CWP-631

Probabilistic micro-earthquake location for reservoir monitoring and characterization

Ran Xuan and Paul Sava
Center for Wave Phenomena, Colorado School of Mines

ABSTRACT
Micro-seismicity is used to monitor fluid migration during reservoir production and hydro-fracturing operations. This is usually done with sparse networks of seismic sensors located in boreholes. The data used for micro-earthquake monitoring are corrupted by noise which reduces the signal-to-noise ratio to values as low as 0.1. Monitoring methods based on traveltime picking of various wave modes (P or S) cannot deal with this level of noise and require extensive user interaction. An alternative class of methods uses time reversal to focus micro-earthquake information at the source position. These methods can handle noisier signals, but are also costlier to run. The technique advocated in this paper exploits time-reversal within the general framework of Bayesian inversion. Given an assumption about the possible locations of micro-earthquakes, we use recorded data to evaluate the feasibility of micro-earthquakes occurring at various locations in the Earth. The method takes into account imaging imperfections due to unknown components of the model or acquisition array aperture. We simulate wavefields corresponding to possible sources distributed in the model and evaluate their match with the wavefield reconstructed from real data recorded in the field. In this regard, the method operates like a pattern-recognition procedure and can exploit a wide variety of techniques designed for this purpose. We use simple cross-correlation to take advantage of the speed and robustness of this technique. The wavefields reconstructed at various locations are used to scan over time the wavefield constructed from field data, thus our method is able to identify not only the position of the micro-earthquakes but also their onset times. The final outcome of this automated process is a map of probabilities indicating the confidence of micro-earthquake occurrence at various positions and times.

Key words: wave equation, imaging

1 INTRODUCTION

High-pressure fluid injected into oil and gas reservoirs causes time-invariant stress and strain changes. When the stress exceeds a threshold characterizing the resistance of rock materials to stress, micro-seismicity is triggered by the release of pressure along pre-existing fractures or through the creation of new fractures (Maxwell & Urbancic, 2001; Duncan, 2005). Precisely locating micro-seismic events can be used to monitor the hydraulic fracturing and for reservoir characterization (Rentsch, 2004).

Most location methods currently used require the identification of seismic arrivals which involves accurate picking of P- and S-wave arrival times. The onset time and the coordinates of the hypocenter of micro-seismic events are given by calculations which require accurate knowledge of the velocity model and of the physical relationships describing wave propagation in the subsurface. The source is located by optimizing a misfit function between measured and calculated quantities (Pujol, 2004; Lay & Wallace, 1995; Thurber & Rabinowitz, 2000). The methods in this category assume that the arrival time of a specific event can be identified on seismic

traces (Gajewski, 2007). However, event identification is not always possible due to the typically low signal-to-noise ratio (SNR) characterizing micro-seismic events. Furthermore, such location methods strongly depend on the picking accuracy of P- and S-wave arrival times and have a low degree of automation (Rentsch, 2007).

In contrast to the methods based on picked arrival times, migration-based methods use the full seismic wavefield to locate micro-seismic events. The main advantage provided by the methods in this category is that they are independent of the picking of specific arrivals and can locate weak events by focusing energy at the source using time-reversal (Gajewski, 2005). The drawback of methods in this category is the high computational cost which limits their practical use. However, this drawback becomes less and less of a problem due to the significant advances in computer hardware, which allow for fast processing of large volumes of data.

Within the context of migration-based methods, we advocate in this paper a new methodology which can be used to automatically identify and locate sources of micro-seismicity. We use time-reversal based on numeric solutions to the acoustic wave-equation to reconstruct seismic wavefields in the space around the possible source position and at times around the possible onset time of seismic events. Ideally, the energy corresponding to a micro-seismic event focuses at some time and some position in space. However, due to the incomplete acquisition, inaccuracies of the velocity model and noise present in data, this focus is hard to identify or may not even form at all.

We address this problem by constructing at various locations in the model and at various times "ideal" wavefields which we would record given the field acquisition geometry and the velocity model in the region under investigation if a micro-earthquake source would be located at that position and time. Then, we employ a pattern matching procedure to compare the reconstructed wavefields based on field data with our simulated wavefields and identify the most probable locations and onset times of micro-seismic events. This process exploits the kinematic properties of the focused wavefields and takes advantage of the distortions caused by the incomplete acquisition in the area under investigation.

Given the uncertainty characterizing the process of micro-earthquake identification, we do not construct our results as precise positions and onset times. Instead, we construct maps of probability indicating the likelihood that a micro-seismic source exists at a given position in space and a given time. We construct our solutions using Bayesian inversion theory which simply states that inversion is a process which refines our state of knowledge. The Bayesian theory discusses about a prior state of knowledge which incorporates our knowledge about our system in absence of any measurements and about a posterior state of knowledge which combines the prior with additional knowledge acquired from new measurements.

We begin the paper with a review of Bayesian inversion theory applied to the problem of micro-earthquake identification and illustrate our technique with synthetic data examples.

2 A METHODOLOGY FOR AUTOMATIC MICRO-SEISMIC LOCATION

2. 1 Bayesian inversion theory

Bayesian inversion theory uses probability density functions as a measurement of a given state of knowledge (Jaynes, 2003). A crucial features of Bayesian theory is that a probability can be assigned to a hypothesis. According to the Bayesian theory, we begin our investigation with a certain state of knowledge which is independent of any subsequent measurement we make. We assign to any possible model acceptable to our prior state of knowledge a likelihood which is quantifies how probable a particular model is to represent reality (Tarantola, 2005). The posterior combines our prior understanding of the model parameters with additional information derived from measurements and assumptions about the physical laws governing the process under investigation. This approach allows us to incorporate any a priori information in the data, parameters and theory. As discussed by Tarantola (2005), the posterior state of information is obtained using the Bayes theorem

$$\sigma(\mathbf{m},\mathbf{d}) = k\frac{\rho(\mathbf{m},\mathbf{d})\Theta(\mathbf{m},\mathbf{d})}{\mu(\mathbf{m},\mathbf{d})}\ , \tag{1}$$

where $\mathbf{m}$ represents the model, $\mathbf{d}$ represents the data, $\rho(\mathbf{m},\mathbf{d})$ represents the a priori state of information, $\Theta(\mathbf{m},\mathbf{d})$ represents the theoretical state of information, $\mu(\mathbf{m},\mathbf{d})$ represents the homogeneous state of information, k represents a normalization constant, and $\sigma(\mathbf{m},\mathbf{d})$ represents the a posteriori state of information.

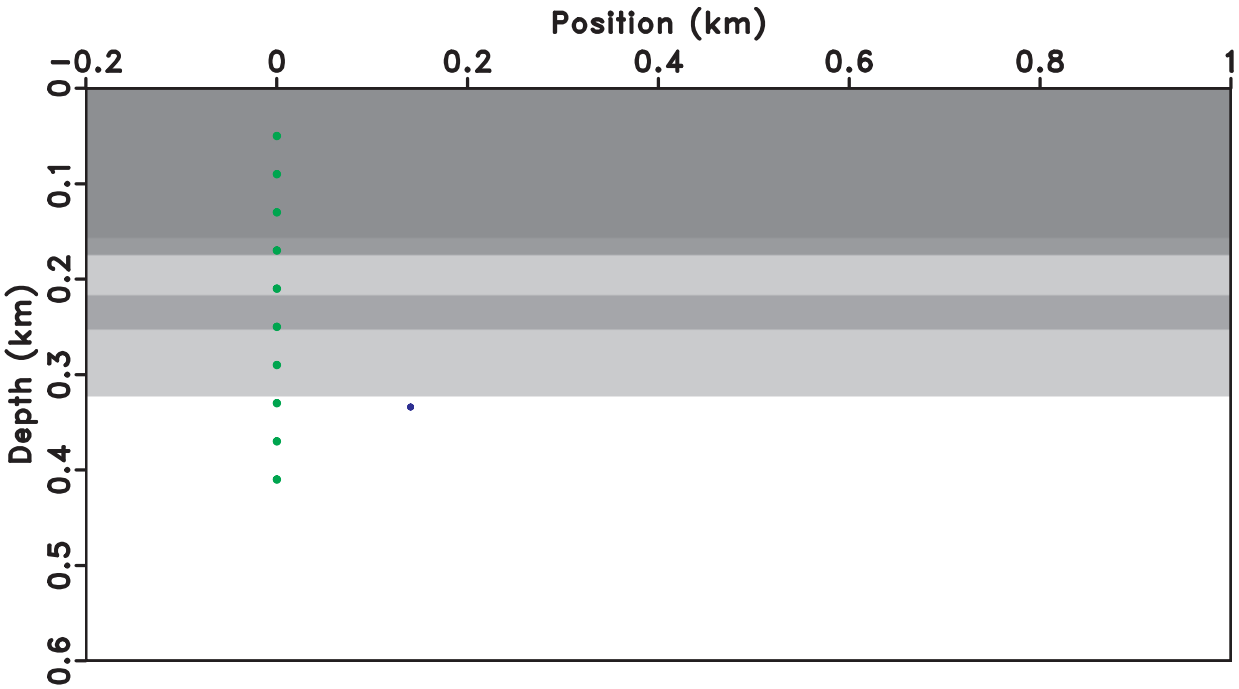
According to equation 1, the a posteriori state of information $\sigma(\mathbf{m},\mathbf{d})$ is computed by combining the theoretical function $\Theta(\mathbf{m},\mathbf{d})$ with the prior information $\rho(\mathbf{m},\mathbf{d})$. The solution to the inversion problem is given by $\sigma_{\mathfrak{M}}(\mathbf{m})$ which represents the posterior marginal probability on the model space and is obtained by integrating the a posteriori probability density function over data:

$$\sigma_{\mathfrak{M}}(\mathbf{m}) = \int_{\mathfrak{D}} \sigma(\mathbf{m},\mathbf{d})d\mathbf{d}\ , \tag{2}$$

