

Machine-learning-facilitated earthquake and anthropogenic source detections near the Weiyuan Shale Gas Blocks, Sichuan, China

PengCheng Zhou^{1,2,3}, William L. Ellsworth³, HongFeng Yang^{2*}, Yen Joe Tan^{2,3}, Gregory C. Beroza³, MinHan Sheng⁴, and RiSheng Chu⁴

¹Key Laboratory of Ocean and Marginal Sea Geology, South China Sea Institute of Oceanology, Innovation Academy of South China Sea Ecology and Environmental Engineering, Chinese Academy of Sciences, Guangzhou 510301, China;

²Earth System Science Programme, Faculty of Science, The Chinese University of Hong Kong, Sha Tin, Hong Kong 999077, China;

³Department of Geophysics, Stanford University, Stanford, CA 94305, USA;

⁴State Key Laboratory of Geodesy and Earth's Dynamics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430077, China

Key Points:

- We obtain P and S phase arrival times using PhaseNet, a machine-learning phase picker.
- An earthquake catalog for the November 2015 to November 2016 Weiyuan, Sichuan, earthquake sequence is constructed from raw seismic data.
- Our expanded earthquake catalog identifies the spatiotemporal evolution patterns of many events.

Citation: 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. <http://doi.org/10.26464/epp2021053>

Abstract: Seismic hazard assessment and risk mitigation depend critically on rapid analysis and characterization of earthquake sequences. Increasing seismicity in shale gas blocks of the Sichuan Basin, China, has presented a serious challenge to monitoring and managing the seismicity itself. In this study, to detect events we apply a machine-learning-based phase picker (PhaseNet) to continuous seismic data collected between November 2015 and November 2016 from a temporary network covering the Weiyuan Shale Gas Blocks (SGB). Both P- and S-phases are picked and associated for location. We refine the velocity model by using detected explosions and earthquakes and then relocate the detected events using our new velocity model. Our detections and absolute relocations provide the basis for building a high-precision earthquake catalog. Our primary catalog contains about 60 times as many earthquakes as those in the catalog of the Chinese Earthquake Network Center (CENC), which used only the sparsely distributed permanent stations. We also measure the local magnitude and achieve magnitude completeness of M_L0 . We relocate clusters of events, showing sequential migration patterns overlapping with horizontal well branches around several well pads in the Wei202 and Wei204 blocks. Our results demonstrate the applicability of a machine-learning phase picker to a dense seismic network. The algorithms can facilitate rapid characterization of earthquake sequences.

Keywords: induced seismicity; machine learning; Weiyuan Shale Gas Block

1. Introduction

During the past several years, the Southern Sichuan Basin of China has experienced a marked increase in seismicity near hydraulic fracturing (HF) sites, accompanied by increasing seismic hazard as the maximum magnitudes of earthquakes has climbed (Figure 1) (Lei XL et al., 2017, 2019; Meng LY et al., 2019; Sheng MH et al., 2020; Yang HF et al., 2020). Earthquakes associated with the development and production of natural gas from tight formations pose a threat not only to local communities but also to the safe produc-

tion of shale gas (Ellsworth, 2013; Holland, 2013; Clarke et al., 2014; Atkinson et al., 2016; Bao XW and Eaton, 2016; Grigoli et al., 2017; Yang HF et al., 2017). Accurate seismic hazard assessment for potentially induced earthquakes depends critically on how well we know the stress conditions on those faults and the mechanisms of earthquake induction (Walsh III and Zoback, 2016; Schoenball and Ellsworth, 2017; Yu HY et al., 2019).

Various techniques, e.g., focal mechanism inversion, geodetic observations, and geologic surveys, have been widely used to investigate fault structures after earthquakes occur, but all of these approaches have limited resolution of small-scale fault structures, focal depths, and hidden faults (Zhu LP and Helmlinger, 1996; Xu WB et al., 2015; Yang HF, 2015; Zhou PC et al., 2019). This is particularly true for induced earthquake-prone regions, where small

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

Received 25 JUL 2020; Accepted 04 OCT 2021.

Accepted article online 15 NOV 2021.

©2021 by Earth and Planetary Physics.

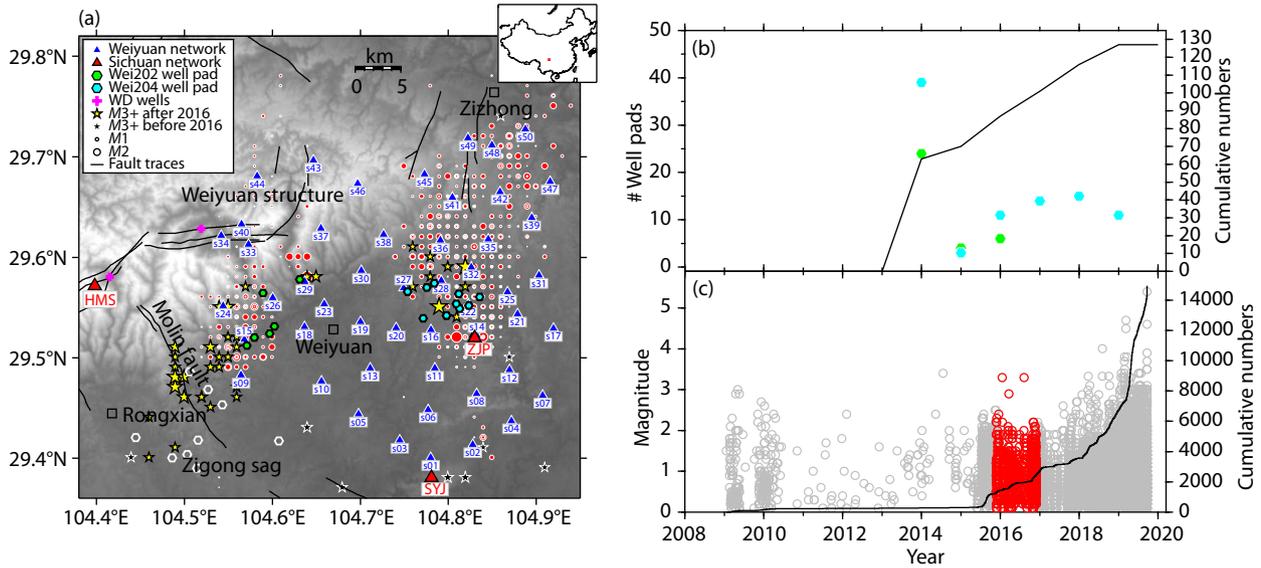


Figure 1. Map view of the study area. (a) Red dots mark epicenters reported by the Chinese Earthquake Network Center (CENC) between 12 November 2015 and 18 November 2016. Black lines represent the surveyed faults (Liang et al., 2019). Seismic stations used by the CENC and in this study are plotted with red and blue triangles, respectively. Shale gas well pads (HF wells) where drilling was completed by 2017 are shown with filled hexagons (green, Wei202 block; cyan, Wei204 block). The open white hexagons mark the locations of some HF well pads drilled after 2017 that are suspected to be related to earthquakes near the Molin fault (Yang HF et al., 2020). The magenta cross marks the location of a Wastewater Disposal (WD) well with an injection depth of 4 km. The stars mark $M \geq 3$ earthquakes (black, before our study period; yellow, after our study period). (Inset) Location map of the Sichuan Basin; the red square indicates the study region. (b) Numbers of wells completely annually and cumulative well count (black curve). (c) Magnitude-time plot of the CENC catalog in the study region, and cumulative count of earthquakes. Earthquakes occurring in our study period in red.

faults, including many that are unmapped, have proved capable of hosting small-to-moderate-size earthquakes (Schoenball et al., 2018). Because many of the potentially seismogenic faults have limited or no surface expression, their detailed geometry is difficult to determine. High-precision earthquake hypocenters have the potential to illuminate the activating faults before damaging earthquakes occur (Rubin et al., 1999; Waldhauser and Ellsworth, 2002; Kao H and Shan SJ, 2007; Peng ZG and Zhao P, 2009; Yang HF et al., 2009; Schoenball et al., 2018). During hydraulic fracturing operations, or later during production, a near-real-time precise earthquake catalog can thus inform both operators and authorities about evolving hazard conditions (Lee et al., 2019; Langenbruch et al., 2020; Schultz et al., 2020).

A combination of the number and accuracy of picked phase arrivals and the distribution of stations determines the quality of an earthquake catalog (Lee and Stewart, 1981). The template matching method (e.g., Gibbons and Ringdal, 2006; Peng ZG and Zhao P, 2009; Yang HF et al., 2009; Skoumal et al., 2015; Zhang M and Wen LX, 2015; Shelly et al., 2016) has proven to be a powerful means to expand seismicity catalogs but relies on the initial number of earthquakes and on the quality and completeness of recorded template waveforms. Unsupervised search for earthquakes using data mining methods are highly effective but currently impractical for real-time application (Yoon et al., 2015, 2019). Conventional automatic algorithms (e.g., short-term average/long-term average STA/LTA method; Allen, 1978) work well for impulsive P-wave arrivals but suffer from limited efficiency and accuracy when signals are emergent, or noise levels are high. Moreover, in

densely instrumented areas where the data flow outpaces seismic analyst data processing rates, many events must be skipped, and S waves may go unanalyzed.

Machine-learning-based techniques have now matured to outperform routine analysis methods (e.g., Ross et al., 2018; Mousavi et al., 2019b; Wang J et al., 2019; Zhu WQ and Beroza, 2019). PhaseNet (Zhu WQ and Beroza, 2019) is a deep-neural-network-based arrival-time picking method. Although it was trained on northern California earthquakes, it has since demonstrated success in picking accurate arrival times of both P and S waves in diverse regions and signal environments, which has made it an attractive choice for routine earthquake detection (Zhang M et al., 2019; Liu M et al., 2020; Park et al., 2020; Wang MM et al., 2020; Tan YJ et al., 2021).

In addition to robust arrival time picking, real-time seismic monitoring and catalog construction depend crucially on accurate event association and location techniques (Johnson et al., 1997; Zhang M et al., 2019; Johnson, 2020). Rapid communication of accurate information between network monitoring and oil field operations, particularly when small earthquakes occur in rapid succession, is fundamentally important for safety (Schultz et al., 2020). Phase association methods follow a variety of strategies to group candidate picks into events that best fit the predicted arrival times (e.g., Stewart, 1977; Ester et al., 1996; Johnson et al., 1997; Grigoli et al., 2018; Bergen and Beroza, 2018; Ross et al., 2019). Among them, the Rapid Earthquake Association and Location (REAL) method (Zhang M et al., 2019) can determine the location directly with sufficient accuracy for decision making, given a

sufficient number of phase picks and a reasonably accurate velocity model. Simultaneous solutions of both velocity and hypocentral parameters are required to achieve precise earthquake locations (e.g., VELEST; Kissling et al., 1994). Recent applications of REAL in several regions have demonstrated its ability to characterize earthquake sequences rapidly from raw seismic data, in combination with a machine-learning phase picker (e.g., PhaseNet), and a sequential workflow (Liu M et al., 2020; Park et al., 2020; Wang RJ et al., 2020; Tan YJ et al., 2021; Wong et al., 2021).

