RESEARCH ARTICLE

Key Points:
- We perform a probabilistic Bayesian moment tensor inversion (including uncertainties) of largest earthquakes in The Geysers geothermal field
- Empirically estimated noise covariance matrix accounts for interdependences of data errors and notably different focal mechanisms
- One of the events is predominantly double-couple, while others involve crack opening and imply a low Poisson ratio or fluid extraction

Correspondence to:
M. Mustać,
marija.mustac@anu.edu.au

The Variability and Interpretation of Earthquake Source Mechanisms in The Geysers Geothermal Field From a Bayesian Standpoint Based on the Choice of a Noise Model

Marija Mustać[1] [ID], **Hrvoje Tkalčić[1]**, and **Alexander L. Burky[2]**

[1]Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia, [2]Department of Geosciences, Princeton University, Princeton, NJ, USA

Abstract Moment tensor (MT) inversion studies of events in The Geysers geothermal field mostly focused on microseismicity and found a large number of earthquakes with significant non-double-couple (non-DC) seismic radiation. Here we concentrate on the largest events in the area in recent years using a hierarchical Bayesian MT inversion. Initially, we show that the non-DC components of the MT can be reliably retrieved using regional waveform data from a small number of stations. Subsequently, we present results for a number of events and show that accounting for noise correlations can lead to retrieval of a lower isotropic (ISO) component and significantly different focal mechanisms. We compute the Bayesian evidence to compare solutions obtained with different assumptions of the noise covariance matrix. Although a diagonal covariance matrix produces a better waveform fit, inversions that account for noise correlations via an empirically estimated noise covariance matrix account for interdependences of data errors and are preferred from a Bayesian point of view. This implies that improper treatment of data noise in waveform inversions can result in fitting the noise and misinterpreting the non-DC components. Finally, one of the analyzed events is characterized as predominantly DC, while the others still have significant non-DC components, probably as a result of crack opening, which is a reasonable hypothesis for The Geysers geothermal field geological setting.

Plain Language Summary Earthquakes are a result of rocks breaking under stress, but their mechanisms can be more complicated in volcanic and geothermal areas. A good example is The Geysers geothermal field, an area of highly fractured rocks where electricity production activities were related to occurrence of small earthquakes. To investigate the complexities of several largest earthquakes in The Geysers, we analyzed the ground motion recorded on stations throughout California. We used a probabilistic method to compute the source mechanisms, which gives the uncertainty of the results and reduces the effect of noise on the calculations. Our results show a variety of earthquake mechanisms, from a simple earthquake, through an event that includes crack opening within rocks, to even more complicated events that might involve fluids moving away from the source area.

1. Introduction

The Geysers geothermal field is the world's largest producer of geothermal electricity. It is located in northern California, in a broad zone of transform faulting. The very existence of this vapor-dominated field is a result of specific geological conditions involving permeable rocks that are suspected to be underlain by a magma body (e.g., Brook, 1981). Exploitation of the field since the early 1960s and subsequent water injection to maintain steam pressure have been followed by a significant increase in seismic activity. Microearthquakes in the area were attributed to water injection and steam extraction (e.g., Majer et al., 2007). Previous studies of source mechanisms have shown that about half of the microearthquakes have significant non-DC components (e.g., Martínez-Garzón et al., 2017; Ross et al., 1999). The non-DC components are thought to be the most unstable part of source inversions, as they are sensitive to station coverage, inadequacies in the Green's functions due to hypocenter mislocation and imperfect structural models, and noise in the seismic waveforms (e.g., Kuge & Lay, 1994; Křížová et al., 2013; Saraò et al., 2001). Hence, it is important to perform a thorough study to check the reliability of these components. A hierarchical Bayesian moment tensor (MT) inversion

technique proposed in Mustać and Tkalčić (2016, 2017) provides parameter uncertainties that show solution robustness, which makes the method suitable for this task. Non-DC components of earthquakes in The Geysers can be a result of electricity production, complex local geology, or they can be spurious—a consequence of inadequate modeling, technique deficiency, or poor azimuthal coverage. To better understand them, we shall first discuss the geological setting in more detail.

The Geysers field is situated in the central belt of the Franciscan assemblage. The rugged and mountainous terrain of the reservoir has numerous variously orientated fractures and contains intrusions of silicic batholith (Thompson, 1989). Recently, Boyle and Zoback (2014) used focal mechanisms and found a bimodal distribution of fracture orientations within the reservoir (between N40°E and N80°E), and orientations of N30°E–N40°E below it. Shear-wave splitting showed fracture orientations between north and N60°E (Elkibbi & Rial, 2005). Introduction of magma into the crust in late Tertiary and Quaternary can be related to east-southeast extension that accompanied northward propagation of the San Andreas transform system in this region (McLaughlin, 1981). Geophysical features that delineate the buried magmatic body are an extensive negative Bouguer gravity anomaly up to ~25 mGal (Isherwood, 1981) and a delay in teleseismic P wave arrival times by up to 1 s (Iyer et al., 1979). They indicate the presence of a body of partial melt at midcrustal depths that acts as a heat source for the reservoir (Blakely & Stanley, 1993). This complex structure involving stacks of permeable and impermeable highly fractured rocks results in specific vapor-dominated conditions (McLaughlin, 1981). A number of geodetic studies found that the reservoir is subsiding (e.g., Lofgren, 1981; Mossop & Segall, 1997, 1999) and contracting in horizontal directions (Denlinger & Bufe, 1982).

The Geysers field is surrounded by the inactive Collayomi fault on the northeast and active Maacama fault on the west (McLaughlin, 1981). The maximum horizontal stress within the field has an orientation of N26°E and is approximately equal to the vertical stress, resulting in a mixture of strike-slip and normal faults (Boyle & Zoback, 2014). Earthquakes in The Geysers mostly occur at depths between 2 and 5 km. Their locations do not seem to be associated with any of the fault systems, rather with the electricity production activities (Bufe et al., 1981; O'Connell & Johnson, 1988; Romero et al., 1994). Moreover, significant non-DC components were found in a large number of source studies. Kirkpatrick et al. (1996) used P and S wave amplitudes and polarities to obtain MT solutions for 2,552 events. Almost half of the earthquakes required significant ISO and compensated linear vector dipole (CLVD) components. The ISO parts were up to 30% of the mechanism and had both positive and negative signs. Similar results were obtained by Ross et al. (1996, 1999). Martínez-Garzón et al. (2017) confirmed previous results and observed that larger magnitude events ($M_W > 2$) had mostly strike slip and reverse mechanisms, while normal faulting events displayed a lower magnitude and had larger volumetric components. Most recent studies using waveform data have found even larger ISO components (up to 47%) for certain earthquakes (Boyd et al., 2015; Johnson, 2014b).

As the seismicity in The Geysers field is dominated by microearthquakes, most source studies use local networks to focus on microearthquake activity and require very detailed structural models of The Geysers field. The complexities in geological conditions are reflected in strong lateral variations in seismic velocities, the v_P/v_S ratio and attenuation. O'Connell and Johnson (1991) obtained vertically heterogeneous one-dimensional P and S wave velocity models. They found a low v_P/v_S ratio at depths of maximum steam production and high v_P/v_S above it. The high v_P/v_S ratio was interpreted as a fluid-saturated condensation zone, as Castagna et al. (1985) found that v_P/v_S ratio is sensitive to fluid saturation and peaks at 100% saturation because v_P is increased and v_S is reduced. Three-dimensional models of compressional wave velocities show correlations with surface geology (e.g., Julian et al., 1996; Romero et al., 1994; Zucca et al., 1994). A zone of low v_P/v_S ratio appears to characterize the geothermal reservoir as a consequence of the difference between water vapor in the reservoir and liquid water in the surrounding rock (e.g., Julian et al., 1996). Apart from that, P wave and S wave attenuation have high values in the reservoir (Romero et al., 1997). Zucca et al. (1994) found that P wave attenuation increases with depth above the reservoir (Q_P decreases) and decreases within the reservoir. Attenuation increase is related to increase in lithostatic pressure that closes cracks, and a sharp change in the behavior of attenuation with depth is interpreted by a sharp drop in water saturation at the top of the reservoir. Furthermore, production activities might be changing the material properties with time (Foulger et al., 1997) and the changing heterogeneous velocity field needs to be taken into account in local MT studies.

Larger earthquakes have recently been studied by Boyd et al. (2015), who performed a regional full-waveform MT inversion of 53 events with moment magnitudes larger than 3. They found that, in comparison with the
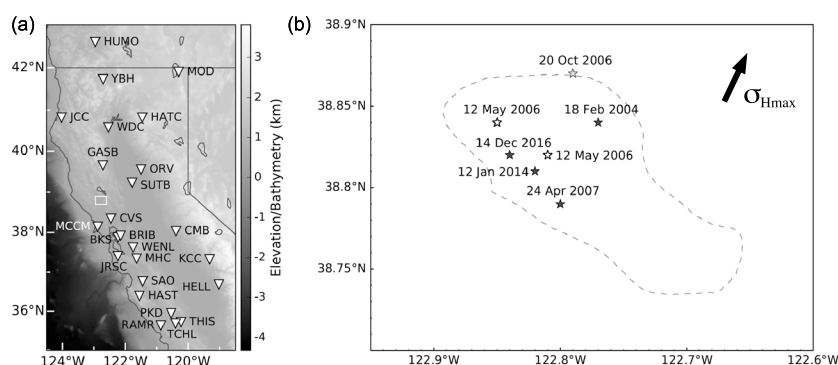
Figure 1. (a) Relief map of California showing stations used in the inversions (inverted triangles). Note that a different combination of stations was used for different events. (b) Map of The Geysers geothermal field (dashed line) with $M_W \geq 4.5$ earthquakes analyzed in this study (black stars), in a previous study (gray star), or not analyzed (white stars), as explained in the text. This area is denoted in Figure 1a by a white rectangle. The arrow shows the direction of the maximum horizontal stress according to Boyle and Zoback (2014).

rest of California, The Geysers earthquakes have significantly higher ISO components with a mean of ~33%. Guilhem et al. (2014) used local waveforms and retrieved focal mechanisms similar to the Berkeley Seismological Laboratory (BSL) regional solutions for a number of earthquakes. The same study found that the first-motion data indicate larger non-DC components. We presented results for the 20 October 2006 $M_W = 4.6$ event (gray star in Figure 1b) in a previous paper and found a predominantly non-DC mechanism indicating crack opening and possible fluid extraction (Mustać & Tkalčić, 2017). This solution was obtained after we took into account noise correlations, and it had a considerably larger CLVD component than the solution obtained without data noise considerations.

In this paper, we examine the non-DC components of largest magnitude events in The Geysers field using a hierarchical Bayesian MT inversion technique and data from selected seismic stations at regional distances. The technique provides an ensemble of solutions that characterize the uncertainties. Furthermore, we include an empirically estimated noise covariance matrix to account for correlations present in the waveform data and compare it to results that neglect the correlations by means of the Bayesian evidence. To reduce the effect of time-dependent complex local structure, we use stations at regional distances and low period data.

We first estimate the influence of station coverage and inversion assumptions on the existence of the ISO component of the MT in synthetic experiments. Subsequently, we study earthquakes with $M_W \geq 4.5$. One of the largest events ($M_W = 4.71$) occurred on 12 May 2006 at 10:37:29 UTC but was closely followed by a nearby $M_L = 4.15$ event at 10:38:36 UTC (white stars in Figure 1b). Surface waves of these two events, which dominate the frequencies used in this study, overlap, and we omit them from our analysis. We examine the remaining four $M_W \geq 4.5$ earthquakes that occurred on 18 February 2004, 24 April 2007, 11 January 2014, and 14 December 2016 (black stars in Figure 1 and bold text in Table 1). We also compute the Bayesian evidence for the purpose of model comparison, that is, to evaluate solutions obtained for different noise covariance matrices and different numbers of noise parameters.

