

Spatial-temporal evolution of injection-induced earthquakes in the Weiyuan Area determined by machine-learning phase picker and waveform cross-correlation

Wing Ching Jeremy Wong¹, JinPing Zi¹, HongFeng Yang^{1*}, and JinRong Su²

¹Earth System Science Programme, Faculty of Science, The Chinese University of Hong Kong, Sha Tin, Hong Kong 999077, China;

²Earthquake Monitoring Center, Sichuan Earthquake Agency, Chengdu 610041, China

Key Points:

- We detected and located earthquakes from January 2018 to February 2019 using the machine learning method and double-differencing technique.
- The machine learning method and waveform cross-correlation enabled systematic phase picking and reduced inconsistencies in catalog picking.
- Hypocenter patterns demonstrate clear spatial-temporal clustering with depth constrained within 5 km.
- Our high-resolution earthquake relocations and waveform similarity analysis have revealed an unmapped fault that may have led to the September 2019 M_w 5.2 event.

Citation: Wong, W. C. J., Zi, J. P., Yang, H. F., and Su, J. R. (2021). Spatial-temporal evolution of injection-induced earthquakes in the Weiyuan Area determined by machine-learning phase picker and waveform cross-correlation. *Earth Planet. Phys.*, 5(6), 485–500. <http://doi.org/10.26464/epp2021055>

Abstract: Anthropogenic induced seismicity has been widely reported and investigated in many regions, including the shale gas fields in the Sichuan basin, where the frequency of earthquakes has increased substantially since the commencement of fracking in late 2014. However, the details of how earthquakes are induced remain poorly understood, partly due to lack of high-resolution spatial-temporal data documenting the evolution of such seismic events. Most previous studies have been based on a diffusive earthquake catalog constructed by routine methods. Here, however, we have constructed a high resolution catalog using a machine learning detector and waveform cross-correlation. Despite limited data, this new approach has detected one-third more earthquakes and improves the magnitude completeness of the catalog, illuminating the comprehensive spatial-temporal migration of the emerging seismicity in the target area. One of the clusters clearly delineates a potential unmapped fault trace that may have led to the M_w 5.2 in September 2019, by far the largest earthquake recorded in the region. The migration of the seismicity also demonstrates a pore-pressure diffusion front, suggesting additional constraints on the inducing mechanism of the region. The patterns of the highly clustered seismicity reconcile the causal link between the emerging seismicity and the activity of hydraulic fracturing in the region, facilitating continued investigation of the mechanisms of seismic induction and their associated risks.

Keywords: hydraulic fracturing; Sichuan Basin; machine learning; injection-induced earthquake; Weiyuan earthquake

1. Introduction

Anthropogenic activities inducing seismicity have been widely reported in recent decades. In particular, subsurface fluid injection (or withdrawal) activities, including enhanced geothermal recovery (Majer et al., 2007; Lee et al., 2019), carbon sequestration (Zoback and Gorelick, 2012), underground gas storage (Cesca et al., 2014; Foulger et al., 2018; Zhou PC et al., 2019; Jiang GY et al., 2020, 2021), wastewater disposal (Ellsworth, 2013) and hydraulic fracturing of shale gas exploration (Clarke et al., 2014; Weingarten

et al., 2015; Bao XW and Eaton, 2016; Yang HF et al., 2020; Zhou PC et al., 2021), have all been associated with seismic events. The subsequently induced earthquake rates caused by these industrial activities can rise as high as ten times the background seismicity; induced earthquakes can be of sufficient magnitude to be felt locally and even cause significant destruction (Grigoli et al., 2017; Yang HF et al., 2017; Atkinson et al., 2020).

Hydraulic fracturing, also known as fracking, is a directional well-bore-completion method used to recover hydrocarbons from unconventional oil and gas reservoirs. High-pressure fluids are injected into the targeted shale layers to create new fracture networks, enhancing fluid flow within impermeable yet porous strata. After the opening of the fracture network, the injected fluid is allowed to flow back so as to extract hydrocarbons (Gale et al., 2014). Giv-

Correspondence to: H. F. Yang, hyang@cuhk.edu.hk

Received 10 MAR 2021; Accepted 11 AUG 2021.

Accepted article online 02 NOV 2021.

©2021 by Earth and Planetary Physics.

en hydraulic fracturing’s limited injection volumes and discrete fracturing patches (unlike wastewater injection and enhanced geothermal recovery with continuous high-pressure injection), it was commonly asserted as recently as 2015 that fracking would present negligible risk of triggering damaging earthquakes. The microseismicity expected during fracking procedures was predicted to be limited to magnitudes ranging from -4 to 0 . However, evidence has been accumulating that hydraulic fracturing induces reactivation of large pre-existing faults, thus leading to destructive earthquakes (Lei XL et al., 2019; Schultz et al., 2020; Atkinson et al., 2020; Yang HF et al., 2020).

By far, a significant portion of the most intense earthquakes attributable to hydraulic fracturing have been located in the Sichuan Basin. In recent years, two major shale gas fields, in the Changning and Weiyuan areas, have witnessed induced earthquakes of magnitude five or above (Lei XL et al., 2019; Sheng MH et al., 2020). These earthquakes have led to casualties and to estimated economic losses of up to 50 million RMB (Lei XL et al., 2019).

Rising concerns over the safety and seismic risks of the area have led to extensive research into the properties and the mechanisms of the aforementioned significant earthquakes. Despite several detailed studies of these earthquakes, the locations of their hypocenters remain debatable, particularly for the largest event that occurred in early September, 2019, in the Weiyuan Gas Field (Figure 1, Table 1). According to previous analyses, the focal depth of the 2019 M_w 5.2 earthquake exhibits variation from two to five

kilometers (Wang MM et al., 2020; Lei XL et al., 2020; Sheng MH et al., 2020). Although location disputes and other differences are common in earthquake studies, as geophysical inversion with different observations and model parameters may have divergent results, knowing with precision the hypocenter locations of induced earthquakes is pivotal in determining how induction takes place (Yang HF et al., 2020; Lei XL et al., 2020; Sheng MH et al., 2020).

It is worrisome that the faults on which these relatively large earthquakes occurred remain unclear. For instance, the pre-existing fault leading to the 7 September M_w 5.2 Weiyuan earthquake remains controversial. Wang MM et al. (2020) analyzed the earthquake’s focal mechanism and incorporated the seismic reflection profile to outline the potential fault geometry of the event. Their results suggested that the earthquake with 5 km depth activated the pre-existing basement fault of a structural wedge at the Weiyuan anticline. However, according to the interpretation of the reflection profile along the northeastern Weiyuan region, there are also two detachment faults with similar geometry at shallower depths, comparable to the injection depth of 2.5 km. Another analysis from Sheng MH et al. (2020) used previously well-located events from a local dense array to conduct relative relocation, and determined the depth of the largest event to be 2.9 km. The event location coincides with the aforementioned upper detachment fault. The September 2019 M_w 5.2 earthquake is not the only one whose source location remains disputed; differences among reported source locations of the February 2019 Rongxian earth-

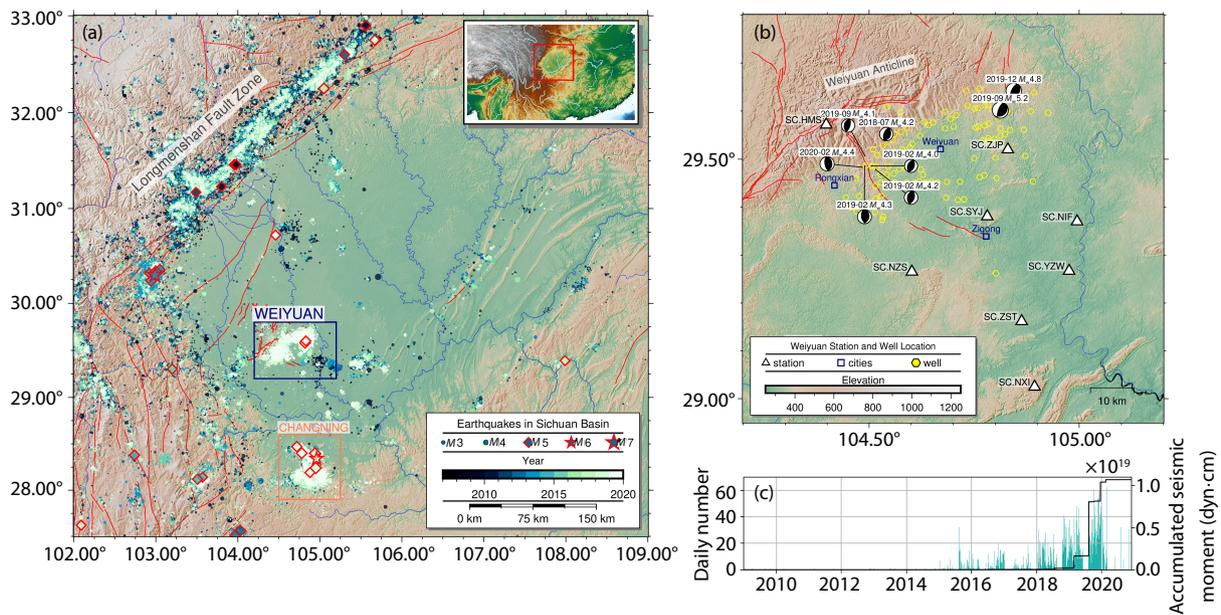


Figure 1. Background seismicity of the Sichuan Basin; station and well coverage in Weiyuan area; focal mechanism of the magnitude 4+ events and seismicity rate in the Weiyuan area. (a) Background seismicity of the Sichuan Basin. Events are colored by date and sized with magnitude. Events with magnitude larger than 5 are outlined by red diamond and star. Red lines delineate the major faults. Blue and orange rectangles highlight two recent major seismic clusters related to hydraulic fracturing in the Weiyuan and Changning regions respectively. (b) Stations and recent major earthquakes in the Weiyuan area. White triangles, blue squares, and yellow hexagons indicate locations of the seismic stations, major cities, and well locations respectively. Locations of 153 wells are derived from Google Earth satellite and field visits from 2015–2020. (c) Seismicity rate of the Weiyuan area. The bar and line represent the daily event counts and the cumulative moment released in the Weiyuan area. The recent M_w 5.2 and M_w 4.9 events in 2019 September and December released significant seismic moment in the area.

Table 1. Major 2019 earthquakes in the Weiyuan Region; differing source parameters from CENC catalog and various studies.