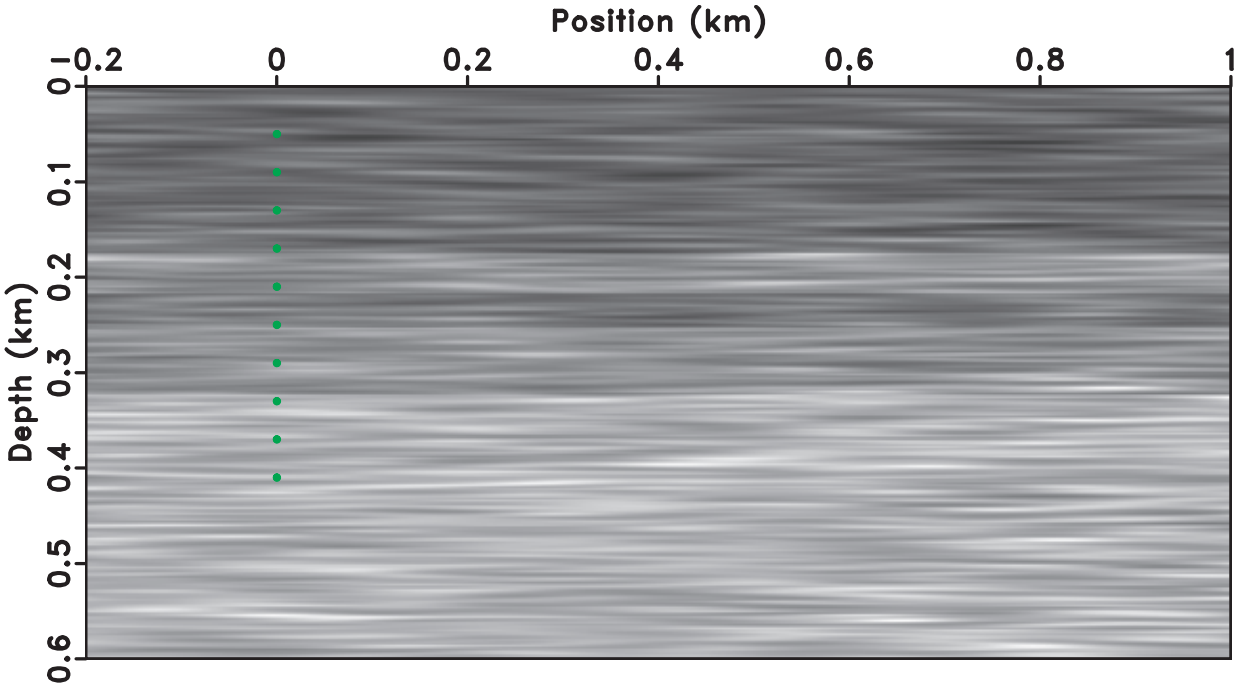
where $\mathfrak{D}$ is the data space. When the uncertainty associated with the kernel function is negligible, $\sigma_{\mathfrak{M}}(\mathbf{m})$ is

$$\sigma_{\mathfrak{M}}(\mathbf{m}) = \nu\rho_{\mathfrak{M}}(\mathbf{m})\rho_{\mathfrak{D}}(G(\mathbf{m}),\mathbf{d}^{obs})\ , \tag{3}$$

where $\rho_{\mathfrak{M}}(\mathbf{m})$ is the prior state of information on the model space, $\rho_{\mathfrak{D}}(\mathbf{d})$ is the prior information on the data



(a)



(b)

Figure 1. (a) The velocity model used for the kernel function and (b) the velocity model with random variability used to simulate data. The random velocity in panel (b) is the sum of the velocity in model (a) with correlated Gaussian noise with a magnitude of 40% of the background. The background velocity model is sampled every 2 m and ranges between 1900 and 2400 m/s. The acquisition array consists of 10 receivers placed at $x = 0$ km with a spacing of 40 m. We simulate an explosive seismic source at $x = 0.136$ km and $z = 0.333$ km.

space, and ν is a normalization constant. A summary of Bayesian inversion theory is presented in Appendix A.

We apply the Bayesian inversion theory to solve the problem of micro-earthquake identification. The locations corresponding to micro-seismic events are provided by incorporating information from recorded seismograms with prior information on the events, e.g. possible locations and onset time of the events, as well as physical relations between the model and data.

The model parameters used for the micro-earthquake location problem are $\mathbf{m} = \{\mathbf{x}, t\}$, where $\mathbf{x} = \{x, y, z\}$ describes the position of a micro-seismic in space and t describes the onset time. The data $\mathbf{d}$ represent the wavefields reconstructed from recorded seismograms in the space around the most probable micro-earthquake source locations and onset times. In this context, it is irrelevant whether the data are acquired in a borehole, on the surface, or using a combination of borehole and surface arrays. The data parameters $D(\mathbf{x}, t)$ are voxels of the 4-D wavefield. Each model has only one datum because there is one wavefield characterizing one micro-seismic event in space and time. The kernel function $G(\mathbf{m})$ describes the theoretical relationship between the data and model. As discussed earlier, we evaluate data for all possible models by taking into consideration the velocity in the region under investigation and the acquisition geometry used for the field experiment. In our case, the kernel function is based on the two-way acoustic wave-equation and is constructed using the following steps:

- First, for a given model ($\mathbf{m}$) characterized by the onset time t and location $\mathbf{x}$, we propagate waves forward from the source using the velocity $c(\mathbf{x})$ according to the two-way acoustic wave-equation

$$\frac{1}{c^2}\ddot{U} = \nabla^2 U + f(\mathbf{m}, t) \;, \tag{4}$$

where $f(\mathbf{m}, t)$ is the considered source wavelet and $U(\mathbf{x}, t)$ represents the seismic wavefield simulated forward in time.

- Second, the forward wavefield is recorded at receiver coordinates $\mathbf{x}_0$ arbitrarily located in space:

$$R(\mathbf{x}_0, t) = U(\mathbf{x} = \mathbf{x}_0, t) \;, \tag{5}$$

where $R(\mathbf{x}_0, t)$ represents the recorded seismogram.

- Third, the recorded seismogram is time reversed and propagated backward in time from the acquisition array to reconstruct the wavefield at all locations in space and all times within the window considered for our prior model. This step is also based on the two-way acoustic wave-equation

$$\frac{1}{c^2}\ddot{D} = \nabla^2 D + R(\mathbf{x}_0, -t) \;, \tag{6}$$

where $D(\mathbf{x}, t)$ are the data parameters as explained earlier.

For a given model, we refer to the data calculated using this procedure as the predicted data D^{pre}. In contrast, we refer to the wavefields reconstructed from field seismograms as the observed data D^{obs}.

We illustrate our method with a 2D synthetic example. We assume that we use the velocity shown in Figure 1(a) as the background velocity to construct the predicted data. We consider the velocity shown in Figure 1(b) as the true velocity which is used to simulate the field seismograms. We simulate data by considering a micro-seismic source at $\mathbf{m}_t = \{x_t, z_t, t_t\}$, where the subscript t indicates the "true" position of the source. The wavefield simulated for this event is calculated using,

$$\frac{1}{v}\ddot{U} = \nabla^2 U + f(\mathbf{m}_t, t) \;, \tag{7}$$

where $v(\mathbf{x})$ is the velocity shown in Figure 1(b). The simulated synthetic seismogram is shown in Figure 2. We assume that the velocity v is unknown during the inversion process.

2. 2 The prior information on the model

We assume that the hypocenter and the onset time of micro-seismic events may be located with equal probability within a certain space domain and time interval. Thus, we represent this prior knowledge on the model space as a boxcar function in space and time. The prior probability $\rho_{\mathfrak{M}}(\mathbf{m}) = 1$ within this region of space and time and 0 elsewhere. We uniformly sample the model space within this region. The number of models located in this space is $n_x \times n_z$, where n_x and n_z represent the numbers of models along the x and z axes, respectively. In our example we sample the model parameters every 4 m, and some of the locations for which we simulate data using our procedure are shown in Figure 3. By selecting a finite number of possible source positions, we both exploit our knowledge about the true source location and reduce computational time. The assumption we make here is that the data (reconstructed wavefield corresponding to the recorded seismograms) characterize with equal probability all models within our selected model space, as illustrated in Figure 1(a).

2. 3 The theoretical information on model and data

For all models within our selected region of space, we compute the corresponding wavefield by the forward/backward procedure outlined in the preceding section. In our case, we assume that each model corresponds to one dataset and that all simulated data are equally probable to explain the recorded data. This construction, which corresponds to the illustration in Figure 1(b), applies the kernel functions G linking theoretically individual members of the model and data spaces.

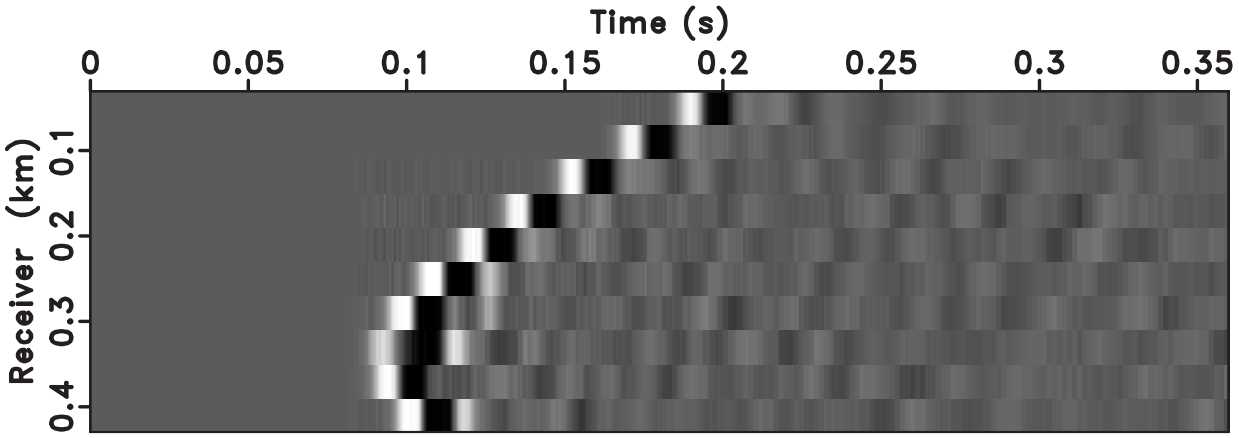


Figure 2. The synthetic seismogram simulated using the velocity model shown in Figure 1(b). We consider an absorbing surface. The source used to simulate this event is a Ricker wavelet with a dominant frequency of 50 Hz. A total of 2000 time steps with a time increment of 0.2 ms are calculated.

