Lawrence Berkeley National Laboratory
LBL Publications

Title
Rapid estimation of earthquake locations using waveform traveltimes

Permalink
https://escholarship.org/uc/item/5f90m3kh

Journal
Geophysical Journal International, 217(3)

ISSN
0956-540X

Authors
Vasco, DW
Nakagawa, Seiji
Petrov, Petr
et al.

Publication Date
2019-06-01

DOI
10.1093/gji/ggz114

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Rapid estimation of earthquake locations using waveform traveltimes

D.W. Vasco , Seiji Nakagawa, Petr Petrov and Greg Newman

Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA. E-mail: dwvasco@lbl.gov

SUMMARY

We introduce a new approach for locating earthquakes using arrival times derived from waveforms. The most costly computational step of the algorithm scales as the number of stations in the active seismographic network. In this approach, a variation on existing grid search methods, a series of full waveform simulations are conducted for all receiver locations, with sources positioned successively at each station. The traveltime field over the region of interest is calculated by applying a phase picking algorithm to the numerical wavefields produced from each simulation. An event is located by subtracting the stored traveltime field from the arrival time at each station. This provides a shifted and time-reversed traveltime field for each station. The shifted and time-reversed fields all approach the origin time of the event at the source location. The mean or median value at the source location thus approximates the event origin time. Measures of dispersion about this mean or median time at each grid point, such as the sample standard error and the average deviation, are minimized at the correct source position. Uncertainty in the event position is provided by the contours of standard error defined over the grid. An application of this technique to a synthetic data set indicates that the approach provides stable locations even when the traveltimes are contaminated by additive random noise containing a significant number of outliers and velocity model errors. It is found that the waveform-based method out-performs one based upon the eikonal equation for a velocity model with rapid spatial variations in properties due to layering. A comparison with conventional location algorithms in both a laboratory and field setting demonstrates that the technique performs at least as well as existing techniques.

Key words: Earthquake location; Earthquakes; Hypocentre; Seismicity

1 INTRODUCTION

Currently, there are a wide range of approaches and algorithms for earthquake location, involving both deterministic and stochastic techniques, many of which are described in the recent paper by Wuestefeld *et al.* (2018) and the earlier comprehensive review by Lomax *et al.* (2009). The most relevant work for this study is grid search methods that aim to find the hypocentre directly (Sambridge & Kennett 1986; Nelson & Vidale 1990; Dreger *et al.*1998; Lomax *et al.*2001; Langet *et al.*2014). In spite of its long history, dating back to the work of Geiger (1910), the topic of earthquake location is still of great importance and the subject of active research.

The accuracy and efficiency are two important issues that must be addressed by any proposed technique if it is to be considered useful. They are particularly pressing in the monitoring of activities related to fluid flow such as hydrofracturing, waste-water injection, geothermal activities and the geological storage of carbon dioxide (Grigoli *et al.*2017). For example, many of these activities take place in areas with complicated velocity structures and are controlled by features such as fracture zones, faults and/or thin high permeability layers that contain abrupt variations in material properties. Errors in such velocity models or errors in modelling the wave propagation from a source to a given receiver can lead to incorrect event locations, even when double-difference techniques are used (Michelini & Lomax 2004). High-frequency asymptotic ray theory, and results based upon the associated eikonal equation, can become inaccurate in the presence of rapid spatial variations in material properties (Vasco 2018; Vasco & Nihei 2019). As a result, some have advocated waveform-based location techniques that go beyond conventional ray theory (Angus *et al.*2014). Such approaches can significantly increase the computation time required for locating events (Grigoli *et al.*2018), even with the use of the one-way wave equation proposed by Angus *et al.* (2014).

Alternative approaches for incorporating more accurate wave propagation methods have invoked time-reversal and back propagation to focus the seismic energy onto the source location (Ishii *et al.*2005; Larmat *et al.*2006, 2008). The same computational issues arise with these approaches though it is possible to incorporate approximations, such as the use of a 2-D Helmholtz equation (Boschi *et al.*2018). Because waveforms can vary significantly from one event to another, a full waveform calculation is required for each event. Furthermore, the use and interpretation of waveforms involves much greater data handling and requires a much more accurate background velocity structure to correctly model the phases following the first arrival.

Here we adopt a location algorithm that allows for the efficient use of waveform modelling methods to calculate traveltimes. As with the grid search method of Nelson & Vidale (1990), the technique utilizes traveltime fields from sources positioned at each station. These traveltime fields are then stored and used in the location of any event within the volume. If the events are similar in size, so that the frequency content does not change significantly, then the traveltimes for each wavefield only need to be calculated once. The approach gains efficiency when the number of events exceeds the number of stations in a given seismic network. As such, it is applicable for networks of any size from a small set of borehole stations to a large global network.

2 METHODOLOGY

In this section we describe the technique and illustrate it using a synthetic set of arrival times. The synthetic test case is based upon a realistic velocity model from the Cranfield site (Lu *et al.*2012) in Mississippi (Fig. 1). The

elastic model was derived from a logged well and was extended to a layered structure and defined over a 3-D grid. Three sets of receivers are used in the test from three simulated wells. The receivers are spaced 50 m apart within each well. The projections of the wells onto an east–west plane are shown in Fig. 1. The structure contains abrupt variations in seismic velocity that are characteristic of many sedimentary formation and is intended to highlight possible differences between a ray-based location method and one based upon waveform calculations. The technique is valid for any given velocity model and does not require any particular receiver geometry. We discuss each of the main steps of the algorithm in the sub-sections that follow.

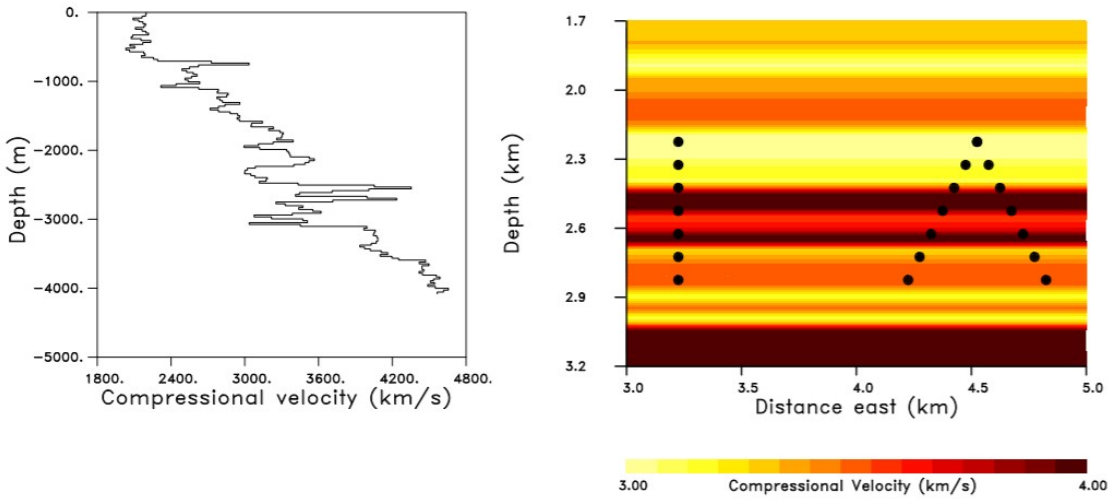

Figure 1. Left-hand panel: compressional velocity variation with depth for the illustrative model containing 150 layers. Right-hand panel: East–west vertical cross-section through the velocity model. The filled circles indicate the projection of the 21 down-hole receivers onto the vertical plane.

2.1 First step: calculate receiver traveltime fields

The location algorithm assumes that some method for computing seismic traveltimes exists for a source at any position within a volume of interest. That is, a field $T(\mathbf{x};\ \mathbf{x}_m)$ is available via some numerical procedure, providing the traveltime to a point $\mathbf{x}$, for a source at point $\mathbf{x}_m$. As in Nelson & Vidale (1990), the calculation may invoke the finite-difference solution of the eikonal equation (Vidale 1988; Podvin & Lecomte 1991; Zelt & Barton 1998; Sethian 1999)

$$\nabla T \cdot \nabla T = \frac{1}{c^2}, \qquad (1)$$

where $c(\mathbf{x})$ describes the seismic velocity variation over the volume of interest. The traveltime field was calculated using the method of Zelt & Barton (1998). A cross-section through an eikonal-based traveltime field, corresponding to the layered test model, is shown in Fig. 2. The test event used to generate the traveltimes, located at (x, y, z) = (3.58 km, 2.41 km, 2.82 km), will be used in this section to illustrate our approach.