Time (YYYY/MM/DD HH:MM) (UTC)	Lon.(°)	Lat.(°)	Depth (km)	Magnitude	Source
2019/02/25 05:15	104.49	29.47	1.0	M_w 4.3	Yang HF et al., 2020
	104.49	29.48	1.7	M_w 4.53	Lei XL et al., 2020
	104.5	29.47	5.0	M_s 4.9	CENC Catalog
2019/09/07 22:42	104.82	29.58	2.5	M_w 4.92	Lei XL et al., 2020
	104.931	29.530	5.0	M_w 5.2	Wang MM et al., 2020
	104.812	29.602	2.9	M_w 4.97	Sheng MH et al., 2020
	104.82	29.58	10.0	M_s 5.4	CENC Catalog

quakes also have not been unresolved. Yang HF et al. (2020) combined the inversion by InSAR surface deformation and seismic data to conclude that the M_w 4.3 Rongxian earthquake was located at a depth of 1 km and adjacent to the Molin fault, and that its two foreshocks were at greater depth (2.7 km), occurring on a different fault. However, the depth of the M_w 4.3 Rongxian earthquake has been suggested as >2 km (Lei XL et al. 2020; Yi GX et al., 2020).

Furthermore, there has been limited analysis of the foreshocks and aftershocks of the events that help constrain the potential fault geometry. Lei XL et al. (2020) relocated earthquake sequences of the major events; however, the resulting foreshocks and aftershocks did not display any well-aligned clusters. Therefore, the activated fault remains not well delineated, and the inducing mechanism is yet to be resolved.

The delineation of fault geometry, especially at greater depths, demands high-resolution earthquake locations. However, the location proposed by previous methods was derived from catalog phase arrivals (Yang HF et al., 2020; Lei XL et al., 2020; Wang MM et al., 2020), which are routinely determined from P and/or S arrivals. As a consequence, the location's accuracy depends critically on the heterogeneities of the velocity model and the uncertainties of the picked arrival times and station distribution. As shown in one of the results (Yang HF et al., 2020), the hypocenters after double-difference relocation remain diffusive, with depths reaching 20 kilometers, and display rough estimates of the distribution of earthquake clusters (Figure 2a, b), making it difficult to differentiate whether they are induced or tectonic earthquakes.

In order to search for potential fault geometry that might be illuminated by detailed knowledge of the micro-seismicity, we picked seismic arrivals by use of a detector that employs machine learning. We then associated and located earthquakes in the Weiyuan area from January 2018 to March 2019 using the permanent station network. Furthermore, we employed waveform cross-correlation to enhance the accuracy of differential times and relocated the earthquakes using the double-difference method (Waldhauser and Ellsworth, 2000). Furthermore, we used available event catalog data to calibrate the local magnitude calculations of all the earthquakes; the resulting new expanded catalog of events was then analyzed for their spatiotemporal distribution as well as their potential for fault surface delineation. Lastly, we re-confirmed that the earthquakes with shallow depths in the Wei-

uan area were induced by hydraulic activities.

2. Geological Setting and Seismic Data

2.1 Overview of Geology, Hydraulic Fracturing Activities in Weiyuan Area

The Sichuan Basin was discovered as the largest shale gas reservoir in China in the 1990s and is still the major petroleum producer to date (Lei XL et al., 2013). The Weiyuan Gas Field is one of the major gas reservoirs in the southwestern Sichuan Basin and has high productivity at present. According to government reports, the Weiyuan Shale Gas Field had a targeted shale gas production volume of 2 billion m^3 in 2015. The shale gas field belongs to a large anticline dome structure that lies on the southeastern slope of the Leshan–Longnūsi paleo-uplift and the western side of the Early Cambrian "Mianyang–Changning" intra-cratonic sag. The rich and high-quality shale gas formation in the area is found at depths of 1.5 to 4.5 km and between the upper Ordovician Wufeng formation. The Longmaxi–Wufeng formation shale layer has an average thickness of 100 meters (Liang X et al., 2019).

The "Wei-201 well" drilled in 2010 in the Laochang Village, Xinchang Town, Weiyuan County, was the first shale gas evaluation well in China. During completion of the well and the first hydraulic fracturing operation, the earthquake rate in the Weiyuan area increased drastically (Figure 1c). Systematic shale gas hydraulic fracturing in horizontal wells commenced in 2014 (Lei et al., 2017; Yang HF et al., 2020). At least four horizontal wells were drilled and completed after a year of development. Recent satellite imagery has shown active well development in the Weiyuan area (Figure 1b).

2.2 Overview of Emerging Seismicity in the Sichuan Basin and the Weiyuan Region

Although frequent and large earthquakes occur along the Longmenshen Fault Zone at the northwestern part of the Sichuan Basin, natural seismicity within the Sichuan Basin has been scarce and limited, according to the earthquake catalog (Figure 1a). Before the recently emerging seismic activities, there were minor swarms of earthquakes related to the salt mining industries located at the southeastern side of the Weiyuan area. As reported by the modern catalog from the China Earthquake Network Center (CENC), no earthquakes of magnitude larger than five were ever recorded within the Weiyuan Region before 2018. Since the introduction of unconventional shale gas exploration in 2015,

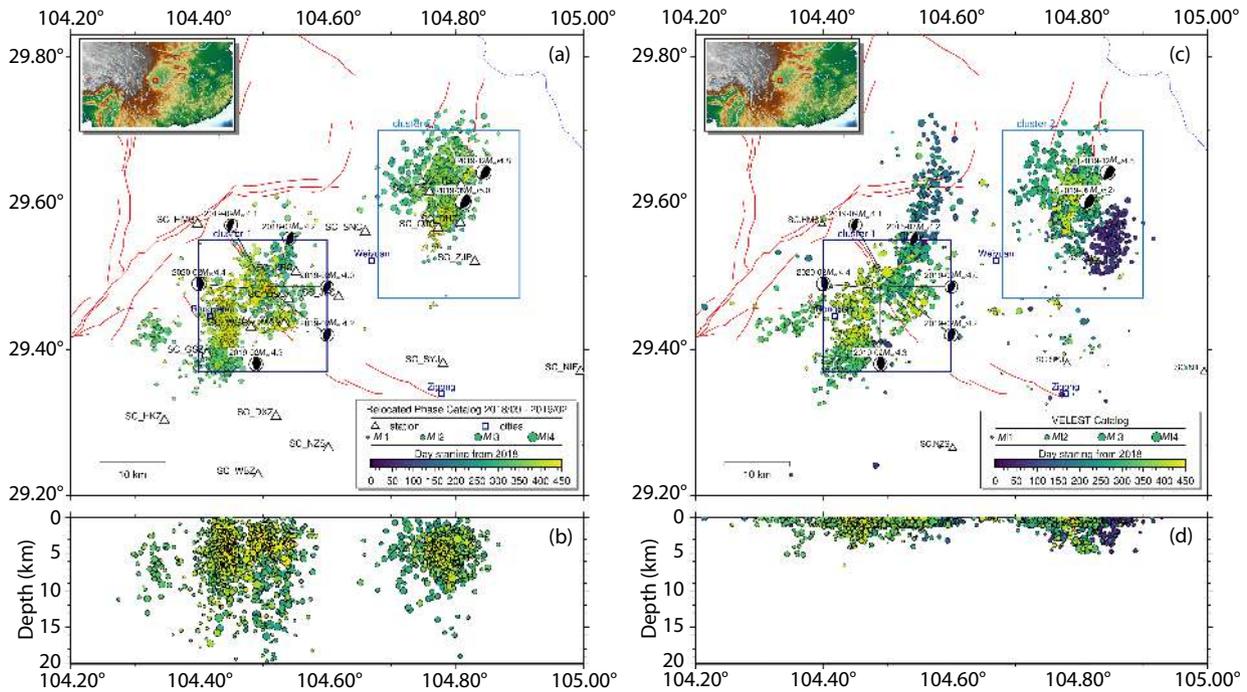


Figure 2. Earthquake catalog (a, b) phase data from 2018-09 to 2019-02 in previous catalog (Retrieved from Yang HF et al., 2020); and (c, d) from new machine learning catalog. (a) Red lines delineate mapped faults; triangles locate the stations used in the catalog; colored circles show the relocated events.

however, seismicity has surged in the Weiyuan Area (Figure 1c) (Zhou PC et al., 2021). As recently as in 2019, four M_w 4+ earthquakes and one of M_w 5.2 occurred along the southern edge of the Weiyuan anticline structure. The corresponding seismicity rate can now reach 50 events per day.

2.3 Seismic and Drilling Data

Event waveforms from the permanent local seismographic network were collected with nine short-period seismometers from 2018 to September 2019 and used in this study. The seismic network has an average spacing of 30 kilometres and is located at the southern flatland of the Weiyuan anticline. The permanent station network provides primary spatial and azimuth coverage over the targeted Weiyuan region and the Molin fault area (Figure 1b). Event waveforms from the nine local stations have a section length of five minutes starting from the event origin time. In total, we have 31,847 sets of event waveform data, covering 20 to 30 percent of all daily data. We have also detected and identified drilling locations in the Weiyuan area from 2015 to 2020, based on satellite photos and our field visits (Figure 1b).

3. Methods

3.1 Machine Learning Detector and Detection

Machine learning applications have been widely adopted in seismology, with particularly outstanding performance in seismic phase detection and picking accuracy (Kong et al., 2019). A deep neural network with convolutional layers can recognize the complex shape of an earthquake arrival phase and select the corresponding onset (Perol et al., 2018). We used the publicly available machine learning detection package “PhaseNet”, which is a U-

shaped convolutional neural network modified with 1-D time series data input, to analyze the seismic data. PhaseNet’s deep neural network consists of four convolutional down-sampling and up-sampling stages. In each stage, the data are convoluted with a 1D filter, and each neuron is activated by the rectified linear unit (ReLU) to recognize the spatial pattern. The model outputs continuous probability distribution of P arrivals, S arrivals, and noise levels with the same dimensions as the input. The PhaseNet model is trained with over 700,000 labeled seismic arrivals from the Northern California Earthquake Data Center. The training datasets consist of input from various seismic instruments, including broad-band seismometers, short-period seismometers, and accelerometers. Limited data pre-processing is applied with only a linear trend removed during the training of PhaseNet, and thus the training sets consist of a portion of low signal-to-noise-level templates. The PhaseNet deep learning model has proven its robustness, accurately detecting arrivals, even in noisy conditions, when tasked with analyzing the characteristics of P and S arrivals and the ambient noise pattern in seismic recordings (Zhu WQ and Beroza, 2019). Our study applied PhaseNet as the phase detector to select P and S arrivals from the continuous waveform.

3.2 Earthquake Location and Magnitude Calibration

The detected P and S phases are associated with the grid-search method using REAL — the Rapid Earthquake Association Locator (Zhang M et al., 2019). The travel-time grids use the modified velocity profile derived by ambient-noise topography, and shallow depth data from well-logging records (Meng XB et al., 2018) (Figure S1). Absolute location using VELEST (Kissling et al., 1995) is applied after the initial grid search hypocenter location. VELEST is a distributed Fortran program to locate earthquakes using travel-

time iterative inversion based on predefined velocity profiles. The resulting locations are the initial reference model which is then subjected to further analysis. In our analysis, VELEST-computed events with large azimuth gaps ($> 270^\circ$) were classified as poorly-determined and were discarded. The accuracy of the catalog is further improved by using the hypoDD relocation method incorporated with phase arrival cross-correlation shift correction (Waldhauser and Ellsworth, 2000).