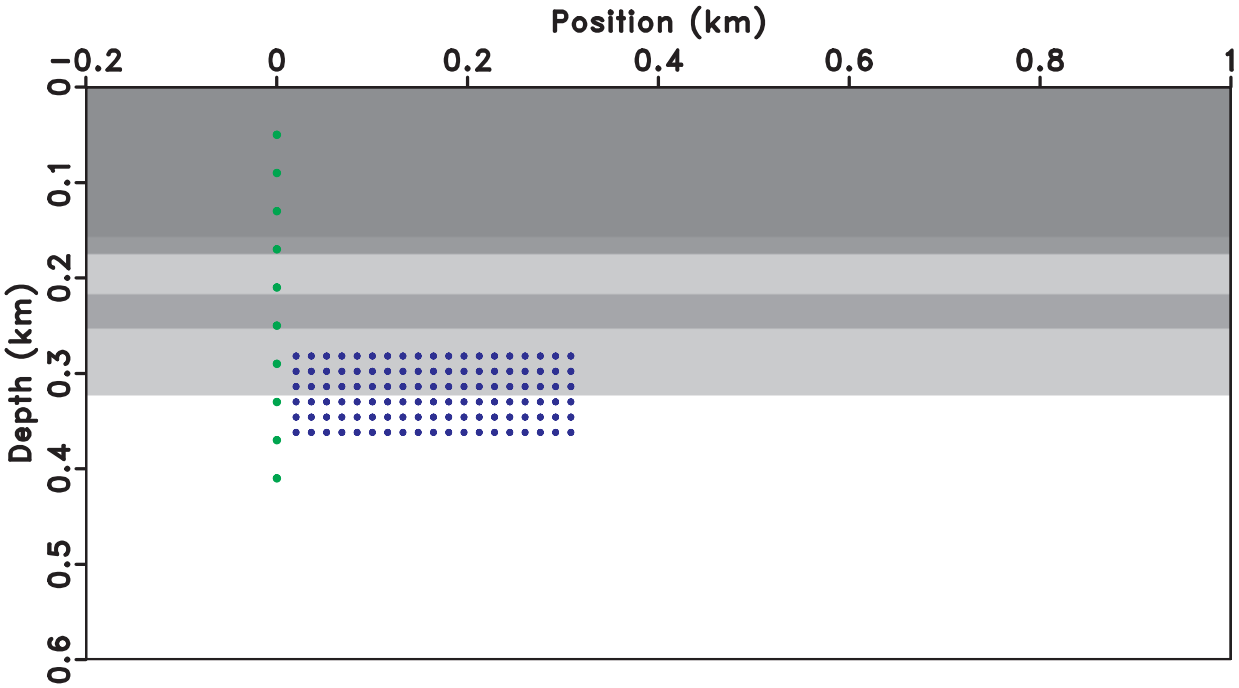


Figure 3. A schematic representation of the a priori information on the model space. We assume that the micro-earthquakes occur in the region indicated by the dense dots between $x = 0 - 0.35$ km and $z = 0.25 - 0.35$ km. A similar discrete set of possible onset times are defined along the time axis.

We assemble all those calculations in a database of wavefields which is used to compare with data reconstructed from the "field" seismograms. Figure 4 illustrates the wavefield database at all locations indicated by dots in Figure 3. The size of the cubes in Figure 4 is two wavelengths in space and two periods in time.

The kernel functions constructed with this procedure are functions of the velocity, the acquisition geometry and the source type. We refer to the predicted data calculated from the models defined by $\rho_{\mathfrak{M}}(\mathbf{m})$ for the fixed velocity model, layout of receivers and source type as the reconstructed wavefield database. This database only needs to be calculated once as long as we keep the subsurface velocity and acquisition geometry fixed.

If two sources are located at the same position but have different onset times, the refocused fields of these two events are the same. It is sufficient to compute data for one onset times and then compare these simulated wavefields with the wavefields reconstructed from field seismographs at various times. Thus we can not only identify the location of the source, but we can also identify the onset time.

In an ideal case, if the sources are completely surrounded by receives, the wavefronts reconstructed by time reversal should focus into a point at the source location and at the onset time. However, typical acquisition arrays are limited to a relatively small region of space. In the example shown in Figure 3, 10 receivers are placed along a borehole located at $x = 0$ km. For limited aperture of receivers, the wavefronts do not focus at the source location.

As illustrated in Figure 5, the cross-range resolution C_r of the refocused field at $t = 0$ s increases when the range decreases, but the range resolution does not change much. The relationship between the cross-range resolution C_r and the range L is $C_r = \lambda_0 L/a$, where λ_0 is the wavelength and a is the array aperture (Born & Wolf, 2003). The wavefront shape at $t = 0$ s is controlled by the range and the cross-range for a fixed array aperture and wavelength. It is apparent not only that the focus is incomplete, but that the shape of the reconstructed wavefields changes in space corresponding to the relative positions to the acquisition array as well as the velocity distribution in space.

2. 4 The prior information on the data

As discussed earlier, we refer to data as reconstructed wavefields at all locations in space and within a given time window. We loop over the models provided by $\rho_{\mathfrak{M}}(\mathbf{m})$ and look for the times and locations in space where data reconstructed from recorded seismograms match the simulated data at locations where we assume that micro-seismic events are located. We can conceptually describe the process as measuring the similarity between multi-dimensional images. Various methods can be used to compare the similarity between the predicted data and the observed data. These methods include cross-correlation, sum-of-squared differences, and ratio image uniformity, which are three of the most common area-based methods used in medical image registration to compare the images of the same modality (Zitova & Flusser, 2003). In this work, we use normalized cross-correlation (N_{cc}) to measure the similarity between theoretical wavefields and wavefields reconstructed from field seismograms. By definition, we compute normalized cross-correlation by the expression

$$N_{cc}(\mathbf{m}) = \frac{\sum_{\mathbf{x}-M_{\mathbf{x}}}^{\mathbf{x}+M_{\mathbf{x}}} \sum_{t-M_t}^{t+M_t} (D^{obs} - \overline{D^{obs}})(G(\mathbf{m}) - \overline{G(\mathbf{m})})}{\sqrt{\sum_{\mathbf{x}-M_{\mathbf{x}}}^{\mathbf{x}+M_{\mathbf{x}}} \sum_{t-M_t}^{t+M_t} (D^{obs} - \overline{D^{obs}})^2 \sum_{\mathbf{x}-M_{\mathbf{x}}}^{\mathbf{x}+M_{\mathbf{x}}} \sum_{k=t-M_t}^{t+M_t} (G(\mathbf{m}) - \overline{G(\mathbf{m})})^2}}, \tag{8}$$

where the overline denotes average values, D^{obs} are the observed data, and $G(\mathbf{m})$ are the predicted data. The size of the averaging window is $(2M_{\mathbf{x}}+1)\times\times(2M_t+1)$, where $\mathbf{x} = \{x, y, z\}$. The normalized cross-correlation is evaluated at the locations of every model $\mathbf{m}$. The values of N_{cc} are between -1 and $+1$. Using this equation, we pinpoint the model which generates the highest N_{cc}. In terms of Bayesian inversion, this process of identification of the most probable model corresponds to the conjunction of the prior and theoretical probabilities, as indicated in Figures 1(c)-1(d). The size of the window used in the weighting function affects the value of $N_{cc}(\mathbf{m})$ and its identification power.

Four snapshots of the wavefield reconstructed from the seismogram given by Figure 2 are shown in Figures 6(a)-6(d).

The reconstructed wavefields are symmetric about to the borehole axis, which creates an ambiguity in the identification of micro-earthquake sources. In general, eliminating this ambiguity required additional azimuthal coverage, e.g. from additional boreholes or from the surface. In our synthetic example, we eliminate this ambiguity by the choice of the prior information.

We assume that there is no error associated with the theoretical information, but that errors are introduced by the measurement uncertainty of the observed data. We also assume that the error is distributed in a Gaussian distribution and we give the highest probability to the model which generates the smallest data misfit. The probability density function based on the normalized cross-correlation given by equation 8 is

$$\sigma_{\mathfrak{M}}(\mathbf{m}) = \rho_{\mathfrak{M}}(\mathbf{m}) e^{-\frac{(1-Ncc)^2}{2C_d}}, \tag{9}$$

where C_d is the measurement uncertainty. The larger the value of C_d, the flatter the distribution of $\sigma_{\mathfrak{M}}(\mathbf{m})$. When C_d is small, $\sigma_{\mathfrak{M}}(\mathbf{m})$ tends toward a delta function. We expect that for the distribution of $\sigma_{\mathfrak{M}}(\mathbf{m})$, the radius of the contour with a value of 0.9 is half the wavelength because the location error of the N_{cc} method is