Table 1
Locations and Magnitudes of M ≥ 4.5 Earthquakes in The Geysers Region, as Reported by BSL

Earthquake	Longitude (deg)	Latitude (deg)	Depth (km)	M_W
18 Feb 2004	**−122.77**	**38.84**	**3.39**	**4.58**
12 May 2006	−122.81	38.82	4.53	4.71
12 May 2006	−122.85	38.84	2.04	4.15 (M_L)
20 Oct 2006	−122.79	38.87	3.46	4.56
24 Apr 2007	**−122.80**	**38.79**	**2.48**	**4.46**
12 Jan 2014	**−122.82**	**38.81**	**2.61**	**4.53**
14 Dec 2016	**−122.84**	**38.82**	**1.48**	**5.01**

2. Method

We utilize the probabilistic Bayesian inversion method for the seismic MT, as presented in Mustać and Tkalčić (2016, 2017). The inversion gives a posterior probability distribution of model parameters **m** given the data **d**, denoted by $p(\mathbf{m}|\mathbf{d})$. The posterior probability distribution depends on the prior probability distribution of the model parameters $p(\mathbf{m})$ and the likelihood function $p(\mathbf{d}|\mathbf{m})$, as defined by the Bayes' theorem (Bayes & Price, 1763)

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}. \tag{1}$$

The denominator $p(\mathbf{d})$, also known as the Bayesian evidence, normalizes the posterior probability distribution in a way that its integral over the model space equal to unity. We examine the maximum a posteriori probability (MAP) solution for all events.

The prior probability distribution incorporates our previous knowledge about the model gathered from different observations. The prior enables one to reject implausible solutions from a family of solutions that fit the data. The model parameters **m** consist of six coefficients a_n corresponding to six elementary MTs (Cotton & Coutant, 1997; Kikuchi & Kanamori, 1991; Mustać & Tkalčić, 2016), and the full MT is a linear combination of elementary tensors. We utilize a uniform prior probability distribution for all parameters and define it over a large set of values. The prior for coefficients a_n is based on a previous value for the scalar moment M_{00} (e.g., from a study using wave polarities or GPS data). They can have values between $\pm 170\% \ M_{00}$, with an additional constraint that the scalar moment M_0 needs to be between 30 % M_{00} and 170% M_{00}. The noise is defined as a percent of data root-mean-square (RMS) and can reach values up to 5 times larger than RMS.

The likelihood function introduces information obtained from the data to the posterior probability distribution. It is defined using the misfit between the modeled seismograms $\mathbf{G}(\mathbf{m})$ and the data **d**

$$p(\mathbf{d}|\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C_D}|}} \ \exp\left[-\frac{1}{2}(\mathbf{G}(\mathbf{m}) - \mathbf{d})^T \mathbf{C_D}^{-1}(\mathbf{G}(\mathbf{m}) - \mathbf{d})\right]. \tag{2}$$

In the above equation N is the number of data points, $\mathbf{C_D}$ is the data covariance matrix, and $|\mathbf{C_D}|$ its determinant. Although the $\mathbf{C_D}$ matrix formally includes the theory error that stems mainly from uncertainties in the centroid location and the velocity model used to create the Green's functions, we do not include this part when estimating the $\mathbf{C_D}$. We minimize the effect of these theory errors through the use of long-period waveforms (20–50 s) that are less sensitive to small-scale heterogeneities.

We construct $\mathbf{C_D}$ as block diagonal, with each block corresponding to one seismogram component, and perform inversions using a diagonal and a two-attenuated-cosine $\mathbf{C_D}$. The diagonal matrix accounts for noise amplitudes but not its correlations. In our previous applications (Mustać and Tkalčić, 2016, 2017), the $\mathbf{C_D}$ matrix was empirically estimated using ambient noise autocorrelation traces. It is constructed as a Toeplitz matrix of the two-attenuated-cosine approximation of the average autocorrelation (equation (4) in Mustać & Tkalčić, 2017). Using the autocorrelation to estimate the $\mathbf{C_D}$ matrix is similar to the approach of Gouveia and Scales (1998), but with a different normalization factor that penalizes correlations for longer lags.

To explore the effect of local seismicity and large teleseismic events on noise, here we examine autocorrelations calculated for 24 h long seismograms recorded on stations used in this study. We choose a day without local earthquakes of magnitudes above 2, but containing a M_W 6.5 teleseismic earthquake in the Northern California earthquake data center (NCEDC) catalogue. Although it was a seismically quiet day, 43 earthquakes occurred in a 500 km radius, with a maximum magnitude of 1.64. The seismograms are processed in the same way as we process the data: the instrument responses are removed and they are band-pass filtered between 20 and 50 s. We compute autocorrelations for each hour, with a maximum lag of 220 s. Autocorrelations of vertical components are consistent among stations, with a period of ~20 s (Figure 2). They can be related to local Earth structure beneath the stations (e.g., Phạm & Tkalčić, 2017; Sens-Schönfelder & Wegler, 2006; Sun & Kennett, 2016). The autocorrelations of horizontal components vary considerably throughout the day. On most stations, they have a period of ~40 s, except in hours containing the teleseismic event (12–14) and when a local event is prominent on the seismograms (e.g., for hours 1, 8, and 9), when the period is around 20 s. This suggests that signal-generated noise has shorter period autocorrelation, and we use the same, shorter period, two-attenuated-cosine $\mathbf{C_D}$ for all components.
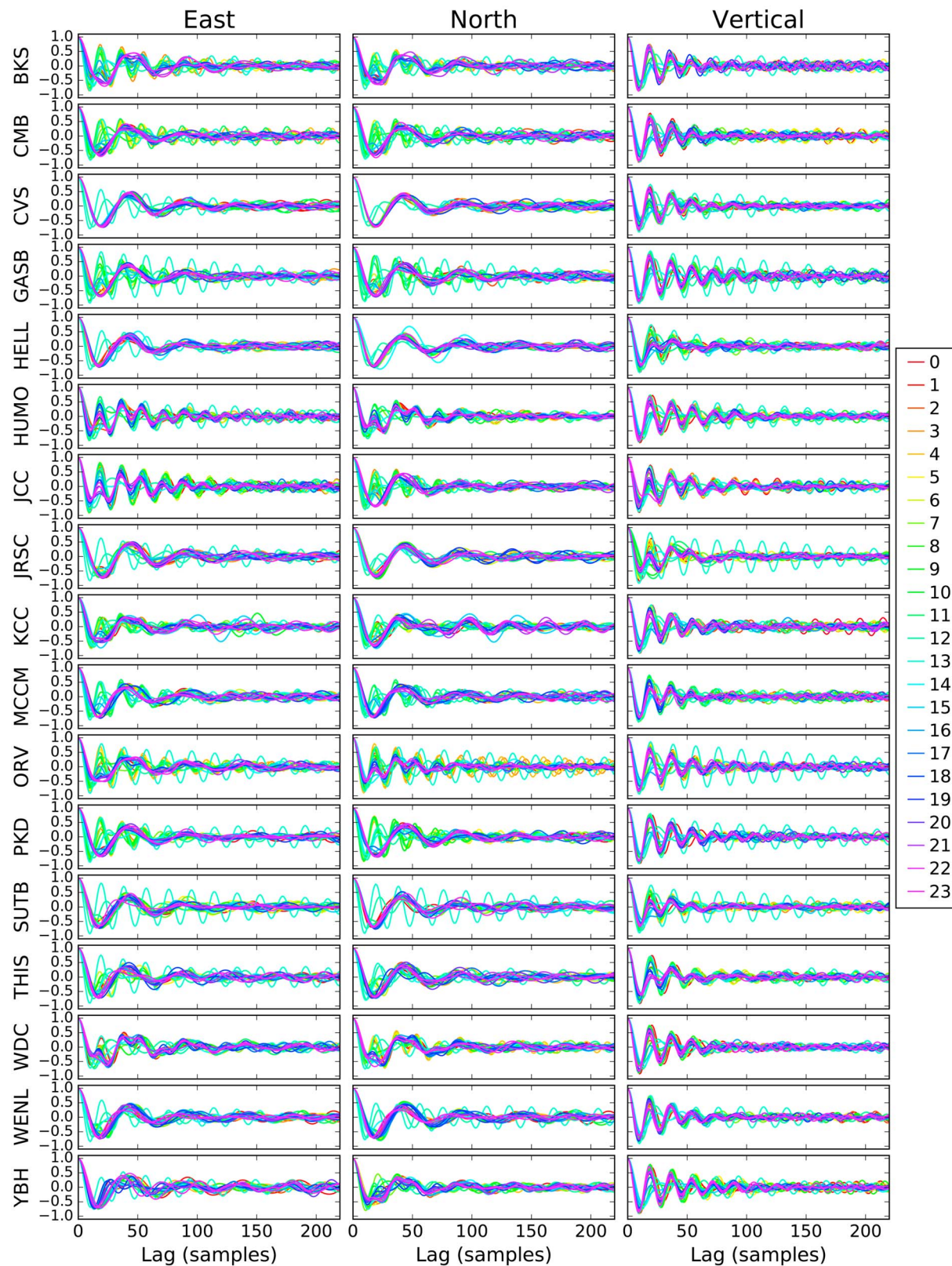
Figure 2. Autocorrelations of the three-component seismograms on 23 June 2015 from stations used in the inversions. They are calculated for each hour of the day, with a maximum lag of 220 s. The colorbar shows the hour.
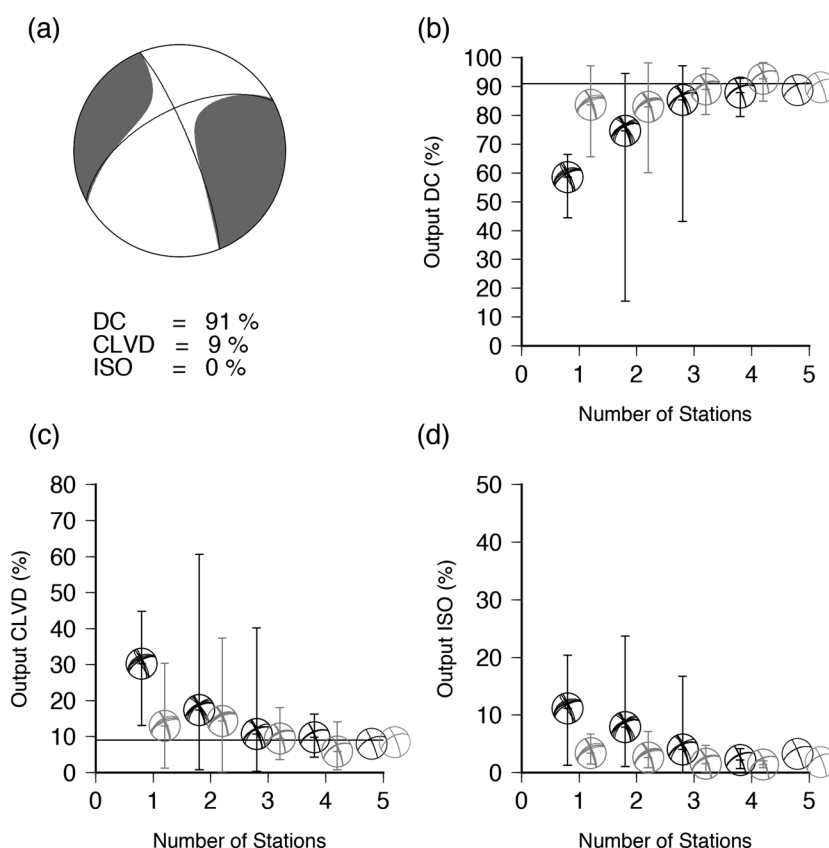
Figure 3. (a) Input mechanism for the synthetic test with a deviatoric input source. Percents of (b) DC, (c) CLVD, and (d) ISO components of the MAP solution from a Jackknife test using two groups of five stations. Horizontal lines show the DC, CLVD, and ISO input values. The fault plane solutions show nodal planes from all inversions with the corresponding number of stations and are centered at the mean value, while the error bars indicate the percent minimum and maximum values.