In this study, we investigate earthquakes that occurred in the region surrounding the Weiyuan Shale Gas Blocks (SGB), Sichuan, China, between November 2015 and November 2016, using a temporary network of 50 stations (Figure 1). The deployment occurred during a time when small-magnitude earthquakes were increasing in number, but three years before the deadly M_w 4–5 class earthquakes that occurred in 2019 (Sheng MH et al., 2020; Wang MM et al., 2020; Yang HF et al., 2020). Here we demonstrate the feasibility of rapidly generating a catalog using a dense network and machine-learning algorithms. To generate the base for building a high-resolution earthquake catalog, we follow the workflow of Liu M et al. (2020), starting with phase detection and picking by applying the machine learning picker PhaseNet (Zhu WQ and Beroza, 2019) followed by event association with REAL (Zhang M et al., 2019) and hypocentral location using VELEST (Kissling et al., 1994). Our REAL catalog contains 79,395 events, constructed using 981,284 P and 859,576 S arrival time readings. Based on our VELEST-located earthquakes, we are able further to refine and reveal the seismicity migration patterns surrounding several HF wells.

2. Geologic Settings, Shale Gas Production and Seismic Data

The Weiyuan SGB is located in the Weiyuan–Rongxian counties

and Zizhong City in the Sichuan Province. The geology is characterized by the Weiyuan anticline, an NE–SW trending dome structure with many mapped faults trending in the NE direction on the north flank and several faults trending in the SE direction on the south flank (Figure 1) (Liang X et al., 2019). Between 1980 and 2015, the region within 25-km from Weiyuan County was struck by only 12 earthquakes with magnitudes greater than 3, according to the China Earthquake Network Center (CENC) routine catalog (Figure 1); most of these 12 events were located to the south of our study region, near Zigong City. The occurrence rate of smaller earthquakes ($M < 3$) was as low as about one earthquake per week before April 2015 and after 2009, when the CENC routine catalog became available. However, frequent earthquakes (~12 earthquakes per day) have occurred in the Weiyuan SGB since May 2015, at pace with the development of HF wells in the region (Figure 1). Shale gas resource is abundant in the upper Ordovician Wufeng Formation to lower Silurian Longmaxi Shale Formation (Lei XL et al., 2017; Wu X et al., 2019). By 2015, more than 100 HF well pads were projected, and among them, 17 had been drilled, with typical completion depths of around 4.8 km below ground level (~400 m) in the Wei202 block (near station s15 in Figure 1) and 5.8 km below ground level (~350 m) for Wei204 block (near station s22 in Figure 1) (Safety and Environmental Protection Technology Research Institute of CNPC and Beijing CNPC Construction Project Labor Safety and Health Pre Evaluation Co., Ltd, 2015; Dong DZ et al., 2018). Figure 1 shows HF wells in production or being drilled by late 2015.

A temporary seismic network of 50 stations was deployed in the area from November 2015 to November 2016 (Figure 2a), covering the Weiyuan SGB (Figure 1). Each station was equipped with a Guralp CMG-40T seismograph that was set with a sampling rate of 100 Hz. The average station spacing was 5 km. The network was in normal service starting from late November 2015. However, in the late stage of deployment of the network, some stations failed due

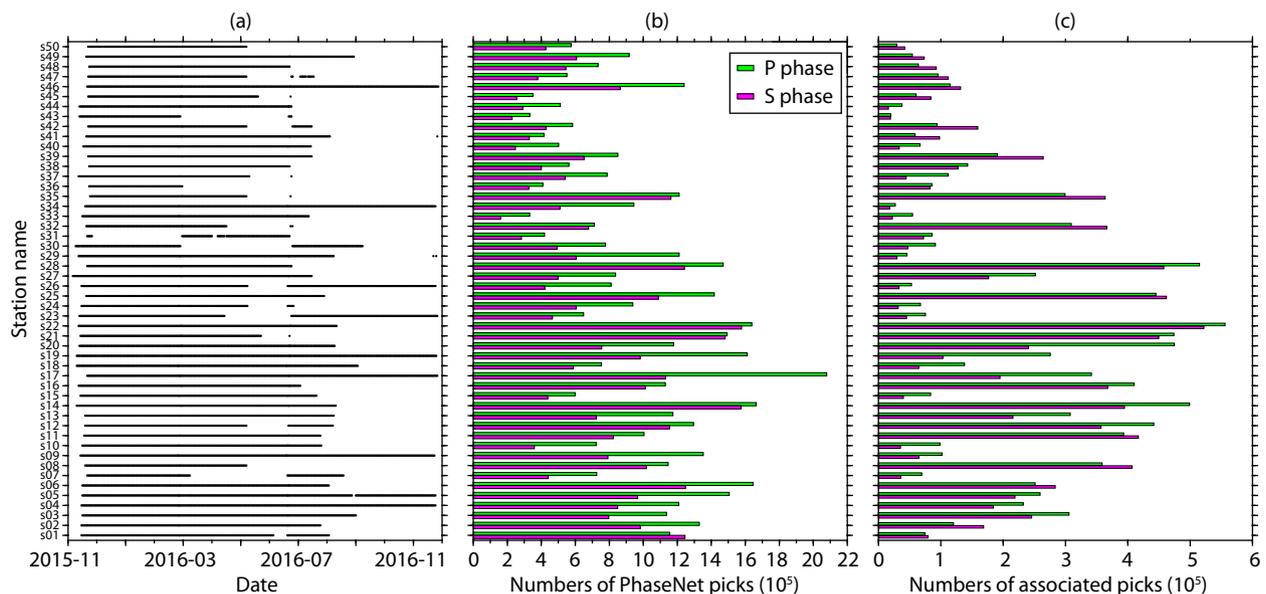


Figure 2. Seismic network operations and phase picks. (a) The x axis shows the days with reference to 2016-01-01; the y axis shows the station names. The continuity of each horizontal lines indicates data availability. (b) Numbers of P/S picks after running PhaseNet. (c) Numbers of P/S picks after association. About 75% of the PhaseNet picks were discarded.

to environmental factors (mostly rain and water), resulting in fewer working stations after June 2016.

The seismic monitoring system in the Weiyuan region relied on the Sichuan network with more distant regional stations (Yang HF et al., 2020). Among them, only three stations are located within 30 km of Weiyuan County, which is in the center of our study region (Figure 1a). CENC uses these and more distant stations for constructing the routinely reported catalog in the Sichuan Basin. During the operation period of the temporary seismic network, 1,217 earthquakes were reported with a magnitude range of 0.1–3.3 in the study area falling within Latitude [29.35°, 29.85°] and Longitude [104.5°, 105°] (Figure 1).

3. Construction of the Event Catalog

We applied a sequential workflow to process to one year’s continuous seismic data to demonstrate how machine learning methods can expand the catalog in this region, even without a priori information from the standard catalog.

3.1 Events Detection

We use a 30-s time window to divide the continuous data into segments for input to PhaseNet. Neighboring segments overlap by 50% to enhance the completeness of picking by avoiding edge effects. To PhaseNet, we input three-component seismic waveforms of each segment; PhaseNet classifies each sample point with a probability that it is either a P wave, S wave, or noise; the three probabilities sum to unity. The noise is defined as the complement to P and S waves in the data (Zhu WQ and Beroza, 2019). Peaks in the P-wave and S-wave probability distributions corres-

pond to the P and S arrival times (Figure 3). Conservatively, we set a probability of 0.5 as the acceptance threshold for both P and S picks (e.g., Figure 4). Smaller values might lead, especially in noisy data, to a larger number of false or bad picks (e.g., Liu M et al., 2019; Wang RJ et al., 2020).

The detected phases consist of 4,860,125 P-wave arrival times and 3,567,610 S-wave arrival times (Figure 2b). The number of detections at each station correlates with the station’s uptime (Figures 2a–b). Complete P and S picking from one month’s continuous data from a single station, performed on a MacBook laptop computer (3.1 GHz Dual-core Intel Core i5 processor), takes approximately 2 CPU hours.

3.2 Events Association

Next, we use the Rapid Earthquake Association and Location (REAL) package (Zhang M et al., 2019) to create events out of the raw-pick data stream. REAL scans the input picks in time order, searching for associations of both P- and S-phases to form events on a 3D grid. Here we use a grid spacing of 2 km. The algorithm projects arrival times from each station onto the grid and searches for the grid point with the largest number of coincident arrivals. If several grid point have the same (maximum) number of picks, we select the one with the smallest root-mean-square misfit (Zhang M et al., 2019). The optimal stacking point of arrival times corresponds to the initial hypocenter. Associated arrivals are declared in an event and thus removed, and the process steps forward in time with the oldest remaining arrival time.

In the original implementation, REAL associated arrival picks to a particular earthquake by counting the number of picks within the

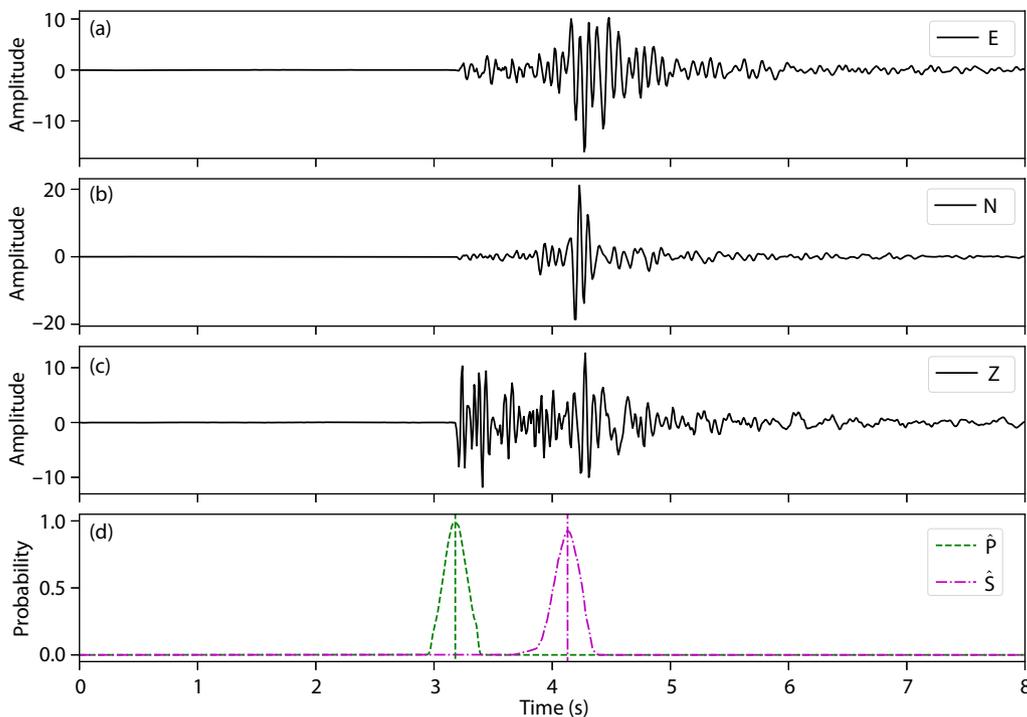


Figure 3. An example of the output of PhaseNet for a M_L 1.0 event from the routine catalog. The upper panels (a–c) are the “ENZ” components of seismograms. The panel (d) presents the probability distributions for P waves (P) and S waves (S). The estimated arrival time is at the peak probability.