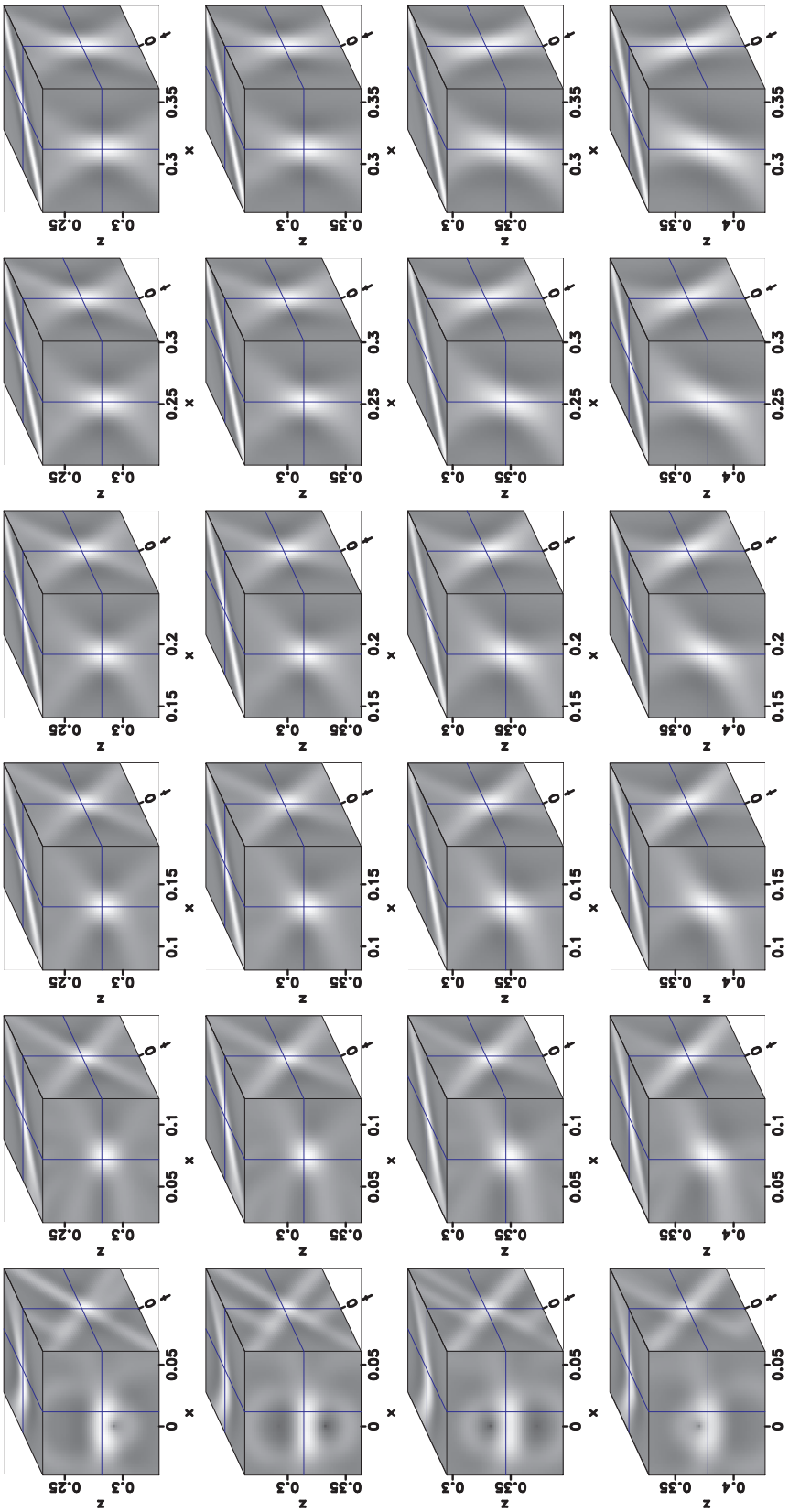


Figure 4. Seismic wavefield database. The center of each cube is the source location of the predicted data.

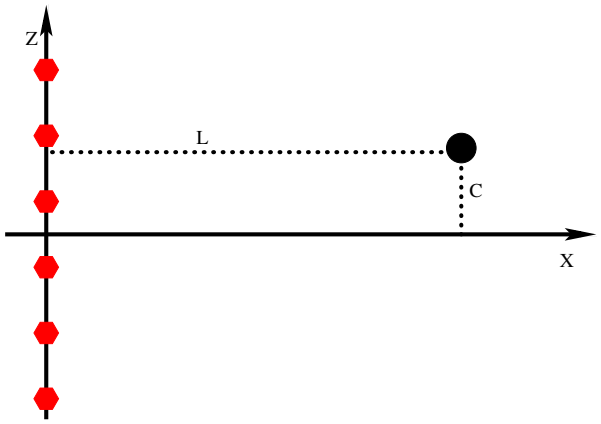


Figure 5. Schematic of the range and cross-range parameters characterizing the resolution of the refocused field. The regular polygons represent receivers and the circle represents the source location. The shortest distance between the source and the array is called range L. The shortest distance between the source and the line which is perpendicular to the center of the array is called cross-range C.

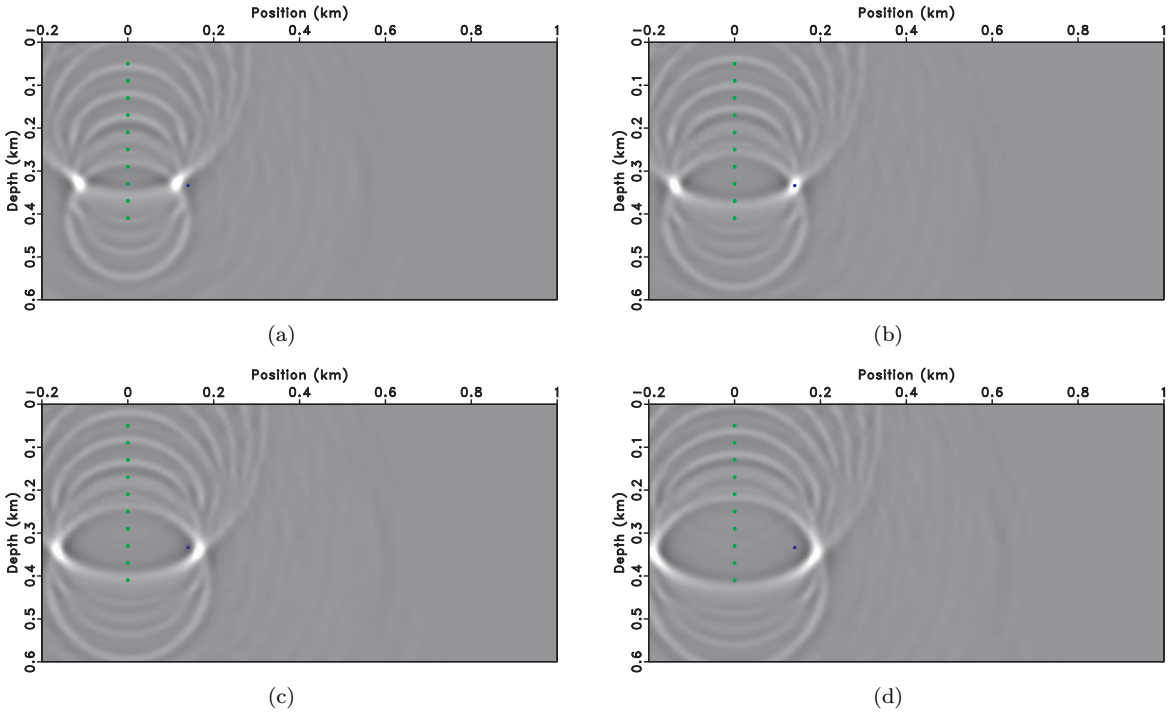


Figure 6. Four snapshots of the observed data. The wave field defocuses when it moves away from the source location. It is hard to choose between (a) and (b) by manual picking.

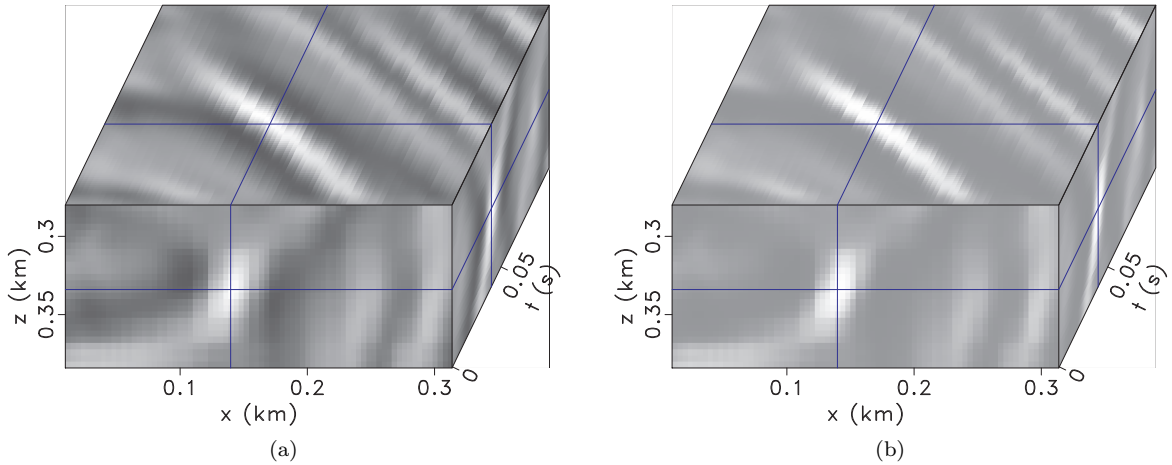


Figure 7. (a) Normalized cross-correlation. (b) Probability distribution based on the normalized cross-correlation value. The measurement uncertainty controls the width of the probability distribution around the maximum value.

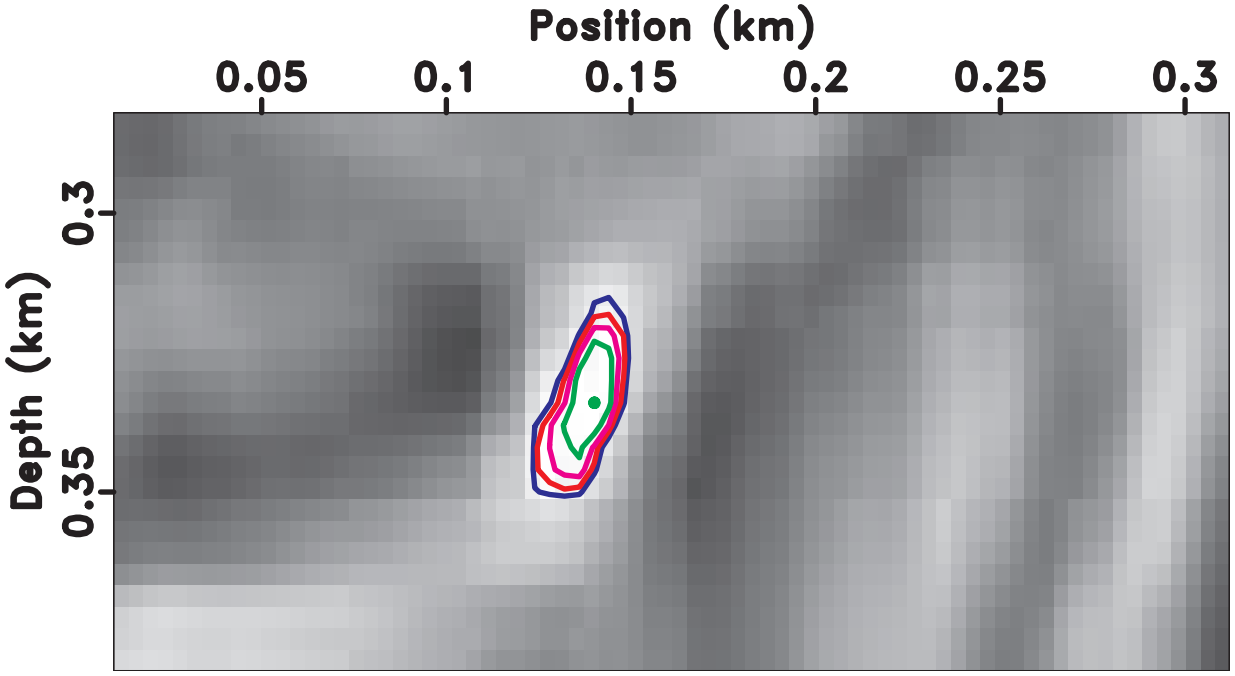


Figure 8. One snapshot of the probability density function, where the dot is the hypocenter of the micro-earthquake. The outer contour has probability of 0.8 and the contour interval is 0.05.