In addition to its shape, the covariance matrix has a factor corresponding to the noise magnitude. We treat it as a free parameter in the inversion, either as a common noise parameter for all stations or an individual noise variance at each station (Mustać & Tkalčić, 2017). Individual noise parameters for each station modify the stations' contribution in the inversion. We use the Bayesian evidence, as defined in Appendix A, for the purpose of model comparison, that is, to compare solutions obtained with different noise covariance matrices and different numbers of noise parameters.

3. Synthetic Tests

Since the Bayesian inversion is computationally expensive, we do not want to use more than a dozen stations. Moreover, we wish to explore how the noise model applied in the study affects the solutions. In theory, waveform data from one station are sufficient to determine the MT, but including data from more stations gives a better azimuthal coverage and consequently provides a better constraint on the solution (e.g., Šílený et al., 1996; Šílený, 2009). To explore the dependence of the DC, CLVD, and ISO component percentages on the azimuthal coverage, we perform a Jackknife test with two groups of five stations. The percentages of DC and CLVD components are determined assuming that the same principal stresses produce their radiation, while the percent of the ISO component is defined by the trace of the MT (e.g., Jost & Herrmann, 1989; Knopoff & Randall, 1970). We choose the locations of stations that were previously used to obtain the mechanism of the 20 October 2006 event in Mustać and Tkalčić (2017): BKS, GASB, HATC, MHC, and RAMR for the first group and BRIB, JRSC, HAST, SUTB, and WENL for the second group. We do not invert for the centroid location, only for the MT components and the noise. The synthetic seismograms and the Green's functions are calculated using the Gil7 structural model (e.g., Pasyanos et al., 1996). We add correlated noise such that its RMS is of 50% of the data RMS on all stations. It is created by multiplying uncorrelated Gaussian noise with a Cholesky
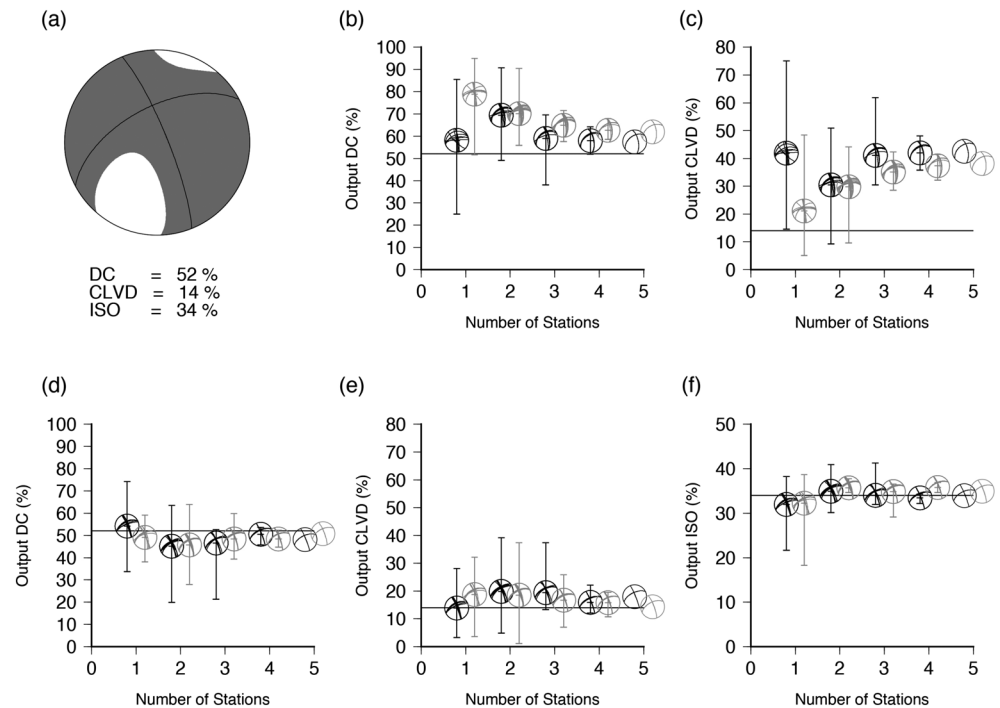
Figure 4. (a) Input mechanism for the second (deviatoric inversion) and third (full MT inversion) synthetic test. Percents of (b) DC and (c) CLVD components of the MAP solution in deviatoric inversions. Percents of (d) DC, (e) CLVD, and (f) ISO components of the MAP solution in full MT inversions. Horizontal lines show the DC, CLVD, and ISO input values. The fault plane solutions show nodal planes from all inversions with the corresponding number of stations and are centered at the mean value, while the error bars indicate the percent minimum and maximum values.

decomposition (e.g., Dettmer et al., 2004) of the two-attenuated-cosine covariance matrix. The inversion, where the noise is treated as one of the free parameters, is performed assuming the same kind of correlation.

First, we want to test if spurious ISO component exists due to the imperfect azimuthal coverage. We create the synthetic data using a deviatoric, predominantly DC, mechanism (Figure 3a) and perform the inversions assuming a full MT (DC+CLVD+ISO). On one hand, the nodal planes are well resolved in all inversions, even when waveform data from a single station are used (Figure 3). On the other hand, the uncertainties of non-DC components are large in inversions with one and two stations. The percents of each source component are not well resolved for certain stations (mostly from group 1, shown in black). As the number of stations increases, the non-DC component percentages approach the input values and the spurious ISO component becomes negligible (Figure 3d).

Second, we add a significant ISO component to the previous mechanism (Figure 4a) to test if we can correctly recover it in the inversion. When we assume a deviatoric mechanism in the inversions, the ISO component mostly gets mapped into the CLVD component when data from three or more stations are used (Figure 4c). The amount of the DC component is only slightly overestimated, but the nodal planes are slightly rotated from the input mechanism (Figure 4b). In inversions with one or two stations, there is more variability in the solutions, as in the previous test.

Finally, we perform inversions of the same data, now assuming a full MT. The nodal planes are again well resolved in all inversions (Figures 4d–4f). There is significant variability in percents of the DC, CLVD, and ISO components in inversions with one to three stations, but they are well resolved when four or five stations are used. On average, the variance reduction (VR = $1 - \int (\mathbf{G}(\mathbf{m}) - \mathbf{d})^2 / \int \mathbf{d}^2$) is a few percent higher in inversions assuming a full MT than in those assuming a deviatoric MT.

4. Inversions With Field Data

The 24 April 2007 and 11 January 2014 earthquakes had moment magnitudes of 4.46 and 4.53, estimated by the BSL, and were both smaller than the $M_W = 4.56$ earthquake from 20 October 2006 presented in our

previous paper (Mustać & Tkalčić, 2017). The 18 February 2004 earthquake had a slightly higher magnitude of $M_W = 4.58$ but was recorded on fewer stations, while the most recent, 14 December 2016 event, was the largest ($M_W = 5.01$). Considering the results from Mustać and Tkalčić (2017) where the algorithm was unsuccessful in retrieving a plausible centroid location due to the earthquake's shallow depth, we do not invert for the earthquake location but fix it to the NCEDC hypocenter. Furthermore, the locations estimated by the United States Geological Survey had the same epicenter and a depth less than a kilometer away from NCEDC locations, so we consider the location parameters robust. We only invert for the MT components and the noise level. We perform four inversions for each event: with both a diagonal and the two-attenuated-cosine covariance matrix, and with common and multiple noise parameters in both cases.

When processing the waveform data, we remove the instrument response and band-pass filter the data between 20 and 50 s using a one-pass Butterworth filter. The Green's functions (i.e., the elementary seismograms) are computed using the Gil7 structural model (e.g., Pasyanos et al., 1996) and filtered in the same way as the data. The Gil7 model is chosen because it gave a better fit than the SoCal model (Dreger & Helmberger, 1990) in linearized inversions. For each earthquake, we initially perform a linearized inversion and find the best time shifts between the data and the synthetics. The time shifts account for inadequacies of a one-dimensional model for particular ray paths and are subsequently used in the hierarchical Bayesian MT inversions.

We use 10 stations when inverting the 24 April 2007, 11 January 2014, and 14 December 2016 earthquake data and five stations for the 18 February 2004 earthquake, as less stations were operational at that time and it was challenging to choose data with good signal-to-noise ratios (SNRs). Particular stations are chosen based on a combination of high SNR and good azimuthal coverage (including stations north and south of the earthquake) for each event, their locations are shown in Figure 1a. The analyzed earthquakes are shown with black stars in Figure 1b and in bold font in Table 1.

4.1. The 18 February 2004 Earthquake

This earthquake had the largest ISO component (45%) obtained by Boyd et al. (2015) of the three examined there and in our study. Our MAP results with a diagonal covariance matrix (light and dark blue beachballs in Figure 5a) have an even larger ISO component of 58%. When a two-attenuated-cosine covariance matrix was applied to account for interdependence of data errors, the ISO component dropped to 44% in the inversion with a common Bayesian noise parameter (light red color) and 42% with multiple noise parameters (dark red color). A similar reduction in the ISO component can be seen for the remaining earthquakes. For this event, it could be a consequence of a better fit the diagonal $\mathbf{C_D}$ gives for the part of the signal preceding the largest amplitude on the radial and vertical components (Figure 5b). This could be an influence of the P wave coda that is interpreted as signal by the diagonal $\mathbf{C_D}$. Although body waves predominantly have higher frequency content, P wave coda has been observed in this frequency range (e.g., Neele & Snieder, 1991). In general, we conclude that the waveform fit for most traces is better when a diagonal $\mathbf{C_D}$ is used in the inversion, possibly because it overfits noise in the data. The noise value (given by numerical values on the right in Figure 5b) is highest for station JRSC in inversions with both the diagonal $\mathbf{C_D}$ and the two-attenuated-cosine $\mathbf{C_D}$.

A striking difference can be seen in the obtained nodal planes. Solutions obtained with a two-attenuated-cosine $\mathbf{C_D}$ have strike angles rotated by approximately 90° from strike angles obtained with a diagonal $\mathbf{C_D}$. The mechanism obtained with the two-attenuated-cosine covariance matrix is similar to what Johnson (2014b) had found for some earthquakes in the northwestern part of The Geysers. Normal mechanism earthquakes are common for larger events (e.g., Boyd et al., 2015), probably as a result of local stress variations. This is also consistent with results of Martínez-Garzón et al. (2013), who found a normal faulting regime at depths between 2 and 3.7 km (and strike-slip regime above and below). The MAP solutions with the two-attenuated-cosine $\mathbf{C_D}$ for all events are given in Table 2. Nodal planes of the whole ensemble of solutions (black lines in Figure 5a) show more variability when the two-attenuated-cosine $\mathbf{C_D}$ is used.

