

A systematic analysis of correlation-based seismic location methods

Lei Li,^{1,2,3,4} Dirk Becker,³ Hao Chen,^{1,2,4} Xiuming Wang^{1,2,4} and Dirk Gajewski³

¹*Institute of Acoustics, Chinese Academy of Sciences, 100190 Beijing, China. E-mail: chh@mail.ioa.ac.cn*

²*University of Chinese Academy of Sciences, 100049 Beijing, China*

³*Institute of Geophysics, University of Hamburg, D-20146 Hamburg, Germany*

⁴*Beijing Engineering Research Center for Offshore Drilling Exploration and Measurement, 100190 Beijing, China*

Accepted 2017 October 11. Received 2017 October 6; in original form 2017 June 29

SUMMARY

Waveform-based seismic location methods can reliably and automatically image weak seismic sources, such as microseismic events and microtremors. Besides the classical diffraction stacking operator which is based on the one-way traveltimes, correlation-based imaging methods are another subcategory of waveform-based methods using differential traveltimes. In this work, we systematically analyse the existing correlation-based methods and propose a novel hybrid correlation stacking method, which belongs to waveform-based relative location methods. The double differential traveltimes from an event pair (i.e. a master event and a target event) to a receiver pair is used to stack the corresponding double correlation waveforms in this new approach. We generalize the correlation-based methods using a unified formula by describing cross-correlation stacking with beamforming. A thorough analysis of these imaging operators using synthetic and field data examples reveals their different characteristics of imaging resolution and level of redundancy, and a moderate level of redundancy can ensure both the accuracy and stability of correlation-based imaging methods, while an extremely high or low level of redundancy will hinder their performance in locating weak seismic events. The examples also demonstrate the potential disadvantage of using multiple phases with inaccurate velocity models for waveform-based location methods.

Key words: Interferometry; Persistence, memory, correlations, clustering; Time-series analysis; Body waves; Induced seismicity.

1 INTRODUCTION

The seismic source location problem is encountered at different scales and applications in seismology, such as tremor and earthquake location (earthquake seismology; e.g. Stein & Wysession 2003), microseismic monitoring in geothermal, oil and gas reservoirs (exploration seismology; e.g. Maxwell 2014), and rock burst monitoring in mines and tunnels (engineering seismology; e.g. Gibowicz & Kijko 1994). The traditional methods to locate seismic events are traveltimes inversions originating from linearized inversion proposed by Geiger (1912). These methods involve searching for the location that fits the observed traveltimes or differential traveltimes best by iterative inversion algorithms. As a counterpart of the conventional traveltimes inversion methods, modern waveform-based source location methods have been proposed due to their robustness and automatism for detecting and locating weak seismic events (e.g. Cesca & Grigoli 2015, and references therein). Waveform-based location methods do not require phase picking nor phase identification and can detect and locate more events with low signal-to-noise ratio (SNR) data. Instead of inverting the location with traveltimes information only, waveform-based methods image the source by focusing or back-projecting the waveforms into discrete grid points with a certain imaging operator, which is constructed

with related traveltimes information. There are also various alternative terminologies for the waveform-based location methods, such as migration-based methods, back-projection methods, beamforming (BF; delay and sum), coherence scanning, etc. (Maxwell 2014; Cesca & Grigoli 2015).

The waveform-based methods can be classified into two categories (Pesicek *et al.* 2014; Poiata *et al.* 2016): the first one is time reversal methods which exploit full waveforms (e.g. McMechan 1982; Gajewski & Tessmer 2005; Artman *et al.* 2010), these methods are time consuming and require imaging conditions to focus the source energy, the second one is the so called migration or stacking methods which mainly utilize primary seismic phases only (e.g. Kao & Shan 2004; Baker *et al.* 2005). Most methods in the second category are based on the diffraction stacking (DS) operator, which treats a seismic event as a diffraction point and stacks the waveforms along theoretical one-way traveltimes curves (e.g. Gajewski *et al.* 2007; Zhebel *et al.* 2010; Price *et al.* 2015). During the past 10 yr, these DS approaches have been successfully used to locate natural earthquakes (e.g. Kao & Shan 2007) and induced seismicity associated with mining operations (Gharti *et al.* 2010; Grigoli *et al.* 2013), geothermal exploitation (Sick & Joswig 2017), and oil and gas reservoirs (Anikiev *et al.* 2014; Zeng *et al.* 2014; Staněk *et al.* 2015).

Cross-correlation stacking is a relatively new imaging operator for locating seismic sources, and it can be regarded as a subsidiary of seismic interferometric imaging (Schuster *et al.* 2004). Instead of stacking traveltimes curves as in DS, cross-correlation stacking stacks the cross-correlograms along differential traveltimes curves. Actually, the idea of utilizing differential traveltimes on pair-wise receivers from common events has emerged as the master-station method in earthquake seismology since the 1990s (e.g. Zhou 1994). Font *et al.* (2004) modified the master-station method to maximum intersection method by including differential traveltimes from all unique station pairs. Analogous to double-difference (DD) method (Waldhauser & Ellsworth 2000), Zhang *et al.* (2010) proposed the station-pair DD method and successfully located non-volcanic tremors. Wang *et al.* (2016) proposed the interferometry traveltimes inversion and pointed out that elimination of source origin time can make the method more stable under velocity model disturbance. Li *et al.* (2016) developed a virtual field optimization method and verified its superiority in locating sources with traveltimes containing large picking errors. In fact, these recently proposed methods all share the same essence of exploiting differential traveltimes at pair-wise stations from common events. Recently, several researchers implemented this idea into waveform-based methods with cross-correlation techniques from seismic interferometry, and applied them to synthetic data (Grandi & Oates 2009; Li *et al.* 2015; Trojanowski & Eisner 2017), microtremor and volcanic tremor data (Behzadi *et al.* 2015; Droznin *et al.* 2015), mining induced seismicity (Dales *et al.* 2017) and earthquake location (Poiata *et al.* 2016; Ruigrok *et al.* 2017). Several different terminologies are used for this method, such as ‘interferometric imaging’ and ‘interferometric locator’ in Li *et al.* (2015) and Dales *et al.* (2017), ‘cross-correlation beamforming’ in Ruigrok *et al.* (2017), we refer to it as cross-correlation stacking here for unity. By doubling the correlation process, Li *et al.* (2017a,b) extended the previous single cross-correlation stacking to a double-correlation method and a more generalized high-order correlation method. They demonstrated their superior performance in lateral imaging resolution for locating tremors with sparse stations.

For both traveltimes inversion methods and waveform-based methods, velocity models are needed to calculate the theoretical traveltimes, and the location results are directly and strongly dependent on the quality of the given velocity model. To address this issue, the relative location method (Fitch 1975; Spence 1980) is a commonly used approach, which introduces well-located master events and thus compensates the implicit velocity errors and anomalies (e.g. lateral velocity heterogeneities) by introducing additional constraints from the master events. The relative location methods are based on differential traveltimes of an event pair (i.e. a master event and a target event) at common stations. The DD method (Waldhauser & Ellsworth 2000) is a well-known and widely used relative location method in earthquake seismology, and was recently utilized to locate downhole microseismic events (Tian *et al.* 2014, 2016). Poliannikov *et al.* (2011, 2013) proposed an interferometric hydrofracture microseismic location method based on seismic interferometry, which used master events in the neighbouring fracture under the condition of single well monitoring, and they latter unified the DD method and interferometric location into a Bayesian framework-based relative location scheme. Grechka *et al.* (2015, 2016) proposed paraxial ray-based relative location for microseismicity, they suggested including multiple master events to ensure the location accuracy when accounting for large distances between master and target events, as well as rapid velocity variations. Fortunately, the relative location technique can also be incorporated with waveform-based meth-

ods. By introducing master events, Grigoli *et al.* (2016) modified the standard DS to master-event waveform stacking with additional traveltimes correction terms evaluated at each station. Li *et al.* (2016) extended the cross-correlation stacking to relative interferometric imaging by replacing correlation waveforms at pair-wise receivers from the common target event with those at common receivers from pair-wise events. For unity and clarity, we refer to it as relative correlation stacking (RCS) here.

In this work, we compare the existing correlation-based location methods and propose a novel hybrid correlation method, which doubles the correlation process in RCS. Then, we generalize these methods into a unified formula using the relationship between the traditional BF algorithm and correlation stacking. We also analyse the principles of these correlation-based imaging operators and compare their different properties associated with imaging resolution and level of redundancy. Finally, realistic synthetic examples show the performance of these correlation-based imaging methods, and a field data example of mining induced seismicity demonstrates their feasibility in locating weak seismic events.

2 METHOD

In this section, we first describe the characteristic function (CF) used in this work, then introduce the basic theories of correlation-based imaging methods studied here, including a novel hybrid correlation stacking (HCS) method which exploits double differential traveltimes from pair-wise events to pair-wise receivers. Then, a generalization of correlation-based imaging methods is proposed by describing cross-correlation stacking with BF algorithm. Finally, a qualitative analysis of imaging operators for these methods is given by a numerical example with a 2-D homogeneous model.

2.1 Characteristic function

Several CFs of original waveforms can be used in waveform-based methods to extract the signal concealed in the data with low-SNR, as well as to compensate polarity changes in the original waveforms resulting from source radiation pattern. For example, waveform envelopes (Gharti *et al.* 2010; Zeng *et al.* 2014), short-term average to long-term average (STA/LTA) traces (Drew *et al.* 2013; Grigoli *et al.* 2014; Verdon *et al.* 2017), slope-detection function (Tan & He 2016) and kurtosis (Poiata *et al.* 2016) are commonly used CFs based on each individual channel, semblances (Chambers *et al.* 2014) and semblance-weighted values (Eaton *et al.* 2011; Zhang & Zhang 2013) are functions characterizing multichannel coherency. All of above CFs have been incorporated successfully to waveform-based location methods with dense seismic networks.

In this study, we use the STA/LTA traces as the input data of waveform-based location methods. Although CFs based on multichannel coherence are better choices when accounting for low-SNR seismograms (e.g. surface microseismic monitoring of deep reservoirs), the key point in this study is to evaluate the performance of different imaging operators, and the STA/LTA traces are good enough to test the methods here. In order to better extract multiple phases, we utilize separate CFs, that is, CF_H and CF_V for horizontal and vertical components (Grigoli *et al.* 2013). Based on the method in Allen (1978), the horizontal and vertical energy of original seismograms are used to calculate their STA/LTA traces

$$CF_{H/V}(t) = \frac{STA_{H/V}(t)}{LTA_{H/V}(t)}, \quad (1)$$

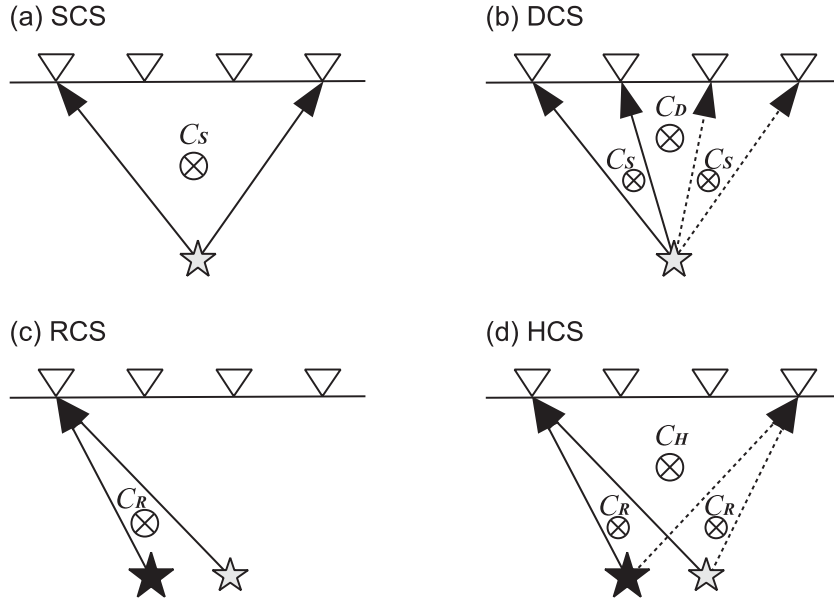


Figure 1. Schematic diagrams of the correlation-based methods: (a) SCS; (b) DCS; (c) RCS; (d) HCS. The reverse triangles are receivers, grey stars are the target events and black stars are master events. The summing junction symbol denotes cross-correlation operation.

$$\text{STA}_{H/V}(t) = \frac{\sum_{\tau=t}^{t+n_s-1} u_{H/V}^2(\tau)}{n_s}, \quad (2)$$

$$\text{LTA}_{H/V}(t) = \frac{\sum_{\tau=t-n_l}^{t-1} u_{H/V}^2(\tau)}{n_l}, \quad (3)$$

where t and τ are time sample indices, $u_H^2 = u_x^2 + u_y^2$ and $u_V^2 = u_z^2$ are horizontal and vertical energy, u_x , u_y and u_z are the three components, n_s and n_l are the time sample lengths of the short and long time windows, respectively. In general, horizontal energy u_H contains strong S -wave energy and vertical energy u_V contains strong P -wave energy (e.g. see Figs 2 and 6). If we just consider single phase, only the corresponding CF (i.e. CF_H or CF_V) is used, otherwise, both CFs are utilized in the stacking process.