All the associated and located events were estimated using calibrated local magnitudes (Equation (1)). We defined the local magnitude equation based on the linear relation of the maximum logarithmic amplitude of horizontal ground motion A_{peak} to the logarithmic epicentral distance, with the correction k on attenuation and geometric spreading and C constant

$$M_{L,\text{reference}} = \log_{10} A_{\text{peak}} + k \log_{10} R + C, \quad (1)$$

$$\log_{10} A_{\text{peak}} = \frac{\log_{10} A_{\text{peak,East}} + \log_{10} A_{\text{peak,North}}}{2}. \quad (2)$$

We determine the k and C values by using the catalog events; we then calculate our local magnitude estimation for each event.

3.3 Fault Plane Fitting

The fault plane is determined by the location of the hypocenter, using a least-squares solution based on the plane equation. Consider a plane $d_i = m_1 + m_2 x_i + m_3 y_i$ with N_{th} hypocenter equations. The plane equation has the form of

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & y_N \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}. \quad (3)$$

By the least-squares method, we can estimate the coefficients of the plane:

$$\mathbf{m}^{\text{est}} = [\mathbf{G}^T \mathbf{G}]^{-1} \mathbf{G}^T \mathbf{d}. \quad (4)$$

The standard derivation of the least-squares planar fit is calculated by:

$$\text{Cov}(\mathbf{m}^{\text{est}}) = \sigma_d^2 [\mathbf{G}^T \mathbf{G}]^{-1}. \quad (5)$$

To determine the fault trace, the related events showing lineation are subset from the catalog, using temporal range and latitude and longitude bins, determined visually. In our study, we have selected the events located in the northeastern part of the Weiyuan area from January to April, 2018.

3.4 Pressure Diffusion Front

The pore pressure induced seismicity has a characteristic migration pattern following the diffusion front, as stated in Equation (6):

$$R = \sqrt{4\pi Dt}. \quad (6)$$

which presents the pore pressure front that results from solving the diffusion equation assuming a homogeneous, isotropic medium. Full derivation is attached in the supplementary material. R is the radius of the pore pressure diffusion-triggering front, based on the injection lapse time t and the scalar hydraulic diffusivity D . The hydraulic diffusivity is the permeability divided by the dynam-

ic viscosity of the fluid (Shapiro, 2015). Seismic events are more likely to occur within the relaxation zone; the diffusion front given by Equation (6) suggests the upper bound of seismicity migration. Since the injection data are limited, we select the first earthquake of each seismic cluster to approximate the injection site and initiation.

4. Results

4.1 Machine Learning Detection and Event Association

Although the PhaseNet machine learning detector was trained with data subjected to minimal pre-processing, we employed standard time-series pre-processing to increase the signal-to-noise ratio before the detection. Linear trend removal, tapering, and highpass 1 Hz were applied. To avoid phases at the boundaries, the pre-processed waveforms were then sliced into 30-second windows with 50 percent overlap. Data in each 30-second window were normalized and fed into PhaseNet for arrival prediction. To insure picking accuracy, this study used only phase picks with probabilities larger than 0.5. In total, this approach identified 232,050 P arrivals and 161,720 S arrivals.

We conducted a travel time grid search using the REAL package (Zhang M et al., 2019). The travel time grid consists of search ranges with station-event distances of up to 1.5 degrees (167 km) with a grid size of 0.02° , and event depths ranging from surface to 30 km with a grid size of 1.5 km. Due to the limited number of stations, a minimum of four P phase arrivals, and six total phase arrivals (P + S), and a residue tolerance of 0.5 seconds in phase arrivals, are configured for an event association. Figure S2 demonstrates one of the associated events with corresponding phase picks. The requirements mentioned above are generally looser than those of previous studies based on more comprehensive station coverage, such as the study of induced earthquakes using a nodal seismic array in Raton Basin, USA (e.g., Wang RJ et al., 2020).

In total, the machine learning phase picking yielded 18,172 P picks and 11,996 S picks, based on which we have associated 3,522 events. The associated picks consist of less than 10 percent of the detected phases. The limited number of associated phases is due in part to the limited station coverage over the major earthquake clusters; data from the closest four of the nine stations account for 60 percent of the total detections.

Locations of the associated events were further identified by use of absolute VELEST location results (Kissling et al., 1995); to minimize event location errors, we accepted only events for which VELEST station azimuth gaps were 270° or less. Overall, 3,061 earthquakes were detected and located in the event waveforms detected in data from the permanent stations between January 2018 and to February 2019 (Figure 3). Compared to the catalog from the China Earthquake Network Center (CENC), our approach has detected 1,458 additional earthquakes, accounting for 47 percent of the revised total; compared to the Sichuan Earthquake Agency's catalog from November 2018 to February 2019, our machine learning methods have added 335 unique earthquakes, which account for 20 percent of the revised total of 1,652 earthquakes.

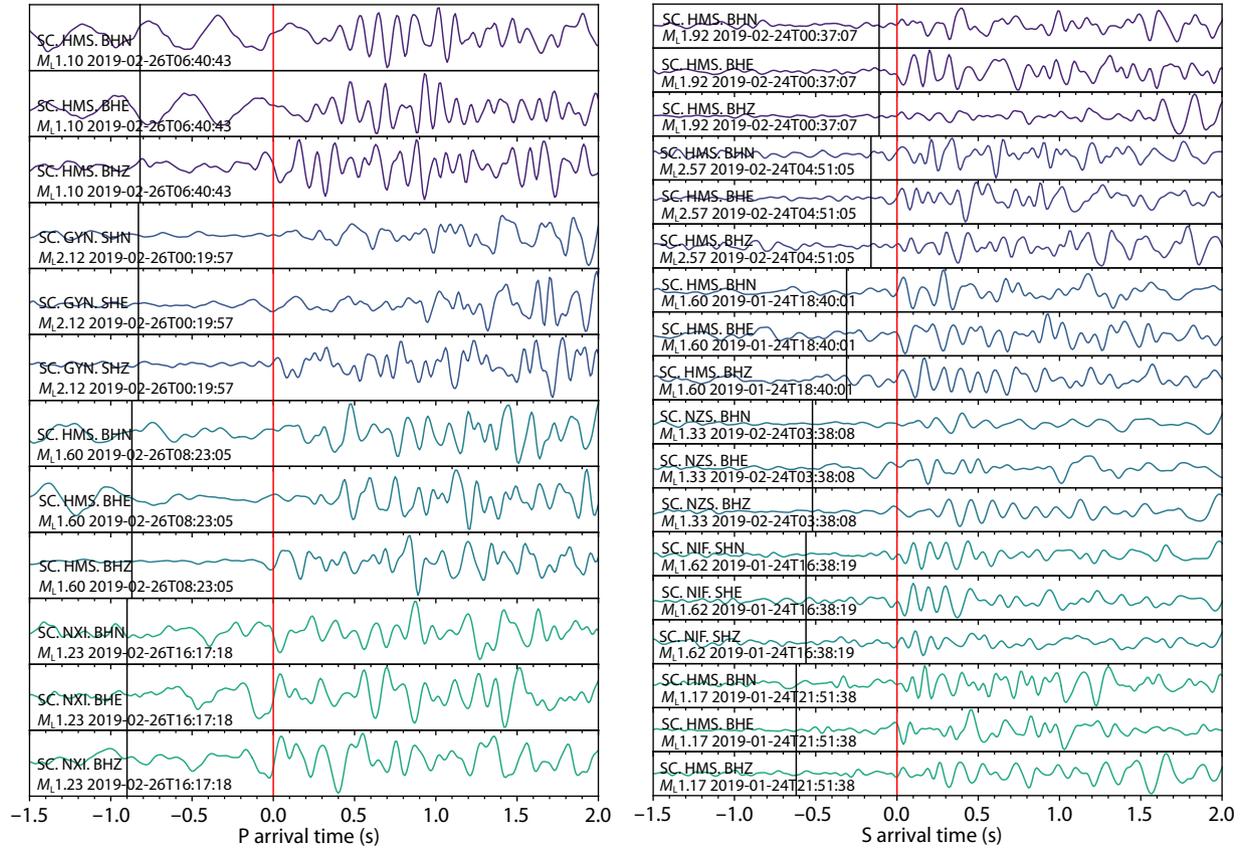


Figure 3. Comparison of arrival phase picks by PhaseNet and manual picks by the Sichuan Earthquake Agency. (a) P wave arrival pick comparison; (b) S wave arrival pick comparison. Red line is the phase pick by PhaseNet; black line is the manual pick.

4.2 Comparison with Catalog Phase Picks

To assess the performance of the machine learning detection used in the study, we compared our machine learning detection and picking results with the catalog phase picks. There are some obvious differences between the arrival times picked by PhaseNet and in the catalog, the latter consisting of either manual picks or picks based on Short Term Average/Long Term Average. The catalog picks demonstrate both effusive and impulsive onsets as P wave arrivals, as shown in the waveform plot in the Figure 3. In contrast, the PhaseNet picks exhibit a consistent picking among the arrivals.

To assess the capability of machine learning detection qualitatively, we computed the arrival time differences between picks based on the machine learning approach and the catalog picks (Figure 4). Compared to catalog picks, PhaseNet picks exhibit systematic delays, with median values of 0.60 seconds in P arrivals and 0.88 seconds in S arrivals (Figure 5). These systematic differences are due to variant consistency in picking arrival onsets of manual picks and PhaseNet.

The accuracy of PhaseNet detection is further illustrated using arrival cross-correlation correction. To obtain a relative pick adjustment, we conducted cross-correlation on each phase arrival pair whose corresponding event separation was less than 20 km. In principle, the cross-correlation time shift of an accurate arrival-pair should be minimal, while the time shift for a less accurate arrival-pair would be finite. However, mis-picked arrival pairs result

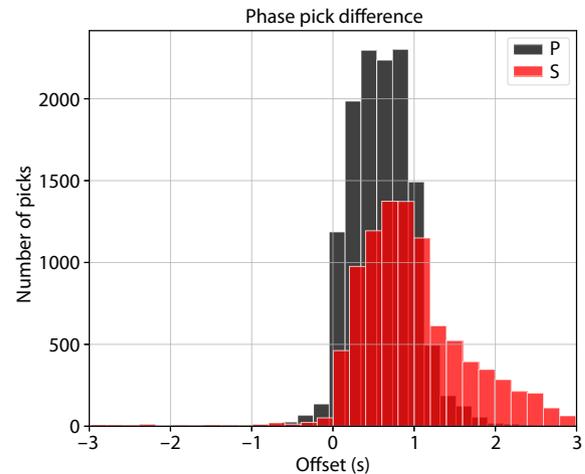


Figure 4. Histogram of phase pick time difference between machine learning detector and catalog phase pick. Offset is the machine learning phase pick minus catalog phase pick time. Black indicates P wave pick offset; red indicates S wave pick offset.

in either a random time shift or a pseudo-minimal time shift. Therefore, the time-shift distribution for cross-correlated arrival pairs picked by an accurate phase picker would be narrow and concentrated, while results from an imprecise phase picker would exhibit a more widely distributed time shift.