To get a better grasp on the mechanism of this earthquake, we also show the whole ensemble of solutions on the lune source-type diagram in Figure 5c. Ensembles from both inversions with a diagonal $\mathbf{C_D}$ (blue circles) are overlapping and are close to the opening crack mechanism. Ensembles from inversions with a two-attenuated-cosine $\mathbf{C_D}$ (red circles) cover a larger part of the lune plot, again indicating higher uncertainties. If we look at this the other way around, smaller uncertainties in solutions with a diagonal $\mathbf{C_D}$ could simply be a consequence of underestimating the noise in those inversions. The diagonal elements of the covariance matrix are not sufficient to properly account for the noise. In turn, the diagonal $\mathbf{C_D}$ interprets a part of the noise as the ISO component.
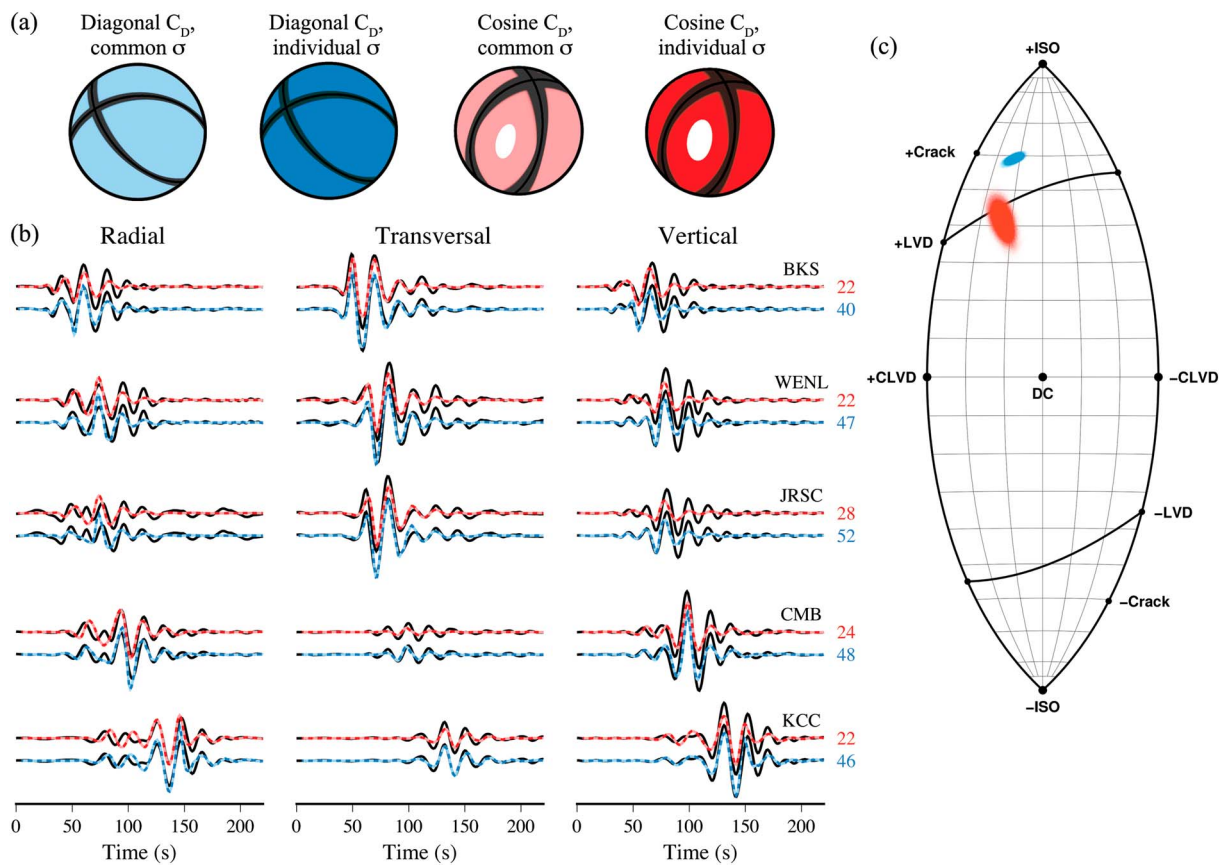
Figure 5. (a) Comparison of solutions with different assumptions for the noise covariance matrix for the 18 February 2004 event. The maximum a posteriori probability (MAP) solutions (shown in colors corresponding to a particular inversion) are overlain with nodal planes of the whole ensemble of solutions with density proportional to the posterior probability distribution. (b) Waveform data are shown in black, and synthetic seismograms from the corresponding MAP solutions are colored as in Figure 5a. Seismograms from inversions with multiple noise parameters are plotted with a dashed line. Numbers next to the waveforms show the MAP value for noise as percent of data RMS in inversions with multiple noise parameters. (c) Ensembles of solutions from the four inversions on a lune source-type diagram, colored as in Figure 5a with density proportional to the posterior probability distribution.

To examine this finding further, we compare the average autocorrelation of the standardized residuals from all four inversions in Figure 6a. Standardized residuals are the residuals $\mathbf{G}(\mathbf{m}) - \mathbf{d}$ scaled with the inverse of the lower triangular matrix of the Cholesky decomposition of $\mathbf{C_D}$ (e.g., Dettmer et al., 2007). We would expect them to be a delta function with a maximum for zero lag in the case of uncorrelated residuals, that is, for an adequate noise model incorporated in the $\mathbf{C_D}$ matrix. In inversions with the diagonal $\mathbf{C_D}$ (blue lines), they are still correlated and similar to the assumed correlation (black line). The correlation is reduced in inversions with the two-attenuated-cosine $\mathbf{C_D}$ (red lines). Similar behavior can be seen in the inversions of the remaining three (Figures 6b–6d) earthquakes. In other words, the inversion is fitting the data standardized by the decomposition of the covariance matrix $\mathbf{C_D}$ (e.g., Hallo & Gallovič, 2016).

Table 2
Values of M_{ij} Components of the MAP Solution With a Two-Attenuated-Cosine $\mathbf{C_D}$ and Multiple Noise Parameters for All Examined Earthquakes, and Their Variance Reductions

Earthquake	M_{xx}	M_{xy}	M_{xz}	M_{yy}	M_{yz}	M_{zz}	VR (%)
18 Feb 2004	1.74	1.37	0.20	4.81	1.13	−0.10	69.67
24 Apr 2007	−1.35	−2.08	0.54	2.69	1.17	−0.20	72.10
12 Jan 2014	−1.24	−1.79	0.27	5.20	0.78	−0.57	66.87
14 Dec 2016	−8.91	−13.80	2.61	27.09	5.22	4.89	69.15

Note. M_{ij} components are in Aki and Richards (2002) convection with units of 10^{15} Nm.
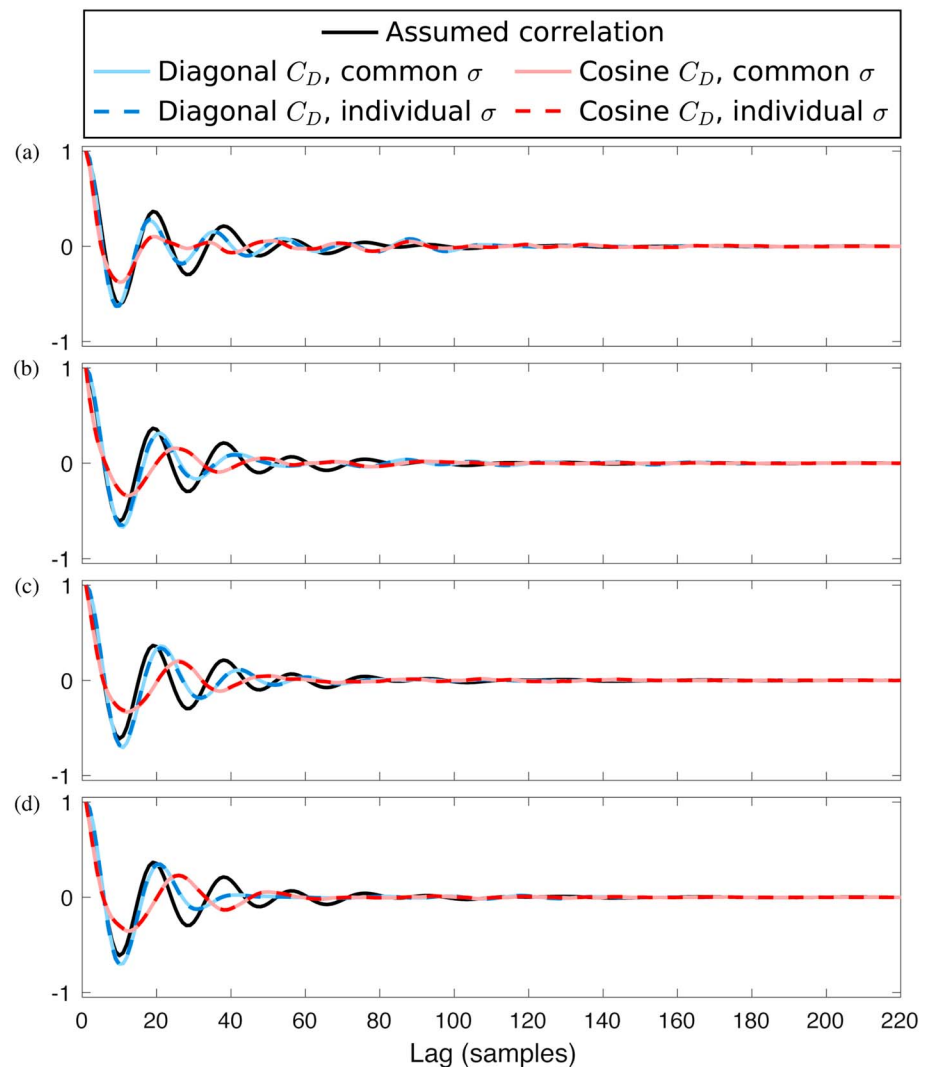
Figure 6. Autocorrelations of standardized residuals from the (a) 18 February 2004, (b) 24 April 2007, (c) 11 January 2014, and (d) 14 December 2016 earthquake inversions. Black line shows the correlation assumed in the inversion, and the remaining colors are as in Figure 5 (light blue for the inversion with a diagonal C_D and a common noise parameter, dark blue for multiple noise parameters, light red for the inversion with the empirical C_D and a common noise parameter, and dark red for multiple noise parameters).

Solutions obtained with a two-attenuated-cosine C_D and multiple noise parameters are close to the line connecting the DC mechanism and the tensional crack. For them, the earthquake mechanism can be explained by a combination of shear and tensile dislocations. In a geological setting with numerous fractures such as The Geysers, this is a realistic scenario. Crack opening might occur because of high-pressure fluids present in the rocks. This effect might be increased with water injected into the geothermal field to keep the high levels of electricity production.

4.2. The 24 April 2007 Earthquake

For this event, we have used data from 10 BDSN stations (Figure 7a). The highest noise values are obtained for stations HUMO, HELL, and YBH when using a diagonal C_D, and, similarly, stations HUMO, HELL, and SUTB when using the two-attenuated-cosine C_D. In general, the noise values are larger for more distant stations.

