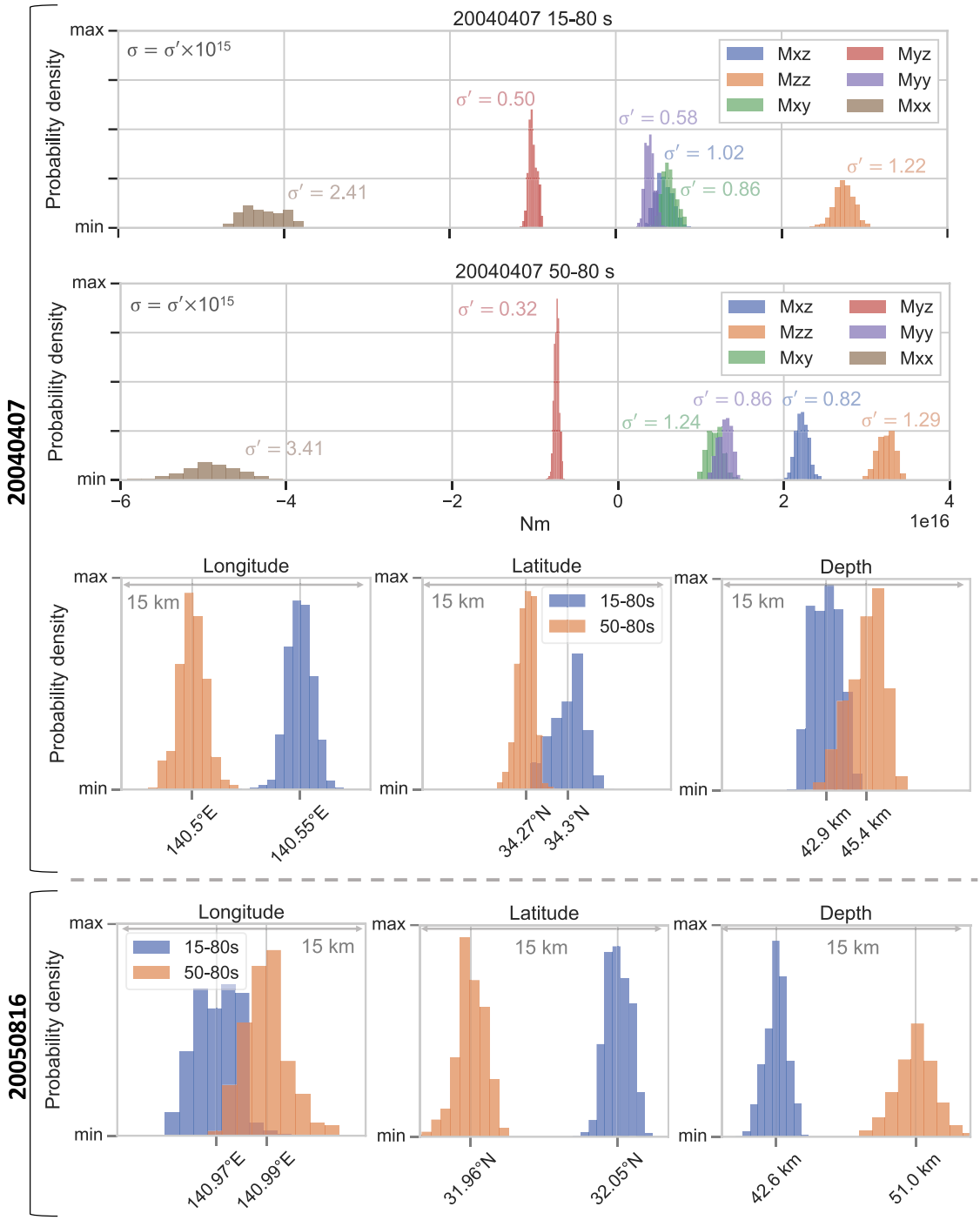


**Figure 10.** Comparison of marginal probability densities inferred with 15–80 and 50–80 s data for events 20040407 and 20050816. Marginal probability densities for moment tensor elements of event 20040407 are shown in the top two graphs, plotted on the same scale. Also shown are the standard deviation values for each parameter. Marginal probability densities for location parameters for events 20040407 and 20050816 are shown below. Generally, shorter-period inversion constrains the inversion parameters better. This is especially true for the diagonal elements ($M_{xx}$, $M_{yy}$, $M_{zz}$) of the moment tensor.

nism or location responsible for such a variability. For event 20050816, for example, we see notorious trade-offs between depth and $M_{yz}$ and $M_{zz}$ at long periods, which are successfully reduced when the inversion is performed with shorter-period data (Figure 12). More parameters for event 20050816 are shown in Figure S9 in Supporting Information S1.

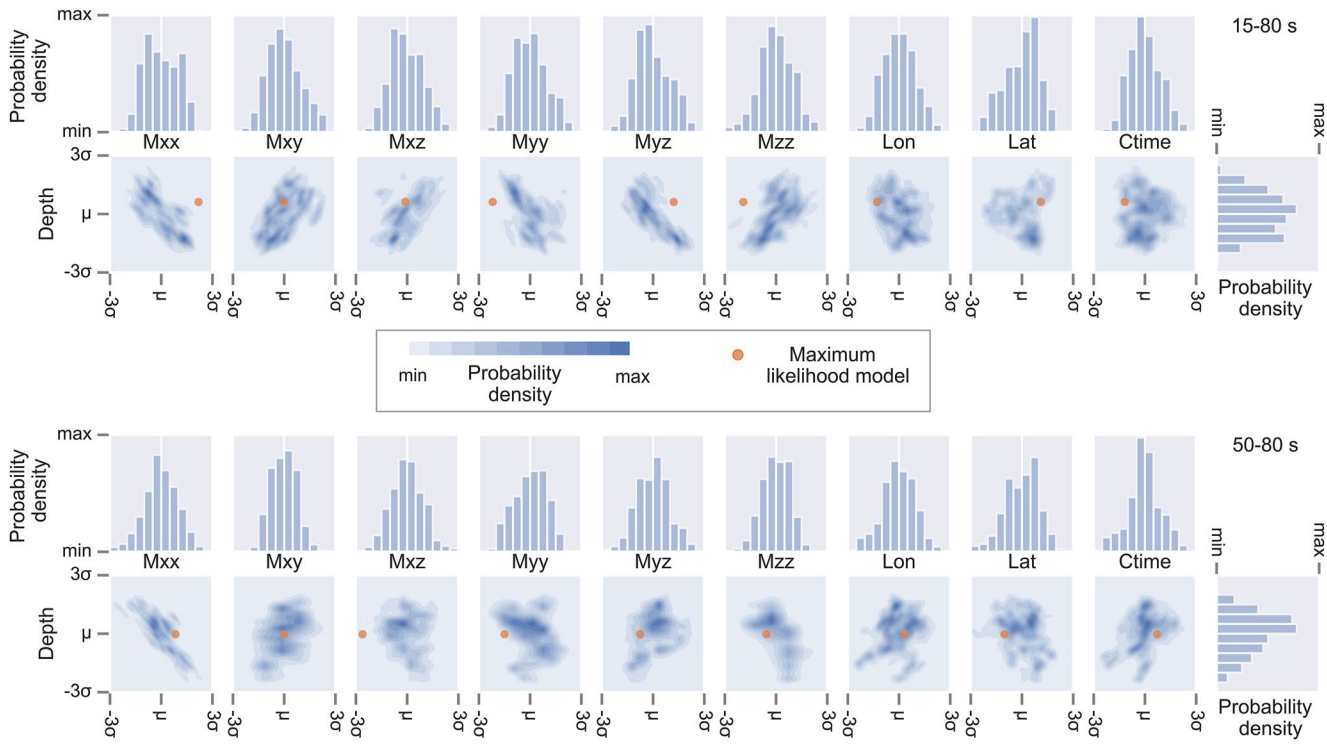


**Figure 11.** Comparison of selected trade-offs and marginal probability density functions for 15–80 s (top) and 50–80 s (bottom) period data inversion of event 20040407. The limits for each parameter depend on the corresponding standard deviations, and are set from $\mu - 3\sigma$ to $\mu + 3\sigma$, where $\mu$ is the mean and $\sigma$ is the standard deviation of the distributions. When shorter-period data are used, the posteriors are more complicated (e.g., $M_{yz}$, $M_{xx}$) and sometimes have several maxima. Because the multi-parameter distributions are not exactly Gaussian, maximum-likelihood models do not correspond to the mean values and for some parameters lie outside of the main probability densities. Depth trades off with the majority of the moment tensor components. For most of the parameters the trade-offs remain consistent for short- and long-period data inversions, however it is not always true (e.g., $M_{yz}$ and $M_{zz}$). Although for this particular event, trade-offs seem to be stronger when short-period data are used, it is not a general feature across other events.

### 6.2. Waveform Fit

In Figure 13, we present the waveform fit for two events inferred with 15–80 s period data. Compared to the GCMT solution, the fit improves by 4% for event 20050816% and 7% for event 20040407 for the unconstrained inversion. The maximum-likelihood solution from the constrained inversion usually gives slightly worse fit compared to the unconstrained one, the difference is 1% for the event 20040407 presented in the lower panel of Figure 13. Although the numeric waveform fit improvement is relatively small, its effect on the source mechanism is significant. An improvement of 4% for event 20050816 means 21% increase in DC component (Figure 8) and a depth relocation of 7 km (Figure 9). Hence, large variations in source parameters are hidden in the subtle waveform differences, which are possible to extract only by virtue of a good Earth model. Waveform data for other events are available at the Zenodo repository (Simutė, 2022).

### 6.3. Statistical Analysis

Here we present all events inverted in their shortest acceptable period band (see Section 5.4) and discuss those features of the results, which manifest across all the inversions. Results, in terms of the maximum-likelihood solutions, are presented in Table S3 in Supporting Information S1.

#### 6.3.1. Moment Tensor Type Across Different Inversions

The inversion results from short-period data (15–80 and 20–80 s) show a significant increase in the DC component, when compared to the GCMT solutions (Figure 14). In the case of unconstrained inversion, DC component increases by up to 33% (event 20040407), with isotropic component not exceeding 12%. The source mechanisms from the constrained inversion have even stronger DC component. It increases by up to 48%, and reaches more than 85% for all but one earthquake.

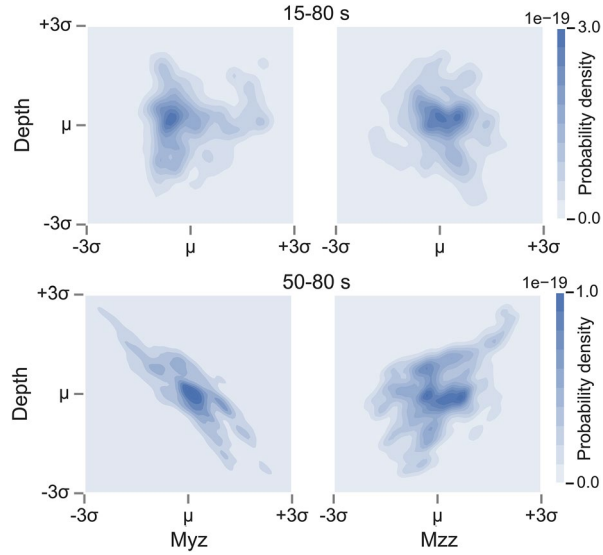


**Figure 12.** Comparison of selected trade-offs for 15–80 s (top) and 50–80 s (bottom) period data inversions of event 20050816. The limits for each parameter depend on the corresponding standard deviations and are set from $\mu - 3\sigma$ to $\mu + 3\sigma$, where $\mu$ is the mean and $\sigma$ is the standard deviation of the Gaussian-approximated distributions. For this event, the notorious trade-offs between depth and $M_{yz}$ and $M_{zz}$ are reduced when the inversion is performed with shorter-period data.