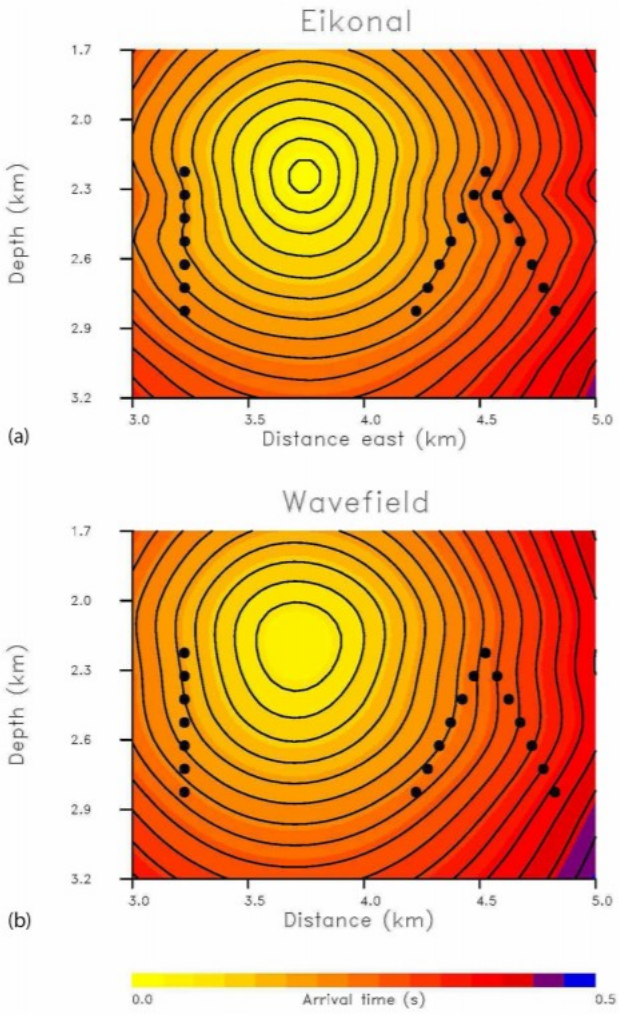

Figure 2. (a) East–west vertical slice through traveltime field computed using the finite-difference solution of the eikonal eq. (1), described in Zelt & Barton (1998). (b) Corresponding vertical cross-section through a traveltime field derived from a finite difference solution of the elastodynamic equation of motion (2).

In this study, the traveltimes come from what is typically a much more time consuming procedure: the numerical solution of the elastodynamic equation of motion (Virieux 1986; Masson *et al.*2006; Masson & Pride 2011; Petrov & Newman 2012)

$$\rho\ddot{\mathbf{u}} = \nabla\left(\lambda\nabla\cdot\mathbf{u}\right) + \nabla\cdot\mu\left[\nabla\mathbf{u} + \left(\nabla\mathbf{u}\right)^{T}\right], \tag{2}$$

where λ(**x**) is the Lame parameter, μ(**x**) is the shear modulus and ρ(**x**) is the density. The general approach involves modelling the entire wavefield and applying a traveltime estimation algorithm. The algorithm used to determine the numerical waveform traveltimes is the same one that is applied to the observed data. This was the procedure we followed in analysing the synthetic test, where a finite-difference code for solving the full elastodynamic wave equation (Masson *et al.*2006; Masson & Pride 2011) was

used to determine the traveltime fields shown in Fig. 2. The finite difference grid spacing of 50 m is identical to that used for solving the eikonal eq. (1). The source function was a sinusoidal function modulated by a Gaussian with a dominant frequency peaked around 10 Hz.

The traveltimes were derived from the radial component of the numerical waveforms by a simple threshold criterion. That is, the traveltime was defined as the instant when the amplitude exceeds 1 per cent of the peak amplitude of the first arriving energy at a given location on the finite difference grid. The spatial variations of the traveltimes for the two methods are compared in Fig. 2. Note that the eikonal-based and waveform-based traveltime estimates differ for this particular velocity model, most likely due to the presence of thin high-velocity layers. As we shall see, traveltimes obtained from full wavefield solutions have some advantages when the velocity model has rapidly varying features, such as the layers shown in Fig. 1. In such models the assumption of smoothly varying elastic properties underlying the eikonal equation no longer hold and its solutions may not be accurate (Vasco & Nihei 2019).

Given a network of M stations, we can consider M traveltime fields, each defined over a 3-D grid. For a source at the location of the m-th station, $\mathbf{x}m$, denote the traveltime field over the grid as $T(\boldsymbol{\jmath}; \mathbf{x}_m)$, where $\boldsymbol{\jmath}$ is a 3-D index vector whose entries range over the nodes of the grid $\mathbf{J}$. For a fixed velocity model, the traveltime field for each station must only be calculated once and may then be used for any number of events. If the velocity model is updated, as during iterative tomographic imaging, then the traveltime fields need to be recalculated after each update.

2.2 Second step: use time-reversal or back-propagation from receivers

Each seismic event will produce a collection of observed arrival times, T_m^o, for some subset of stations in the network denoted by ε. We can invoke reciprocity and time reversal to calculate the possible positions of the incoming wave front at various times in the past. Specifically, for each element $\boldsymbol{\jmath}$ of the grid we subtract the traveltime field $T(\boldsymbol{\jmath}; \mathbf{x}_m)$ from the observed arrival time

$$T_m(\boldsymbol{\jmath}) = T_m^o - T(\boldsymbol{\jmath}; \mathbf{x}_m). \tag{3}$$

All of the back-propagated traveltime fronts, the values of $T_m(\boldsymbol{\jmath})$ for all of the stations that recorded an arrival, will attain roughly the same value at the event location. Random and/or systematic variations, such as those due to noise in the observed arrival times and an incorrect velocity model, will introduce errors into the location estimates.

2.3 Third step: estimate traveltime dispersion over the grid

Because the shifted and back-projected traveltime fields approach the event time at the source location, a measure of central tendency, such as the mean, re-weighted mean or median of the shifted traveltimes (3), will approximate the origin time of the event at the gridpoint nearest to the source. For illustration we will consider the mean value of the back-propagated times, as it has a simple and explicit expression

$$\bar{T}(\boldsymbol{\jmath}) = \frac{1}{N_e} \sum_{m \in \mathcal{E}} T_m(\boldsymbol{\jmath}), \tag{4}$$

where N_e denotes the number of observed arrival times for the given event. However a more robust measure of central tendency, such as the median, is often more desirable in the face of non-Gaussian noise and outliers (Huber 1981). Fig. 3 presents a horizontal slice through the grid $\bar{T}(\boldsymbol{\jmath})$ at the depth of the source, which is indicated by the filled circle. The contour line of 100.0 s, the origin time for this synthetic test, intersects the source location.

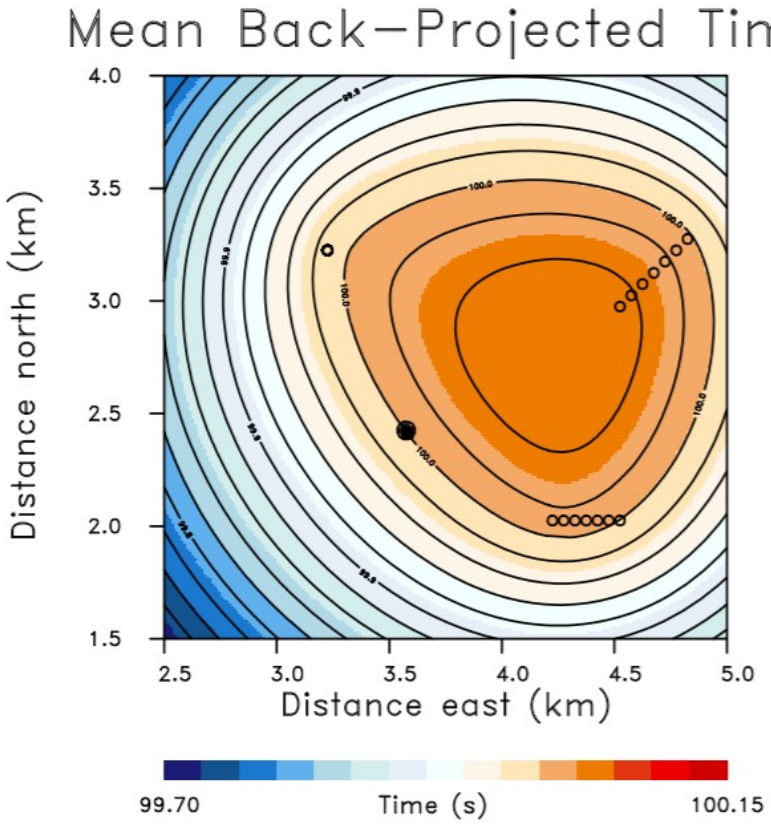

Figure 3. Mean value of the 21 shifted and back projected times from all of the receivers. The vertical projection of the stations are indicated by the open circles. The vertical projection of the event hypocentre is indicated by the filled circle. The horizontal slice through the field lies at 2.82 km, the depth of the event.

At the event location the back-propagated arrival times converge to the origin time and measures of dispersion about this value attain a minimum value over the grid. The sample standard error

$$s(\boldsymbol{\jmath}) = \sqrt{\frac{1}{N_e - 1} \sum_{m \in \mathcal{E}} \left[T_m(\boldsymbol{\jmath}) - \bar{T}(\boldsymbol{\jmath}) \right]^2} \tag{5}$$

or the more robust average deviation

$$a(\boldsymbol{\jmath}) = \frac{1}{N_e - 1} \sum_{m \in \mathcal{E}} \left| T_m(\boldsymbol{\jmath}) - \bar{T}(\boldsymbol{\jmath}) \right|, \tag{6}$$

provide easily computed measures of dispersion at each grid location. The average deviation, $a(\boldsymbol{\jmath})$, the l_1 norm of the residual or error vector, is plotted in Fig. 4 for the eikonal-based and waveform-based location algorithms. The horizontal cross-sections in Fig. 4 are taken at the depth of the event, z = 2.82 km. The contours can be used to assess the degree of uncertainty in the event location.