less than half of the wavelength. Figures 7(a)-7(b) show a comparison between the normalized cross-correlation and the associated probability given by equation 9. One snapshot of Figure 7(b) is shown in Figure 8 and the radius of the contour with a value of 0.9 is about 20 m. We can conclude that a micro-seismic event is located within this contour with a probability of 0.9. We compute normalized cross-correlation by comparing the pre-calculated wavefield database with the observed data, as shown in Figure 7(a). The maximum N_{cc} value is 0.991 at $x = 132$ m, $z = 334$ m, $t = 0.044$ s. However, the N_{cc} value is 0.985 at $x = 136$ m, $z = 334$ m, $t = 0.042$ s. These two values are very close to each other and it would be unwise to pick one without taking the other into consideration. Therefore, instead of selecting a single model as the answer, we construct a probability distribution characterizing various models and their likelihood to represent sources of micro-earthquakes.

3 SYNTHETIC EXAMPLE

In the preceding section, we demonstrate micro-earthquake location in the case when only one seismic event occurs in a region during a time interval. However, a more common situation is that multiple micro-seismic events occur more-or-less at the same time and overlap in space during a given time interval. In this section we demonstrate our methodology on another example with multiple overlapping events. The motivation for such an experiment derives from field observations about how fractures propagate during fluid injection. By spatially mapping the signals emitted by micro-earthquakes, we can monitor the extensions of fractures (Keppler et al., 1982; Shapiro et al., 2006). Numerous field observations indicate that the seismic zone surrounds the fractures and forms an elongated pattern (Phillipsa, 1998; Rutledge & Phillips, 2001; House, 1987).

In this numeric experiment, we assume that the micro-seismic events are distributed along a fracture located at $z = 0.326$ km, as indicated in Figure 9. The opening of this fracture simulates high-pressure steam injected into the borehole located at $x = 0$ km. The triggering time of micro-seismic event varies linearly with the distance from the injection point. The differences between the onset times of nearby events is greater than one period. For simplicity, we assume that the magnitude of all micro-earthquakes is the same, and that the sources are characterized by Ricker wavelets with a dominant frequency of 50 Hz. Figure 10 shows the simulated seismogram characterizing all the events. The seismogram is contaminated with two types of noise. The first type of noise is due to model heterogeneities which we assume that we do not know in the imaging process. The second type of noise is corresponding to the experiment being run in a noisy environment. The frequency band of the random noise is $0-140$ Hz overlapping with the micro-earthquake band. The SNR for this example is as low as 0.1. We assume that measurement uncertainties are 0.64 because of the low SNR. We simulate 20 micro-earthquakes and attempt to locate all of them with our methodology, as seen in Figures 11(a)-11(h). The contours shown indicate the probability of earthquakes occurrence at various times and range from 0.8 to 0.95 with a contour interval of 0.05.

Figure 11(a) shows the results obtained for a micro-earthquake occurring at $\mathbf{m}_0 = \{x, z, t\} = \{0.012$ km, 0.326 km, 0.02 s$\}$. This point is included in the contour corresponding to 0.95 confidence, indicating the success of our procedure. Figure 11(b) shows two events, $\mathbf{m}_4 = \{0.04$ km, 0.314 km, 0.16 s$\}$ and $\mathbf{m}_5 = \{0.052$ km, 0.326 km, 0.24 s$\}$ located within the contour with a confidence level of 0.9. This example shows that our procedure may not have the resolution to distinguish among events that are located too close to one-another. The true location of the second events is indicated by the contours shown in Figure 11(c). However, this apparent mismatch is not a big problem because both regions indicate the presence of micro-earthquakes and they overlap in space and time. Other choice of parameters, e.g. in the definition of the wavefield database, might increase the spatial resolution of the identification. This example also seems to indicate the presence of another event located inside the contour centered at about $\{x = 0.17$ km, $z = 0.35$ km$\}$. This is simply an artifact since no event occurs at this position and time in our simulation. Figure 11(d) shows the event $\mathbf{m}_8 = \{0.072$ km, 0.33 km, 0.32 s$\}$ located within the contour with a confidence value of 0.9.

Figure 11(e) shows that the event at $\mathbf{m}_{12} = \{0.112$ km, 0.33 km, 0.52 s$\}$ is not located within the contours obtained by our inversion. In this case, our method fails. This phenomenon is likely due to the fact that multiple micro-earthquakes overlap in this region creating false foci that masquerade as real events. Figure 11(f) shows the event $\mathbf{m}_{14} = \{0.132$ km, 0.326 km, 0.62 s$\}$ located within the contour with a confidence value of 0.9. Figure 11(g) shows the event at $\mathbf{m}_{15} = \{0.144$ km, 0.334 km, 0.68 s$\}$ located within the contour with a confidence value of 0.8. Although we successfully identify this event and the experiment seems to place it at the right position, we have lower confidence in its position. The lower confidence is likely due to the low SNR in the seismogram, which is approximately 0.1. Finally, Figure 11(h) shows the event at $\mathbf{m}_{19} = \{0.212$ km, 0.33 km, 1.02 s$\}$ located well within the contour with a confidence value of 0.95. This is another example of successful identification of a seismic source.

Although in this example we do not locate all existing micro-earthquakes, we regard this experiment as a success because we are getting a good representation of the direction and speed of fracture propagation in a difficult setting characterized by low SNR and sparse data acquisition.

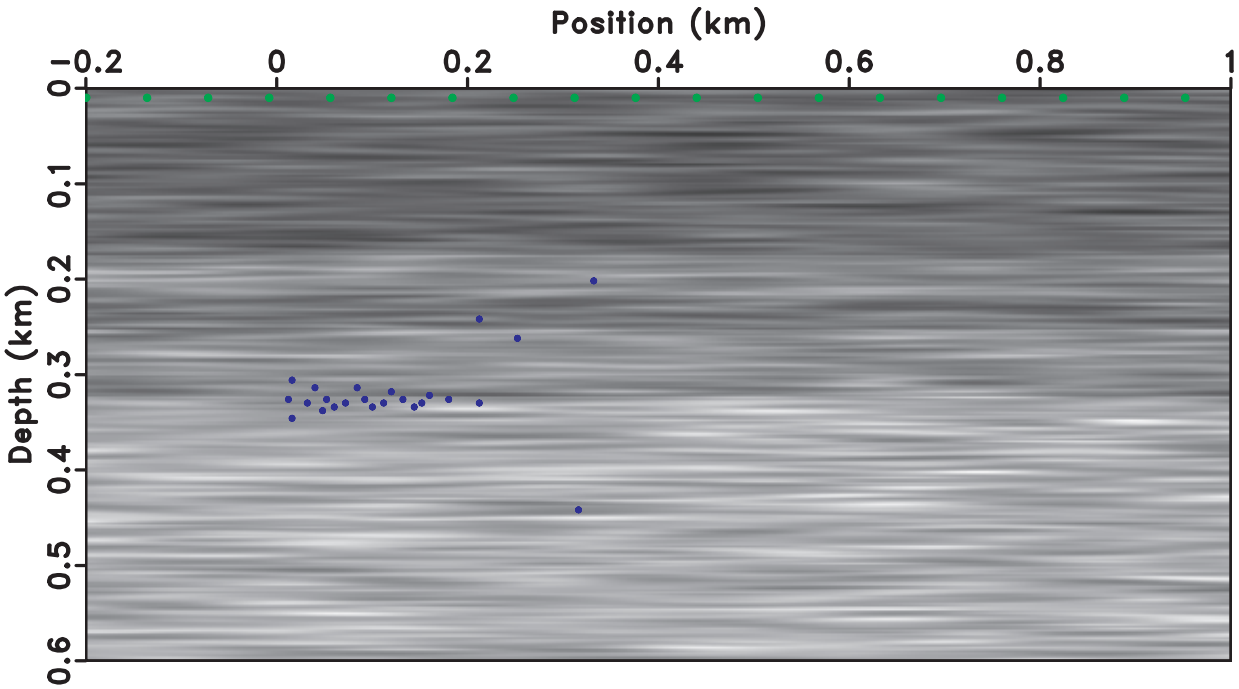


Figure 9. Synthetic model. The dots located at about $z = 0$ km represent receivers, the other dots represent the micro-seismic events.

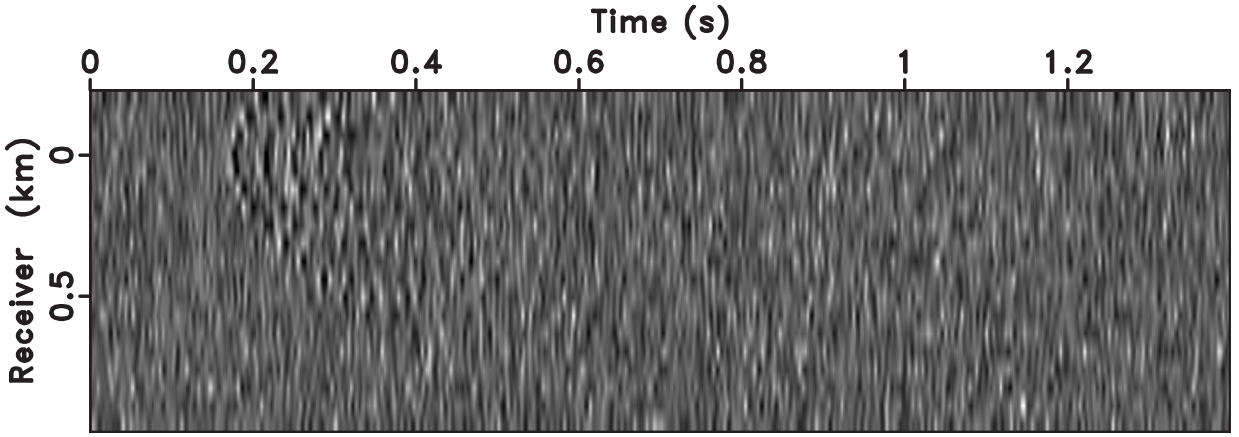


Figure 10. Synthetic seismogram. A total of 7000 time steps with an increment of 0.2 ms are calculated.