The results obtained with longer-period data (30–80 and 50–80 s) follow a different trend. Maximum-likelihood source mechanisms as well as locations stay similar to the GCMT values (Figure 15). In the case of 30–80 s inversion, moment tensors from the unconstrained inversion show only a slight increase in DC component (2%–3%) with isotropic component not exceeding 2%. Source solutions inferred with 50–80 s period data (Figure 15) have negligible deviations from the GCMT solutions, both in terms of moment tensor and location. This means that inversions at long periods either have weaker constraints from the data and are dominated by the prior or that the likelihood term at longer periods favors the same solution as that given by the GCMT.

An overview of depth relocation across different inversions is presented in Figure 16. Similar to the source mechanisms, the largest deviations in depth come from the shortest-period band inversions. Events inferred with 50–80 s period data maintain almost the same depth as given by the GCMT catalog.

### 6.3.2. Principal Component Analysis

The value of probabilistic inversion lies in the statistical inferences which can be made from the ensembles of the accepted models. We can extract the moments, such as the mean value or the variance of the distribution, of the separate model parameters, and after marginalization easily visualize the 1-D probability density distributions (Figure 10). To study inter-parameter trade-offs, it is common to plot 2-D marginal probability densities (Figure 11). However, anything higher than two dimensions becomes difficult to visualize and to study. In order to better understand the posterior distribution of the multi-dimensional parameter space, we propose to perform a principal component analysis (PCA).

PCA works by introducing a new coordinate system of the model space, which emphasizes the variation in the distribution. Each new axis, called principal component (PC), is a linear combination of the physical model parameters. Depending on the type of observations or model parameters, PCA can be performed either on the covariance or on the correlation matrix.

The covariance matrix is defined as:

$$cov(\boldsymbol{X}) = \mathcal{E}\left[\left(\boldsymbol{X} - \boldsymbol{\mu}\mathbf{1}_N^T\right)\left(\boldsymbol{X} - \boldsymbol{\mu}\mathbf{1}_N^T\right)^T\right] = \begin{pmatrix} \sigma_{11} & \sigma_{12} & . & \sigma_{1K} \\ \sigma_{21} & \sigma_{22} & . & \sigma_{2K} \\ . & . & . & . \\ \sigma_{K1} & \sigma_{K2} & . & \sigma_{KK} \end{pmatrix}, \quad (12)$$

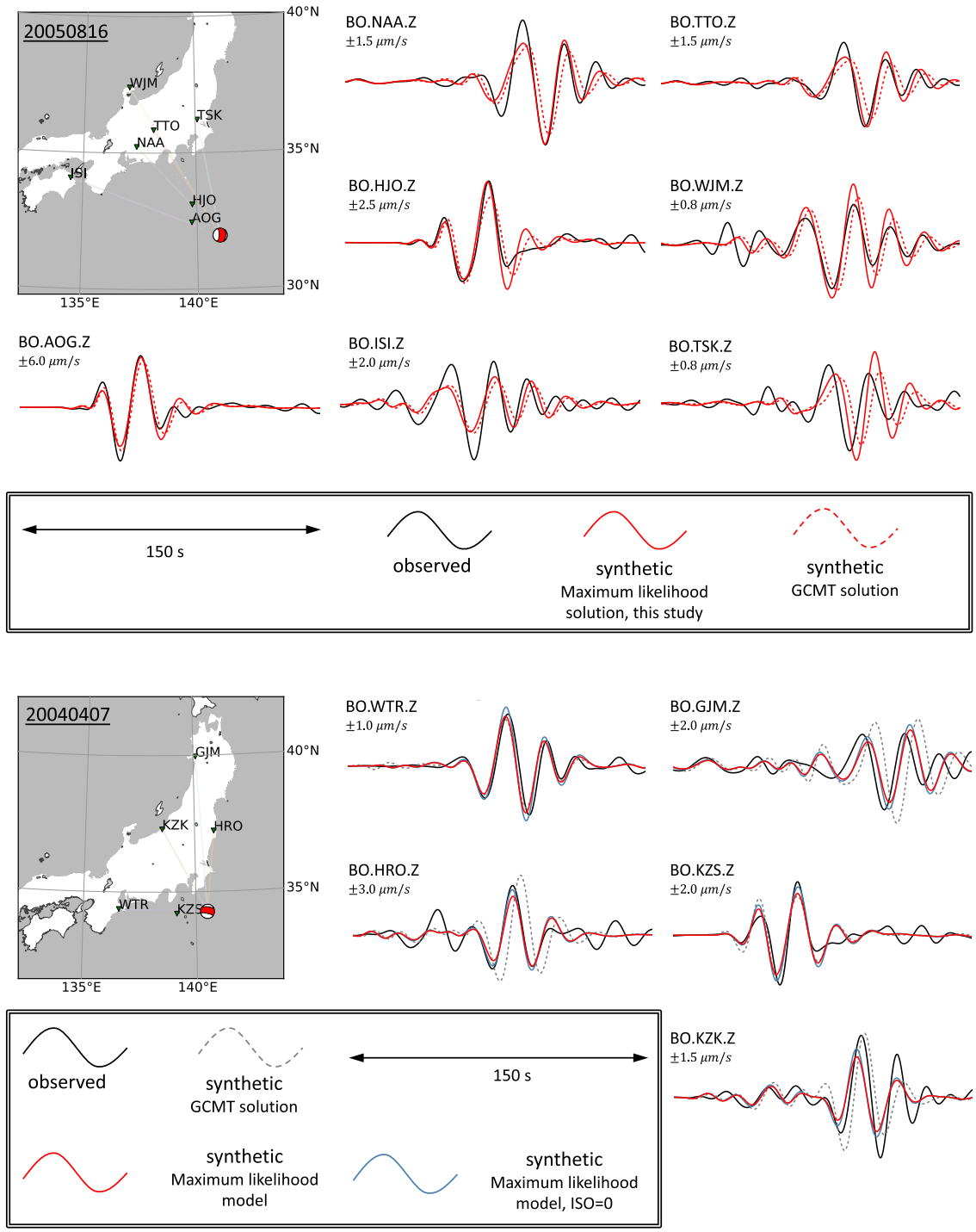


**Figure 13.** (Top) Waveform fit between the observed data (black), synthetic data for the maximum-likelihood model of 15–80 s period unconstrained inversion (solid red) and synthetics for the Global Centroid-Moment-Tensor (GCMT) solution (red dashed) for event 20050816. (Bottom) Waveform fit between the observed data (black), synthetic data for the maximum-likelihood model of 15–80 s period unconstrained inversion (solid red), synthetic data for the maximum-likelihood model from the constrained inversion (blue) and synthetics for the GCMT solution (gray dashed) for event 20040407.

where $\boldsymbol{X}$ is the ensemble of sampled model parameter vectors. It is a $K \times N$ matrix, where $K$ is the number of parameters, and $N$ is the number of samples. The model mean vector $\boldsymbol{\mu}$ is multiplied by an $N \times 1$ vector of ones, and $\mathcal{E}$ denotes the expected value of the product in the square brackets. Because our model parameters are incommensurable, that is, they have diverse physical units, such as s or Nm, and variances with orders of magnitude

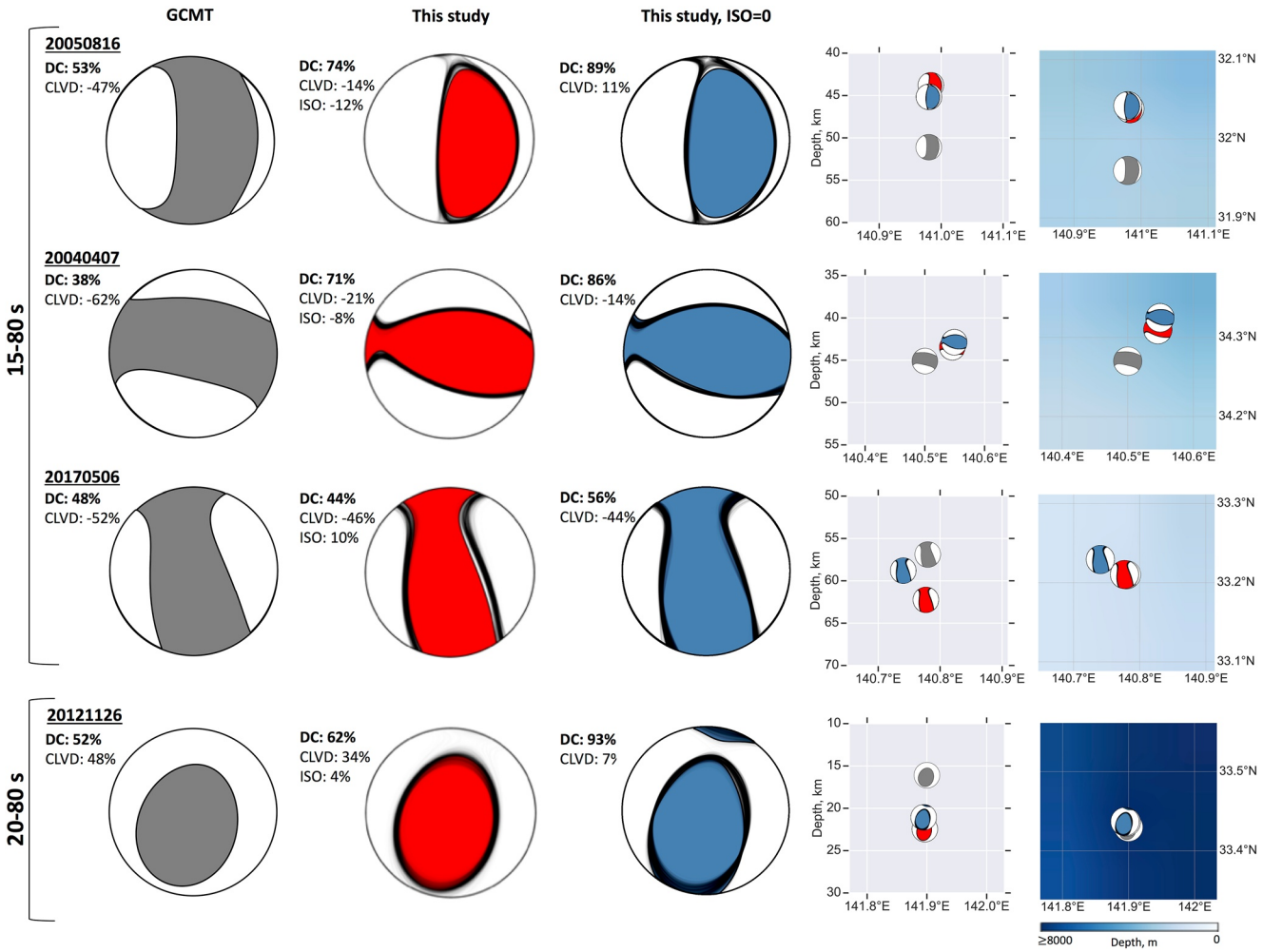


**Figure 14.** Moment tensor ensembles inferred with 15–80 s (top three) and 20–80 s period data (bottom one). Gray beachballs correspond to the Global Centroid-Moment-Tensor (GCMT) solution, red beachballs correspond to an unconstrained inversion, and blue ones correspond to the inversion, where we assume no isotropic component. Gray lines represent every 100th model of the ensemble, with the maximum-likelihood models colored. On the right, we show the maximum-likelihood locations, following the same color code. The horizontal relocation from the GCMT solution amounts to several kilometers.

ranging from $1 \times 10^{-4}$ to $1 \times 10^{30}$, we work with a correlation matrix (Chave, 2017). The correlation matrix is characterized by the normalized covariances, but nevertheless, retains the inter-parameter trade-offs:

$$corr(\boldsymbol{X}) = \begin{pmatrix} 1 & \rho_{12} & . & \rho_{1K} \\ \rho_{21} & 1 & . & \rho_{2K} \\ . & . & . & . \\ \rho_{K1} & \rho_{K2} & . & 1 \end{pmatrix}, \tag{13}$$

where $\rho_{jk} = \sigma_{jk} / \sqrt{\sigma_{jj}\sigma_{kk}}$.