2.2 Single correlation stacking

The standard cross-correlation stacking operator is based on the differential traveltime from the target source to pair-wise receivers (see Fig. 1a). The cross-correlograms (i.e. the gather of cross-correlation waveforms) is first generated by cross-correlating characteristic functions of pair-wise receivers (eq. 4) and then stacked along the corresponding differential traveltime curves (eq. 5). In order to distinguish it from double correlation stacking (DCS) below, we refer to it as single correlation stacking (SCS) in this paper. The formulae of SCS in time domain read as

$$C_S^{ij}(\tau) = \sum_{t=0}^{t_{\max}} CF_i(t) CF_j(t + \tau), \quad (4)$$

$$\begin{aligned} S_S(\mathbf{x}) &= \sum_{i=1}^N \sum_{j=i+1}^N \sum_{\tau=0}^{T_{\max}} C_S^{ij}(\tau) \delta[\tau - (\tau_{i,x} - \tau_{j,x})] \\ &= \sum_{I=1}^M C_S^I(d\tau_{I,x}), \end{aligned} \quad (5)$$

where C_S^{ij} and C_S^I are the single correlation waveforms of characteristic functions at receiver pair $I = \{i, j\}$, CF_i is the CF of the considered component at receiver i , $S_S(\mathbf{x})$ is the SCS value at source position \mathbf{x} , δ is the Dirac delta function and it plays the role of the imaging operator, t_{\max} and T_{\max} are the length of time samples in the original CFs and the cross-correlation waveforms, N and M are the number of receivers and all unique receiver pairs, $d\tau_{I,x} = \tau_{i,x} - \tau_{j,x}$ is the term of the differential traveltime of considered seismic phase. Note that the unknown source origin time is cancelled out in the correlation process, which means the location process is independent of the origin time and the traditional 4-D volume imaging problem of DS (Anikiev *et al.* 2014) reduces to a 3-D problem. On one hand, the decoupling of the origin time from the location process alleviates the origin time-depth trade-off results from wave attenuation (Eisner *et al.* 2013; Anikiev *et al.* 2014) or strong seismic coda (Price *et al.* 2015). Wang *et al.* (2016) also argued that the traveltime variation resulting from very small changes of source positions may be overwhelmed by the origin time variation, especially for medium- and small-scale seismicity, so the origin time could probably hinder the accurate inversion of source positions. On the other hand, the origin time is a minor parameter and can be estimated by roughly referring to the measured seismograms (e.g. intervals of different event segments) and relevant underground operations, it can also be calculated posteriorly according to the observed arrival time and theoretical traveltime. In this work, we are only concerned about the accuracy of spatial source parameters (i.e. \mathbf{x}) and abandon analysing the origin time.

Another issue worth noting is the selection of seismic phases (i.e. primary P waves and S waves) involved in the imaging process. For example, if we consider both P waves and S waves in SCS, the complete term of differential traveltime becomes

$$d\tau_{I,x}^{PS} = \begin{bmatrix} (\tau_{i,x}^P - \tau_{j,x}^S) (\tau_{i,x}^P - \tau_{j,x}^P) \\ (\tau_{i,x}^S - \tau_{j,x}^S) (\tau_{i,x}^S - \tau_{j,x}^P) \end{bmatrix}, \quad (6)$$

where $\tau_{i,x}^P - \tau_{j,x}^S$ are differential traveltimes of P wave at receiver i and S wave at receiver j , other terms have similar meanings. Although the combination of P waves and S waves can alleviate the

origin time-depth trade-off and improve the location resolution by providing more constraints (Grigoli *et al.* 2013), they will also introduce extra uncertainty resulting from the additional velocity models (Eisner *et al.* 2010). Gharti *et al.* (2010) pointed out that the contribution of the strong phase to the weak phase stacking can directly cause a bias of the maximum imaging value. In an extreme case, when stacking the seismograms of an explosive source along traveltimes curves of slower S waves, it is likely to obtain a maximum at a deeper position than the real one. Dales *et al.* (2017) pointed out that using the dominant phase only should be the most efficient approach. Our previous study has also shown that involving more correlation terms cannot ensure the improvement of the imaging result, since it may also introduce additional interferences and result in artefacts in the final imaging profile (Li *et al.* 2015). Therefore, we suggest using single phase only in waveform-based methods, especially when reliable multiple velocities or accurate recognition of multiple phases are inaccessible.

2.3 Double correlation stacking

After doubling the correlation process, namely cross-correlating the single cross-correlograms C_S in eq. (4) once more, SCS is extended to DCS (Li *et al.* 2017a). Correspondingly, the double differential traveltimes are needed to focus the double correlation waveforms over receiver triplets and quadruplets (see Fig. 1b). The formulae of DCS can be derived based on that of SCS

$$C_D^{IJ}(\tau) = \sum_{t=0}^{T_{\max}} C_S^I(t) C_S^J(t + \tau), \quad (7)$$

$$S_D(\mathbf{x}) = \sum_{I=1}^M \sum_{J=I+1}^{2T_{\max}-1} C_D^{IJ}(\tau) \delta[\tau - (d\tau_{I,x} - d\tau_{J,x})]$$

$$= \sum_{I=1}^M C_D^{IJ}(d\tau_{IJ,x}), \quad (8)$$

where C_D^{IJ} is the double correlation waveform, S_D is the DCS value, $d\tau_{IJ,x} = d\tau_{I,x} - d\tau_{J,x}$ is the double differential traveltime at receiver pairs I and J . Undoubtedly, a higher level of redundancy is extracted in DCS compared to SCS, but the computational effort is also significantly increased (approximately $(M-1)/2$ times more) if we choose all unique receiver triplets and quadruplets. An alternative scheme is to select one or several receivers as the reference receivers, and then only the double correlation waveforms over these receiver triplets are calculated and stacked (Li *et al.* 2017a).

2.4 Relative correlation stacking

Combining the idea of relative location by introducing a master event m , whose location and origin time are known, we can obtain RCS method (Li *et al.* 2016). The RCS operator is based on the differential traveltime from the two events (i.e. the master event m and target event x) to common receivers (see Fig. 1c). The unknown origin time τ_0 of target event remains in RCS, so the imaging process involves a 4-D searching volume and the stacking process has to be repeated to search for the true τ_0 . As a consequence, there is a similar origin time-depth trade-off as in DS, though the additional information of the master event can help to constrain the location process. We compute the RCS values as

$$C_R^i(\tau) = \sum_{t=0}^{t_{\max}} CF_t^m(t) CF_t^x(t + \tau), \quad (9)$$

$$S_R(\mathbf{x}, \tau_0) = \sum_{i=1}^N \sum_{\tau=0}^{T_{\max}} C_R^i(\tau) \delta[\tau - (\tau_0 + (\tau_{i,m} - \tau_{i,x}))]$$

$$= \sum_{i=1}^N C_R^i(\tau_0 + d\tau_{i,mx}), \quad (10)$$

where C_R^i is the relative correlation waveform, CF^m and CF^x are CFs of master event m and target event x , S_R is the RCS value, δ is still the Dirac delta function and $d\tau_{i,mx} = \tau_{i,m} - \tau_{i,x}$ is the differential traveltimes from the two events to common receivers. The accuracy of the location of master events is essential for the location of target events, and the quality of their seismograms also affects the location process for waveform-based methods. Therefore, we should select events with high location accuracy and high-SNR as master events.

2.5 Hybrid correlation stacking

Inspired by the double correlation idea, we propose to double the relative correlation process (see eq. 9) and obtain a novel HCS method (see Fig. 1d). Inserting eq. (9) into eq. (7), we can derive the formula of HCS as

$$C_H^{ij}(\tau) = \sum_{t=0}^{T_{\max}} C_R^i(t) C_R^j(t + \tau), \quad (11)$$

$$S_H(\mathbf{x}) = \sum_{i=1}^N \sum_{j=i+1}^{2T_{\max}-1} C_H^{ij}(\tau) \delta[\tau - (d\tau_{i,mx} - d\tau_{j,mx})]$$

$$= \sum_{i=1}^N C_H^{ij}(d\tau_{ij,mx}), \quad (12)$$

where C_H^{ij} is hybrid correlation waveform generated by correlating relative correlation waveforms C_R once more (see Figs 1 and 2 for more detail), S_H is the HCS value, $d\tau_{ij,mx} = d\tau_{i,mx} - d\tau_{j,mx}$ is the double differential traveltime from the two events to a receiver pair $\{i, j\}$. After doubling the relative correlation process, the unknown origin time in RCS (see eq. 10) is cancelled and the level of redundancy is increased. The vanishing of τ_0 can not only save computational effort, but also help to relieve the effects of the origin time-depth trade-off in RCS. Recently, Guo & Zhang (2017) proposed a double-pair DD location method, which exhibited smaller relative location uncertainty than original DD method. Our waveform-based HCS method shares the same essence with the double-pair DD method. Note that there is a ‘double’ in ‘double differential traveltime’ for HCS operator, as well as in the name of DD method. For HCS, the double differential traveltime means the difference between differential traveltime from the event pair to a common receiver and that to another common receiver, while in DD method double-difference stands for the difference between observed and theoretical differential traveltime from the event pair to a common receiver. The observed (differential) traveltime is expressed explicitly in traveltime inversion methods, but it is implicitly utilized associated with waveforms.

2.6 Generalization of correlation-based imaging methods

BF is a general term for phase-shifted summations over different stations in a monitoring array. The traditional BF algorithm stacks the energy of the summation over delayed signals from all stations (see eq. A1 in Appendix). The essence of BF algorithm is actually based on the DS operator. An important difference between BF and

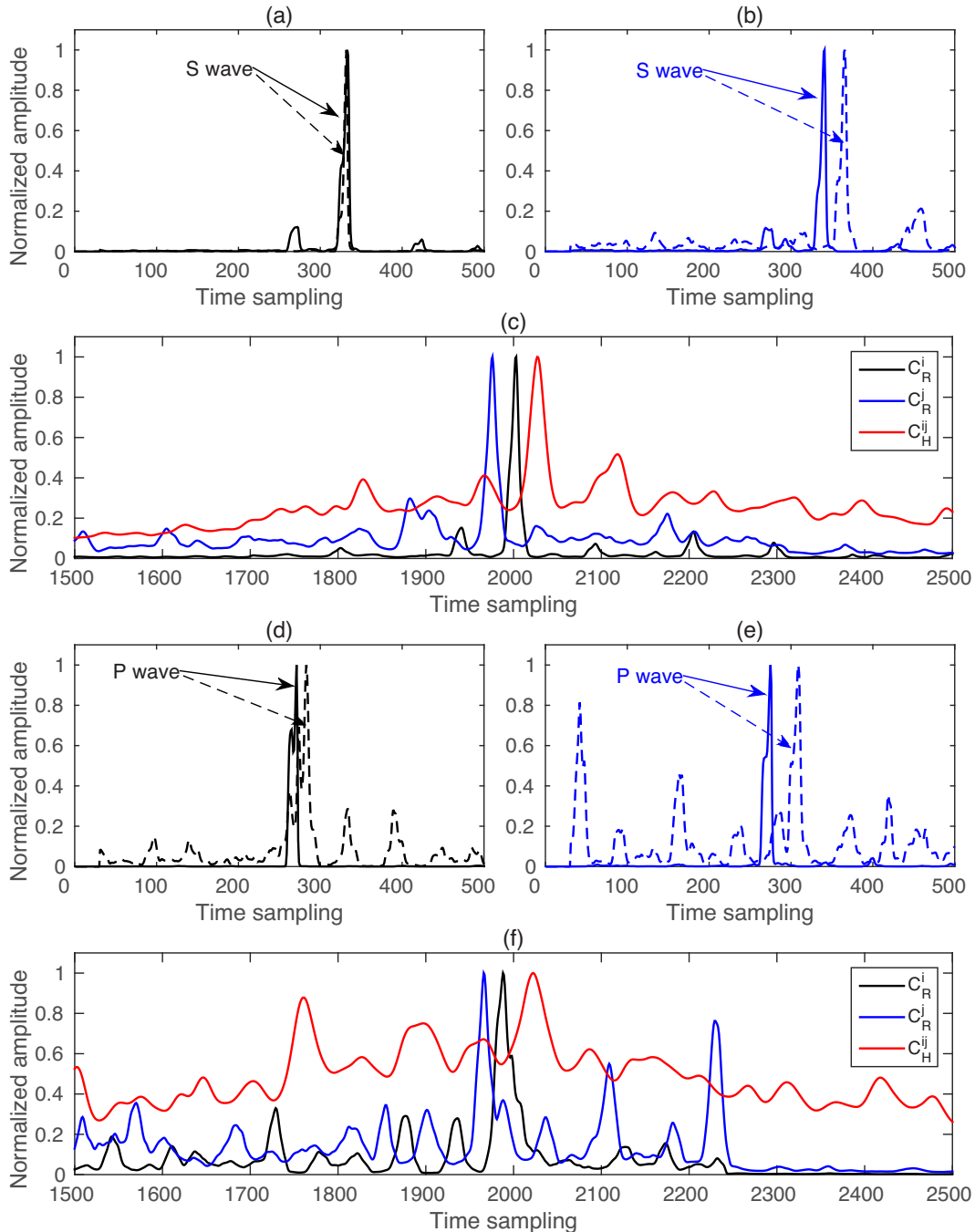


Figure 2. Generation process of hybrid correlation waveforms. (a,b) STA/LTA traces of horizontal components for a master event (solid black and blue lines) and a target event (dashed black and blue lines) at a station pair i and j ; (c) solid black and blue lines are relative correlation waveforms generated from the correlation of STA/LTA traces for an event pair at a common station i or j (see eq. 9), solid red line is hybrid correlation waveforms generated from the correlation of relative correlation waveforms at a station pair i and j (see eq. 11); (d–f) corresponding waveforms as in (a)–(c) but for vertical components. All the waveforms are normalized according to their maximum values. The noise from the input STA/LTA traces will be inherited after the first correlation process in RCS, and it is exaggerated in the second correlation process in HCS, which decreases the resolution of the final correlation waveforms (solid red lines in panels c and f), as well as the resolution of final imaging profile (see Fig. 7d).

other common stacking techniques is that there is a distinct summation over time samples in BF, which resembles the summation process in correlation algorithm. Consequently, the result of BF is equal to the sum of autocorrelations and appropriately delayed cross-correlations of the waveforms (see eq. 13, detailed derivation see Appendix), and the cross-correlation term dominates the BF image (Birchfield & Gillmor 2002; Poiata *et al.* 2016; Li *et al.* 2017a;

Ruigrok *et al.* 2017).

$$\begin{aligned}
 S_{\text{beam}}(\mathbf{x}) &= \sum_{t=0}^{t_{\max}} \left(\sum_{i=1}^N C F_i(t + \tau_{i,\mathbf{x}}) \right)^2 \\
 &\approx 2S_S + \sum_{t=0}^{t_{\max}} \sum_{i=1}^N (C F_i(t + \tau_{i,\mathbf{x}}))^2.
 \end{aligned} \tag{13}$$

Table 1. Comparison of imaging operators for waveform-based location methods. N is the number of receivers and M is the number of all unique receiver pairs.