In our study, the waveform was first subjected to a bandpass 2–

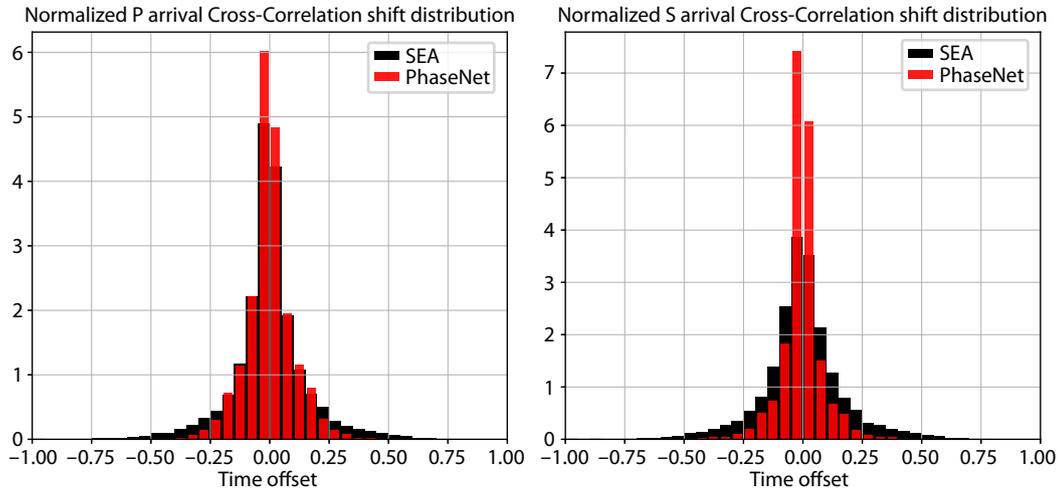


Figure 5. Histogram of phase pick correction between machine learning detector and catalog phase pick by waveform cross-correlation. Black indicates P wave pick correction; red indicates S wave pick correction.

15 Hz filter and sliced 0.5 s before and 1.5 s after arrival picks for the cross-correlation; the only cross-correlation results considered were those exhibiting coherency of 0.7 or above.

The resulting arrival cross-correlation shift distribution revealed that the machine learning detector used in our study delivered a consistent improvement. In both P and S arrival picks, the time-shifting distributions of the PhaseNet picks are more confined than those of the manual picks. PhaseNet picks were revealed to exhibit a significantly lower variance in correction shift than catalog picks: PhaseNet's standard deviations were 0.099 s for P and 0.083 s for S, compared to the corresponding 0.158 s and 0.172 s for the catalog picks (Figure 6). These results suggest that the PhaseNet picker is significantly more reliable than the catalog picking method.

To optimize the machine learning picking, we tested a higher acceptance of the PhaseNet phase pick output, lowering the probability level cutoff from 0.5 to 0.3. Visual inspection of these lower-probability phase picks confirms the expected significant increase in ambiguity of arrival onset selection. Although the number of phase pick (401,689 P, 333,914 S) and the consequent number of events associated (7000+) by grid-search methods doubled with the loosened criteria, the final number and accuracy of events located after double-difference relocation remains mostly unchanged. Both 0.5 and 0.3 PhaseNet picking acceptance led to relocation of nearly a thousand events. We concluded that acceptance of lower probability picking by the machine learning methods does not significantly improve the final results; accordingly we retain our original criterion: limiting machine learning phase picks to those at probability levels of 0.5 or greater.

4.3 Double Differencing Relocation

We relocated the events using the double-difference relocation algorithm hypoDD (Waldhauser and Ellsworth, 2000). In addition to differential times derived from cross-correlation, we also used phase arrivals to calculate the differential times if their cross-correlation coefficients were lower than the threshold, with PhaseNet's picking probabilities used as the phase weighting. The config-

urations of the hypoDD inversion are listed in Table S1.

We relocated 1,350 events in the Weiyuan area, comprising one-third of the events located. Compared to the absolute locations from VELEST, the relocated catalog shows enhanced clustering and lineation of the hypocenters (Figure 7). The catalog also displays multiple swarms that clearly emerged and migrated during the period, correlating with the continuous development and operations of hydraulic fracturing activities in the region. Compared to the Sichuan Earthquake catalog, the event depths of our relocated catalog have shifted shallower, with a median displacement of 4.53 km in depth and 2.27 km horizontally (Figure 8).

Addressing the significant loss of events that occurred during our relocation processes, we applied different parameters to maximize the number of events relocated. We then tested the accuracy of the results. To recover most of the discarded results during the inversion, we followed the configuration used to analyse the initial hydraulic fracturing stage in the Guy-Greenbrier area, Arkansas, United States, for which data from three seismic stations were available (Yoon et al., 2017). The air-quakes were retrained during relocation iterations according to their configurations. Our number of relocated earthquakes doubled with use of the air-quake reserved configurations; however, most of the additional earthquakes tended to be at the subsurface (within 1 km depth) and their locations exhibited large uncertainties. Inclusion of these ill-defined shallow events altered the deeper hypocenter lineation. Therefore, with limited station spatial and azimuth coverage, we considered removing ill-relocated events during the iterations.

To further examine the reliability of the relocation results and to obtain a more exhaustive catalog of event locations, we subset and relocated a major event cluster located at the northeastern side of the Weiyuan county, where the largest M_w 5.2 event occurred (Figure 7). We applied and tested multiple hypoDD inversion configurations to the targeted cluster (Table S2). Since all of the stations are located at the southern and southeastern side of the cluster but with large separations between them (> 60 km), to test the stability of our results we tried a bootstrap sampling method, limiting stations and corresponding phase data in the in-

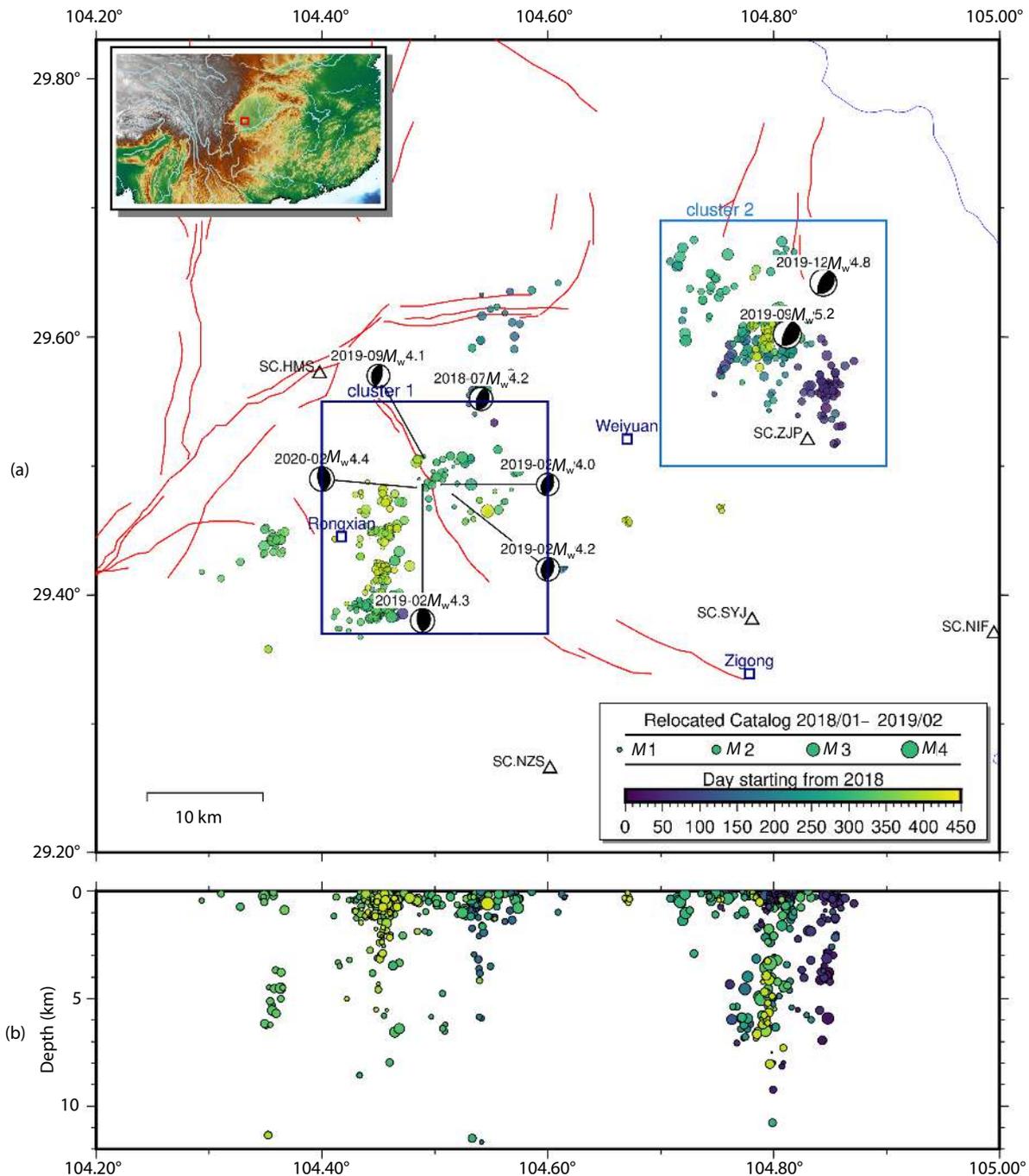


Figure 6. (a) Relocation results. The color and size refer to the times of occurrence and the magnitudes of the events. (b) Depth plot of the relocated events. In total, 1002 events were relocated.

version. Overall, the resulting catalog displayed decreased resolution; using the closest four stations did not demonstrate significant improvement (Figure S3). Therefore, the hypocenter locations appear to be better confined by use of all available station data; restricting the analysis to data from fewer stations does not appear to have any advantages.