The first PC, or an eigenvector with the corresponding largest eigenvalue, is responsible for the largest variance of the distribution. In our case this corresponds to the greatest uncertainty axis. The best-constrained direction of the multi-dimensional distribution is represented by the smallest PC, indicating the narrowest extent of the distribution. The exact share of the explained variance is proportional to the size of the eigenvalue. If there are no dominant eigenvalues and all of them are of a comparable size, it means that all the PCs, or the effective parameters, are resolved with a similar (un)certainty. The correlation matrix is then close to being proportional to

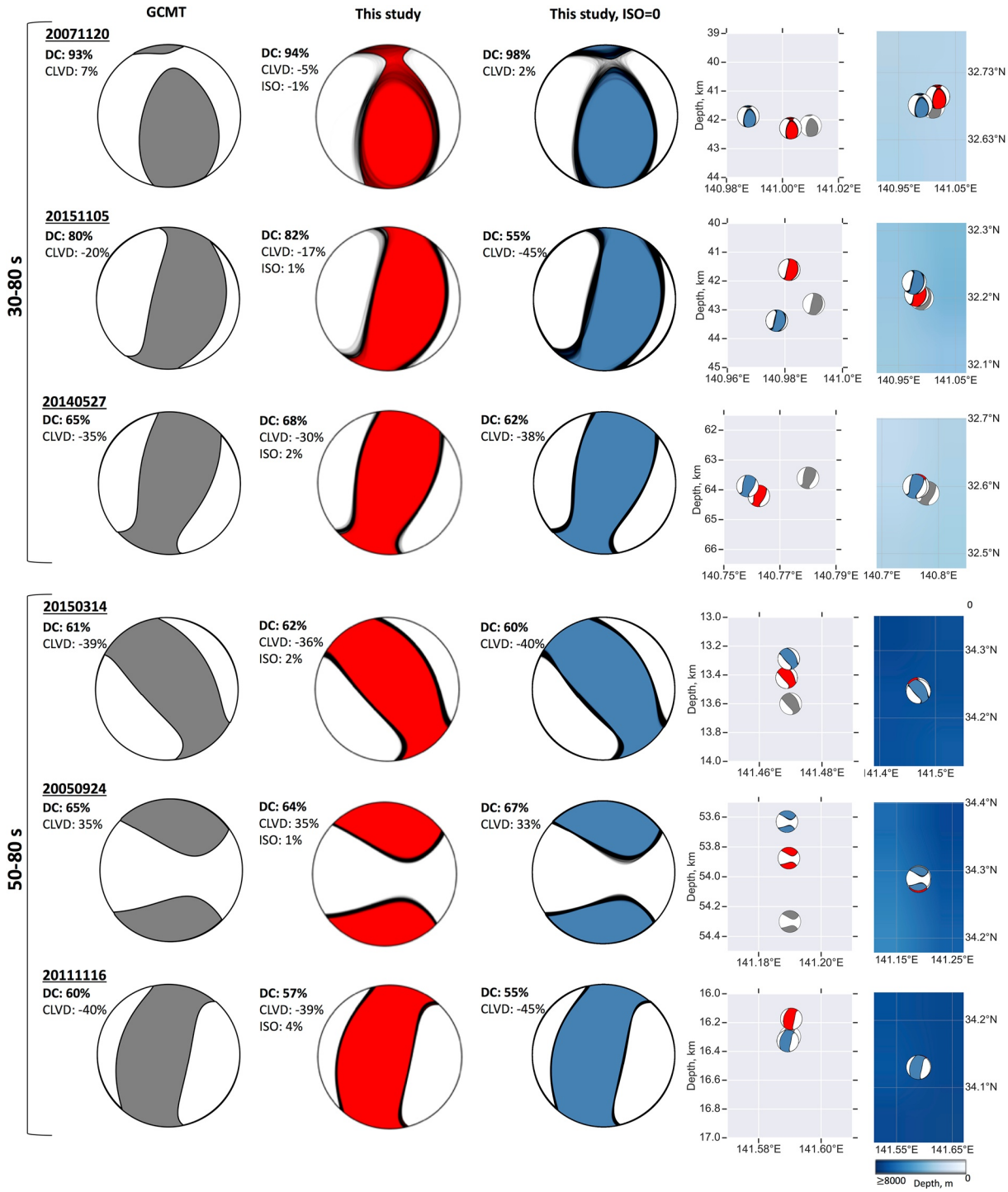


**Figure 15.** Moment tensor ensembles inferred with 30–80 s (top three) and 50–80 s period data (bottom three). Gray beachballs correspond to the Global Centroid-Moment-Tensor solution, red beachball correspond to an unconstrained inversion, and blue ones correspond to the inversion, where we assume no isotropic component. Gray lines represent every 100th model of the ensemble, with the maximum-likelihood models colored. On the right, we show the maximum-likelihood locations, following the same color code.

the identity matrix, resembling a multi-dimensional sphere (Chave, 2017). In a data set of observations, if some threshold variance, typically set between 70% and 90%, can be explained by much fewer PCs than a number of original parameters, one could retain only the most important PCs, and benefit from the reduced dimensionality of the problem (Chave, 2017).

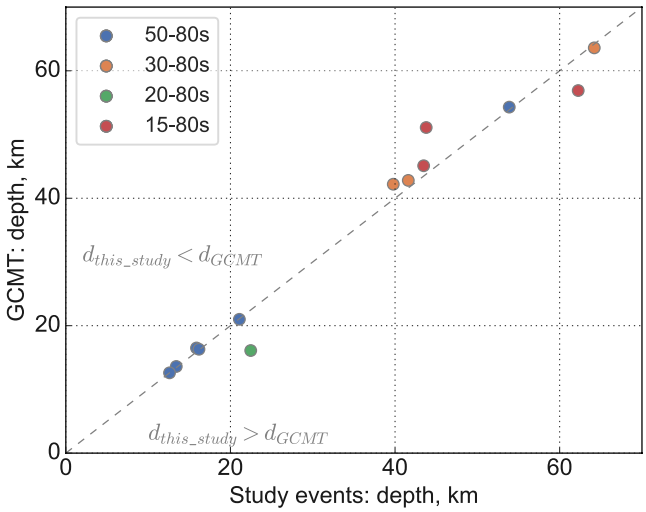


**Figure 16.** Comparison of earthquake depth in Global Centroid-Moment-Tensor (GCMT) and unconstrained inversion of our study, color-coded by the period band of the inversion. Events inverted with the shorter-period data tend to deviate more from the GCMT solution that those inverted with longer periods.

We study PCs for the unconstrained inversion of each study event at its corresponding shortest acceptable period band as presented in Section 6.3.1. The left column of Figure 17 illustrates the cumulative explained variance by the PCs for each event. We see that events inverted at longer periods (30–80, 50–80 s) reach the 70% variance threshold with the first three and 90% with the first five PCs, while those inverted at shorter periods (15–80, 20–80 s) need four for 70% and seven for 90% variance threshold. In other words, the magnitudes of the PCs from long-period inversions extend over a greater range than those from the shorter-period inversions (right plot of Figure 17). From this we can deduce that the shape of multi-dimensional distribution for longer-period inversions diverge from the sphere more than for the short-period inversions. This means that for an event inverted at long periods, some axes (e.g., PC 0) are much more difficult to constrain than others (e.g., PC 9), while at shorter periods, they can all be constrained more equally.

In order to ascertain whether these certainty and uncertainty axes are similar across all the inversions, we investigate similarity of the PCs among the different events by computing the dot product between PCs associated (a) with uncertainty (PC 0–PC 4) and (b) certainty (PC 5–PC 9) within the same period band inversions. Putting a threshold of 30°, we find no consistency among the axes corresponding to either the smallest or the largest eigenvalues. In other words, PCs are pointing to different directions for each event and no linear combination of physical parameters can be generalized to be the least- or the best-constrained direction in the model space. This suggests that we cannot easily reduce the dimension of the problem by ignoring some parameters, because the eigenvectors are different for each event. Because the data and the azimuthal coverage are similar for all the study events, it is likely that the source mechanism or the centroid location are responsible for the fact, that each event is characterized by a very different set of PCs.

## 7. Discussion

### 7.1. Increase of Double-Couple Component

Accounting for heterogeneous Earth structure enables us to use shorter-period data (15–80, 20–80 s) in regional source inversion. This leads to a higher double-couple component compared to GCMT solutions, and it suggests

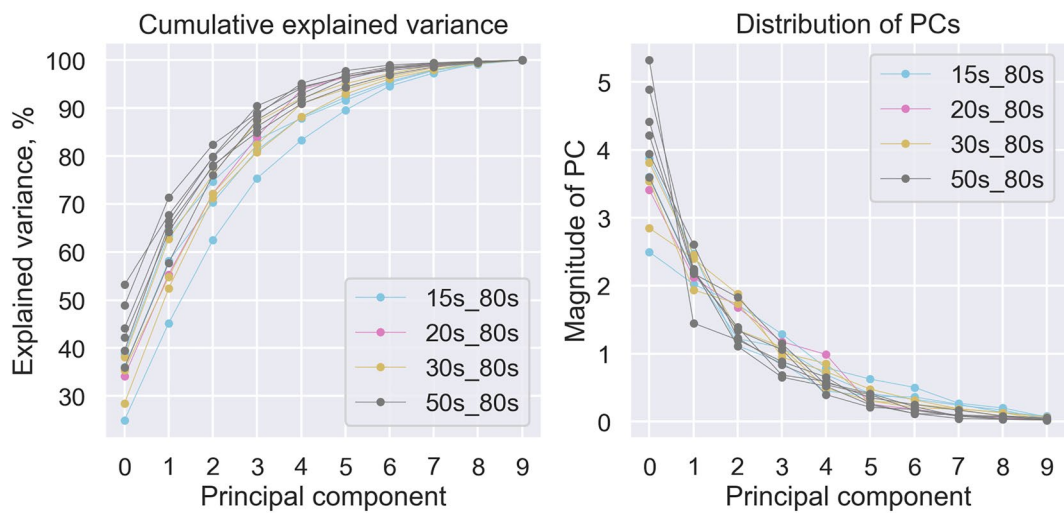


**Figure 17.** (Left) Cumulative explained variance with respect to the principal components (PCs) of the posterior correlation matrix of all study events within their acceptable period band. (Right) Distribution of eigenvalues, or PCs of the posterior correlation matrix.

that at least part of the CLVD component, given by the GCMT catalog, is likely to be an artifact caused by forward modeling errors. This is supported by the comparison of posterior distributions for the 3-D full-waveform inversion model and its 1-D lateral average, shown in Figure 8. A similar trend of increasing DC component when 3-D Earth structure is taken into account has been shown by Hejrani et al. (2017) and X. Wang and Zhan (2019) for earthquakes in the Papua New Guinea and Los Angeles regions, respectively.