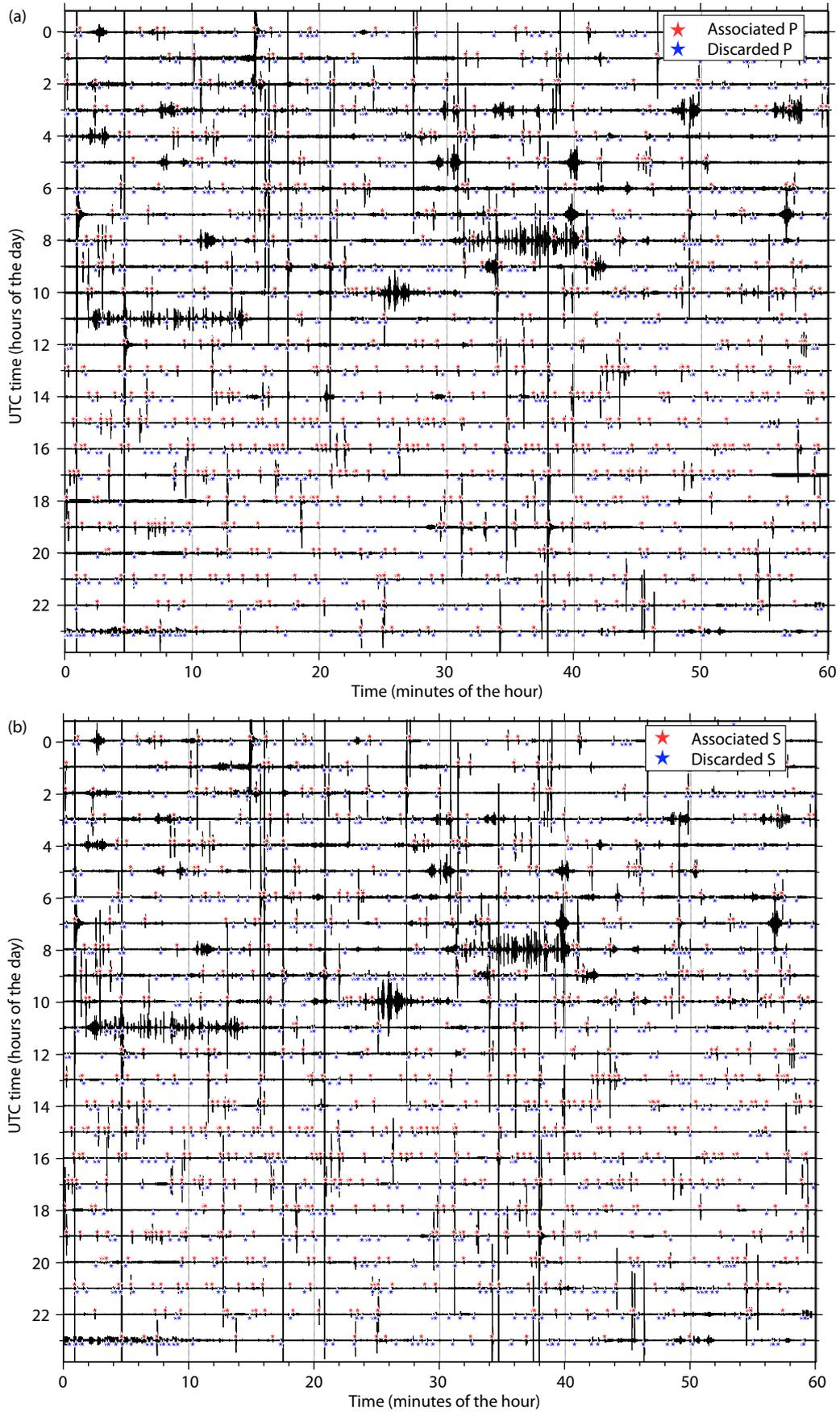


Figure 4. An example of PhaseNet picks on 1-day-long data from station s22 on 2016-03-07. (a) Associated (red) and discarded (blue) P picks on vertical component. (b) Associated (red) and discarded (blue) S picks on N component.

theoretical travel-time windows. Occasionally, when events overlapped in space and time, a pick could be associated to multiple events, which is obviously problematic. To make the result more robust against a large event splitting into multiple small events with small residuals, we use a modified REAL that assigns individual picks to the event with the greatest number of picks (Tan YJ et al., 2021). Consequently, this version of REAL keeps only the most reliable event within a time window (e.g., 1 s around P waves and 1.6 s around S waves), thus potentially missing events that occur closely in space and time.

We precalculated the theoretical P and S travel-time tables using the TauP Toolkit (Crotwell et al., 1999) and a 1-D velocity model (Lei XL et al., 2017). The search intervals for the event location are 0.02° over a 0.2° by 0.2° (in latitude and longitude) horizontal range and 2 km for a depth range above 20 km, respectively. The grid is centered at the station that recorded the earliest P phase. We set the association threshold to 5 P picks and a total of 10 P or S picks. So at least five stations are needed to declare an event. Events occurring close to the boundary of the study area or with a station azimuthal gap larger than 250° are deleted in the association procedure. We also remove S picks that occur less than 0.5 seconds after P, as described by Zhang M et al. (2019).

To evaluate the rate of false positives, we visually checked the associated 703 P and 710 S phase picks on station s22 for one day's data (data from 2016-03-07; Figure 4), finding that PhaseNet picked phase arrivals times that show subsample-level agreement with manual picks, with average differences of 0.002 and 0.007 s for P and S phases, respectively (Figure 5a–b). Moreover, most of the associated phases show picking probabilities over 0.7 (Figure 5c), which validates setting a probability of 0.5 as the acceptance threshold for both P and S picks, as our approach to dis-

carding bad picks on noisy data.

The REAL catalog contains 79,395 events from 12 November 2015 through 18 November 2016 (Figure 6). 25% of the PhaseNet picks are successfully associated, consisting of 981,284 P-wave arrival times and 859,576 S-wave arrival times (Figures 2c and 5d). The computation time depends on the number of picks recorded by our network. For one day of data, it takes about 10 minutes on a MacBook laptop computer (3.1 GHz Dual-core Intel Core i5 processor) to finish associating ten thousand picks.

3.3 Local Magnitude Estimation

For each event, we measure the maximum amplitudes of horizontal components after deconvolving the instrument response from the raw waveforms and then convolving with the theoretical Wood-Anderson seismometer response (Zhang M et al., 2019). The measured waveform window starts 0.5 sec before P arrivals and is twice the S-P interval in length (e.g., Figure 7a). If an arrival time is missing, either P or S, we use the predicted travel times. We compute source-to-station distances using the locations from REAL, and the attenuation table of Richter (1958) to determine the local magnitude (M_L). Then we average the measurements at individual stations to estimate the magnitude of each event.

The estimated magnitudes range from $M_L -0.5$ to 3.5. Just two earthquakes were larger than $M_L 3$: an $M_L 3.3$ earthquake on 2016-01-07, and an $M_L 3.5$ on 2016-07-26. Moreover, we compare the event waveforms for the $M_L > 2.8$ earthquakes, which were all reported by CENC but with generally lower magnitudes. The waveforms and peak amplitudes of $M_L > 2.8$ earthquakes, with similar epicentral distances between 14 and 16 km recorded on stations s22, suggest that their magnitudes are reliably estimated. Location uncertainties are significantly reduced; the RMS (root mean

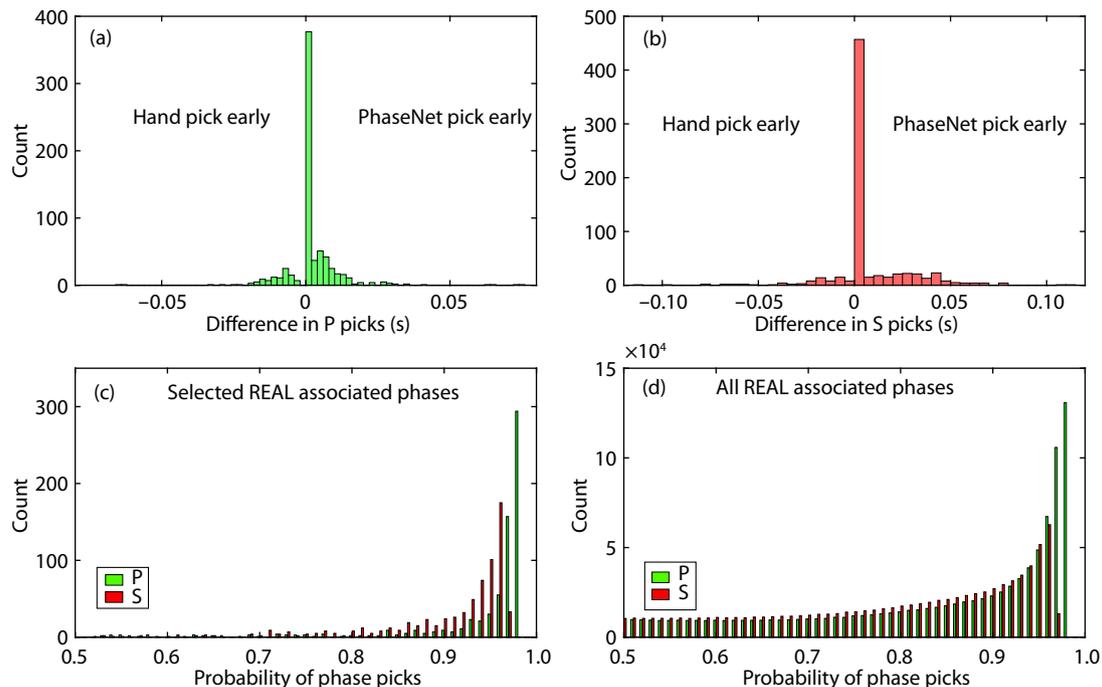


Figure 5. Statistics of associated PhaseNet picks. (a) and (b) compare the hand picks to the PhaseNet picks of P and S, as demonstrated in Figure 4. (c) presents probabilities of picks from PhaseNet shown in (a) and (b). (d) shows probabilities of picks from PhaseNet shown in Figure 2c.

