submitted to *Geophys. J. Int.*

Automated seismic waveform location using Multichannel Coherency Migration (MCM)–II. Application to induced and volcano-tectonic seismicity

Peidong Shi,[1*] Andy Nowacki,[1] Sebastian Rost[1] and Doug Angus[2]

[1] *School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK*

[2] *ESG Solutions, Kingston, ON K7K 7K2, Canada*

Received 2018 **** **; in original form 2017 October 20

SUMMARY

Locating microseismic events is essential for many areas of seismology including volcano and earthquake monitoring and reservoir engineering. Due to the large number of microseismic events in these settings, an automated seismic location method is required to perform real time seismic monitoring. The measurement environment requires a precise and noise-resistant event location method for seismic monitoring. In this paper, we apply Multichannel Coherency Migration (MCM) to automatically locate microseismic events of induced and volcano-tectonic seismicity using sparse and irregular monitoring arrays. Compared to other migration-based methods, in spite of the often sparse and irregular distribution of the monitoring arrays, the MCM can show better location performance and obtain more consistent location results with the catalogue obtained by manual picking. Our MCM method successfully locates many triggered volcano-tectonic events with local magnitude smaller that 0, which demonstrates its applicability on locating very small earthquakes. Our synthetic event location example at a carbon capture and storage site shows that continuous and coherent drilling noise in industrial settings will pose great challenges for source imaging. However, automatic quality control techniques including filtering in the frequency domain and weighting are used to automatically select high quality data, and can thus effectively reduce the effects of continuous drilling noise and improve source imaging quality. The location performance of the MCM method for synthetic and real microseismic datasets demonstrates that the MCM method can perform as a reliable and automatic seismic waveform analysis tool to locate microseismic events.

Key words: Earthquake source observations – computational seismology – time-series analysis – earthquake monitoring and test-ban treaty verification.

1 INTRODUCTION

Microseismic or passive seismic monitoring has been used extensively in monitoring geo-industrial applications (e.g., hydraulic fracturing, carbon dioxide storage and mining setting (Power et al. 1976; Verdon et al. 2011; Gibowicz & Kijko 2013; Shi et al. 2018a)) as well as hazard monitoring (e.g., volcano-seismology and slope stability (Wilks et al. 2017; Xu et al. 2011)). As a cost-effective monitoring technique, microseismic monitoring is used to demonstrate storage security of carbon capture and storage (CCS) (Verdon et al. 2010; Shi et al. 2018c). It is also an effective method for monitoring volcanoes and forecasting potential eruptions (McNutt 1996; Lavallée et al. 2008). Microseismic monitoring can provide geomechanical deformation information induced by fluid injection or flow, which can be used to evaluate rock failure processes in the reservoir of a carbon storage site or volcanic edifice.

Noise is an inevitable feature of recorded seismic data. Typically, random noise is assumed to be stationary with a Gaussian distribution, whereas real noise is often non-stationary and so does not conform to a single Gaussian distribution (Birnie et al. 2016; Yuan et al. 2018b). With these features, seismic data with real noise are often more challenging for seismic processing and more difficult to deal with than Gaussian or white noise. For CCS, microseismic monitoring is often conducted during carbon dioxide injection. Therefore, the ambient noise due to the fluid flow and injection exists all the time during the injection process, especially for monitoring arrays which are deployed close to the injection well. Local drilling with associated continuous drilling noise can also affect the recorded seismic data significantly. The injection and drilling noise are

⋆ Corresponding author: Peidong Shi. Email: eepsh@leeds.ac.uk. Fax: +44 113 343 5259

continuous and are often coherent across many of the receivers. They can form a great challenge for microseismic event location in CCS (Barkved et al. 2002; Knudsen et al. 2006; Birnie et al. 2016, 2017). Therefore, suitable ways to reduce or remove real noise and obtaining accurate event location results are required.

For CCS and volcano seismicity, a large number of seismic events can happen within a short period, which can be very difficult and time-consuming to locate by manual arrival time picking (Yuan et al. 2018a). In addition, the ever increasing monitoring data volume and larger monitoring arrays also put great demands on automatic seismic location algorithms for efficient microseismic monitoring. The traditional arrival time based location methods require phase identification and picking, thus are not suitable for automatic event location. Although there are ways to perform an automatic arrival time picking (Bai & Kennett 2000; Maggi et al. 2009), manual picking is still required to increase the picking reliability when the signal-to-noise ratio of seismic data is low or the arrivals of seismic events are overlapped. There have been various migration-based location methods developed to automatically locate seismic events using recorded waveforms (Kao & Shan 2007; Gharti et al. 2010; Drew et al. 2013; Grigoli et al. 2013a,b; Zhebel & Eisner 2014; Langet et al. 2014; Cesca & Grigoli 2015; Grigoli et al. 2016). Compared with arrival time based methods where the arrival times are determined by manual picking, automated waveform based location methods do not need phase picking and association, thus are more efficient and have the ability to identify more seismic events. Small, more numerous seismic events which cannot be picked manually or automatically can be effectively identified by fully utilizing the recorded full waveforms. Thus the automated waveform based location methods can help add more insights into the fracturing process and natural earthquakes. By using the waveforms and the matched filter technique, Peng & Zhao (2009) detected a large number of missing aftershocks along the Parkfield section of the San Andreas fault and used the newly detected seismic events to understand the postseismic deformation around the rupture zone associated with the mainshock of the 2004 Parkfield earthquake. However, the matched filter technique requires reliable waveform templates. Therefore, this technique is not suitable for research areas where there is no available event catalogue.

Migration-based methods have the potential to be applied as real time location schemes, yet the location reliability and accuracy of these methods is often unsatisfactory in presence of strong noise. Location accuracy is very important in terms of correctly imaging the fracture process and geometry, which can be used to reveal the source mechanism and deformation orientation. Large location errors during microseismic monitoring of CCS and volcano seismicity may contribute to huge economic loss or larger risk as the injection may be terminated prematurely if the induced fracture length has been exaggerated or volcano activity is underestimated because of mislocation of volcano seismicity. The other problem which often challenges migration-based location methods is the station coverage and distribution. Sparse monitoring stations hinder the utilization of waveform coherency for migration-based location methods, which causes poor noise-resistance and location performance. Irregular station distribution will reduce imaging resolution and lead to blurred location results. However, due to the restrictions of the actual deployment environment and cost, practical monitoring arrays are often sparse and irregularly distributed especially for natural earthquake monitoring arrays. Therefore, an automatic and precise seismic location method which can work on sparse and irregular monitoring arrays as well as efficiently with dense and/or regular networks is in great demand.

Shi et al. (2018b) proposed a fully automated seismic location method based on waveform coherency. This automated location method utilizes Multichannel Coherency Migration (MCM) and is suitable for locating induced seismicity and natural earthquakes. Different to traditional migration-based location methods which locate the source by stacking waveforms of characteristic functions, MCM calculates the multichannel coherency among stations and stacks the coherency to reveal the source location and origin time. By utilizing multichannel waveform coherency, MCM exhibited excellent location performance with high resolution and outstanding noise resistance. The multichannel coherency has also been utilized to improve the horizontal imaging resolution in seismic interpretation (Yuan et al. 2017). Compared to traditional migration-based location methods, MCM can extract more effective information from seismic waveforms, which give it the ability to locate microseismic and resist interference with noise and other non-related events. The theory and synthetic tests of the multidimensional MCM event-location method can be found in Shi et al. (2018b). Here, we demonstrate that the MCM location method can be used to automatically locate both injection induced and volcano-tectonic microseismic events especially when the monitoring array is sparse and/or irregularly distributed. We also compare and discuss the location results with other commonly used migration methods under different real noise levels using sparse and irregular monitoring arrays. First, as a feasibility study, we use the MCM to locate two volcano-tectonic earthquakes at the Uturuncu Volcano in Bolivia using a sparse monitoring array and also compare the location results with published event locations in the catalogue. We then apply the MCM to automatically locate triggered earthquakes following the M_w 8.8 Maule earthquake at Uturuncu (Jay et al. 2012) using four hours of continuous waveform data. Then, synthetic seismic data of an irregularly distributed monitoring array with real drilling noise were used to evaluate the location performance of different methods for induced seismicity. In order to obtain a satisfactory location result, quality control methods to remove the coherent drilling noise are explored and discussed. Finally, location performance and imaging resolution in different directions of different migration-based methods are analysed and discussed in detail.

2 THEORY AND COMPUTATIONAL EFFICIENCY ANALYSIS

In this section, we will briefly introduce the 2-dimensional MCM (for a more detailed description and the multidimensional MCM see Shi et al. (2018b)). For MCM, at a particular imaging point k and origin time t_0, the correlation coefficient between the waveforms of two different stations is calculated by:

$$r_{ij} = \frac{\sum_{t=t_0}^{t_0+t_w} \left[d_i(t+t_{ki}) - \overline{d_i(t+t_{ki})}\right] \left[d_j(t+t_{kj}) - \overline{d_j(t+t_{kj})}\right]}{(N_t - 1)\, \sigma_i \sigma_j}, \tag{1}$$

where r_{ij} is the correlation coefficient (i.e. coherency) between the waveforms at station i and j, d_i and d_j are the two input waveforms within the selected time window for station i and j, t_w is the coherency analysis time window for a particular seismic phase, N_t is the number of time samples in the time window, t_{ki} and t_{kj} are traveltimes of a particular seismic phase from imaging point k to the station i and j, σ is the standard deviation of the corresponding signal and the overlines denote averages.

After calculating correlation coefficients for all possible station pairs, the stacking function can be expressed as

$$p(x,y,z,t_0) = \frac{1}{N(N-1)} \left(\sum_{i<j}^{N} |r_{ij}^P| + \sum_{i<j}^{N} |r_{ij}^S| \right). \tag{2}$$

where r_{ij}^P and r_{ij}^S represent the waveform coherency of P- and S-phases for station pair ij, N is the number of stations and the number of unique receiver pairs equals $N(N-1)/2$, $p(x,y,z,t_0)$ is the final 4D imaging function and stores the stacked waveform coherency at position (x,y,z) and origin time t_0 (Shi et al. 2018b).

The 4D migration volume contains all the information about source location and origin time. Locations (x_s,y_s,z_s) and origin times t_{0s} of seismic events can be identified by finding the maximum value above a preset coherency threshold within certain time periods

$$p(x_s,y_s,z_s,t_{0s}) = max_{t0\in[t_1,t_2]}\{p(x,y,z,t_0) \geq p_c\}. \tag{3}$$

As an automated seismic location method, only a few input parameters, i.e. length of coherency analysis time window t_w and coherency threshold p_c, are required for MCM in an event location process. The length of coherency analysis time window t_w should be equal to or larger than the approximate period of seismic phases (Shi et al. 2018b). A longer time window is suggested in order to suppress the interference of noise and other incoherent phases when seismic data contain strong noise or coda waves. The coherency threshold p_c is determined according to the background noise level. A higher coherency threshold can help identify seismic events which have more probability to be real seismic events, but will also decrease the number of identified seismic events. It is worth noting that the migration process and the event identification process are two totally independent processes. So it is easy to adaptively adjust the coherency threshold according to a migration volume and choose a suitable threshold which can fulfil the requirements of the application.

For event locations based on manual picking, the computational efforts depend on the number of earthquake events in the time period. The more events there are, the more expensive it is to pick and locate them. However, for MCM, the computational cost is independent of the number of events. The computational cost is only related to the number of imaging points, the number of searched origin times and the number of stations. The whole MCM procedure is highly parallelizable, and the migration process is quite independent on different scales (from imaging point level to origin time level). Therefore, parallel computing in MCM can be performed on different imaging points or different origin times according to actual requirements. Very little communication is required for MCM when performing parallel computations, e.g. maximum migration values of different origin times when performing parallel computing on different origin times or migration values of different imaging points when performing parallel computing on different imaging points. We implement MCM using the Message Passing Interface (MPI) and analyse its computational efficiency on a high performance cluster (Figure 1). Both P- and S-waves are used in the MCM calculation and the number of time samples within the P/S time window is 100. Figure 1 (a) shows the computational times for different numbers of imaging points (Ns) and origin times (Nt) used in the MCM. As can be seen in the figure, the computational cost increases linearly with the number of imaging points and origin times, which demonstrates that the MCM workload scales essentially perfectly. Figure 1 (b) shows the computational times for different numbers of stations (N) used in the MCM. As we can see in the figure, the computational cost increases rapidly with the number of stations. Actually, the computational cost is proportional to the number of unique station groups: $N \times (N-1)/2$ (as Figure 1 (b) blue line shows), which is in accordance with the theory of MCM (Shi et al. 2018b). Figure 1 (c) shows the computational times and speedup ratios when different numbers of computing cores (Nc) are used. As expected the computational times (black line) decrease dramatically when more cores are used in the computation. The speedup ratios (blue line) are calculated by dividing the computational times of different cores by the computational time of a single core. Due to the high scalability of the MCM process, the speedup ratio of MCM is very close to the theoretical speedup ratio (red dashed line). Accordingly, we assume that the computational time is proportional to the number of imaging points, the number of origin times and the number of unique station pairs, and the speedup ratio equals the theoretical speedup ratio. Therefore, the computational time $t = k \times Ns \times Nt \times N \times (N-1)/Nc$, where k is a coefficient related to computer architecture. Using the data of Figure 1 (a-c), we obtain a coefficient of $k = 1.5 \times 10^{-7}$ second with the current settings. If real time processing is required, the MCM calculation time should be less than the length of the data. Here, we assume that the sampling interval for searched origin times is 0.1 second, and thus we have 10 origin times to process for each second of recorded seismic data. Therefore, for a real time processing, the required cores should fulfil $Nc \geq 10 \times k \times Ns \times N \times (N-1)$. Figure 1 (d) shows the required cores for real time processing when different numbers of stations and imaging points are used in MCM. The real time processing is expensive, but is still feasible with the current computer resources when the number of stations is not very large (e.g. $Ns \leq 40$). For sparse surface monitoring arrays, the number of deployed stations is usually smaller than 20. The number of imaging points can be reduced to less than 300K when locating seismic events in a small region. Therefore, real time processing is completely feasible in this situation. For example, for the Uturuncu dataset which we will discuss in detail in the next section, 68 cores are needed to conduct real time processing (shown as the red dot in Figure 1 (d)). Here, we only implement MCM using CPUs and MPI. Because the whole MCM process is highly parallelizable and the computation of MCM can be simultaneously processed in large blocks of data, we anticipate much larger speedup ratios when using Graphics Processing Units (GPU), which we are currently exploring.

3 LOCATION OF SHALLOW SEISMICITY AT UTURUNCU VOLCANO

Uturuncu is a long-dormant stratovolcano in Bolivia, which has an elevation of about 6000 m (Jay et al. 2012). Recent studies of surface deformation, fumarolic activity and the earthquake rate of Uturuncu show signs of unrest and potential of eruption again, which calls for close monitoring (Pritchard & Simons 2004; Sparks et al. 2008; Jay et al. 2012). As shown in Figure 2, 15 three-component seismometers have been temporarily deployed surrounding the inflating Uturuncu from April 2009 to April 2010 (Pritchard 2009). The farthest station is located about 25 km from the volcano summit. The seismometers have a sampling rate of 50 samples/s, which means the highest effective frequency of the recorded data is 25 Hz. Nine seismometers are short-period instruments and six seismometers are intermediate-period instruments. The tectonic setting of Uturuncu and the catalogue for these events located by manual picking can be found in Jay et al. (2012). We apply the MCM on the recorded continuous waveform data to show the potential of this method in a volcano-tectonic settings, using a sparse seismic network common in such environments.

3.1 Locating two local volcano-tectonic microearthquakes

First, we apply four different waveform migration methods to locate two local volcano-tectonic earthquakes at the Uturuncu and compare the location results. The magnitudes of these two local volcano-tectonic earthquakes are below M_L 1.0. The depths of the two shallow volcano-tectonic earthquakes are above the sea level. We use four different waveform migration techniques, i.e. envelope (Kao & Shan 2007; Gharti et al. 2010), STA/LTA (Drew et al. 2013; Grigoli et al. 2013b), kurtosis (Langet et al. 2014) and MCM (Shi et al. 2018b), to compare the performance in this setting. For STA/LTA migration, the short-term time window has been chosen to be 4 seconds and the long-term time window is 40 seconds. The time window for calculating kurtosis is 4 s. For MCM, a coherent analysis time window of 6 s and a two-channel based coherency scheme are used to locate the seismic events. The coherency threshold of MCM is set to 0.13. Because the monitoring array is very sparse, we set the weighting factors of all stations to 1, which means each trace is equally treated and used for migration. The spatial and temporal intervals used in the source imaging are 100 m and 0.08 s respectively. Because the vertical component data show distinct arrivals of P-waves, we only utilize the direct P-wave to conduct MCM for the vertical component data. Similarly for the north-south and east-west components, we only utilize the direct S-wave to image the events. The coherency of the three component data are then added together to obtain the final imaging values of a particular origin time and space point. The location results of the migration methods are compared to the locations in the catalogue. The velocity model used in the event location is the same layered model as described in Jay et al. (2012).

Figure 3(a) shows the recorded three-component waveforms at station UTCA for the first event, whose local magnitude M_L is 0.63 (Jay et al. 2012). The direct P-wave and S-wave of this event are distinguishable in the recorded waveforms, but the waveforms contain extended coda. The whole waveform train containing direct waves and coda waves for this event is about 6 seconds. Figure 4 shows the vertical and horizontal profiles of the migration results for the four different waveform migration methods using all available data. The depth (Z-axis) is measured relative to the sea level. The layer at depth 0.5 km which shows a velocity increase can be seen clearly in the migration profiles. The catalogue location (Jay et al. 2012) of this event is displayed as a star in the figure for comparison. For the envelope and STA/LTA migration, the source energy is not well focused. Thus the event location results of these two methods are not reliable, probably because the envelope and STA/LTA cannot identify the event onset from the recorded waveforms. For kurtosis and MCM, the source energy is well focused, thus the location results are more useful. The event location result of the MCM shows better agreement to the location in the catalogue. The location deviations of the MCM result relative to the event in the catalogue are 0.584, 0.557 and 0.469 km in the X, Y and Z directions, respectively (Table 1). Figure 5(a) shows the stacking function of the MCM method at the position of the most coherent point. The stacking function jumps to the maximum value at about one coherent analysis time window earlier than the published origin time of the event and drops down to the noise level quickly. The estimated origin time of the MCM method can be determined from the maximum coherency time, the analysis time window and the period of the direct waves. This is in agreement with Shi et al. (2018b). We will discuss this later in detail in the discussion section. The maximum coherency value is only about 0.16. A longer coherent analysis time window tends to decrease the overall waveform coherency as more data including noise are put into the coherent analysis. However, a longer time window is beneficial for obtaining a stable migration result. The coherency of the coda wave is also included to benefit the source imaging. For this volcano earthquake dataset, tests show that the analysis time window needs to be at least 1 s to eliminate the influence of the noise and pure coda waves. We used a time window of 6 s for both events to make the migration results more stable.

Table 1 shows the quantitative location results of the different migration methods and the comparison with the catalogue location. The origin time of this event for the MCM method in the table is estimated using the maximum coherent time plus the coherent analysis time window following Shi et al. (2018b). The event location of the MCM method shows the best correlation to the event location in the catalogue with less deviation in the location and origin time. Predicted P- and S-wave arrival times for this event in the catalogue and the event located by MCM are compared on record sections in Figure 6. The direct P- and S-wave arrivals correspond well with the predicted P- and S-wave arrival times for MCM location in most stations. Therefore the location determined by MCM of this event is acceptable.

Figure 3(b) shows the recorded three-component waveforms at station UTCA for the second event, whose local magnitude M_L is -0.29. The direct P-wave can be well identified in the vertical component and the direct S-wave can be well identified in the north-south and east-west components. The coda waves following the direct P- and S-waves are obvious. Figure 7 shows the vertical and horizontal profiles of the migration results for the four different waveform migration methods. As with the migration results of the previous event, the envelope and STA/LTA migration methods do not focus the source energy appropriately. The migration results of the kurtosis and MCM method are quite similar. The horizontal locations of this event using the kurtosis and MCM method are consistent with the catalogue location with only little deviation. However, the located event depths of both kurtosis and MCM method are deeper than the event depth in the catalogue (1.72 km and 1.92 km deeper respectively). Nevertheless, compared to the horizontal location of the seismic event, the event depth is often not well constrained by the recorded data especially for surface arrays. The trade-off between event depth and origin time often makes event depth determination problematic and more difficult (Eisner et al. 2010). Figure 5(b) shows the stacking function of the MCM method at the position of the most coherent point. Table 2 shows the quantitative location results of the different migration methods and the comparisons with the catalogue. The location results of the MCM correspond very well with the catalogue in the horizontal directions (with very small deviations of 0.166 km and 0.181 km in the X and Y directions, respectively). Predicted P- and S-wave arrival times for this event in the catalogue and the event located by MCM are further compared on record sections in Figure 8. Probably because of the strong heterogeneity in the subsurface, the recorded waveforms at some stations are not very coherent with the waveforms at other stations. However, the migration results of the MCM method are not seriously affected and seem still reliable. From the record sections of the vertical component (Figure 8 first row), we can clearly see the recorded direct P-wave arrivals show better consistency with the theoretical P-wave arrival times in the record section of the MCM method. This further demonstrates the reliability of the MCM location results.