### 7.2. Power of Subtle Waveform Differences

Significant changes in the source parameters, both in terms of the moment tensor and the location, may occur despite relatively small waveform differences. In one of the examples we showed that an overall waveform fit improvement of 4% led to 21% increase in DC component together with a depth relocation of 7 km. This implies that some source parameters or combinations thereof are highly sensitive to subtle waveform differences and Earth structure, and can therefore only be resolved with a reliable 3-D Earth model.

### 7.3. Effective Source Solutions

In this study we have presented effective, that is, period-dependent, point-source solutions, inferred with different-period data. In the following, we discuss means of how varying frequency data can lead to effective source parameters.

#### 7.3.1. Source Complexity

Spatial and temporal complexity of the source have been previously suggested to explain the discrepancy between short- and long-period source mechanisms (e.g., Frankel, 2013; Grandin et al., 2015; Wallace et al., 1982). For example, it has been argued that asperities on the fault surface generate shorter-period seismic waves, while the overall faulting episode is represented by the long-period waveforms (Wallace et al., 1982). However, a typical $M_w$ 5 event, such as those selected for our study, should not exceed several seconds in half-duration and a few kilometers fault surface radius, which makes it unlikely for 15 s period data to constrain a subfaulting episode, as the periods are much longer than the expected duration of the earthquake (Ekström et al., 2012; Eshelby & Peierls, 1957; Hanks, 1977).

#### 7.3.2. Near-Source Anisotropy

Another possibility, which could explain effective source solutions, is related to the near-source heterogeneities and specifically, anisotropy. Anisotropy arises from different, scale-dependent mechanisms (e.g., Backus, 1962; Kawakatsu et al., 2009; N. Wang et al., 2013). Such a frequency-dependent nature of observed anisotropy has indeed been reported beneath Japan (Wirth & Long, 2010) and other subduction zone regions (e.g., Fouch & Fischer, 1998; Greve & Savage, 2009). Anisotropy in the immediate vicinity of the source can affect wave propagation in such a way that a purely isotropic event might appear to have excited shear waves, while a shear earthquake might appear to have had a non-DC component (e.g., Kawasaki & Tanimoto, 1981; Li et al., 2018; Vavryčuk, 2004). The effective source solutions might arise from the fact that the fine-scale anisotropy is only captured by the shorter-period waves, while longer-period waves sample an effective medium over larger scales. Therefore, at short-period inversions (15–80 s), where the fine-scale anisotropy around the source is accounted for, the apparent CLVD component decreases, while at long periods (50–80 s), it remains relatively high.

#### 7.3.3. Information Content

Long-period data may not contain enough information to properly constrain the source parameters of relatively small-magnitude earthquakes. First, data are more correlated at longer-periods and hence, carry less independent information. This results in inherent trade-offs between moment tensor components, such as $M_{zz}$ and $(M_{xx} + M_{yy})$, which trade off in the case of long-period surface wave inversion (Dufumier & Rivera, 1997; Fitch et al., 1981). Second, the amplitude spectrum of small-magnitude events tends to decrease with increasing period (Aki, 1967), and hence, the signal-to-noise ratio at long periods may therefore, be diminished.

Long-period data used in this study, that is, body and surface waves in the 50–80 s period band, are comparable to the data used to construct the GCMT catalog. The reported minimum periods of body and surface waves, used to constrain the earthquakes chosen for our study, are 40 and 50 s, respectively (The Global CMT Project, 2021). Although the azimuthal coverage in the GCMT inversion may be more complete than in our study, due to the

available teleseismic data, the lack of seismic receivers in the north-western Pacific Ocean is evident (Ekström et al., 2012). Similarity between data periods could potentially explain why our long-period source solutions are in a close agreement with the GCMT predictions.

### 7.4. Effect of the Prior

The source parameters constrained with long-period data are often very close to the prior model, that is, the GCMT solution, which we use for the majority of the inversions. In the Bayesian framework, the fact that the posterior closely resembles the prior can mean two things: either the likelihood is very similar to the prior and hence prefers the same solution, or the data constraints are weak, and the posterior is dominated by the prior probability density distribution.

Although stochastic algorithms are direct search methods, theoretically giving us a chance to obtain probability density functions of a full model space, its practical implementation might be very expensive, especially in a high-dimensional space. With the aim to speed up the convergence and to alleviate computational costs, we introduced a modified version of HMC. If well tuned, HMC can be a very efficient sampling algorithm for its gradient-based approach of proposing the samples. The derivative should in principle be computed at every point along the Hamiltonian trajectory. However, this is expensive, and in our formulation we suggest to approximate the derivative around a prior mean in model space, instead (Fichtner & Simutė, 2018). The derivative is exact for the parameters linearly related to data (i.e., the moment tensor components), but it is an approximation for non-linear parameters (i.e., centroid). We note that this approximation only concerns the samples drawn to be proposed, and not the acceptance criterion, which should still ensure that even with this approximation, only relevant models are accepted. Furthermore, if we let our sampler run for an infinite amount of time, the way samples are proposed would not matter. However, during the finite run time, the posterior might actually be biased toward the point in model space around which the approximation is performed, which in our case is often prior mean.

### 7.5. Data and Modeling Errors

In this study we pragmatically treated observed data noise and forward modeling errors. We conservatively estimated data noise from the pre-signal recordings, and assumed it to be normally distributed without spatial or temporal correlation. This is a simplification, as data errors are expected to be correlated, especially at longer periods. For this conceptual demonstration we also did not account for the forward modeling errors. Efforts to properly incorporate modeling errors and data noise covariances in stochastic inversions were made by for example, Staehler and Sigloch (2014, 2017), Vackár et al. (2017), Hallo and Gallovič (2016), Duputel et al. (2012).

### 7.6. Isotropic Component

Due to its poor resolvability, the isotropic component is typically ignored in the source inversion studies (e.g., Hejrani et al., 2017; Lee et al., 2011; Liu et al., 2004; Staehler & Sigloch, 2014) and in most earthquake catalogs. Examples include the catalog of the Southern California Seismic Network (SCSN) (Clinton et al., 2006; Hutton et al., 2000), the Global CMT project (Dziewoński et al., 1981; Ekström et al., 2012), the NIED earthquake catalog in Japan (Kubo et al., 2002). The resolution of the isotropic component is limited by the fact that the $M_{zz}$ term does not excite Love waves, and in the Rayleigh wave excitation it trades off with the $(M_{xx} + M_{yy})$ term at long periods (Dufumier & Rivera, 1997; Fitch et al., 1981). Furthermore, body waves for both the isotropic component and the vertical CLVD are excited without the azimuthal dependence, which complicates their recovery with data recorded only at the surface (Dufumier & Rivera, 1997; Fitch et al., 1981; Kawakatsu, 1996). The isotropic component therefore, tends to be constrained to prevent unexplained waveform differences from being mapped into an additional parameter (e.g., Vavryčuk, 2004).

In this study we performed unconstrained inversion, thereby allowing the source parameters to be freely determined by the data. Through a reliable Earth model we improved the waveform fit, and owing to both body and surface wave measurements over a wider frequency range (15–80 s), the trade-offs between the $M_{zz}$ and the $(M_{xx} + M_{yy})$ terms might have been alleviated. As shown in Section 6.1.3, including shorter-period data in waveform inversion, indeed facilitates the resolution of the diagonal components. Due to a greater range of take-off angles, deeper events could further benefit the resolution of the isotropic component.

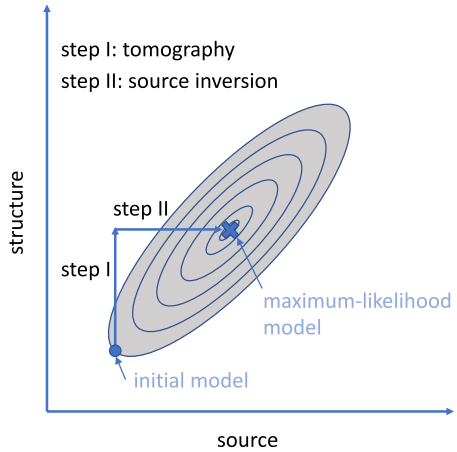


**Figure 18.** An idealized sketch of source and structure inversions in a two-step procedure.

### 7.7. Station Coverage

While hosting numerous seismic events, the choice of the Izu-Bonin trench as a source region is also challenging due to an azimuthal gap in the south-east. Although a good station coverage is typically sought after for a better source parameter resolution, it has been shown that waveform inversion of the fault plane solution is not significantly affected by the lack of azimuthal coverage (e.g., Cesca et al., 2006; Dreger & Helmberger, 1993). Furthermore, no linear combination of physical parameters was found in this study (Section 6.3.2) to be consistently worse resolved across all the events, which would be expected if station coverage was playing a limiting factor in source parameter resolvability. Lastly, our solutions at longer periods converge toward the GCMT solutions, suggesting that either stations used in both cases were similarly distributed (as discussed in Section 7.3.3), or that a 3-D structural model enhances the resolvability of source parameters, as also indicated by Donner et al. (2020).

### 7.8. Trade-Offs Between Structure and Source

Because of the trade-offs between Earth structure and earthquake source parameters, it is challenging to constrain them independently (e.g., Hejrani et al., 2017; Hjörleifsdóttir & Ekström, 2010; Morales-Yáñez et al., 2020). In this study, we proceed with a two-step approach: first we constrain the structure, then, with a suitable Earth model at hand, we aim to recover improved source parameters. In an idealized scenario, where the probability density distribution between the source and the structure parameters would be Gaussian, this procedure would look rather straightforward (Figure 18). However, the real world presents many complications. First, we do not have much constraints on the model space of the Earth structure except for the least-squares solution. Second, such a unimodal distribution might be an oversimplification. Lastly, the probability density distribution is likely to be frequency-dependent, with increasing complexity at increasing frequencies. Therefore, locating a global minimum in the source—structure space might actually be a very difficult task, especially at shorter periods. Such a two-step approach could potentially favor the original source solutions, used in the tomographic study. We attempted to reduce such a dependency by using the NIED solutions for structural inversion and the GCMT solutions for comparison and as prior means in the moment tensor inversions.

## 8. Conclusions

We presented results of a probabilistic seismic source inversion for 13 small-to-moderate magnitude offshore earthquakes at the Izu-Bonin trench. The inversions were conducted using fully heterogeneous, radially anisotropic Green's functions and the HMC sampling algorithm. We simultaneously inferred centroid location, centroid time and six independent moment tensor components, and ran a separate inversion constraining the isotropic component to zero for comparison purposes. With the goal to use a sufficiently good Earth model, we varied the minimum inversion period and limited ourselves to those period bands, for which the data fit between the observed data and the maximum-likelihood solution was sufficiently good.

Accounting for 3-D Earth structure at short periods (15–80, 20–80 s) generally leads to an increase in DC component compared to the GCMT solution (Figure 14). This suggests that at least some part of the non-DC component in the GCMT catalog might be apparent—resulting from unmodeled Earth structure. Events inverted at longer periods (e.g., 50–80 s) (Figure 15), do not show a significant change in mechanism or centroid location and stay close to the GCMT predictions. We have presented several possible mechanisms to explain the effective solutions, the most likely being near-source propagation effects, which cannot be resolved by long-period data, or weakening data constraints with increasing periods.