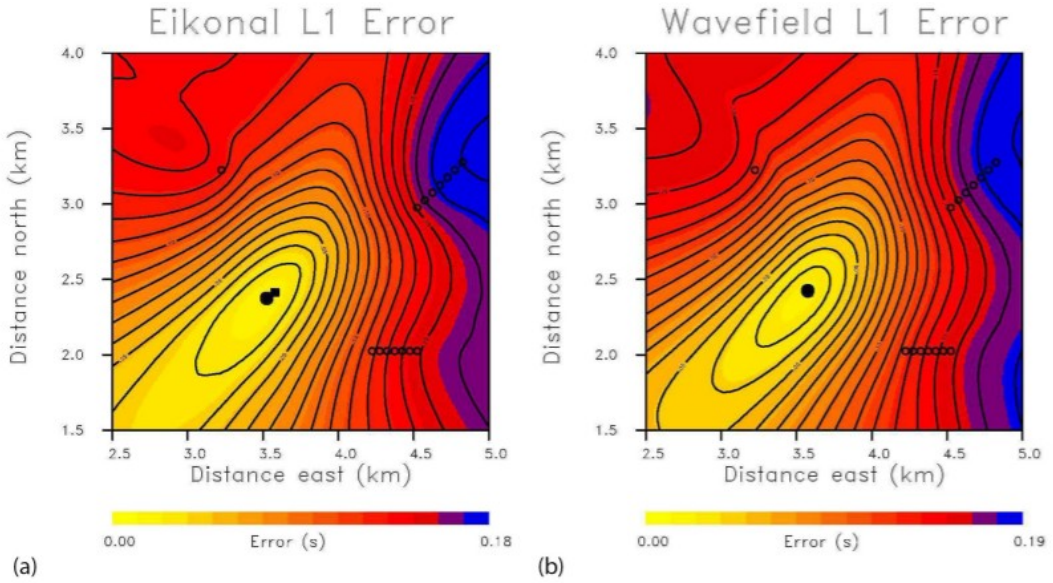

Figure 4. Left-hand panel: Horizontal slice through the 3-D distribution of the average deviation $a(\boldsymbol{\jmath})$. In this case the traveltimes were calculated using the model shown in Fig. 1 and the eikonal eq. (1). Right-hand panel: Horizontal cross-section through the average deviation calculated using traveltimes from a solution of the elastodynamic equation of motion. The actual event locations are denoted by the filled squares while the solutions that minimizes $a(\boldsymbol{\jmath})$ are indicated by the filled circle.

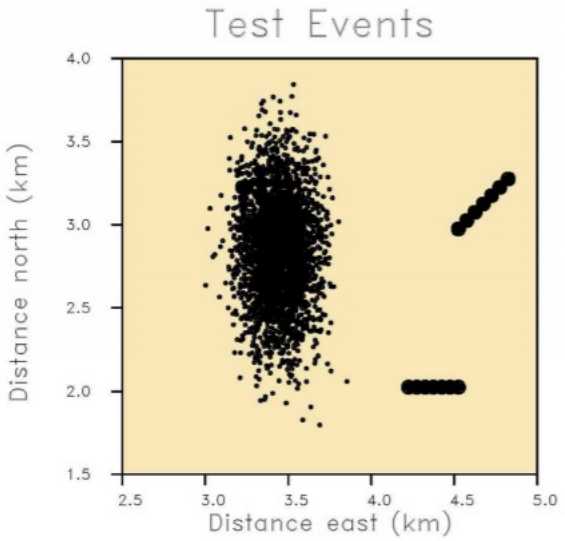

Figure 5. Surface projection of the epicentres of 3000 test events (small open circles) used to study the effects of modelling and random errors on location estimates produced by the new algorithm. The surface projections of the 21 borehole stations are indicated by the filled circles.

Table 1. Computation times.

Approach	Procedure	CPU Time
Eikonal	Traveltime field calculations	3.33 s /Station
	Event locations	0.15 s/Event
Waveform	Traveltime field calculations	814.38 s/Station
	Event locations	0.16 s/Event

2.4 Final step: determine the region where the dispersion is a minimum

The event location is given by the value of JJ at which the statistical dispersion about the mean or median is a minimum, either

$$\min_{\boldsymbol{\jmath}\in\mathbf{J}} s(\boldsymbol{\jmath}) \tag{7}$$

or

$$\min_{\boldsymbol{\jmath}\in\mathbf{J}} a(\boldsymbol{\jmath}) \tag{8}$$

can be used, though the latter is less influenced by non-Gaussian outliers. Because the grid $\mathbf{J}$ is at most 3-D, we can minimize $s(\boldsymbol{\jmath})$ or $a(\boldsymbol{\jmath})$ using a simple grid search. The locations estimated by a grid search for the eikonal- and wavefield-based algorithms are shown in Fig. 4, indicated by the filled circle. The actual location used to generate the synthetic arrival times is denoted

by the filled square. The eikonal-based solution displays a horizontal location error of about 100 m due to the limitations of the high-frequency approximation in this layered model.

The measure of dispersion that is minimized in order to locate the event should be tailored to the statistical properties of the traveltimes. That is, we would like to determine the location that maximizes the likelihood of the observed traveltimes, obtaining a maximum likelihood estimate (James *et al.*2017). The appropriate estimator to use, specifically the correct likelihood function to maximize, depends upon the statistical properties of the errors in the observations (Hoel 1971). For example, the sample standard error $s(\boldsymbol{J})$ given by eq. (5) is an appropriate function when the errors in the observations are from a Gaussian distribution (James *et al.*2017). Hence, it may not be the best function to use because traveltime errors have been studied by several authors and are found to be non-Gaussian (Jeffreys 1932; Bolt 1960; Buland 1984, 1986; Pulliam *et al.*1993). The main issues are that traveltime estimates contain a significant number of outliers due to the picking of the wrong onset for the first arrival. In fact, a chi-square test applied to compressional wave traveltime residuals from the International Seismological Centre (ISC) found that the hypothesis that their cumulative distribution and that of a Gaussian distribution are the same could be rejected with a confidence level of 99 per cens (Pulliam *et al.*1993). A histogram of the traveltime residuals reveals that the tails of the distribution are simply too broad to be represented by a Gaussian (Vasco *et al.*1994).

There are other probability distributions that may be used to model traveltime residuals. For example, it is not possible to reject the hypothesis that the cumulative distribution function of the ISC traveltime residuals and that of a two-sided exponential probability distribution are the same (Pulliam *et al.*1993). The two-sided exponential distribution is given by the probability distribution

$$\Pi\left(|\rho| \leq \frac{\delta\rho}{2}\right) = 2^{-N/2}\sigma^{-N}\exp\left[-\frac{\sqrt{2}}{\sigma}\sum_{i=1}^{N}|\rho_i|\right]\delta\rho, \tag{9}$$

where N is the number of residuals and σ is the standard error associated with the residuals. The distribution (9) has broad tails that can match those typically observed in traveltime residuals. It can be shown that the maximum likelihood estimator associated with residuals containing errors following a two-sided exponential distribution is obtained by minimizing $a(\boldsymbol{J})$, the absolute deviation of the residuals (Menke 2018). Thus, minimizing the average deviation function $a(\boldsymbol{J})$ will usually provide a more reliable location estimate than the minimum of the sample standard error $s(\boldsymbol{J})$. This is the L_1 norm of the traveltime residual vector and its use will reduce the influence of

outliers on the location estimates (Nelson & Vidale 1990; Pulliam *et al.*1993; Shearer 1997; Schweitzer 2001; Shearer 2009).

3 APPLICATIONS

In this section we test the approach for locating events on three different data sets in order to assess the reliability of the algorithm and the utility of waveform arrival times. The first test involves locating events using synthetic traveltimes containing additive two-sided exponential random values. We compare locations that rely upon traveltimes computed using the eikonal equations with our approach based on traveltimes from a full wavefield simulation. The second test involves high-quality traveltimes from a laboratory experiment where the creation of a fracture is monitored. Finally, compressional and shear arrival times from an event associated with the creation of a hydraulic fracture are used to estimate the location and origin time. These data appear to contain several outliers and are representative of arrival times used for locating seismic events.

3.1 A synthetic test with traveltime picking errors

The analysis in this sub-section is a continuation of the Cranfield illustration introduced earlier. In particular, the velocity model shown in Fig. 1 is the basis for calculating traveltimes, using both the eikonal equation and the extraction of arrival times from numerical solutions of the elastodynamic equation of motion. A stochastic algorithm generated 3000 artificial test events arranged around a fracture plane. The probability of an event decreases with distance from the plane, leading to a clustering around the hypothetical fracture (Fig. 5). A finite-difference solution to the elastodynamics equation of motion is used to calculate the traveltimes to seven observation points in each of the three simulated wells (Fig. 5) for a total of 21 stations in the local network. The synthetic traveltimes, with added random noise to simulate errors, are then used in the location algorithm presented above. Because the velocity model does not change, we only need to conduct 21 forward simulations in order to generate the traveltime fields from each observation point to all points in the grid, the quantities $T(\boldsymbol{\jmath},\boldsymbol{x}_m)$ in expression (3). The traveltime fields $T(\boldsymbol{\jmath},\boldsymbol{x}_m)$ are stored and used to compute the mean and median of the backpropagated times. The median is a more robust measure of central tendency than the mean value given in eq. (4). However, because non-Gaussian behaviour is indicated by any significant differences between the median and mean values, it is useful to compute and compare both quantities. The computational requirements of the eikonal and waveform location algorithms are shown in Table 1.