Method	Imaging operator	Traveltime information	Basic imaging patterns in 2-D (3-D) scenario	Level of redundancy
SCS	$\delta[\tau - (\tau_{i,x} - \tau_{j,x})]$	differential traveltime from the common source to a receiver pair	hyperbolae (hyperboloids) intersections	M
DCS	$\delta[\tau - (d\tau_{I,x} - d\tau_{J,x})]$	double differential traveltime from the common source to receiver triplets and quadruplets	hyperbolae (hyperboloids) intersections and more redundancy	$M(M-1)/2$
RCS	$\delta[\tau - (\tau_0 + (\tau_{i,m} - \tau_{i,x}))]$	differential traveltime from a source pair to the common receiver	circular arcs (spherical surfaces) intersections	N
HCS	$\delta[\tau - ((\tau_{i,m} - \tau_{i,x}) - (\tau_{j,m} - \tau_{j,x}))]$	double differential traveltime from a source pair to a receiver pair	hyperbolae (hyperboloids) intersections	M
DS	$\delta[\tau - (\tau_0 + \tau_{i,x})]$	traveltime from a source to a receiver	circular arcs (spherical surfaces) intersections	N

Similarly, if we replace CF with cross-correlation waveforms C_S and C_R , the formula can be regarded as single correlation beamforming (SCBF, eq. 14) and relative correlation beamforming (RCBF, eq. 15). The corresponding dominating term changed to the DCS and HCS. In other words, the double correlation and hybrid correlation method are actually the simplified forms of SCBF and RCBF, respectively (see Appendix for more details):

$$\sum_{t=0}^{T_{\max}} \left(\sum_{I=1}^M C_S^I(t + \tau_{I,x}) \right)^2 \approx 2S_D + \sum_{t=0}^{T_{\max}} \sum_{I=1}^M (C_S^I(t + \tau_{I,x}))^2 \quad (14)$$

$$\sum_{t=0}^{T_{\max}} \left(\sum_{i=1}^N C_R^i(t + \tau_0 + d\tau_{i,mx}) \right)^2 \approx 2S_H + \sum_{t=0}^{T_{\max}} \sum_{i=1}^N (C_R^i(t + \tau_0 + d\tau_{i,mx}))^2. \quad (15)$$

The summation over time samples introduces much more stacks and decreases the vertical resolution in the final imaging volumes, and it explains the lower vertical imaging resolution of double correlation methods compared to single correlation methods. By unifying the similar formulae of these methods and describing cross-correlation stacking with BF algorithm, we can generalize correlation-based source imaging methods as a unified formula

$$S_{\text{corr}} = \sum_{I,J} \sum_t W_I(t) W_J(t + \Delta t_{IJ}) \approx \frac{1}{2} \sum_t \left[\sum_I W_I(t) \right]^2, \quad (16)$$

$$W = \begin{cases} CF, & \text{if } S_{\text{corr}} = S_S \text{ or } S_R, \\ C_S, & \text{if } S_{\text{corr}} = S_D, \\ C_R, & \text{if } S_{\text{corr}} = S_H, \end{cases} \quad (17)$$

where I and J indicate receivers or receiver pairs, W denotes the CFs of original waveforms or cross-correlation waveforms.

2.7 Imaging operator analysis

Based on above introduction and generalization of correlation-based imaging methods, now we can analyse the basic properties of their imaging operators. As introduced before, the four different correlation-based imaging methods utilize different traveltime information, which construct different imaging patterns and are summarized in Table 1. We will illustrate the imaging patterns and level of redundancy in more detail with a numerical example below.

For waveform-based location methods, the CFs are stacked or back-projected according to different imaging operators. Correspondingly, the imaging patterns are determined by the distribution of time residues between theoretical and observed traveltime information contained in the imaging operators. For simplicity and without loss of generality, illustrations of imaging patterns for these methods are based on a 2-D homogeneous model (Figs 3 and 4). In this example, four receivers are arranged at the surface and only primary P -wave is considered, the target source is set at the centre of the model and the master event for RCS and HCS is set at (250 m, 150 m). Figs 3(a)–(d) show the distributions of multiplication of time residues based on the imaging operators in Table 1. The time residues of all grid points for different receivers (in RCS), receiver pairs (in SCS and HCS), receiver triplets or quadruplets (in DCS) are multiplied. The logarithm is used to expose the very small difference in the time residues. These distributions of time residues exactly reflect the imaging patterns of the corresponding methods shown in Figs 4(a)–(d), which are generated with eqs (5), (8), (10) and (12), respectively. In this simple example, we just use the original waveforms as the input and only four receivers are considered, which are certainly much fewer than in the real cases of monitoring induced seismicity, but it is effective to expose the essence of these imaging operators and the limited number of receivers has no impact on the conclusion. According to the imaging operators in Table 1, we can also infer their imaging patterns from a mathematical point of view. For SCS and HCS, the variables in the imaging operators are the same differential traveltime term: $\tau_{i,x} - \tau_{j,x}$, which correspond to hyperbolae (hyperboloids) intersections in 2-D (3-D) scenario. Artifacts, which resemble the hyperbolic (hyperboloidal) forms, will be generated due to the non-uniqueness of the equal distance from the source(s) to the receiver pair. For RCS, the variable is the single traveltime term $\tau_{i,x}$ and thus it has the same imaging patterns like DS, which are circular arcs (spherical surfaces) intersections in 2-D (3-D) scenario. Although RCS and DS share the same basic imaging patterns and level of redundancy, the latter lack the additional constraints from the master event, which could improve the location results significantly. The double differential traveltime terms in the imaging operator of DCS include single differential traveltime terms in SCS and other differential traveltime terms of receiver triplets and quadruplets. For general 3-D models with more receivers, deformed hyperboloids and spherical surfaces intersections with highly focused source energy will be present in the final imaging profiles. A distinct difference between hyperbolae intersection and circular arcs intersection is that the latter exhibits much higher vertical imaging resolution. In general, the lateral imaging

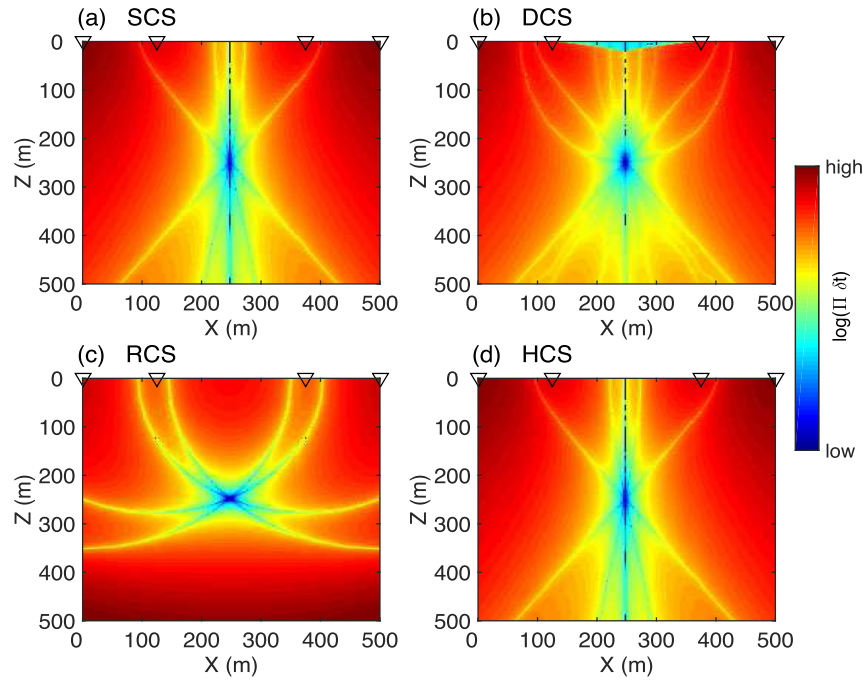


Figure 3. Distribution of traveltimе residues based on the imaging operators for correlation-based methods: (a) SCS; (b) DCS; (c) RCS; (d) HCS. The reverse triangles are receivers. The positions with lowest residue values denote the source locations.

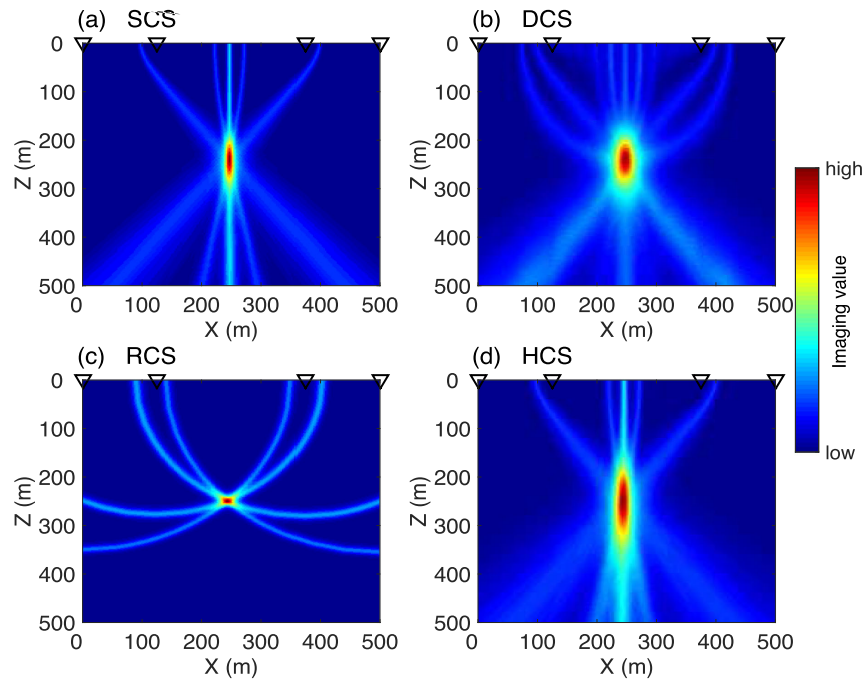


Figure 4. Imaging results of the correlation-based methods: (a) SCS; (b) DCS; (c) RCS; (d) HCS. The reverse triangles are receivers. The positions with highest imaging values denote the source locations.

resolution will be constrained well provided that the receivers cover enough lateral directions with respect to the epicentre of the subsurface source. The imaging patterns shown here just present the basic imaging resolution, since the final imaging resolution is also affected by other factors, such as receiver geometry and SNR of the data. The analysis of basic imaging patterns also reflects the redundancy exploited in these methods. If we use the number of stacks to represent the level of redundancy, the RCS and DCS extract the least and most redundancy, respectively. In this case, the number of receiver $N = 4$, the number of all unique receiver pairs $M = 6$, so

the level of redundancy in SCS, DCS, RCS and HCS are 6, 15, 4 and 6, respectively.

3 SYNTHETIC TESTS

We apply above correlation-based methods to a realistic synthetic data set based on the local and temporary HAMNET network for monitoring mining induced seismicity (Bischoff *et al.* 2010). The HAMNET network consists of 15 three-component surface stations and we manually generate 200 target events with random

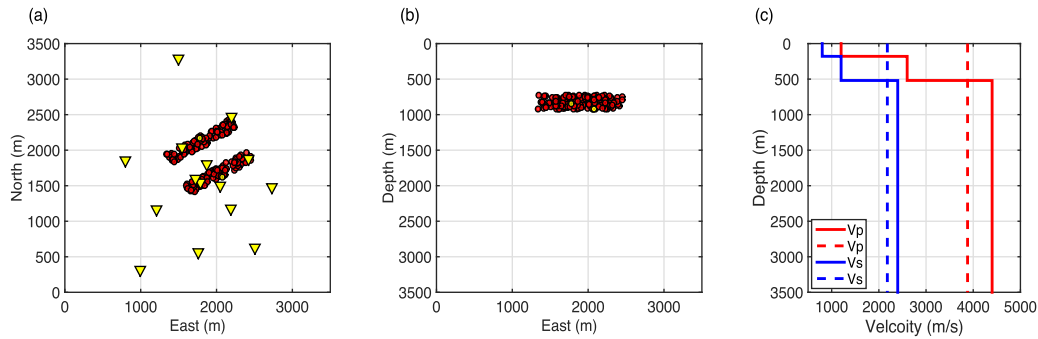


Figure 5. The true location of the 200 synthetic events and the velocity models used in this work. (a) Top view of acquisition surface, yellow reverse triangles are stations, red dots are target events and two yellow dots are master events for each 100 target events within the same line; (b) east-depth plane; (c) velocity models, solid lines are the layered model used for the synthetic events, dashed lines are the homogeneous model used for the synthetic examples and field data set. The reference point (east, north) = (0, 0) corresponds to east = 41117 and north = 5721611 in the UTM coordinate system.

