Locating microseismic sources with a single seismometer channel using coda wave interferometry

Youqian Zhao¹, Andrew Curtis¹, and Brian Baptie²

ABSTRACT

A novel source location method based on coda wave interferometry (CWI) was applied to a microseismic data set of mining-induced events recorded in Nottinghamshire, England. CWI uses scattered waves in the coda of seismograms to estimate the differences between two seismic states. We used CWI to estimate the distances between pairs of earthquake locations, which are then used jointly to determine the relative location of a cluster of events using a probabilistic framework. We evaluated two improvements to this location technique: These account for the impact of a large difference in the dominant wavelength of a recording made on different instruments, and they standardize the selection of parameters to be used when implementing the method. Although the method has been shown to produce reasonable estimates on larger earthquakes, we tested the method for microseismic events with shorter distinguishable codas in recorded waveforms, and hence, fewer recorded scattered waves. The earthquake location results are highly consistent when using different individual seismometer channels, showing that it is possible to locate event clusters with a single-channel seismometer. We thus extend the potential applications of this cost-effective method to seismic events over a wider range of magnitudes.

INTRODUCTION

Finding relative locations of seismic events is essential for discriminating earthquake fault and auxiliary planes from the sequences of aftershocks or foreshocks (Got et al., 1994), studying earthquake interaction and recurrence (Chen et al., 2013), and monitoring stress state and induced (micro)seismicity (Ellsworth, 2013). Instead of obtaining the relative locations of a cluster of events from their absolute locations, which are typically subject to inaccuracies in subsurface velocity models, they are preferably estimated directly from traveltime differences between early arriving body waves, for example, using joint hypocenter determination (Douglas, 1967; Dodge et al., 1995) or master event location (Deichmann and García-Fernandez, 1992). The double-difference (DD) method of Waldhauser and Ellsworth (2000) locates clusters of earthquakes distributed over larger distances than the dominant recorded seismic wavelength when a dense seismic station network is available. However, in areas with fewer stations or unfavorable event-to-seismometer azimuthal coverage, the performance of the DD method deteriorates. In such circumstances, a recently developed source location method based on coda wave interferometry (CWI) (Snieder, 2006) would be particularly useful if proven to be reliable over the range of earthquake magnitudes of interest.

Coda refers to the later part of a seismogram, generated by multiply scattered waves. Seismic coda is disregarded in many seismological applications due to its complex appearance with few uniquely identifiable arrivals. However, the coda is extremely sensitive to small changes in the seismic system. Aki (1969) uses the spectrum of coda from local earthquakes to estimate seismic moments, and more recently Snieder et al. (2002) develop CWI to use phase information to infer differences between pairs of sources or changes in the medium. CWI has been used to determine the relative velocity changes in the earth’s subsurface (Poupinet et al., 1984) and other solid materials and to monitor changes in the interior of volcanoes (Wegler et al., 2006). A third range of application of CWI emerged is to estimate the source separation between earthquakes with similar source mechanisms (Snieder and Vrijlandt, 2005; Robinson et al., 2011). The resulting source separations are used by Robinson et al. (2013) to determine the relative location of earthquakes, in which they demonstrate that CWI appears to outperform conventional location methods when the number of seismic stations is low.

CWI requires that the recorded coda contains waves that leave the source in many directions, and which are then multiply scattered toward the seismometers by heterogeneities in the medium. Robin-
son et al. (2013) test the method on events with long codas — hence, on data that contained many scattered arrivals per event. In our implementation of this novel technique, we apply the method to a microearthquake data set from a colliery in Nottinghamshire, England, for which the codas are relatively short. We thus test the location technique in a new region, and for sources for which the theoretical requirements of CWI are less obviously fulfilled. Additionally, the magnitude range is such that the method might have significant common industrial as well as academic and hazard-related applications. We find that source separation estimates are highly consistent and the earthquake location results agree to within estimated uncertainties when using different individual seismometer channels. Therefore, it is possible to locate these earthquakes with a single-channel seismometer. We also discuss two issues that arise during the implementation and provide solutions because they may be encountered by other authors in the future.