The Uturuncu example shows that MCM can be used as a practical and precise seismic location method for automated volcano-tectonic and natural earthquake monitoring. As MCM utilizes the waveform coherency across different stations, it performs better under high noise conditions and can obtain a more accurate location result compared to other migration-based location methods. Sparse monitoring arrays will decrease imaging resolution and cause location uncertainties. The utilization of multichannel coherency information in MCM can greatly expand available information used for location and improve imaging resolution (Shi et al. 2018b), which is critical for seismic event location using sparse monitoring arrays.

3.2 Locating triggered events on four hours of continuous waveform data

The M_w 8.8 Maule earthquake on 27 February 2010 (at 06:34 UTC) triggered hundreds of earthquakes at Uturuncu with the passage of surface waves and the overtone phases of surface waves (Jay et al. 2012). Those triggered seismic swarms are recorded by the deployed Uturuncu monitoring arrays. According to Jay et al. (2012), the triggered events occurred with the onset of the Love and Rayleigh waves, and the earthquake rate reaches a maximum value of two events per minute with the passage of the Rayleigh wave overtones. We apply the MCM to automatically locate these triggered earthquakes using four hours of continuous data (06:00:00 to 10:00:00 UTC), which recorded most of the triggered events. The recorded waveform data at station UTCA are shown in Figure 9. As shown in the enlarged part of Figure 9, there are many small magnitude events which can be very difficult and time consuming to pick manually. As the triggered earthquakes start immediately after the surface wave train, many events occurred in a short time period with very close or overlapping waveform trains. Therefore, it will be very difficult to pick and associate different phases to a particular event. In addition, because of interference of noise and coda waves, it is also very difficult to accurately pick the P- and S-wave arrival times of small seismic events. The manual picking accuracy is highly dependent on human experience. The manual picking errors will inevitably cause location errors. As the MCM does not require picking and phase identification and the location accuracy of the MCM does not depend on event magnitude (waveform amplitude), it is very suitable to be used to automatically locate those dense triggered microseismic events.

The surface waves and surface wave overtones of the M_w 8.8 Maule earthquake not only trigger many seismic events in this area but also forms a big challenge for migration imaging using waveforms. Here, we filter waveforms using a frequency band of 4.2 - 21.6 Hz to exclude the influence of surface waves and low frequency noise. Because the sample rate (50 samples/s) is low, we suggest to use a long time window for coherency analysis in the MCM. Using a longer time window can improve the imaging stability and quality in noisy situations. We adopt a four second time window for both P- and S-waves in the MCM to resist the interference of noise and coda waves. Similarly, we only utilize the direct P-wave to conduct MCM for the vertical component data, and only utilize direct S-wave for the horizontal component data. The coherency value of the P-wave for the vertical component data and coherency values of the S-wave for the two horizontal component data are then stacked together to form the final migration value. For conventional waveform migration methods which stack amplitudes or characteristic functions of amplitudes, S-phases are often assigned higher weighting factors because S-phases tend to have higher amplitude. However, the S-phase often interferes with coda and converted waves, thus tends to have lower waveform coherency across different stations. In contrast, the P-phase which arrives first often has higher coherency despite its lower amplitude. Therefore, we assign a weighting factor of 0.6 to the P-phase of the vertical component and factors of 0.2 to the S-phases of each of the two horizontal components (east-west and north-south), noting that the MCM is insensitive to amplitudes.

The imaging area is 18 km, 15 km and 8 km in north-south, east-west and vertical directions, respectively (as shown in Figure 2 white rectangle area). The imaging point interval is 200 m in all different directions. Therefore, there are 283,556 imaging points in total. The time interval for searching for origin times is 0.08 s. The total number of searched origin times in the four hours is about 180,000. We assume that two earthquakes will not occur at the same time or within a few origin time samples (0.08 s). Therefore, at each searched origin time, we only save the imaging point which has the maximum coherency value. Figure 10 shows the variation of the maximum coherency value with different origin times in the four hours. When an earthquake occurs, at the correct origin time, the coherency values of each imaging point will all rise due to the arrival of the long waveform trains including direct, converted and coda waves. Therefore, we can observe many local peaks rising from the background noise in Figure 10, which potentially correspond to seismic events. We use a coherency threshold of 0.1. By identifying the maximum value of each local peak, we can find the location and origin time of each seismic event. We identify 560 local peaks in Figure 10, which are viewed as potential seismic events. We then check each potential seismic event using the corresponding record sections of these potential seismic events and verify 322 seismic events which have clear phase arrivals. The verified seismic events are shown as red dots in Figure 10. Although there are many events which do not show clear P- and S-phase arrivals, they may still be real seismic events, because the signal-to-noise ratio (SNR) for these events may be small (smaller than 1). Since the MCM has the ability to resist strong noise, it is not surprising that it can successfully identify seismic events below the noise level. The problem is that although identified by MCM, the weak seismic events with SNR below 1 cannot be effectively verified through their record sections at the present. By adopting stricter parameters (such as a higher coherency threshold, higher source prominence and longer origin time gap), we can also reduce the number of unverifiable seismic events and improve the proportion of confirmed seismic events. However, this would inevitably result in losing some small real seismic events which cannot be effectively verified by inspection of the record sections at the present. Therefore, further studies about detecting and verifying seismic events (especially events with low SNR) from migration traces/volumes are still needed. Here, since the verifying process is very quick and easy, in order to identify as many seismic events as possible, we adopt relatively relaxed parameters to identify the local maxima in the coherency time slice (Figure 10).

The existing catalogue has 114 seismic events in total in this four hour time period in this area, which are located by manual picking. For those 114 seismic events, 112 events (98.25%) have been successfully located by the MCM. In addition, the MCM has also automatically located 210 more seismic events than the existing catalogue, which have been verified on the record sections. By checking the corresponding record sections, we find that the MCM not only automatically locates many more triggered seismic events than the catalogue, but also the origin time estimates of most events are more accurate than the existing catalogue under the current velocity models. This demonstrates that MCM is an efficient and reliable automatic location method. Figure 11 shows the locations of the 322 verified seismic events and the 114 current catalogue events. In the figure, we can see that the distribution of automatically located seismic events is consistent with that of the events in the catalogue. There are two main earthquake clusters. One is located in the northern part of the study area and close to the volcano. This earthquake cluster occurred earlier (6 am to 8 am), and the events are mainly triggered by the surface waves (Love and Rayleigh waves) of the Maule earthquake. The other earthquake cluster occurred from 8 am to 10 am and is located in the southern part of the study area. The seismic events are mainly triggered by the surface wave overtones of the Maule earthquake.

Figure 12 shows a seismic event (referred to as event 1) which is both located by the MCM and the manual picking (catalogue). The MCM location result has a similar horizontal location as the catalogue result, but is deeper than the catalogue event. From the corresponding record sections (Figure 13), we can clearly see that the predicted arrival times of the MCM results have a much better correspondence with the P- and S-phase arrivals, especially for the S-phases of the horizontal components. This demonstrates that the MCM location results are reliable and have a better estimation of the origin times of seismic events. Figure 14 shows the migration profiles and record sections of a newly identified seismic event (referred to as event 2) by MCM, which is not in the existing catalogue. The source energy focuses nicely in the migration volume. The record sections which show clear P- and S-wave arrivals also indicate a real microseismic event occurred. It is worth noting that although event 2 is lower in event magnitude and has smaller amplitudes than event 1, the waveform coherency (0.22) of event 2 is higher than that of event 1 (0.17). For MCM, the waveform coherency not only depends on the amplitude (relating to event magnitude and SNR), but is also influenced by the interference of coda waves, converted waves and arrivals from other events. Thus, it is not surprising that a small seismic event can have higher waveform coherency and focussing of migration energy than a larger seismic event when the small event is less affected by interference of other non-coherent waves. This characteristic makes MCM very suitable for locating microseismic events.

Many more seismic events have been identified by MCM in this four hour time period than the published catalogue, which greatly complements the catalogue. We provide our extended catalogue in the supplementary material. Figure 15 shows the number of triggered seismic events within the four hours. With this more complete catalogue, we find that rates of triggered events rise shortly after the passage of surface waves or surface wave overtones. Different to Jay et al. (2012), who conclude that rates of triggered events increased to a peak value of two events per minute with the passage of the X2/X3 Rayleigh wave overtones, we find that earthquake rates reach a peak value of about five events per minute after the passage the G1/R1 surface waves (Figure 15). An increase of seismicity after the passage of X2/X3 is noticeable, but only reaches about three events per minute.

4 AQUISTORE SYNTHETIC DATA WITH REAL NOISE

Synthetic waveform data with added Gaussian noise is often used in testing the performance of location algorithms. However in reality, the real noise field is not white, stationary or Gaussian (Birnie et al. 2016). Several noise studies have shown that seismic noise is often variable in space and time, leading to increased difficulty in source imaging (Birnie et al. 2017). In this section, we apply the MCM location algorithm to the Aquistore noise dataset to examine the location performance in the presence of real seismic noise. The Aquistore noise data have been extracted from a permanent surface array installed at the Aquistore carbon dioxide storage site (Roach et al. 2015; Birnie et al. 2016, 2017). The monitoring data used here were recorded by the surface array during the drilling and construction phase of the injection and observation wells prior to CO_2 injection. Therefore, significant drilling noise and non-stationary noise were recorded in the dataset. No injection-related or induced seismic events are recorded in this period, which makes the recorded time-series an excellent dataset for investigating the effect of real seismic noise on seismic location. Figures 16 (a) and (b) show the surface array geometry and velocity model of the Aquistore area. The surface array consists of 50 buried geophones (34 in North-south direction and 16 in East-west direction) with a sampling frequency of 500 Hz.

We generate waveform seismic data using the propagator matrix technique of Zhu & Rivera (2002) for both a shallow and a deep event (Figures 16 (c-e)). The shallow and deep events are located at a depth of 2.55 km and 3.15 km, respectively. The deep event has been placed in a thin and relatively low velocity layer. There are also many thin layers above and below the deep event (Figures 16 (b-e)), which may cause difficulty in imaging the deep event. We use the shallow and deep events to examine the influence of complex velocity model on the migration result. For both the shallow and deep event, a $45°$ dip-slip double-couple source with 40 Hz peak frequency is used to give a specified radiation pattern. The recorded real noise (Birnie et al. 2016) is added to the synthetic data to mimic as closely as possible a 'real' dataset with varying signal-to-noise ratios. The SNR is defined by the ratio of the maximum amplitude between signal and noise. This kind of semi-synthetic dataset enables a quantitative evaluation of the location errors in the presence of different realistic noise scenarios and has been employed to evaluate the monitoring performance of a dedicated seismic monitoring array (López-Comino et al. 2017). The synthetic data and noise data are shown in Figure 17. After adding noise, the arrival of the direct P-wave cannot be easily recognized. Stations 18-24 and 41-43 are deployed near the injection and observation well (as shown in Figure 16(a)), and thus are seriously contaminated by drilling noise (Figure 17c). The non-stationarity and spatial variability of the noise will make event location more difficult.

4.1 Location results for shallow event

We compare the location results of different migration methods using waveform envelope, STA/LTA and kurtosis as characteristic functions and also the MCM method for different SNRs. The same monitored real noise of different levels have been added into the synthetic dataset to make the semi-synthetic datasets of different SNRs. The SNRs are chosen to be infinite (noise free), 1, 0.5, 0.25 and 0.025 respectively. We then analyse the influence of SNRs to location results and compare the performance of different migration methods under different SNR situations. Figure 18 compares the migration results for the four different methods when the SNR is 1 (for a complete comparison of different SNRs, see supplementary material Figures S.1-S.3). When the SNR is larger than 0.25, the MCM exhibits the best resolution and location performance in both the horizontal and vertical directions. Due to the use of the derivative (Langet et al. 2014), the kurtosis seems to have better resolution in the XZ profile (as shown in Figure 18). However the location results of kurtosis migration are often biased due to the trade-off between depth and origin time.

The results also show that receiver distribution influences the results of the locations. Compared to the X direction, the image in the Y direction relies on fewer geophones leading to increased location uncertainty in that direction, and therefore the envelope, STA/LTA and kurtosis methods show poorer resolution in the Y direction (as shown in Figure 18). However the MCM still maintains very good resolution in the Y direction. The location results of the envelope, STA/LTA and kurtosis methods are often biased in the Z direction. When the SNR is below 0.25, the MCM fails to locate the source, because the noise recorded during drilling and construction of the injection well is pervasive over all the traces, especially notable in the traces which are close to the injection well. The drilling and construction noise coming from the injection well is continuous in time, and so leads to continuous coherent noise on all the traces. When the SNR is below 0.25, the drilling noise dominates the wavefield in all the traces. The continuous (both in space and time) and coherent drilling noise contributes to the failure of the MCM method when the SNR is below 0.25. The other methods also fail to locate the source accurately because of strong noise contamination. When the SNR is 0.025, all the methods fail to locate the source. However, approaches have been developed to 'whiten' the noise and hence to reduce the influence of coherent noise (Birnie et al. 2017).

The automatic weighting scheme can be integrated into the multidimensional MCM flexibly (Shi et al. 2018b), which gives MCM the ability to conduct automatic quality control of the input data. We devise an automatic quality control scheme to deal with the drilling noise and surmount the SNR limit in the presence of continuous drilling noise. The automatic quality control scheme comprises weighting and filtering. The weighting factors (Shi et al. 2018b, equation 3) are determined by evaluating the amplitude of each trace. Because the continuous drilling noise will normally contaminate a whole trace, here we use an average absolute amplitude ratio to discriminate very noisy traces and apply a weighting coherency calculation scheme to all traces. The absolute amplitude ratio of a trace is defined as the ratio of the average absolute amplitude of the trace to the average absolute amplitude of all traces ($a_i = \overline{|\mathbf{d}_i|}/\overline{|\mathbf{D}|}$, a_i is the absolute amplitude ratio of the i-th trace, $\mathbf{d}_i$ is the waveform amplitudes of the i-th trace and $\mathbf{D}$ is the amplitudes of all traces). Figure 19 shows the absolute amplitude ratios of different stations for the noisy datasets with different

SNRs. For traces which are highly contaminated by continuous drilling noise, the energy of this trace will be much larger than the average energy over the whole traces, which will contribute to a high absolute amplitude ratio (as shown in Figure 19). Through inspecting the absolute amplitude ratios of all traces, we can identify high quality traces and thus stabilize the migration result. We set an absolute amplitude ratio limit of 1.5. Above this limit, the weighting factor of this trace will be set to 0, otherwise the weighting factors are 1. Because our waveform coherency is evaluated through correlation coefficient, the absolute value of the amplitude will not affect the coherency calculation. Therefore, weighting values of 0 or 1 rather than sliding values are assigned to exclude or include traces in the coherency calculation. Through weighting, we select high quality data to conduct migration and exclude traces with very high absolute amplitude ratio (which means extremely low SNR for that trace). As shown in Figure 19, traces 19-24 and 41-43 which are close to the observation and injection wells are highly contaminated by the drilling noise (consistent with Figures 16 and 17). Therefore, after weighting these traces will be excluded from the dataset used for imaging. Before calculating multichannel waveform coherency or characteristic functions, the selected data are filtered in the frequency domain. Because the drilling and construction of the injection well are low-frequency processes, we applied a 6th-order highpass Butterworth filter with a cutoff frequency of 50 Hz to the semi-synthetic data to remove the low frequency drilling noise.

The migration results with automatic quality control scheme (weighting and filtering) are shown in Figure 20 and the SNR before filtering is 0.025 (for a complete comparison of different SNRs, see supplementary materials Figures S.4-S.6). Through the automatic quality control scheme, the imaging quality of the four migration methods becomes better and the imaging resolution also improves especially for low SNR scenarios. The MCM exhibits better location results with higher resolution compared to the other methods for all SNR situations. When the SNR is above 0.025, MCM can locate the source accurately without deviation, while the other three methods all have location deviations. When the SNR is 0.025, only the MCM can locate the source correctly with a minimal deviation of 20 m. With such a low SNR, the STA/LTA method focused at the shallow part of the true source position with very low imaging resolution, while the kurtosis method cannot focus correctly. Because of the non-Gaussian property of the real noise and the sensitivity of the characteristic function of the kurtosis method, the kurtosis method is more susceptible to the array geometry. An irregular and/or sparse monitoring array will tend to bias the location results of the kurtosis method. Thus the location results of the kurtosis method are less stable compared to the other three methods. Figure 21 shows the location errors of the four methods under different SNRs with/without automatic quality control scheme. The MCM method outperforms the other methods at all noise levels when the automatic quality control scheme is applied (Figure 21(b)). The implemented automatic quality control scheme using filtering and weighting can effectively improve the location accuracy for most tested methods.

4.2 Location results for deep event

The location results for the deep event with a SNR of 1 are shown in Figure 22. Since the SNR is relatively high, for consistency and better comparison with the migration results of the shallow event (Figure 18), original data without automatic quality control are used for migration. The velocity model above the deep event is more complicated as it contains thin layers and large velocity contrasts. However, compared to the shallow event, the location results of the deep event are not seriously affected by the complexity of the velocity model. Due to the increase of the velocity in the imaging area, the arrival time differences between the adjacent imaging points become smaller, which is detrimental for distinguishing the phase arrivals. Correspondingly the imaging resolution for all the 4 methods decreases compared to imaging results of the shallow event (as can be seen in the comparison of Figure 18 and 22). The imaging results of the envelope and STA/LTA methods still have large deviations in the vertical direction, while the MCM and kurtosis methods locate the deep event accurately. The imaging results of the MCM exhibit high resolution in the horizontal direction. However, the resolution in the vertical direction deteriorates compared to the results of the shallow event. The degradation of the vertical resolution is related to the chosen length of the time window of the coherence analysis as well as the velocity of the imaging area. Although the same time window is applied in the imaging of the shallow and deep events, the higher velocity of the deep event layer contributes to the reduction of the vertical resolution. Using a smaller time window can improve the imaging resolution, but at the expense of reducing noise suppression ability.

Figure 23 shows the stacking functions of the four methods at the true source location of the deep event. The four methods all exhibit excellent source prominence at the correct origin time. Time windows for both P- and S-phases are simultaneously used in the migration. The pink area around -0.6 second in Figure 23 highlights the time range where P-phases move into the stacking window of the S-phases when searching for origin time. Meanwhile, the pink area around 0.7 second highlights the time range where S-phases move into the stacking window of the P-phases. For the stacking functions of the envelope and STA/LTA methods, a notable peak can be observed at these times. However, the MCM can effectively suppress this kind of disturbance and avoid identifying unrealistic events.

From the results of the Aquistore dataset, we can see that MCM can be used as an effective migration method to automatically locate microseismic events induced by fluid injection or hydraulic fracturing. Although drilling or injection noise can pose big challenges for source imaging, different ways can be adopted to acquire reliable seismic location results. Irregularly distributed monitoring arrays will lead to unbalanced imaging resolution in different directions. However, due to the utilization of multichannel waveform coherency, MCM can acquire higher and much more balanced imaging resolution in different directions compared to other migration-based methods. As traveltime differences between adjacent imaging points in low velocity zone are larger, the imaging results in the low velocity zone (i.e. a shallow event) are better than those in a high velocity zone (i.e. at greater depths), and the source imaging resolution in the low velocity zone is also higher than that in the high velocity zone.