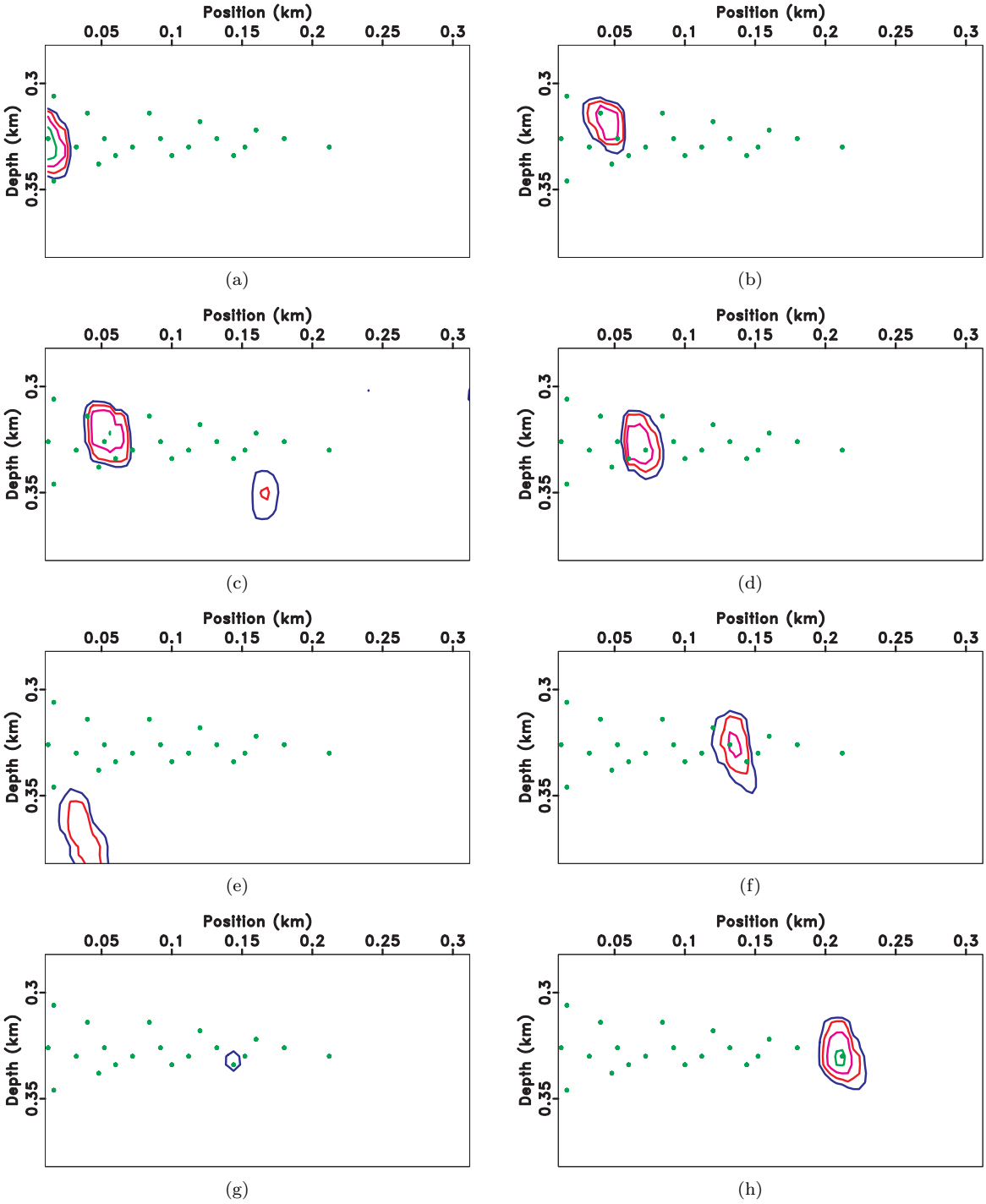


Figure 11. Snapshots of the probability distribution of micro-seismic events. The value of the outer contour is 0.8, the contour interval is 0.05. The dots represent the locations of micro-earthquakes.

4 CONCLUSION

We present a method for automatic micro-earthquake location using Bayesian inversion theory. Our method exploits the unique shape of wavefields partially refocused at the source position. This method does not require picking of arrival times but relies on wavefield focusing obtained by time reversal. Our method not only identifies the spatial location of the source, but also specifies the onset time of the source. Furthermore, since we are using a probabilistic approach to micro-earthquake location, our method provides confidence maps of micro-earthquake locations which can be used for risk assessment.

Synthetic data examples demonstrate the robustness of the method and its applicability to situations when micro-earthquakes occur in close succession in a small region in space and are recorded in a noisy environment with a SNR as low as 0.1.

5 ACKNOWLEDGMENTS

This work was supported by ExxonMobil URC. We are grateful to Mike Payne, Jie Zhang, Rongrong Lu, Alex Martinez, and Anupama Venkataraman at ExxonMobil URC for access to data, technical discussions and financial support of this project.

REFERENCES

Born, M., and Wolf, E., 2003, Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light: , Cambridge University Press.

Duncan, M., 2005, Is there a future for passive seismic?: First Break, **23**, 111–115.

Gajewski, D., 2005, Reverse modelling for seismic event characterization: Geophys.F.Int, **163**, 276–284.

Gajewski, D., 2007, Localization of seismic events by diffraction stacking: SEG Expanded Abstracts, **26**, 1287–1291.

House, L., 1987, Locating microearthquakes induced by hydraulic fracturing in crystalline rock: Geophysical Research Letters, **14**, no. 9, 919–921.

Jaynes, E., 2003, Probability theory: The logic of science: , Cambridge University Press.

Keppler, H., Pearson, C., Potter, R., and Albright, J., 1982, Microearthquakes induced during hydraulic fracturing at the Fenton Hill HDR site: the 1982 experiments: 1983 Geothermal Resources Council's annual meeting, pages 1–9.

Lay, T., and Wallace, T., 1995, Modern global seismology: , Academic Press.

Maxwell, S., and Urbancic, T., 2001, The role of passive microseismic monitoring in the instrumented oil field: The Leading Edge, pages 636–639.

Mosegaard, K., and Tarantola, A., 2002, Probabilistic approach to inverse problems: , Academic Press.

Phillipsa, W., 1998, Induced microearthquake patterns and oil-producing fracture systems in the Austin chalk: Tectonophysics, **289**, 153–169.

Pujol, J., 2004, Earthquake location tutorial: Graphical approach and approximate epicentral location techniques: Seismological Research Letters, **75**, 63–74.

Rentsch, S., 2004, Location of seismicity using gaussian beam type migration: SEG Expanded Abstracts, **23**, 354–357.

Rentsch, S., 2007, Fast location of seismicity:a migration-type approach with application to hydraulic-fracturing data: Geophysics, **72**, no. 1, S33–S40.

Rutledge, J., and Phillips, W., 2001, Hydraulic stimulation of natural fractures as revealed by induced microearthquakes, carthage cotton valley gas field, east texas: Geophysics, pages 1–37.

Shapiro, S., Dinske, C., and Rothert, E., 2006, Hydraulic-fracturing controlled dynamics of microseismic clouds: Geophysical Research Letter, **33**, no. 14, 1–5.

Tarantola, A., 2005, Inverse Problem Theory and Methods for Model Parameter Estimation: , Society for Industrial and Applied Mathematics.

Thurber, C., and Rabinowitz, N., 2000, Advances in seismic event location: , Modern Approaches in Geophysics.

Zitova, B., and Flusser, J., 2003, Image registration methods: a survey: Image and Vision Computing, **21**, 977–1000.

6 APPENDIX A

This appendix summarizes the main elements of Bayesian inversion theory. Bayesian inversion theory characterizes our state of knowledge in a probabilistic manner through the use of probability density functions (PDFs) linking model and data parameters (Jaynes, 2003). We can define two states of knowledge. The *prior* characterizes our knowledge about the model and data parameters before any measurements are taken. This information is based on our general knowledge of the phenomenon under investigation and on the distribution of model parameters and assumptions about reliability of our measurements. The *posterior* characterizes our knowledge about the model and data parameters after measurements are made and data are processed and interpreted. Both states of knowledge are characterized by PDFs linking model and data parameters. Ideally, the measurements help refine the prior into the posterior state of knowledge which provides a tighter connection between model and data parameters. In describing Bayesian inversion theory, we need to discuss the fol-