The initial hypoDD relocation method resulted in significant loss of relatively large magnitude earthquakes. Their loss appears to be due to the limited similarity of arrivals between the large and small magnitude earthquakes, as their source time functions are different. Therefore, for earthquakes of magnitude 2.5 and above,

we used phase arrivals instead of the combined approach of phase and cross-correlation relocation. The retention rate of the resulting catalog reached 80%. To validate the accuracy of locations of the relatively large earthquakes, we selected the 2019 February Rongxian earthquake sequence and compared results of our new method with the results reported by a study in which focal mechanisms were determined (Yi GX et al., 2020) (Table 2). In general the reported locations are consistent with each other; reported depths of these three earthquakes differ from between 1.4 and 2 km (Table 2). However, we should note that our depths refer to the hypocenters, while the depths in Yi GX et al. (2020) are

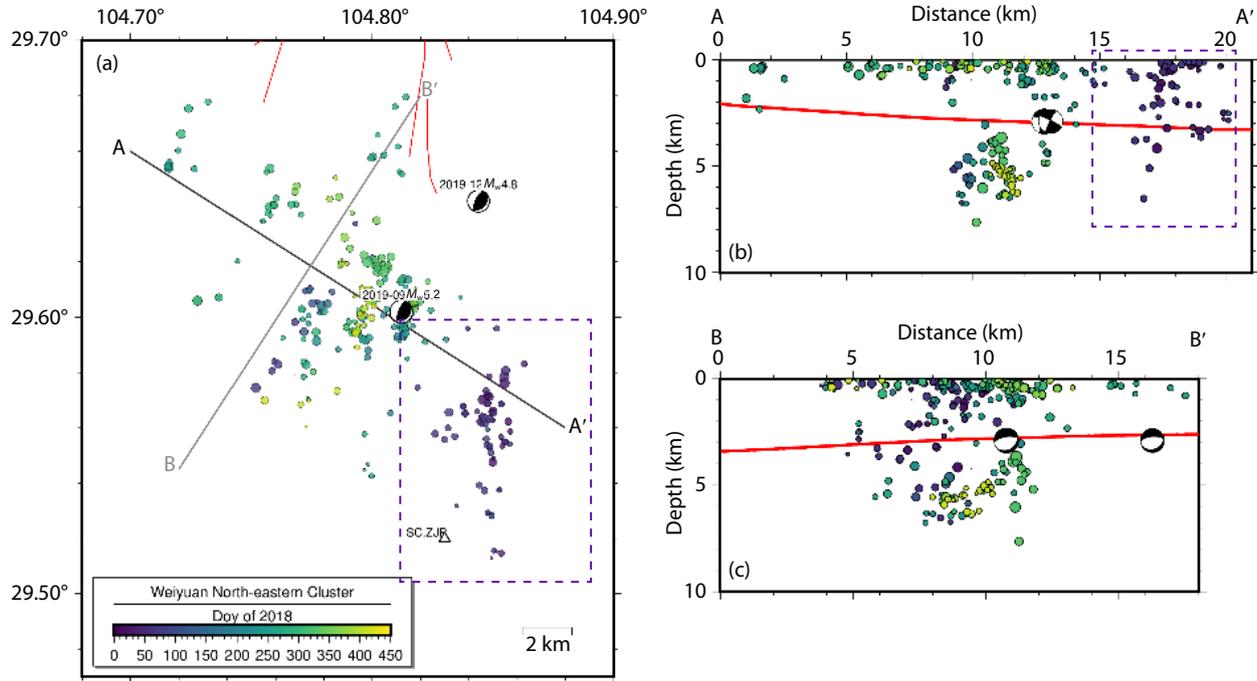


Figure 7. (a) Detailed view on the northeastern seismic cluster. Events are colored by dates and sized by magnitude. (b) Hypocenter depth projection along AA' and (c) hypocenter depth projection along BB'.

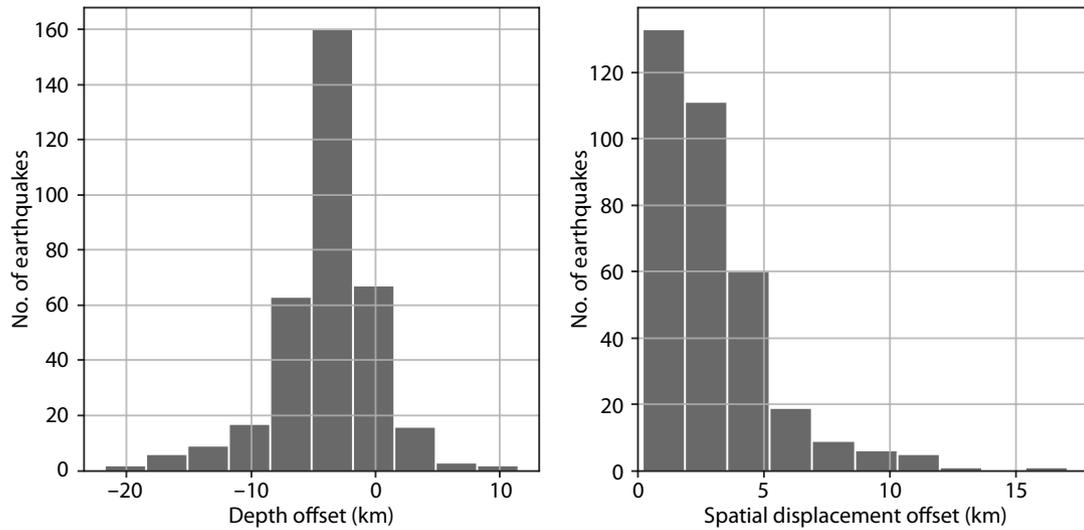


Figure 8. Hypocenter offset comparison between relocation and catalog.

Table 2. Comparison of locations of relatively large magnitude earthquakes reported by this study and one previous study.

Date and time	M_L	Lat.(°)	Lon.(°)	Depth	Source
2019/02/25T00:40:27	4.3	29.508	104.499	1.2	This study
		29.486	104.503	2.5	Yi GX et al. (2020)
2019/01/23T06:40:37	3.8	29.505	104.500	0.8	This study
		20.492	104.494	3	Yi GX et al. (2020)
2019/01/07T11:13:06	3.5	29.446	104.469	2.7	This study
		29.443	104.454	4	Yi GX et al. (2020)

centroid depths. Nevertheless, our hypocenter depths are much shallower than the network reports, which were typically at 10 km or greater for these earthquakes.

4.4 Magnitude and Completeness of the Catalog

In Equation (1), correction parameters k and C are determined using the least squares method, where k is the distance correction parameter representing the geometric spreading and attenuation, and C is the base level in the magnitude calculation. Since we do not have the complete responses of the instruments used in the permanent network, we applied a 1 Hz highpass filter directly to remove low-frequency noise and collect the peak ground motion amplitude of the filtered east and north component waveform (Equation (2)), and thus obtained the local magnitude for each detected event.

Using 1,480 reference events and the least-squares method, k is determined as 2.001; C is -4.867 (Figure S4). The final local magnitude of an event is the median of all the calibrated magnitudes from the nine permanent stations. The calibrated local magnitude results are consistent with the reported events, with a standard deviation of 0.13 and only one outlier event, underdetermined by 0.7 (Figure S5).

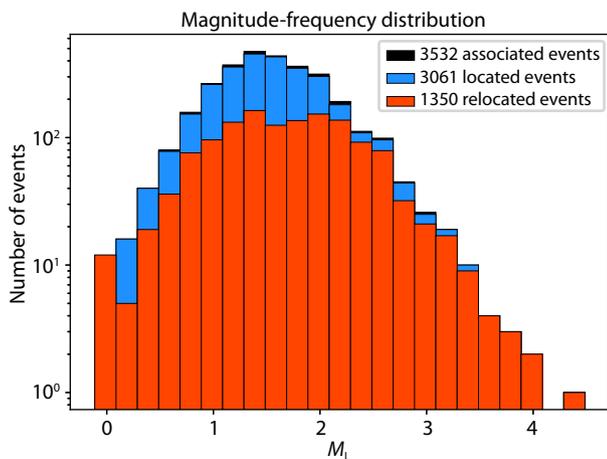


Figure 9. Magnitude frequency distribution of detected events. Black, blue, and red distributions refer to grid-search association, absolute location, and relocation results. The magnitude completeness and the b value of the catalog are 1.75 and 1.114 respectively.

Using the machine learning detector, we have detected earthquakes as small as M_L 0. We have calculated the b value of our newly acquired catalog. The magnitude and frequency distribution of our located earthquakes, calculated using the python software EQCORRSCAN (Chamberlain et al., 2018), suggests that our catalog becomes incomplete below magnitude 1.75 (Figure 9 & 10). The corresponding b value for catalog events of magnitude larger than 1.75 is 1.114 using the method proposed by Weimer and Wyss (2000). We conclude that despite the limited event waveforms and the sparse permanent station network, our approach can deliver high catalog magnitude completeness.

5. Discussion

5.1 Machine Learning and Waveform Cross-correlation Method in Constructing an Event Catalog

The machine learning method and waveform cross-correlation application not only improve the spatial-temporal resolution of the earthquakes but also suggest a feasible method to be applied routinely to all datasets available in the Sichuan Basin. Use of the approach described here can provide essential event information to assess the hydraulic fracturing risk in real time, including monitoring the crucial comparison of earthquake depths to the depth of the hydraulic fracturing layer. The low cost of computation using the machine learning method is advantageous. PhaseNet processing of a year's data from all nine stations was completed within 6 hours, using a 40-CPU-core computer. The efficiency of this machine learning application could be further improved using the latest GPU implementation. In addition, the linear scaling between dataset size and neural network operations in the method we have described greatly reduces runtime with expanding datasets, especially when compared to the template matching method (Yang et al., 2009; Peng and Zhao, 2009; Meng et al., 2018), whose exponential scaling between runtime and data and template size (Perol et al., 2018) results in comparatively slow and limited performance on such large and growing datasets. A great advantage of the machine learning application is that it can be scaled up economically to take advantage of all the available data in the region.

The machine learning method has proved its robustness and reliability in detecting emerging induced seismic swarms compared to the traditional method of moving-average detection and manual picking. With implementation of waveform cross correla-

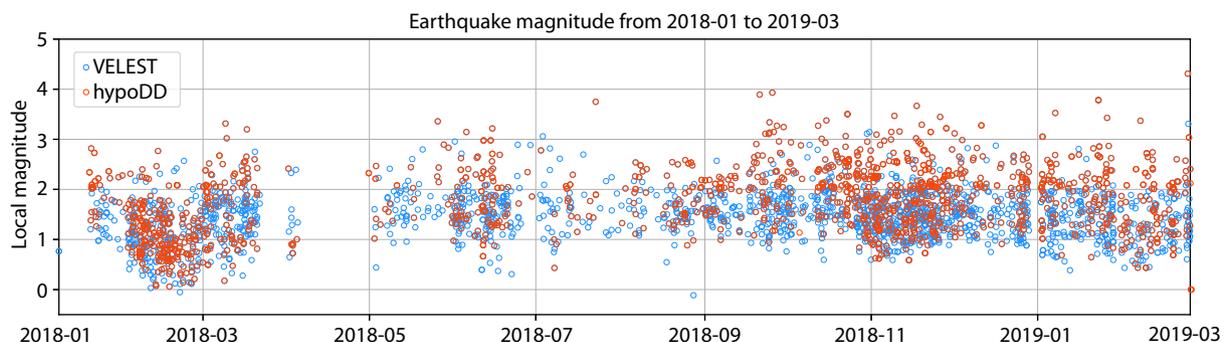


Figure 10. Time series of event-magnitude plot. Red and blue circles refer to hypoDD results and VELEST results respectively.

tion, such improved arrival differential times have led to a better relocation of earthquakes using the sparse seismic stations in the region. Since March 2019, 14 additional local seismic stations have been deployed in the Rongxian–Weiyuan region, significantly improving the network coverage and lowering the magnitude threshold of earthquake detection. With more precise arrival picks from the machine learning method, we can anticipate significant improvements in the accuracy of earthquake catalog locations.