DATA

The British Geological Survey (BGS) deployed a temporary seismic recording network of four 3C (NOLA, NOLD, NOLE, and NOLF) and three Z-component (NOLB, NOLC, and NOLG) stations from February 2014, after some small earthquakes were detected in and around New Ollerton, Nottinghamshire, since the middle of December 2013 (Figure 1a). This is a region where historical seismic activity is related to coal mining. We selected 50 microearthquakes with the magnitudes 0.7–1.8 ML, with the criteria that (1) the signal-to-noise ratio is sufficiently high for the first arrivals to be identified, (2) recorded waveforms contain a distinguishable coda, and (3) the waveform of each event has a maximum correlation coefficient larger than some threshold with waveforms of at least four other events. Dominant frequencies among the different seismic station channels vary between approximately 2.5 and 10 Hz. All waveforms were filtered to 2–20 Hz before processing.

ESTIMATING SOURCE SEPARATION WITH CWI

CWI constrains the separation between a pair of sources by comparing the two seismogram codas recorded at each station. The theory is based on scattering path summation (Snieder, 1999), whereby the total wavefield at a given location is written as the superposition of waves traveling along all possible trajectories

\[
u^1(t) = \sum A_T(t),
\]

where \(u^1\) is the total wavefield from event 1, \(T\) represents a wave trajectory, and \(A_T\) is the contribution to the total wavefield of waves

![Figure 1. Map of the source region and typical waveforms. (a) Thorsby colliery, New Ollerton, Nottinghamshire, England. The star shows the area around which the microearthquakes are likely to have occurred, the triangles are temporary seismic stations, and the rectangles indicate locations of subsurface mining galleries. Panels (b-d) show the waveforms of two earthquakes in event group 3 recorded by the N channel of station NOLF, where panel (b) shows the whole waveforms, panels (c and d) show a 1 s window around the first arrival, and one of the time windows (4 s) in the coda, respectively, both indicated by boxes in panel (b). A comparison between panels (c and d) indicates that the coda is more sensitive to source locations than early arrivals.](image)
that travel along trajectory $T$. All scattered wave trajectories can be divided into the path from the source to the first scatterer and the path followed thereafter. If two nearby and similar sources are compared, CWI assumes that the paths to the first scatterers change, but those subsequent paths do not because they depend mainly on the medium rather than the source: The latter paths simply redistribute the small changes in energy arriving at the first scatterers over space and time throughout the coda. As a consequence, the dominant difference in recorded waveforms is in coda arrival times (Snieder, 2006). Equation 1 represents the waveform of one source recorded at an arbitrary station, so that of another closely located source with identical mechanism recorded at the same station is $u^2(t) = \sum_j A_j (t - \tau_j)$, where $\tau_j$ is the traveltime difference of waves traveling along trajectory $T$ due to the difference in source position. Because the two sources are located close together, the two waveforms will be similar, which can be quantified by the normalized crosscorrelation of the two waveforms in a time window defined with a central time $t$ and a half-width $t_\text{lw}$, calculated for a sequence of time windows in the coda:

$$t_s = \frac{\int_{t-t_\text{lw}}^{t+t_\text{lw}} u^1(t')u^2(t' + t_\text{lw})dt'}{\sqrt{\int_{t-t_\text{lw}}^{t+t_\text{lw}} u^1(t')^2dt' \int_{t-t_\text{lw}}^{t+t_\text{lw}} u^2(t')^2dt'}}.$$  

(2)

In each time window, the distribution of the traveltime differences $\tau_j$ contains information about the source separation $\delta$ (which can be in any direction). Snieder (2006) estimates the standard deviation of the traveltime difference $\sigma_t$ from the maximum of the correlation coefficient $R_{\text{max}}$. The value $\sigma_t$ can then be related to the source separation $\delta$ by $\sigma_t^2 = (1/2)(\delta^2/v^2)$ for an isotropic source in an acoustic medium (where $v$ is the velocity), or by $\sigma_t^2 = [6/a^2 + 7/b^2]/[7(2/a^2 + 3/b^2)]\delta^2$ (where $a$ and $b$ are the P- and S-wave velocities) for a double-couple source in an elastic medium (Snieder and Vrijlandt, 2005). Because the waves arriving in different time windows have traveled along different paths, the separation results of each time window are independent and are used to estimate uncertainty.

We conducted a series of synthetic experiments using multiply scattered waveforms generated with the acoustic Foldy (1945) method (Galetti et al., 2013) in a 16 $\times$ 16 km$^2$ medium with 150 random point scatterers within a velocity of 3000 m/s, with receivers throughout. The 20 sources are located around the center of the medium with Ricker wavelet time functions of equal length, with small intersource distances compared with the dominant wavelength, ensuring similarity in their synthetic waveforms. The results show that individual intersource separations computed from different coda time windows from any one receiver fluctuated by up to 12.8% of their mean across all time windows and that mean estimates for any one receiver lie within the uncertainty bounds of any other single receiver. Such results provide confidence, but they do not include real-world effects, such as true heterogeneity, double-couple sources, or elastic effects, for which we turn to the real data.