5 DISCUSSION

For the synthetic data case, the stacked coherency trace will normally exhibit a flat top as we have discussed in detail in Shi et al. (2018b). However, for real data, such flat tops may not exist because of strong interference from noise and coda waves (as shown in Figure 5). Typically, one records and takes the time and position which has the maximum coherency value as the origin time and location of a seismic event. However due to a systematic bias between the origin time and the maximum coherent time, calibration is needed in order to obtain an accurate estimation of the origin time of the seismic event. As shown in Figure 24, similar to synthetic data, the coherency will start to rise at one coherent analysis time window (referred to as Tw) before the correct source origin time (referred to as T0). For real seismic data, because of subsurface heterogeneity, there are many coda waves following the direct P- and S-phases. Those coda waves often show lower coherency compared to direct phases and have high amplitudes compared to background noise. Here, we assume the coda waves are incoherent. Therefore, the maximum coherency value will appear one period (referred to as T) of the direct phase after the rise of the coherency. That is to say the maximally coherent time (referred to as Tmax) is $Tw - T$ ahead of the correct source origin time. So the calibration equation for source origin time is $T_0 = Tmax + Tw - T$. Since it is easy to obtain a good estimate of the period of the direct phase, we can perform the origin time calibration easily and efficiently. If the direct phases have high SNR and coda waves are partly coherent, we may see a small flat top around Tmax or the maximum coherency value appears around the theoretical maximally coherent time (as shown in Figure 24 the dashed line). Thus in this situation, the estimate of the origin time will be affected and shows a small deviation. However, according to our experience of processing the Uturuncu dataset, after calibration we can have a good estimation of source origin times. Finally, in the stacked coherency trace, waveform coherency will decrease to background noise level at one coherent analysis time window after the maximally coherent time.

For the Uturuncu dataset, only a few stations are available for source location, which negatively affects the MCM imaging. However compared to other migration-based methods, in spite of the very sparse monitoring array, MCM still obtains more reliable and precise location results by the use of multichannel waveform coherency. A dense array with wide aperture and azimuth coverage will greatly improve the imaging quality, especially when a long analysis time window is used. When the stations are widely spread, the traveltime differences to different stations will be large. Thus the migration result of the MCM will be better and the influence of the continuous coda waves can also be reduced. High frequency information in the recorded data is important for improving the imaging resolution. For the Uturuncu dataset, because of a low sampling rate, the highest effective frequency is limited to 25 Hz. The volcano-tectonic earthquakes often contain high-frequency content above this cut-off frequency. The insufficient sampling of the waveform data (as can be seen in the Figure 3(b)) has limited the imaging resolution and quality. Despite the sparse recording array and the lack of high frequency content, the MCM still obtains reliable event locations. When possible, we recommend volcano monitoring arrays record at at least 100 Hz to facilitate future automatic volcano-tectonic event determination.

For natural earthquakes, strong coda waves are often observed in the seismograms. The strong coda waves can have a significant influence on the envelope of the waveforms and can also seriously affect the event detection using approaches such as STA/LTA and kurtosis. Thus the location performance of the envelope, STA/LTA and kurtosis migration methods will be negatively affected by the coda waves. Only when the coda waves of different stations are long-lasting and coherent, they will have a negative impact on the MCM migration. The continuous coherent coda waves will make the MCM source imaging ambiguous. One way to deal with coherent coda waves is to increase the analysis time window for the coherency calculation. By using a longer time window, the whole waveform train can be included in the coherent analysis and the direct P- or S-waves as well as the coda waves are utilized to image the source event. Thus in this way the coherency of the coda waves can be fully utilized to improve the event location, however at the expense of reducing the imaging resolution. If coda waves are incoherent, a short analysis time window is suggested to improve waveform coherency value and imaging resolution.

For event location at the Uturuncu, although the same velocity model used in obtaining the event catalogue is utilized for location here, the event locations of the waveform migration method are different to the event location in the catalogue, especially in event depth. The discrepancy may come from different types of information being used in the event location. For events in the catalogue, only the arrival times of the direct P- and S-waves obtained by manual picking are used in the event location. However for waveform migration methods, the recorded waveforms from different stations are directly used to locate the seismic event. MCM automatically identify the maximum coherent time according to recorded traces and the predicted phase arrival times are thus slightly different from the manually picked arrival times. Regardless of velocity model, the location result of the arrival time based methods will be affected by the accuracy of manual picking, especially for low magnitude events. The location result of migration based methods is mainly affected by the signal-to-noise ratio and medium heterogeneity, which influence the recorded waveforms.

In the record sections (Figures 6 and 8), the recorded direct P- and S-wave arrivals at some stations do not show a good consistency with the theoretical arrival times. And despite most direct P-wave arrivals corresponding very well to the theoretical arrival times, the recorded S-waves often arrive earlier than the theoretical S-wave arrival times. This discrepancy likely comes from the velocity model used in the event location. Here we just applied a layered velocity model with a constant v_p/v_s ratio of 1.75. In reality, the subsurface can have strong lateral velocity heterogeneity as well as varying v_p/v_s ratio. The S-wave velocity model obtained by ambient noise tomography (Jay et al. 2012) reveals the velocity heterogeneity in the Uturuncu area. If the velocity model is very rough, it is worthwhile to adopt a method which can simultaneously locate the source and update the velocity model. In this way, we can improve the event location accuracy and obtain a more precise velocity model at the same time. By adjusting both the event location and velocity model iteratively, the location results can match the arrivals of seismic phases more precisely. Nevertheless, this is beyond the scope of this study.

For general waveform location methods based on the stacking of characteristic functions, the imaging resolution in different directions is highly dependent on the array distribution. More geophones in a certain spatial direction increases resolution in that direction. However, if one direction is better sampled than the other directions, the imaging results will be dominated by the waveform stacking in that direction. Thus the imaging resolution in other directions (especially in the perpendicular direction) will be degraded (as can be seen in the comparison between the first and second rows in Figure 18). If we want to achieve equal resolution in different directions when locating the source, evenly distributed geophones are required. However, MCM utilizes the coherency between all possible receiver pairs, therefore the information from different directions can achieve a better balance improving the MCM locations compared to the other methods. For an irregularly distributed monitoring array, assuming n_p stations have been deployed in the predominant direction, whereas n_c stations ($n_p > n_c$) are deployed in the non-predominant direction. The contribution from non-predominant direction to the whole migrated volume for MCM ($2n_pn_c/\left[(n_p+n_c)(n_p+n_c-1)\right]$) is always higher than that for conventional migration-based methods ($n_c/(n_p+n_c)$). For MCM, due to the use of multichannel waveform coherency across all the stations, the effective information from non-predominant direction can occupy a higher proportion in migration compared to other conventional single-channel-based location methods. Therefore, the imaging results of MCM are less affected by the irregular distribution of the receivers, and the imaging resolution in different directions are well balanced.

As shown in the imaging results of the Aquistore real noise data, the location results of the envelope and STA/LTA methods often show large deviations in depth. This is because the characteristic functions such as envelope and STA/LTA cannot represent the arrival times of the P- and S-phases accurately. For envelope and STA/LTA, the maximum value of the characteristic function often appears later than the correct arrival times of the P- and/or S-waves. For example, if the source time function is a Ricker wavelet, the maximum value of the envelope is located at the peak amplitude of the P- and S-phases, not at the accurate arrival times of the P- and S-phases, i.e. a half-period later. The characteristic function represents a transformation on the original waveform, and the transformation on recorded waveforms of different stations can have different effects because of noise, source radiation pattern, instrument response, etc. Thus the delayed times corresponding to the correct arrival times can be different for different traces. This will lead to a trade-off between the location depth and the origin time of the event. Finally, both the depth and origin time of the location results can be biased. For the kurtosis method, due to the application of the derivative of kurtosis (Langet et al. 2014), it can represent the arrival times of the P- and S-waves more accurately. Thus less deviations in depth are observed in the location results. However, the kurtosis method is more affected by noise and irregular array geometry. The location results are not stable compared to the other methods, and deviations in depth can easily appear when it cannot represent the arrival time correctly. For MCM, due to the use of multichannel waveform coherency among different traces, the maximum coherency value will appear at the correct arrival times of the P- and S-waves and waveform coherency will decrease rapidly when they deviate from the correct source location (Shi et al. 2018b). There is less trade-off between location depth and origin time. Therefore, the MCM can accurately identify the source location and also the origin times with higher resolution.

Continuous coherent noise such as drilling noise remains a challenge for MCM. The coherent noise which is continuous both in space and time will lead to high coherency values between all receiver pairs, thus contributing to the failure of MCM when the coherent noise level is too high. Removing the continuous coherent noise is key to overcoming this problem. If the coherent noise in the recorded data falls into a specific frequency band, we can use frequency filtering or frequency-wavenumber filtering to remove the coherent noise and improve the imaging quality. For microseismic monitoring, the main coherent noise such as the drilling noise and injection noise are often low frequency noise (less than tens of Hz), while the dominant frequency of the microseismic signals are often relatively very high (from tens to thousands of Hz). Therefore, this kind of low frequency noise can be separated and removed from the microseismic dataset by filtering. Automatic quality control techniques such as weighting and filtering are effective ways to mitigate the effects of noise and improve imaging quality.

6 CONCLUSIONS

In this paper, we applied the MCM method (Shi et al. 2018b) to locate microseismic events in a reservoir and a volcanic setting in the presence of realistic noise. The location results of triggered volcano-tectonic earthquakes demonstrate the feasibility of using MCM method to locate natural earthquakes recorded by sparse arrays. The MCM can automatically locate many triggered events which are difficult and time consuming to manually pick. The MCM has the ability to locate microseismic events which are otherwise often neglected by researchers. Using MCM, we can efficiently obtain a more complete catalogue, which can help us better understand the subsurface earthquake process. The newly obtained seismic catalogue at Uturuncu using MCM can be found in the supplementary material. The predicted arrival times of P- and S-phases at different stations are also attached, which can be used for further studies such as relocation. Compared to other migration based methods, MCM shows more reliable location results and performs better in high noise, sparse monitoring array and strong coda situations. The Aquistore real noise case demonstrates the excellent imaging performance of the MCM in the presence of strong realistic noise. Even though strong coherent noise exists in all traces, the MCM can still locate the source accurately. Usual quality control techniques such as the frequency filtering and weighting are feasible ways to remove coherent drilling or injection noise, the latter of which we employ in an automatic way. Compared to the other methods, the location results of the MCM have higher resolution and are more stable.

Computational efficiency tests of the MCM show that the MCM is highly scalable and parallelizable. The parallel MCM code can achieve a high speedup ratio easily, which gives MCM the ability to perform real time processing. Seismic location with sparse and/or irregularly distributed monitoring array is problematic and difficult. MCM can expand the effective information used for locating by calculating multichannel waveform coherency across different stations, thus in this way improving the location performance with sparse array. When the monitoring array is irregularly distributed, MCM imaging resolution in different directions can also be well balanced due to the use of pairwise handling among all available stations. Compared to other single-channel-based location methods, the location result of MCM is less affected by the irregular and/or sparse distribution of the receivers, and the imaging resolutions in different directions are higher and well balanced. The MCM code is open source and can be downloaded from `https://github.com/speedshi/seisloc`. The MCM code is written in FORTRAN and further developments of the MCM software will be released in the future.

7 ACKNOWLEDGEMENTS

We would like to thank Claire Birnie and the Petroleum Technology Research Centre (PTRC) for access to Aquistore Data. Aquistore is an independent research and monitoring project managed by the PTRC which intends to demonstrate that storing liquid carbon dioxide (CO_2) deep underground (in a brine and sandstone water formation), is a safe, workable solution to reduce greenhouse gases (GHGs). Seismic data for the Uturuncu volcano-tectonic earthquakes were obtained via the IRIS DMC from stations in 'The life cycle of Andean volcanoes:Combining space-based and field studies (ANDIVOLC/Cornell)' Network run by the Cornell University (doi: 10.7914/SN/YS_2009). A. Nowacki is supported by a Leverhulme Early Career Fellowship and the Natural Environment Research Council (grant NE/R001154/1).

REFERENCES

Bai, C.-y. & Kennett, B., 2000. Automatic phase-detection and identification by full use of a single three-component broadband seismogram, *Bull. Seismol. Soc. Am.*, **90**(1), 187–198.

Barkved, O., Gaucher, E., Hornby, B., Kristiansen, T., & Maisons, C., 2002. Analysis of seismic recordings during injection using in-well permanent sensors, in *64th EAGE Conference & Exhibition*.

Birnie, C., Chambers, K., Angus, D., & Stork, A. L., 2016. Analysis and models of pre-injection surface seismic array noise recorded at the Aquistore carbon storage site, *Geophys. J. Int.*, **206**(2), 1246–1260.

Birnie, C., Chambers, K., & Angus, D., 2017. Seismic arrival enhancement through the use of noise whitening, *Phys. Earth Planet. In.*, **262**, 80–89.

Cesca, S. & Grigoli, F., 2015. Chapter two-full waveform seismological advances for microseismic monitoring, *Adv. Geophys.*, **56**, 169–228.

Drew, J., White, R. S., Tilmann, F., & Tarasewicz, J., 2013. Coalescence microseismic mapping, *Geophys. J. Int.*, **195**(3), 1773–1785.

Eisner, L., Hulsey, B., Duncan, P., Jurick, D., Werner, H., & Keller, W., 2010. Comparison of surface and borehole locations of induced seismicity, *Geophys. Prospect.*, **58**(5), 809–820.

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. J. & Kijko, A., 2013. *An introduction to mining seismology*, vol. 55, Elsevier.

Grigoli, F., Cesca, S., Amoroso, O., Emolo, A., Zollo, A., & Dahm, T., 2013a. Automated seismic event location by waveform coherence analysis, *Geophys. J. Int.*, **196**(3), 1742–1753.

Grigoli, F., Cesca, S., Vassallo, M., & Dahm, T., 2013b. 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., Krieger, L., Kriegerowski, M., Gammaldi, S., Horalek, J., Priolo, E., & Dahm, T., 2016. Automated microseismic event location using master-event waveform stacking, *Sci. Rep.*, **6**, 25744.

Jay, J. A., Pritchard, M. E., West, M. E., Christensen, D., Haney, M., Minaya, E., Sunagua, M., McNutt, S. R., & Zabala, M., 2012. Shallow seismicity, triggered seismicity, and ambient noise tomography at the long-dormant Uturuncu Volcano, Bolivia, *Bull. Volcanol.*, **74**(4), 817–837.

Kao, H. & Shan, S.-J., 2007. Rapid identification of earthquake rupture plane using Source-Scanning Algorithm, *Geophys. J. Int.*, **168**(3), 1011–1020.

Knudsen, S., Havsgård, G., Berg, A., & Bostick, T., 2006. Flow-induced noise in fiber-optic 3C seismic sensors for permanent tubing-conveyed installations, in *68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006*.

Langet, N., Maggi, A., Michelini, A., & Brenguier, F., 2014. Continuous Kurtosis-based migration for seismic event detection and location, with application to Piton de la Fournaise Volcano, La Réunion, *Bull. Seismol. Soc. Am.*, **104**(1), 229–246.

Lavallée, Y., Meredith, P., Dingwell, D., Hess, K., Wassermann, J., Cordonnier, B., Gerik, A., & Kruhl, J., 2008. Seismogenic lavas and explosive eruption forecasting, *Nature*, **453**(7194), 507.

López-Comino, J., Cesca, S., Kriegerowski, M., Heimann, S., Dahm, T., Mirek, J., & Lasocki, S., 2017. Monitoring performance using synthetic data for induced microseismicity by hydrofracking at the Wysin site (Poland), *Geophys. J. Int.*, **210**(1), 42–55.

Maggi, A., Tape, C., Chen, M., Chao, D., & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.*, **178**(1), 257–281.

McNutt, S. R., 1996. Seismic monitoring and eruption forecasting of volcanoes: a review of the state-of-the-art and case histories, in *Monitoring and mitigation of volcano hazards*, pp. 99–146, Springer.

Peng, Z. & Zhao, P., 2009. Migration of early aftershocks following the 2004 Parkfield earthquake, *Nat. Geosci.*, **2**(12), 877.

Power, D. V., Schuster, C. L., Hay, R., Twombly, J., et al., 1976. Detection of hydraulic fracture orientation and dimensions in cased wells, *J. Pet. Technol.*, **28**(09), 1–116.

Pritchard, M., 2009. The life cycle of Andean volcanoes:Combining space-based and field studies, *International Federation of Digital Seismograph Networks. Other/Seismic Network. 10.7914/SN/YS_2009*.

Pritchard, M. & Simons, M., 2004. An InSAR-based survey of volcanic deformation in the central Andes, *Geochem. Geophys. Geosyst.*, **5**(2).

Roach, L. A., White, D. J., & Roberts, B., 2015. Assessment of 4D seismic repeatability and CO2 detection limits using a sparse permanent land array at the Aquistore CO2 storage site, *Geophysics*, **80**(2), WA1–WA13.

Shi, P., Angus, D., Nowacki, A., Yuan, S., & Wang, Y., 2018a. Microseismic Full Waveform Modeling in Anisotropic Media with Moment Tensor Implementation, *Surv. Geophys.*, **39**(4), 567–611.

Shi, P., Angus, D., Rost, S., Nowacki, A., & Yuan, S., 2018b. Automated seismic waveform location using Multichannel Coherency Migration (MCM)–I. Theory, *Geophys. J. Int.*.

Shi, P., Yuan, S., Wang, T., Wang, Y., & Liu, T., 2018c. Fracture Identification in a Tight Sandstone Reservoir: A Seismic Anisotropy and Automatic Multisensitive Attribute Fusion Framework, *IEEE Geosci. Remote Sens. Lett.*, **15**(10), 1525–1529.

Sparks, R. S. J., Folkes, C. B., Humphreys, M. C., Barfod, D. N., Clavero, J., Sunagua, M. C., McNutt, S. R., & Pritchard, M. E., 2008. Uturuncu volcano, Bolivia: Volcanic unrest due to mid-crustal magma intrusion, *Am. J. Sci.*, **308**(6), 727–769.

Verdon, J., Kendall, J.-M., White, D., & Angus, D., 2011. Linking microseismic event observations with geomechanical models to minimise the risks of storing CO2 in geological formations, *Earth Planet. Sci. Lett.*, **305**(1), 143–152.

Verdon, J. P., Kendall, J.-M., White, D. J., Angus, D. A., Fisher, Q. J., & Urbancic, T., 2010. Passive seismic monitoring of carbon dioxide storage at Weyburn, *The Leading Edge*.

Wilks, M., Kendall, J.-M., Nowacki, A., Biggs, J., Wookey, J., Birhanu, Y., Ayele, A., & Bedada, T., 2017. Seismicity associated with magmatism, faulting and hydrothermal circulation at Aluto Volcano, Main Ethiopian Rift, *J. Volcanol. Geoth. Res.*.

Xu, N., Tang, C., Li, L., Zhou, Z., Sha, C., Liang, Z., & Yang, J., 2011. Microseismic monitoring and stability analysis of the left bank slope in Jinping first stage hydropower station in southwestern China, *Int. J. Rock Mech. Min. Sci.*, **48**(6), 950–963.

Yuan, S., Wang, S., Ma, M., Ji, Y., & Deng, L., 2017. Sparse Bayesian learning-based time-variant deconvolution, *IEEE Trans. Geosci. Remote Sens.*, **55**(11), 6182–6194.

Yuan, S., Liu, J., Wang, S., Wang, T., & Shi, P., 2018a. Seismic waveform classification and first-break picking using convolution neural networks, *IEEE Geosci. Remote Sens. Lett.*, **15**(2), 272–276.

Yuan, S., Wang, S., Luo, C., & Wang, T., 2018b. Inversion-Based 3-D Seismic Denoising for Exploring Spatial Edges and Spatio-Temporal Signal Redundancy, *IEEE Geosci. Remote Sens. Lett.*, (99), 1–5.

Zhebel, O. & Eisner, L., 2014. Simultaneous microseismic event localization and source mechanism determination, *Geophysics*, **80**(1), KS1–KS9.

Zhu, L. & Rivera, L. A., 2002. A note on the dynamic and static displacements from a point source in multilayered media, *Geophys. J. Int.*, **148**(3), 619–627.