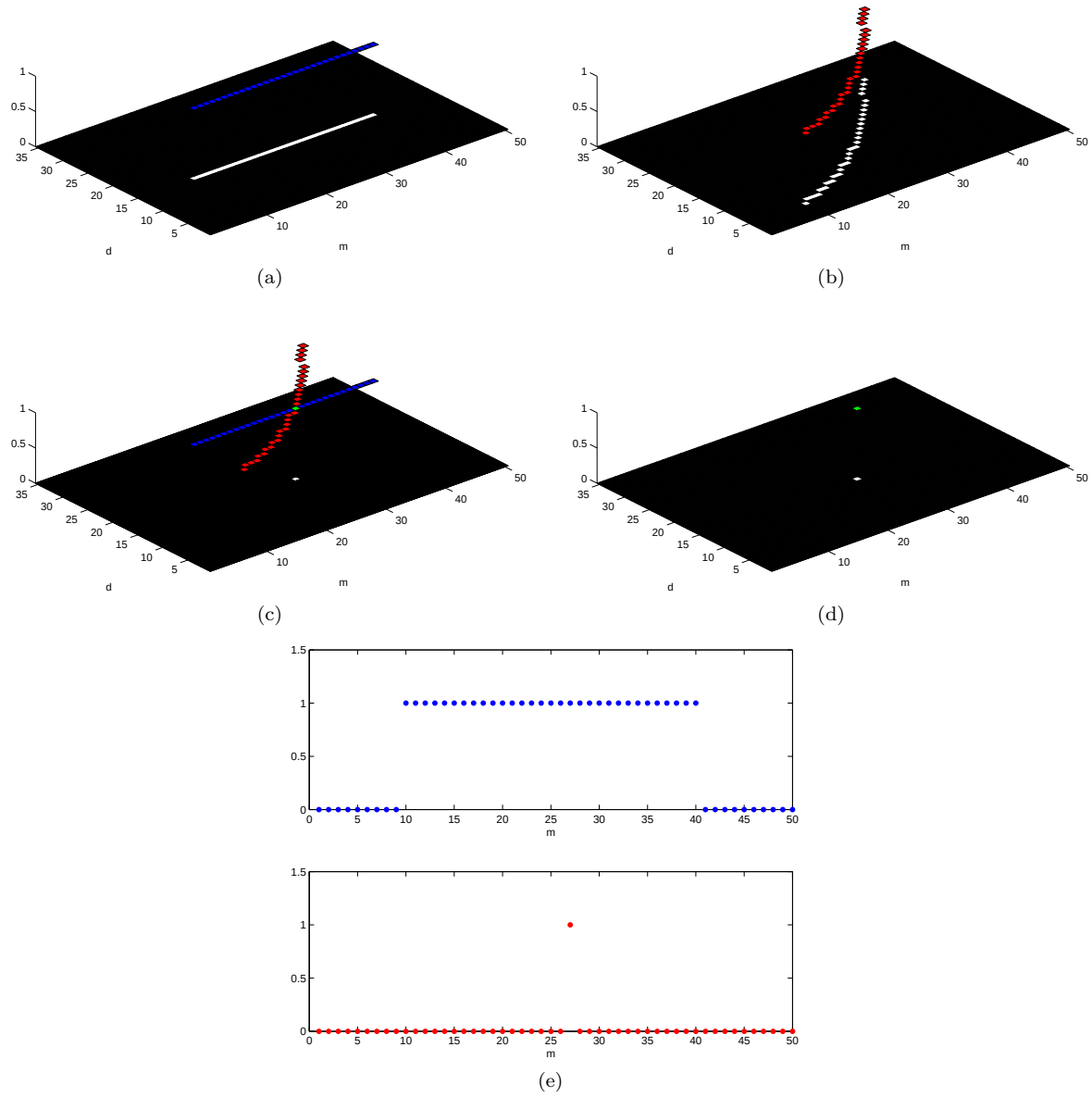


Figure A-1. (a) A priori probability density function $\rho(\mathbf{m}, \mathbf{d})$ represented by a delta function corresponding to $\rho_{\mathfrak{M}}(\mathbf{m})$ which is also represented by a step function. (b) Theoretical probability density function $\Theta(\mathbf{m}, \mathbf{d})$ representing the physical relationship given by $G(\mathbf{m})$ and assumed to be exact. (c) Conjunction between $\rho(\mathbf{m}, \mathbf{d})$ and $\Theta(\mathbf{m}, \mathbf{d})$. (d) A posteriori probability density function $\sigma(\mathbf{m}, \mathbf{d})$ indicating that the solution to the Bayesian inversion problem is given by a single model. (e) A comparison between the marginal probabilities showing $\rho_{\mathfrak{M}}(\mathbf{m})$ in the upper panel and $\sigma_{\mathfrak{M}}(\mathbf{m})$ in the lower panel.

lowing elements: model ($\mathbf{m}$) and model space, data ($\mathbf{d}$) and data space, the prior state of knowledge (a priori PDF $\rho(\mathbf{m},\mathbf{d})$), the theoretical relations between model and data (theoretical PDF $\Theta(\mathbf{m},\mathbf{d})$), and the posterior state of knowledge (a posteriori PDF $\sigma(\mathbf{m},\mathbf{d})$).

6. 1 Model and data spaces

The model space ($\mathfrak{M}$) is composed of a set of individual models which can be denoted by $\mathbf{m}_i$ ($i \in \{1,2,...,N\}$). The model space represents all the models that can possibly characterize the physical phenomenon under investigation.

The data space $\mathfrak{D}$ is composed of all instrumental responses to the investigated models which can be denoted by $\mathbf{d}_i$ ($i \in \{1,2,...,M\}$). The data space represents all data that can possibly be recorded and characterize the physical phenomenon under investigation.

6. 2 A priori probability density $\rho(\mathbf{m},\mathbf{d})$

We describe $\rho(\mathbf{m},\mathbf{d})$ by a joint PDF over the model and data spaces. This prior information is independent of any measurement. Furthermore, in the special case when the prior information on the model space $\rho_{\mathfrak{M}}(\mathbf{m})$ is independent of the prior information on the data space $\rho_{\mathfrak{D}}(\mathbf{d})$, the a priori probability density function $\rho(\mathbf{m},\mathbf{d})$ is (Tarantola, 2005)

$$\rho(\mathbf{m},\mathbf{d}) = \rho_{\mathfrak{D}}(\mathbf{d})\rho_{\mathfrak{M}}(\mathbf{m}) \,, \tag{A-1}$$

where $\rho_{\mathfrak{D}}(\mathbf{d})$ describes the measurement uncertainty of the observed data, and $\rho_{\mathfrak{M}}(\mathbf{m})$ illustrates our confidence in the chosen model parameters.

Two examples of a priori probability density functions are shown in Figures 1(a) and 2(a). The vertical axis represents the value of the probability that a particular model characterizes the physical system under investigation. In Figure 1(a), $\rho(\mathbf{m},\mathbf{d}) = \rho_{\mathfrak{M}}(\mathbf{m})\delta(\mathbf{d} - D^{obs})$, where D^{obs} denote the observed data, and the a priori probability in the model space is a step function. In this case, no error is associated with the observed data. In Figure 2(a), $\rho(\mathbf{m},\mathbf{d}) = \rho_{\mathfrak{M}}(\mathbf{m})e^{-\frac{(\mathbf{d}-D^{obs})^2}{2C_d}}$, where $\rho_{\mathfrak{M}}(\mathbf{m})$ is a tapered step function, and C_d is the measurement uncertainty of the data. We assume that the data are distributed following a Gaussian distribution.

6. 3 Homogeneous probability density $\mu(\mathbf{m},\mathbf{d})$

In order to remove the effect of the discretization of the model and data parameters, we define a homogeneous PDF $\mu(\mathbf{m},\mathbf{d})$ whose role is to balance the discrete model and data spaces. By definition, the homogeneous probability distribution states that a probability assigned to each region of the space is proportional to the volume of the region (Mosegaard & Tarantola, 2002). In the special case when the homogeneous probability densities on the model and data spaces are independent, $\mu(\mathbf{m},\mathbf{d})$ is expressed as

$$\mu(\mathbf{m},\mathbf{d}) = \mu_{\mathfrak{D}}(\mathbf{d})\mu_{\mathfrak{M}}(\mathbf{m}) \,, \tag{A-2}$$

where $\mu_{\mathfrak{M}}(\mathbf{m})$ is the homogeneous probability density on the model space and $\mu_{\mathfrak{D}}(\mathbf{d})$ is the homogeneous probability density on the data space.

For example, in spherical coordinates, the volume of a standard region is $dV(r,\theta,\varphi) = r^2\sin\theta dr d\theta d\varphi$. The homogeneous probability distribution is $\mu = kr^2\sin\theta$, where k is an arbitrary constant. For 3-D Cartesian coordinates, $dV(x,y,z) = dxdydz$ and μ is a constant.

6. 4 Theoretical probability density $\Theta(\mathbf{m},\mathbf{d})$

The theoretical probability density $\Theta(\mathbf{m},\mathbf{d})$ describes the correlations between the model and data parameters. The defined relationship corresponds to a physical law and may incorporate uncertainties associated with theory, for example due to assumptions or simplifications of the physical law or due to imperfect knowledge of the physical parameters underlying it. The theoretical probability density $\Theta(\mathbf{m},\mathbf{d})$ can be expressed as

$$\Theta(\mathbf{m},\mathbf{d}) = \theta(\mathbf{d}|\mathbf{m})\mu_{\mathfrak{M}}(\mathbf{m}) \,, \tag{A-3}$$

where $\theta(\mathbf{d}|\mathbf{m})$ describes the probability distribution of data for the given kernel function and model, and $\mu_{\mathfrak{M}}(\mathbf{m})$ is the marginal probability defined in the preceding section. In Figure 1(b), $\theta(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d} - G(\mathbf{m}))$, where $G(\mathbf{m})$ is the kernel function which describes the physical relationship between $\mathbf{d}$ and $\mathbf{m}$. No error is associated with this kernel function. In Figure 2(b), $\theta(\mathbf{d}|\mathbf{m}) = e^{-\frac{(\mathbf{d}-G(\mathbf{m}))^2}{2C_t}}$, where C_t describes the theory uncertainty. The error associated with this kernel function follows a Gaussian distribution.

6. 5 A posteriori probability density $\sigma(\mathbf{m},\mathbf{d})$

The conjunction between the a priori state of information $\rho(\mathbf{m},\mathbf{d})$ and the theoretical probability density function $\Theta(\mathbf{m},\mathbf{d})$, which describes the theoretical relationship between model and data, provides the a posteriori state of information $\sigma(\mathbf{m},\mathbf{d})$:

$$\sigma(\mathbf{m},\mathbf{d}) = k\frac{\rho(\mathbf{m},\mathbf{d})\Theta(\mathbf{m},\mathbf{d})}{\mu(\mathbf{m},\mathbf{d})}, \tag{A-4}$$

where k is a normalization constant that serves the purpose of keeping the area under the graph of $\sigma(\mathbf{m},\mathbf{d})$ constant. The expression for k is

$$k = \frac{1}{\int_{\mathfrak{X}} d\mathfrak{X}\rho(\mathbf{m},\mathbf{d})\Theta(\mathbf{m},\mathbf{d})/\mu(\mathbf{m},\mathbf{d})}, \tag{A-5}$$

where $\mathfrak{X} = (\mathfrak{D},\mathfrak{M})$ is the joint space of the data and the model.