With a north-northwest striking nodal plane and an oblique-slip mechanism, MAP solutions from all four inversions have a similar DC part (Figure 7b), which agrees well with the regional stress field. Variability in nodal planes is similar in all inversions, but the waveform fit is once again better for the solutions obtained with a diagonal noise covariance matrix. However, this might be dubious for records that are noisy in the frequency
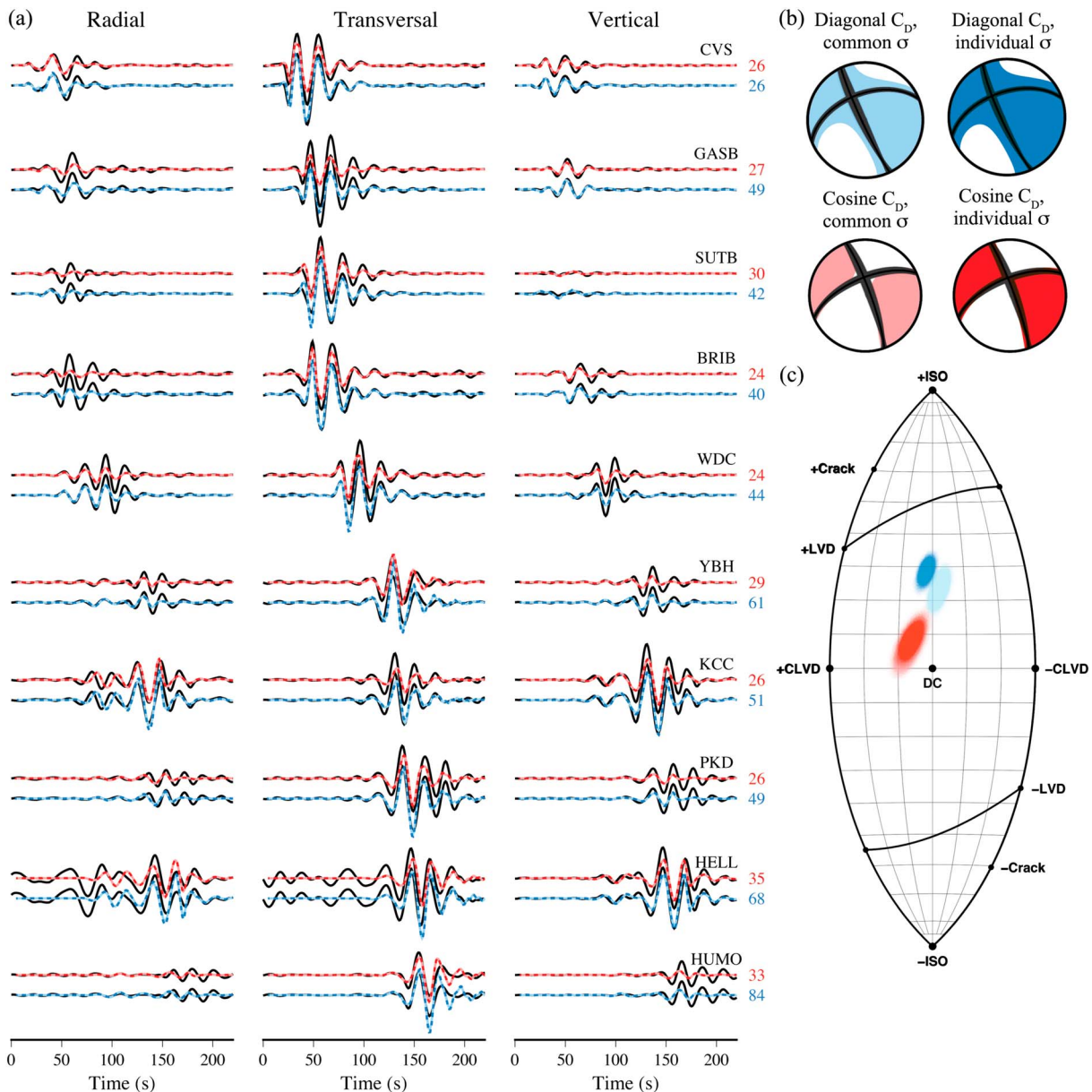
Figure 7. (a) Waveform data and synthetic seismograms corresponding to the MAP solutions of four inversions for the 24 April 2007 event, together with the noise values and colored as in Figure 5. (b) Comparison of MAP solutions and ensemble nodal planes. (c) Ensembles of solutions shown on a lune source-type diagram.

range considered in this study (e.g., HELL). As the inversions with the two-attenuated-cosine covariance matrix interpret more features of the waveform as noise, the scalar moment is around 40% lower when compared to solutions with a diagonal covariance matrix.

When we examine the non-DC parts of the solutions, MAP solutions with a diagonal C_D are similar to the linearized inversion result obtained by Boyd et al. (2015), which was 31% isotropic. Inversions with the two-attenuated-cosine C_D once more show larger uncertainties and a considerably lower ISO component (Figure 7c). The CLVD component is also somewhat larger, but the MAP solutions are predominantly DC (63–69%).

4.3. The 11 January 2014 Earthquake

Once again the waveform fit is better in inversions with a diagonal C_D for most stations (Figure 8a). Inversions with both covariance matrices and multiple noise parameters have the highest noise value for stations YBH
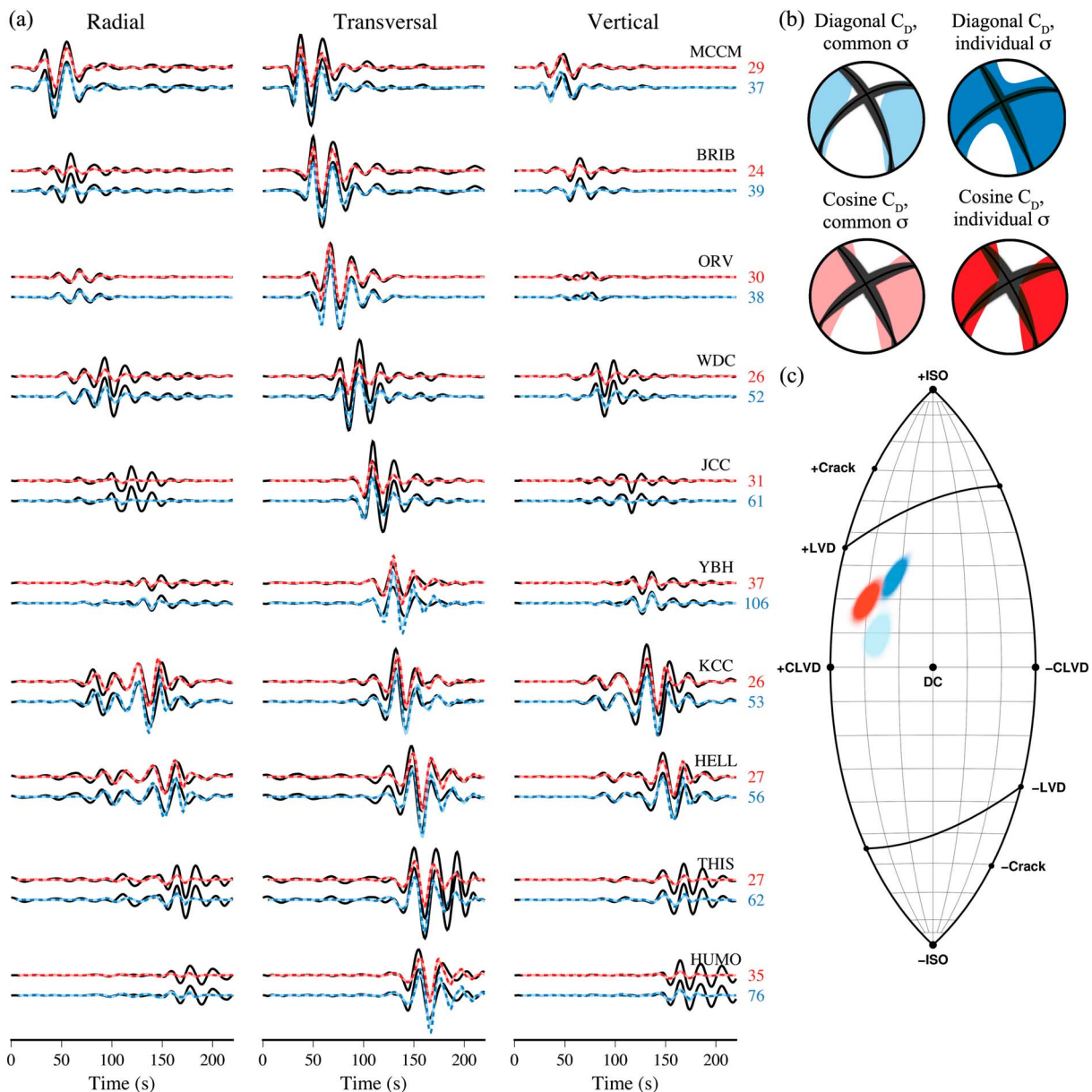
Figure 8. (a) Waveform data and synthetic seismograms corresponding to the MAP solutions of four inversions for the 11 January 2014 event, together with the noise values and colored as in Figure 5. (b) Comparison of MAP solutions and ensemble nodal planes. (c) Ensembles of solutions shown on a lune source-type diagram.

and HUMO. These noise values are followed by those of THIS and JCC in the inversion with a diagonal $\mathbf{C_D}$, and stations JCC and ORV in the inversion with a two-attenuated-cosine $\mathbf{C_D}$.

The DC components of all solutions for this event again have a nodal plane consistent with the regional stress field, but their uncertainties are larger than for the previous event (Figure 8b). However, this is not the dominant component of the mechanism (Figures 8b and 8c). MAP solutions obtained with a diagonal noise covariance matrix are ~40% DC, and the solutions with a two-attenuated-cosine $\mathbf{C_D}$ are only ~25% DC and 52–54% CLVD. All inversions yield significant positive CLVD and ISO components for this earthquake, but the CLVD component is larger when the two-attenuated-cosine $\mathbf{C_D}$ is used (Figure 8c). Some trade-off between the CLVD and ISO components can be seen for all inversions except the simplest one with a diagonal $\mathbf{C_D}$ and a common noise parameter for all stations (light blue color).
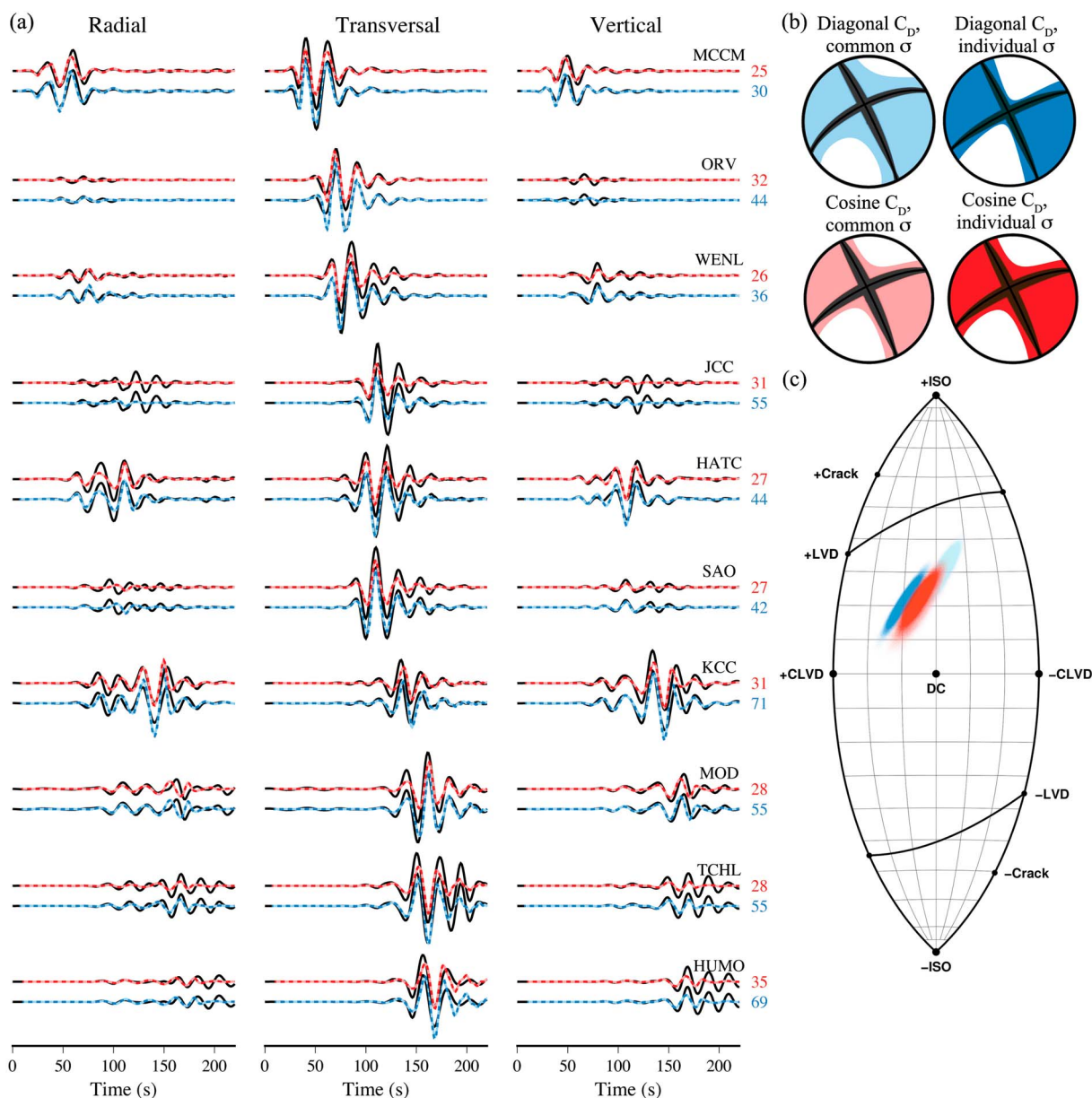
Figure 9. (a) Waveform data and synthetic seismograms corresponding to the MAP solutions of four inversions for the 14 December 2016 event, together with the noise values and colored as in Figure 5. (b) Comparison of MAP solutions and ensemble nodal planes. (c) Ensembles of solutions shown on a lune source-type diagram.