The performance of any practical location algorithm should degrade gradually in the presence of realistic traveltime errors. The two-sided exponential distribution, discussed in the Methodology section, provides a reasonable probability model for generating random deviations with properties similar to traveltime errors. We use this distribution to generate

errors in our test data set. We adopt the approach for generating two-sided exponentially distributed random deviates given in Press *et al.* (1992). Such a distribution can simulate the large deviations that are due to picking an incorrect first pulse. Initially, only small errors, with magnitudes of 1 percent of the peak variability of the traveltimes, were added to the calculated times prior to the location. The peak variability is the largest deviation from the average of the arrival times. The distribution of location errors incurred by the algorithms using the eikonal and wavefield methods to compute traveltimes are shown in Fig. 6. The approach based upon traveltimes calculated using wavefield simulations and arrival time extraction from the wavefields produces mislocation errors that peak at 50 m (Fig. 6, right-hand panel), the spacing of the simulation grid. The number of mislocation errors larger than 50 m drops rapidly and very few exceed 100 m. Because the eikonal equation is an approximation that assumes a smoothly varying model, and the velocity model contains many fine layers with abrupt boundaries (Fig. 1), there may be considerable modelling error for this case. This suggestion is supported by the fact that the location errors associated with an eikonal equation based approach peak at 100 m, twice the error of the wavefield-based locations, and that the tail of the distribution extends to errors of several hundred meters (Fig. 6).

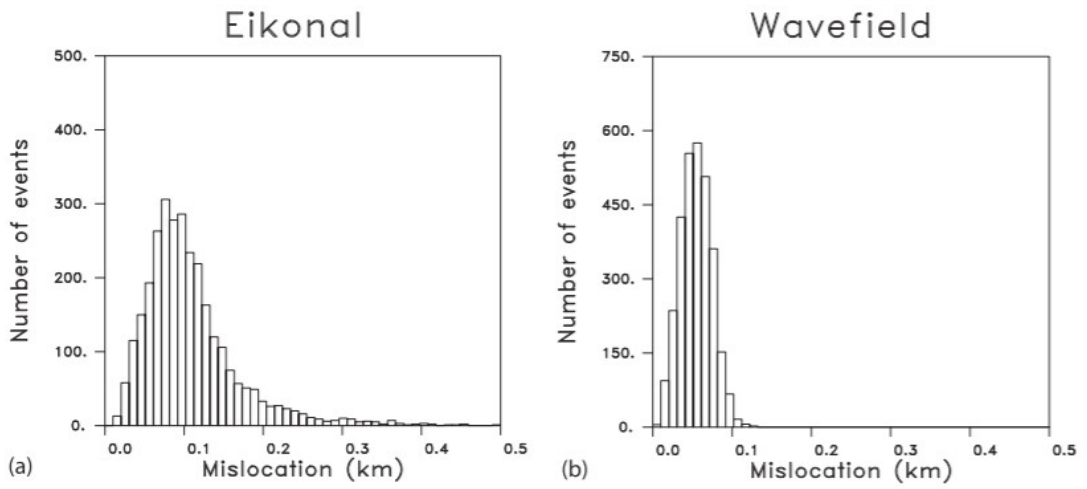

Figure 6. Left-hand panel: Histogram of the location errors associated with the location algorithm and the use of traveltime estimates based upon the eikonal equation. Right-hand panel: Histogram of the location errors that resulted from applying the new location algorithm and using wavefield traveltime estimates.

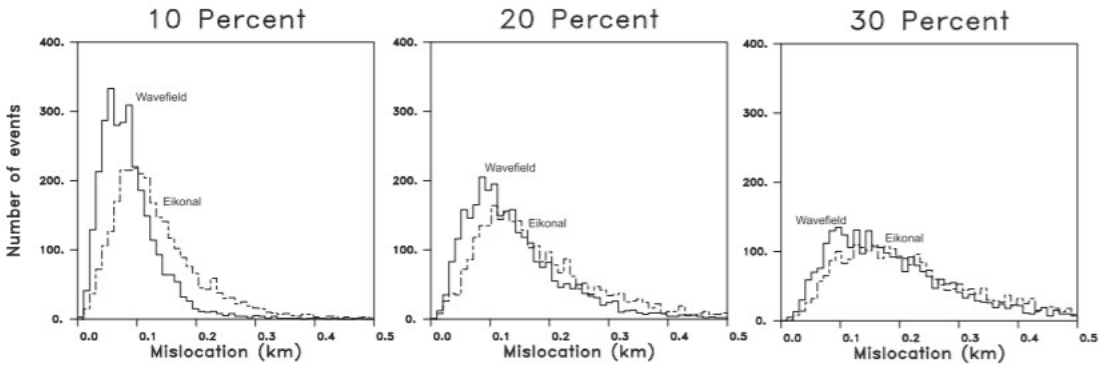

Figure 7. Histograms of error estimates from the location of the 3000 events shown in Fig. 5. The peak magnitudes of the errors as a percentage of the traveltime signal are indicated. The dashed lines denote the histograms associated with locations based upon traveltimes calculated using the eikonal equation. The solid lines indicate location errors corresponding to an algorithm that uses wave field derived arrival times.

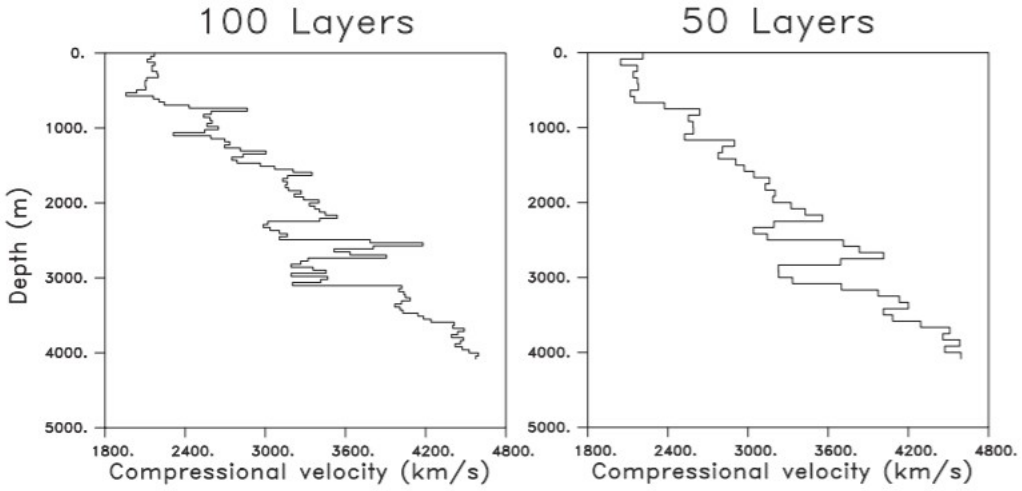

Figure 8. Two velocity models constructed by averaging the 150 layer model shown in Fig. 1 using fewer layers. Left-hand panel: Model representing vertical velocity variation using 100 layers. Right-hand panel: a 50 layer velocity model.

Let us now increase the root-mean-squared magnitude of the errors and examine the resulting location errors. The effects of increasing errors are evident in Fig. 7, where we plot histograms for various levels of added error. As the range of the deviations increases, the distributions develop longer and longer tails with increasing mislocation errors. The shape of the mislocation histograms resemble log-normal distributions after the errors reach 20 per cent of the peak traveltime variation. For error magnitudes that are 10 per cent of the traveltime signal the peak of the wavefield locations is still around 50 m with a secondary peak at 100 m. The peak of the eikonal-based mislocation distribution is at 100 m, with a much larger tail to values approaching 400 m. After the traveltime errors reach 20 per cent of the peak traveltime signal the distribution of wavefield mislocation errors peaks at 75 m and the tail extends to 400 m. The eikonal-based locations still have larger mislocation errors and the peak of the distribution has shifted to values somewhat greater than 100 m. When errors reach 30 per cent of the peak traveltime signal both distributions have long tails, though the wavefield-

based histogram has a peak that is about 50 m smaller than the eikonal-based location errors.

These results are likely influenced by the spacing of the grid used to model the traveltime field. We did not use any interpolation methods to try and obtain subgrid spacing accuracy. That is, we did not try and interpolate between a local network of gridpoints surrounding the minimum of $a(\boldsymbol{\gamma})$ in order to refine the location. Such adjustments could improve the location precision and reduce the data misfit.