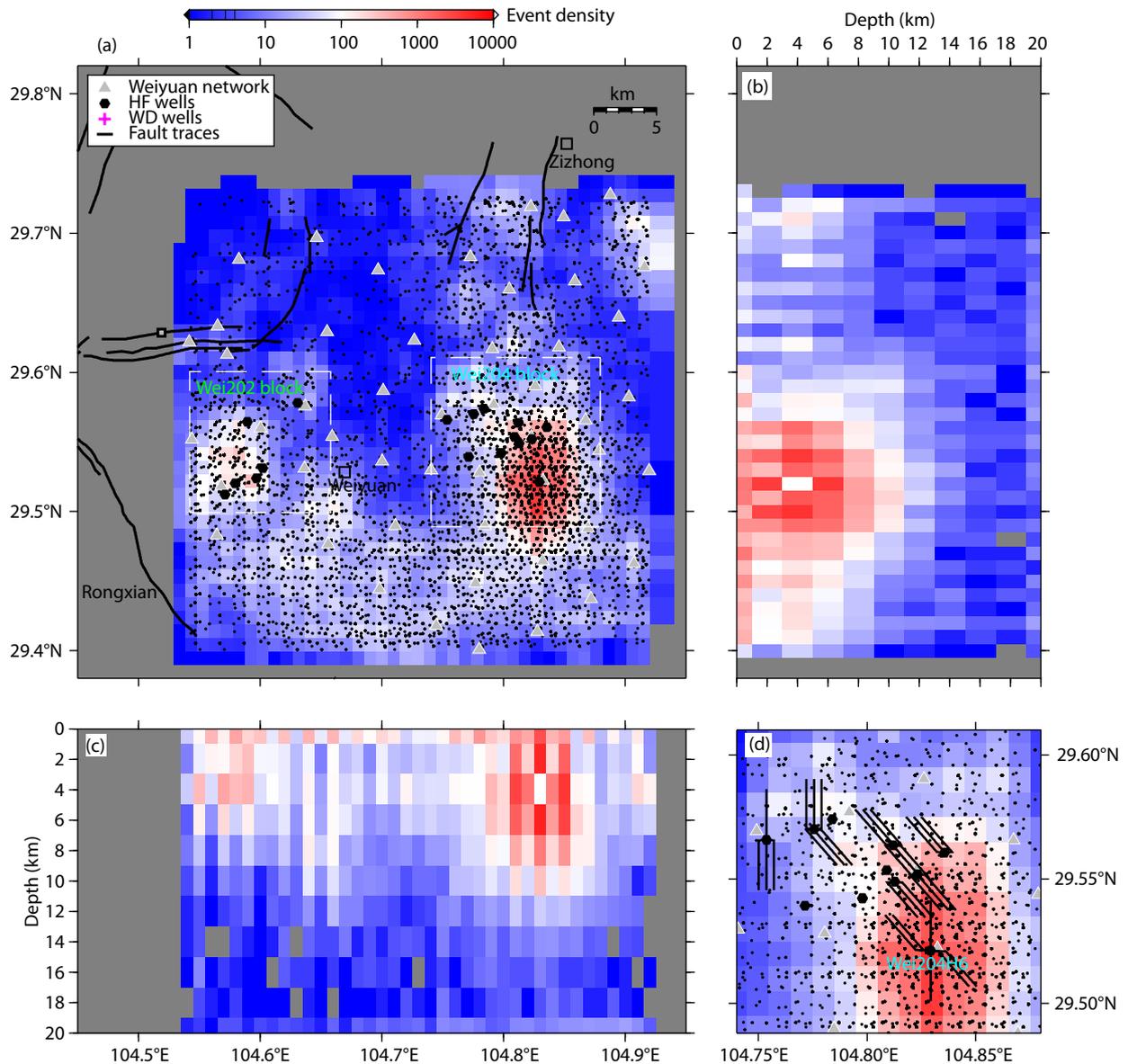


Figure 6. Distribution of the REAL catalog events. (a) The background colors indicate density of events (numbers in 0.01×0.01 degree area); black dots indicate event locations. (b) Latitude-depth view. (c) Longitude-depth view. (d) A zoomed map view for Wei204 block in (a); black hexagons mark the locations of HF well pads; black lines connecting them show the traces of horizontal well branches. However, regulatory data are lacking.

square) residuals fall to below 0.3 s, which is an improvement of almost one order of magnitude (Figure 7).

In addition to ordinary earthquakes, we detected 7,979 explosions ($M_L < 0.5$) in this part of our study region (below latitude 24.49°) that occurred exclusively between 2015-12-22 and 2016-01-20 (Figures 8a). They are primarily nighttime (local time 8:00 pm to 8:00 am) occurrences (Figure 8c), and are confirmed by reports from a 3D seismic survey conducted in the southern part of our study region starting on December 21, 2015 (INFOPETRO, 2016).

After removing explosion events, we retain 71,416 events in the magnitude range of -0.5 to 3.5 . Multiple event-rate peaks can be observed in 2016-01, 2016-03, 2016-04, and 2016-05. The events in the REAL catalog show temporal clustering behavior for the en-

tire period except for September 2016 (Figure 9a). Overall, by using the temporary array we are able to locate about 60 times more events than are in the CENC catalog (Figure 9b). Note that the seismicity rate trends of the two catalogs generally agree well with each other except after October 2016, when the CENC seismicity rate climbs to its peak while our REAL catalog seismicity rate flattens. The numbers of our located events show a three-fold increase than CENC catalog after October 2016, but the removal of 41 of the 50 stations of our temporary network (Figure 2a) results in decreased event locatability; thus the REAL catalog seismicity rate drops abnormally at the very end of our observation period (Figure 9b).

Maps of the REAL catalog epicenters, with estimated magnitudes unevenly distributed in both time and space, can be seen in

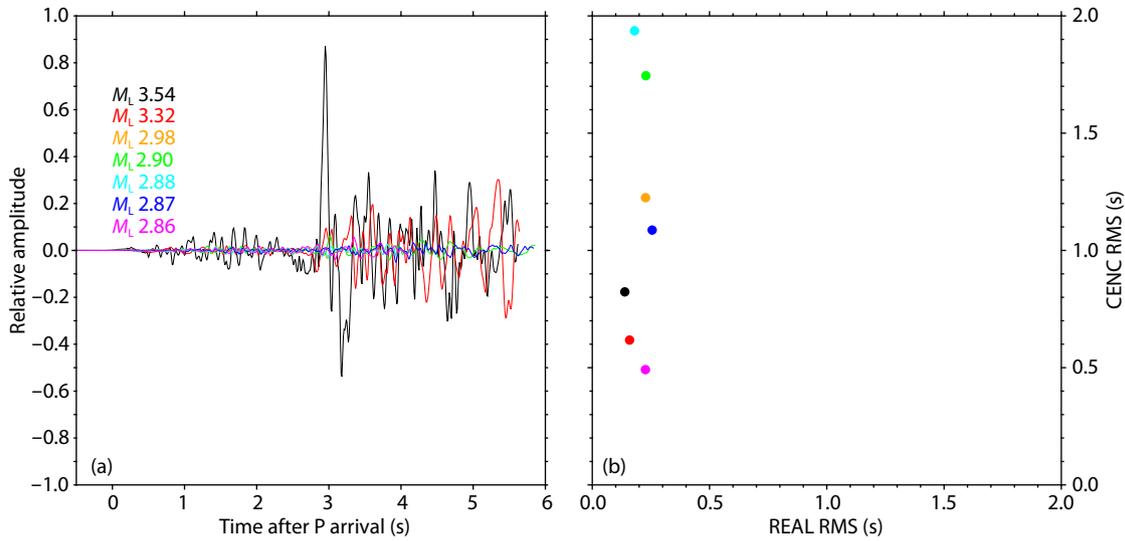


Figure 7. Comparisons of earthquakes with magnitude over M_L 2.8. (a) Event waveforms on stations with similar distances between 14 and 16 km. (b) RMS travel time residuals of all phase picks for each earthquakes shown in (a).

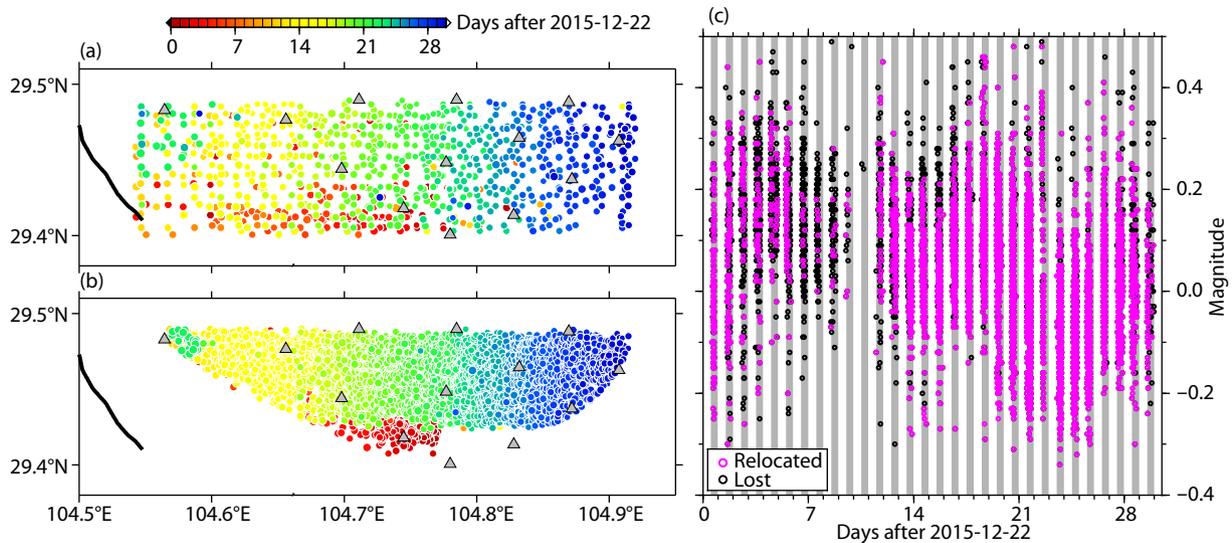


Figure 8. Maps of explosions and their occurrence times. (a) Explosions in REAL catalog. (b) VELEST relocated explosions. (c) Magnitude-time plot of explosions located by REAL (black and magenta circles) and VELEST (magenta circles), respectively. The explosions occurred almost exclusively in the night time, as the shaded areas show.

Figures 9 and 10. The smaller events are scattered across our study region (Figures 10a–b). Note that most of the larger magnitude earthquakes are located close to HF wells (Figures 10c–d). Moreover, earthquakes with M_L over 2.5 are concentrated near the wells in the Wei202 block (the west region in Figure 6d). In contrast, the Wei204 block features very dense concentrations of smaller events (east region in Figures 6d and 10a–b).

3.4 1-D Velocity Model Refinement

Here we present a modified 1-D velocity model of the shallow structure by using VELEST (Kissling et al., 1994) and our picked P and S arrivals. In VELEST, the non-linear ‘coupled hypocenter-velocity model problem’ is linearized and iteratively solved by inversion of the damped least-square matrix of traveltime partial derivatives (Kissling et al., 1994). We solve for the S- and the P-wave ve-

locity models independently by using VELEST since many P- and S-wave arrival times of high quality were picked by PhaseNet (Zhu WQ and Beroza, 2019).

A starting model was derived from combinations of several documented velocity models with different layered interfaces that were potentially applicable to our study. Zhao Z et al. (1997) used body wave travel times to develop a 1-D P/S model to approximate the average crustal structure in the Longmenshan Fault Zone. Lei XL et al. (2017) adopted a fine-layered 1-D velocity model to earthquake relocations near Changning, ~70 km south of our study region, based on seismic ambient noise tomography results from Wang X et al., (2013). Particularly for the shallow structures in the Weiyuan region, Meng XB et al. (2018) derived a 1-D P-velocity model from borehole well logging data, and Zeng Q et al.

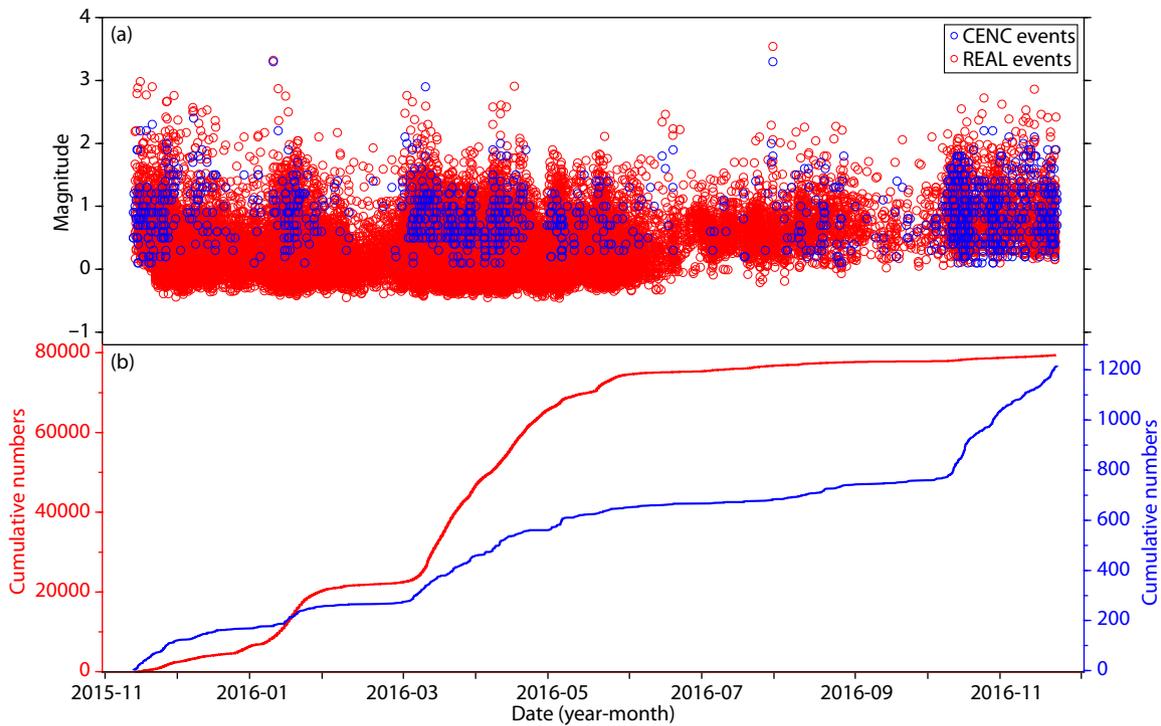


Figure 9. Number of events in different catalogs with time. (a) Magnitude-time plot of the CENC routine catalog (blue) and REAL catalog (red). (b) Cumulative count of events in the CENC and REAL catalogs.