The mechanism of this earthquake could still be a combination of shear and tensile faulting. However, for rocks with a Poisson ratio of 0.25, such events lie on or above the line connecting the DC and +Crack point on the lune diagram (Tape & Tape, 2013), and most of the solutions from the ensemble obtained with a two-attenuated-cosine C_D are under this line. This could indicate that the event occurred in an area with rocks having a lower Poisson ratio, which might be an effect of open cracks in the rocks (e.g., Walsh, 1965) or rocks under tension (e.g., Homand-Etienne & Houpert, 1989). Alternatively, it could mean that crack opening during the earthquake was complicated, or that the ISO component was reduced due to fluid extraction.

4.4. The 14 December 2016 Earthquake

The most recent event had the highest moment magnitude of 5.01 (Table 1). The retrieved noise values are in general higher for more distant stations, with the exception of high values for ORV and JCC (Figure 9a). Station ORV has the lowest amplitude on radial and vertical components, probably because it lies in the direction of
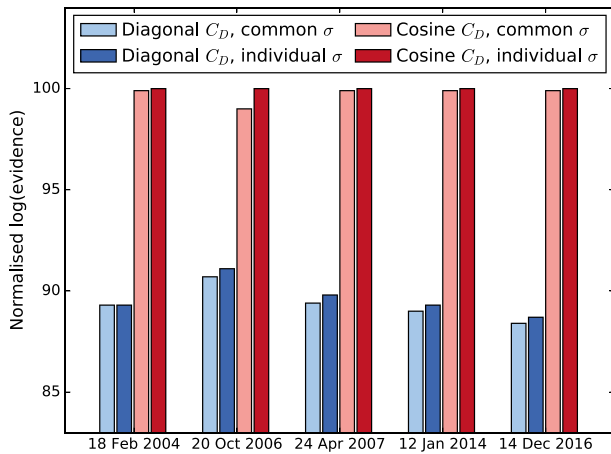
Figure 10. Evidence values for inversions of the 18 February 2004, 20 October 2006, 24 April 2007, and 11 January 2014 earthquakes, normalized with the highest value and colored as in the previous figures.

one of the nodal planes, while JCC has somewhat complex later part of the radial and vertical seismograms, possibly due to site effects.

The DC part is similar among the four inversions and also similar to the previous two earthquakes (Figure 9b). Nodal plane uncertainties are smaller than for the 2014 event, but this event has the highest uncertainties in the non-DC components and the largest trade-off between CLVD and ISO components (Figure 9c). Both inversions with a two-attenuated-cosine $\mathbf{C_D}$ yield similar solutions, while the inversion with a diagonal $\mathbf{C_D}$ and common noise parameters has a higher ISO component, and the one with individual noise parameters yield a higher CLVD component.

4.5. Bayesian Evidence as a Way to Evaluate the Noise Models

When we compare waveform fit between data and synthetic seismograms obtained from the MAP inversions using the empirically estimated two-attenuated-cosine covariance matrix and the diagonal covariance matrix, the latter one is often better. From a standpoint of a traditional inversion, where the best waveform fit defines the best solution, we would say that a simple diagonal matrix gives a better result. However, Mustać and Tkalčić (2017) demonstrated in a synthetic experiment that the diagonal covariance matrix can overfit the noise. Therefore, a quantitative measure is desired to help guide a choice of the noise covariance matrix.

To address this issue, we return to the Bayes' theorem presented in equation (1) and use the denominator $p(\mathbf{d})$, called the Bayesian evidence, to compare solutions obtained with different noise covariance matrices. Multidimensional integration over the prior is necessary to compute the evidence but computationally expensive. Thus, we use an approximation of the evidence (e.g., Beck & Yuen, 2004; Papadimitriou et al., 1997; Sambridge et al., 2006)

$$p(\mathbf{d}|k) = p(\mathbf{d}|\hat{\mathbf{m}},k)p(\hat{\mathbf{m}}|k)(2\pi)^{k/2}|\mathbf{C'_M}|^{1/2}, \qquad (3)$$

where k is the number of parameters in the inversion, $p(\mathbf{d}|\hat{\mathbf{m}},k)$ is the likelihood of the MAP model $\mathbf{m}$, $p(\hat{\mathbf{m}}|k)$ its prior value, and $\mathbf{C'_M}$ is the posterior model covariance matrix. Additional details and the derivation of the evidence for this particular problem are given in Appendix A.

Results for all events are similar (Figure 10). Evidence values are higher when the empirical two-attenuated-cosine covariance matrix is used, meaning that these are the preferred results from a Bayesian point of view. The difference is around 10% of the highest evidence value. Apart from that, inversions with multiple noise parameters have slightly higher evidences than the ones with a common noise parameter. This difference is small when the ensembles of solutions are overlapping (e.g., for the diagonal $\mathbf{C_D}$ matrix solutions for the 18 February 2004 event).

5. Discussion and Conclusions

In synthetic experiments using the BDSN stations and a source located in The Geysers region, we have shown that a full MT inversion creates a negligible ISO component for deviatoric mechanisms, and that the non-DC components can be retrieved well when using data from four or more stations. Furthermore, we have examined four $M_W \geq 4.5$ earthquakes from The Geysers region using the hierarchical Bayesian MT inversion method and compared results with a diagonal and an empirical two-attenuated-cosine noise covariance matrix. In addition to inverting for the noise level, the two-attenuated-cosine covariance matrix can account for the observed interdependence of data errors through its nondiagonal terms. The results show a variety of source mechanisms.

Three of the earthquakes examined here (24 April 2007, 11 January 2014, and 14 December 2016) have a nodal plane consistent with the regional stress field. The 18 February 2004 earthquake is rather intriguing: it has a large ISO component of ∼60% in inversions with the diagonal $\mathbf{C_D}$ and ∼40% with the two-attenuated-cosine $\mathbf{C_D}$. Non-DC components obtained with the two-attenuated-cosine $\mathbf{C_D}$ still require a volume-increasing process occurring simultaneously or sequentially with shear faulting. The most obvious candidate for volume increase in this region of highly fractured rocks is crack opening, perhaps related to water injection.

The two-attenuated-cosine C_D had the effect of reducing the ISO component for the first three events but had a smaller influence on the 14 December 2016 event, possibly due to a higher SNR. It gave a predominantly double-couple solution for the 24 April 2007 event. For the 11 January 2014 event, it also increased the CLVD component and yielded a solution that requires a lower Poisson ratio, a more complicated crack opening process or an additional process that reduces the volumetric component, such as fluid extraction.

Retrieved mechanisms with positive volumetric components agree well with observations of Boyd et al. (2015). However, these results are unexpected in a region that is undergoing volumetric contraction, as determined in geodetic studies. A mechanism that explains positive volumetric components was recently proposed by Johnson (2014a). It involves cracks that extend into wing cracks oriented in the direction of the most compressive stress. Additional volume increase could be achieved by interaction of nearby cracks, a process that is likely to occur in a highly fractured environment such as this (e.g., Ashby & Sammis, 1990, give an approximate physical model for this process).

Results obtained when using the empirical two-attenuated-cosine C_D are difficult to justify in a standard framework because they yield a poorer waveform fit compared to results with the diagonal C_D. However, the Bayesian evidence used for model comparison favors these results, perhaps because noise changes the phase of the signal together with the amplitude. The largest discrepancy in solutions is obtained for the 18 February 2004 event, where the two-attenuated-cosine C_D yielded strike angles rotated by about 90° relative to results with a diagonal C_D. This might be due to its large volumetric component and also because of fewer stations used. All stations are south of the earthquake, which is a significant reduction in coverage. The effect of azimuthal coverage has not been examined for these station and event locations, but the synthetic experiment presented in Mustać and Tkalčić (2016) showed that stations from a similar azimuthal range can constrain the solution when waveform data are used.

Additional factors that can affect the results are the shallow depths of these events and the use of a one-dimensional structural model to compute the Green's functions. Including a composite structural model was shown to be significant for events in complex geological environments (e.g., the unusual CLVD mechanism of an earthquake in the Bárðarbunga volcano, Iceland; Tkalčić et al., 2009).

Considering that we do not explicitly include theory errors when constructing the C_D matrix, these errors get mapped into the noise variance parameters. Yagi and Fukahata (2011) separated the modeling and observational error in their analysis of the C_D matrix in a finite fault study, but they assumed an uncorrelated, Gaussian error of Green's functions. Our noise parameters are higher than what we would expect from solely ambient noise (as they are 1/SNR) and have values more similar to what Hallo and Gallovič (2016) obtained when considering Green's functions' uncertainties. They included uncertainties that stem from randomly perturbed velocity models and later simplified the computation by approximating the Green's function covariance using time shifts from the mean velocity model. Both works are a step forward in including the Green's function uncertainties, but a complete method would require access to uncertainties of the tomographic velocity and attenuation models, which are not routinely calculated in seismic imaging. However, recent efforts to accompany regional tomographic inversions by uncertainty (e.g., Bodin & Sambridge, 2009; Kim et al., 2016; Young et al., 2011), hold keys to performing the MT inversions with a more complete description of the covariance matrix.

In summary, the DC components of 24 April 2007, 11 January 2014, and 14 December 2016 are consistent with the regional stress field, and the 18 February 2004 earthquake is more affected by local stress variations. The 24 April 2007 earthquake is characterized as predominantly double couple, while the 18 February 2004 and 11 January 2014 events are a result of combined shear and tensile faulting.

Our study illustrates that analyzing the effects of noise prior to the MT inversion and properly handling them in the inversion is of a critical importance because the results and subsequent interpretation can vary significantly. At this stage, despite the fact that inversions with the two-attenuated-cosine C_D yielded a lower ISO component than inversions with the diagonal C_D for all examined earthquakes, we cannot make a general statement with regards to whether the inclusion of the data noise model acts to suppress the spurious non-double-couple components or it acts to emphasize true MT components (including the non-double-couple radiation in the source). It is likely that both can occur as a result of the isolation of the effects of noise. Since our noise model with an empirical C_D is favored by the Bayesian evidence, it means that a part of the ISO component obtained in standard inversions is merely an artifact of fitting the noise in the

data. A caution should be exercised when interpreting inversions where the effects of noise in the data and their interdependence had not been taken into account.