Constraining isotropic component to zero is a common practice in order to stabilize the inversion and to prevent unexplained waveform differences from mapping into the additional parameter. Here, we observe that owing to the 3-D Earth model, isotropic component remains weak even when unconstrained.

The diagonal components of the moment tensor, which are responsible for the volumetric change, are significantly less constrained, compared to the off-diagonal ones (Figure 10 and Figure S8 in Supporting Information S1). From the events which could be inverted in all period bands, we also see that shorter-period data constrain the source parameters better than the long-period data (Figure 10 and Figure S8 in Supporting Information S1). Posterior probability density distributions of shorter-period data inversions appear multimodal and are more complicated than those of long-period. This illustrates non-uniqueness of short-period source inversions and highlights the need for stochastic approaches.

We note that finding a solution, which has a significantly better waveform fit than that provided by the GCMT catalog, is a challenging task. In our examples, the waveform fit only improves by several percent (Figure 13). However, a small change in the waveforms brings about a significant change in the source solution. In other words, large variations in source parameters are hidden in the subtle waveform differences, which are possible to extract only by virtue of a good Earth model.

To aid the study of the multi-dimensional posterior, we perform PCA. We include all the study events at their shortest acceptable inversion period band, with the aim to retrieve the best- or the least-constrained direction in model space. By comparing the PCs (eigenvectors) corresponding to either the smallest or the largest eigenvalues, we find no consistency among different events from the same period-band inversion group. Each event is characterized by a very different set of PCs, and no linear combination of physical parameters stands out as the least- or the best-constrained direction in model space.

In this conceptual study, we detailed the methodology for a probabilistic source inversion using 3-D Green's functions and presented a proof-of-concept catalog of source solutions. Such an approach allows us to better constrain source characteristics and comes with the ensemble statistics, such as uncertainty limits and inter-parameter trade-offs. Inferred source parameters contribute to our understanding of the regional seismotectonics and earthquake physics and can also be fed back into and, potentially, improve tomographic studies.

## Data Availability Statement

All seismic waveform data used in this study are freely available from the Full Range Seismograph Network of Japan (F-Net, http://www.fnet.bosai.go.jp), the Broadband Array in Taiwan for Seismology (BATS, http://bats.earth.sinica.edu.tw), the Korea National Seismograph Network (http://www.kma.go.kr/weather/earthquake/internationallist.jsp), and the China National Seismic Network, the New China Digital Seismograph Network, the Northeast China Extended Seismic Array, the Global Seismograph Network, and the Korean Seismic Network, made available by the IRIS Data Management Center (http://ds.iris.edu/ds/nodes/dmc/). The centroid moment tensors were obtained from the Global Centroid-Moment-Tensor Catalog (www.globalcmt.org, http://ds.iris.edu/spud/momenttensor) and National Research Institute for Earth Science and Disaster Prevention Seismic Moment Tensor Catalog (https://www.fnet.bosai.go.jp/). Seismic wave propagation software, SES3D, used to model the waveforms is available on https://cos.ethz.ch/software/production/ses3d.html. Seismic tomography model of the Japanese islands, developed and used in this study, can be downloaded from https://cos.ethz.ch/research/CSEM.html. Processed observed as well as synthetic waveform data for all the study events are available at https://doi.org/10.5281/zenodo.7384213 (Simutė, 2022).

**Acknowledgments**
We express our gratitude to Göran Ekström, Andreas Zunino, Jean-Paul Montagner, Pascal Bernard, and Claudio Satriano for useful discussions and insights about this study. We also thank the Editors and the three anonymous reviewers for their constructive comments, which helped to improve the manuscript. This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s1040 and by the European Unions Horizon 2020 research and innovation program through an ERC Starting Grant (The Collaborative Seismic Earth Model, Grant 714069).

## References

Afanasiev, M., Boehm, C., van Driel, M., Krischer, L., Rietmann, M., May, D. A., et al. (2018). Modular and flexible spectral-element waveform modelling in two and three dimensions. *Geophysical Journal International*, *216*(3), 1675–1692. https://doi.org/10.1093/gji/ggy469

Aki, K. (1967). Scaling law of seismic spectrum. *Journal of Geophysical Research*, *72*(4), 1217–1231. https://doi.org/10.1029/JZ072i004p01217

Aki, K., & Richards, P. (2002). *Quantitative seismology*. University Science Books.

Alder, B. J., & Wainwright, T. E. (1959). Studies in molecular dynamics. I. General method. *Journal of Chemical Physics*, *31*(2), 459–466. https://doi.org/10.1063/1.1730376

Aleardi, M., Salusti, A., & Pierini, S. (2020). Transdimensional and Hamiltonian Monte Carlo inversions of Rayleigh-wave dispersion curves: A comparison on synthetic datasets. *Near Surface Geophysics*, *18*(5), 515–543. https://doi.org/10.1002/nsg.12100

Babuška, V., & Cara, M. (1991). *Seismic anisotropy in the Earth*. Kluwer Academic Publishers.

Backus, G. E. (1962). Long-wave elastic anisotropy produced by horizontal layering. *Journal of Geophysical Research*, *67*(11), 4427–4440. https://doi.org/10.1029/jz067i011p04427

Backus, G. E., & Mulcahy, M. (1976a). Moment tensors and other phenomenological descriptions of seismic sources—I. Continuous displacements. *Geophysical Journal of the Royal Astronomical Society*, *46*(2), 341–361. https://doi.org/10.1111/j.1365-246x.1976.tb04162.x

Backus, G. E., & Mulcahy, M. (1976b). Moment tensors and other phenomenological descriptions of seismic sources—II. Discontinuous displacements. *Geophysical Journal of the Royal Astronomical Society*, *47*(2), 301–329. https://doi.org/10.1111/j.1365-246X.1976.tb01275.x

Bayes, T., & Price, R. (1763). An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. *arXiv e-prints*. arXiv:1701.02434 [stat.ME].

Beyreuther, M., Barsch, R., Krischer, L., Wassermann, J., & Behr, Y. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, *81*(3), 47–58. https://doi.org/10.1785/gssrl.81.3.530

Biswas, R., & Sen, M. (2017). 2D full-waveform inversion and uncertainty estimation using the reversible jump Hamiltonian Monte Carlo. In *SEG expanded abstracts* (pp. 1280–1285).

Blom, N., Gokhberg, A., & Fichtner, A. (2020). Seismic waveform tomography of the central and eastern Mediterranean upper mantle. *Solid Earth*, *11*(2), 669–690. https://doi.org/10.5194/se-11-669-2020

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., et al. (2016). Global adjoint tomography: First-generation model. *Geophysical Journal International*, *207*(3), 1739–1766. https://doi.org/10.1093/gji/ggw356

Burgos, G., Capdeville, Y., & Guillot, L. (2016). Homogenized moment tensor and the effect of near-field heterogeneities on nonisotropic radiation in nuclear explosion. *Journal of Geophysical Research: Solid Earth*, *121*(6), 4366–4389. https://doi.org/10.1002/2015JB012744

Byerly, P. (1928). The nature of the first motion in the Chilean earthquake of November 11, 1922. *American Journal of Science*, *16*(93), 232–236. https://doi.org/10.2475/ajs.s5-16.93.232

Cesca, S., Buforn, E., & Dahm, T. (2006). Amplitude spectra moment tensor inversion of shallow earthquakes in Spain. *Geophysical Journal International*, *166*(2), 839–854. https://doi.org/10.1111/j.1365-246X.2006.03073.x

Chave, A. D. (2017). *Computational statistics in the earth sciences: With applications in MATLAB* (pp. 281–343). Cambridge University Press.

Clinton, J. F., Hauksson, E., & Solanki, K. (2006). An evaluation of the SCSN moment tensor solutions: Robustness of the Mw magnitude scale, style of faulting, and automation of the method. *Bulletin of the Seismological Society of America*, *96*(5), 1689–1705. https://doi.org/10.1785/0120050241

Cua, G., & Heaton, T. (2007). The virtual seismologist (vs) method: A Bayesian approach to earthquake early warning. In P. Gasparini, G. Manfredi, & J. Zschau (Eds.), *Earthquake early warning systems* (pp. 97–132). Springer. https://doi.org/10.1007/978-3-540-72241-0_7

Dettmer, J., Benavente, R., Cummins, P. R., & Sambridge, M. (2014). Trans-dimensional finite-fault inversion. *Geophysical Journal International*, *199*(2), 735–751. https://doi.org/10.1093/gji/ggu280

Donner, S., Mustać, M., Hejrani, B., Tkalčić, H., & Igel, H. (2020). Seismic moment tensors from synthetic rotational and translational ground motion: Green's functions in 1-D versus 3-D. *Geophysical Journal International*, *223*(1), 161–179. https://doi.org/10.1093/gji/ggaa305

Dreger, D. S. (2003). 85.11—TDMT_INV: Time domain seismic moment tensor INVersion. In W. H. Lee, H. Kanamori, P. C. Jennings, & C. Kisslinger (Eds.), *International handbook of earthquake and engineering seismology, part B* (Vol. 81, p. 1627). Academic Press. https://doi.org/10.1016/S0074-6142(03)80290-5

Dreger, D. S., & Helmberger, D. (1991). Source parameters of the Sierra Madre Earthquake from regional and local body waves. *Geophysical Research Letters*, *18*(11), 2015–2018. https://doi.org/10.1029/91GL02366

Dreger, D. S., & Helmberger, D. V. (1993). Determination of source parameters at regional distances with three-component sparse network data. *Journal of Geophysical Research: Solid Earth*, *98*(B5), 8107–8125. https://doi.org/10.1029/93JB00023

Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, *195*(2), 216–222. https://doi.org/10.1016/0370-2693(87)91197-x

Dufumier, H., & Rivera, L. (1997). On the resolution of the isotropic component in moment tensor inversion. *Geophysical Journal International*, *131*(3), 595–606. https://doi.org/10.1111/j.1365-246x.1997.tb06601.x

Duputel, Z., Agram, P. S., Simons, M., Minson, S. E., & Beck, J. L. (2014). Accounting for prediction uncertainty when inferring subsurface fault slip. *Geophysical Journal International*, *197*(1), 464–482. https://doi.org/10.1093/gji/ggt517

Duputel, Z., Rivera, L., Fukahata, Y., & Kanamori, H. (2012). Uncertainty estimations for seismic source inversions. *Geophysical Journal International*, *190*(2), 1243–1256. https://doi.org/10.1111/j.1365-246X.2012.05554.x

Durek, J. J., & Ekström, G. (1996). A radial model of anelasticity consistent with long-period surface wave attenuation. *Bulletin of the Seismological Society of America*, *86*, 144–158.