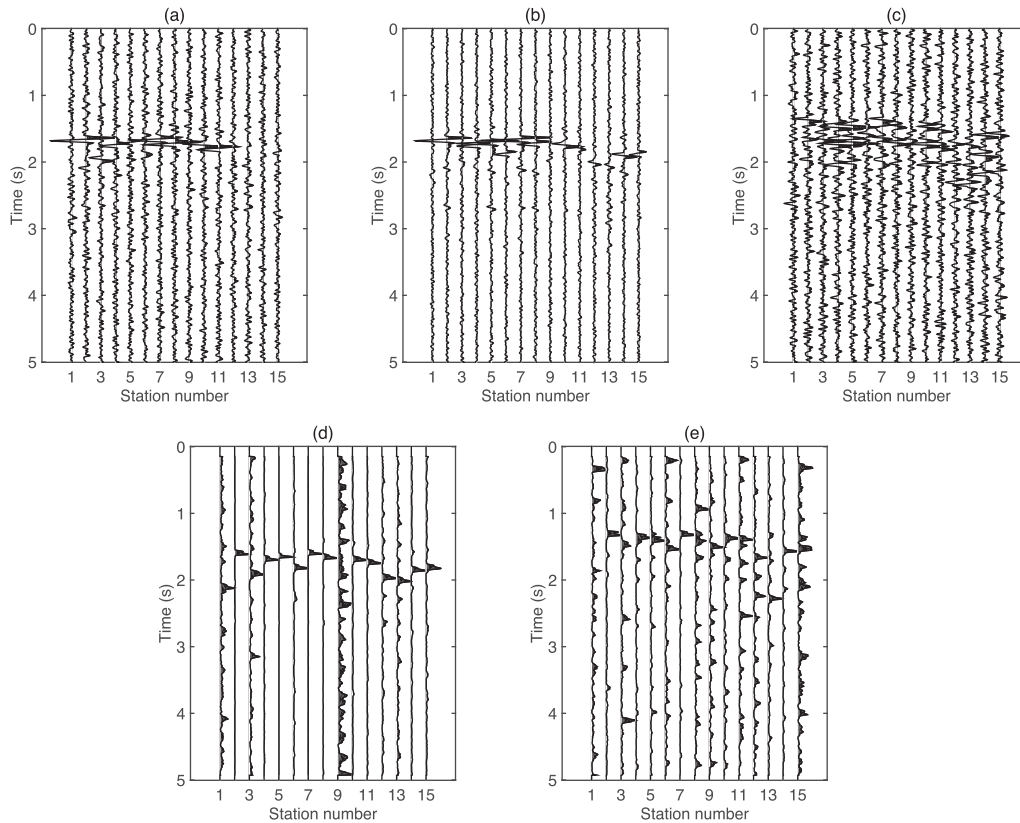


Figure 6. Three components of bandpass filtered (2–20 Hz) waveforms (a–c) and STA/LTA traces (d,e) of horizontal and vertical components for a sample synthetic event. (a–c) Waveforms of east, north and vertical components with normal distributed noise, waveforms are normalized to the maximum amplitude of each component; (d) STA/LTA traces of horizontal components (a,b); (e) STA/LTA traces of vertical components (c). STA/LTA traces are calculated according to eqs (1)–(3).

double-couple source mechanisms (the strike, dip, and rake of which are randomly distributed between $[70^\circ, 110^\circ]$, $[25^\circ, 65^\circ]$, and $[-70^\circ, -110^\circ]$, respectively) and an M_w magnitude of -0.1 . These events are randomly clustered along two parallel lines (Figs 5a and b) and there are 100 events in each line. A 1-D layered model (Fig. 5c) and the software Qseis (Wang 1999) are used to generate the synthetic waveforms, which are then contaminated by normal distributed noise with a standard deviation equal to 2 percent of the maximum amplitude for all channels (Figs 6a–c). The time sampling is 5 ms and the theoretical traveltime table is calculated by the Eikonal solver package FDTIMES (Podvin & Lecomte 1991) with the grid spacing of 25 m. For relative meth-

ods, a master event is selected arbitrarily for each 100 target events within the same line (yellow dots in Figs 5a and b), and no noise is added to the waveforms of master events, indicating the events with high-SNR are selected as master events. The target imaging volume is $3.5 \text{ km} \times 3.5 \text{ km} \times 3.5 \text{ km}$. The STA/LTA traces (Figs 6d and e) are calculated according to eqs (1)–(3). Selection of processing parameters, including filtering parameters, length of short and long time windows, and recognition of seismic phases contained in the waveforms, depends on source characteristics and monitoring purposes, and this can be done by pre-processing of raw data. Fig. 2 shows the generation process of hybrid correlation waveforms.

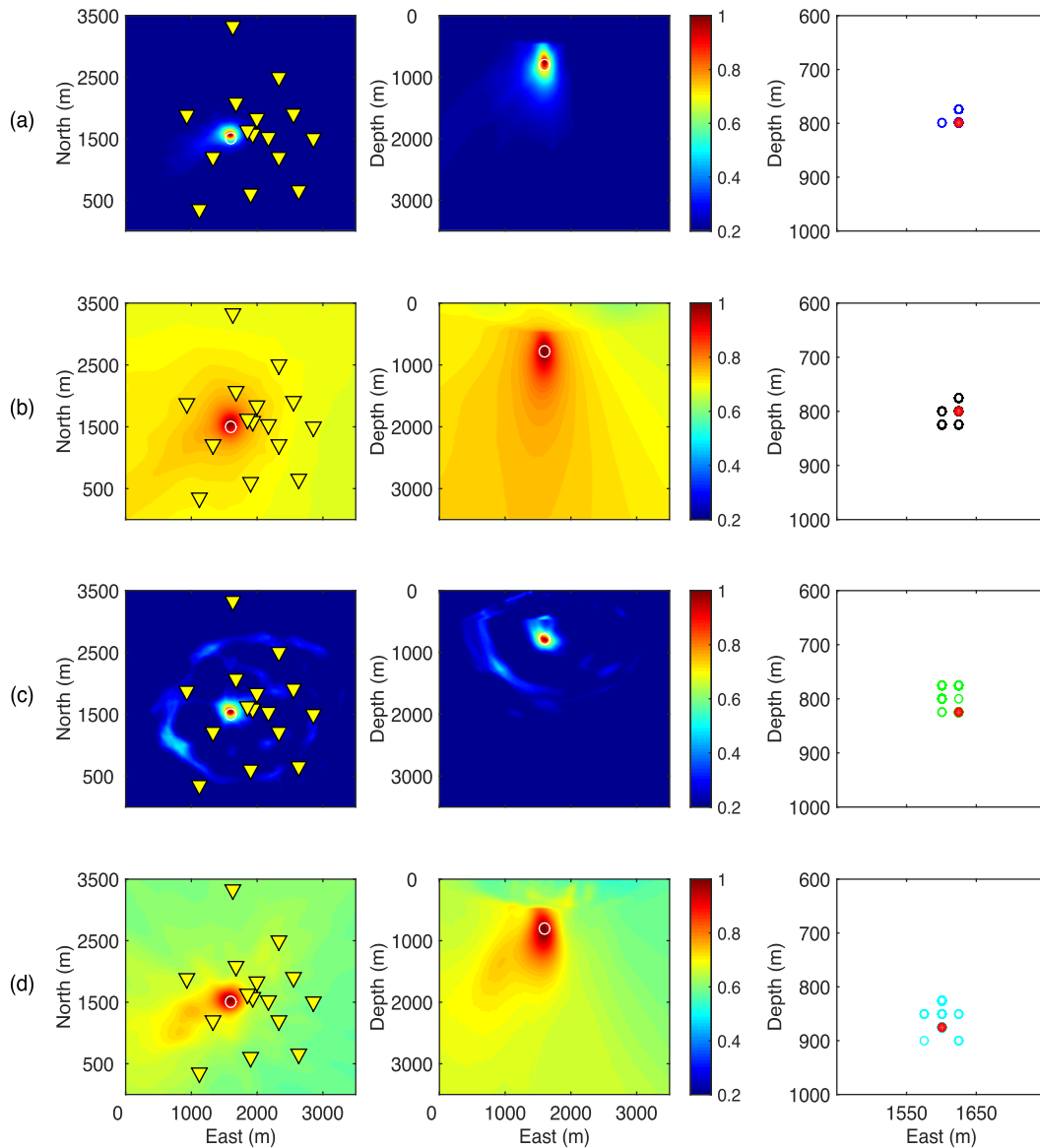


Figure 7. Imaging results (the left and middle columns) and the Jackknife test results (the right column) of the sample synthetic event shown in Fig. 6: (a) SCS; (b) DCS; (c) RCS; (d) HCS. Only strong S waves are considered here. White circles are the true locations of the sample event. The imaging results are squared and normalized according to individual maximum values. The red dots of the Jackknife test are the locations using all stations and circles are locations by randomly removing two stations at each time. Note the relatively low imaging resolution in panels (b) and (d), which indicates the noise from the input waveforms will be inherited and exaggerated in the final imaging profile after double correlation process (see Figs 2c and f). The Jackknife test also reveals the higher uncertainty of double correlation-based methods.

Since the horizontal components have strong and dominant S waves, we first consider S waves only, namely, using the CF of horizontal energy CF_H in eq. (1) only. The imaging results of the sample event in Fig. 6 are shown in Fig. 7. The computation time of SCS, DCS, RCS and HCS for this single event are about 5, 250, 130 and 6 s, respectively (serial codes on a computer with four cores of Intel i5-2500 3.30 GHz and 8 GB RAM), and their levels of redundancy are 105, 5460, 15 and 105, respectively. The DCS needs more computation time compared to SCS since it involves much more stacks, while HCS increases the computational efficiency significantly compared to RCS by avoiding searching the origin time. As analysed in Section 2.7, the lateral imaging resolution for these methods are quite well due to the proper lateral coverage of surface monitoring stations, and RCS exhibits the best vertical imaging resolution compared to the other three methods, the imaging patterns

of which are mainly constructed by deformed hyperboloids intersections. The higher level of redundancy in DCS and HCS compared to their counterparts (i.e. SCS and RCS, respectively) introduces more noise after the second correlation process using non-negative STA/LTA traces (see Fig. 2), and thus decreases the imaging resolution with this dense monitoring network. Then we locate this event by considering both P waves and S waves, namely, utilizing multiple traveltime information along with two separate CFs in eqs (1)–(3). The imaging functions (i.e. eqs 5, 8, 10 and 12) discussed in Section 2 change from the stacking along differential traveltime curves for single phase to stacking of products of those along differential traveltime curves for multiple phases. Fig. 8 shows the corresponding imaging results and the vertical imaging resolution for all methods are slightly improved.

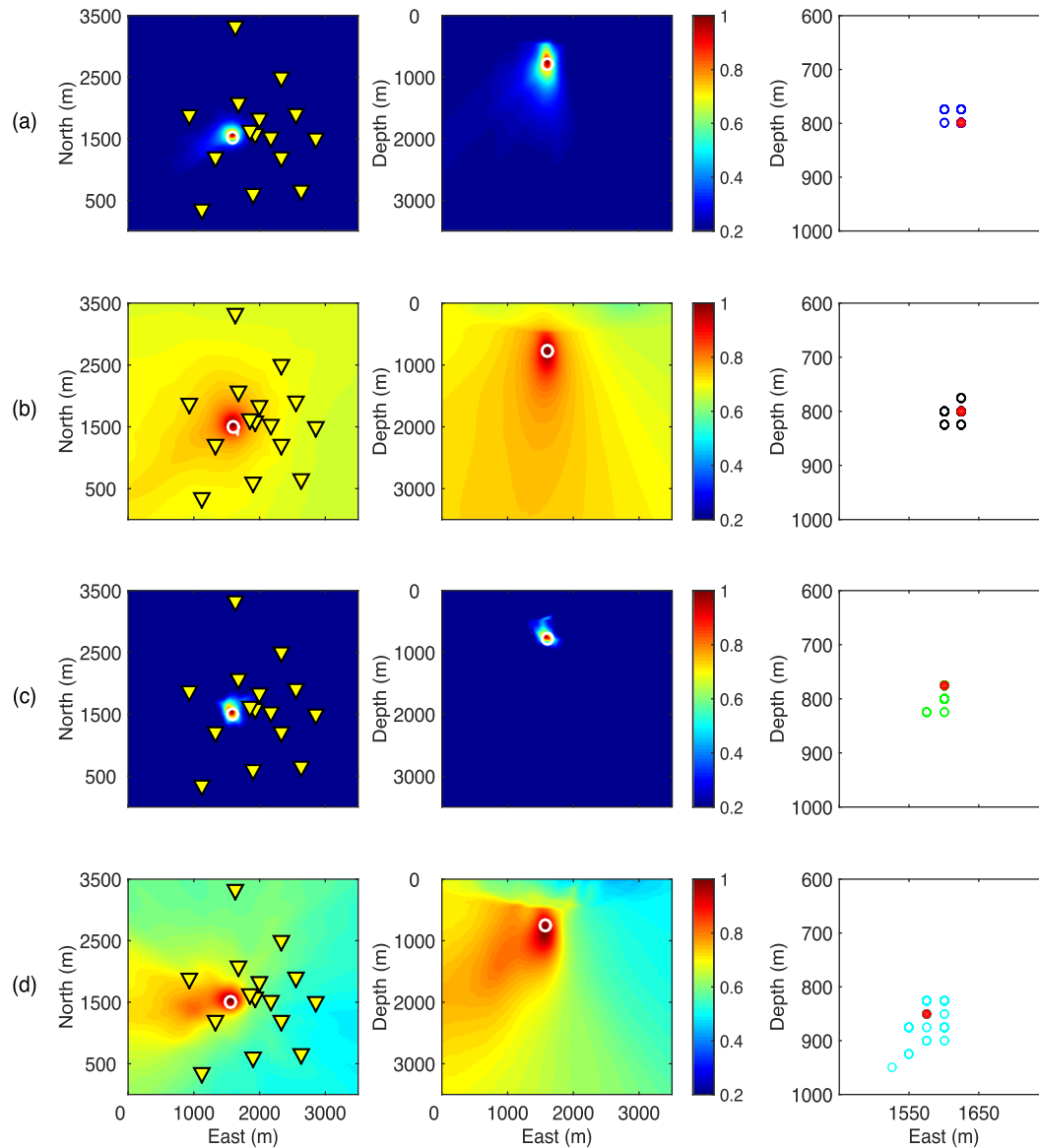


Figure 8. Imaging results (the left and middle columns) and the Jackknife test results (the right column) of the sample synthetic event shown in Fig. 6: (a) SCS; (b) DCS; (c) RCS; (d) HCS. Both P waves and S waves are considered here. White circles are the true locations of the sample event. The imaging results are squared and normalized according to individual maximum values. The red dots of the Jackknife test are the locations using all stations and circles are locations by randomly removing two stations at each time.

In order to estimate the location uncertainty, we use a Jackknife test to locate the same sample event 200 times by removing two stations at each iteration. The results of Jackknife tests for considering single phase and multiple phases are shown in the right column of Figs 7 and 8. The standard deviation of Euclidean distance among the 200 locations is taken as the uncertainty of the event (Li *et al.* 2016b). When using single phase only, the uncertainties of this event for SCS, DCS, RCS and HCS are 8.39, 20.05, 26.89 and 27.61 m, respectively. The corresponding uncertainties in Fig. 8 are 14.07, 18.25, 15.74 and 34.42 m, respectively. The uncertainties are quite small and are about the size of one grid spacing. The uncertainties of double correlation methods are larger than their counterparts, which is consistent with the imaging results, although HCS is origin time-independent compared with RCS. Using multiple phases decreases the uncertainty of DCS but increases the uncertainty of HCS, which indicates the combined effects of multiple phases and high level of redundancy cannot ensure more stable location results.