Increased completeness and accuracy of earthquake catalogs will provide a comprehensive framework for assessing risks associated with induced earthquakes. Rapid assessment of emerging seismic swarms is critical to the detection of earthquakes induced by injection and hydraulic fracturing. Estimating the risks from earthquakes induced by anthropogenic activities is an active field of research (Lee et al., 2019). The latest assessment scheme is based on a real-time estimation of seismic risk by use of the Gutenberg–Richer relation (Gutenberg and Richter, 1956) and the nonhomogeneous Poisson process of induced events (Shapiro et al., 2010). Our method offers a reliable and complete earthquake catalog that includes small induced events, which are vital both statistically and physically to assessment of the risks of induced earthquakes.

5.2 Confined and Clustered Earthquakes Between January 2018 and February 2019

Based on both their absolute locations and our computed relocations, we report that most of the earthquakes studied here, from January 2018 to February 2019, were confined to depths of 5 kilometres or less. This conclusion is in contrast with the published catalog depths, which range as deep as 20 kilometres. However, our retrospective investigation of the locations of these earthquakes is consistent with the preliminary locations reported during the subsequent period when additional temporary and permanent stations had been deployed. Early analysis of data from these augmented local arrays suggests that the induced seismicity in the Weiyuan area is of limited depth (Zhou PC et al., 2021; Zi JP et al., 2021), in agreement with — and we believe confirming the plausibility of — the results of our reanalysis of the limited earlier data.

We suggest that this agreement confirms the ability of our machine learning detection and waveform cross-correlation method to determine earthquake locations effectively and robustly even when station coverage and data have been limited.

Natural earthquakes are known to occur at depths beyond tens of kilometres. Our finding that the earthquakes reanalyzed here occurred at much shallower depths strongly supports the growing evidence that these events are likely to have been induced by hydraulic fracturing activities.

Recent seismicity in the Weiyuan area has also been highly clustered both spatially and temporally. Although the availability of the event waveforms used in the present analysis are biased to periods with higher seismicity rates, our catalog clearly reveals the emergence and migrated clustering of events both along the Molin fault and in the northeastern Weiyuan area (Figure 2c). These clusters have generally lasted for a few weeks and had a

spatial extent of 10 to 20 kilometres. The size and the duration of these seismic clusters suggests correlation with active fracking activities within or in proximity to the clusters.

5.3 Potential Fault Geometry Leading to the 2019 M_w 5.2 Earthquake

The improved detail resulting from implementation of machine learning picking and waveform cross correlation relocating reveals multiple lineations in the emerging seismic clusters in the Weiyuan region. For example, the northeastern seismic clusters across the study period exhibit multiple lineations trending southwest to northeast (Figure 7). These earthquake lineations suggest fault activation over the hydraulic fracturing inducing period. The orientations of the lineations are consistent with the principal stress geometry in the Weiyuan area. According to well-logging analysis, the orientation of the local maximum principal stress is roughly N90°E (Meng XB et al., 2018). The maximum principal stress axis estimated from earthquake focal mechanism of the major fracking sites in the Weiyuan area is oriented sub-horizontal with an azimuth of 106° (Lei XL et al., 2020). The lineation illuminated by the seismicity suggests that the fault is optimally oriented with respect to the regional stress field and could be readily reactivated.

Repeated fault failures following induced seismicity are common and have been reported extensively. For example, hydraulic fracturing activities in Duverney, Canada, have induced repeated fault failure over a basement-rooted fault that cuts through the injection layer and the seismogenic region (Eyre et al., 2019). However, multiple inducing mechanisms — including pore-pressure diffusion, poroelastic stress transfer, and the coupled aseismic slip — can lead to repeated failure of a pre-existing fault (Atkinson et al., 2020). Differentiating these inducing mechanisms requires extensive wealth of high resolution spatial and temporal seismic and well operation data. Limited catalog resolution and well operation data prevent further elucidation of the inducing mechanism in the Weiyuan area at this time. Until additional stations are deployed, the evolution of seismicity in this critical region cannot be assessed quantitatively.

Nonetheless, our retrospective analysis has demonstrated that data from induced seismicity can illuminate unmapped fault geometry. The relocated seismicity in the northeastern section is clustered on planar structures and also coincides with the later M_w 5.2 earthquake in September 2019 (Figure 7). The disclosure of this unmapped fault surface has significant potential in future analysis of the risks of further induced earthquakes, including assessment of potential rupture scenarios.

5.4 Potential Pore Pressure Diffusion Inducing Mechanism for an Individual Cluster

One of the earthquake swarms that occurred in the northeastern side of the Weiyuan region in February 2019 was confined along a fault surface (events circled in purple box in Figure 7), thus allowing us to investigate the swarm's spatial–temporal evolution. The hypoDD high resolution location results display a distinctive spatial–temporal evolution coinciding with the theoretical diffusion front, with hydraulic diffusivity of 1 m²/s (Figure 11), implying that

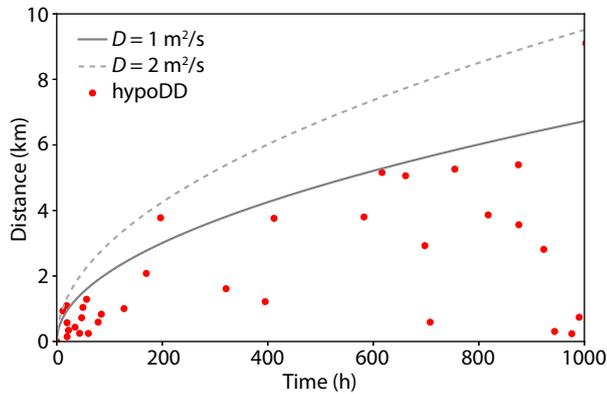


Figure 11. Seismicity migration of the northeastern cluster in February 2019, and the corresponding pore-pressure front. The diffusion front is calculated with Equation (6). First recorded event in the cluster is selected as the initial time and location of injection.

this seismic cluster of earthquakes might have been caused by pore-pressure diffusion.

Although one of the clusters has suggested strong evidence of a pore-pressure diffusion inducing mechanism, the inducing mechanisms of other emerging seismicity clusters in the region remain unclear; the possible relationship between hydraulic fracturing activities and the large earthquakes in 2019 are particularly perplexing (Yang HF et al., 2020; Yang and Yao, 2021). Future work is needed with better network coverage to elucidate the potential mechanisms leading to these earthquakes.

6. Conclusion

In our study, machine learning has demonstrated promising results in detecting a large number of seismic phases from limited available waveform data. The phase picks by PhaseNet have shown consistency with and improvement on the conventional catalog's manual picks. Our approach is characterized as follows. Grid-search method using local velocity profile modified from Zhao et al. (1997) and Meng et al. (2018) is used to associate these phases. Absolute location with VELEST is performed with the associated phase, and the earthquakes with large azimuth gap and hypocenter uncertainty are discarded (Kissling et al., 1995). To further improve hypocenter location, cross-correlation on the phase-pick waveform is conducted. The corresponding phase-pair shifting and correlation coefficient are inputted in the double differencing relocation algorithm. The resulting catalog has an average error in hypocenter determination of less than 200 meters. The final catalog consists of 4018 earthquakes, 25% of which are unique, not included in the Sichuan Earthquake Agency catalog.

The hypocenter depths of the Weiyuan earthquakes studied here do not display the diffusive range suggested by the previous catalog. Most of the earthquakes relocated by our method are confined within 5 kilometres depths. The relocated catalog clearly displays at least 6 emerging clusters in the Weiyuan area from 2018 to March 2019, mirroring the rapid development of hydraulic fracturing activities in the study area. In addition, after high precision relocation by the double-difference algorithm, it becomes clear

that earthquakes in the northeastern part of the Weiyuan region exhibited preliminary fault lineation. The location of the fault lineation coincides with the centroid location of the 2019 M_w 5.2 event, by far the largest Weiyuan event. This newly described lineation suggests a possible unmapped fault geometry associated with the rupture event. The improved catalog and detailed structures of the emerging seismicity facilitate continued investigation of the consequences of active hydraulic fracturing in the Sichuan Basin.

Acknowledgments

This study was supported by the National Key R&D Program of China (2018YFC1504501); the Hong Kong Research Grants Council (No. 14303721 and N_CUHK430/16); and the Faculty of Science, CUHK.

References

- Atkinson, G. M., Eaton, D. W., and Igonin, N. (2020). Developments in understanding seismicity triggered by hydraulic fracturing. *Nat. Rev. Earth Environ.*, 1(5), 264–277. <https://doi.org/10.1038/s43017-020-0049-7>
- Bao, X. W., and Eaton, D. W. (2016). Fault activation by hydraulic fracturing in western Canada. *Science*, 354(6318), 1406–1409. <https://doi.org/10.1126/science.aag2583>
- Cesca, S., Grigoli, F., Heimann, S., González, Á., Buforn, E., Maghsoudi, S., Blanch, E., and Dahm, T. (2014). The 2013 September–October seismic sequence offshore Spain: a case of seismicity triggered by gas injection?. *Geophys. J. Int.*, 198(2), 941–953. <https://doi.org/10.1093/gji/ggu172>
- Chamberlain, C. J., Hopp, C. J., Boese, C. M., Warren-Smith, E., Chambers, D., Chu, S. X., Michailos, K., and Townend, J. (2018). EQcorrscan: repeating and near-repeating earthquake detection and analysis in Python. *Seismol. Res. Lett.*, 89(1), 173–181. <https://doi.org/10.1785/0220170151>
- Clarke, H., Eisner, L., Styles, P., and Turner, P. (2014). Felt seismicity associated with shale gas hydraulic fracturing: the first documented example in Europe. *Geophys. Res. Lett.*, 41(23), 8308–8314. <https://doi.org/10.1002/2014GL062047>
- Ellsworth, W. L. (2013). Injection-induced earthquakes. *Science*, 341(6142), 1225942. <https://doi.org/10.1126/science.1225942>
- Eyre, T. S., Eaton, D. W., Garagash, D. I., Zecevic, M., Venieri, M., Weir, R., and Lawton, D. C. (2019). The role of aseismic slip in hydraulic fracturing-induced seismicity. *Science Advance*, 5(8), eaav7172. <https://doi.org/10.1126/sciadv.aav7172>
- Foulger, G. R., Wilson, M. P., Gluyas, J. G., Julian, B. R., and Davies, R. J. (2018). Global review of human-induced earthquakes. *Earth-Sci. Rev.*, 178, 438–514. <https://doi.org/10.1016/j.earscirev.2017.07.008>
- Gale, J. F. W., Laubach, S. E., Olson, J. E., Eichhubl, P., and Fall, A. (2014). Natural fractures in shale: a review and new observations. *AAPG Bull.*, 98(11), 2165–2216. <https://doi.org/10.1306/08121413151>
- Grigoli, F., Cesca, S., Priolo, E., Rinaldi, A. P., Clinton, J. F., Stabile, T. A., Dost, B., Fernandez, M. G., Wiemer, S., and Dahm, T. (2017). Current challenges in monitoring, discrimination, and management of induced seismicity related to underground industrial activities: A European perspective. *Rev. Geophys.*, 55(2), 310–340. <https://doi.org/10.1002/2016RG000542>
- Gutenberg, B., and Richter, C. F. (1956). Earthquake magnitude, intensity, energy, and acceleration (Second paper). *Bulletin of the Seismological Society of America*, 46(2), 105–145. <https://doi.org/10.1785/BSSA0320030163>
- Jiang, G. Y., Qiao, X. J., Wang, X. Q., Lu, R. Q., Liu, L., Yang, H. F., Su, Y. D., Song, L. L., Wang, B. S., and Wong, T. F. (2020). GPS observed horizontal ground extension at the Hutubi (China) underground gas storage facility and its application to geomechanical modeling for induced seismicity. *Earth Planet. Sci. Lett.*, 530, 115943. <https://doi.org/10.1016/j.epsl.2019.115943>
- Jiang, G. Y., Liu, L., Barbour, A. J., Lu, R. Q., and Yang, H. F. (2021). Physics-based evaluation of the maximum magnitude of potential earthquakes induced