Because the theory of CWI requires identical source mechanisms, we classified the 50 New Ollerton microearthquakes into three groups with 33, 10, and 7 events, respectively. Source similarity in each group is assured by waveform similarity, assessed by crosscorrelation. For each group, we picked the first arrivals of waveforms recorded by the same channels of seismic stations and aligned them. We assume that the differences in the coda are due only to varying source locations; this is not true if the seismic velocity also changes. Although the two types of changes might be discriminated (Snieder et al., 2002), this is beyond our scope. Figure 1 shows the similarity of waveforms within one group, and it compares early arrivals and coda of the same pair of waveforms.

We computed intersource separations with data from each individual station. Coda windows begin at 26 s, and we assumed that the P- and S-wave velocities around the sources are 4088 and 2298 m/s. Figure 2a shows the 21 intersource separations from the seven events in group 3. Each data point is the mean of results calculated for all available time windows from each usable channel in each waveform pair, with the standard deviation of each mean displayed in Figure 2b. The results show that source separations and uncertainties are consistent between individual independent stations. Before passing the separation data to the location process, those pairs whose means were estimated to be smaller than their standard deviations were rejected.

**SOURCE LOCATION**

Our source location algorithm estimates relative location from the separation estimates and their uncertainties. It also accounts for a known bias — an increasing tendency toward underestimation of larger true source separations due to cycle skipping in the correlation of coda in equation 2. To quantify this bias, Robinson et al. (2011) apply a conditional probabilistic density function (pdf) $P(\delta_t | \tilde{\delta}_{\text{CWI}})$, the probability of the true separation being $\delta_t$ given that the estimate from CWI is $\tilde{\delta}_{\text{CWI}}$, where the tilde above the separations indicates that the quantities are normalized by the dominant wavelength $\lambda_d$ in recorded data. The source separation estimates from CWI are always smaller than $\lambda_d$, so $\delta_t$ and $\tilde{\delta}_t$ are smaller than one. According to Bayes’ theorem, the posterior probability of $\delta_t$ is proportional to the likelihood of observing $\delta_{\text{CWI}}$ given the true separation $\delta_t$ multiplied by the prior on $\delta_t$:

$$P(\delta_t | \tilde{\delta}_{\text{CWI}}) \propto P(\tilde{\delta}_{\text{CWI}} | \delta_t) \times P(\delta_t).$$  

(3)

The prior is used to incorporate information about source separation or event location known prior to the location process, which here is considered to be a uniform distribution with wide bounds. The likelihood function $P(\tilde{\delta}_{\text{CWI}} | \delta_t)$ is approximated by a positively bounded Gaussian pdf by establishing empirical functions for its mean $\mu = ...$
\[ \mu(\tilde{\delta}_i) = a_1 - \frac{a_2 \tilde{\delta}_i a_3}{a_2 \tilde{\delta}_i a_3 + a_3 \tilde{\delta}_i a^2 + 1}, \]

\[ \sigma(\tilde{\delta}_i) = c + b_1 \frac{b_2 \tilde{\delta}_i b_3 + b_3 \tilde{\delta}_i b_4}{b_2 \tilde{\delta}_i b_3 + b_3 \tilde{\delta}_i b_4 + 1}. \]  

(4)

where \(a_1, \ldots, a_3\) are found to be 0.4661, 48.9697, 2.4693, and 1.1619, respectively; in \(\sigma = \sigma(\tilde{\delta}_i)\), \(b_1, \ldots, b_5\), \(c\) are 0.1441, 101.0376, 120.3864, 2.8430, 6.0823, and 0.017, respectively, and the location results do not change significantly with small changes in the parameter values. Equation 3 holds for each earthquake pair, given the separation data from different time windows for each channel used. Robinson et al. (2013) incorporate the separations between multiple event pairs by multiplying the formulas for all available event pairs, assuming they are independent of each other:

\[ P(e_1, \ldots, e_n|\tilde{\delta}_{CWI}) = c \prod_{i=1}^{n} P(e_i) \times \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} P(\tilde{\delta}_{CWI,ij}|e_i, e_j), \]  

(5)

where \(c\) is a constant, \(n\) is the number of events, \(e_i = (x_i, y_i, z_i)\) is the event \(i\) location; within the last term, we use \(\delta_{ij} = \|e_i - e_j\|_2\) for source separation \(\delta_{ij}\) between the \(i\)th and \(j\)th earthquakes, and the prior \(P(e_i)\) only contains the relative event locations. The maximum of the joint posterior pdf occurs at the most probable combination of the events locations. Taking the negative logarithm of equation 5, the multiplication is converted to summations and the optimization is solved as a minimization problem using a conjugate-gradient method.