The methods are then applied to truncated waveform segments corresponding to the 200 synthetic events, which have the same length of 5 s. The location results for these methods are shown in Fig. 9 and the detailed location errors are presented in Fig. 10. The DS method is also included for comparison and is represented by the red lines in Fig. 10. In general, both absolute and relative methods perform well in this ideal case, both the lateral and vertical positions are located very well, and 80 per cent of the events are located within 75 m from their true locations. The extremely high level of redundancy in DCS introduces extra noise and uncertainty, which are responsible for the relatively lower location accuracy. Then we locate the 200 events by considering both P waves and S waves. According to the location errors in Fig. 11, location results of most events are still good after including P waves contained in more noisy vertical energy (Figs 6c and e). Although using multiple phases introduces more constraints and tends to improve the location results, the location of using single phase is as accurate as that of using multiple phases in this ideal case.

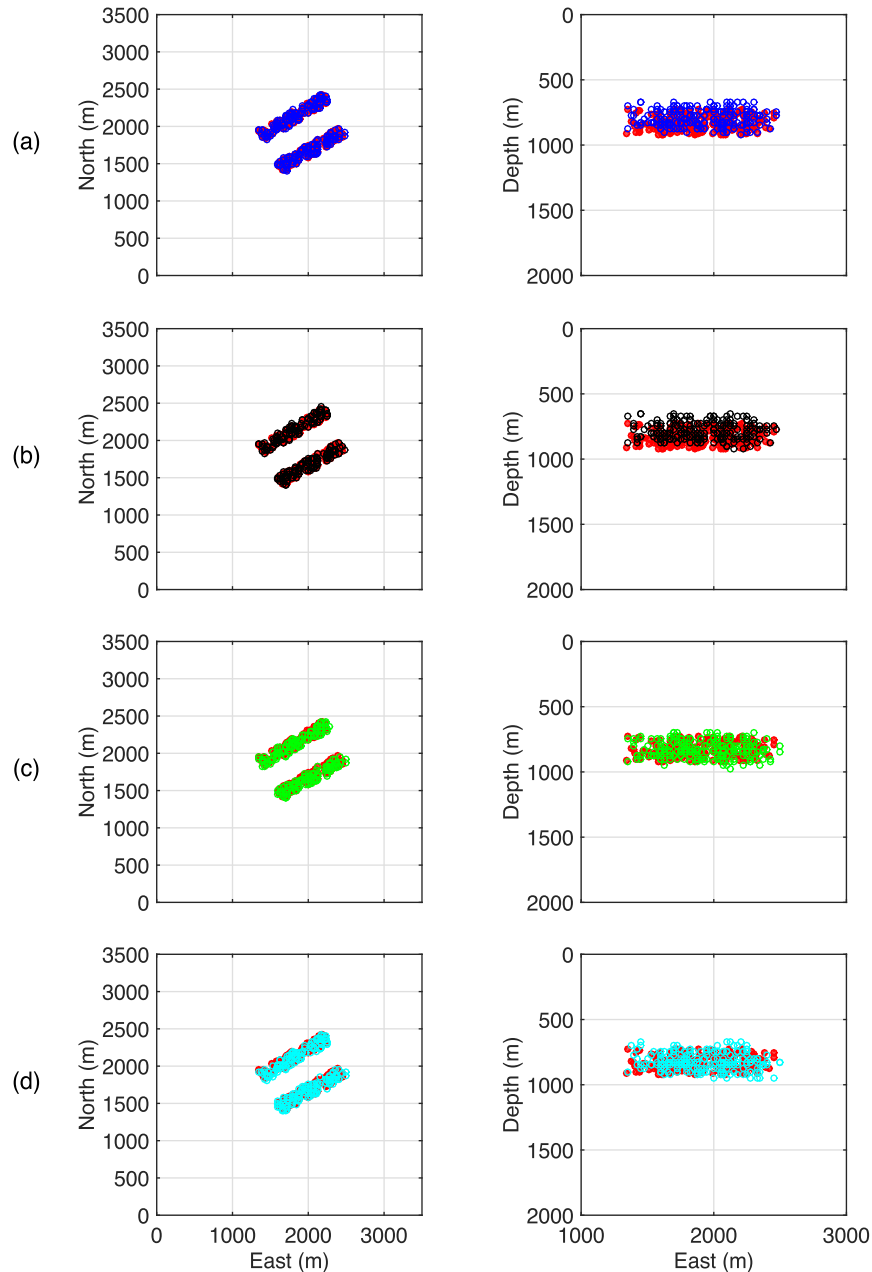


Figure 9. Location results of the 200 synthetic events using strong S waves only. The red dots are true locations and circles with blue, black, green, and cyan colours are results of SCS (a), DCS (b), RCS (c) and HCS (d), respectively.

Then we test the robustness of these methods with an inaccurate homogeneous model (see dashed lines in Fig. 5c). The location results using S waves only and using both P waves and S waves are shown in Figs 12 and 13, respectively. Compared to the results using the true layered model (Figs 10 and 11), the lateral positions are still well resolved for all correlation-based methods due to the good lateral coverage of the monitoring stations. Utilizing multiple phases with an inaccurate velocity model just deteriorates the results for all methods (see Figs 10 and 13), especially for origin time-dependent and low level of redundancy based DS method. This demonstrates the potential disadvantages of using multiple phases with unreliable multiple velocity models. As expected, the relative methods perform much better since they are less sensitive to traveltime errors, which is contributed by the cancellation of nearly equal traveltimes (coincident ray paths) of the master and target events.

Finally, we test these methods with more noisy synthetic waveforms. The standard deviations of the noise are increased to 5 per cent and 8 per cent of the maximum amplitude. We locate the 200 events with different noise levels by using S waves only and both P waves and S waves. Here the true layered model is used and the location errors are shown in Fig. 14. Using the dominant phase only can even obtain better results than using multiple phases, since interferences from multiple phases will overweight their additional constraints when assumed multiple phases are weak or submerged in noisy data. When using multiple phases in noisy data, the methods with low level of redundancy (i.e. DS and RCS) perform worse than others, while an extremely high level of redundancy (i.e. DCS) just deteriorates the location results severely (see Figs 14c and d) and most of which converge to computational boundary due to the very low resolution of double correlation waveforms. The results

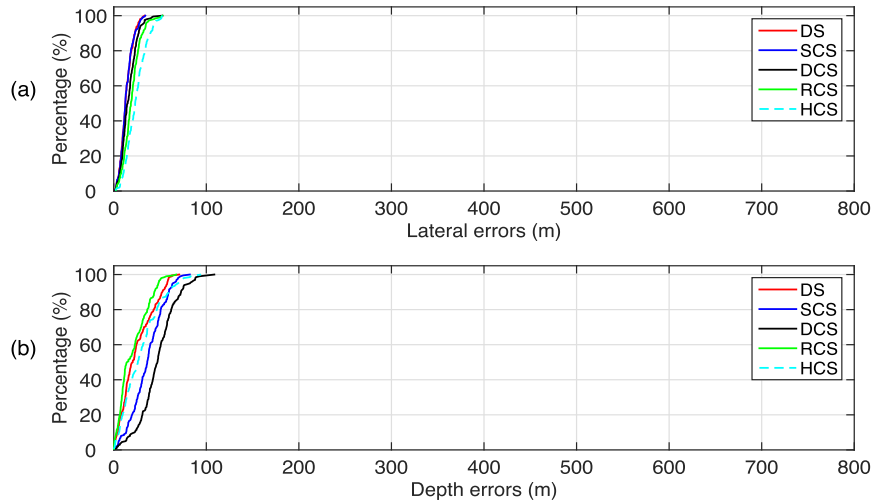


Figure 10. Location errors of the 200 synthetic events for different methods with the true layered model using strong S waves only. (a) Relation between lateral errors and percentage of event number; (b) relation between depth errors and percentage of event number.

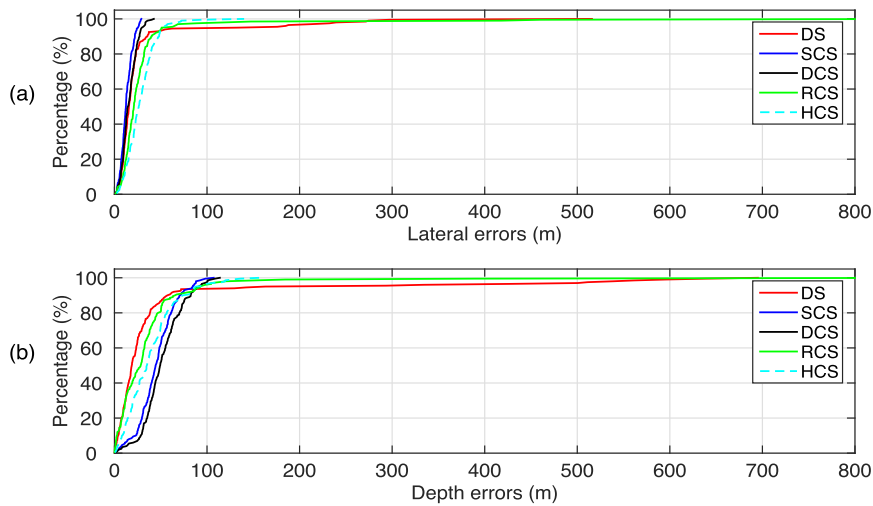


Figure 11. Location errors of the 200 synthetic events for different methods with the true layered model using both P waves and S waves. (a) Relation between lateral errors and percentage of event number; (b) relation between depth errors and percentage of event number.

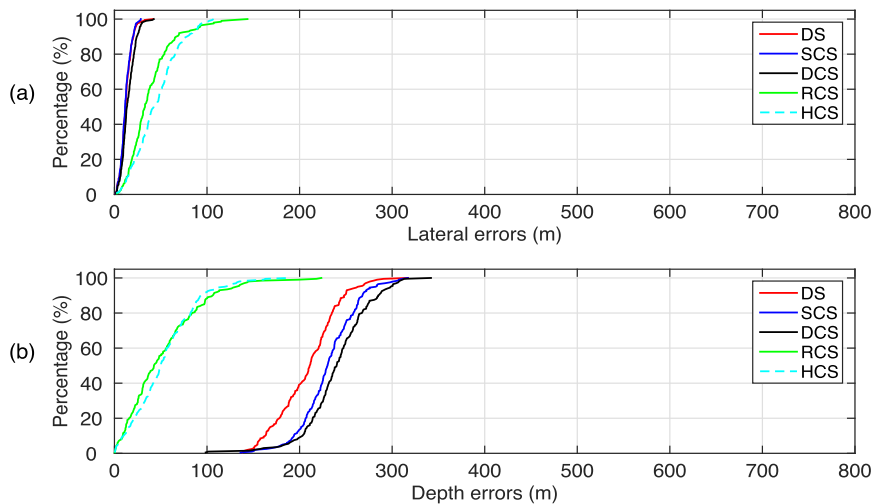


Figure 12. Location errors of the 200 synthetic events for different methods with the homogeneous model using strong S waves only. (a) Relation between lateral errors and percentage of event number; (b) relation between depth errors and percentage of event number.

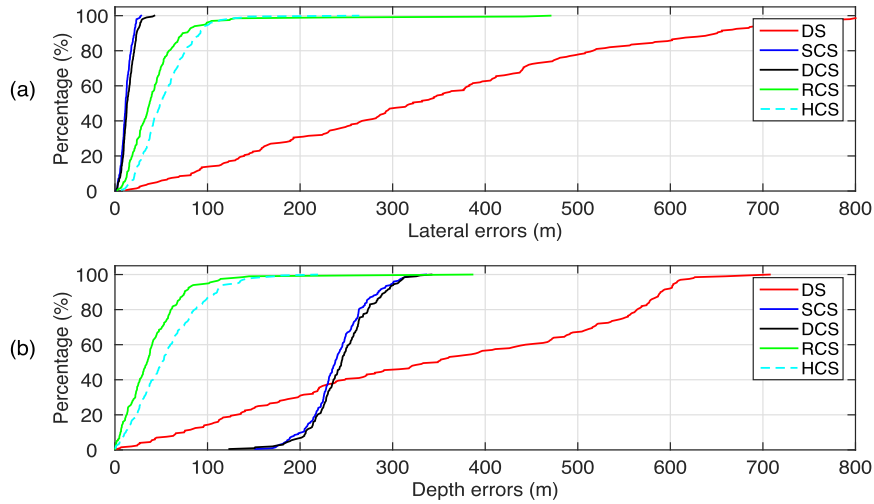


Figure 13. Location errors of the 200 synthetic events for different methods with the homogenous model using both *P* waves and *S* waves. (a) Relation between lateral errors and percentage of event number; (b) relation between depth errors and percentage of event number.

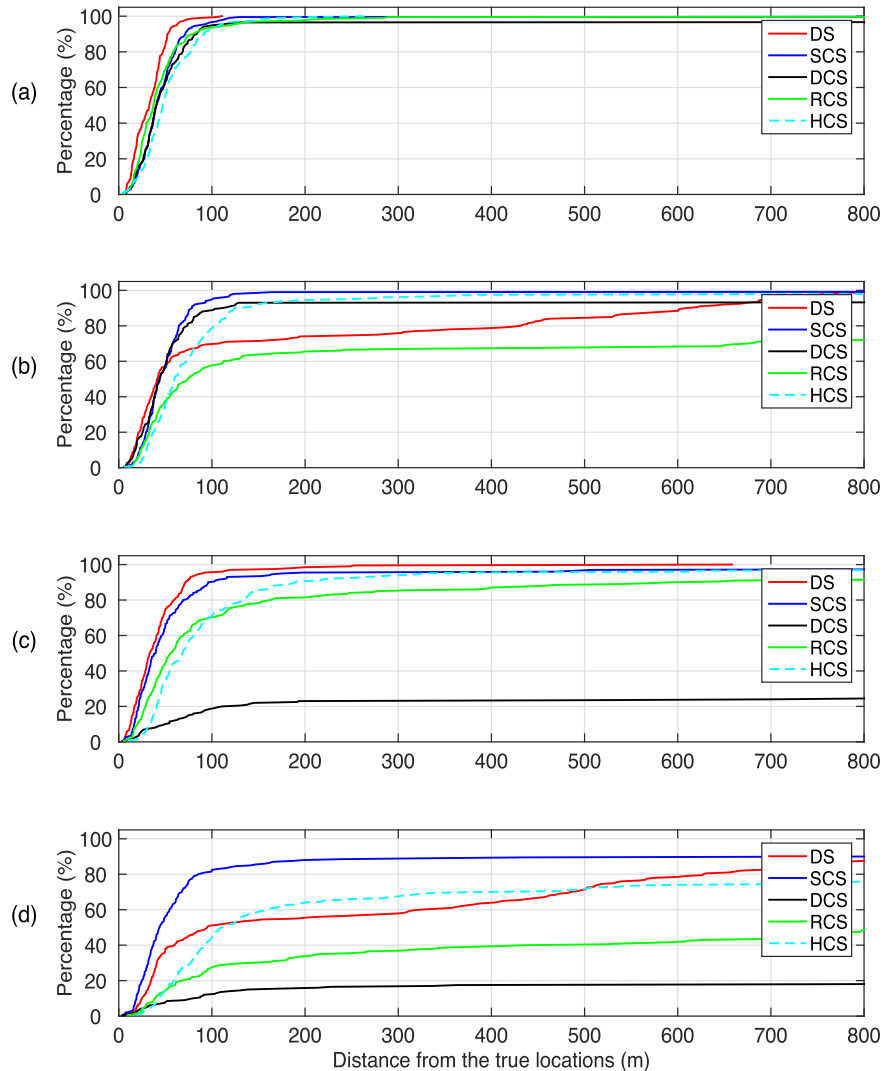


Figure 14. Location errors of the 200 synthetic events with two noise levels with the true layered model using strong *S* waves only (a,c) and both *P* waves and *S* waves (b,d). (a,b) Relation between location errors and percentage of event number for events with noise level 5 per cent; (c,d) Relation between depth errors and percentage of event number for events with noise level 8 per cent.