- by the Hutubi (China) underground gas storage. *J. Geophys. Res.*, 126(4), e2020JB021379. <https://doi.org/10.1029/2020JB021379>
- Kissling, E., Kradolfer, U., and Maurer, H. (1995). Program VELEST User's Guide—short Introduction. Zurich: Institute of Geophysics.
- Kong, Q. K., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., and Gerstoft, P. (2019). Machine learning in seismology: turning data into insights. *Seismol. Res. Lett.*, 90(1), 3–14. <https://doi.org/10.1785/0220180259>
- Lee, K. K., Ellsworth, W. L., Giardini, D., Townend, J., Ge, S., Shimamoto, T., Yeo, I. W., Kang, T. S., Rhie, J., and Sheen, D. H. (2019). Managing injection-induced seismic risks. *Science*, 364(6442), 730–732. <https://doi.org/10.1126/science.aax1878>
- Lei, X. L., Ma, S. L., Chen, W. K., Pang, C. M., Zeng, J., and Jiang, B. (2013). A detailed view of the injection-induced seismicity in a natural gas reservoir in Zigong, southwestern Sichuan Basin, China. *J. Geophys. Res.: Solid Earth*, 118(8), 4296–4311. <https://doi.org/10.1002/jgrb.50310>
- Lei, X. L., Huang, D. J., Su, J. R., Jiang, G. M., Wang, X. L., Wang, H., Guo, X., and Fu, H. (2017). Fault reactivation and earthquakes with magnitudes of up to Mw4.7 induced by shale-gas hydraulic fracturing in Sichuan Basin, China. *Scientific Reports*, 7(7971), 1–12. <https://doi.org/10.1038/s41598-017-08557-y>
- Lei, X. L., Wang, Z. W., and Su, J. R. (2019). The December 2018 M_L 5.7 and January 2019 M_L 5.3 earthquakes in South Sichuan Basin induced by shale gas hydraulic fracturing. *Seismol. Res. Lett.*, 90(3), 1099–1110. <https://doi.org/10.1785/0220190029>
- Lei, X. L., Su, J. R., and Wang, Z. W. (2020). Growing seismicity in the Sichuan Basin and its association with industrial activities. *Sci. China Earth Sci.*, 63(11), 1633–1660. <https://doi.org/10.1007/s11430-020-9646-x>
- Liang, X., Liu, S. G., Wang, S. B., Deng, B., Zhou, S. Y., and Ma, W. X. (2019). Analysis of the oldest carbonate gas reservoir in China—new geological significance of the dengying gas reservoir in the Weiyuan Structure, Sichuan Basin. *J. Earth Sci.*, 30(2), 348–366. <https://doi.org/10.1007/s12583-017-0962-y>
- Peng, Z., and Zhao, P. (2009). Migration of early aftershocks following the 2004 Parkfield earthquake. *Nature Geoscience*, 2(12), 877–881. <https://doi.org/10.1038/ngeo697>
- Perol, T., Gharbi, M., and Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Sci. Adv.*, 4(2), e1700578. <https://doi.org/10.1126/sciadv.1700578>
- Majer, E. L., Baria, R., Stark, M., Oates, S., Bommer, J., Smith, B., and Asanuma, H. (2007). Induced seismicity associated with enhanced geothermal systems. *Geothermics*, 36(3), 185–222. <https://doi.org/10.1016/j.geothermics.2007.03.003>
- Meng, X. B., Chen, H. C., Niu, F. L., Tang, Y. C., Yin, C., and Wu, F. R. (2018). Microseismic monitoring of stimulating shale gas reservoir in SW China: 1. An improved matching and locating technique for downhole monitoring. *J. Geophys. Res.: Solid Earth*, 123(2), 1643–1658. <https://doi.org/10.1002/2017JB014488>
- Schultz, R., Skoumal, R. J., Brudzinski, M. R., Eaton, D., Baptie, B., and Ellsworth, W. (2020). Hydraulic fracturing-induced seismicity. *Rev. Geophys.*, 58(3), e2019RG000695. <https://doi.org/10.1029/2019RG000695>
- Shapiro, S. A., Rotherth, E., Rath, V., and Rindschwentner, J. (2002). Characterization of fluid transport properties of reservoirs using induced microseismicity. *Geophysics*, 67(1), 212–220. <https://doi.org/10.1190/1.1451597>
- Shapiro, S. A., Dinske, C., Langenbruch, C., and Wenzel, F. (2010). Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations. *Lead. Edge*, 29(3), 304–309. <https://doi.org/10.1190/1.3353727>
- Shapiro, S. A. (2015). *Fluid-Induced Seismicity*. Cambridge: Cambridge University Press.
- Sheng, M. H., Chu, R. S., Ni, S. D., Wang, Y., Jiang, L. M., and Yang, H. F. (2020). Source parameters of three moderate size earthquakes in Weiyuan, China, and their relations to shale gas hydraulic fracturing. *J. Geophys. Res.*, 125(10), e2020JB019932. <https://doi.org/10.1029/2020JB019932>
- Waldhauser, F., and Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: method and application to the Northern Hayward Fault, California. *Bull. Seismol. Soc. Am.*, 90(6), 1353–1368. <https://doi.org/10.1785/0120000006>
- Wang, M. M., Yang, H. F., Fang, L. H., Han, L. B., Jia, D., Jiang, D. Q., and Yan, B. (2020). Shallow faults reactivated by hydraulic fracturing: the 2019 Weiyuan earthquake sequences in Sichuan, China. *Seismol. Res. Lett.*, 91(6), 3171–3181. <https://doi.org/10.1785/0220200174>
- Wang, R. J., Schmandt, B., Zhang M., Glasgow, M., Kiser, E., Rysanek, S., and Stairs, R. (2020). Injection-induced earthquakes on complex fault zones of the raton basin illuminated by machine-learning phase picker and dense nodal array. *Geophys. Res. Lett.*, 47(14), e2020GL088168. <https://doi.org/10.1029/2020GL088168>
- Weingarten, M., Ge, S., Godt, J. W., Bekins, B. A., and Rubinstein, J. L. (2015). High-rate injection is associated with the increase in U.S. mid-continent seismicity. *Science*, 348(6241), 1336–1340. <https://doi.org/10.1126/science.aab1345>
- Wiemer, S., and Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western United States, and Japan. *Bull. Seismol. Soc. Am.*, 90(4), 859–869. <https://doi.org/10.1785/0119990114>
- Yang, H. F., Liu, Y. J., Wei, M., Zhuang, J. C., and Zhou, S. Y. (2017). Induced earthquakes in the development of unconventional energy resources. *Sci. China Earth Sci.*, 60(9), 1632–1644. <https://doi.org/10.1007/s11430-017-9063-0>
- Yang, H. F., Zhou, P. C., Fang, N., Zhu, G. H., Xu, W. B., Su, J. R., Meng, F. B., and Chu, R. S. (2020). A shallow shock: the 25 February 2019 ML 4.9 earthquake in the Weiyuan Shale gas field in Sichuan, China. *Seismol. Res. Lett.*, 91(6), 3182–3194. <https://doi.org/10.1785/0220200202>
- Yang, H., Zhu, L., and Chu, R. (2009). Fault-Plane Determination of the 18 April 2008 Mount Carmel, Illinois, Earthquake by Detecting and Relocating Aftershocks. *Bull. Seismol. Soc. Am.*, 99(6), 3413–3420. <https://doi.org/10.1785/0120090038>
- Yang, H. F., and Yao, S. L. (2021). Shallow destructive earthquakes. *Earthquake Science*, 34(1), 15–23. <https://doi.org/10.29382/eqs-2020-0072>
- Yi, G. X., Long, F., Liang, M. J., Zhao, M., and Wang, S. W. (2020). Geometry and tectonic deformation of seismogenic structures in the Rongxian-Weiyuan-Zichong region, Sichuan Basin: insights from focal mechanism solutions. *Chin. J. Geophys.*, 63(9), 3275–3291. <https://doi.org/10.6038/cjg202000095>
- Yoon, C. E., Huang, Y. H., Ellsworth, W. L., and Beroza, G. C. (2017). Seismicity during the initial stages of the guy-greenbrier, arkansas, earthquake sequence. *J. Geophys. Res.: Solid Earth*, 122(11), 9253–9274. <https://doi.org/10.1002/2017JB014946>
- Zhang, M., Ellsworth, W. L., and Beroza, G. C. (2019). Rapid earthquake association and location. *Seismol. Res. Lett.*, 90(6), 2276–2284. <https://doi.org/10.1785/0220190052>
- Zhao, Z., Fan, J., Zheng, S. H., Hasegawa, A., and Horiuchi, S. (1997). Precision determination of the crustal structure and hypocentral locations in the Longmenshan thrust nappe belt. *Acta Seismol. Sin.*, 19(6), 615–622.
- Zhou, P. C., Yang, H. F., Wang, B. S., and Zhuang, J. C. (2019). Seismological investigations of induced earthquakes near the Hutubi underground gas storage facility. *J. Geophys. Res.: Solid Earth*, 124(8), 8753–8770. <https://doi.org/10.1029/2019JB017360>
- Zhou, P. C., Ellsworth, W. L., Yang, H. F., Tan, Y. J., Beroza, G. C., Sheng, M. H., and Chu, R. S. (2021). Machine-learning-facilitated earthquake and anthropogenic source detections near the Weiyuan Shale Gas Blocks, Sichuan, China. *Earth Planet. Phys.*, 5(6), 501–519. <https://doi.org/10.26464/epp2021053>
- Zhu, W. Q., and Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.*, 216(1), 261–273. <https://doi.org/10.1093/gji/ggy423>
- Zoback, M. D., and Gorelick, S. M. (2012). Earthquake triggering and large-scale geologic storage of carbon dioxide. *Proc. Natl. Acad. Sci. USA*, 109(26), 10164–10168. <https://doi.org/10.1073/pnas.1202473109>

Supplementary Materials

S1 Derivation of the Pore-pressure Front

The hydraulic fracturing operation is approximated by a point source perturbation in an infinite heterogenous anisotropic poroelastic fluid-saturated medium. The pressure evolution due to the injection is described by the following diffusion equation:

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x_i} \left[D_{ij} \frac{\partial p}{\partial x_j} \right], \tag{A1}$$

where D_{ij} is the hydraulic diffusivity tensor of the material. The hydraulic diffusivity can be pressure-dependent in some situations; the equation then becomes non-linear. Considering a simplified case with pressure-independent, isotropic hydraulic diffusive and homogeneous material, the following is the pressure evolution solution in spherical coordinates satisfying the diffusion equation, given a time-harmonic pressure perturbation $p_0 \exp(i\omega t)$ on a small spherical surface of radius a at origin.

$$p(r, t) = p_0 e^{-i\omega t} \frac{a}{r} \exp \left[(i-1)(r-a) \sqrt{\frac{\omega}{2D}} \right], \tag{A2}$$

where ω is the angular frequency of the injection pore-pressure perturbation.