3.2 A synthetic test with velocity model errors

Another common source of error is associated with the use of an incorrect velocity model. Velocity models are almost always approximate and involve some assumptions or averaging. For example, sonic log variations are often spatially averaged along the length of the well when estimating the compressional and shear velocity variations as a function of depth. The model in Fig. 1 was constructed in this fashion by averaging sonic log variations in order to derive a 150 layer model. While it is difficult to fully characterize the wide variety of possible modelling errors, we can consider a simple and common issue. Specifically, we consider the case in which the velocity model is under parametrized and smaller scale spatial variations are averaged into larger-scale features. The reference model for this test case is the 150 layer model plotted in Fig. 1.

We consider two coarser representations constructed using 100 layers and 50 layers (Fig. 8). In these tests, the times are generated using the fine scale model shown in Fig. 1 along with a numerical code to calculate the waveforms and a picking routine to identify the first arrivals. These values are then used as the arrival times to be matched via the location algorithm described above. Both the eikonal equation and the waveform codes are used for the earthquake location algorithm but with traveltime fields $T(\boldsymbol{\gamma};x_m)$ calculated using the coarser scale models shown in Fig. 8. The 3000 test events plotted in Fig. 5 are used to generate the distributions of locations errors shown in Fig. 9 for both velocity models. Comparing these results with the error-free distributions in Fig. 6 we see that the location errors are greater for both the eikonal and waveform estimates. Using the 100 layer model results in mean errors peaking at around 190 m for the eikonal-based location algorithm and of 150 m for the waveform-based approach. When 50 layers are used to represent the velocity model, the errors for the eikonal-based algorithm do not increase by much and the distribution remains close to that of the 100 layer model. Similarly, the waveform-based location method produces a distribution that is very close to the previous 100 layer model results. The error distribution peaks at about 140 m with the differences due to the variability within the histogram.

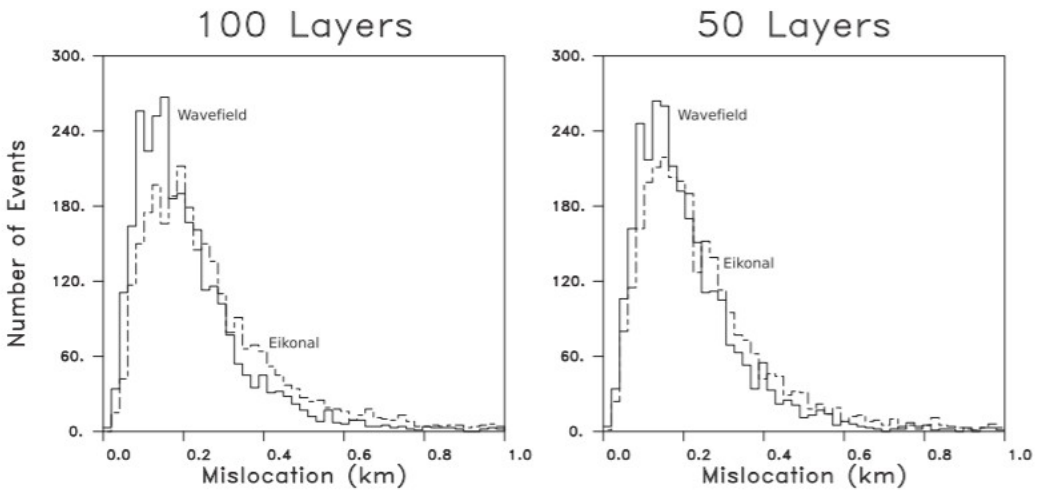

Figure 9. Histograms of the magnitude of the mislocation vectors for locations done with the reduced velocity models shown in Fig. 8. Left-hand panel: Both the wavefield (solid line) and eikonal (dashed line) mislocations for the 100 layer model. Right-hand panel: Wavefield and Eikonal errors for locations based upon the 50 layer model.

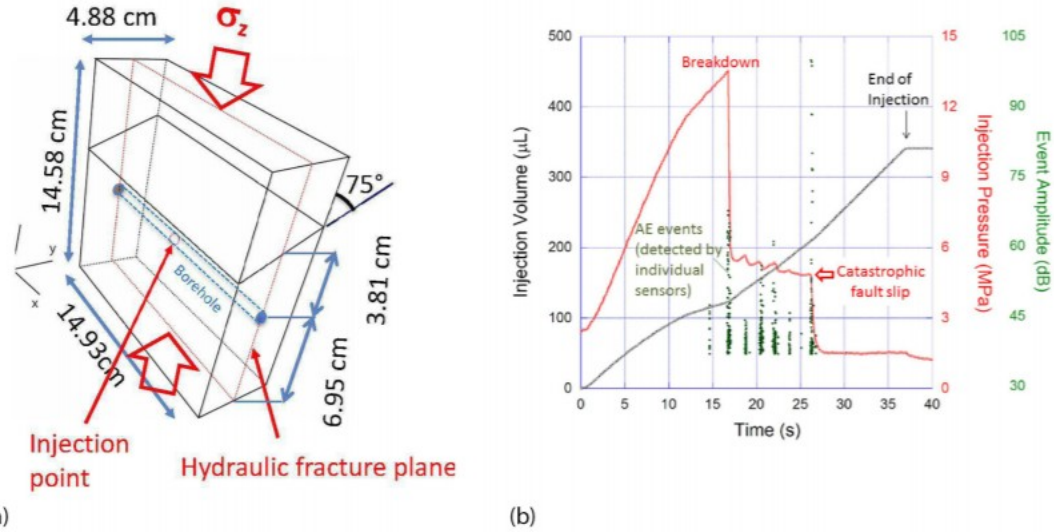

Figure 10. (a) Schematic diagram of the acrylic block used in the laboratory experiment. The borehole used to inject the fluid is labelled as is the hydraulic fracture plane. (b) Fluid volume, pressure and acoustic emission sequence associated with the injection into the acrylic block.

3.3 A laboratory injection test

Our second test involves a data set from a laboratory injection experiment in an acrylic block subjected to uniaxial loading. This setting is not as tightly controlled as a numerical test, yet better constrained than the field experiment described below. The geometry of the experiment is shown in Fig. 10, and the block is approximately 14.93 cm long, 14.58 cm high and 4.88 cm thick. A 6.4 mm diameter borehole containing stainless steel tubing is located at 6.95 cm above the base of the block. The tubing has a pair of 1 mm diameter perforations at the centre, and the injection section is sealed by a pair of O-rings that are located approximately 1 cm from the sides of the block. During the experiment pure glycerol (viscosity 1000 cP) containing red

dye was injected into the borehole via a screw pump. For the duration of the experiment the block was subjected to an axial stress of 6.9 MPa.

The time variation of the injection volume and pressure are shown in Fig. 10 along with the sequence of detected seismic events. The breakdown occurred 17 s from the start of the injection at a maximum injection pressure of 13.5 MPa. In order to guide the propagation direction of an induced fracture, a notch was introduced at the top side of the centre of the borehole. Once the fracture was initiated from this notch, it propagated upwards towards a saw-cut surface oriented 75° from the vertical (17 to 26 s). This dipping surface was flattened and roughened by using a 220-grit sanding paper. As soon as the fracture reached the pre-cut surface, the sheared surface was destabilized, resulting in catastrophic slip. Note that in Fig. 10, the detection threshold for the acoustic emission transducers was set at 37 dB, and the transducers were triggered multiple times when an incoming wave had a large amplitude and a long duration. Most events did not trigger enough of the transducers, that is at least six instruments, to provide a useful location.

The locations of the 12 transducers that recorded the arrivals from the acoustic emission events are shown in Fig. 11. These transducers are 4.7-mm wide and 5.8-mm tall micro acoustic emission sensors (PICO-Z, Score Dunegan) with a frequency band of 200–800 kHz for less than 20 dB drops in the sensitivity from the peak frequency of 650 kHz. The transducers are glued to the block surface via Loctite superglue. The transducers were used in both active mode to conduct seismic wave transmission tests between the transducers and in passive mode to detect and record acoustic emission signals. The collected waveforms from both active and passive measurements were processed and recorded by a digital multichannel acoustic emission measurement system (AMSY-6, Vallen Systeme). The locations of the acoustic emission sources were determined from the traveltime of the first-arriving compressional waves. We do not expect detectable change in the velocity of the homogeneous and intact matrix of the acrylic block during the injection. Also, because the induced fracture was less than 100 microns and filled with highly viscous glycerol, it should have little impact on the propagation of the observed high-frequency compressional waves. Traveltime fields from the 12 transducer locations were used to construct the functions $T(\boldsymbol{\jmath},\boldsymbol{x}_m)$ needed for the location algorithm. The numerical grid used to compute the traveltime fields had a spacing of 0.5 mm and was 100 (x) by 300 (y) by 300 (y), extending slightly beyond the boundaries of the physical block. The spacing was considered adequate because the tranducers themselves were several millimetres in lateral extent.