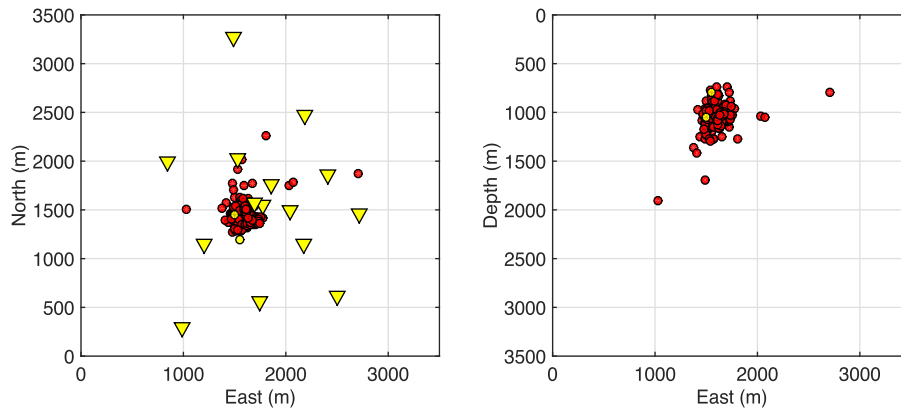


Figure 15. The location of the 200 field events by travelt ime inversion. Yellow reverse triangles are stations, red dots are target events and two yellow dots are master events for each 100 target events with the same magnitude range. The reference point (east, north) = (0, 0) corresponds to east = 41117 and north = 5721611 in the UTM coordinate system.

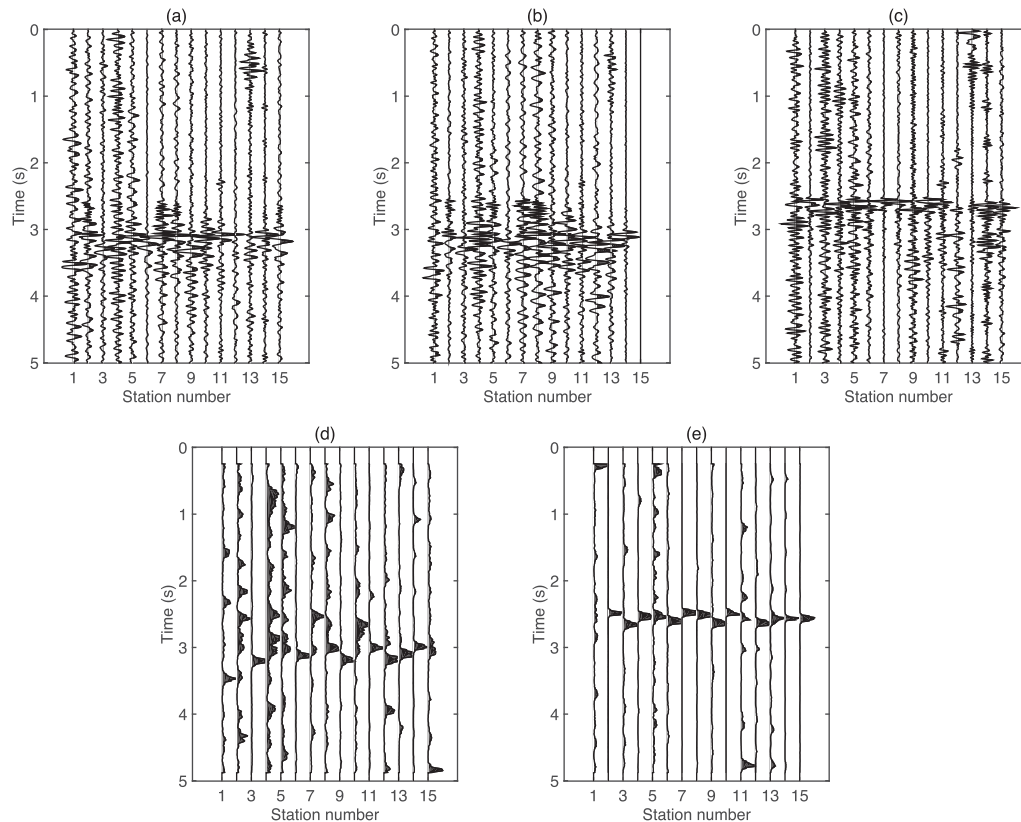


Figure 16. Three components of original seismograms (a–c) and STA/LTA traces (d,e) of horizontal and vertical components for a sample field event. (a–c) Waveforms of east, north and vertical components, waveforms are normalized to the maximum amplitude of each component; (d) STA/LTA traces of horizontal components (panels a and b); (e) STA/LTA traces of vertical components (panel c). STA/LTA traces are calculated using the band-pass filtered (5–30 Hz) seismograms according to eqs (1)–(3).

suggest that DCS is not suitable for very noisy data and methods with a moderate level of redundancy (i.e. SCS and HCS) can ensure the location accuracy no matter multiple phases are utilized or not.

4 APPLICATION TO FIELD DATA

These correlation-based imaging methods are then applied to locate 200 field events associated with mining induced seismicity monitored by the HAMNET network. We select 100 strong events (M_L magnitude > 0.5) and 100 weak events (M_L magnitude =

–0.8) from the data set (Fig. 15). For relative methods, a master event is selected by referring to location accuracy of absolute methods for each 100 target events with the same magnitude range. We set the target imaging volume as $5 \text{ km} \times 5 \text{ km} \times 5 \text{ km}$ with the grid spacing of 50 m. Converse to the synthetic example, there are relatively strong P -wave energy in the vertical components for most selected events, we use P waves only at first and then compare the result with that of using multiple phases. Fig. 16 shows the seismograms and STA/LTA traces of a sample field event, which has a magnitude of –0.8. We use the same homogeneous velocity model (see dashed lines in Fig. 5c) as in travelt ime inversion to

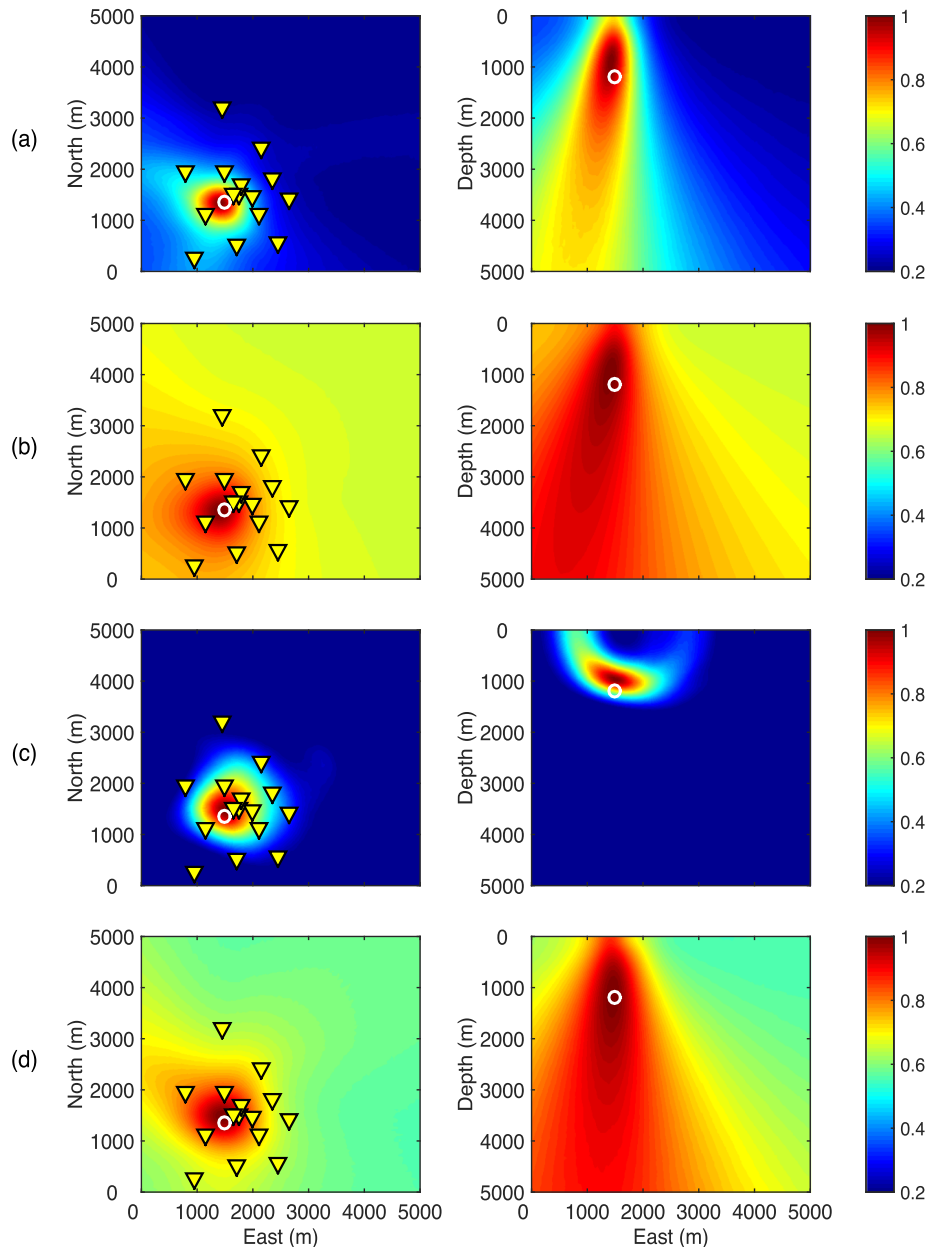


Figure 17. Imaging results of the sample field event shown in Fig. 16. Only strong P waves are considered here. White circles are the locations of the sample event by travelt ime inversion: (a) SCS; (b) DCS; (c) RCS; (d) HCS. The imaging results are squared and normalized according to individual maximum values.

locate these events with correlation-based methods. Fig. 17 shows the imaging results of the sample event. Similar to imaging results in Fig. 7, RCS exhibits the best vertical imaging resolution, the higher level of redundancy in DCS and HCS results in lower imaging resolutions compared to their counterparts. The uncertainties of this sample event for SCS, DCS, RCS and HCS are 86.56, 87.07, 85.97 and 96.26 m, respectively. The uncertainties are larger than in the synthetic examples and their differences are smaller, which are caused by the simplified homogeneous model and the larger size of grid spacing.

Then we compare the location results of the 200 events between travelt ime inversion (Fig. 15) and correlation-based methods. In general, when using single phase only, there are good agreements between the source locations and clustering of travelt ime inversion and correlation-based methods (Fig. 18). The location biases for different methods with respect to travelt ime inversion are shown

in Fig. 19. The relative methods still perform better than absolute methods for most of the events. We can find that about 80 per cent of the events are located within 300 m from the manual locations when using the dominant P waves only (Fig. 19b), and they perform better than DS in this case. There is smaller difference between the results of relative and absolute location methods in the field data example (Fig. 19) than in the synthetic example (Figs 12 and 13). Here we use the location results from picking-based travelt ime inversion as references, which also contain possible errors, thus the comparison is not as fair as in the synthetic example. Another possible reason is that the accuracy of master events in the field data example is not ensured, and it would affect the results of relative location methods. Using multiple phases with a simplified homogeneous model deteriorates results of most events (Fig. 19d), which is consistent with the result in the synthetic example (Fig. 14). Both the synthetic and field data examples suggest that locating weak seismic sources