It should also be noted that in this study we did not investigate possible effects of theory errors (e.g., uncertainty in the model of Earth's structure), but the inclusion of these effects through the noise covariance matrix is at the cutting edge of research in this area and will be investigated in future studies.

Appendix A: The Bayesian Evidence

The evidence $p(\mathbf{d})$ is defined as an integral over $p(\mathbf{d}|\mathbf{m})p(\mathbf{m})$ over the model parameters $\mathbf{m}$. We give the expression for the Bayesian evidence for the most general case of a two-attenuated-cosine covariance matrix and different noise parameters for each station, and with the source location as a model parameter. Since we treat the noise as a free parameter in the inversion, the posterior (and hence the evidence) depends on the noise hyperparameters σ and the number of parameters in the inversion k. The integral has the form

$$p(\mathbf{d}|k,\sigma) = \int p(\mathbf{d}|\mathbf{m},k,\sigma)p(\mathbf{m})p(\sigma|k)\mathrm{d}\mathbf{m}, \tag{A1}$$

where $p(\sigma|k)$ is the prior distribution of the noise parameters. We consider the noise parameters σ_i on each station i to be independent. If n_r stations are used, the prior distribution of σ is defined as

$$p(\sigma) = \prod_{i=1}^{n_r} p(\sigma_i). \tag{A2}$$

The parameters σ_i are constrained by root-mean-square (RMS) of the noise at that particular station $\mathrm{RMS}(u_i)$. The prior on σ_i is uniform and has the form

$$p(\sigma_i) = \begin{cases} 1/\Delta\sigma_i & \text{if } 0 < \sigma_i < 5 * \mathrm{RMS}(u_i) \\ 0 & \text{otherwise,} \end{cases} \tag{A3}$$

where $\Delta\sigma_i = 5 * \mathrm{RMS}(u_i)$. To obtain the exact value of the evidence (independent of σ), we would need to integrate equation (A1) over the noise parameters. This is a demanding task to do analytically, so we approximate the integral with its value at the maximum *a posteriori* probability (MAP) of noise parameters, $\hat{\sigma}$. This returns us to equation (A1) without the dependence on σ, but we need to keep this approximation in mind.

We further simplify the computation using the MAP solution of the model parameters, as in equation (3). Since the evidence has numerically large values, we compute its logarithm

$$\log\left[p(\mathbf{d}|k)\right] = \log\left[p(\mathbf{d}|\hat{\mathbf{m}},k)\right] + \log\left[p(\hat{\mathbf{m}}|k)\right] + \frac{k}{2}\log(2\pi) + \frac{1}{2}\log\left[|\mathbf{C'_M}|\right]. \tag{A4}$$

In the most general case, the number of parameters in the inversion is $k = N_L + 6 + N_n$, where N_L is the number of sampled discrete centroid locations and N_n the number of noise parameters in the inversion, which is equal to the number of stations n_r.

The likelihood, defined in equation (2), depends on k through the number of noise parameters. Its logarithm evaluated at $\hat{\mathbf{m}}$ is

$$\log\left[p(\mathbf{d}|\hat{\mathbf{m}},k)\right] = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{C_D}| - \frac{1}{2}\left[\mathbf{G}(\hat{\mathbf{m}}) - \mathbf{d}\right]^T \mathbf{C_D^{-1}}(\mathbf{G}(\hat{\mathbf{m}}) - \mathbf{d}). \tag{A5}$$

The matrix $\mathbf{C_D}$ is block diagonal, with each block $\mathbf{C_i}$ corresponding to a particular station. We have

$$\mathbf{C_i} = \sigma_i^2\mathbf{R}, \tag{A6}$$

where $\mathbf{R}$ is the matrix of correlations for three-component seismograms. The matrix $\mathbf{R}$ consists of three $N_0 \times N_0$ matrices $\mathbf{R_0}$ that parameterize noise correlations for one seismogram component (radial, transversal, and vertical), and N_0 is the number of points in a single seismogram

$$\mathbf{R} = \begin{pmatrix} \mathbf{R_0} & & 0 \\ & \mathbf{R_0} & \\ 0 & & \mathbf{R_0} \end{pmatrix}. \tag{A7}$$

Finally, we compose the matrix $\mathbf{C_D}$ as

$$\mathbf{C_D} = \begin{bmatrix} \sigma_1^2 \mathbf{R} & & 0 \\ & \ddots & \\ 0 & & \sigma_{n_r}^2 \mathbf{R} \end{bmatrix}. \tag{A8}$$

To compute the expression A5, we need the determinant of this matrix, which is equal to the product of the determinants of matrices down its diagonal. Hence,

$$|\mathbf{C_D}| = \left[\prod_{i=1}^{n_r} (\sigma_i^2)^{3N_0} \right] \cdot |\mathbf{R_0}|^{3n_r}. \tag{A9}$$

Because of its block-diagonal structure, the inverse of matrix $\mathbf{C_D}$ is also simple to obtain

$$\mathbf{C_D^{-1}} = \begin{bmatrix} \frac{1}{\sigma_1^2}\mathbf{R}^{-1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_{n_r}^2}\mathbf{R}^{-1} \end{bmatrix}. \tag{A10}$$

To compute the second term in equation (A4), we need to determine the form of the prior distribution $p(\mathbf{m})$. In the most general case, the model vector $\mathbf{m}$ consists of the location and MT parameters. As they are independent in our study, we can write

$$p(\mathbf{m}) = p(\mathbf{a})p(\mathbf{r}), \tag{A11}$$

where $\mathbf{a}$ is a vector with elements a_n that define the seismic MT and $\mathbf{r}$ is the location vector. MT elements are constrained by the initial estimate of the scalar moment M_{00} in the following way:

$$p(a_n) = \begin{cases} 1/\Delta a_n & \text{if } -(1+\epsilon) \cdot M_{00} < a_n < (1+\epsilon) \cdot M_{00} \\ 0 & \text{otherwise,} \end{cases} \tag{A12}$$

where $\Delta a_n = 2 \cdot (1+\epsilon) \cdot M_{00}$. We consider a_n independent and include an additional constraint on the value of the scalar moment

$$p(M_0|\mathbf{a}) = \begin{cases} 1/\Delta M_0 & \text{if}(1-\epsilon) \cdot M_{00} < M_0 < (1+\epsilon) \cdot M_{00} \\ 0 & \text{otherwise,} \end{cases} \tag{A13}$$

with $\Delta M_0 = 2\epsilon M_{00}$. Thus, we can write for p($\mathbf{a}$)

$$p(\mathbf{a}) = \left[\prod_{i=1}^{6} p(a_n) \right] p(M_0|\mathbf{a}). \tag{A14}$$

After some trial and error, ϵ was set to 70% for these earthquakes. Due to the necessity of having elementary seismograms computed, the centroid is a discrete variable with N_L possible values, which makes the prior for the centroid location equal to

$$p(\mathbf{r}) = 1/N_L. \tag{A15}$$

Putting together the last four equations we obtain for the prior

$$p(\mathbf{m}) = \frac{1}{N_L \cdot (\Delta a_n)^6 \cdot \Delta M_0}. \tag{A16}$$

The second term in equation (A4) now has the form

$$\log \left[p(\hat{\mathbf{m}}|k) \right] = -\log(N_L) - 6 \cdot \log(\Delta a_n) - \log(\Delta M_0). \tag{A17}$$

That leaves us with the final component in equation (3), the determinant of the posterior model covariance matrix, $|\mathbf{C'_M}|$. For the MAP solution, where the location is fixed, we have a linear relationship with the model

parameters and the data, that is, we can compute $\mathbf{G}(\mathbf{m})$ as matrix multiplication $\mathbf{Am}$. The matrix $\mathbf{A}$ consists of the six elementary seismograms for each record and has the shape $(3 \cdot n_r \cdot N_0) \times 6$. The matrix $\mathbf{C'_M}$ can be computed as

$$\mathbf{C'_M} = \left(\mathbf{A}^T \mathbf{C_D^{-1}} \mathbf{A} + \mathbf{C_M^{-1}}\right)^{-1}, \tag{A18}$$

where $\mathbf{C_M}$ is the prior model covariance matrix of the form $\mathbf{C_M} = \sigma_m^2 \mathbf{I}$, $\mathbf{I}$ is the identity matrix, and σ_m^2 variances of the prior. We are using uniform priors, which can be seen as a special case of a Gaussian where all σ_m^2 approach infinity. In that case, $\mathbf{C_M^{-1}}$ approaches zero and we compute the logarithm of the determinant of $\mathbf{C'_M}$ only using the first component of the right-hand side of A18.

We have shown how to compute the evidence (i.e., its logarithm) for the most general case. When the noise covariance matrix $\mathbf{C_D}$ is diagonal, the matrix $\mathbf{R}$ is an identity matrix for all stations, and the computations are simpler. In the case of only one noise parameter, all noise variances σ_i^2 have equal values and the total number of parameters is $k = N_L + 6 + 1$.

References

Aki, K., & Richards, P. G. (2002). *Quantitative seismology* (2nd ed.). Sausalito, CA: University Science Books.

Ashby, M., & Sammis, C. (1990). The damage mechanics of brittle solids in compression. *Pure and Applied Geophysics*, *133*(3), 489–521.

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

Beck, J., & Yuen, K. (2004). Model selection using response measurements: Bayesian probabilistic approach. *Journal of Engineering Mechanics*, *130*(2), 192–203.

Blakely, R. J., & Stanley, W. (1993). The Geyser's magma chamber, California: Constraints from gravity data, density measurements, and well information. *Transactions of the Geothermal Resources Council*, *17*, 227–233.

Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, *178*(3), 1411–1436. https://doi.org/10.1111/j.1365-246X.2009.04226.x

Boyd, O. S., Dreger, D. S., Lai, V. H., & Gritto, R. (2015). A systematic analysis of seismic moment tensor at The Geysers geothermal field, California. *Bulletin of the Seismological Society of America*, *105*(6), 2969–2986. https://doi.org/10.1785/0120140285

Boyle, K., & Zoback, M. (2014). The stress state of the northwest Geysers, California geothermal field, and implications for fault-controlled fluid flow. *Bulletin of the Seismological Society of America*, *104*(5), 2303–2312. https://doi.org/10.1785/0120130284

Brook, C. A. (1981). Variability and sources of hydrogen sulfide and other gases in steam at The Geysers. In R. J. McLaughlin, & J. M. Donnelly-Nolan (Eds.), *Research in The Geysers-Clear Lake geothermal area* (Vol. 1141, pp. 193–203). Washington, DC: U.S. Geological Survey Professional Paper.

Bufe, G., Lester, F., & Marks, S. (1981). Seismicity of The Geysers-Clear Lake region, California. In R. J. McLaughlin, & J. M. Donnelly-Nolan (Eds.), *Research in The Geysers-Clear Lake geothermal area* (Vol. 1141, pp. 129–138). Washington, DC: U.S. Geological Survey Professional Paper.

Castagna, J. P., Batzle, M. L., & Eastwood, R. L. (1985). Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. *Geophysics*, *50*(4), 571–581.

Cotton, F., & Coutant, O. (1997). Dynamic stress variations due to shear faults in a plane-layered medium. *Geophysical Journal International*, *128*, 676–688.

Denlinger, R. P., & Bufe, C. G. (1982). Reservoir conditions related to induced seismicity at The Geysers steam reservoir, Northern California. *Bulletin of the Seismological Society of America*, *72*(4), 1317–1327.