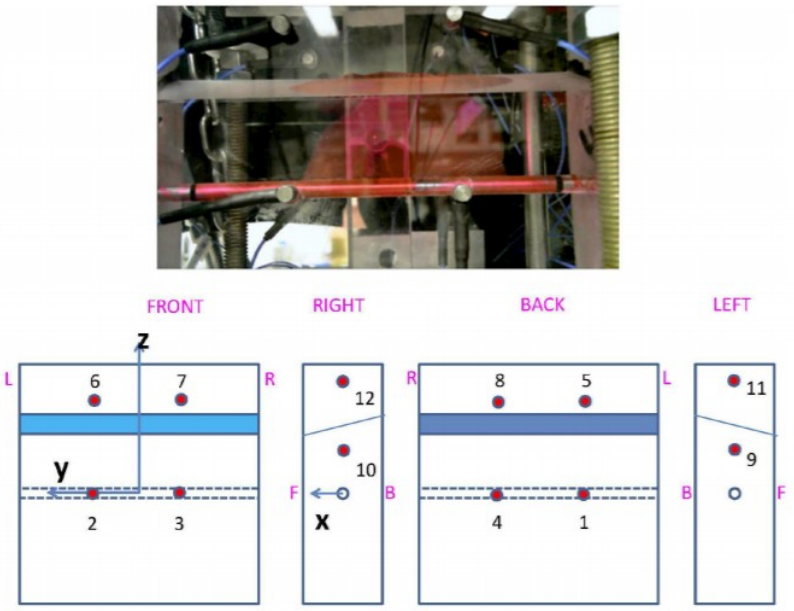

Figure 11. Upper panel: photograph of the acrylic block with the borehole and fracture dyed red. Lower panel: the location of the borehole (dashed lines), acoustic sensors (filled red circles) and dipping plane (blue line).

The arrival times of the observed compressional waves were determined using an Akaike Information Criterion-based first-arrival picking algorithm (Akaike 1974; Maeda 1985; Kuperkoch *et al.*2010). Using the traveltimes and the functions $T(\boldsymbol{\jmath},\boldsymbol{x}_m)$ we can compute the mean or median backpropagated traveltimes over the grid and the average deviation $a(\boldsymbol{\jmath})$. We plot cross-sections through misfit functions for two events in Fig. 12. The cross-sections through the misfit function for Event 4 indicates a relatively well-constrained location. The event was located using 11 traveltime observations providing adequate convergence. Note that the misfit function, $a(\boldsymbol{\jmath})$ given by eq. (6), is not smooth and has some sharp bends and kinks due to the use of the absolute deviation of the residuals. Also, there seems to be less constraints on the depth of the event below 0.0 cm. The poor constraints are due to the lack of tranducers below 0.0 cm as shown in Fig. 11. The misfit function associated with event 17, also plotted in Fig. 12, appears to be well constrained in x, but have a trade-off between y and z positions. This event is only constrained by six observations, leading to some degree of trade-off. However, its position well within the network of transducers, that is above a depth 0.0 cm, leads to a somewhat better-constrained location.

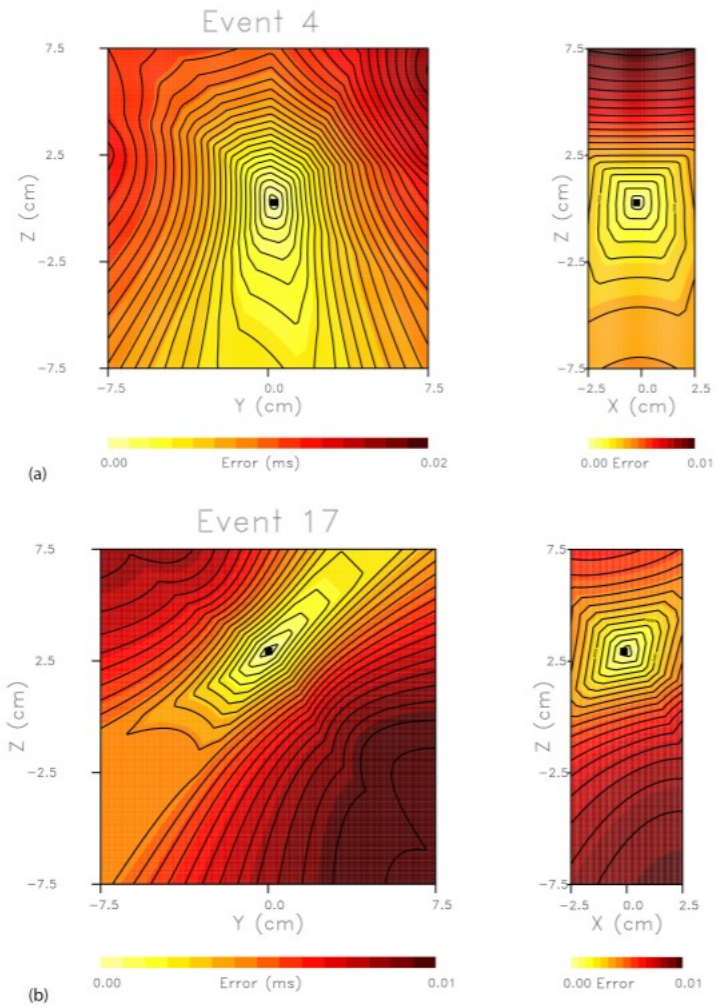

Figure 12. (a) X–Y slice through the absolute deviation error volume $a(J)$ associated with event 4 (Left-hand panel). The cross-sections are through the event location. (Right-hand panel) X-Z cross-section the error volume. (b) X-Y slice through the absolute deviation error volume $a(J)$ corresponding to event 17 (Left-hand panel). The cross-sections are through the event location. (Right-hand panel) X-Z cross-section the error volume.

There were 20 events that had six or more identified arrivals and could be reliably located (Fig. 13). The locations tend to organize along two linear trends that follow the intersection of the fracture with the borehole and with a dipping fracture that was cut into the block lying at a height of about 2.5 cm (Fig. 13). None of the located events appear to lie between the borehole and the saw-cut angled fracture. There are two exceptional events near the edge of the block, one at the right edge along the extension of the upper linear trend and another at the very top of the acrylic block. In general, the events agree with locations produced by a simplex-based location algorithm shown as open circles in Fig. 13. The modified simplex method utilized straight rays under the assumption of a homogeneous medium to calculate the traveltimes from the source to the receiver. The original simplex technique of Nelder & Mead (1965) for optimization in an n-dimensional parameter space progressively updates a simplex of $n + 1$ vertices by the systematic application of reflection, expansion and contraction operations in order to replace the vertex (model) with the highest misfit by one with a

lower misfit. Huang & McColl (1996) proposed improving the convergence of the original method by using a vertex value-weighted centroid, rather than the geometric centroid. The simplex approach is applicable to misfit functions that are not smooth, allowing it to be used for minimizing the median of the traveltime residuals. Furthermore, because it works with a distribution of solutions there is a reduced probability that it will become trapped in a local minimum. The locations produced by our approach do appear to define the two linear patterns that would be expected in this experiment better than the conventional location technique. The first is from the intersection of two planes, the hydraulic fracture plane and the dipping fracture. The second is from the intersection of the hydraulic fracture plane and the top of the borehole.

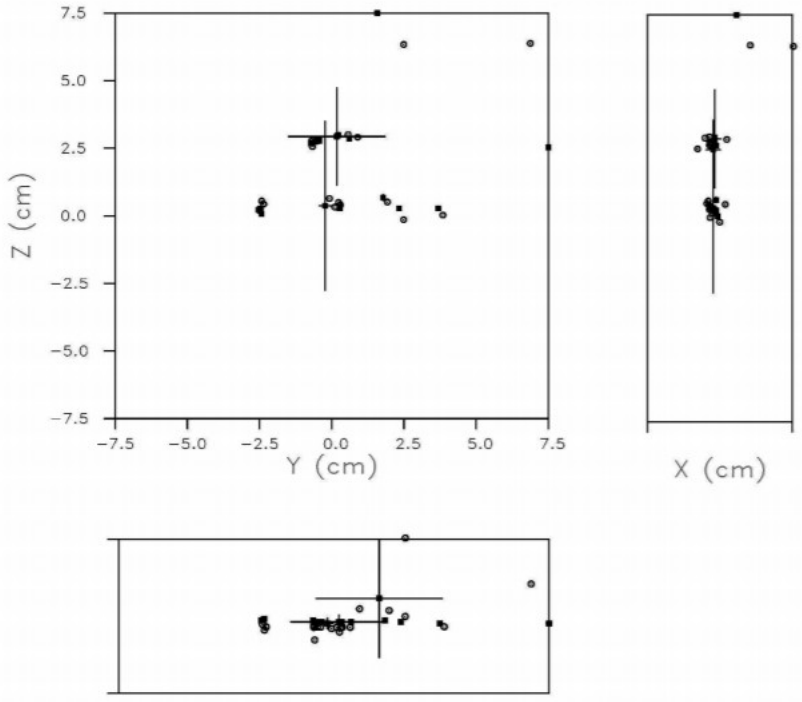

Figure 13. The 20 located events for the injection experiment in the acrylic block. The conventional locations are denoted by the open circles while the locations from the new algorithm are indicated by the filled squares.

3.4 Locating an event associated with the creation of a hydrofracture

The final application involves observations from the monitoring of a hydrofracture in a producing oil field. Four vertical observation wells, each containing 20 seismometers, surround the fracture location, as shown in Fig. 14. As in the Cranfield example, the velocity structure contains layers with sharp boundaries. A layered elastic model is available from well logs in the area, providing compressional and shear velocity as well as density. Velocity contrasts of up to 50 per cent from around 4 to 6 km s^{-1} occur over depth intervals of around 100 m or so. The 1 km (x) by 1 km (y) by 2 km (z) volume was discretized onto a grid with 10 m spacing for the purpose of computing traveltimes associated with each of the 80 receiver locations. Because both

compressional and shear arrivals were available, we included both phases in the location algorithm. That required computing a total of 160 traveltime grids for the 80 stations and both the compressional and shear phases. For a single event, the approach requires greater overhead than a conventional location algorithm. However, the approach is more efficient when several hundred or thousand locations need to be determined.