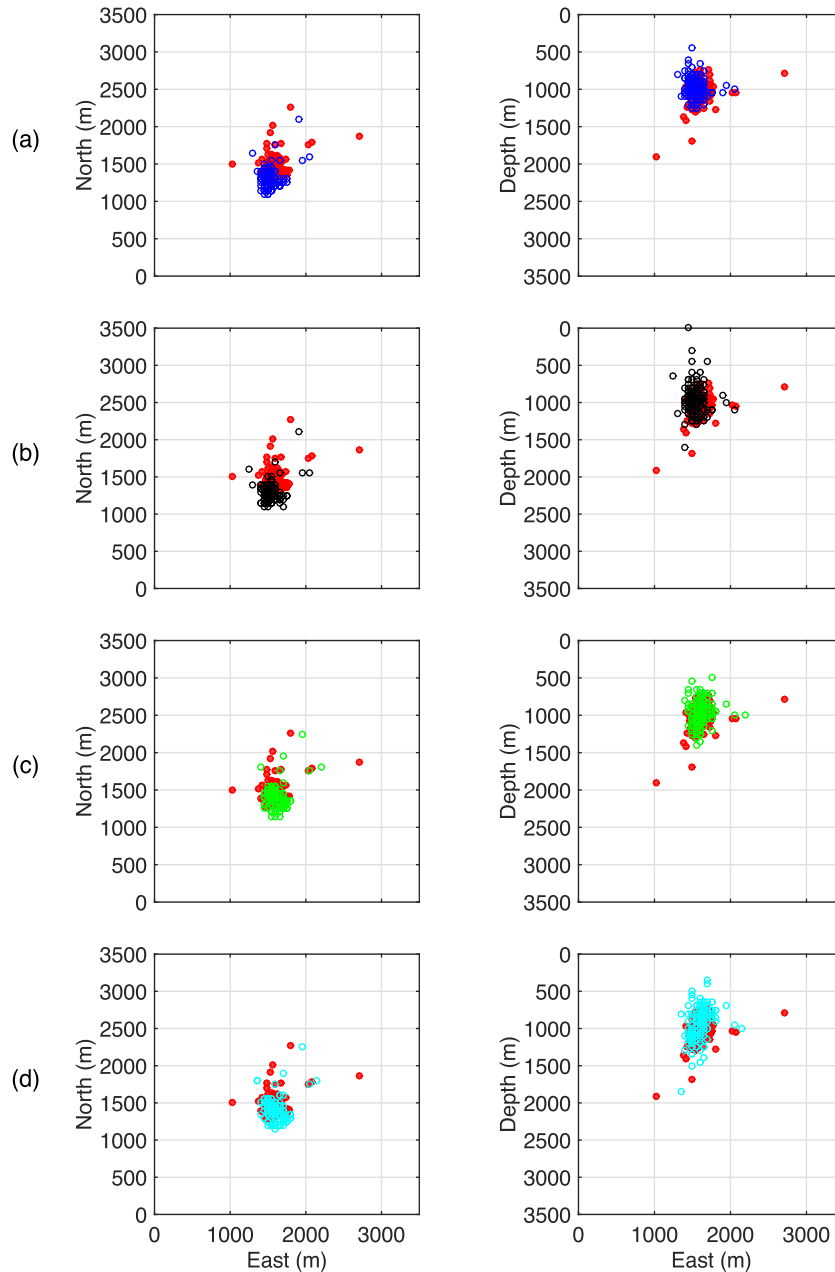


Figure 18. Location results of the 200 field events using strong P waves only. The red dots are location results by traveltime inversion and circles with blue, black, green, and cyan colours are results of SCS (a), DCS (b), RCS (c), and HCS (d), respectively.

by stacking the dominant phase seems to be a practical and safe strategy.

5 DISCUSSION AND CONCLUSIONS

We review and compare correlation-based imaging methods for automated seismic source location, including a novel HCS method which belongs to waveform-based relative location methods. In this method, the double differential traveltime from a source pair (i.e. a master event and a target event) to a receiver pair is used to stack the corresponding double correlation waveforms. The proposed HCS method proves to be a reliable and efficient approach for seismic location, particularly in case of a less reliable velocity model. We summarize these correlation-based methods using a unified formula

and demonstrate their generalized relationship by approximation with the BF algorithm. Then, a thorough analysis of the imaging operators for these methods is given to expose basic characteristics of their imaging resolution and level of redundancy.

Synthetic and field data examples associated with mining induced seismicity and a surface monitoring array consisting 15 stations demonstrate the feasibility of these automated methods. The four correlation-based methods can be categorized in two ways. On one hand, SCS and RCS are single correlation-based methods, DCS and HCS are involved with double correlation waveforms. The double correlation-based methods extract more redundancy while exhibiting lower imaging resolution and higher location uncertainty. The extremely high level of redundancy in DCS makes it unsuitable for very low-SNR data. In addition, SCS and HCS have much higher computational efficiency than the other two methods. On the other

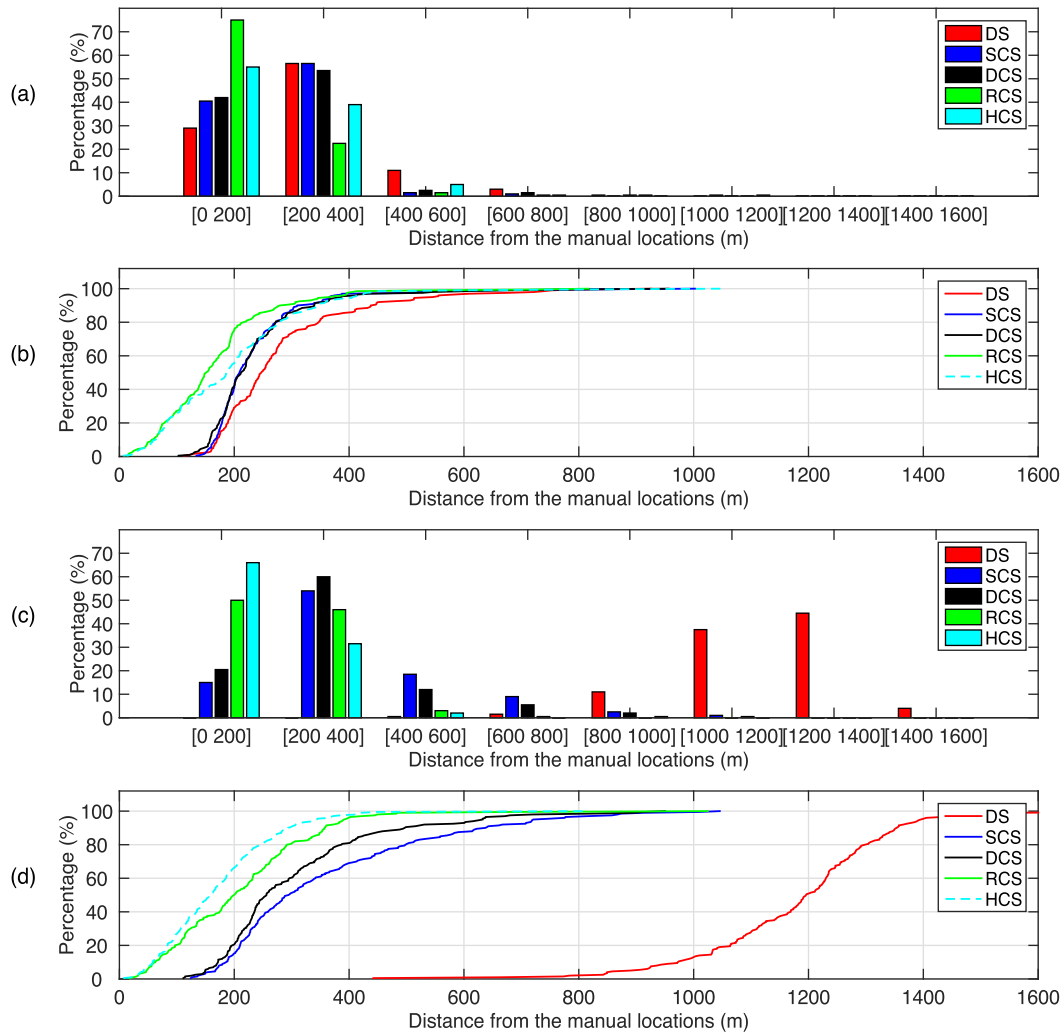


Figure 19. Location biases of the 200 field events for different methods with the homogeneous model. (a) and (c) are histograms of location biases using strong P waves only and both P waves and S waves, respectively; (b) and (d) are the corresponding relations between location biases and percentage of event number.

hand, SCS and DCS are absolute location methods, RCS and HCS are relative location methods. The relative location methods are good alternatives when considering the velocity uncertainty, since they are less sensitive to velocity errors due to extra constraints from the adjacent, well-located master events. Moreover, the examples also demonstrate the potential disadvantages of using multiple phases, since the interferences from multiple phases will outweigh their additional constraints when assumed multiple phases are weak or submerged in noisy data, or inaccurate multiple velocity models are involved. To sum up, a moderate level of redundancy (e.g. in SCS) can ensure both the accuracy and stability of location results, while an extremely high (e.g. in DCS) or low (e.g. in DS) level of redundancy will hinder the performance of waveform-based methods in locating weak seismic events. It is worth pointing out that the performance of these methods may be different when considering a sparser monitoring array with fewer stations or a denser array with more stations. For instance, the difference of both the computational efficiency and imaging resolution between single and double correlation methods will vary as the number of station changes.

For relative methods, the accuracy of the location of the master events has dominating effects on the location of target events. We use a different master event for a different subset of target events in this work, a modified version could include multiple mas-

ter events for a single target event, in which the contribution of multiple master events are properly weighted (Grechka *et al.* 2016; Grigoli *et al.* 2016). Moreover, when combined with waveform-based methods, the quality of their seismograms is also strongly related to the location result by affecting the stacking and imaging process, which needs further investigation. Generally speaking, multiple, well-located events with high-SNR should be selected as master events for waveform-based relative methods.

The location accuracy and the location uncertainty estimated by Jackknife tests are analysed to compare different methods. Actually, the location uncertainty generally means the certainty (stability or precision) of the location result. Although a Bootstrap or Jackknife method can be used to extract some statistical information of the location results (Grigoli *et al.* 2013; Li *et al.* 2016b), and the imaging resolution or distribution (see Fig. 4) also provides partial knowledge about the location uncertainty (Kao & Shan 2004; Anikiev *et al.* 2014), we argue location uncertainty for waveform-based location methods still need further study since the velocity uncertainty should also be considered, and it is the main contribution to location uncertainty (Gesret *et al.* 2015). Moreover, from a physical point of view, the prevailing wavelength or the size of the Fresnel volume (along with the size of grid spacing) should also be considered for waveform-based location methods using band-limited waveforms.

They determine the upper limit of the imaging resolution or uncertainty of a source image.

The traditional location procedure involves using manual phase picking and thus exhibits a challenge for real-time monitoring of induced seismicity, and the waveform-based methods provide automated alternatives for the processing routine (Grigoli *et al.* 2017). However, the computation efficiency is also an issue for these methods. For instance, the computation time of DCS and RCS is much larger than their counterparts, and this issue is even more exposed when encountering large monitoring arrays with hundreds of receivers. A promising solution is incorporating stochastic global optimization algorithms with waveform-based methods (Gharti *et al.* 2010; Li *et al.* 2017c). These algorithms can help to converge the imaging process very fast and reserve the event clustering when hundreds of events are considered.

ACKNOWLEDGEMENTS

We are grateful to HAMNET Project and Monika Bischoff for providing the data set. We thank the Applied Seismics Group Hamburg for continuous discussion and sponsors of the Wave Inversion Technology (WIT) Consortium. We thank Francesco Grigoli and an anonymous reviewer for their helpful comments and suggestions, as well as Ivan Abakumov for carefully verifying the formulae. The work is also sponsored by National Natural Science Foundation of China (Grant Nos. 11374322 and 11574347). Lei Li is also grateful to China Scholarship Council (CSC) for partially funding this work. The work of Dirk Becker is funded by the BMBF 'Geotechnologien' project SIMULTAN (Grant No. BMBF03G0843B).

REFERENCES

- Allen, R.V., 1978. Automatic earthquake recognition and timing from single traces, *Bull. seism. Soc. Am.*, **68**(5), 1521–1532.
- Anikiev, D., Valenta, J., Staněk, F. & Eisner, L., 2014. Joint location and source mechanism inversion of microseismic events: Benchmarking on seismicity induced by hydraulic fracturing, *Geophys. J. Int.*, **198**(1), 249–258.
- Artman, B., Podladtchikov, I. & Witten, B., 2010. Source location using time-reverse imaging, *Geophys. Prospect.*, **58**(5), 861–873.
- Baker, T., Granat, R. & Clayton, R.W., 2005. Real-time earthquake location using Kirchhoff reconstruction, *Bull. seism. Soc. Am.*, **95**(2), 699–707.
- Behzadi, M., Gajewski, D. & Vanelle, C., 2015. Seismic event localization based on cross-correlation stacking, WIT Report No. 18, pp. 217–230.
- Birchfield, S.T. & Gillmor, D.K., 2002. Fast Bayesian acoustic localization, in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Orlando, Florida.
- Bischoff, M., Fischer, L., Wehling-Benatelli, S., Fritschen, R., Meier, T. & Friederich, W., 2010. Spatio-temporal characteristics of mining induced seismicity in the eastern Ruhr-area, in *Cahiers du Centre Europe en de Geodynamique et de Seismologie*, vol. 30, eds Ritter, J. & Oth, A., The European Center for Geodynamics and Seismology (ECGS).
- Cesca, S. & Grigoli, F., 2015. Full waveform seismological advances for microseismic monitoring, *Adv. Geophys.*, **56**, 169–228.
- Chambers, K., Dando, B.D., Jones, G.A., Velasco, R. & Wilson, S.A., 2014. Moment tensor migration imaging, *Geophys. Prospect.*, **62**(4), 879–896.
- Dales, P., Audet, P., Olivier, O. & Mercier, J.-P., 2017. Interferometric methods for spatio-temporal seismic monitoring in underground mines, *Geophys. J. Int.*, **210**(2), 731–742.
- Drew, J.W., Robert, S., Tilmann, F. & Tarasewicz, J., 2013. Coalescence microseismic mapping, *Geophys. J. Int.*, **195**(3), 1773–1785.
- Droznin, D.V., Shapiro, N.M., Droznina, S. Ya., Senyukov, S.L., Chebrov, V.N. & Gordeev, E.I., 2015. Detecting and locating volcanic tremors on