The posterior information $\sigma(\mathbf{m},\mathbf{d})$ is computed

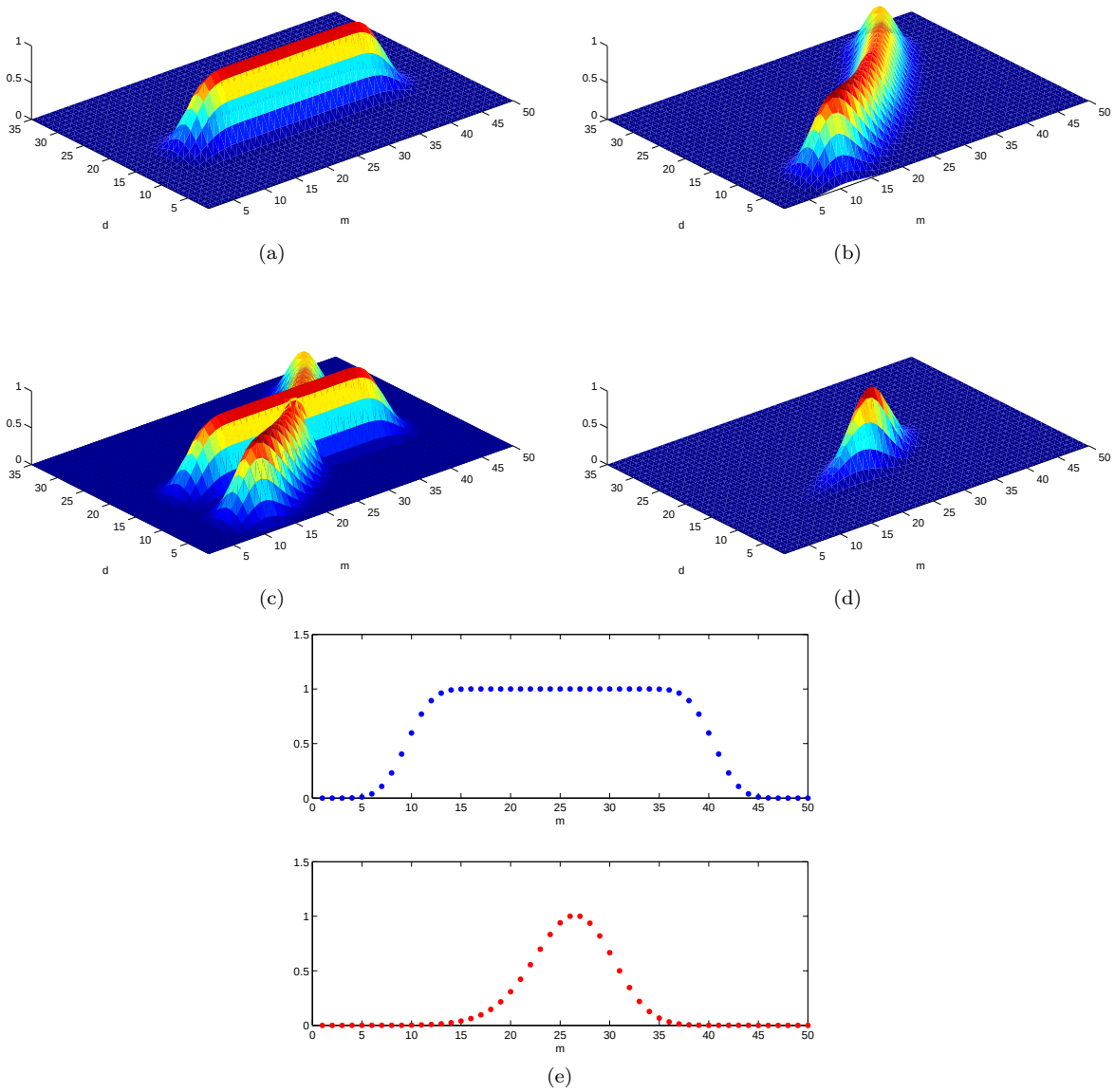


Figure A-2. (a) A priori probability density function $\rho(\mathbf{m}, \mathbf{d})$ represented by a Gaussian function corresponding to $\rho_{\mathfrak{M}}(\mathbf{m})$ represented by a tapered step function. (b) Theoretical probability density function $\Theta(\mathbf{m}, \mathbf{d})$ representing the physical relationship given by $G(\mathbf{m})$ and assumed to be uncertain. (c) Conjunction between $\rho(\mathbf{m}, \mathbf{d})$ and $\Theta(\mathbf{m}, \mathbf{d})$. (d) A posteriori probability density function $\sigma(\mathbf{m}, \mathbf{d})$ indicating that the solution to the Bayesian inversion problem is given by several models corresponding to the uncertainties of the prior and theoretical probability density functions. (e) A comparison between the marginal probabilities showing $\rho_{\mathfrak{M}}(\mathbf{m})$ in the upper panel and $\sigma_{\mathfrak{M}}(\mathbf{m})$ in the lower panel.

based on observations. In the special case when both the model space and the data space are described in Cartesian coordinates, the posterior information is proportional to the conjunction between the prior information and the theoretical information:

$$\sigma(\mathbf{m},\mathbf{d}) \propto \rho(\mathbf{m},\mathbf{d})\Theta(\mathbf{m},\mathbf{d}). \tag{A-6}$$

Figures 1(c) and 2(c) show the conjunction between the $\rho(\mathbf{m},\mathbf{d})$ and $\Theta(\mathbf{m},\mathbf{d})$. In Figure 1(c), $\sigma(\mathbf{m},\mathbf{d}) \propto \delta(D^{obs} - G(\mathbf{m}))$ which is shown in Figure 1(d). Figure 2(c) gives the conjunction between the PDFs shown in Figures 2(a) and 2(b). Its physical meaning is that the observed data are exactly the same as the data calculated from the kernel function. In Figure 2(c), $\sigma(\mathbf{m},\mathbf{d}) \propto e^{-\frac{(D^{obs}-G(\mathbf{m}))^2}{2C}}$, where $C = C_d + C_t$. Figure 2(d) shows $\sigma(\mathbf{m},\mathbf{d})$. The probability distribution in Figure 2(d) shows how well the data calculated from the kernel function explain the observed data.

6. 6 The marginal probability density $\sigma_{\mathfrak{M}}(\mathbf{m})$

In order to provide a solution to the inversion problem, we have to transfer the information provided by $\sigma(\mathbf{m},\mathbf{d})$ to the model space. We obtain the marginal probability density in the model space $\sigma_{\mathfrak{M}}(\mathbf{m})$ by projecting $\sigma(\mathbf{m},\mathbf{d})$ onto the model space:

$$\sigma_{\mathfrak{M}}(\mathbf{m}) = \int_{\mathfrak{D}} \sigma(\mathbf{m},\mathbf{d}) d\mathbf{d} \; . \tag{A-7}$$

The probability density function provided by $\sigma_{\mathfrak{M}}(\mathbf{m})$ indicates what models satisfy at the same time our prior knowledge on the distribution of model parameters, the theoretical relationships between model and data and the measurement uncertainties. A comparison between $\rho_{\mathfrak{M}}(\mathbf{m})$ and $\sigma_{\mathfrak{M}}(\mathbf{m})$ is shown in Figures 1(e) and 2(e). The distribution on $\sigma_{\mathfrak{M}}(\mathbf{m})$ is narrower on the posterior relative to the prior indicating the tighter connection introduced by the recorded data.

6. 7 Negligible theoretical uncertainties

When we assume that no uncertainty is associated with the kernel function $G(\mathbf{m})$ which describes the physical relationship between the data and the model, the theoretical probability density $\Theta(\mathbf{m},\mathbf{d})$ can be expressed as

$$\Theta(\mathbf{m},\mathbf{d}) = \delta(\mathbf{d} - G(\mathbf{m}))\mu_{\mathfrak{M}}(\mathbf{m}) \; , \tag{A-8}$$

where $\mu_{\mathfrak{M}}(\mathbf{m})$ is a constant for models parameterized in Cartesian coordinates. The theoretical probability density $\Theta(\mathbf{m},\mathbf{d}) = 1$ when the data equal the value calculated from the kernel function and zero otherwise.

In this case, the marginal probability in the model space in equation A-7 can be expressed as

$$\sigma_{\mathfrak{M}}(\mathbf{m}) = \int_{\mathfrak{D}} \sigma(\mathbf{m},\mathbf{d}) d\mathbf{d} = k \int_{\mathfrak{D}} \frac{\rho(\mathbf{m},\mathbf{d})\Theta(\mathbf{m},\mathbf{d})}{\mu(\mathbf{m},\mathbf{d})} d\mathbf{d} \; , \tag{A-9}$$

and when $\Theta(\mathbf{m},\mathbf{d})$ is given in equation A-8, with $\rho(\mathbf{m},\mathbf{d}) = \rho_{\mathfrak{D}}(\mathbf{d})\rho_{\mathfrak{M}}(\mathbf{m})$ and $\mu(\mathbf{m},\mathbf{d}) = \mu_{\mathfrak{D}}(\mathbf{d})\mu_{\mathfrak{M}}(\mathbf{m})$, we can write

$$\sigma_{\mathfrak{M}}(\mathbf{m}) = k\rho_{\mathfrak{M}}(\mathbf{m}) \int_{\mathfrak{D}} \frac{\rho_{\mathfrak{D}}(\mathbf{d})\delta(\mathbf{d} - G(\mathbf{m}))}{\mu_{\mathfrak{D}}(\mathbf{d})} d\mathbf{d} \; , \tag{A-10}$$

where $\rho_{\mathfrak{D}}(\mathbf{d})$ is a function of $\mathbf{d}$ and the observed data, and $\mu_{\mathfrak{D}}(\mathbf{d})$ is a constant for data discretized in Cartesian coordinates. Thus, equation A-7 for an ideal kernel function can be written as

$$\sigma_{\mathfrak{M}}(\mathbf{m}) = \nu\rho_{\mathfrak{M}}(\mathbf{m})\rho_{\mathcal{D}}(G(\mathbf{m}), D^{obs}) \; , \tag{A-11}$$

where

$$\nu = \int_{\mathfrak{M}} \rho_{\mathfrak{M}}(\mathbf{m})\rho_{\mathcal{D}}(\mathbf{d}, D^{obs}) d\mathbf{m} \tag{A-12}$$

is a constant. In equation A-11, $\sigma_{\mathfrak{M}}(\mathbf{m})$ is proportional to the product of the prior probabilities on the model and data spaces.