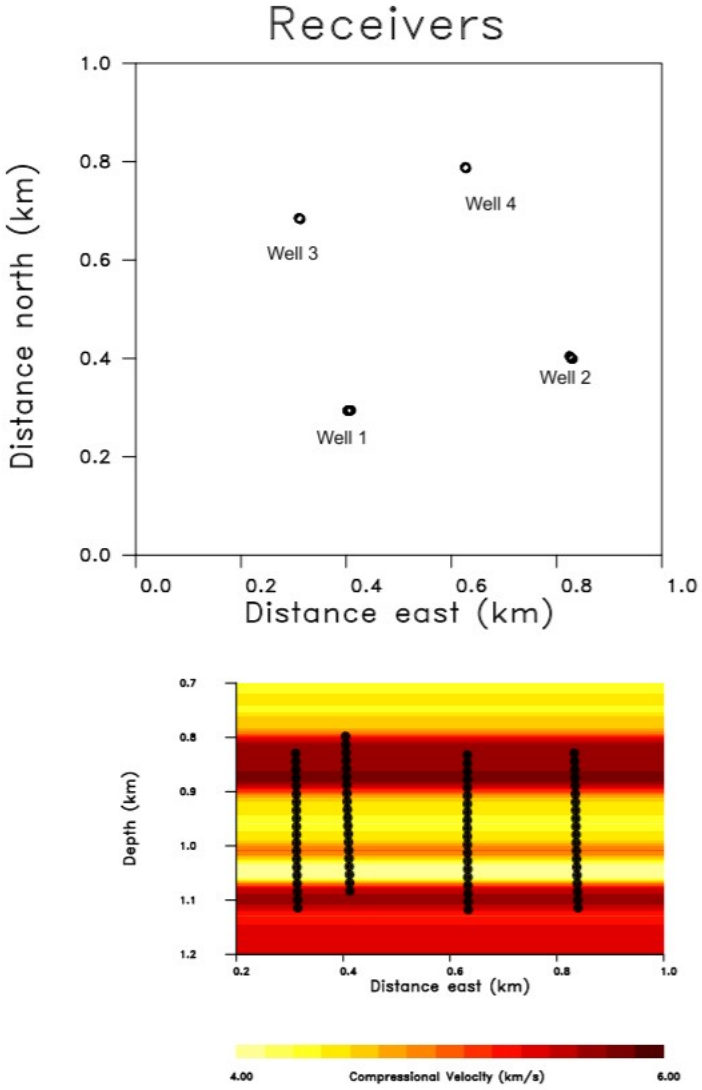

Figure 14. Upper panel: Horizontal positions for the receivers in the four vertical wells that were used to monitor a hydrofracture. Lower panel: East–west vertical slice through the compressional velocity model. The projections of the locations of the 80 receivers in the four vertical wells are plotted as filled circles.

The main purpose of this example is to compare the new approach with a conventional location algorithm on an application to real field data. The backpropagated and shifted times were used to calculate the mean time and the average deviation $a(\boldsymbol{J})$. Horizontal and vertical cross-sections through the average deviation grid are shown in Fig. 15. The event is located in the interior of the array, near the bottom edge of the volume defined by the wells. The calculated location is indicated by the filled circle in the top panel of Fig. 15. The location estimated by a conventional approach, based upon

ray theory, is denoted by a filled square. The conventional approach involved an iterative minimization of the sum of the squares of the residuals using a gradient-based algorithm with a Levenberg–Marquardt approximation to the second derivative terms. The two locations are roughly 10 m apart, which is the spacing of the modelling grid. From Fig. 15 it is evident that there is greater uncertainty in the depth of the event, probably due to its location near the base of the well array. Including both compressional and shear arrivaltimes reduced the trade-off between the origin time and the depth of the event. The fit to the observed traveltimes is shown in Fig. 16. In general both compressional and shear arrivals are matched. However, there are some notable outliers that are offset from the 45° line indicating a good match. These arrivals are significantly early and may be early picks of a precurser to the actual arrival.

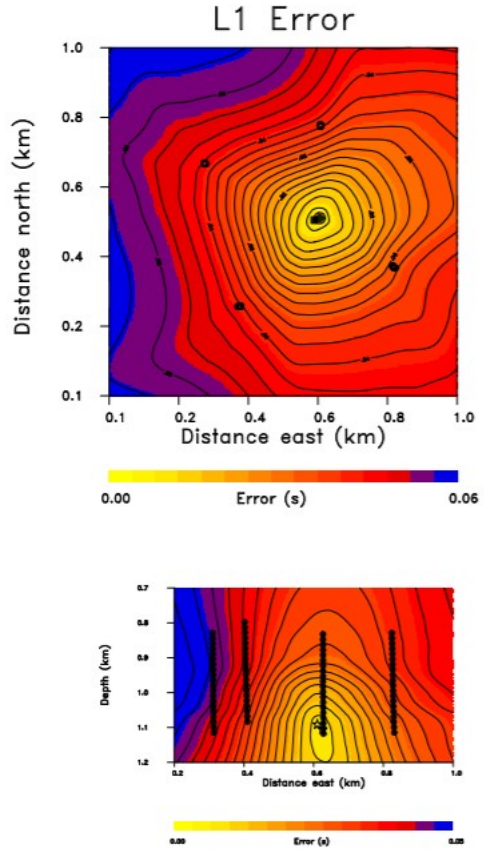

Figure 15. Upper panel: Horizontal slice through the L_1 error volume $a(J)$ given in eq. (6). The section was taken at the estimated depth of the event 1.1 km. The horizontal position estimated by a conventional iterative location algorithm is indicated by the filled square. The location by the new approached base upon tabulated traveltime fields is given by the filled circle. Lower panel: East–west vertical slice through the L_1 error volume. The filled circles are the positions of the 80 receivers. The two locations resulting from the conventional and new location algorithms are shown as unfilled stars.

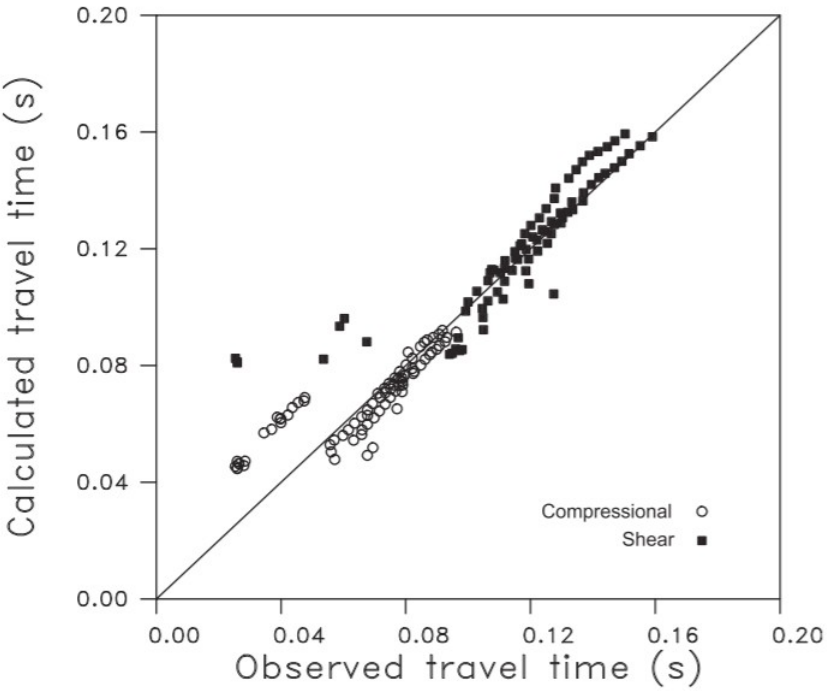

Figure 16. Compressional (open circles) and shear (filled squares) traveltimes for the 80 receivers used in the hydrofracture monitoring experiment. The 45° line corresponds to an exact fit.

4 CONCLUSIONS

The method that we have described, an extension of eikonal-based approach of Nelson & Vidale (1990), is well suited for incorporating waveform modelling into the location algorithm. This can be advantageous in the presence of thin layers or rapid variations in velocity which may cause the eikonal equation to break down (Vasco & Nihei 2019). Seismic events are often associated with such complex velocity structures as found in shear zones, subducting slabs and areas of volcanic activity. Our results indicate that such fine-scale structure and rapid variations in properties can produce location errors of 300 m or more in a realistic elastic model.

The final application demonstrated that the approach can incorporate both compressional and shear arrivals into the location algorithm. It should also be possible to include arrivaltimes and take-off angles derived from three component instruments. The station traveltime fields can be used to trace paths down the gradient of the traveltime field in the direction of the take-off angles. For example, paths from all stations can be used to find the location in the grid where the trajectories converge and the backpropagated times display the least scatter. Take-off angles from a traveltime field calculated using the eikonal equation can deviate significantly from angles calculated from a full numerical solution of the elastodynamic equation. Thus, the use of take-off angle can introduce significant errors in models with rapid variations in elastic properties, such as layers with sharp boundaries.