(2020) inverted the S velocities from ambient noise tomography results. In running VELEST, we modified the model by Zhao Z et al. (1997) by adding several shallow layers above 2 km to the initial model (Figure 8); previous studies had suggested the possibility of significantly lower shallow velocities.

To utilize VELEST in updating the velocity model, we selected 1,600 events with 400 explosions in the south and 1200 earthquakes that are well-distributed over the Wei202 block, the Wei204 block, and the northern region near Zizhong County. The selected events are the best-recorded ones in each subregion, each with at least 20 associated picks whose phase probabilities were greater than 0.7. Starting with the 1D velocity models modified from Lei XL et al. (2017) and Zeng Q et al. (2020), we estimate optimal 1D P- and S-velocity models by averaging over the outputs of 20 experiments on the selected sub-datasets to account for lateral heterogeneities and the trade-off between epicentral locations and station effects. Damping factors for the hypocentral parameters and velocity parameters were selected by optimizing the data misfit reduction and the parameter resolution. Here we heavily damped the station terms. Low-velocity layers have been avoided so as not to introduce instabilities in the inversion process. Convergence to a stable solution is obtained after 30 iterations.

Despite different initial models, the resulting optimal 1D P- and S-velocity models turn out to be very similar both for the layered depths and the velocities, showing minor dependences on initial velocities and numbers of layered interfaces. These confirm both the precision and accuracy of our locations and demonstrate that the final model we derive is robust.

The best-estimated model, outperforming the other ones in relocating explosions back to the shallow surface (Figure 9a), is defined as the one initiated from Lei XL et al. (2017). The optimal model, with much slower velocities in the shallow part and varying V_p/V_s ratios for different layered depths (solid red curves in Figure 11a), is adopted in the absolute relocations of the REAL catalog events in the following part.

3.5 Absolute Relocation

Subsequently, we relocate all of the 79,395 REAL catalog events (Figure 12a) by using VELEST and the updated 1-D velocity model (Figure 11). The well-constrained events are retained with $< 200^\circ$ station gap and < 0.6 -s travel time residual, resulting in 74,921 relocated events (Figures 12b), showing a success rate of 94%. More specifically, 5,692 out of 7,979 (71%) explosions and 69,229 out of 71,416 (97%) non-explosion events are relocated by VELEST. The explosions and other events that are lost were mostly ones that occurred near the margins of our seismic network (Figures 8 and 12).

The VELEST locations show increased spatial clustering (Figure 12b). It is noteworthy that the depths of the explosions in the south are relocated to be very shallow, with a concentration above 1 km (Figure 13b). The mean RMS travel time residuals for those explosions drop from 0.207 s to 0.121 s (Figure 13c). As a result, the sweeping behaviors of explosions are much more obvious in VELEST locations (Figure 8b). After removing the relocated explosions, the VELEST catalog's depths peak at around 3 to 4 km, with a much better depth resolution than the REAL locations (Figures 12 and 13a–b). Moreover, the mean RMS travel time residuals drop from 0.194 s to 0.145 s (Figure 13d).

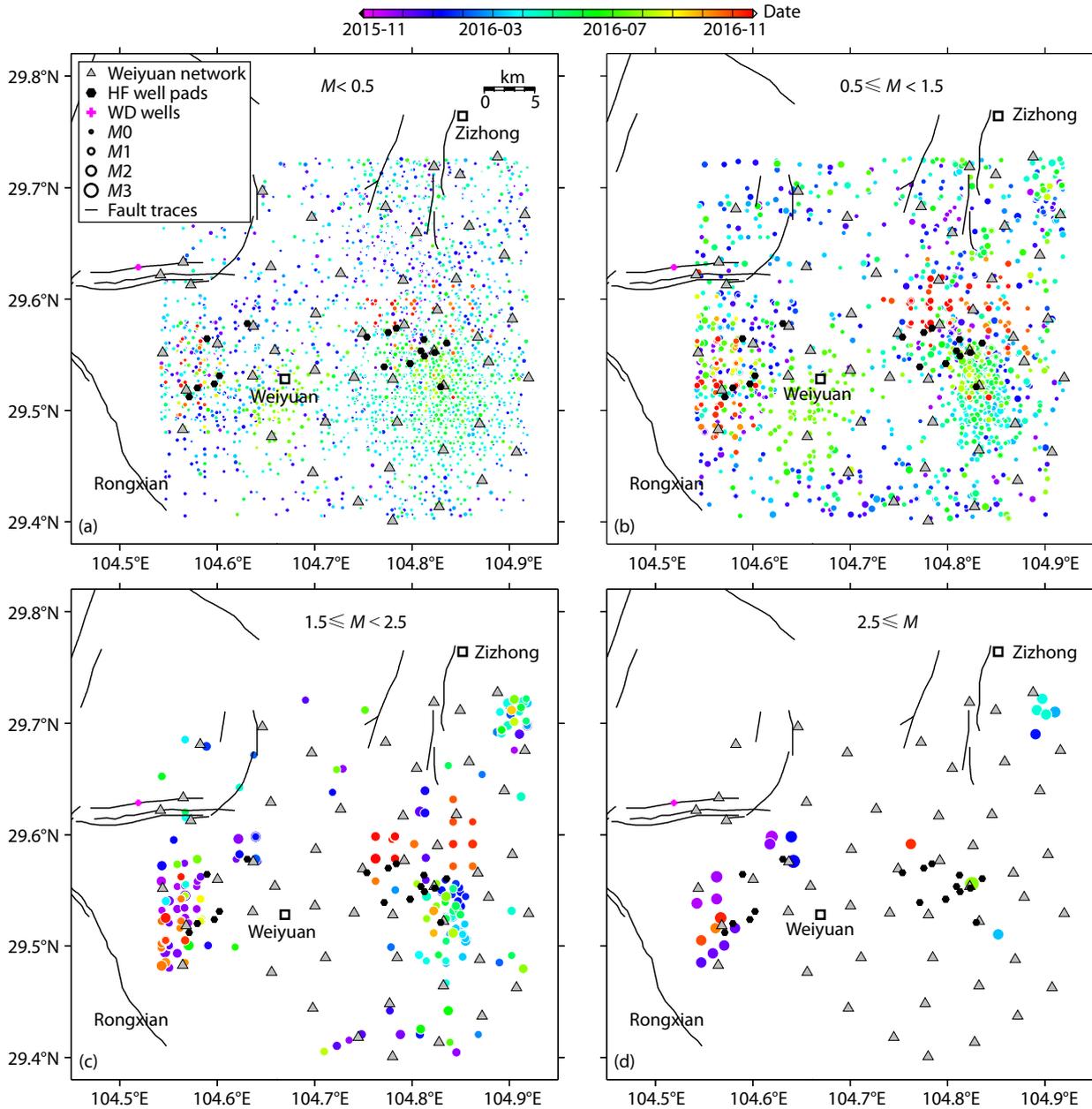


Figure 10. REAL catalog events split into different magnitude ranges. (a)–(d) are for events falling magnitude ranges: <0.5 , $0.5 \leq M < 1.5$, $1.5 \leq M < 2.5$ and >2.5 . The black hexagons indicate HF well pads.

4. Catalog Characterization

Among the obtained 79,395 detections, we identify 7,979 explosions and locate 71,416 non-explosion events; the smallest magnitude is -0.5 (Figures 9 and 12). With an updated 1-D velocity model, we are able to relocate 69,229 out of 71,416 events, which are not evenly distributed but concentrate in a few clusters (Figures 4a and 10). The densest clusters appear in the east of the study region, close to the wells of the Wei204 block, with the second densest cluster near the Wei202 block to its west.

4.1 Explosions

We relocated 5,692 out of 7,979 detected explosions ($M_L < 0.5$) that occurred exclusively between 2015-12-22 and 2016-01-20 (Figure 8a). The explosions serve as a benchmark for the perform-

ances of our earthquake timing, association, and locations (Figure 8). They occur with a semi-diurnal pattern of occurrence, almost exclusively at night (Figure 8c). Their magnitudes range from $M_L -0.4$ to 0.5 , and most have focal depths at 0 km. During the time of the 3D survey, we located an average of 266 events/day; the peak was 461, recorded on 2016-01-15 (Figure 8c). The temporal progression of these small events sweeps systematically from west to east (Figures 8a–b), showing typical migration characteristics of an active source seismic survey. Conveniently, the explosions facilitate the validation of our velocity model (Figure 11). Explosions near the southwestern and southeastern margins of our network (Figures 8a–b), largely occurring at the end of December 2015 and the end of January 2016 (Figures 8c), were discarded due to our application of a more strict threshold of the azimuthal

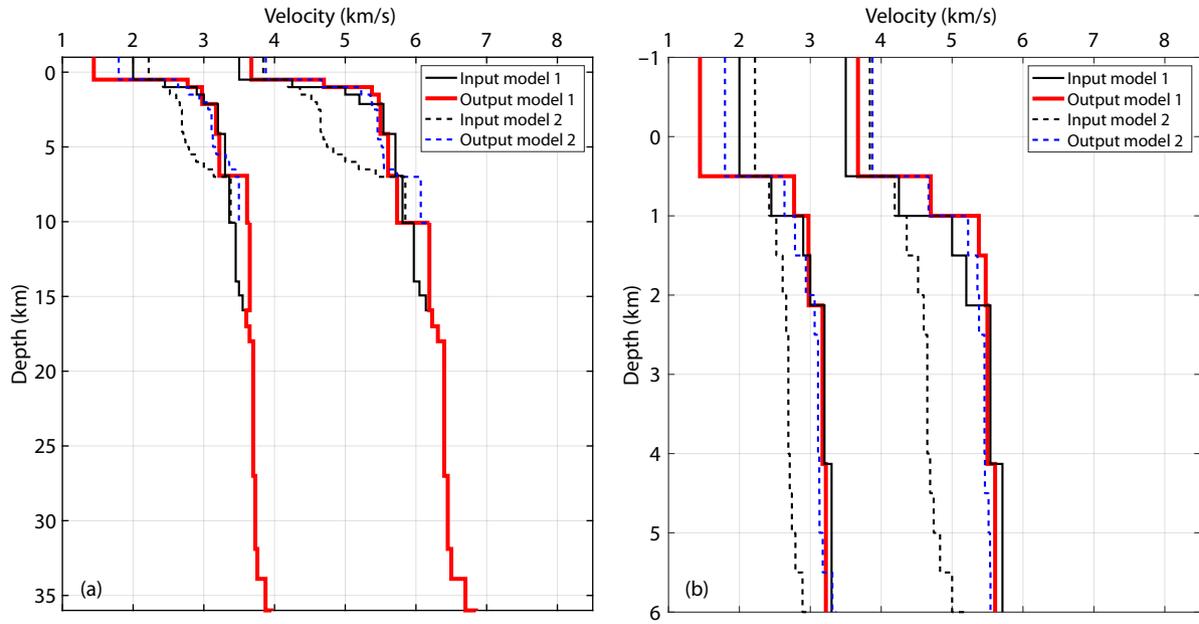


Figure 11. Output models using different initial models. Initial model 1 is modified from Lei XL et al. (2017). Initial model 2 is from Zeng Q et al. (2020). The solid red one is applied in the VELEST locations in this study. (a) Input and output velocity models. (b) Zoomed view of depths above 6 km for the velocity models in (a).