- the Klyuchevskoy group of volcanoes (Kamchatka) based on correlations of continuous seismic records, *Geophys. J. Int.*, **203**(2), 1001–1010.
- Eaton, D.W., Akram, J., St-Onge, A. & Forouhdeh, F., 2011. Determining microseismic event locations by semblance-weighted stacking, in *Proceedings of the CSPG CSEG CWLS Convention*.
- Eisner, L., Hulsey, B.J., Duncan, P., Jurick, D., Werner, H. & Keller, W., 2010. Comparison of surface and borehole locations of induced seismicity, *Geophys. Prospect.*, **58**(5), 809–820.
- Eisner, L., Gei, D., Hallo, M., Opršal, I. & Ali, M.Y., 2013. The peak frequency of direct waves for microseismic events, *Geophysics*, **78**(6), A45–A49.
- Fitch, T.J., 1975. Compressional velocity in source regions of deep earthquakes: an application of the master earthquake technique, *Earth planet. Sci. Lett.*, **26**(2), 156–166.
- Font, Y., Kao, H., Lallemand, S., Liu, C.-H. & Chiao, L.-Y., 2004. Hypocentre determination offshore of eastern Taiwan using the maximum intersection method, *Geophys. J. Int.*, **158**(2), 655–675.
- Gajewski, D. & Tessmer, E., 2005. Reverse modelling for seismic event characterization, *Geophys. J. Int.*, **163**(1), 276–284.
- Gajewski, D., Anikiev, D., Kashtan, B., Tessmer, E. & Vanelle, C., 2007. Localization of seismic events by diffraction stacking, in *SEG Technical Program Expanded Abstracts 2007*, pp. 1287–1291.
- Geiger, L., 1912. Probability method for the determination of earthquake epicenters from the arrival time only, *Bull. St Louis Univ.*, **8**, 56–71.
- Gesret, A., Desassis, N., Noble, M., Romary, T. & Maisons, C., 2015. Propagation of the velocity model uncertainties to the seismic event location, *Geophys. J. Int.*, **200**(1), 52–66.
- Gharti, H.N., Oye, V., Roth, M. & Kühn, D., 2010. Automated microearthquake location using envelope stacking and robust global optimization, *Geophysics*, **75**(4), MA27–MA46.
- Gibowicz, S. & Kijko, A., 1994. *An Introduction to Mining Seismology*, Academic Press Inc.
- Grandi, S. & Oates, S.J., 2009. Microseismic event location by cross-correlation migration of surface array data for permanent reservoir monitoring, in *Proceedings of 71st EAGE Conference & Exhibition incorporating SPE EUROPEC 2009*, pp. X012.
- Grechka, V., De La Pena, A., Schisselé-Rebel, E., Auger, E. & Roux, P., 2015. Relative location of microseismicity, *Geophysics*, **80**(6), WC1–WC9.
- Grechka, V., Li, Z. & Howell, B., 2016. Relative location of microseismic events with multiple masters, *Geophysics*, **81**(4), KS149–KS158.
- Grigoli, F., Cesca, S., Vassallo, M. & Dahm, T., 2013. Automated seismic event location by travel-time stacking: An application to mining induced seismicity, *Seismol. Res. Lett.*, **84**(4), 666–677.
- Grigoli, F., Cesca, S., Amoroso, O., Emolo, A., Zollo, A. & Dahm, T., 2014. Automated seismic event location by waveform coherence analysis, *Geophys. J. Int.*, **196**(3), 1742–1753.
- Grigoli, F., Cesca, S., Krieger, L., Krigerowski, M., Gammaldi, S., Horalek, J., Priolo, E. & Dahm, T., 2016. Automated microseismic event location using master-event waveform stacking, *Sci. Rep.*, **6**, 25744, doi:10.1038/srep25744.
- 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.
- Guo, H. & Zhang, H.J., 2017. Development of double-pair double difference earthquake location algorithm for improving earthquake locations, *Geophys. J. Int.*, **208**(1), 333–348.
- Kao, S. & Shan, S., 2004. The source-scanning algorithm: Mapping the distribution of seismic sources in time and space, *Geophys. J. Int.*, **157**(2), 589–594.
- Kao, S. & Shan, S., 2007. Rapid identification of earthquake rupture plane using Source-Scanning Algorithm, *Geophys. J. Int.*, **168**(3), 1011–1020.
- Li, K.L., Sgatonni, G., Sadeghisorkhani, H., Roberts, R. & Gudmundsson, O., 2017a. A double-correlation tremor-location method, *Geophys. J. Int.*, **208**(2), 1231–1236.
- Li, K.L., Sadeghisorkhani, H., Sgatonni, G., Gudmundsson, O. & Roberts, R., 2017b. Locating tremor using stacked products of correlations, *Geophys. Res. Lett.*, **44**(7), 3156–3164.

- Li, L., Chen, H. & Wang, X.M., 2015. Weighted-elastic-wave interferometric imaging of microseismic source location, *Appl. Geophys.*, **12**(2), 221–234.
- Li, L., Chen, H. & Wang, X.M., 2016a. Relative elastic interferometric imaging for microseismic source location, *J. Geophys. Eng.*, **13**(5), 733–744.
- Li, L., Xie, Y.J., Chen, H., Wang, X.M. & Gajewski, D., 2017c. Improving the efficiency of microseismic imaging with particle swarm optimization, in *Proceedings of 79th EAGE Conference & Exhibition 2017*, pp. WeB411.
- Li, X.B., Wang, Z.W. & Dong, L.J., 2016b. Locating single-point sources from arrival times containing large picking errors (LPEs): the virtual field optimization method (VFOM), *Sci. Rep.*, **6**, 19205, doi:10.1038/srep19205.
- Maxwell, S., 2014. *Microseismic Imaging of Hydraulic Fracturing: Improved Engineering of Unconventional Shale Reservoirs*, Society of Exploration Geophysicists.
- McMechan, G.A., 1982. Determination of source parameters by wavefield extrapolation, *Geophys. J. Int.*, **71**(3), 613–628.
- Pesicek, J.D., Child, D., Artman, B. & Cieřlik, K., 2014. Picking versus stacking in a modern microearthquake location: Comparison of results from a surface passive seismic monitoring array in Oklahoma, *Geophysics*, **79**(6), KS61–KS68.
- 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**(1), 271–284.
- Poiata, N., Satriano, C., Vilotte, J-P., Bernard, P. & Obara, K., 2016. Multi-band array detection and location of seismic sources recorded by dense seismic networks, *Geophys. J. Int.*, **205**(3), 1548–1573.
- Poliannikov, O.V., Malcolm, A.E., Djikpesse, H. & Prange, M., 2011. Interferometric hydrofracture microseism localization using neighboring fracture, *Geophysics*, **76**(6), WC27–WC36.
- Poliannikov, O.V., Prange, M., Malcolm, A.E. & Djikpesse, H., 2013. A unified Bayesian framework for relative microseismic location, *Geophys. J. Int.*, **194**(1), 557–571.
- Price, D., Angus, D., Chambers, K. & Jones, G., 2015. Surface microseismic imaging in the presence of high-velocity lithologic layers, *Geophysics*, **80**(6), WC117–WC131.
- Ruigrok, E., Gibbons, S. & Wapenaar, K., 2017. Cross-correlation beamforming, *J. Seismol.*, **21**, 495–508.
- Schuster, G.T., Yu, J., Sheng, J. & Rickett, J., 2004. Interferometric/daylight seismic imaging, *Geophys. J. Int.*, **157**(2), 838–852.
- Sick, B. & Joswig, M., 2017. Combining network and array waveform coherence for automatic location: examples from induced seismicity monitoring, *Geophys. J. Int.*, **208**(3), 1373–1388.
- Spence, W., 1980. Relative epicenter determination using P-wave arrival-time differences, *Bull. seism. Soc. Am.*, **70**(1), 171–183.
- Staněk, F., Anikiev, D., Valenta, J. & Eisner, L., 2015. Semblance for microseismic event detection, *Geophys. J. Int.*, **201**(3), 1362–1369.
- Stein, S. & Wysession, M., 2003. *An Introduction of Seismology, Earthquakes, and Earth Structure*, Blackwell.
- Tan, Y.Y. & He, C., 2016. Improved methods for detection and arrival picking of microseismic events with low signal-to-noise ratios, *Geophysics*, **81**(2), KS133–KS151.
- Tian, X., Zhang, J. & Zhang, W., 2014. Double difference method for locating microseismic events from a single well, in *SEG Technical Program Expanded Abstracts 2014*, pp. 2193–2197.
- Tian, X., Zhang, J. & Zhang, W., 2016. Cross double-difference inversion for microseismic event location using data from a single monitoring well, *Geophysics*, **81**(5), KS183–KS194.
- Trojanowski, J. & Eisner, L., 2017. Comparison of migration-based location and detection methods for microseismic events, *Geophys. Prospect.*, **65**(1), 47–63.
- Verdon, J.P., Kendall, J., Hicks, S.P. & Hill, P., 2017. Using beamforming to maximise the detection capability of small, sparse seismometer arrays deployed to monitor oil field activities, *Geophys. Prospect.*, **65**, 1582–1596.
- Waldhauser, F. & Ellsworth, W.L., 2000. A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. seism. Soc. Am.*, **90**(6), 1353–1368.
- Wang, L.S., Chang, X. & Wang, Y.B., 2016. Locating micro-seismic events based on interferometry traveltimes inversion, *Chin. J. Geophys. (in Chinese)*, **59**(8), 3037–3045.
- Wang, R.J., 1999. A simple orthonormalization method for stable and efficient computation of Green's functions, *Bull. seism. Soc. Am.*, **89**(3), 733–741.
- Zeng, X., Zhang, H.J., Zhang, X., Wang, H., Zhang, Y. & Liu, Q., 2014. Surface Microseismic Monitoring of Hydraulic Fracturing of a Shale-Gas Reservoir Using Short-Period and Broadband Seismic Sensors, *Seismol. Res. Lett.*, **85**(3), 668–677.
- Zhang, H.J., Nadeau, R.M. & Toksoz, M.N., 2010. Locating nonvolcanic tremors beneath the San Andreas fault using a station-pair double-difference location method, *Geophys. Res. Lett.*, **37**, L13304, doi:10.1029/2010GL043577.
- Zhang, W. & Zhang, J., 2013. Microseismic migration by semblance-weighted stacking and interferometry, in *SEG Technical Program Expanded Abstracts 2013*, pp. 2045–2049.
- Zhebel, O., Gajewski, D. & Vanelle, C., 2010. Localization of seismic events in 3D media by diffraction stacking, in *SEG Technical Program Expanded Abstracts 2010*, pp. 2181–2185.
- Zhou, H.W., 1994. Rapid three-dimensional hypocentral determination using a master station method, *J. geophys. Res.*, **99**(B8), 15 439–15 455.

APPENDIX: GENERALIZATION OF CORRELATION-BASED IMAGING METHODS

The traditional BF algorithm is based on delaying and stacking the energy of the considered signals $s_i(t)$ (i is the receiver index and t is the time sample index), and its result is equal to the sum of autocorrelations and appropriately delayed cross-correlations of the waveforms:

$$\begin{aligned}
 S_{\text{beam}}(\mathbf{x}) &= \sum_{t=0}^{t_{\max}} \left(\sum_{i=1}^N s_i(t + \tau_{i,\mathbf{x}}) \right)^2 \\
 &= 2 \sum_{\substack{i=1 \\ j=i+1}}^N \left\{ \sum_{t=0}^{t_{\max}} s_i(t + \tau_{i,\mathbf{x}}) s_j(t + \tau_{j,\mathbf{x}}) \right\} \\
 &\quad + \sum_{t=0}^{t_{\max}} \sum_{i=1}^N (s_i(t + \tau_{i,\mathbf{x}}))^2 \\
 &= 2 \sum_{\substack{i=1 \\ j=i+1}}^N \left\{ \sum_{t=i_{i,\mathbf{x}}}^{t_{i,\mathbf{x}}+t_{\max}} s_i(t) s_j(t + \Delta\tau_{ij,\mathbf{x}}) \right\} \\
 &\quad + \sum_{t=0}^{t_{\max}} \sum_{i=1}^N (s_i(t + \tau_{i,\mathbf{x}}))^2 \\
 &\approx 2S_S + \sum_{t=0}^{t_{\max}} \sum_{i=1}^N (s_i(t + \tau_{i,\mathbf{x}}))^2, \tag{A1}
 \end{aligned}$$

where the cross-correlation stacking term S_S dominates the BF value. Furthermore, above derivation can be directly forwarded to scenarios considering multiple seismic phases. Similarly, if we replace signal $s_i(t)$ in eq. (A1) with cross-correlation waveforms C_S and C_R , the formula (eq. A1) changes to eqs (14) and (15). The two formulae can be regarded as SCBF and RCBF. The corresponding dominating terms change to the DCS S_D and HCS S_H , respectively. A numerical experiment based on the 2-D homogenous model in Section 2.7 is conducted to verify eqs (14) and (15). The imaging results in Fig. A1 shows that DCS and HCS are actually the simplified forms of SCBF and RCBF, respectively.

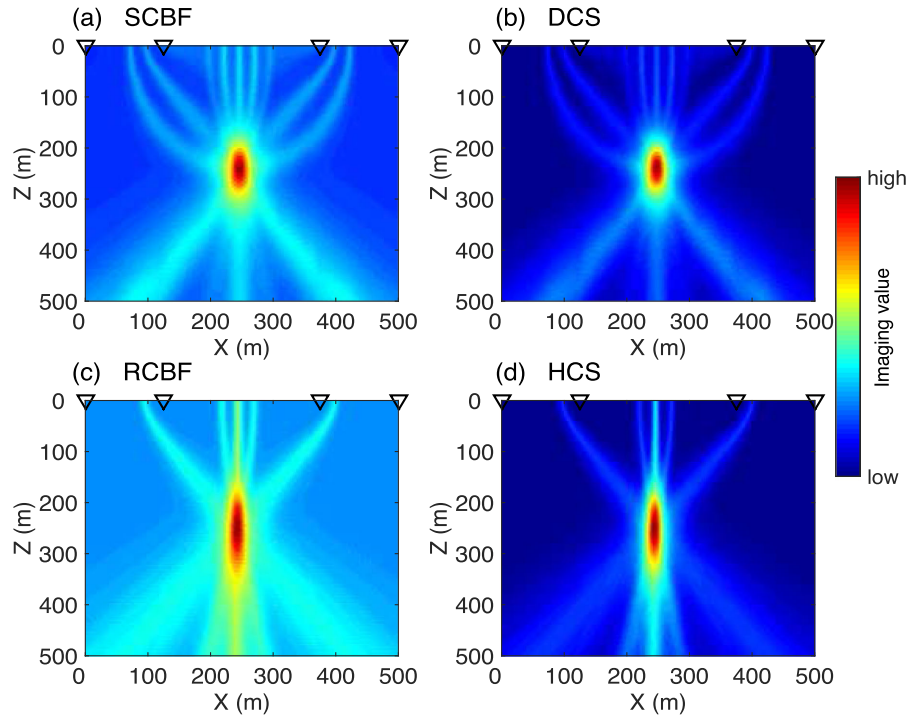


Figure A1. Imaging results of SCBF (a), DCS (b), RCBF (c) and HCS (d).