Dziewoński, A. M., Chou, T.-A., & Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research*, *10*(B4), 2825–2852. https://doi.org/10.1029/jb086ib04p02825

Dziewoński, A. M., & Woodhouse, J. H. (1983). An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981. *Journal of Geophysical Research*, *88*(B4), 3247–3271. https://doi.org/10.1029/jb088ib04p03247

Eisner, L., & Clayton, R. W. (2001). A reciprocity method for multiple-source simulations. *Bulletin of the Seismological Society of America*, *91*(3), 553–560. https://doi.org/10.1785/0120000222

Ekström, G., Nettles, M., & Dziewonski, A. M. (2012). The global CMT project 2004–2010: Centroid moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, *200–201*, 1–9. https://doi.org/10.1016/j.pepi.2012.04.002

Engdahl, E. R., Sindorf, J. G., & Eppley, R. A. (1977). Interpretation of relative teleseismic P wave residuals. *Journal of Geophysical Research*, *82*(36), 5671–5682. https://doi.org/10.1029/JB082i036p05671

Eshelby, J. D., & Peierls, R. E. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *241*(1226), 376–396. https://doi.org/10.1098/rspa.1957.0133

Faccenna, C., Holt, A. F., Becker, T. W., Lallemand, S., & Royden, L. H. (2018). Dynamics of the Ryukyu/Izu-Bonin-Marianas double subduction system. *Tectonophysics*, *746*, 229–238. https://doi.org/10.1016/j.tecto.2017.08.011

Ferreira, A. M. G., Weston, J., & Funning, G. J. (2011). Global compilation of interferometric synthetic aperture radar earthquake source models: 2. Effects of 3-D Earth structure. *Journal of Geophysical Research: Solid Earth*, *116*(B8), B08409. https://doi.org/10.1029/2010JB008132

Ferreira, A. M. G., & Woodhouse, J. H. (2006). Long-period seismic source inversions using global tomographic models. *Geophysical Journal International*, *166*(3), 1178–1192. https://doi.org/10.1111/j.1365-246X.2006.03003.x

Ferreira, A. M. G., & Woodhouse, J. H. (2007). Source, path and receiver effects on seismic surface waves. *Geophysical Journal International*, *168*(1), 109–132. https://doi.org/10.1111/j.1365-246X.2006.03092.x

Fichtner, A. (2021). *Lecture notes on inverse theory*. Cambridge Open Engage. https://doi.org/10.33774/coe-2021-qpq2j

Fichtner, A., Bunge, H.-P., & Igel, H. (2006). The adjoint method in seismology—I. Theory. *Physics of the Earth and Planetary Interiors*, *157*(1–2), 86–104. https://doi.org/10.1016/j.pepi.2006.03.016

Fichtner, A., & Igel, H. (2008). Efficient numerical surface wave propagation through the optimization of discrete crustal models—A technique based on non-linear dispersion curve matching (DCM). *Geophysical Journal International*, *173*(2), 519–533. https://doi.org/10.1111/j.1365-246x.2008.03746.x

Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P. (2009a). Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophysical Journal International*, *179*(3), 1703–1725. https://doi.org/10.1111/j.1365-246x.2009.04368.x

Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P. (2009b). Spectral-element simulation and inversion of seismic waves in a spherical section of the Earth. *Journal of Numerical Analysis, Industrial and Applied Mathematics*, *4*, 11–22.

Fichtner, A., & Simutė, S. (2018). Hamiltonian Monte Carlo inversion of seismic sources in complex media. *Journal of Geophysical Research*, *123*(4), 2984–2999. https://doi.org/10.1002/2017JB015249

Fichtner, A., & Tkalcic, H. (2010). Insights into the kinematics of a volcanic caldera drop: Probabilistic finite-source inversion of the 1996 Bardarbunga, Iceland, earthquake. *Earth and Planetary Science Letters*, *297*(3–4), 607–615. https://doi.org/10.1016/j.epsl.2010.07.013

Fichtner, A., van Herwaarden, D., Afanasiev, M., Simutė, S., Krischer, L., Cubuk-Sabuncu, Y., et al. (2018). The collaborative seismic earth model: Generation I. *Geophysical Research Letters*, *45*(9), 4007–4016. https://doi.org/10.1029/2018gl077338

Fichtner, A., Zunino, A., & Gebraad, L. (2018). Hamiltonian Monte Carlo solution of tomographic inverse problems. *Geophysical Journal International*, *2016*(2), 1344–1363. https://doi.org/10.1093/gji/ggy496

Fichtner, A., Zunino, A., Gebraad, L., & Boehm, C. (2021). Autotuning Hamiltonian Monte Carlo for efficient generalised nullspace exploration. *Geophysical Journal International*, *227*(2), 941–968. https://doi.org/10.1093/gji/ggab270

Fitch, T. J., North, R. G., & Shields, M. W. (1981). Focal depths and moment tensor representations of shallow earthquakes associated with the Great Sumba Earthquake. *Journal of Geophysical Research: Solid Earth*, *86*(B10), 9357–9374. https://doi.org/10.1029/jb086ib10p09357

Ford, S. R., Dreger, D. S., & Walter, W. R. (2009a). Identifying isotropic events using a regional moment tensor inversion. *Journal of Geophysical Research*, *114*(B1), B01306. https://doi.org/10.1029/2008JB005743

Ford, S. R., Dreger, D. S., & Walter, W. R. (2009b). Source analysis of the memorial day explosion, Kimchaek, North Korea. *Geophysical Research Letters*, *36*(21), L21304. https://doi.org/10.1029/2009GL040003

Fouch, M. J., & Fischer, K. M. (1998). Shear wave anisotropy in the Mariana subduction zone. *Geophysical Research Letters*, *25*(8), 1221–1224. https://doi.org/10.1029/98GL00650

Frankel, A. (2013). Rupture history of the 2011 M 9 Tohoku Japan earthquake determined from strong-motion and high-rate GPS recordings: Subevents radiating energy in different frequency bands. *Bulletin of the Seismological Society of America*, *103*(2B), 1290–1306. https://doi.org/10.1785/0120120148

French, S. W., & Romanowicz, B. A. (2014). Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. *Geophysical Journal International*, *199*(3), 1303–1327. https://doi.org/10.1093/gji/ggu334

Fukuyama, E., Ishida, M., Horiuchi, S., Inoue, H., Hori, S., Sekiguchi, S., et al. (2001). *NIED seismic moment tensor catalogue January–December, 2000 (technical note no. 217)*. National Research Institute for Earth Science and Disaster Prevention.

Galitzin, B. (1909). Zur frage der bestimmung des azimuts der epizentrums eines beben. In *C. R. des Seances de l'Assoc. Int. Seism. Zermatt* (pp. 132–141).

Gebraad, L., Boehm, C., & Fichtner, A. (2020). Bayesian elastic full-waveform inversion using Hamiltonian Monte Carlo. *Journal of Geophysical Research: Solid Earth*, *125*(3), e2019JB018428. https://doi.org/10.1029/2019JB018428

Gokhberg, A., & Fichtner, A. (2016). Full-waveform inversion on heterogeneous HPC systems. *Computers & Geosciences*, *89*, 260–268. https://doi.org/10.1016/j.cageo.2015.12.013

Grandin, R., Vallée, M., Satriano, C., Lacassin, R., Klinger, Y., Simoes, M., & Bollinger, L. (2015). Rupture process of the $M_w$ = 7.9 2015 Gorkha earthquake (Nepal): Insights into Himalayan megathrust segmentation. *Geophysical Research Letters*, *42*(20), 8373–8382. https://doi.org/10.1002/2015GL066044

Graves, R. W., & Wald, D. J. (2001). Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions—1: Strong motions. *Journal of Geophysical Research*, *106*(B5), 8745–8766. https://doi.org/10.1029/2000jb900436

Greve, S. M., & Savage, M. K. (2009). Modelling seismic anisotropy variations across the Hikurangi subduction margin, New Zealand. *Earth and Planetary Science Letters*, *285*(1), 16–26. https://doi.org/10.1016/j.epsl.2009.05.035

Hallo, M., Asano, K., & Gallovič, F. (2017). Bayesian inference and interpretation of centroid moment tensors of the 2016 Kumamoto earthquake sequence, Kyushu, Japan. *Earth Planets and Space*, *69*(1), 134. https://doi.org/10.1186/s40623-017-0721-4

Hallo, M., & Gallovič, F. (2016). Fast and cheap approximation of Green function uncertainty for waveform-based earthquake source inversions. *Geophysical Journal International*, *207*(2), 1012–1029. https://doi.org/10.1093/gji/ggw320

Hanks, T. C. (1977). Earthquake stress drops, ambient tectonic stresses and stresses that drive plate motions. *Pure and Applied Geophysics*, *115*(1), 441–458. https://doi.org/10.1007/BF01637120

Hara, S., Fukahata, Y., & Iio, Y. (2019). P-wave first-motion polarity determination of waveform data in western Japan using deep learning. *Earth Planets and Space*, *71*(1), 127. https://doi.org/10.1186/s40623-019-1111-x

Hasegawa, A. (1990). Seismicity: Subduction zone. In *Geophysics* (pp. 1054–1061). Springer US.

Hasegawa, A. (2011). Seismicity, subduction zone. In H. K. Gupta (Ed.), *Encyclopedia of solid Earth geophysics* (pp. 1305–1315). Springer Netherlands. https://doi.org/10.1007/978-90-481-8702-7_14

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, *57*(1), 97–109. https://doi.org/10.1093/biomet/57.1.97

Hayes, G. P. (2018). Slab2—A comprehensive subduction zone geometry model: U.S. Geological Survey data release. https://doi.org/10.5066/F7PV6JNV

Hayes, G. P., Rivera, L., & Kanamori, H. (2009). Source inversion of the W-phase: Real-time implementation and extension to low magnitudes. *Seismological Research Letters*, *80*(5), 817–822. https://doi.org/10.1785/gssrl.80.5.817

Hayes, G. P., Wald, D. J., & Johnson, R. L. (2012). Slab1.0: A three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research: Solid Earth*, *117*(B1), B01302. https://doi.org/10.1029/2011JB008524

Hejrani, B., Tkalcic, H., & Fichtner, A. (2017). Centroid moment tensor catalogue using a 3D continental scale Earth model: Application to earthquakes in Papua New Giunea and the Solomon islands. *Journal of Geophysical Research*, *122*(7), 5517–5543. https://doi.org/10.1002/2017jb014230

Helmberger, D. V., & Enge, G. R. (1980). Modeling the long-period body waves from shallow earthquakes at regional ranges. *Bulletin of the Seismological Society of America*, *70*(5), 1699–1714.