However, the injection in reality is a step function rather than a simple harmonic function. Thus, the boundary condition for the injection point is a rectangular source function of

$$p(t) = \begin{cases} 0, & \text{if } t < 0, \\ p_0, & \text{if } t \geq t_0. \end{cases} \tag{A3}$$

The corresponding power spectrum of the source function is a well-known sinc function with the dominant part of the signal ranges below $2\pi/t_0$. It is assumed that the probability of a large amplitude low frequency triggering event is higher, and is $2\pi/t_0$. Thus, considering that the seismic events are characterized by the high probability front (Shapiro et al., 2002), we obtain

$$r = \sqrt{4\pi D t}. \tag{A4}$$

Table S1. Configuration of hypoDD relocations for Weiyuan events.

IDAT		IPHA		MAXDIST					
3		3		200					
OBSCC		OBSCT		MINDIST		MAXDIST		MAXGAP	
0		4		-999		-999		-999	
ISTART		ISOLV		IAQ		NSET			
2		2		1		5			
NITER	WTCCP	WTCCS	WRCC	WDCC	WTCTP	WTCS	WRCT	WDCT	DAMP
3	0.01	0.005	-999	-999	1.00	0.5	20	-999	120
3	0.01	0.005	-999	-999	1.00	0.5	10	4	80
3	1	0.5	6	3	1.00	0.5	10	3	60
3	1	0.5	5	3	0.1	0.05	9	3	50
3	1	0.5	4	3	0.01	0.005	8	3	50

Table S2. Configuration of hypoDD relocations for northeastern Weiyuan cluster.

IDAT		IPHA		MAXDIST					
3		3		200					
OBSCC		OBSCT		MINDIST		MAXDIST		MAXGAP	
4		4		-999		-999		-999	
ISTART		ISOLV		IAQ		NSET			
2		2		1		5			
NITER	WTCCP	WTCCS	WRCC	WDCC	WTCTP	WTCS	WRCT	WDCT	DAMP
4	0.01	0.005	-999	-999	1.00	0.5	20	-999	150
4	0.01	0.005	-999	-999	1.00	0.5	5	4	100
4	1	0.7	-999	4	0.01	0.01	5	4	90
4	1	0.7	5	1	0.01	0.01	5	4	80
4	1	0.7	5	0.5	0.01	0.01	5	4	70

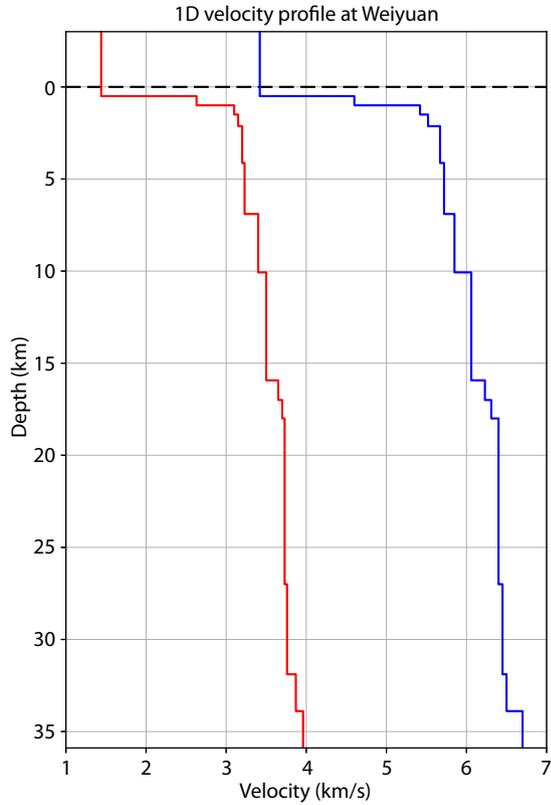


Figure S1. Velocity profile of Weiyuan. Dash line at zero depth refers to sea level.

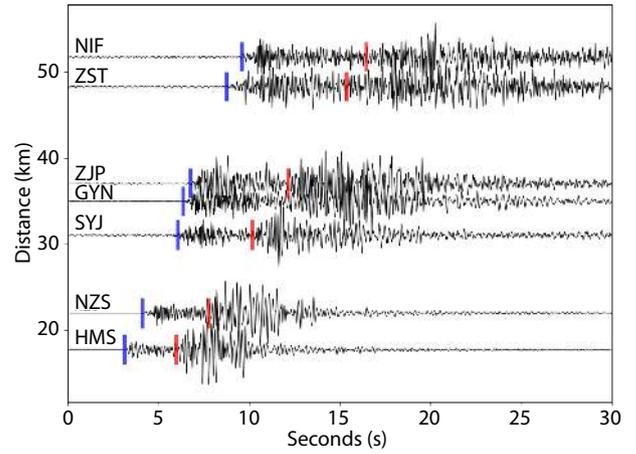


Figure S2. Section plot of the detection using machine learning detector. Vertical component is shown in the plot. Blue bars and red bars refer to the arrival pick by the machine learning detector.

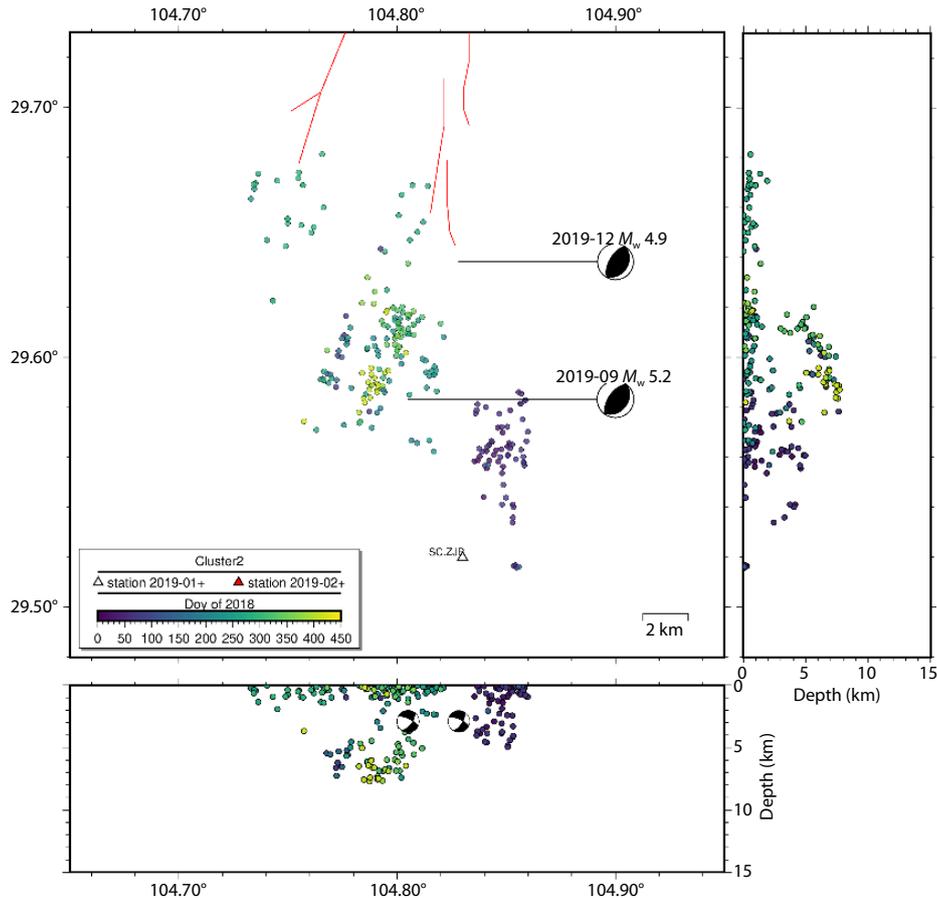


Figure S3. Relocated northeastern Weiyuan cluster using data from the four closest stations. Colored by date of occurrence.

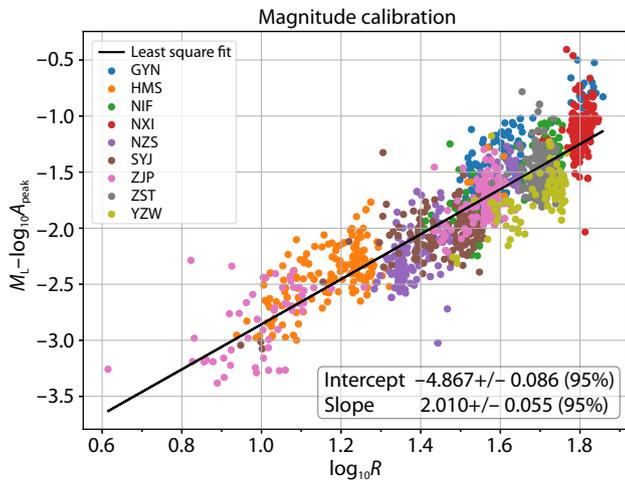


Figure S4. Magnitude calibration results. Scatter plot of the linear fit of the magnitude calibration using local magnitudes from the Sichuan Earthquake agency: 1480 events and maximum peak horizontal motion. The least squares results $k = 2.001$, $C = -4.867$ in the calibration equation.

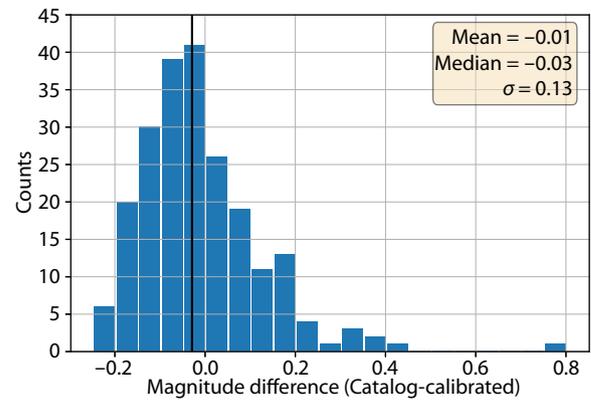


Figure S5. Histogram of local magnitude difference with the reference events. All events' magnitudes, except one, computed by Equation (1) differ less than 0.5 with reference magnitude from the Sichuan Earthquake Agency.