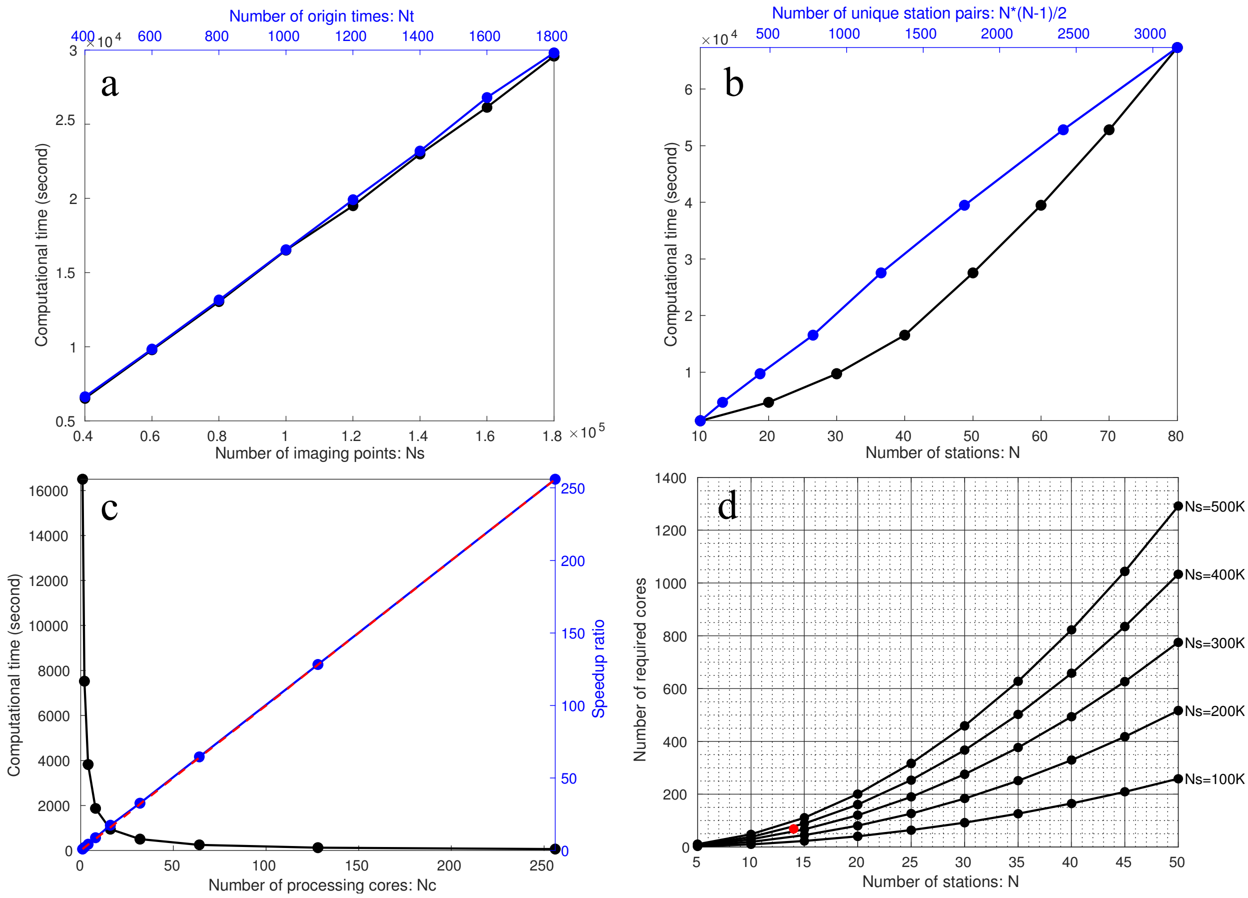


Figure 1. Computational efficiency analysis of the MCM location. The efficiency test is performed on the Intel E5-2670(2.6GHz) processor. (a) The computational times for different numbers of imaging points and searched origin times. Black line and the bottom X-axis show the variation of computational times with the number of imaging points, when the number of origin times and stations are fixed as 1000 and 40. Blue line and the top X-axis show the variation of computational times with the number of origin times, when the number of imaging points and stations are fixed as 100000 and 40. (b) The computational times for different numbers of stations, when the number of origin times and imaging points are fixed as 1000 and 100000. Black line and the bottom X-axis show the variation of computational time with the number of stations. Blue line and the top X-axis show the variation of computational time with the number of unique station pairs. Program runs on one core for (a) and (b). (c) The computational times (black line and left Y-axis) and speedup ratios (blue line and right Y-axis) when different numbers of cores are used. Red dashed line show the theoretical speedup ratios. The number of origin times, imaging points and stations are fixed as 1000, 100000 and 40. (d) The required cores used for real time processing under different numbers of stations and imaging points. Different black lines show the scenarios for different numbers of imaging points. The red dot shows the scenario for the following Uturuncu dataset, where 14 stations are deployed, 283556 imaging points are scanned and 68 cores are required for real time processing.

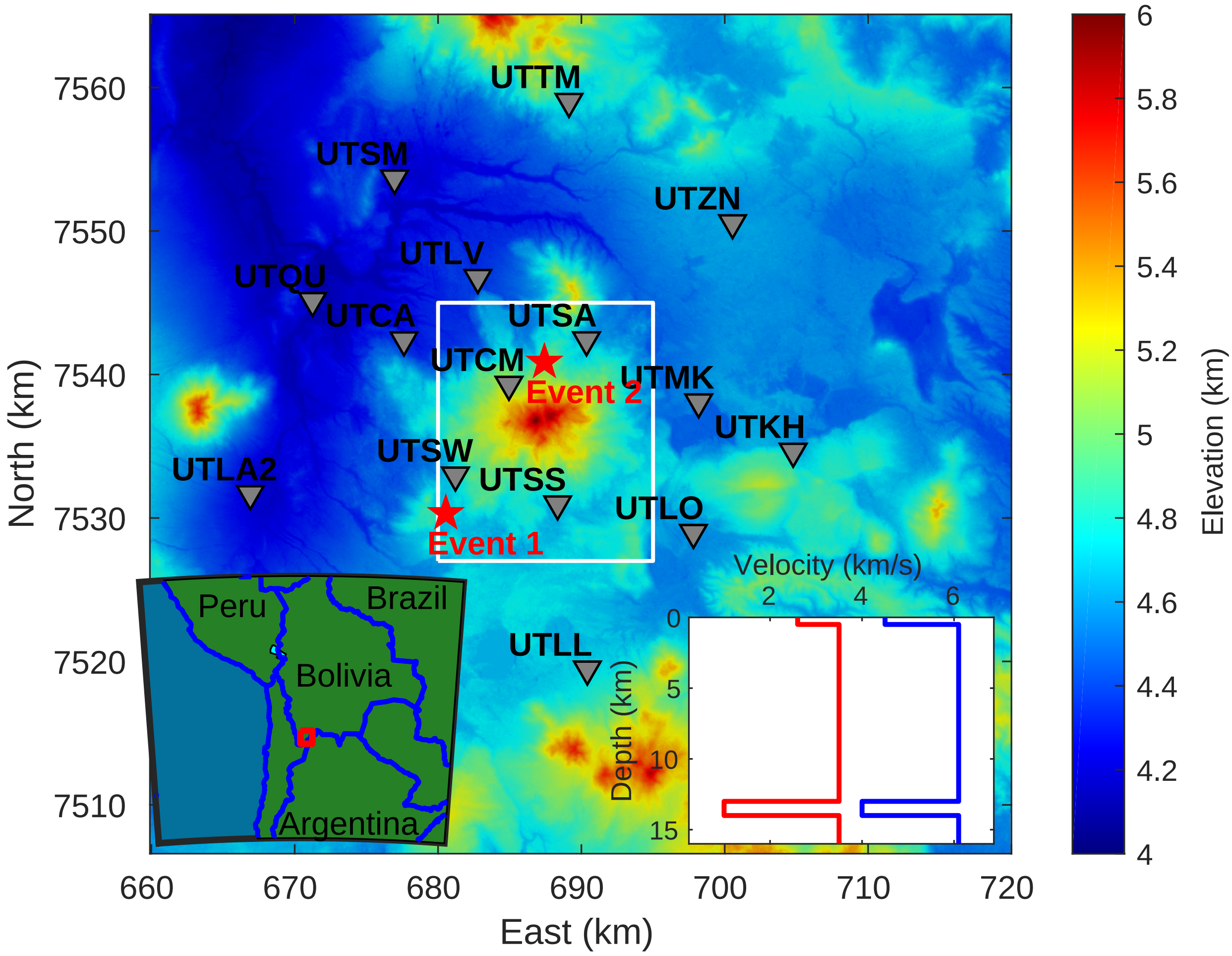
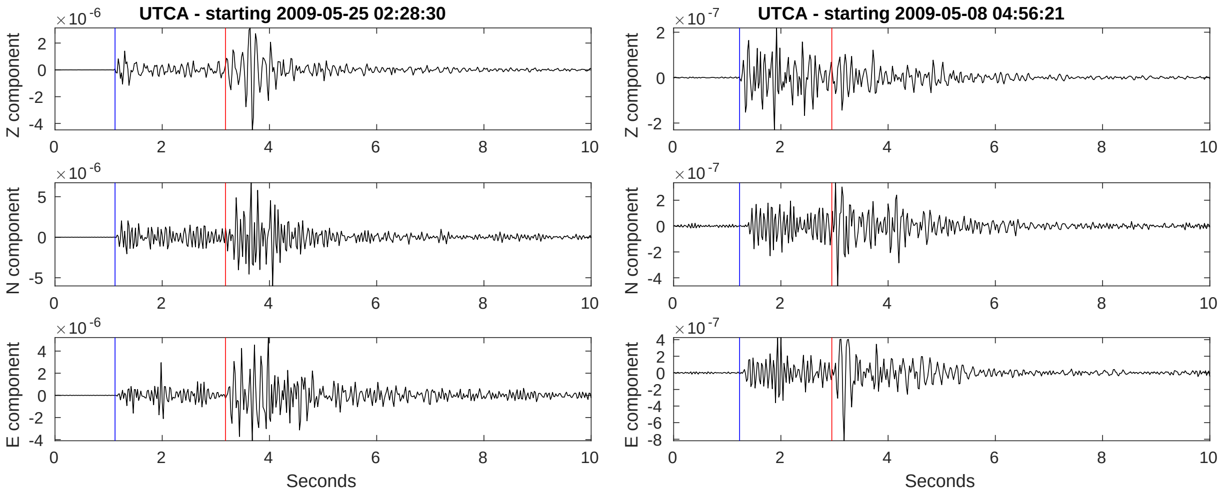


Figure 2. Location of the seismic stations and Uturuncu volcano (UTM zone: 19K). The stations are represented by gray triangles. Two local volcano-tectonic earthquakes in the catalogue are represented by red stars. The Uturuncu is located in the middle of the figure. The color in the figure represents elevation relative to the sea level. The lower left part exhibits a regional map, in which the red rectangle shows the research area. The lower right part exhibits the velocity model used in the event location, in which the red and blue lines show the P- and S-wave velocities. The white rectangle shows the imaging area (shown as Figure 11) for the four hours of continuous data.



(a) (b)

Figure 3. The recorded three component waveforms at station UTCA for the two shallow, local volcano-tectonic earthquake. The blue and red lines show the arrivals of P- and S-waves respectively. (a) Waveforms for the first event. The instrument response has been removed and the waveforms are filtered using a band-pass filter of 5-23 Hz. (b) Waveforms for the second event. The instrument response has been removed and the waveforms are filtered using a bandpass filter of 5-21 Hz.

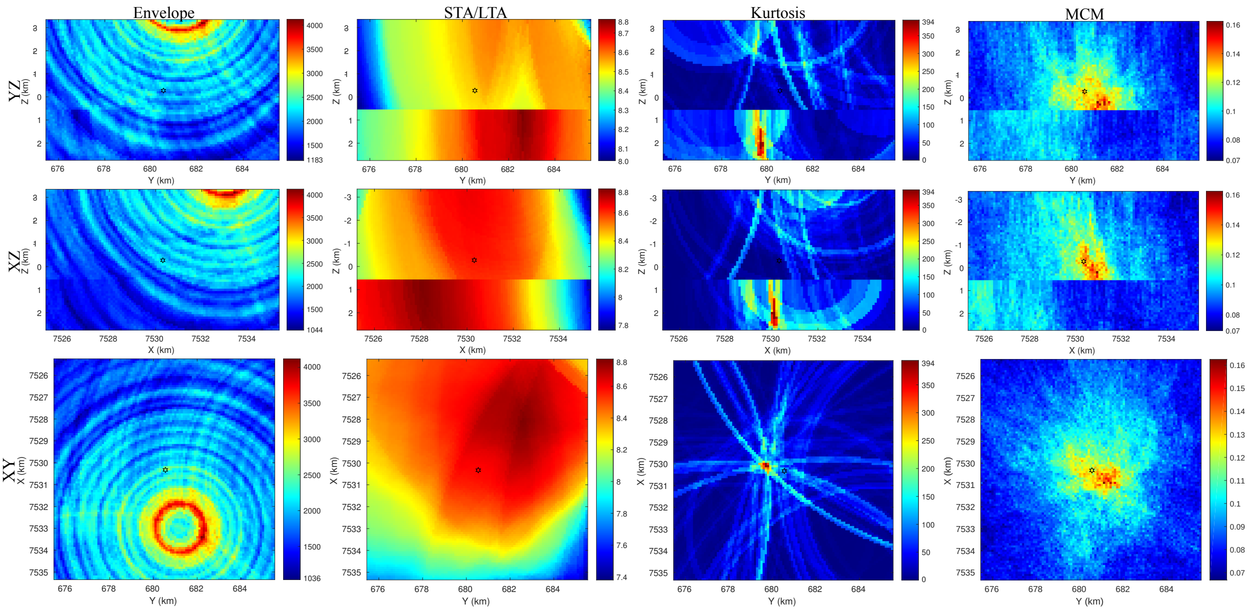


Figure 4. Migration profiles through the maximum migrated value for the first volcano-tectonic earthquake. The dark stars show the corresponding seismic event in the catalogue obtained by manual picking. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows YZ profiles, second row shows XZ profiles, third row shows XY profiles.

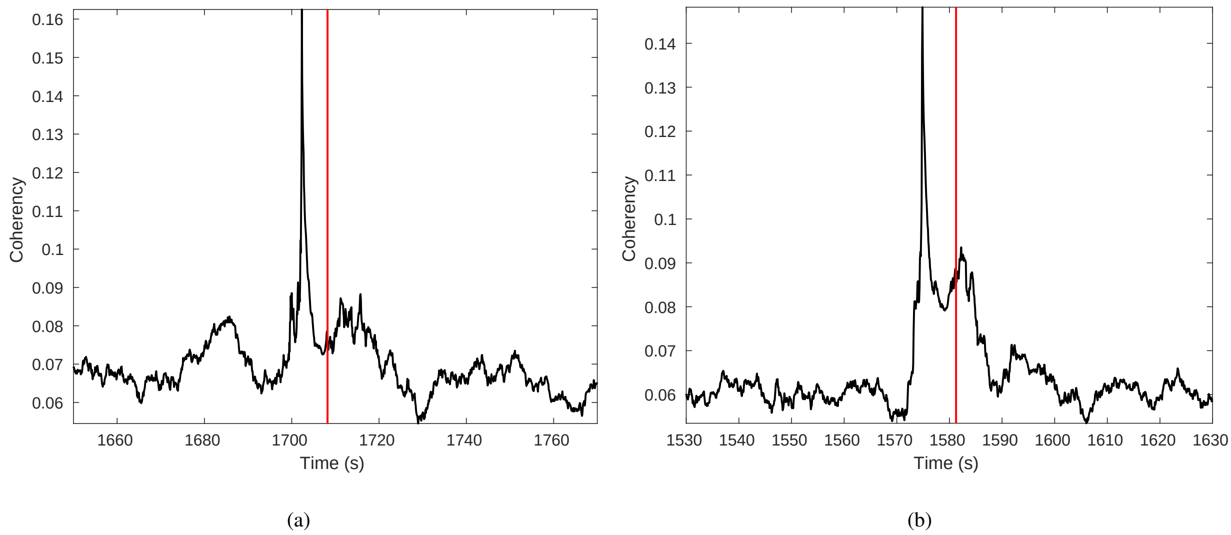


Figure 5. The stacking functions of the MCM method at the position of the maximum migrated value. The red line shows the origin time of the event in the catalogue obtained by manual picking. (a) The stacking function for event 1. The time is relative to 2009-05-25 02:00:00. (b) The stacking function for event 2. The time is relative to 2009-05-08 04:30:00.

Table 1. Location results of different waveform migration methods for the Uturuncu shallow volcano-tectonic earthquake and comparison with the event in the catalogue. The origin time is relative to 2009-05-25 02:00:00 (UTC).

	Event location				Deviation from manual traveltime location			
	X (km)	Y (km)	Z (km)	T_0 (s)	ΔX (m)	ΔY (m)	ΔZ (m)	ΔT_0 (s)
Catalogue	7530.316	680.543	-0.269	1708.2	-	-	-	-
Envelope	7533.4	682.2	-3.2	1709.8	3084	1657	2931	1.6
STA/LTA	7528.0	682.6	1.0	1709.4	2316	2057	1269	1.2
Kurtosis	7530.1	679.7	2.0	1708.4	216	843	2269	0.2
Coherency	7530.9	681.1	0.2	1708.3	584	557	469	0.1

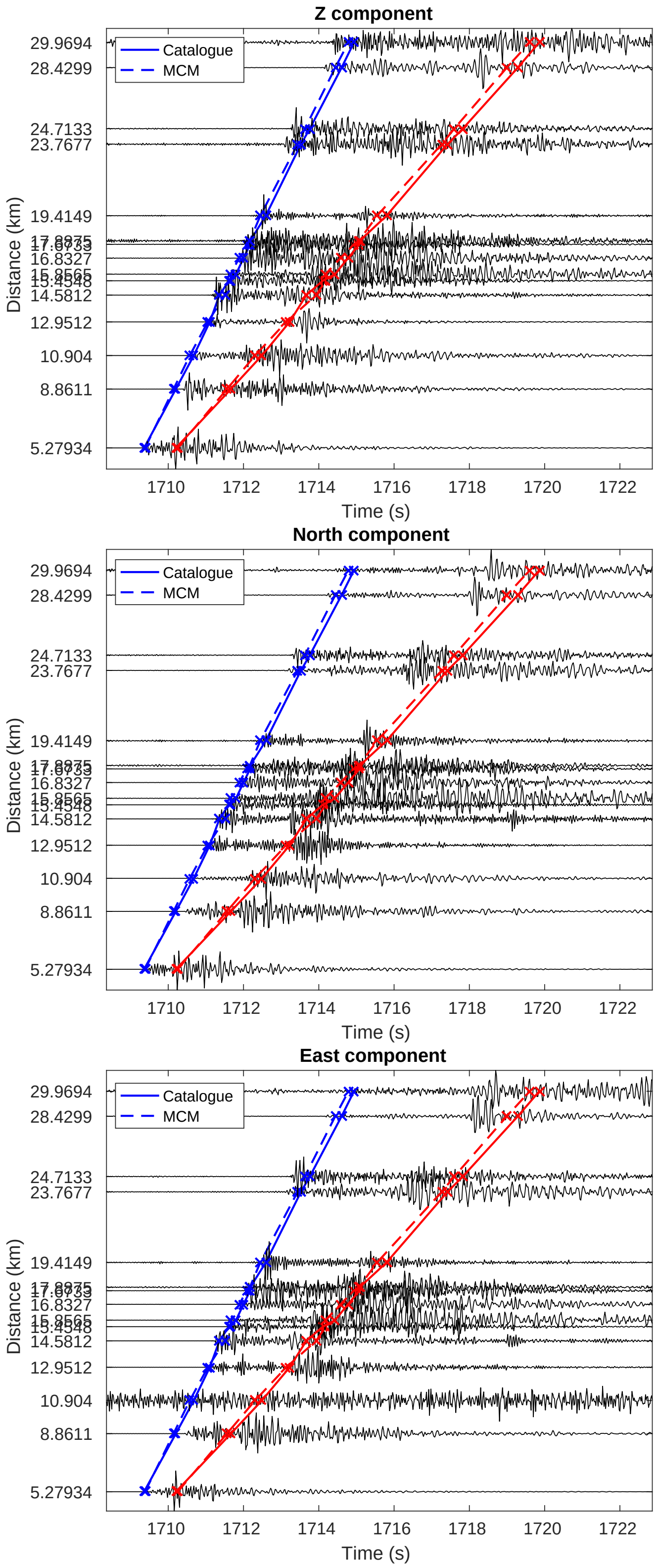


Figure 6. The three component record sections of the first event. The predicted P- and S-wave arrival times for this event in the catalogue and the event located by MCM are marked by solid and dashed lines respectively. The blue and red colors show the arrival times of the direct P- and S-wave respectively. The time in the figure is relative to 2009-05-25 02:00:00 (UTC).