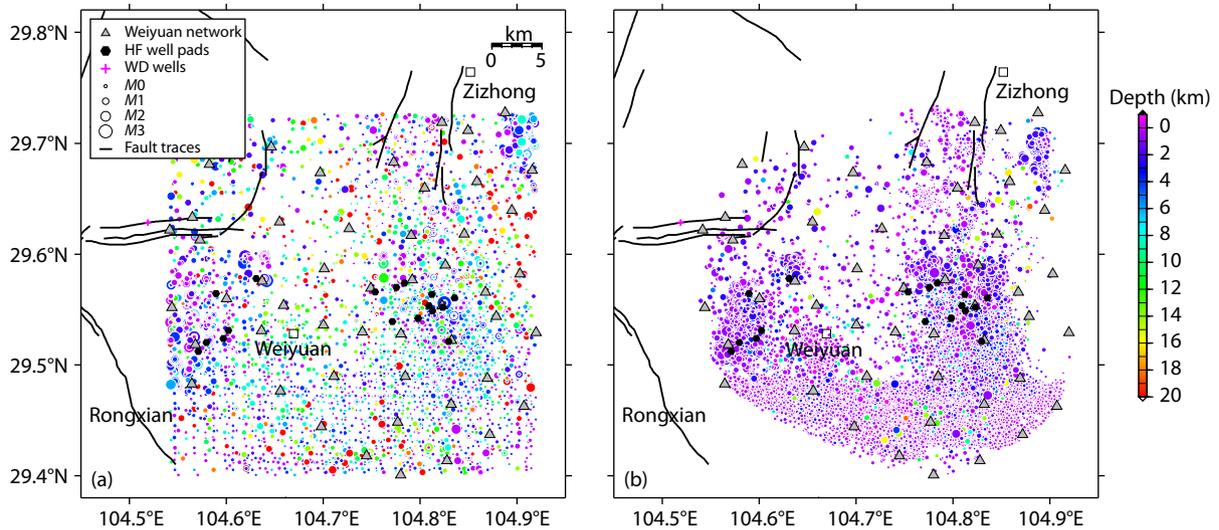


Figure 12. Comparisons of two different catalogs in terms of geographic locations. (a)–(b) show REAL and VELEST catalogs, respectively. The colored circle shows event locations scaled with depth.

gap in retaining the data for VELEST relocations.

4.2 Magnitudes of Earthquakes in Our Catalog

We double-checked all M_L 2.5 earthquakes in our REAL catalog, and found that they were all reported by CENC but with generally lower magnitudes, which could explain the greater number of M 3 class earthquakes in the REAL catalog compared to the CENC catalog (Figure 14). The $M_L > 2.8$ earthquakes are carefully validated with their event waveform inspections (Figure 7).

We compare our REAL catalog with those reported in the CENC catalog specified in our study region. Event pairs with their origin time differences smaller than 3 seconds are considered to be the

same events. Using that criterion, we have successfully detected 1129 out of 1217 CENC earthquakes and are able to relocate 1,067 of them (Figure 14a). For the same earthquakes, magnitudes in our REAL catalog are systematically larger than those in the CENC catalog by an average of 0.5 units (Figures 14b–c). Such a difference in magnitudes is likely due to different definitions, including our use of Richter’s original attenuation table. The mean RMS travel time residuals drop from 0.895 s to 0.199 s and 0.146 s for CENC, REAL, and VELEST relocations, respectively (Figure 15a). The offset between event pairs varies between kilometers to tens of kilometers (Figures 15b–c). We also note that some epicenters in the north from the CENC catalog appear in the south in our REAL catalog and became more tightly clustered (Figures 14a and

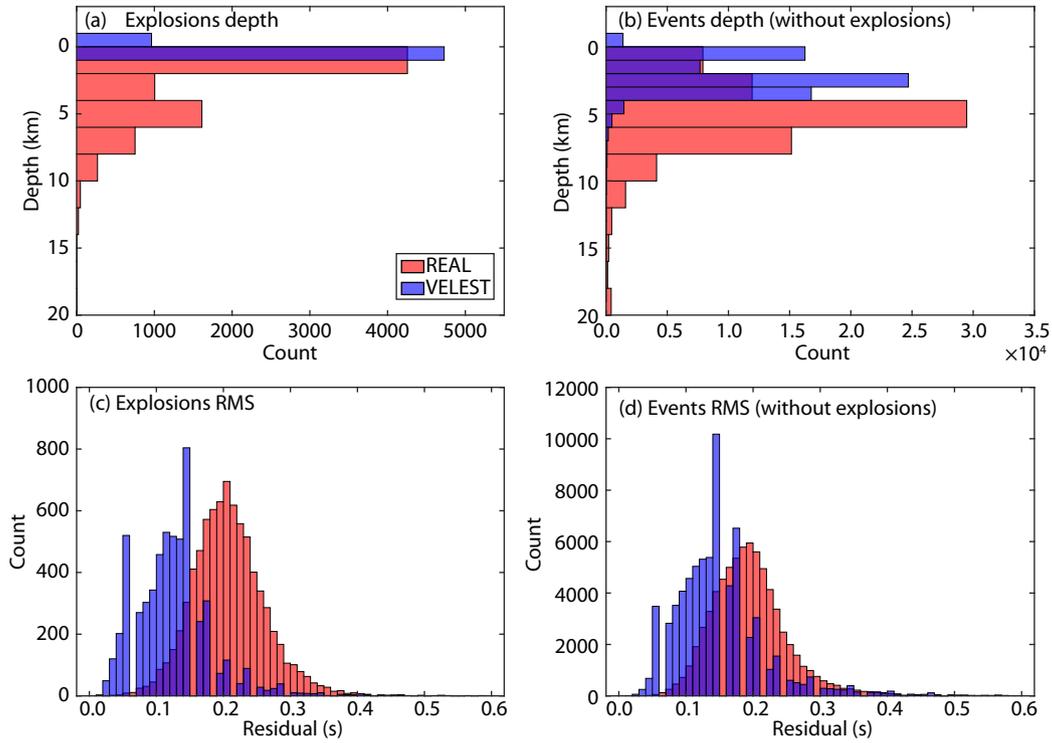


Figure 13. Comparisons of two different catalogs in terms of depth distributions and location uncertainties. (a)–(b) show depths of explosions and non-explosions, respectively. (c)–(d) show RMS travel time residuals of explosions and non-explosions, respectively.

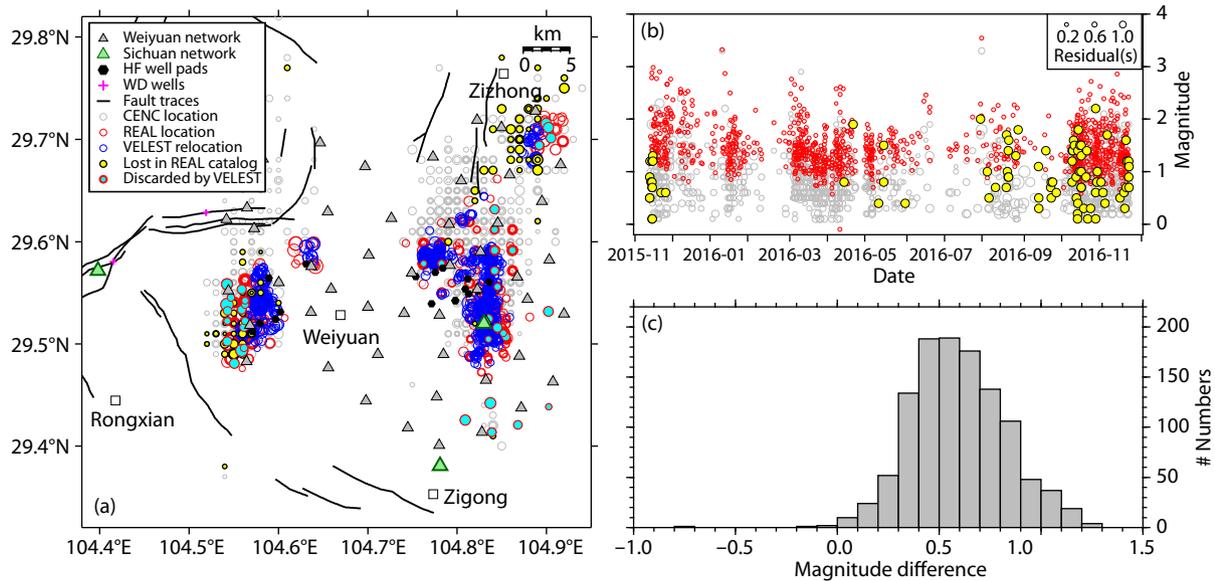


Figure 14. Comparison between event pairs in the REAL and the CENC catalogs. (a) Geographic distributions of seismic networks and event locations. The green triangles indicate locations of the sparsely distributed local network in Sichuan. The yellow dots show events that were reported by CENC but were missed by our association. The cyan dots are REAL associated locations (CENC events) but lost by VELEST relocation processes. (b) Magnitude-time plot for CENC (gray) and REAL (red) events, with their sizes scaled according to RMS residuals. The yellow dots are missed CENC events. (c) Magnitude differences between our estimated magnitudes and CENC routine reported ones for the same events. The magnitude bin is 0.1.

15b–c), which may also contribute to the reported magnitude difference.