Hingee, M., Tkalčić, H., Fichtner, A., & Sambridge, M. (2011). Seismic moment tensor inversion using a 3-D structural model: Applications for the Australian region. *Geophysical Journal International*, *184*(2), 949–964. https://doi.org/10.1111/j.1365-246x.2010.04897.x

Hjörleifsdóttir, V., & Ekström, G. (2010). Effects of three-dimensional Earth structure on CMT earthquake parameters. *Physics of the Earth and Planetary Interiors*, *179*(3), 178–190. https://doi.org/10.1016/j.pepi.2009.11.003

Hutton, K., Woessner, J., & Hauksson, E. (2000). Earthquake monitoring in southern California for seventy-seven years (1932–2008). *Bulletin of the Seismological Society of America*, *100*(2), 423–446. https://doi.org/10.1785/0120090130

Igel, H., & Gudmundsson, O. (1997). Frequency-dependent effects on travel times and waveforms of long-period S and SS waves. *Physics of the Earth and Planetary Interiors*, *104*(1–3), 229–246. https://doi.org/10.1016/s0031-9201(97)00051-4

Igel, H., Nissen-Meyer, T., & Jahnke, G. (2002). Wave propagation in 3D spherical sections: Effects of subduction zones. *Physics of the Earth and Planetary Interiors*, *132*(1), 219–234. https://doi.org/10.1016/S0031-9201(02)00053-5

Jia, Z., Wang, X., & Zhan, Z. (2020). Multifault models of the 2019 Ridgecrest sequence highlight complementary slip and fault junction instability. *Geophysical Research Letters*, *47*(17), e2020GL089802. https://doi.org/10.1029/2020GL089802

Jost, M. L., & Herrmann, R. B. (1989). A student's guide to and review of moment tensors. *Seismological Research Letters*, *60*(2), 37–57. https://doi.org/10.1785/gssrl.60.2.37

Kanamori, H., & Given, J. W. (1981). Use of long-period surface waves for rapid determination of earthquake-source parameters. *Physics of the Earth and Planetary Interiors*, *27*(1), 8–31. https://doi.org/10.1016/0031-9201(81)90083-2

Kanamori, H., & Rivera, L. (2008). Source inversion of W phase: Speeding up seismic tsunami warning. *Geophysical Journal International*, *175*(1), 222–238. https://doi.org/10.1111/j.1365-246X.2008.03887.x

Kasahara, K. (1963). Computer program for a fault-plane solution. *Bulletin of the Seismological Society of America*, *53*(1), 1–13. https://doi.org/10.1785/bssa0530010001

Kawakatsu, H. (1996). Observability of the isotropic component of a moment tensor. *Geophysical Journal International*, *126*(2), 525–544. https://doi.org/10.1111/j.1365-246X.1996.tb05308.x

Kawakatsu, H., Kumar, P., Takei, Y., Shinohara, M., Kanazawa, T., Araki, E., & Suyehiro, K. (2009). Seismic evidence for sharp lithosphere-asthenosphere boundaries of oceanic plates. *Science*, *324*(5926), 499–502. https://doi.org/10.1126/science.1169499

Kawasaki, I., & Tanimoto, T. (1981). Radiation patterns of body waves due to the seismic dislocation occurring in an anisotropic source medium. *Bulletin of the Seismological Society of America*, *71*(1), 37–50. https://doi.org/10.1785/bssa0710010037

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108–124. https://doi.org/10.1111/j.1365-246x.1995.tb03540.x

Khoshkholgh, S., Zunino, A., & Mosegaard, K. (2020). Informed proposal Monte Carlo. *arXiv e-prints*. arXiv:2005.14398.

Knopoff, L. (1961). Analytical calculation of the fault-plane problem. In J. H. Hodgson (Ed.), *Publications of the dominion observatory* (Vol. 24, pp. 309–315). Canada Department of Mines and Technical Surveys.

Koch, M. C., Fujisawa, K., & Murakami, A. (2020). Adjoint Hamiltonian Monte Carlo algorithm for the estimation of elastic modulus through the inversion of elastic wave propagation data. *International Journal for Numerical Methods in Engineering*, *121*(6), 1037–1067. https://doi.org/10.1002/nme.6256

Komatitsch, D., Erlebacher, G., Göddeke, D., & Michea, D. (2010). High-order finite-element seismic wave propagation modeling with MPI on a large GPU cluster. *Journal of Computational Physics*, *229*(20), 7692–7714. https://doi.org/10.1016/j.jcp.2010.06.024

Kong, X., Li, S., Wang, Y., Suo, Y., Dai, L., Géli, L., et al. (2018). Causes of earthquake spatial distribution beneath the Izu-Bonin-Mariana Arc. *Journal of Asian Earth Sciences*, *151*, 90–100. https://doi.org/10.1016/j.jseaes.2017.10.015

Krischer, L., Fichtner, A., Žukauskaitė, S., & Igel, H. (2015). Large-scale seismic inversion framework. *Seismological Research Letters*, *86*(4), 1198–1207. https://doi.org/10.1785/0220140248

Krischer, L., Igel, H., & Fichtner, A. (2018). Automated large-scale full seismic waveform inversion for North America and the North Atlantic. *Journal of Geophysical Research*, *123*(7), 5902–5928. https://doi.org/10.1029/2017JB015289

Kubo, A., Fukuyama, E., Kawai, H., & Nonomura, K. (2002). NIED seismic moment tensor catalogue for regional earthquakes around Japan: Quality test and application. *Tectonophysics*, *356*(1), 23–48. https://doi.org/10.1016/S0040-1951(02)00375-X

Kuge, K., & Kawakatsu, H. (1993). Significance of non-double couple components of deep and intermediate-depth earthquakes: Implications from moment tensor inversions of long-period seismic waves. *Physics of the Earth and Planetary Interiors*, *75*(4), 243–266. https://doi.org/10.1016/0031-9201(93)90004-S

Lee, E.-J., Chen, P., Jordan, T. H., & Wang, L. (2011). Rapid full-wave centroid moment tensor (CMT) inversion in a three-dimensional earth structure model for earthquakes in Southern California. *Geophysical Journal International*, *186*(1), 311–330. https://doi.org/10.1111/j.1365-246X.2011.05031.x

Lentas, K. (2017). Towards routine determination of focal mechanisms obtained from first motion P-wave arrivals. *Geophysical Journal International*, *212*(3), 1665–1686. https://doi.org/10.1093/gji/ggx503

Li, J., Zheng, Y., Thomsen, L., Lapen, T. J., & Fang, X. (2018). Deep earthquakes in subducting slabs hosted in highly anisotropic rock fabric. *Nature Geoscience*, *11*(9), 696–700. https://doi.org/10.1038/s41561-018-0188-3

Liu, Q., Polet, J., Komatitsch, D., & Tromp, J. (2004). Spectral-element moment tensor inversion for earthquakes in Southern California. *Bulletin of the Seismological Society of America*, *94*(5), 1748–1761. https://doi.org/10.1785/012004038

Maiti, S., & Tiwari, R. K. (2009). A hybrid Monte Carlo method based artificial neural networks approach for rock boundaries identification: A case study from the KTB bore hole. *Pure and Applied Geophysics*, *166*(12), 2059–2090. https://doi.org/10.1007/s00024-009-0533-y

Masfara, L. O. M., Cullison, T., & Weemstra, C. (2022). Efficient probabilistic inversion for induced earthquake parameters in 3D heterogeneous media. *Solid Earth Discussions*, *2022*, 1–25. https://doi.org/10.5194/se-2021-156

Mendiguren, J. A. (1977). Inversion of surface wave data in source mechanism studies. *Journal of Geophysical Research*, *82*(5), 889–894. https://doi.org/10.1029/JB082i005p00889

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, *21*(6), 1087–1092. https://doi.org/10.1063/1.1699114

Minson, S. E., Simons, M., Beck, J. L., Ortega, F., Jiang, J., Owen, S. E., et al. (2014). Bayesian inversion for finite fault earthquake source models—II: The 2011 great Tohoku-oki, Japan earthquake. *Geophysical Journal International*, *198*(2), 922–940. https://doi.org/10.1093/gji/ggu170

Morales-Yáñez, C., Duputel, Z., & Rivera, L. (2020). Impact of 3-D Earth structure on W-phase CMT parameters. *Geophysical Journal International*, *223*(2), 1432–1445. https://doi.org/10.1093/gji/ggaa377

Mosegaard, K., & Sambridge, M. (2002). Monte Carlo analysis of inverse problems. *Inverse Problems*, *18*(3), R29–R54. https://doi.org/10.1088/0266-5611/18/3/201

Mosegaard, K., & Tarantola, A. (2002). 16—Probabilistic approach to inverse problems. In W. H. Lee, H. Kanamori, P. C. Jennings, & C. Kisslinger (Eds.), *International handbook of earthquake and engineering seismology, part A* (Vol. 81, pp. 237–265). Academic Press. https://doi.org/10.1016/S0074-6142(02)80219-4

Muir, J. B., & Tkalčić, H. (2020). Probabilistic lowermost mantle P-wave tomography from hierarchical Hamiltonian Monte Carlo and model parametrisation cross-validation. *Geophysical Journal International*, *223*(3), 1643. https://doi.org/10.1093/gji/ggaa397

Mustać, M., & Tkalčić, H. (2016). Point source moment tensor inversion through a Bayesian hierarchical model. *Geophysical Journal International*, *204*(1), 311–323. https://doi.org/10.1093/gji/ggv458

National Research Institute for Earth Science and Disaster Resilience. (2021). Broadband Seismograph Network Laboratory, Network Center for Earthquake, Tsunami and Volcano. Retrieved from https://www.fnet.bosai.go.jp

Neal, R. M. (1993). *Probabilistic inference using Markov chain Monte Carlo methods (technical report no. CRG-TR-93-1)*. Department of Computer Science, University of Toronto.

Neal, R. M. (1996). *Bayesian learning for neural networks*. Springer.

Neal, R. M. (2011). MCMC using Hamiltonian dynamics. In S. Brooks, A. Gelman, G. Jones, & X.-L. Meng (Eds.), *Handbook of Markov Chain Monte Carlo (chapter 5)*. Chapman & Hall/CRC.

Nettles, M., & Ekström, G. (1998). Faulting mechanism of anomalous earthquakes near Bárdarbunga Volcano, Iceland. *Journal of Geophysical Research: Solid Earth*, *103*(B8), 17973–17983. https://doi.org/10.1029/98jb01392

Newrkla, K., Shiddiqi, H. A., Jerkins, A. E., Keers, H., & Ottemöller, L. (2019). Implications of 3D seismic raytracing on focal mechanism determination. *Bulletin of the Seismological Society of America*, *109*(6), 2746–2754. https://doi.org/10.1785/0120190184

Nissen-Meyer, T., Fournier, A., & Dahlen, F. A. (2007). A two-dimensional spectral-element method for computing spherical-Earth seismograms—I. Moment-Tensor source. *Geophysical Journal International*, *168*(3), 1067–1092. https://doi.org/10.1111/j.1365-246x.2006.03121.x

Okamoto, T., Takenaka, H., & Nakamura, T. (2018). Evaluation of accuracy of synthetic waveforms for subduction-zone earthquakes by using a land–ocean unified 3D structure model. *Earth Planets and Space*, *70*(1), 98. https://doi.org/10.1186/s40623-018-0871-z

Omori, F. (1905). Similarity of seismic motions originating at neighboring centers. In *Publ. Earthquake Invest. Com. in foreign languages (Tokyo)* (Vol. *9*, p. 52).

Panning, M., & Romanowicz, B. (2006). A three-dimensional radially anisotropic model of shear velocity in the whole mantle. *Geophysical Journal International*, *167*(1), 361–379. https://doi.org/10.1111/j.1365-246x.2006.03100.x

Patton, H. J., & Randall, G. E. (2002). On the causes of biased estimates of seismic moment for earthquakes in central Asia. *Journal of Geophysical Research: Solid Earth*, *107*(B11), 2302. https://doi.org/10.1029/2001JB000351

Pondrelli, S., Morelli, A., Ekström, G., Mazza, S., Boschi, E., & Dziewonski, A. (2002). European–Mediterranean regional centroid-moment tensors: 1997–2000. *Physics of the Earth and Planetary Interiors*, *130*(1), 71–101. https://doi.org/10.1016/S0031-9201(01)00312-0

Pugh, D. J., White, R. S., & Christie, P. A. F. (2016). A Bayesian method for microseismic source inversion. *Geophysical Journal International*, *206*(2), 1009–1038. https://doi.org/10.1093/gji/ggw186

Romanowicz, B. A. (1982). Moment tensor inversion of long period Rayleigh waves: A new approach. *Journal of Geophysical Research: Solid Earth*, *87*(B7), 5395–5407. https://doi.org/10.1029/JB087iB07p05395

Sambridge, M. S., & Gallagher, K. (2011). Inverse theory, Monte Carlo method. In H. K. Gupta (Ed.), *Encyclopedia of solid earth geophysics* (pp. 639–644). Springer Netherlands. https://doi.org/10.1007/978-90-481-8702-7_192

Sambridge, M. S., & Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, *40*(3), 3-1–3-29. https://doi.org/10.1029/2000RG000089

Scognamiglio, L., Magnoni, F., Tinti, E., & Casarotti, E. (2016). Uncertainty estimations for moment tensor inversions: The issue of the 2012 May 20 Emilia earthquake. *Geophysical Journal International*, *206*(2), 792–806. https://doi.org/10.1093/gji/ggw173

Sen, M. K., & Biswas, R. (2017). Tansdimensional seismic inversion using the reversible jump Hamiltonian Monte Carlo algorithm. *Geophysics*, *82*(3), R119–R134. https://doi.org/10.1190/geo2016-0010.1

Seno, T., & Eguchi, T. (1983). Seismotectonics of the western Pacific region. In T. W. Hilde & S. Uyeda (Eds.), *Geodynamics of the western Pacific-Indonesian region* (pp. 5–40). American Geophysical Union (AGU). https://doi.org/10.1029/GD011p0005

Shang, X., & Tkalčić, H. (2020). Point-source inversion of small and moderate earthquakes from P-wave polarities and P/S amplitude ratios within a hierarchical Bayesian framework: Implications for the Geysers earthquakes. *Journal of Geophysical Research: Solid Earth*, *125*(2), e2019JB018492. https://doi.org/10.1029/2019JB018492

Shuler, A., Ekström, G., & Nettles, M. (2013). Physical mechanisms for vertical-CLVD earthquakes at active volcanoes. *Journal of Geophysical Research: Solid Earth*, *118*(4), 1569–1586. https://doi.org/10.1002/jgrb.50131

Siebert, L., Simkin, T., & Kimberley, P. (2010). *Volcanoes of the world* (3rd ed.). University of California Press.