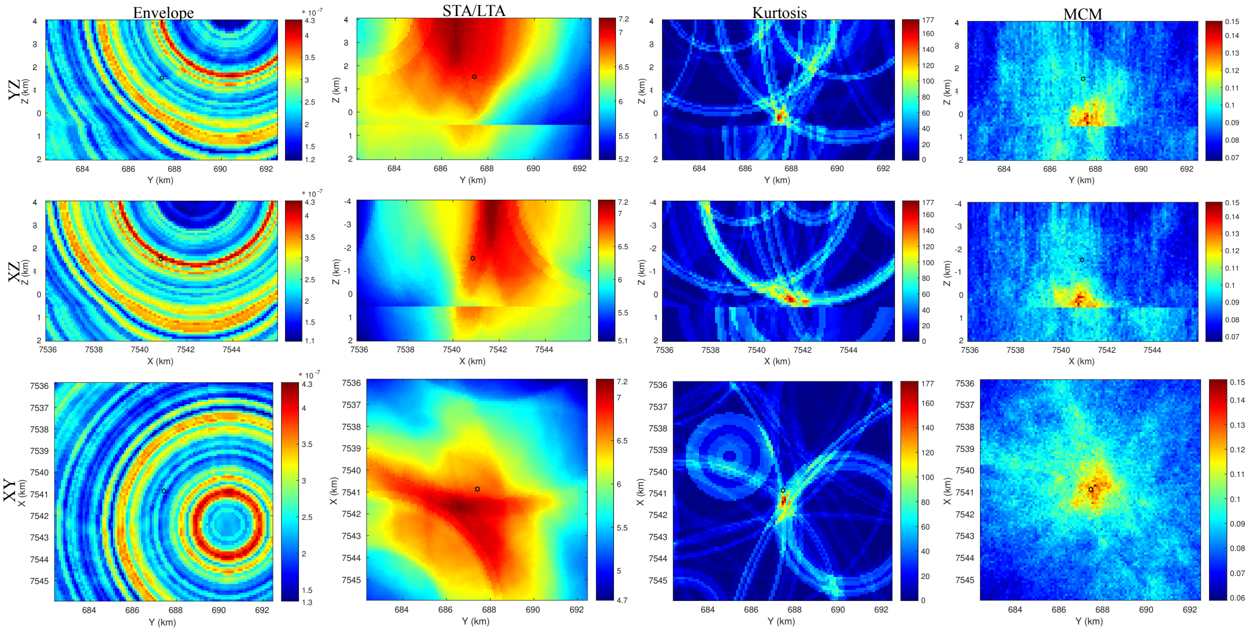


Figure 7. Migration profiles through the maximum migrated value for the second volcano-tectonic earthquake. The dark stars show the corresponding seismic event in the catalogue obtained by manual picking. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows YZ profiles, second row shows XZ profiles, third row shows XY profiles.

Table 2. Location results of different waveform migration methods for the Uturuncu shallow volcano-tectonic earthquake and comparison with the event in the catalogue. The origin time is relative to 2009-05-08 04:30:00 (UTC).

	Event location				Deviation from manual traveltime location			
	X (km)	Y (km)	Z (km)	T_0 (s)	ΔX (m)	ΔY (m)	ΔZ (m)	ΔT_0 (s)
Catalogue	7540.866	687.419	-1.523	1581.3	-	-	-	-
Envelope	7540.9	690.5	-1.6	1582.2	34	3081	77	0.9
STA/LTA	7541.6	686.6	-3.6	1582.9	734	819	2077	1.6
Kurtosis	7541.4	687.4	0.2	1580.9	534	19	1723	0.4
Coherency	7540.7	687.6	0.4	1580.7	166	181	1923	0.6

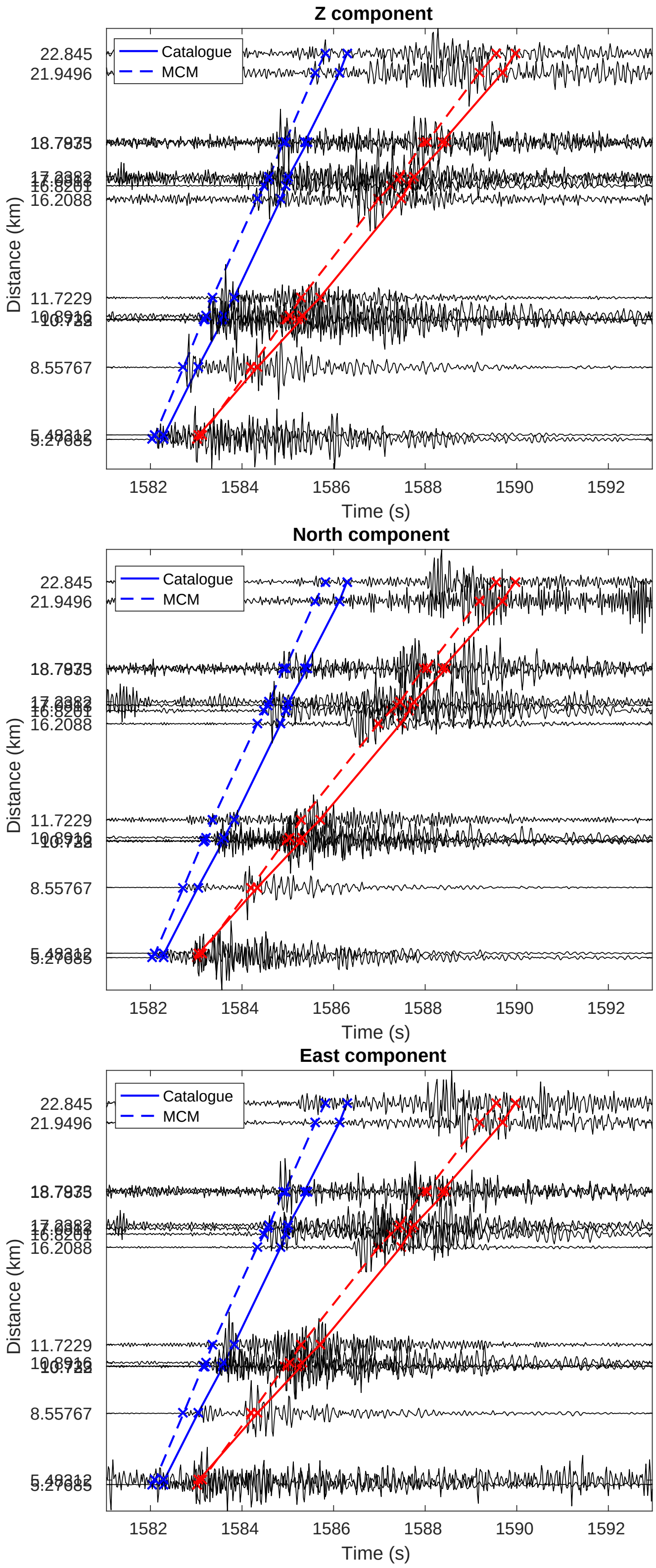


Figure 8. The three component record sections of the second event. The predicted P- and S-wave arrival times for this event in the catalogue and the event located by MCM are marked by solid and dashed lines respectively. The blue and red colors show the arrival times of the direct P- and S-wave respectively. The time in the figure is relative to 2009-05-08 04:30:00 (UTC).

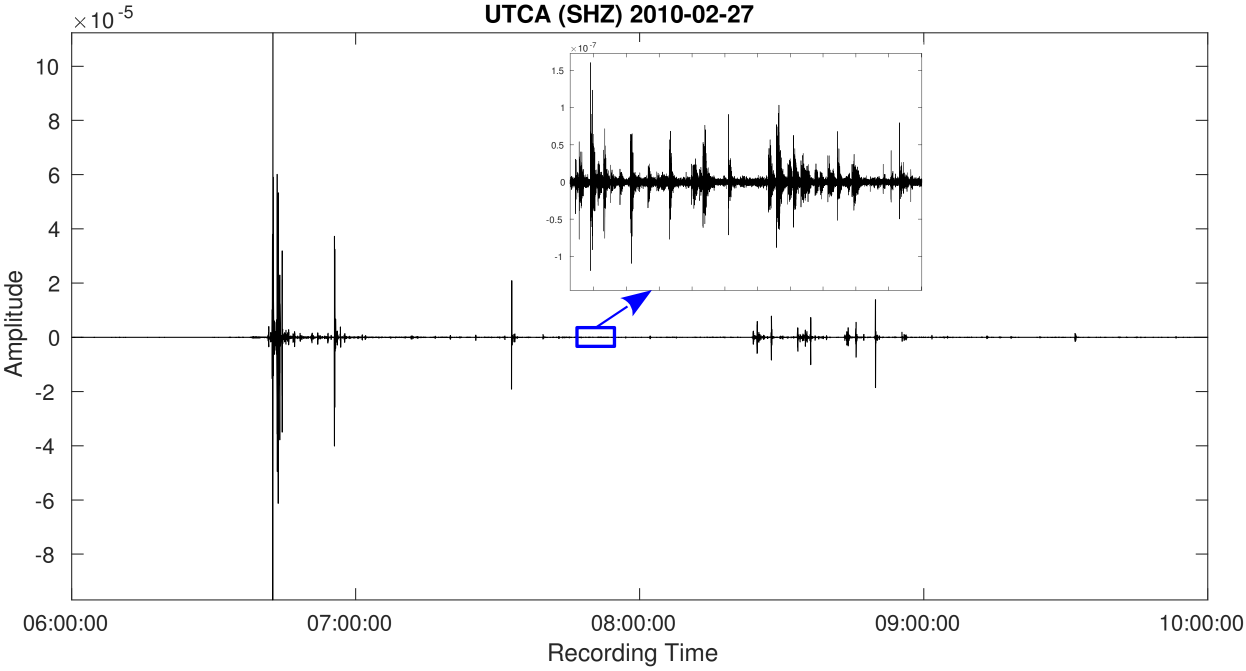


Figure 9. The recorded Z component waveforms at station UTCA. The recording time ranges from 06:00:00 to 10:00:00 (UTC). The waveforms within the blue rectangle are enlarged. The instrument response has been removed and the waveforms are filtered using a bandpass filter of 4.2-21.6 Hz.

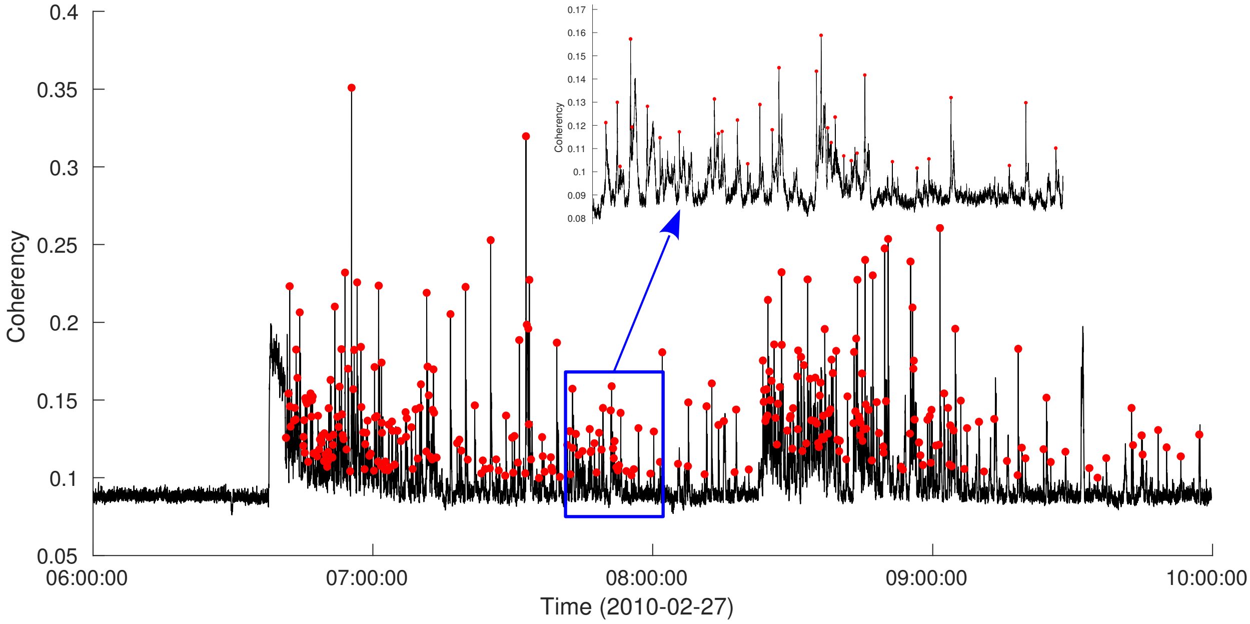


Figure 10. The maximum coherency value at each searched origin time for the four hours of continuous data. The time interval is 0.08 s. The part in the blue rectangle is enlarged. The red points show the 322 verified seismic events.

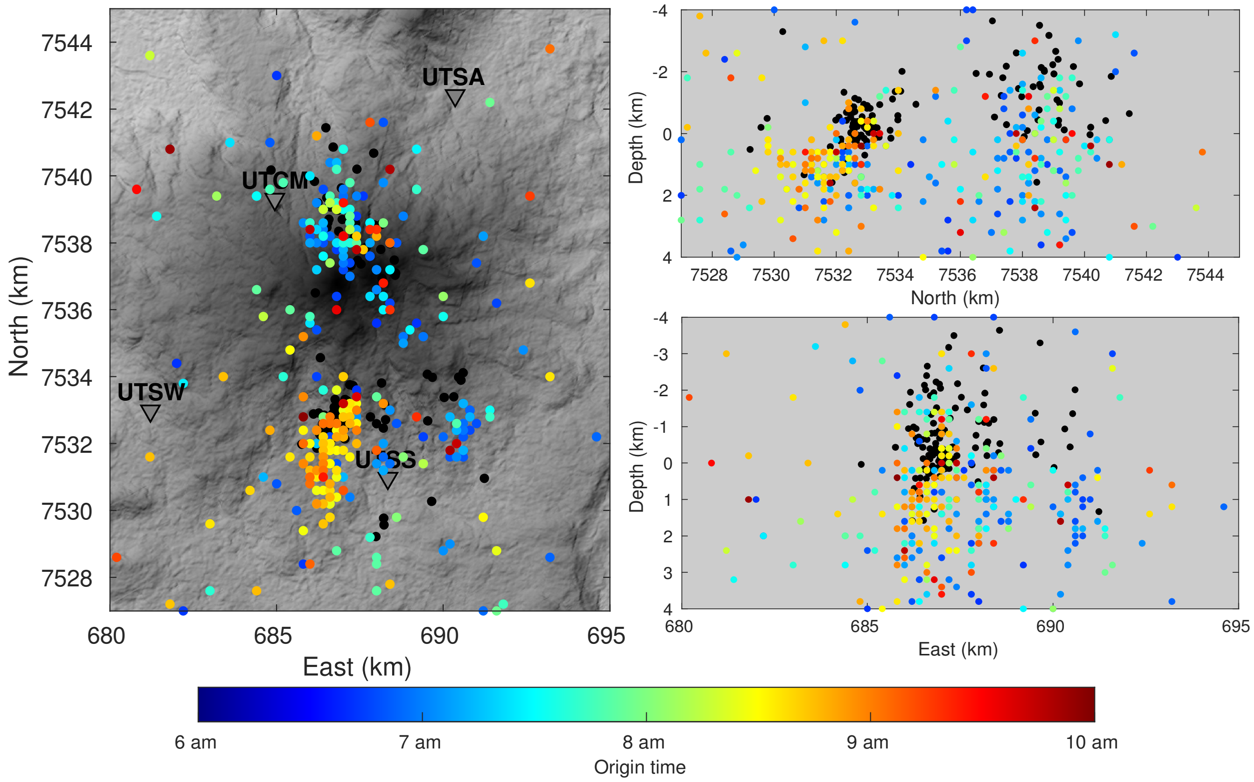


Figure 11. The earthquake locations on the horizontal and vertical profiles. Black dots show the 114 event locations in the existing catalogue. The color-coded dots show the verified 322 event locations for the MCM. The color represents the origin times of earthquake events.

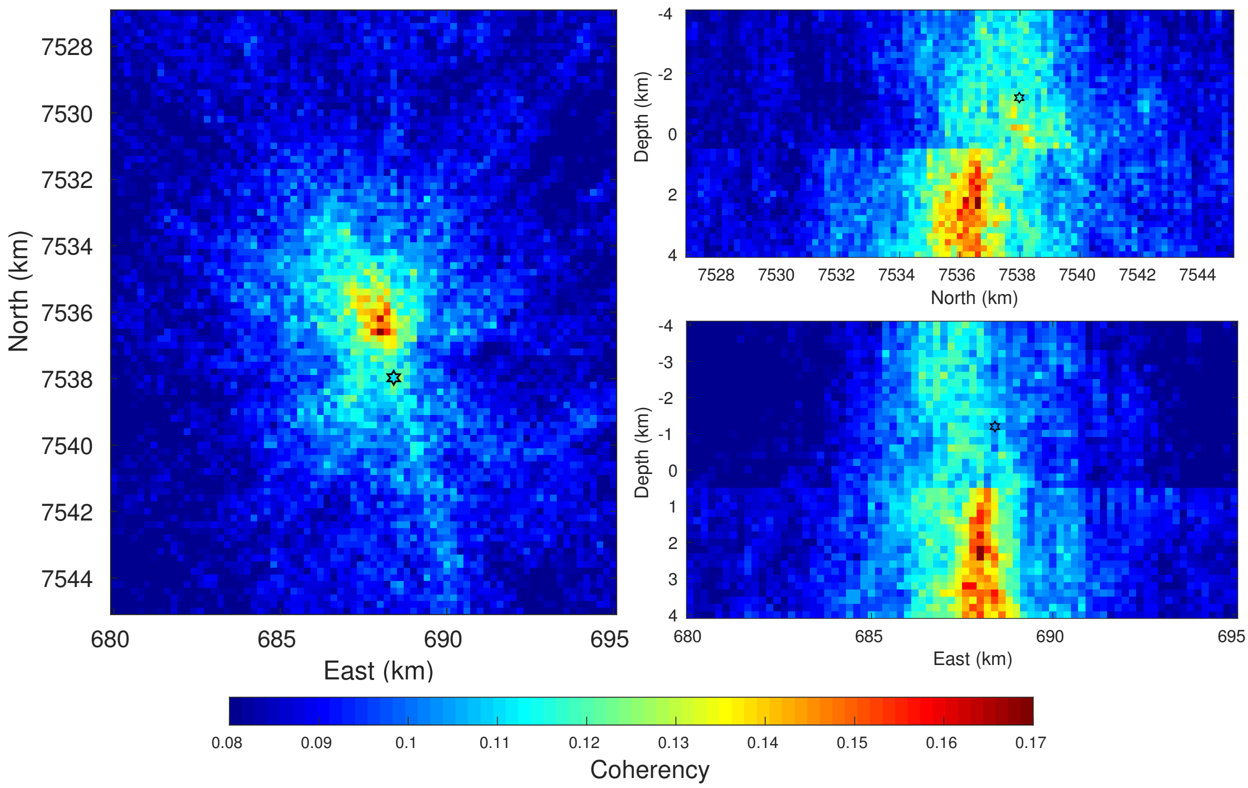


Figure 12. Horizontal and vertical profiles at the maximum value of the migration volume for seismic event 1. Color represents the migration value (coherency). Black star represents the event location in the catalogue.

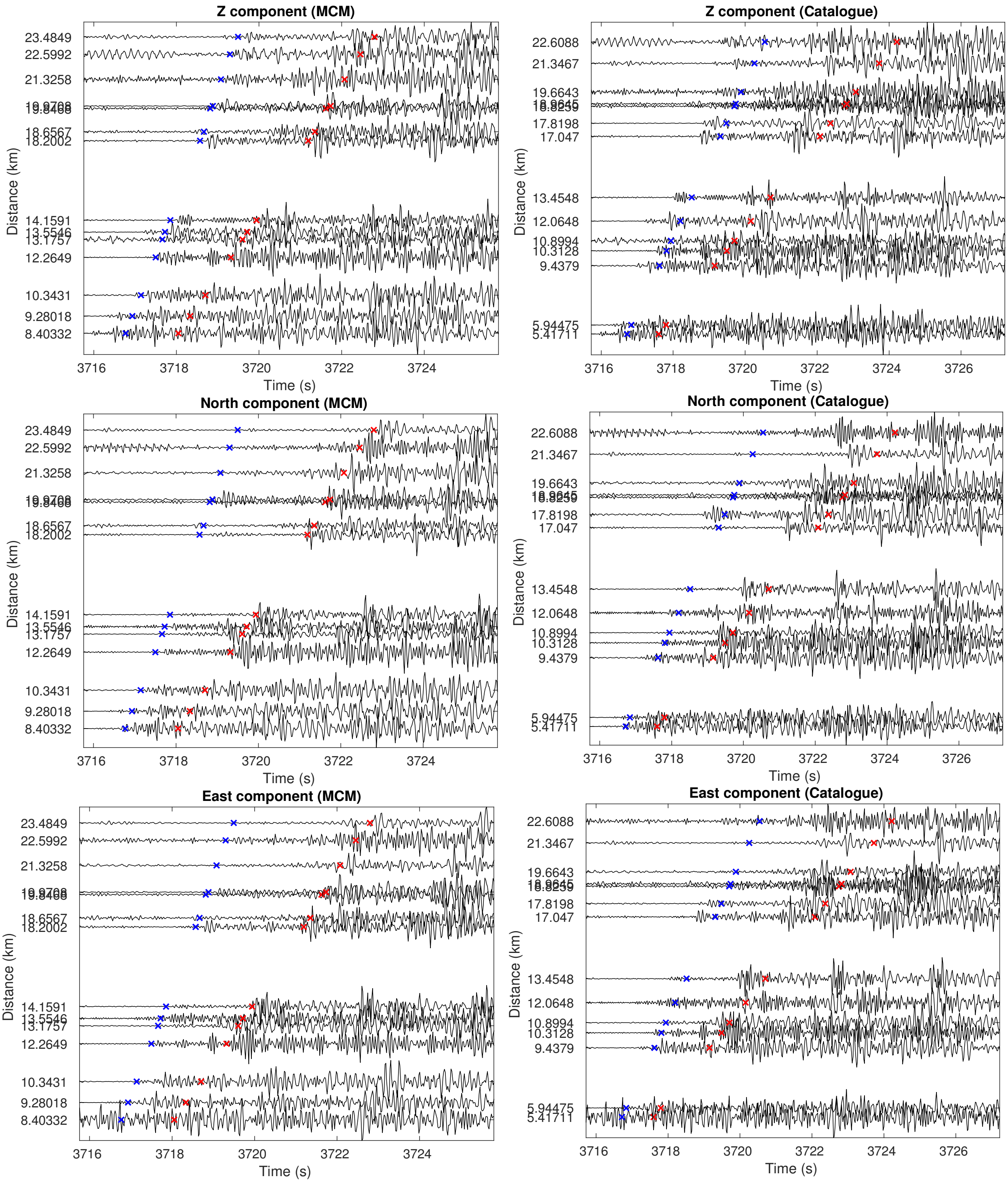


Figure 13. Three component record sections for seismic event 1. The predicted P- and S-wave arrival times are marked by blue and red crosses, respectively. Left panel: record sections for the MCM location result. Right panel: record sections for the catalogue location result. The time in the figure is relative to 2010-02-27 06:00:00 (UTC).