Overall, 88 out of 1,217 (7%) of the CENC earthquakes were lost in

our REAL catalog (Figures 14a–b) due to too large station azimuthal gaps or insufficient numbers of picks. More specifically, out of the 88 missing CENC earthquakes, 16 occurred before 2015–12 when the temporary network was yet to be fully commis-

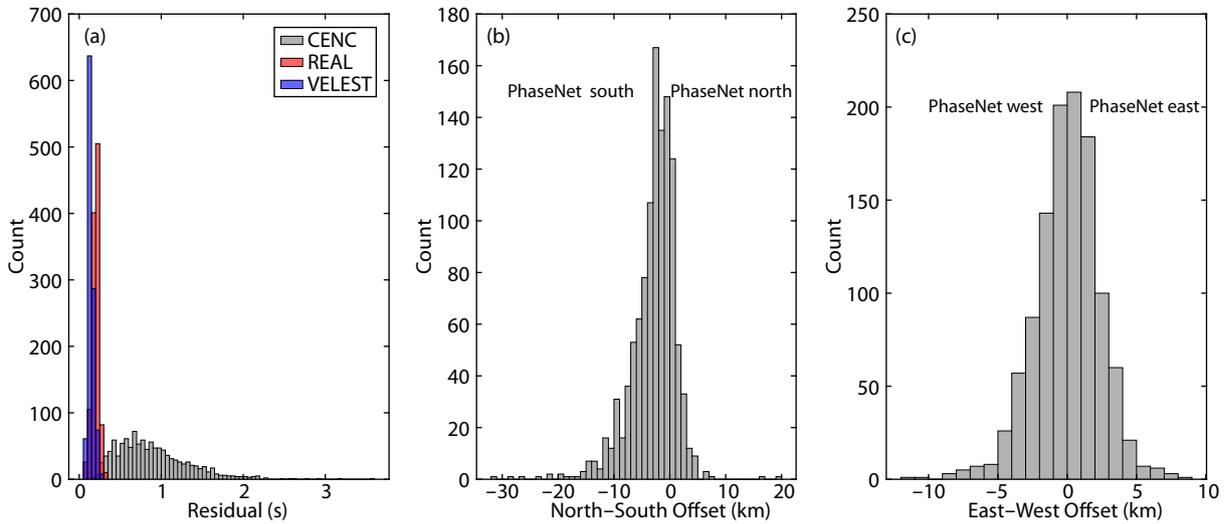


Figure 15. Comparisons of different locations of CENC earthquakes. (a) RMS travel time residuals. (b) North-south offsets of event pairs. (c) East-west offsets of event pairs.

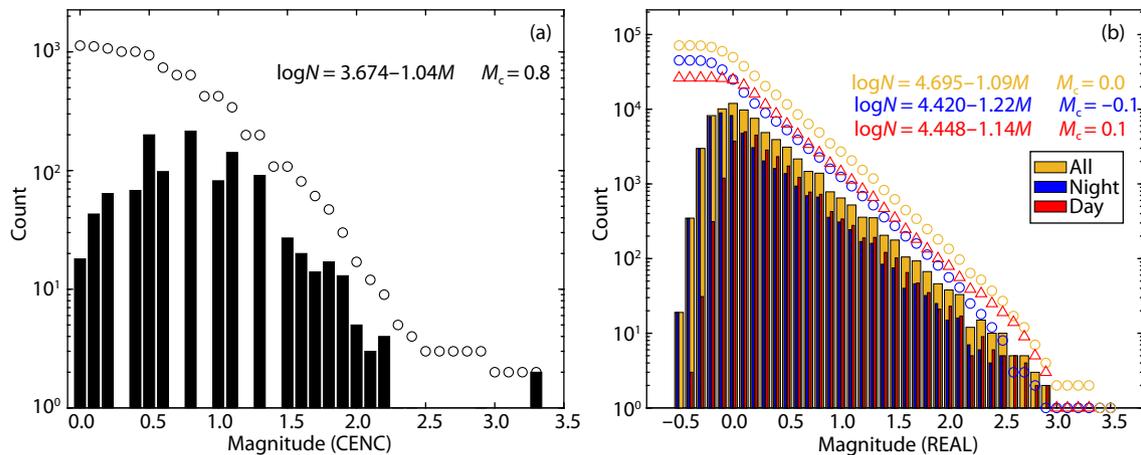


Figure 16. Magnitude size distribution and magnitude of completeness. (a) Estimated magnitude completeness (M_c) for the CENC and REAL catalogs, respectively.

sioned, and 66 occurred after 2016-08, when the temporary network was partly decommissioned (Figure 2a). The remaining six occurred exclusively after 2016-08 during which most temporary stations were removed (Figure 2a). Therefore, our detection success rate can be as high as 99.5% if we focus only on events between 2015-12 and 2016-07 when the temporary network functioned well. With a more strict threshold in retaining the data for VELEST relocations, another 62 earthquakes are lost, their locations distributed near the eastern and western margins of our seismic network (Figure 14a).

We used ZMAP to estimate the magnitude of completeness (M_c) for the REAL and CENC catalogs using the maximum curvature technique for the cumulative frequency–magnitude (Figure 16) (Wiemer and Wyss, 2000; Wiemer, 2001; Mignan and Woessner, 2012; Mignan et al., 2018). After removing explosions, we retain 71,416 REAL events in the magnitude range -0.5 to 3.5 , showing uneven distributions of events between day and night (Figures 17a–b). The M_c value is estimated to be 0, and the b -value is 1.09. In comparison, the M_c value in the CENC catalog is 0.8, and the b -

value is 1.04 (Figure 12a). Although our magnitudes are larger than those in the CENC catalog (Figure 14c), we have a much smaller M_c value (Figure 12b). Such improvement is attributed to our unique privilege of using a temporary network, which was able to detect and locate many more earthquakes than are in the CENC catalog (Figures 12).

To estimate the temporal pattern of M_c , we used a moving window approach and selected the sample window as 500 events overlapping with 125 events. The magnitude bin is 0.1. We compute M_c and the b -value in each window and assign the corresponding time as the middle of the window (Figure 17c–d). The M_c value remained relatively stable around 0 until June of 2016 (Figure 17c), by which time many stations in the north had been removed (see Figure 2). After then the M_c value increased to be around 0.6 (Figure 17c). The b -value ranged from ~ 1.0 during quiet times to > 1.5 during active times (Figures 17d).

4.3 Spatiotemporal Patterns of the Catalog

The REAL and VELEST catalogs exhibit improved spatiotemporal

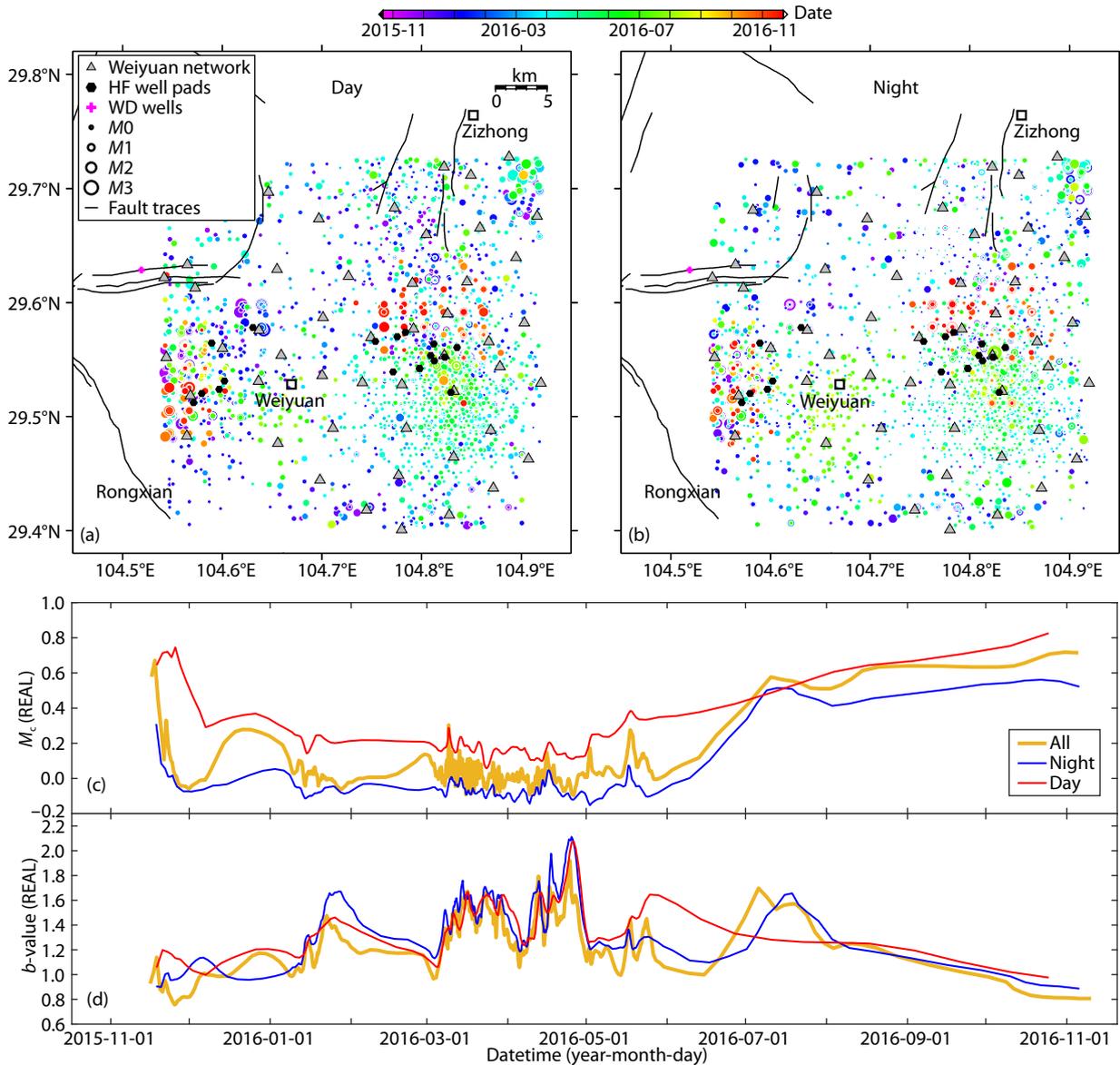


Figure 17. REAL catalog events during day and night. The events (dots) are color-scaled by the occurrence time and scaled with their magnitudes. The black hexagons indicate HF well pads. (a) Events occurring in day time (Local time 8:00 am to 20:00 pm). (b) Events occurring in night time (Local time 20:00 pm to 8:00 am). (c) M_c variations with time. (d) b -value variations with time.

resolution compared with the CENC catalog due chiefly to the increased number of events and better locations (Figures 14 and 15). Most events in our catalog, however, do not locate on mapped faults except for a few in the northeast part of the study region (Figures 12).

To better understand the characteristics of the REAL seismicity, we split the events into different magnitude ranges (Figure 10). The 67,774 events with magnitudes below M_L 0.5 are widely distributed over the study region (Figure 10a). As the magnitude range increases, the events become more localized in space. There are 10,996 events with magnitudes between M_L 0.5 and 1.5, most of which occurred after February 2016, concentrating near the HF wells and in the north part of our study region (Figure 10b). Clustering in the vicinity of the HF wells becomes more pronounced for events with magnitudes over M_L 1.5 (Figures 10c–d). Except

for sporadic events near mapped faults and in the south part of the study region, most of the larger earthquakes are located within ~5 km of the nearest HF wells (Figure 10). One exception is the cluster in the northeast near Zizhong City (Figures 10 and 12).

We detected and located about twice as many events at night time compared to daytime (Figures 17a–b). It is likely a result of variations of cultural noise levels and wind-generated noise (e.g., Wang RJ et al., 2020; Tan YJ et al., 2021). The difference can also be seen from the estimated M_c and b -values for the separate analysis of the two groups of diurnal seismicity (Figures 16b and 17c–d). The earthquakes at nighttime show an estimated M_c of -0.1 , smaller than the value of 0.1 at daytime (Figure 16b). And, the b -values for earthquakes at daytime and night time are estimated to be 1.1 and 1.2 (Figure 16b), respectively. However, the temporal variations of M_c and b -values show more obvious differences during

both the relatively quiet and active periods (Figures 17c–d). Overall, the M_c and b -values of earthquakes in the daytime show more significant temporal variations.