Simutė, S. (2022). Waveform data for centroid moment tensor solutions presented in publication "Bayesian seismic source inversion with a 3-D Earth model of the Japanese islands". *Zenodo*. https://doi.org/10.5281/zenodo.7384213

Simutė, S., Steptoe, H., Gokhberg, A., & Fichtner, A. (2016). Full-waveform inversion of the Japanese islands region. *Journal of Geophysical Research*, *121*(5), 3722–3741. https://doi.org/10.1002/2016jb012802

Smith, G. P., & Ekström, G. (1996). Improving teleseismic event locations using a three-dimensional earth model. *Bulletin of the Seismological Society of America*, *86*(3), 788–796. https://doi.org/10.1785/bssa0860030788

Staehler, S. C., & Sigloch, K. (2014). Fully probabilistic seismic source inversion—Part I: Efficient parameterisation. *Solid Earth*, *5*(2), 1055–1069. https://doi.org/10.5194/se-5-1055-2014

Staehler, S. C., & Sigloch, K. (2017). Fully probabilistic seismic source inversion—Part II: Modelling errors and station covariances. *Solid Earth*, *7*(6), 1521–1536. https://doi.org/10.5194/se-7-1521-2016

Swiss National Supercomputing Center. (2021). CSCS—Swiss National Supercomputing Center. Retrieved from https://cscs.ch

Takemura, S., Kimura, T., Saito, T., Kubo, H., & Shiomi, K. (2018). Moment tensor inversion of the 2016 southeast offshore Mie earthquake in the Tonankai region using a three-dimensional velocity structure model: Effects of the accretionary prism and subducting oceanic plate. *Earth Planets and Space*, *70*(1), 50. https://doi.org/10.1186/s40623-018-0819-3

Takemura, S., Okuwaki, R., Kubota, T., Shiomi, K., Kimura, T., & Noda, A. (2020). Centroid moment tensor inversions of offshore earthquakes using a three-dimensional velocity structure model: Slip distributions on the plate boundary along the Nankai trough. *Geophysical Journal International*, *222*(2), 1109–1125. https://doi.org/10.1093/gji/ggaa238

Tape, W., & Tape, C. (2013). The classical model for moment tensors. *Geophysical Journal International*, *195*(3), 1701–1720. https://doi.org/10.1093/gji/ggt302

Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure and Applied Geophysics*, *128*(1–2), 365–399. https://doi.org/10.1007/bf01772605

The Global CMT Project. (2021). Moment tensor product query. Retrieved from https://ds.iris.edu/spud/momenttensor

Thrastarson, S., van Driel, M., Krischer, L., Boehm, C., Afanasiev, M., van Herwaarden, D.-P., & Fichtner, A. (2020). Accelerating numerical wave propagation by wavefield adapted meshes. Part II: Full-waveform inversion. *Geophysical Journal International*, *221*(3), 1591–1604. https://doi.org/10.1093/gji/ggaa065

Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake Area, central California. *Journal of Geophysical Research: Solid Earth*, *88*(B10), 8226–8236. https://doi.org/10.1029/JB088iB10p08226

Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, *160*(1), 195–216. https://doi.org/10.1111/j.1365-246x.2004.02453.x

Vackár, J., Burjének, J., Gallovič, F., Zahradníkk, J., & Clinton, J. (2017). Bayesian ISOLA: New tool for automated centroid moment tensor inversion. *Geophysical Journal International*, *210*, 693–705. https://doi.org/10.1093/gji/ggx158

Vallée, M., Charléty, J., Ferreira, A. M. G., Delouis, B., & Vergoz, J. (2011). SCARDEC: A new technique for the rapid determination of seismic moment magnitude, focal mechanism and source time functions for large earthquakes using body-wave deconvolution. *Geophysical Journal International*, *184*(1), 338–358. https://doi.org/10.1111/j.1365-246X.2010.04836.x

Vallée, M., & Douet, V. (2016). A new database of source time functions (STFs) extracted from the SCARDEC method. *Physics of the Earth and Planetary Interiors*, *257*, 149–157. https://doi.org/10.1016/j.pepi.2016.05.012

van Herwaarden, D. P., Boehm, C., Afanasiev, M., Thrastarson, S., Krischer, L., Trampert, J., & Fichtner, A. (2020). Accelerated full-waveform inversion using dynamic mini-batches. *Geophysical Journal International*, *221*(2), 1427–1438. https://doi.org/10.1093/gji/ggaa079

Vavryčuk, V. (2004). Inversion for anisotropy from non-double-couple components of moment tensors. *Journal of Geophysical Research: Solid Earth*, *109*(B7), B07306. https://doi.org/10.1029/2003JB002926

Vavryčuk, V. (2015). Moment tensor decompositions revisited. *Journal of Seismology*, *19*(1), 231–252. https://doi.org/10.1007/s10950-014-9463-y

Wallace, T., Given, J., & Kanamori, H. (1982). A discrepancy between long- and short-period mechanisms of earthquakes near the long valley caldera. *Geophysical Research Letters*, *9*(10), 1131–1134. https://doi.org/10.1029/GL009i010p01131

Wang, N., Montagner, J.-P., Fichtner, A., & Capdeville, Y. (2013). Intrinsic versus extrinsic seismic anisotropy: The radial anisotropy in reference Earth models. *Geophysical Research Letters*, *40*(16), 4284–4288. https://doi.org/10.1002/grl.50873

Wang, X., & Zhan, Z. (2019). Moving from 1-D to 3-D velocity model: Automated waveform-based earthquake moment tensor inversion in the Los Angeles region. *Geophysical Journal International*, *220*(1), 218–234. https://doi.org/10.1093/gji/ggz435

Wang, X., & Zhan, Z. (2020). Seismotectonics and fault geometries of the 2019 Ridgecrest sequence: Insight from aftershock moment tensor Catalog using 3-D Green's functions. *Journal of Geophysical Research: Solid Earth*, *125*(5), e2020JB019577. https://doi.org/10.1029/2020JB019577

Wirth, E., & Long, M. D. (2010). Frequency-dependent shear wave splitting beneath the Japan and Izu-Bonin subduction zones. *Physics of the Earth and Planetary Interiors*, *181*(3), 141–154. https://doi.org/10.1016/j.pepi.2010.05.006

Woodhouse, J. H. (1983). The joint inversion of seismic wave forms for lateral variations in Earth structure and earthquake source parameter. In *Physics of the Earth's interior* (Vol. 85, pp. 366–397). North-Holland.

Woodhouse, J. H., & Dziewoński, A. M. (1984). Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms. *Journal of Geophysical Research*, *89*(B7), 5953–5986. https://doi.org/10.1029/jb089ib07p05953

Wu, W., Ni, S., Zhan, Z., & Wei, S. (2018). An SEM-DSM three-dimensional hybrid method for modelling teleseismic waves with complicated source-side structures. *Geophysical Journal International*, *215*(1), 133–154. https://doi.org/10.1093/gji/ggy273

Xu, D., & Li, D. (2008). Molecular dynamics simulation method. In D. Li (Ed.), *Encyclopedia of microfluidics and nanofluidics* (pp. 1391–1398). Springer US. https://doi.org/10.1007/978-0-387-48998-8_1052

Yu, C., Vavryčuk, V., Adamová, P., & Bohnhoff, M. (2019). Frequency-dependent moment tensors of induced microearthquakes. *Geophysical Research Letters*, *46*(12), 6406–6414. https://doi.org/10.1029/2019GL082634

Zahradník, J., Fojtíková, L., Carvalho, J., Barros, L., Sokos, E., & Janský, J. (2015). Compromising polarity and waveform constraints in focal-mechanism solutions; the Mara Rosa 2010 $M_w$ 4 central Brazil earthquake revisited. *Journal of South American Earth Sciences*, *63*, 323–333. https://doi.org/10.1016/j.jsames.2015.08.011

Zanella, G. (2020). Informed proposals for local MCMC in discrete spaces. *Journal of the American Statistical Association*, *115*(530), 852–865. https://doi.org/10.1080/01621459.2019.1585255

Zhao, L., Chen, P., & Jordan, T. H. (2006). Strain Green's tensors, reciprocity, and their applications to seismic source and structure studies. *Bulletin of the Seismological Society of America*, *96*(5), 1753–1763. https://doi.org/10.1785/0120050253

Zhao, L., & Helmberger, D. V. (1994). Source estimation from broadband regional seismograms. *Bulletin of the Seismological Society of America*, *84*(1), 91–104.

Zhu, L., & Helmberger, D. V. (1996). Advancement in source estimation techniques using broadband regional seismograms. *Bulletin of the Seismological Society of America*, *86*(5), 1634–1641. https://doi.org/10.1785/bssa0860051634

Zhu, L., & Zhou, X. (2016). Seismic moment tensor inversion using 3D velocity model and its application to the 2013 Lushan earthquake sequence. *Physics and Chemistry of the Earth*, *95*, 10–18. https://doi.org/10.1016/j.pce.2016.01.002

## References From the Supporting Information

Dahlen, F. A., & Tromp, J. (1998). *Theoretical global seismology*. Princeton University Press.

Dziewoński, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, *25*(4), 297–356. https://doi.org/10.1016/0031-9201(81)90046-7

Karato, S.-I. (1992). On the Lehmann discontinuity. *Geophysical Research Letters*, *19*(22), 2255–2258. https://doi.org/10.1029/92GL02603

Revenaugh, J., & Jordan, T. H. (1991). Mantle layering from ScS reverberations: 3. The upper mantle. *Journal of Geophysical Research: Solid Earth*, *96*(B12), 19781–19810. https://doi.org/10.1029/91JB01487

Shito, A., Karato, S., Nishihara, Y., & Matsukage, K. (2006). Towards mapping the three-dimensional distribution of water in the upper mantle from velocity and attenuation tomography. In *Earth's deep water cycle*, (pp. 225–236). American Geophysical Union (AGU). https://doi.org/10.1029/168GM17