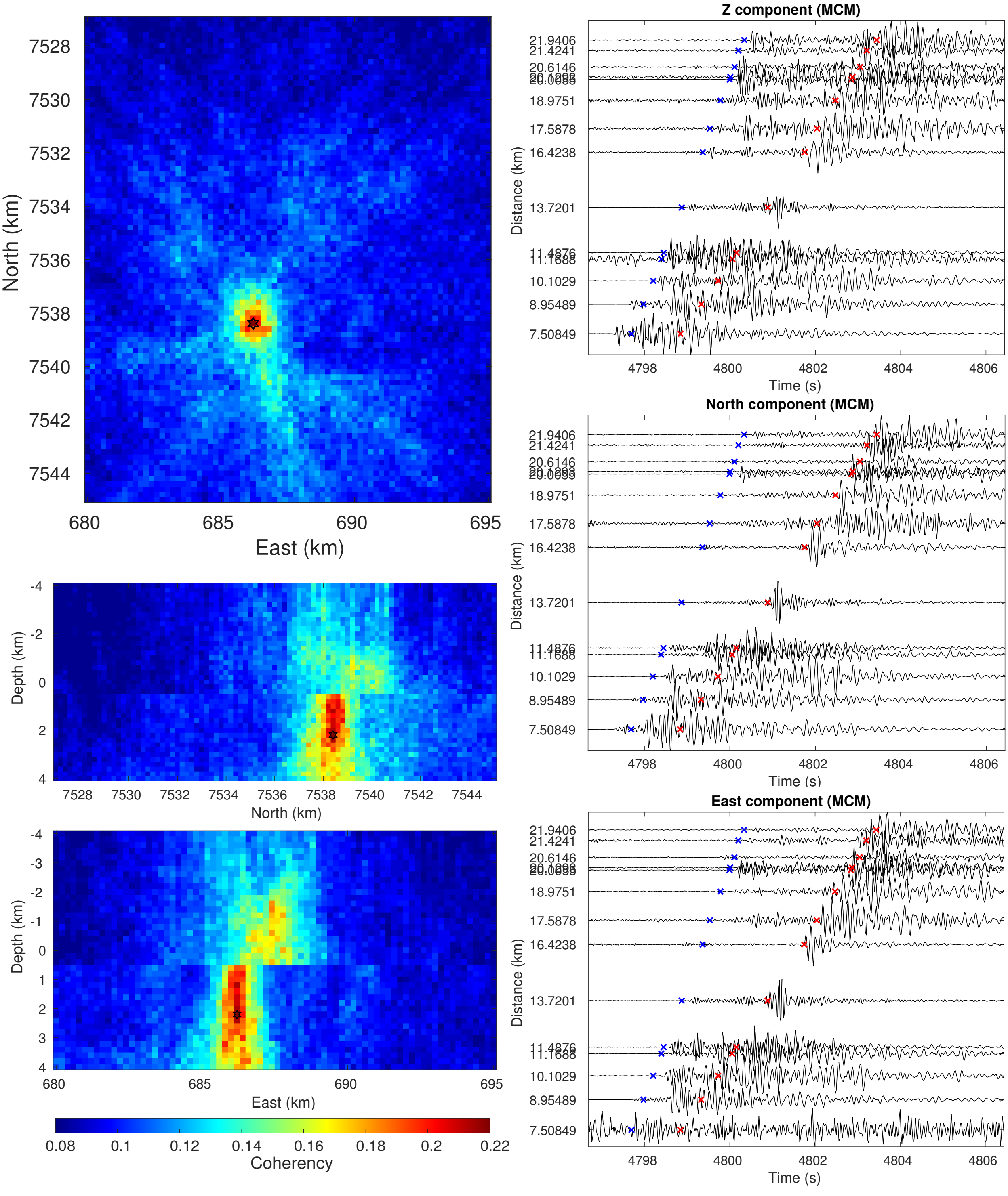


Figure 14. Horizontal and vertical migration profiles and three component record sections for seismic event 2, which is newly detected by MCM and not in the existing catalogue. The predicted P- and S-wave arrival times are marked by blue and red crosses on the record sections, respectively. Left panel: horizontal and vertical profiles at the maximum value of the migration volume. Color represents the migration value and black star shows the final event location. Right panel: record sections for this event.

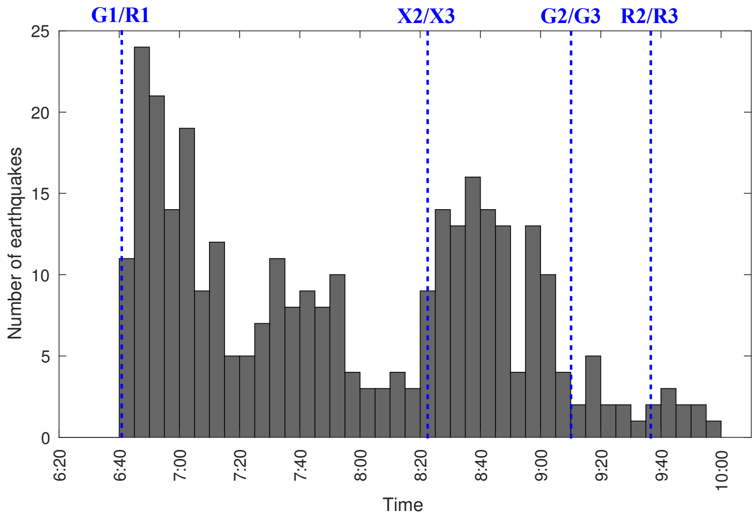


Figure 15. Histogram of Uturuncu triggered events from the M_w 8.8 Maule earthquake for the four hours (6 am to 10 am) in 5-min bins. There are no seismic events from 6:00 am to 06:40 am. Blue dashed lines show the approximate arrival time of surface wave trains. G1/R1 represents the minor-arc Love (G1) and Rayleigh (R1) waves. X2/X3, G2/G3 and R2/R3 represent different surface wave overtones (Jay et al. 2012).

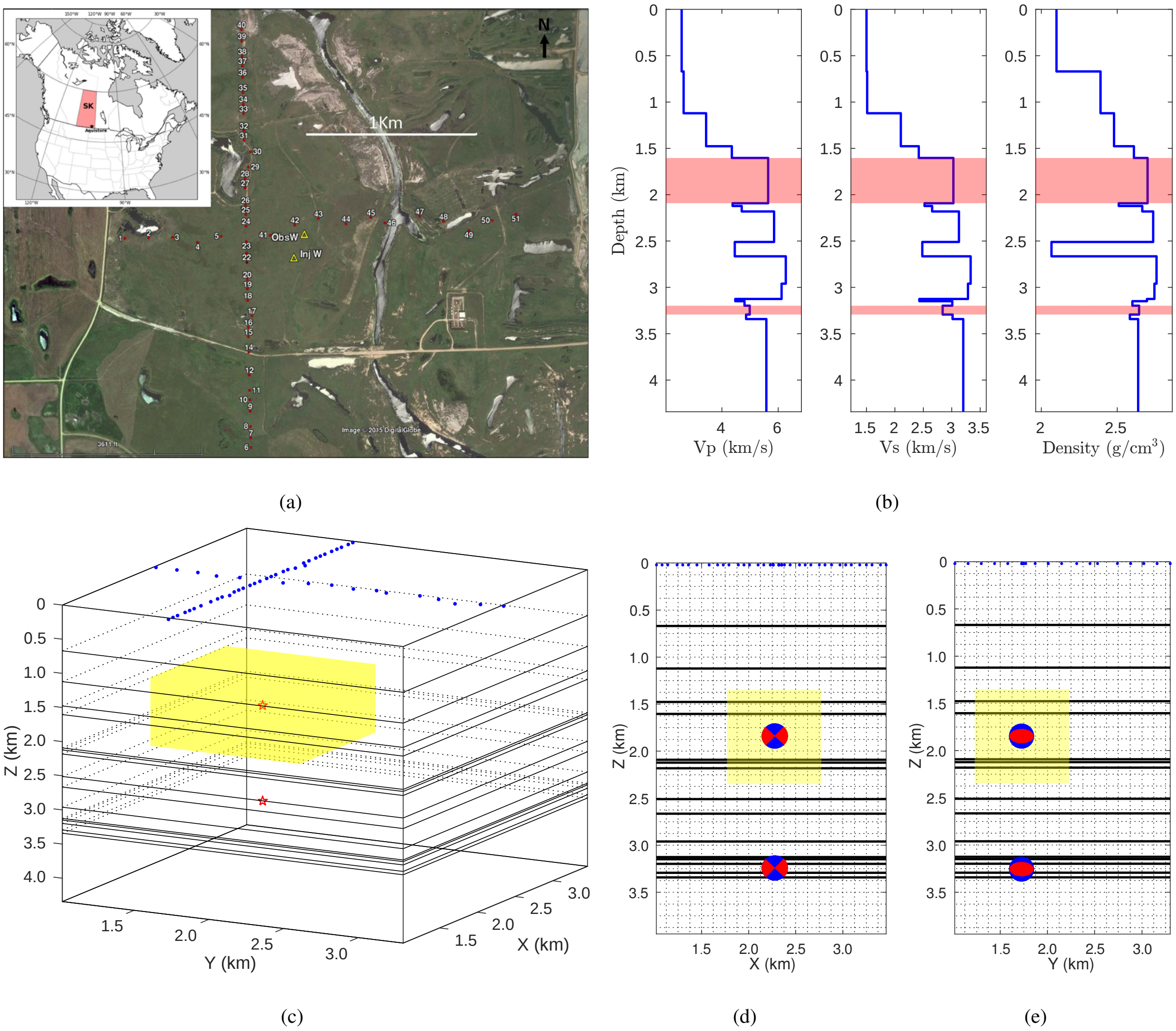


Figure 16. (a) Aquistore permanent seismic array geometry. Geophones are denoted by red dots alongside the station number, while the observation and injection wells are illustrated by yellow triangles. (From Birnie et al. (2016)). (b) P- and S-wave velocity model and density model in Aquistore area. The red color highlights two target layers where the seismic events are located. (c) The numerical model space of the Aquistore area. Vertical (d) XZ and (e) YZ profiles of the numerical model. The red stars shows the locations of two seismic events, whose depth are 1.85 km and 3.25 km respectively. Blue points represent the surface geophones. The yellow color exhibits the imaging area of the shallow event. Source radiation patterns are shown in the vertical profiles using a beach ball with red and blue colors.

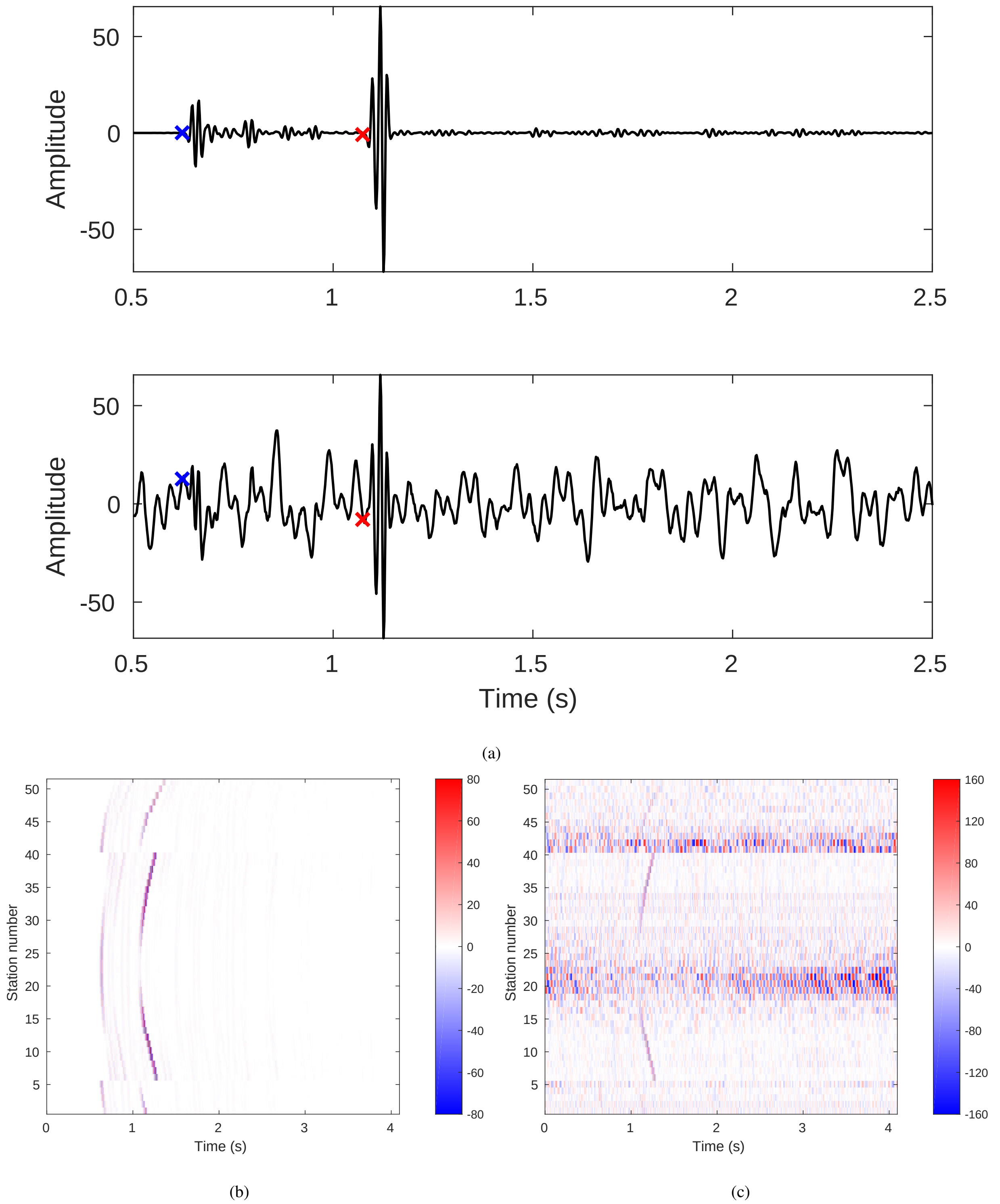


Figure 17. (a) The recorded waveform data at station 30 before (top) and after (bottom) adding real noise. The blue and red crosses show the arrivals of P- and S-phases. (b) The synthetic noise-free seismic profile. (c) The seismic profile after adding real noise. The SNR is 0.5.

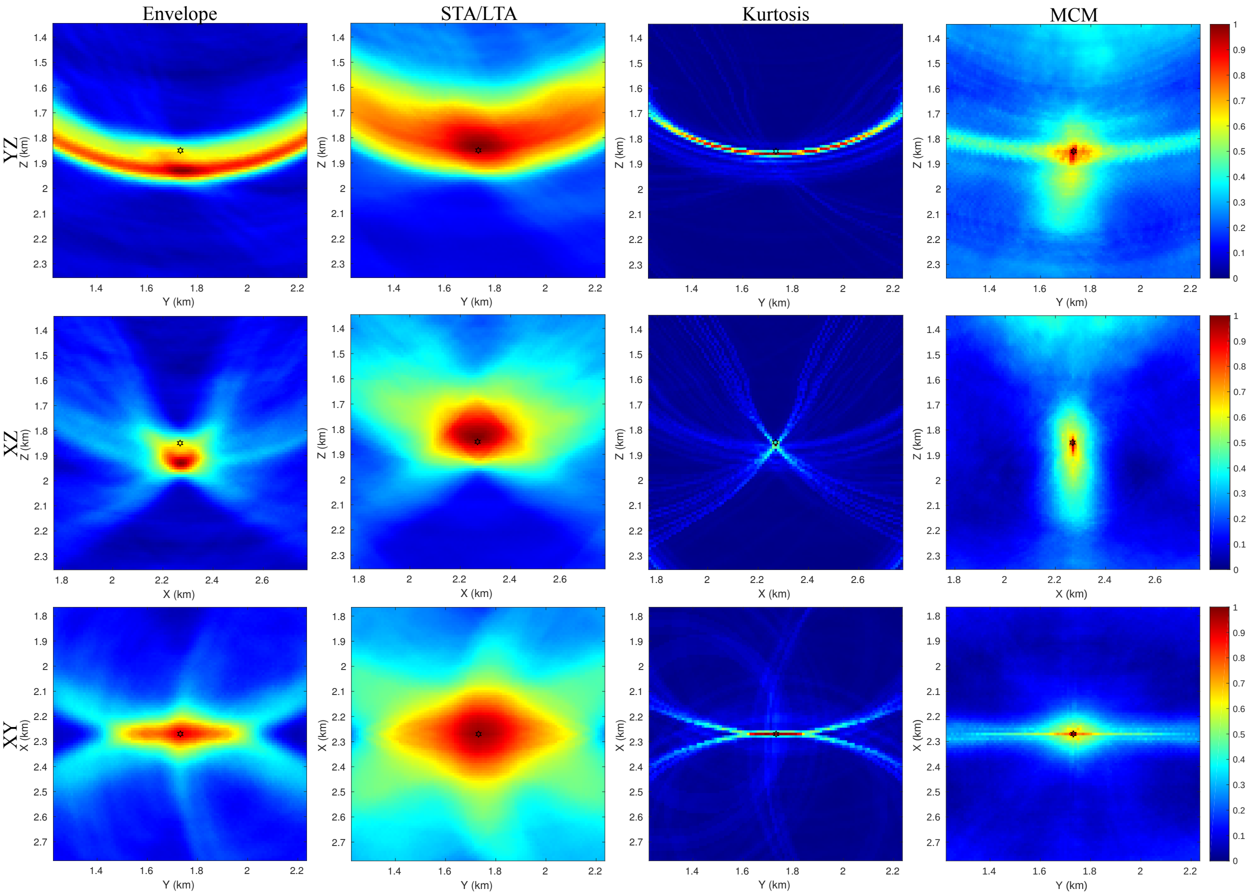


Figure 18. Profiles of the migration results through the true source location for the four methods. The SNR is 1. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows YZ profiles, second row shows XZ profiles, third row shows XY profiles.