Any numerical technique may be used to calculate the traveltime fields that are the basis for the algorithm and it should be straightforward to incorporate effects such as anisotropy, poroelasticity and attenuation into the approach. The only requirement is that numerical methods be available to handle such effects. As the modelling becomes more complicated and time consuming, an approach that scales as the number of stations, rather than as the number of events, should be more efficient when there are a significant number of events. Because the technique works with traveltime fields, instead of backpropagating waveforms, we do not need to recalculate the wave fields for each event. Rather, as long as the size of the events are of the same order of magnitude and similar in frequency content, a single waveform calculation should suffice for each station.

The technique should be applicable to a wide range of monitoring networks from the smale-scale laboratory and borehole arrays treated in this work to large-scale global networks. However, the overhead associated with the calculation of the receiver traveltime fields will grow as the number of stations increase and the volume of interest becomes larger. Thus, form of parallel computation might be required for very large networks of stations. In addition, operations such as the search for the minimum will require greater computational resources for a larger grid. To some degree, these issues may be treated by decimating the representation of the traveltime field and interpolating between active node points when searching for the minimum.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences and Biosciences Division under contract number DE-AC02-05- CH11231. We would like to thank Lane Johnson and Chevron for the traveltime data associated with the hydrofracture.

REFERENCES

Akaike, H., 1974. Markovian representation of stochastic processes and its application to the analyses of autoregressive moving average processes, Ann. Inst. Stat. Math., 26, 363–387.

Angus, D.A., Aljaafari, A., Usher, P. & Verdon, J.P., 2014. Seismic waveforms and velocity model heterogeneity: towards a full-waveform microseismic location algorithm, J. appl. Geophys., 111, 228–233.

Bolt, B., 1960. The revision of earthquake epicentres, focal depths and origin times using a high-speed computer, Geophys. J. R. Astron. Soc., 3, 433–440.

Boschi, L., Molinari, I. & Reinwald, M., 2018. A simple method for earthquake location by surface-wave time reversal, Geophys. J. Int., 215, 1–21.

Buland, R., 1984. Residual statistics, Terra Cognita, 4, 268.

Buland, R., 1986. Uniform reduction error analysis, Bull. seism. Soc. Am., 76, 217–230.

Dreger, D., Uhrhammer, R., Pasyanos, M., Franck, J. & Romanowicz , B., 1998. Regional and far-regional earthquake locations and source parameters using sparse broadband networks: a test on the Ridgecrest sequence, Bull. seism. Soc. Am., 88, 1353–1362.

Geiger, L., 1910. Herdbestimmung bei Erdbeben aus den Ankuftszeiten. Nachrichen von der Gesellschaft der Wissenschaften zu Gottingen, Math. Phys., 1910, 331–349.

Grigoli, F. et al., 2017. Current challenges in monitoring, discrimination, and management of induced seismicity related to underground industrial activities: a European perspective, Rev. Geophys., 55, 310–340.

Grigoli, F., Scarabello, L., Bose, M., Weber, B., Wiemer, S. & Clinton, J.F., 2018. Pick- and waveform-based techniques for real-time detection of induced seismicity, Geophys. J. Int., 213, 868–884.

Hoel, P.G., 1971. Introduction to Mathematical Statistics, John Wiley and Sons.

Huang, Y. & McColl, W.F., 1996. An improved simplex method for function minimization, in Proc. IEEE Int. Conf. Systems, Man, and Cybernetics, pp. 1702–1705 .

Huber, P.J., 1981. Robust Statistics, John Wiley and Sons.

Ishii, M., Shearer, P.M., Houston, H. & Vidale, J.E., 2005. Extent, duration, and speed of the 2004 Sumatra-Andaman earthquake imaged by the HiNet array, Nature, 235, 933–936.

James, G., Witten, D., Hastie, T. & Tibshirani, R., 2017. An Introduction to Statistical Learning, Springer.

Jeffreys, H., 1932. An alternative to the rejection of observations, Proc. R. Soc. Lond., 187, 78–87.

Kuperkoch, L., Meier, T., Lee, J. & Friederich, W., EGLELADOS Working Group, 2010. Automated determination of P-phase arrival times at regional and local distances using higher order statistics, Geophys. J. Int., 181, 1159–1170.

Larmat, C., Montagner, J.-P., Fink, M., Capdeville, Y., Tourin, A. & Clevede, Y., 2006. Time-reversal imaging of seismic sources and application to the great Sumatra earthquake, Geophys. Res. Lett., 33, L19312, doi:10.1029/2006GL026336.

Larmat, C., Tromp, J., Liu, Q. & Montagner, J.-P., 2008. Time reversal location of glacial earthquakes, J. geophys. Res., 113, 1–9.

Langet, N., Maggi, A., Michelini, A. & Brenguier, F., 2014. Continuous kurtosis-based migration for seismic event detection and location, with application to Piton de la Fournaise Volcano, La Reunion, Bull. seism. Soc. Am., 104, 229–246.

Lomax, A., Zollo, A., Capuano, P. & Virieux, J., 2001. Precise, absolute earthquake location under the Somma-Versuvius volcano using a new 3D model, Geophys. J. Int., 146, 313–331.

Lomax, A., Michelini, A. & Curtis, A., 2009. Earthquake location, direct, global-search methods, in Encyclopedia of Complexity and Systems Science, ed. Meyers, R.A., pp. 2449–2473, Springer.

Lu, J. et al., 2012. Complex fluid flow revealed by monitoring CO2 injection in a fluvial formation, J. geophys. Res., 117, 1–13.

Maeda, N., 1985. A method for reading and checking phase times in autoprocessing system of seismic data, Zisin Jishin, 38, 365–379.

Masson, Y.J. & Pride, S.R., 2011. Seismic attenuation due to patchy saturation, J. geophys. Res., 116, 1–17.

Masson, Y.J., Pride, S.R. & Nihei, K.T., 2006. Finite difference modeling of Biot's poroelastic equations at seismic frequencies, J. geophys. Res., 111, 1–13.

Menke, W., 2018. Geophysical Data Analysis: Discrete Inverse Theory, Academic Press.

Michelini, A. & Lomax, A., 2004. The effect of velocity structure errors on double-difference earthquake location, Geophys. Res. Lett., 31, 1–4.

Nelder, J.A. & Mead, R., 1965. A simplex method for function minimization, Comput. J., 7, 308–313.

Nelson, G.D. & Vidale, J., 1990. Earthquake locations by 3-D finitedifference travel times, Bull. seism. Soc. Am., 80, 395–410.

Petrov, P.V. & Newman, G.A., 2012. 3D finite-difference modeling of elastic wave propagation in the Laplace-Fourier domain, Geophysics, 77, T137–T155.

Pulliam, R.J., Vasco, D.W. & Johnson, L.R., 1993. Tomographic inversions for mantle P wave velocity structure based on the minimization of l2 and l1 norms of International Seismological Centre travel time residuals, J. geophys. Res., 98, 699–734.

Podvin, P. & Lecomte, I., 1991. Finite-difference computation of traveltimes in very contrasted velocity models: A massively parallel approach and its associated tools, Geophys. J. Int., 105, 271–284.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1992. Numerical Recipes, Cambridge Univ. Press.

Sambridge, M.S. & Kennett, B.L.N., 1986. A novel method of hypocentre location, Geophys. J. R. astr. Soc., 87, 679–697.

Schweitzer, J., 2001. HYPOSAT - An enhanced routine to locate seismic events, Pure appl. Geophys., 158, 277–289.

Sethian, J.A., 1999. Level Set Methods and Fast Marching Methods, Cambridge Univ. Press.

Shearer, P.M., 1997. Improving local earthquake locations using the L1 norm and waveform cross correlation: application to the Whittier Narrows, California, aftershock sequence, J. geophys. Res., 102, 8269–8283.

Shearer, P. 2009. in Introduction to Seismology, Cambridge Univ. Press.

Vasco, D.W., 2018. An extended trajectory mechanics approach for calculating the path of a pressure transient: derivation and illustration, Water Resour. Res., 54, 1–19.

Vasco, D.W. & Nihei, K.T., 2019. Broad band trajectory mechanics, Geophys. J. Int., 216, 745–759.

Vasco, D.W., Johnson, L.R., Pulliam, R.J. & Earle, P.S., 1994. Robust inversion of IASP91 travel time residuals for mantle P and S velocity structure, earthquake mislocations, and station corrections, J. geophys. Res., 99, 13727–13755.

Vidale, J., 1988. Finite-difference calculation of traveltimes, Bull. seism. Soc. Am., 78, 2062–2076.

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

Wuestefeld, A., Greve, S.M., Nasholm, S.P. & Oye, V., 2018. Benchmarking earthquake location algorithms: a synthetic comparison, Geophysics, 83, KS35–KS47.

Zelt, C.A. & Barton, P.J., 1998. 3D seismic refraction tomography: a comparison of two methods applied to data from the Faeroe Basin, J. geophys. Res., 103, 7187–7210.