Dettmer, J., Dosso, S. E., & Holland, C. W. (2004). Geoacoustic inversion with strongly correlated data errors. *Canadian Acoustics*, *32*, 194–195.

Dettmer, J., Dosso, S. E., & Holland, C. W. (2007). Uncertainty estimation in seismo-acoustic reflection travel time inversion. *Journal of the Acoustical Society of America*, *122*(1), 161–176.

Dreger, D., & Helmberger, D. (1990). Broadband modeling of local earthquakes. *Bulletin of the Seismological Society of America*, *80*(5), 1162–1179.

Elkibbi, M., & Rial, J. (2005). The Geysers geothermal field: Results from shear-wave splitting analysis in a fractured reservoir. *Geophysical Journal International*, *162*(3), 1024–1035. https://doi.org/10.1111/j.1365-246X.2005.02698.x

Foulger, G. R., Grant, C. C., Ross, A., & Julian, B. R. (1997). Industrially induced changes in Earth structure at The Geysers geothermal area, California. *Geophysical Research Letters*, *24*(2), 135–137.

Gouveia, W. P., & Scales, J. A. (1998). Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis. *Journal of Geophysical Research*, *103*(B2), 2759–2779. https://doi.org/10.1029/97JB02933

Guilhem, A., Hutchings, L., Dreger, D. S., & Johnson, L. R. (2014). Moment tensor inversions of M∼3 earthquakes in The Geysers geothermal fields, California. *Journal of Geophysical Research: Solid Earth*, *119*, 2121–2137. https://doi.org/10.1002/2013JB010271

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

Homand-Etienne, F., & Houpert, R. (1989). Thermally induced microcracking in granites: Characterization and analysis. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstract*, *26*(2), 125–134. https://doi.org/10.1016/0148-9062(89)90001-6

Isherwood, W. F. (1981). Geophysical overview of The Geysers. In R. J. McLaughlin, & J. M. Donnelly-Nolan (Eds.), *Research in The Geysers-Clear Lake geothermal area* (Vol. 1141, pp. 83–96). Washington, DC: U.S. Geological Survey Professional Paper.

Iyer, H., Oppenheimer, D., & Hitchcock, T. (1979). Abnormal *P*-wave delays in The Geysers-Clear Lake geothermal area, California. *Science*, *204*(4392), 495–497.

Johnson, L. R. (2014a). A source model for induced earthquakes at The Geysers geothermal reservoir. *Pure and Applied Geophysics*, *171*(8), 1625–1640.

Johnson, L. R. (2014b). Source mechanisms of induced earthquakes at The Geysers geothermal reservoir. *Pure and Applied Geophysics*, *171*(8), 1641–1668.

Jost, M. u., & Herrmann, R. (1989). A student's guide to and review of moment tensors. *Seismological Research Letters*, *60*(2), 37–57.

Julian, B. R., Ross, A., Foulger, G. R., & Evans, J. R. (1996). Three-dimensional seismic image of a geothermal reservoir: The Geysers, California. *Geophysical Research Letters*, *23*(6), 685–688.

Kikuchi, M., & Kanamori, H. (1991). Inversion of complex body waves—III. *Bulletin of the Seismological Society of America*, *81*, 2335–2350.

Kim, S., Tkalčić, H., Rhie, J., & Chen, Y. (2016). Intraplate volcanism controlled by back-arc and continental structures in NE Asia inferred from transdimensional Bayesian ambient noise tomography. *Geophysical Research Letters*, *43*, 8390–8398. https://doi.org/10.1002/2016GL069483

Kirkpatrick, A., Romero, A., Peterson, J., Johnson, L. R., & Ernest, M. (1996). Microearthquake source mechanism studies at The Geysers geothermal field (Technical Report). CA: Lawrence Berkeley Laboratory Report.

Knopoff, L., & Randall, M. (1970). The compensated linear-vector dipole: A possible mechanism for deep earthquakes. *Journal of Geophysical Research*, *75*(26), 4957–4963.

Kuge, K., & Lay, T. (1994). Systematic non-double-couple components of earthquake mechanisms: The role of fault zone irregularity. *Journal of Geophysical Research*, *99*(B8), 15,457–15,467. https://doi.org/10.1029/94JB00140

Křížová, D., Zahradník, J., & Kiratzi, A. (2013). Resolvability of isotropic component in regional seismic moment tensor inversion. *Bulletin of the Seismological Society of America*, *103*, 2460–2473.

Lofgren, B. (1981). Monitoring crustal deformation in The Geysers-Clear Lake region. In R. J. McLaughlin, & J. M. Donnelly-Nolan (Eds.), *Research in The Geysers-Clear Lake geothermal area* (Vol. 1141, pp. 139–148). Washington, DC: U.S. Geological Survey Professional Paper.

Majer, E. L., Baria, R., Stark, M., Oates, S., Bommer, J., Smith, B., & Asanuma, H. (2007). Induced seismicity associated with enhanced geothermal systems. *Geothermics*, *36*(3), 185–222.

Martínez-Garzón, P., Bohnhoff, M., Kwiatek, G., & Dresen, G. (2013). Stress tensor changes related to fluid injection at The Geysers geothermal field, California. *Geophysical Research Letters*, *40*, 2596–2601. https://doi.org/10.1002/grl.50438

Martínez-Garzón, P., Kwiatek, G., Bohnhoff, M., & Dresen, G. (2017). Volumetric components in the earthquake source related to fluid injection and stress state. *Geophysical Research Letters*, *44*, 800–809. https://doi.org/10.1002/2016GL071963

McLaughlin, R. J. (1981). Tectonic setting of pre-Tertiary rocks and its relation to geothermal resources in The Geysers-Clear Lake area. In R. J. McLaughlin, & J. M. Donnelly-Nolan (Eds.), *Research in The Geysers-Clear Lake geothermal area* (Vol. 1141, pp. 3–24). Washington, DC: U.S. Geological Survey Professional Paper.

Mossop, A., & Segall, P. (1997). Subsidence at The Geysers geothermal field, N. California from a comparison of GPS and leveling surveys. *Geophysical Research Letters*, *24*(14), 1839–1842.

Mossop, A., & Segall, P. (1999). Volume strain within The Geysers geothermal field. *Journal of Geophysical Research*, *104*(B12), 29,113–29,131.

Mustać, M., & Tkalčić, H. (2016). Point source moment tensor inversion through a Bayesian hierarchical model. *Geophysical Journal International*, *204*(1), 311–323.

Mustać, M., & Tkalčić, H. (2017). On the use of data noise as site-specific weight parameter in a hierarchical Bayesian moment tensor inversion: The case study of The Geysers and Long Valley caldera earthquakes. *Bulletin of the Seismological Society of America*, *107*(4), 1914–1922.

Neele, F., & Snieder, R. (1991). Are long-period body wave coda caused by lateral heterogeneity? *Geophysical Journal International*, *107*(1), 131–153.

O'Connell, D. R., & Johnson, L. R. (1988). Second-order moment tensors of microearthquakes at The Geysers geothermal field, California. *Bulletin of the Seismological Society of America*, *78*(5), 1674–1692.

O'Connell, D. R., & Johnson, L. R. (1991). Progressive inversion for hypocenters and *P* wave and *S* wave velocity structure: Application to The Geysers, California, geothermal field. *Journal of Geophysical Research*, *96*(B4), 6223–6236.

Papadimitriou, C., Beck, J., & Katafygiotis, L. (1997). Asymptotic expansions for reliability and moments of uncertain systems. *Journal of Engineering Mechanics*, *123*(12), 1219–1229.

Pasyanos, M., Dreger, D. S., & Romanowicz, B. (1996). Toward real-time estimation of regional moment tensors. *Bulletin of the Seismological Society of America*, *88*(5), 1255–1269.

Phạm, T.-S., & Tkalčić, H. (2017). On the feasibility and use of teleseismic *P* wave coda autocorrelation for mapping shallow seismic discontinuities. *Journal of Geophysical Research: Solid Earth*, *122*, 3776–3791. https://doi.org/10.1002/2017JB013975

Romero, A. E., Kirkpatrick, A., Majer, E. L., & Peterson, J. E. (1994). Seismic monitoring at The Geysers geothermal field. *Transactions of the Geothermal Resources Council*, *18*, 331–338.

Romero, A. E., McEvilly, T. V., & Majer, E. L. (1997). 3-D microearthquake attenuation tomography at the Northwest Geysers geothermal region, California. *Geophysics*, *62*(1), 149–167.

Ross, A., Foulger, G. R., & Julian, B. R. (1996). Non-double-couple earthquake mechanisms at The Geysers geothermal area, California. *Geophysical Research Letters*, *23*(8), 877–880.

Ross, A., Foulger, G. R., & Julian, B. R. (1999). Source processes of industrially-induced earthquakes at The Geysers geothermal area, California. *Geophysics*, *64*(6), 1877–1889.

Sambridge, M., Gallagher, K., Jackson, A., & Rickwood, P. (2006). Trans-dimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, *167*, 528–542.

Saraò, A., Panza, G., Privitera, E., & Cocina, O. (2001). Non-double-couple mechanisms in the seismicity preceding the 1991–1993 Etna volcano eruption. *Geophysical Journal International*, *145*(2), 319–335. https://doi.org/10.1046/j.1365-246X.2001.01375.x

Sens-Schönfelder, C., & Wegler, U. (2006). Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia. *Geophysical Research Letters*, *33*, L21302. https://doi.org/10.1029/2006GL027797

Šílený, J. (2009). Resolution of non-double-couple mechanisms: Simulation of hypocenter mislocation and velocity structure mismodeling. *Bulletin of the Seismological Society of America*, *99*, 385–394.

Šílený, J., Campus, P., & Panza, G. (1996). Seismic moment tensor resolution by waveform inversion of a few local noisy records—I. Synthetic tests. *Geophysical Journal International*, *126*, 605–619.

Sun, W., & Kennett, B. (2016). Receiver structure from teleseisms: Autocorrelation and cross correlation. *Geophysical Research Letters*, *43*, 6234–6242. https://doi.org/10.1002/2016GL069564

Tape, W., & Tape, C. (2013). The classical model for moment tensors. *Geophysical Journal International*, *195*(3), 1701–1720.

Thompson, R. C. (1989). Structural stratigraphy and intrusive rocks at The Geysers geothermal field. *Transactions of the Geothermal Resources Council*, *13*, 481–485.

Tkalčić, H., Dreger, D., Foulger, G., & Julian, B. (2009). The puzzle of the 1996 Bárdarbunga, Iceland, earthquake: No volumetric component in the source mechanism. *Bulletin of the Seismological Society of America*, *99*(5), 3077–3085.

Walsh, J. (1965). The effect of cracks on the compressibility of rock. *Journal of Geophysical Research*, *70*(2), 381–389. https://doi.org/10.1029/JZ070i002p00381

Yagi, Y., & Fukahata, Y. (2011). Introduction of uncertainty of Green's function into waveform inversion for seismic source processes. *Geophysical Journal International*, *186*(2), 711–720. https://doi.org/10.1111/j.1365-246X.2011.05043.x

Young, M., Rawlinson, N., Arroucau, P., Reading, A. M., & Tkalčić, H. (2011). High-frequency ambient noise tomography of southeast Australia: New constraints on Tasmania's tectonic past. *Geophysical Research Letters*, *38*, L13313. https://doi.org/10.1029/2011GL047971

Zucca, J. J., Hutchings, L. J., & Kasameyer, P. W. (1994). Seismic velocity and attenuation structure of The Geysers geothermal field, California. *Geothermics*, *23*(2), 111–126.