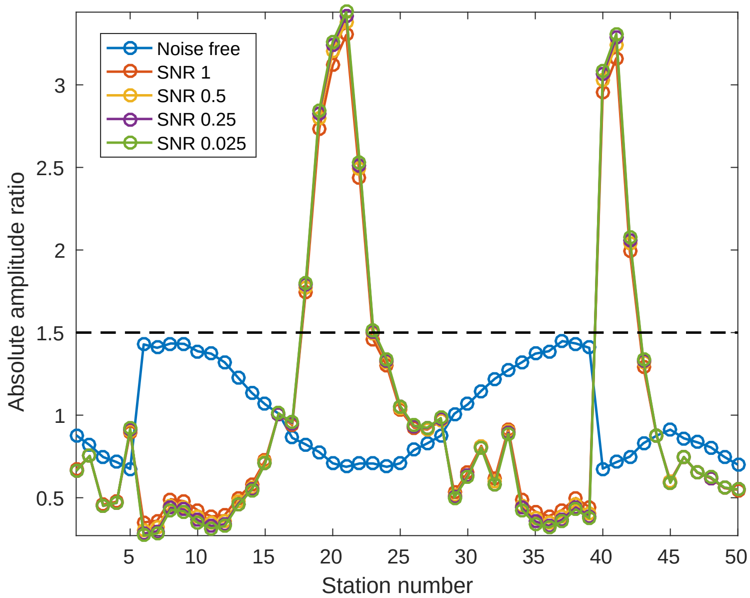


Figure 19. The absolute amplitude ratios for different stations under different SNR scenarios. The absolute amplitude ratio of different traces is defined as the ratio of the average absolute amplitudes of a trace to the average absolute amplitude of all traces. The black dashed line shows an absolute amplitude ratio of 1.5.

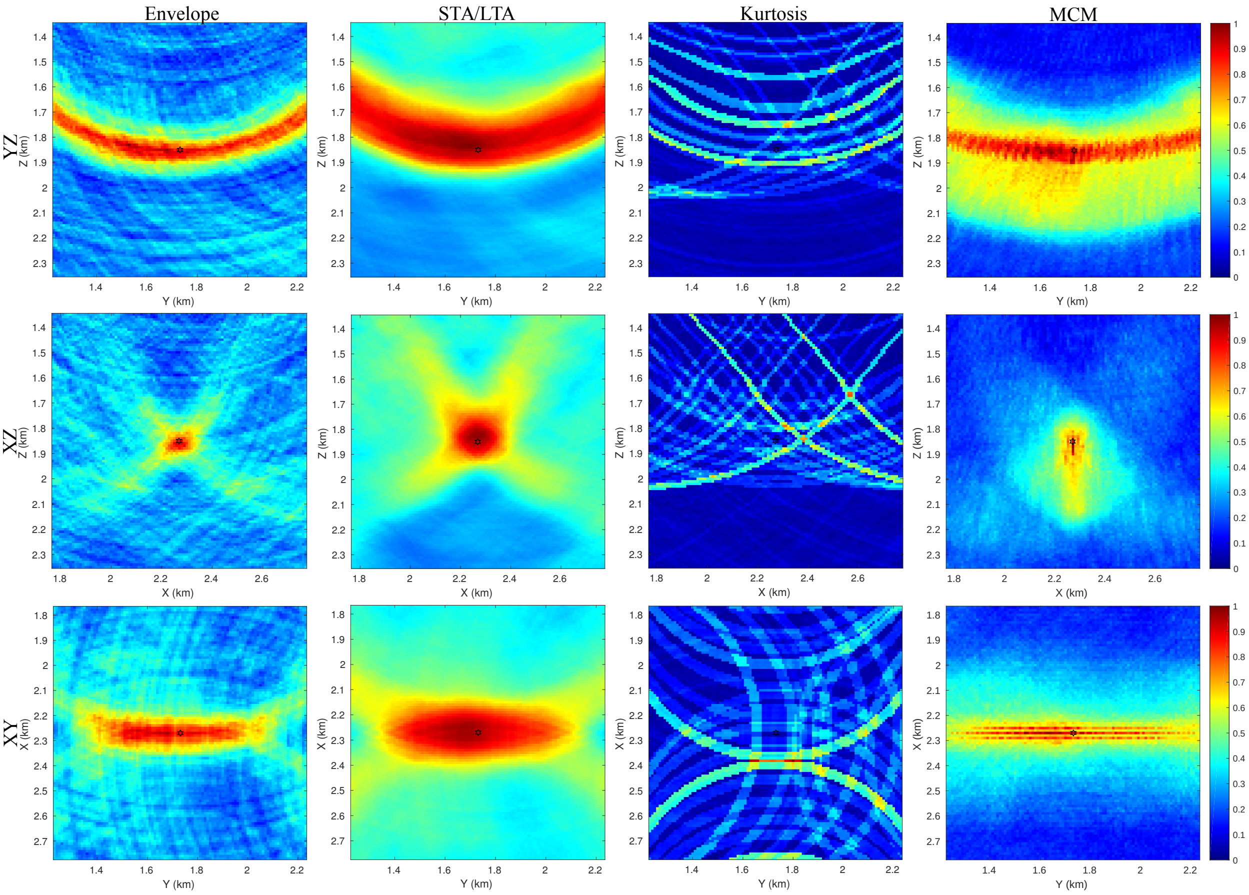


Figure 20. Profiles of the migration results through the true source location with automatic quality control scheme (weighting and filtering). The SNR is 0.025. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows YZ profiles, second row shows XZ profiles, third row shows XY profiles.

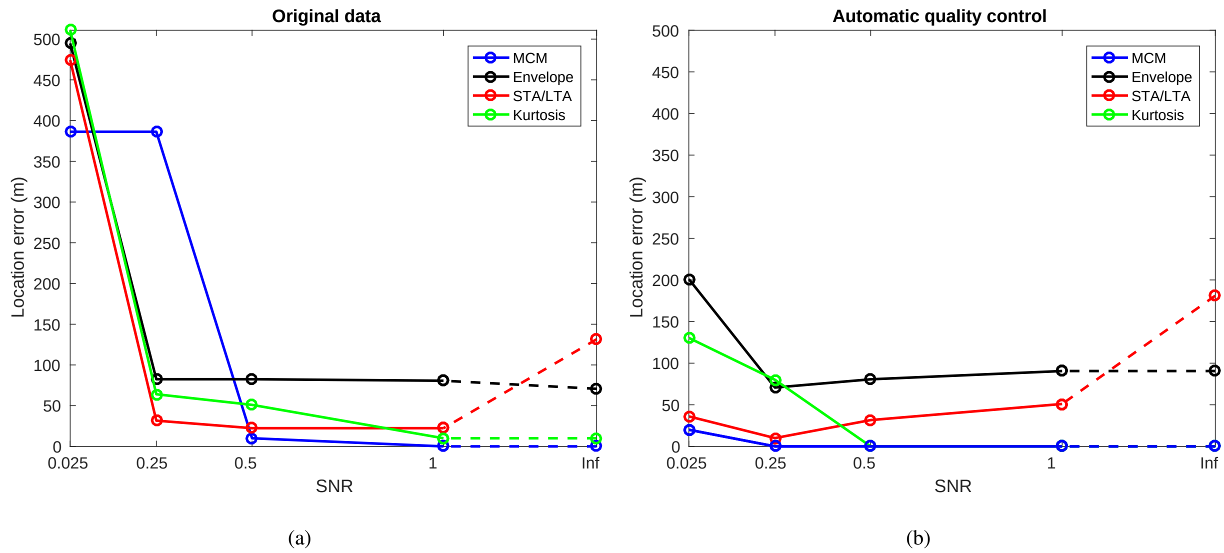


Figure 21. The location errors of the four methods under different SNRs with (a) original data and (b) automatic quality control scheme (weighting and filtering).

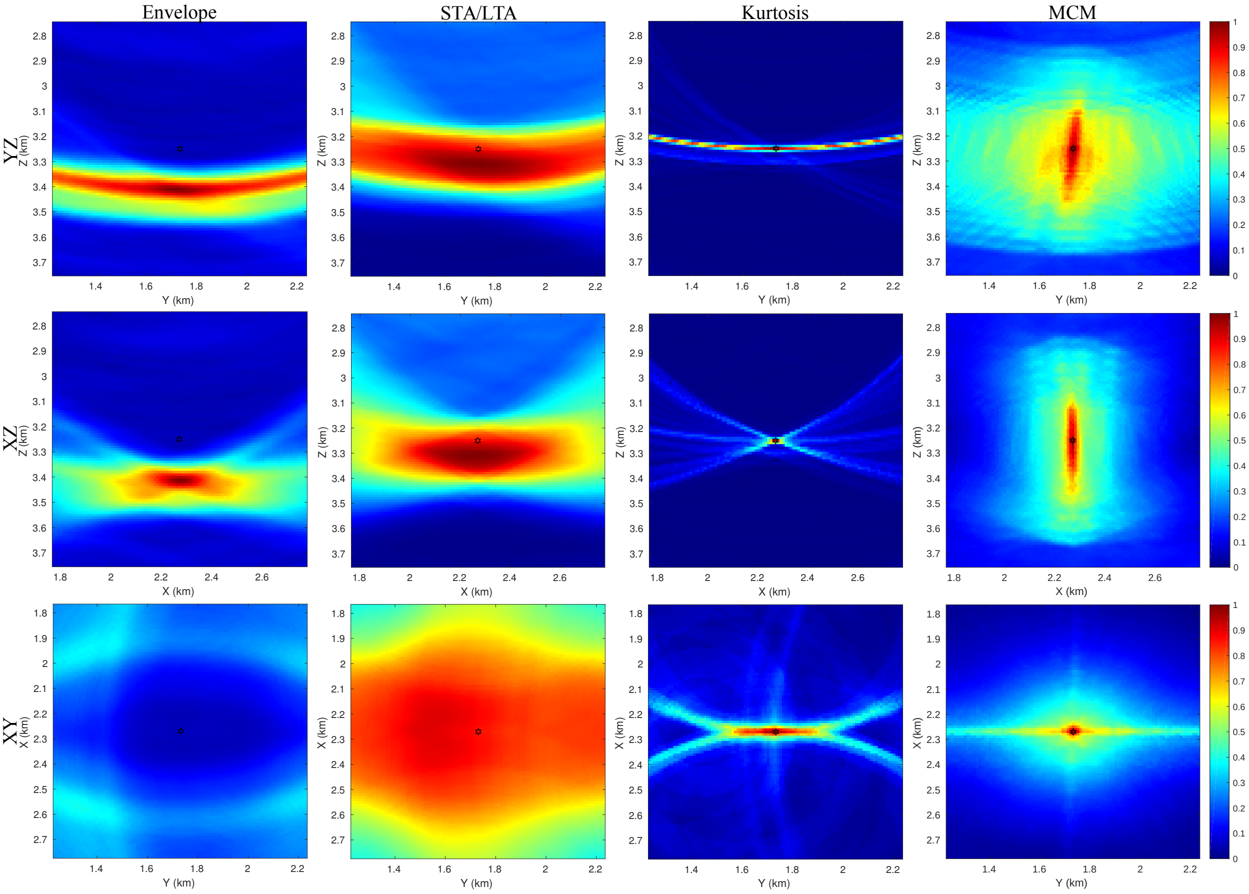


Figure 22. Migration profiles through the true source location of the deep event. The SNR is 1. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows YZ profiles, second row shows XZ profiles, third row shows XY profiles.

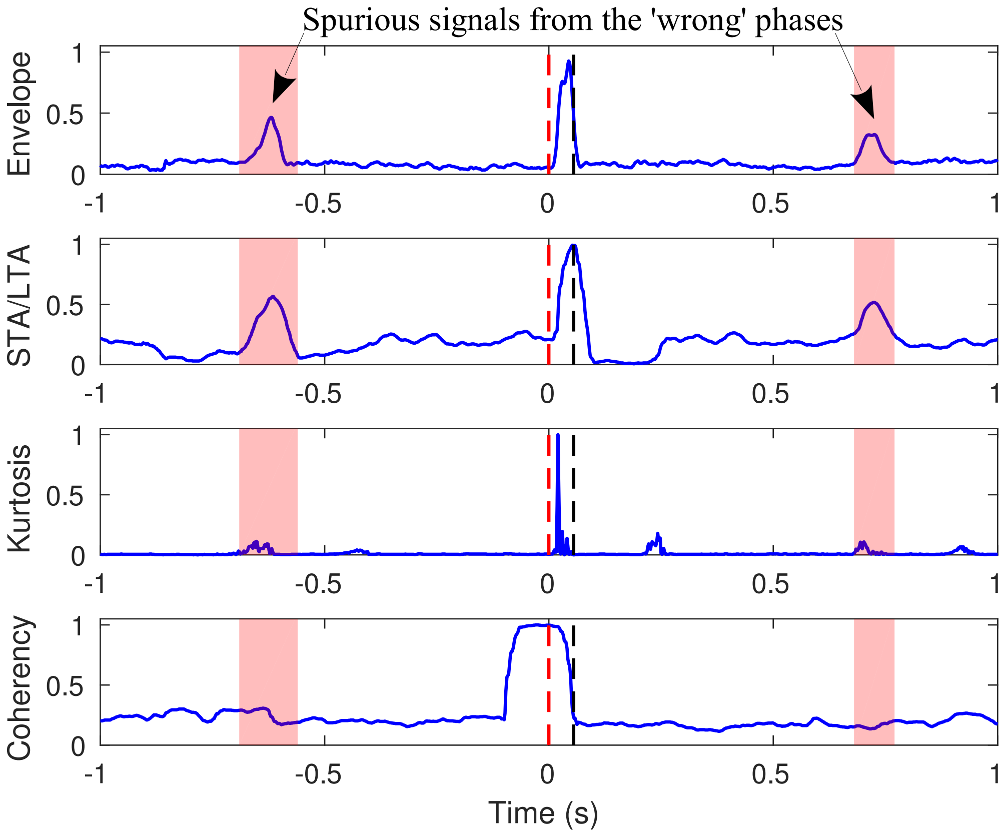


Figure 23. The stacking functions of the four methods at the true source location of the deep event for the Aquistore noise data. The red dashed lines show the origin time of the source time function and the black dashed lines show the end time of the source time function. The pink areas around -0.6 s and 0.7 s highlight the time range where P-/S-phases move into the stacking window of the S-/P-phases when searching for origin time. The SNR is 1.

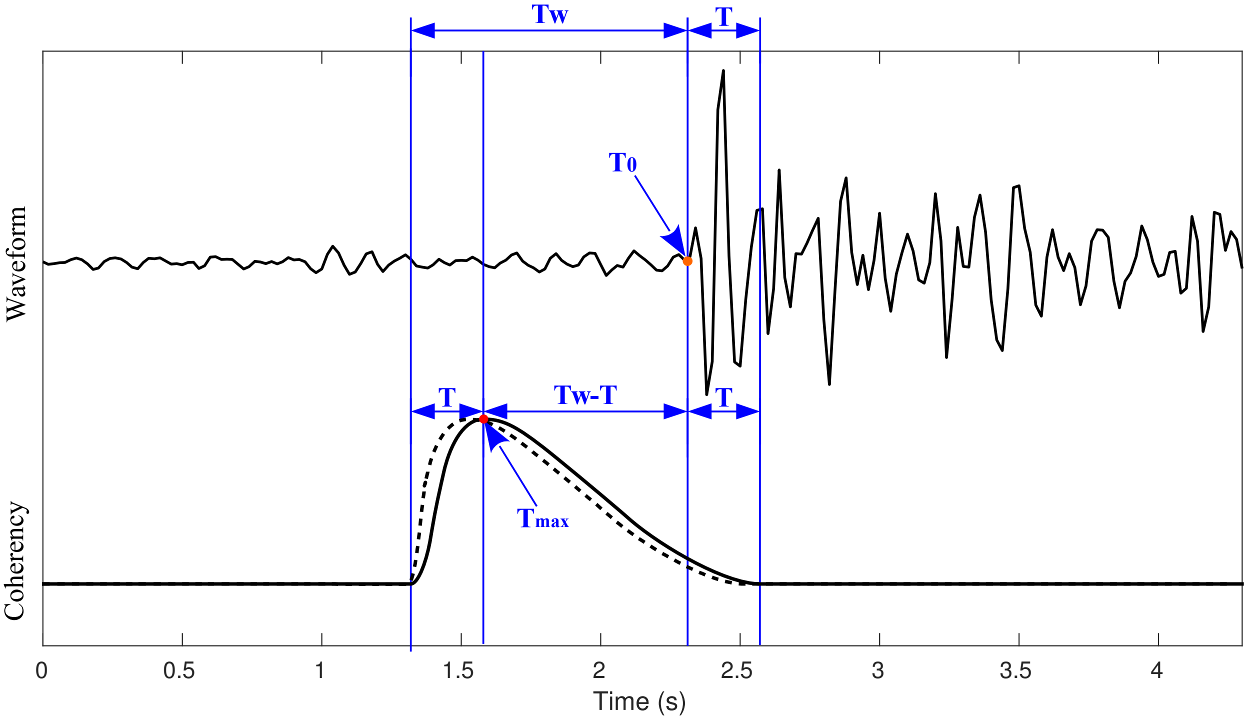


Figure 24. Schematic diagram showing recorded waveforms and the corresponding stacked coherency trace. Tw is the length of coherent analysis time window, and T is the period of direct wave. The orange dot shows the arrival time of direct wave, and the red dot shows the maximum coherency value at the stacked coherency trace. For the stacked coherency trace, the solid line shows the maximum coherency value appearing at T time after the rise of waveform coherency, and the dashed line shows the maximum coherency value appearing within T time after the rise of waveform coherency.

Supplementary Materials

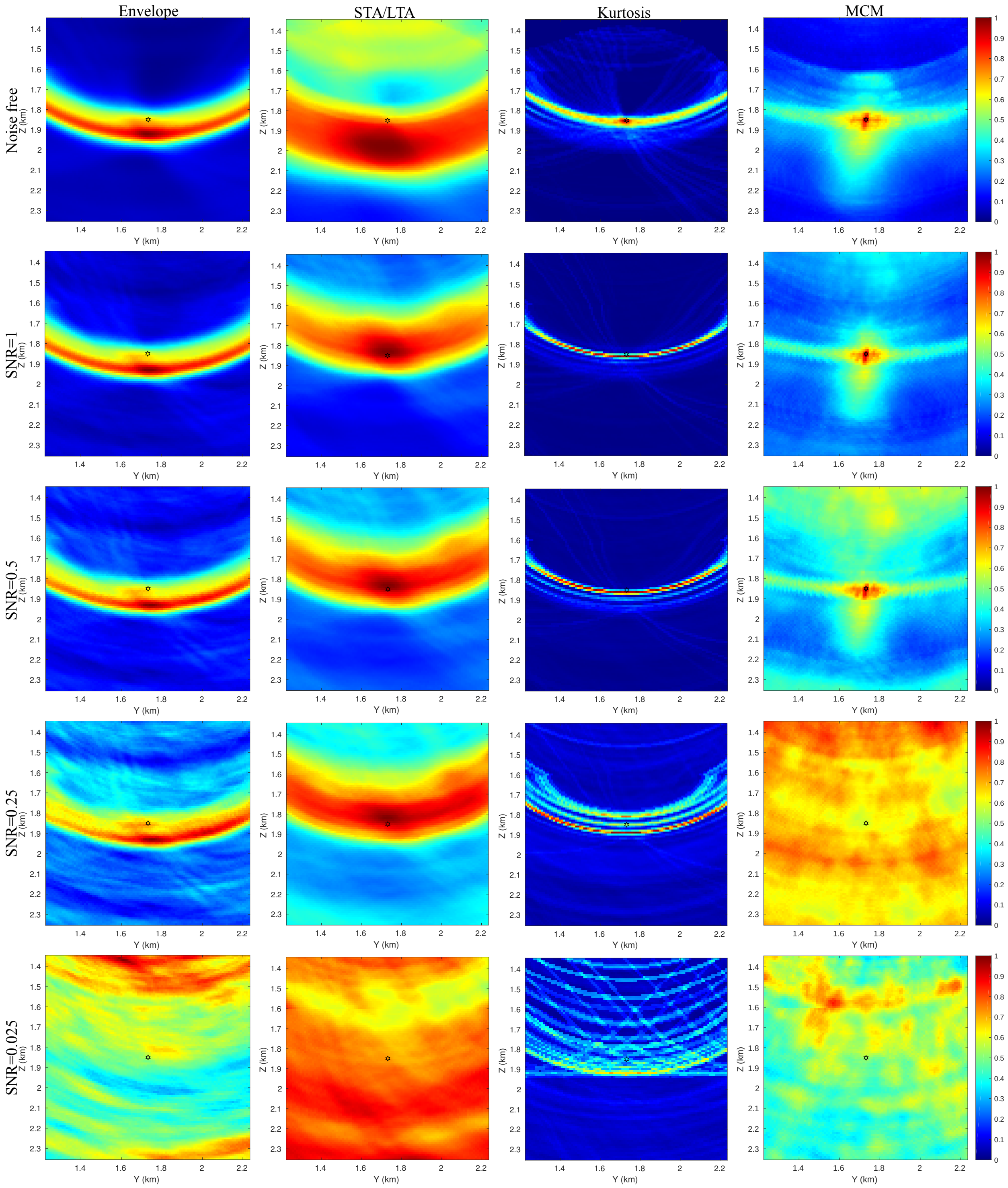


Figure S.1. Vertical profiles (YZ profiles) through the true source location of the migration results under different SNRs for the four methods. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows the results when data is free of noise, second row for SNR is 1, third row for SNR is 0.5, fourth row for SNR is 0.25, fifth row for SNR is 0.025.