We made two improvements to this location method. First, in each iteration of the minimization process, current event locations are used to compute the interevent separations \(\delta_i\), which are then normalized to give the value of the joint posterior. However, the dominant frequency, hence the dominant wavelength, of microearthquakes often extends over a large range, and it is also subject to the limitations of the recording instruments. In this study, the dominant wavelength among different channels varies between 190 and 760 m. Using the average dominant wavelength over different channels/stations therefore introduces inaccuracy to the location process. To this end, when conducting multiple-channel location, we apply an individual likelihood for each channel so that the interevent separations computed during the iterations are normalized by the actual wavelengths. Thus, when data from \(m\) channels are used, equation 5 is modified to

\[ P(e_1, \ldots, e_n|\tilde{\delta}_{CWI}) = c \prod_{i=1}^{n} P(e_i) \times \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} P(\tilde{\delta}_{CWI,ij}|e_i, e_j). \]  

(6)

where \(k\) is the index over \(m\) channels. Second, when specifying the number and length of coda time windows for CWI, we want the separation estimates to be consistent among different channels. Using data from different station channels, we find that the most consistent set of windows varies. Instead of using a trial-and-error method to fix these two fundamental parameters, we conduct a systematic search for an optimal combination that results in the lowest uncertainty of separations calculated between all events using CWI.

We compute a separation-uncertainty matrix for each channel (e.g., Figure 3): For each combination of parameters the matrix element is \(\Omega_{ij} = \sum_N \sigma_{ij}/N\), where \(\sigma_{ij} = \sqrt{\sum_l (\delta_{ij,k} - \tilde{\delta}_{ij})^2}/l\) is the standard deviation of separation estimates from \(l\) coda time windows for events \(i\) and \(j\). \(\delta_{ij,k}\) is the separation estimate of the \(k\)th window and \(\tilde{\delta}_{ij}\) is their mean, and \(N\) is the number of event pairs on that channel. The value of \(\Omega_{ij}\) reflects the uncertainty of separation estimates derived from one recording channel for each source pair. Averaging over multiple source pairs gives a final uncertainty estimate for that combination of parameters (we require at least four estimates to calculate a reasonable standard deviation). Figure 3 shows a typical matrix for the New Ollerton earthquakes, indicating that using four 4 s windows will give the most consistent separation estimates. The values of \(\Omega\) appear to increase with the number of windows when windows are longer than 4.5 s, but this is due to the microseismic codas being short: With more than three such windows, newly added windows are no longer suitable for CWI. This systematic approach thus frees us from the vagaries of trial-and-error methods and from using undefined or subjective quality metrics, and it allows the CWI technique to be automated.

For each of the three groups, minimization of the negative logarithm of the joint location pdf is conducted 10 times with a different randomly distributed initial location. Figure 4a shows the location of group 3 with all available data from 11 seismic station channels, averaged over the results of the 10 experiments, and the variability between the 10 is indicated by 95% confidence ellipses. The average coordinate variability is only 12.7 m, and the minimizations therefore seem likely to have converged to the same (global) minimum, given the uncertainties in the source separation data. We located the events...
Our application of the CWI-based source location method to the New Ollerton earthquakes demonstrates that this method is applicable to microseismicity and to industrial as well as academic purposes. It also shows that the location results with individual stations are highly consistent, as long as the recorded seismograms have a sufficient length of coda. This computationally inexpensive method therefore adds a new technique to the arsenal for seismological applications that require accurate locations of local earthquakes without dense seismometer arrays.

ACKNOWLEDGMENT

We thank the BGS for providing the seismic data of the New Ollerton earthquakes. Y. Zhao is grateful for funding provided by the School of GeoSciences, University of Edinburgh and BGS. We also thank G. Meles for his help at an early stage.

REFERENCES