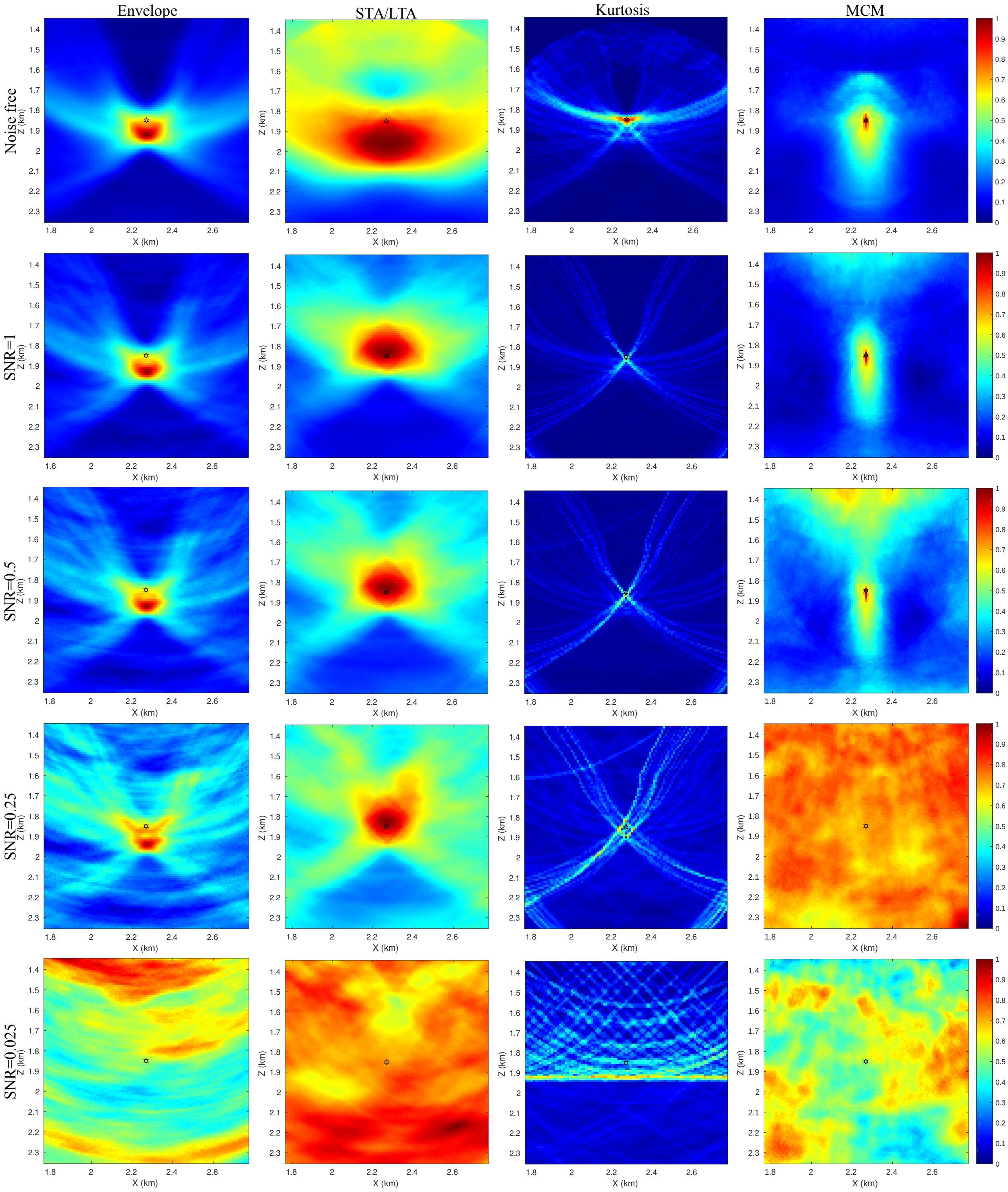


Figure S.2. Vertical profiles (XZ profiles) through the true source location of the migration results under different SNRs for the four methods. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows the results when data is free of noise, second row for SNR is 1, third row for SNR is 0.5, fourth row for SNR is 0.25, fifth row for SNR is 0.025.

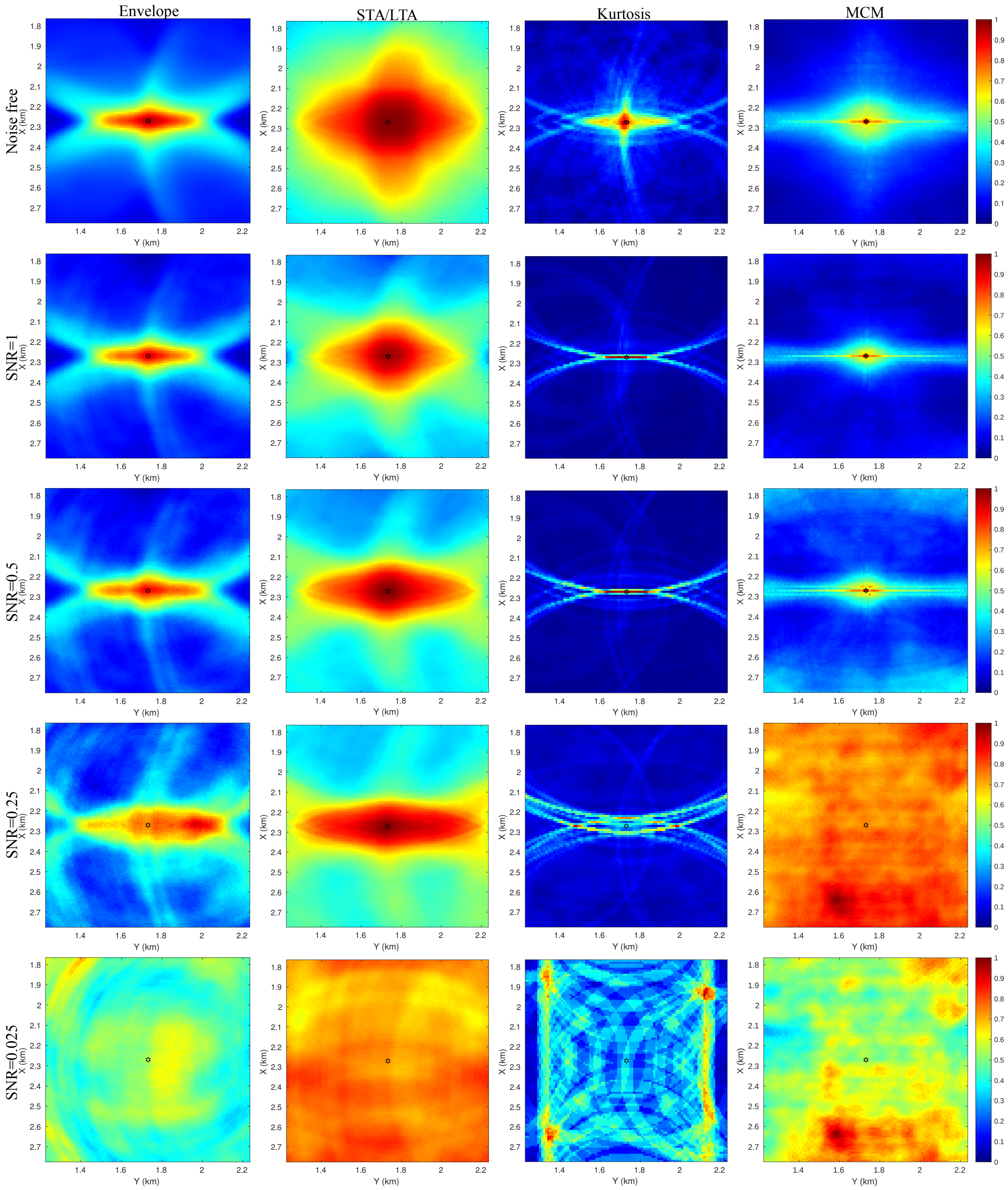


Figure S.3. Horizontal profiles (XY profiles) through the true source location of the migration results under different SNRs for the four methods. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows the results when data is free of noise, second row for SNR is 1, third row for SNR is 0.5, fourth row for SNR is 0.25, fifth row for SNR is 0.025.

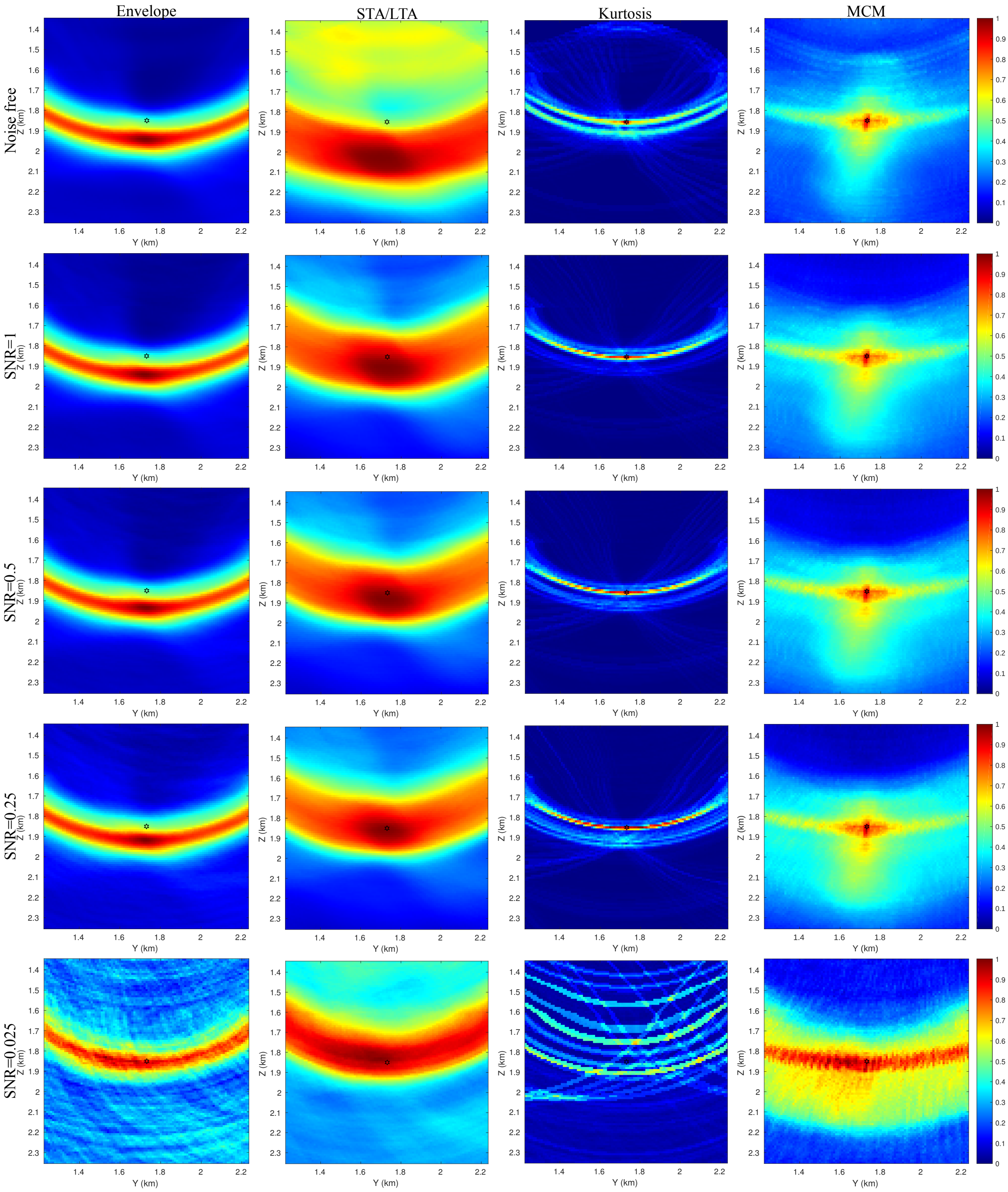


Figure S.4. Vertical profiles (YZ profiles) through the true source location of the migration results with automatic quality control scheme (weighting and filtering) under different SNRs. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows the results when data is free of noise, second row for SNR is 1, third row for SNR is 0.5, fourth row for SNR is 0.25, fifth row for SNR is 0.025.

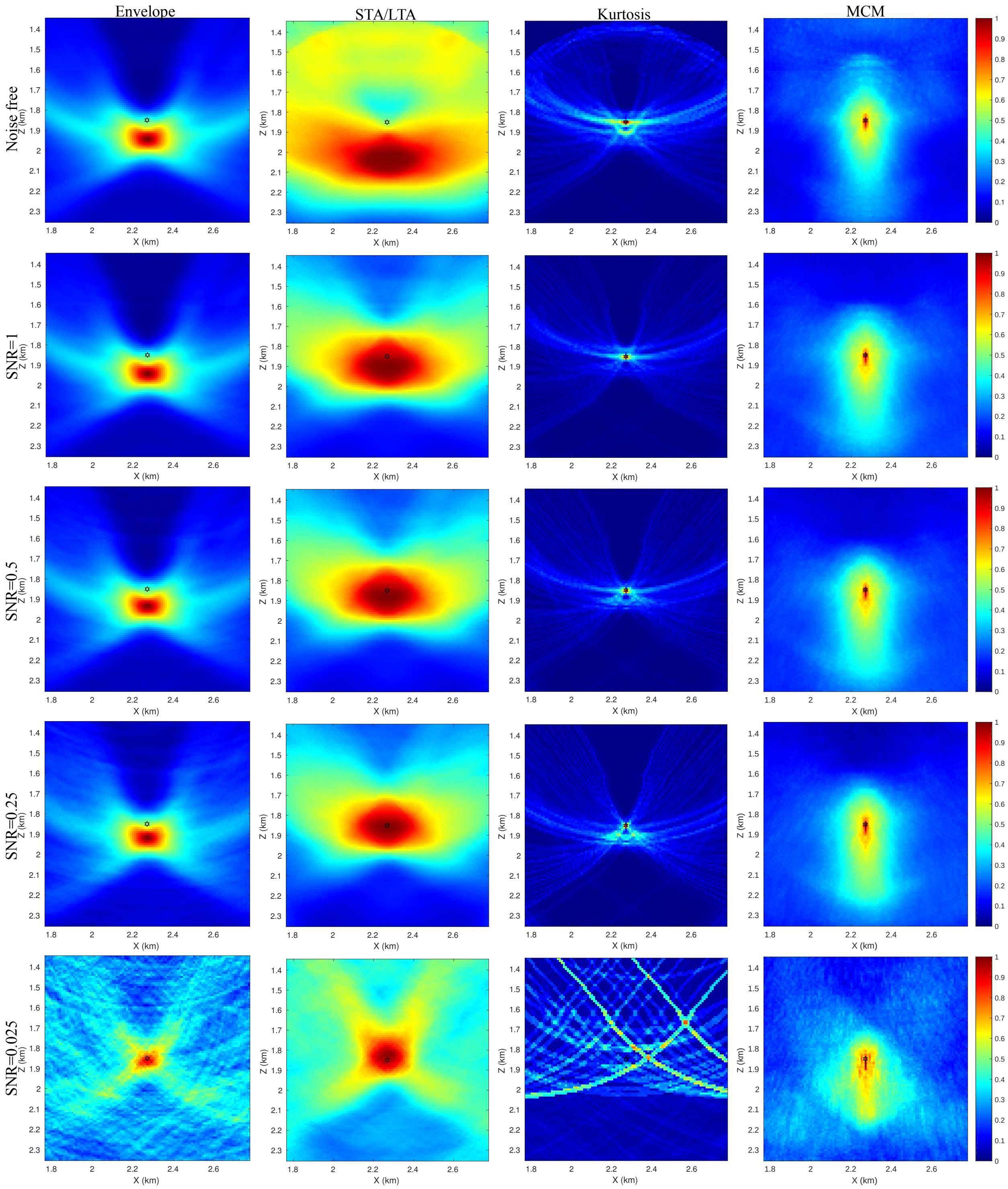


Figure S.5. Vertical profiles (XZ profiles) through the true source location of the migration results with automatic quality control scheme (weighting and filtering) under different SNRs. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows the results when data is free of noise, second row for SNR is 1, third row for SNR is 0.5, fourth row for SNR is 0.25, fifth row for SNR is 0.025.

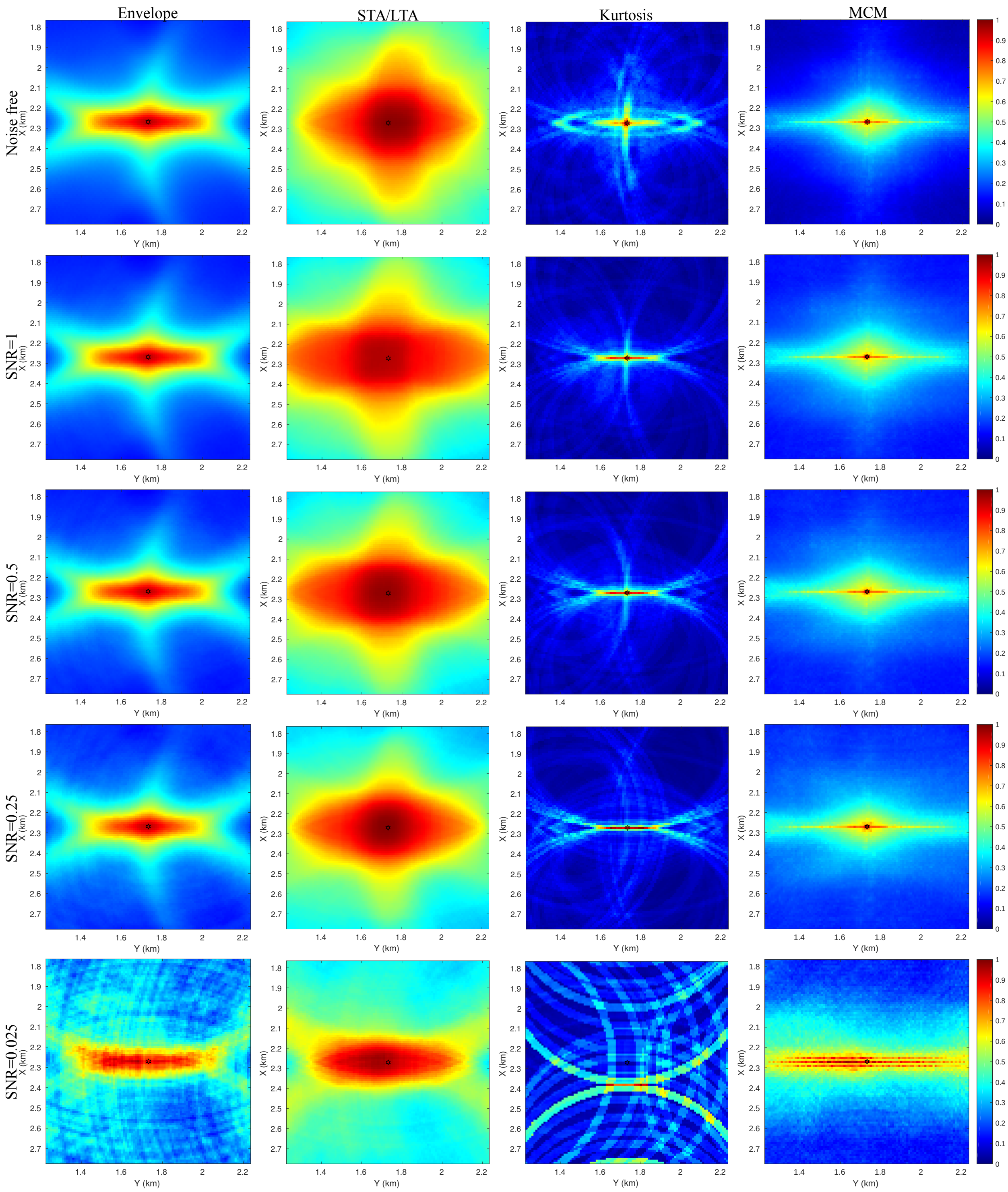


Figure S.6. Horizontal profiles (XY profiles) through the true source location of the migration results with automatic quality control scheme (weighting and filtering) under different SNRs. The dark star in the center shows the true source location. The first column shows results of envelope, second column for STA/LTA, third column for kurtosis, fourth column for MCM. The first row shows the results when data is free of noise, second row for SNR is 1, third row for SNR is 0.5, fourth row for SNR is 0.25, fifth row for SNR is 0.025.