The spatiotemporal clustering patterns are obvious for events with a magnitude below M_L 1.5 but less so for larger magnitude events (Figures 12). Overall, the larger magnitude events near the HF wells show relatively much weaker temporal clustering than do smaller events (Figure 18a–b). The VELEST catalog shows multiple distinct clusters by considering their occurrence times and distance to nearby HF wells (Figures 12b and 18a–b). With a depth concentration between 2 and 4 km, they are located well within the seismic network (Figure 12b); thus, their hypocenters are better constrained and discussed in detail.

The seismicity in Wei202 block consists of 4,749 total events. Most

seismicity in Wei202 block is densely located atop the horizontal branches of the six HF wells (Figure 18a). The seismicity in the northeastern part of the Wei202 block (Figure 18a), with the largest magnitude of M_L 3.3, shows close proximity to the upward wells of Wei202H3, suggesting a causal link. The magnitude–time plots show that the seismicity rate is initially high in 2015–11, the beginning of the Weyuan earthquake sequence (Figure 18c). Later on, the seismicity continues to increase, and we see a gently decreased level of seismic activity until 2016–04. The seismicity rate remains at low levels (~1 event per day) between 2016–05 and 2016–07. After 2016–08, clusters of seismicity emerge near the well toe of Wei202H6 and the well head of Wei202H4.

The seismicity in the Wei204 block has the largest population of events, a total of 60,276. The magnitude–time plot shows several stages of abrupt seismicity rate increase in early 2016–01, late

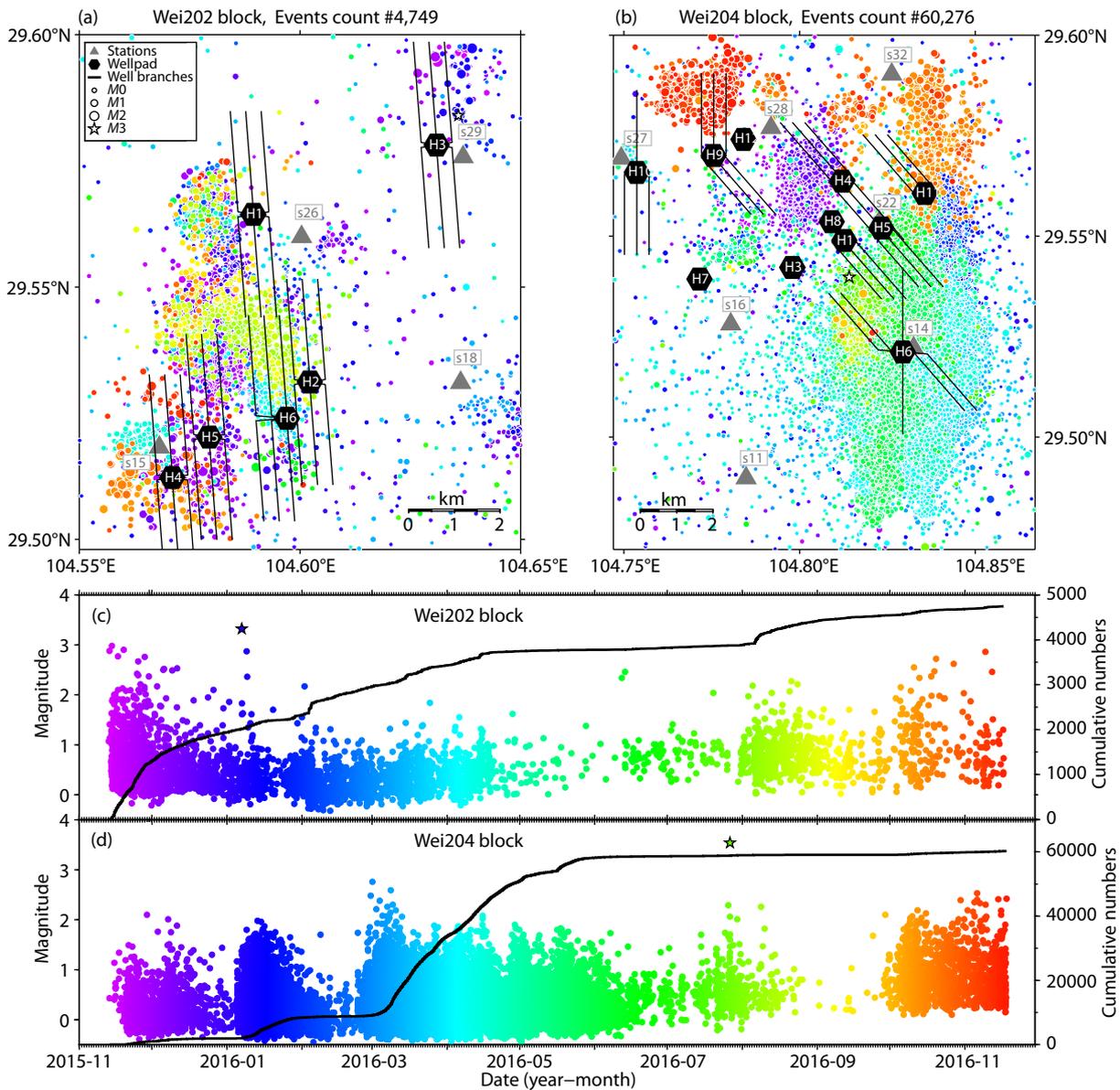


Figure 18. VELEST catalog seismicity in shale gas blocks. (a)–(b) Geographic maps of seismicity distributions for Wei202 and Wei204 blocks, respectively. The events are color-scaled by the occurrence time and scaled according to their magnitudes. The color scales can be seen in (c)–(d) which show magnitude–time plots and cumulative numbers of events for Wei202 and Wei204 blocks, respectively.

2016-02, and early 2016-10, with rapid decay of seismicity near the onset regions (Figure 18b). We found events densely located near the Wei204 block between 2016-03 and 2016-05 (Figure 18b). The b -values are larger than 1, and M_c is below 0 during this time interval (Figures 17c–d). Given the strong spatial correlation between populated seismicity and wells, it is highly likely that the groups of seismicity between 2016-03 and 2016-05 are linked to HF activities of different horizontal well branches from Wei204H6 (Figures 18b and 18d). The largest earthquake in our observed Weiyuan sequence is located near the well toes of the upward horizontal wells from Wei204H6. Additionally, we are able to identify three distinctly grouped events extending around upward wells of Wei204H4, Wei204H9 and Wei204H11. Each group of seismicity lasted only for several weeks and occurred separately in the very beginning and end of our study period (Figures 18b and 18d). The location, orientation, and timing of the seismicity suggest that the two groups of events were probably induced by HF stimulation of wells Wei204H4, Wei204H9 and Wei204H11, respectively.

5. Discussion and Conclusions

We have built a comprehensive earthquake catalog from raw continuous seismic data using a machine-learning phase picker (Zhu WQ and Beroza, 2019) and a sequential workflow of phase association and location algorithms (Zhang M et al., 2019; Liu M et al., 2020). In this study, we are able to report that the processing of 1-year datasets from 50 stations can be done in about 750 CPU hours. Moreover, parallel processing and GPU-based execution will be even faster, costing only minutes (Liu M et al., 2020).

Machine learning and seismology benefit from each other (Kong QK et al., 2019). An important ingredient for improving machine learning is the assembly of large labeled data sets. In seismology, we have such large labeled seismic data sets, e.g., waveforms with manually or automatically picked and manually reviewed arrival times (Mousavi et al., 2019a). Seismology is one of the best fields for developing and testing machine learning algorithms, which will eventually drive the development of machine learning. PhaseNet demonstrates high precision, particularly for S waves (Zhu WQ and Beroza, 2019). The database of accurate P and S arrival times allows us to improve both P- and S-wave velocity models and is especially useful for further polarization-based and S-wave-based seismic analysis (Zhu WQ and Beroza, 2019).

In the phase picking step, the sensitive detection performance of PhaseNet can facilitate REAL in successful association and initial location of weak events. PhaseNet also detects explosions with clear P or S arrivals. The local velocity model we used (Lei XL et al., 2017) was initially developed for the Changning region, which is within ~100 km from our study region. Consequently, the velocities of the deeper depths are likely suitable for our study region, but the shallow velocities might not be very accurate, as can be seen from our velocity model validations (Figure 11). Using such a reasonably good velocity model, the depths of more than half of the 7,979 explosions are recovered to be 0 km (Figure 13a). Suspected shots are typically treated as a nuisance during construction of seismic catalogs; however, as we have shown, they can be quite valuable for the validation of location accuracy since they

provide us the information of their depths from the surface and benefit our velocity model calibrations (Figures 11 and 13). In this study, we knew of their existence (Figure 8) (INFOPETRO, 2016) and could remove them during post-processing. Therefore, explosions were of minimal concern to the accuracy of sequential workflow during the phase association and relocation procedures (e.g., Park et al., 2020). Some events that occur closely in space and time might be missing in this type of analysis, since REAL keeps only the most reliable one within a time window (Zhang M et al., 2019).

Compared to continuous-waveform-based earthquake detection methods (e.g., Zhang M and Wen LX, 2015; Kao H and Shan SJ, 2007), REAL reduces the computational load by using discrete time phase picks. REAL possesses the main advantages of both pick-based and waveform-based detection and location methods (Zhang M et al., 2019), enabling it to work successfully with sparse station coverage. We detected relatively fewer events after August 2016 near the Wei204 block, a bias that is likely to have been due to the removal of stations. Only 10 stations were in service after that time in this area (Figure 2). Station coverage in the east also became poor due to the removal of stations. Even so, we detected and located some event clusters in the north of the Wei204 block (Figure 12).

The phase data and initial hypocenters produced by this workflow serve as the initial step toward the construction of a high-quality and robust earthquake catalog, to be followed by precision relocation using a range of modern tools, e.g., hypoDD (Waldhauser and Ellsworth, 2002). Although the number of detections greatly increased over routine analysis, we likely missed events for a variety of reasons, including insufficient number of picks, events spaced too closely in space and time, or events with large station gaps and/or travel time residuals. Small numbers of picks (especially S picks) or large timing errors can result in large location errors since the depth-origin trade-off most strongly affects locations based on P waves (Zhu WQ and Beroza, 2019).

We suggest that much of the microseismicity in the earthquake sequence may be a result of HF stimulation in Wei202 and Wei204 blocks. The combination of sensitive detection and precise location of microseismicity, and a preliminary database of HF well information, has allowed us to distinguish groups of events induced by HF. The multiple well influences on seismicity pose significant challenges for seismic hazard mitigation, where different actions would be required for different wells.

In the CENC catalog, it is difficult to associate earthquakes with HF wells due to the small number of earthquakes and uncertain locations (Figures 14a). For the same earthquakes in the two catalogs, significant location improvements were obtained (Figures 14a–b and 15a). Our results improve the spatial resolution of the ~1,200 earthquakes reported in the CENC catalog (Figure 14a), showing potential to better separate the seismicity near HF wells from natural tectonic events (Figure 12b). Although we found many small earthquakes clustered near HF wells (Figures 18a–b), it is still challenging to understand the connection, if any, with shale gas development and production without a higher precision catalog. The presence of multiple closely-spaced HF wells poses significant challenges for seismic hazard mitigation since different ac-

tions might be required depending on the causal connection (Figures 18a–b). Thus, we are motivated to refine further the VELEST catalog by using relative relocation methods to unfold their detailed spatiotemporal patterns in our companion paper.

In sum, we have constructed a basic earthquake catalog for the Weiyuan earthquake sequence from 12 November 2015 to 18 November 2016 from raw data without manual review. We have detected and located about 60 times as many events as the routine CENC catalog. We have also used explosion and earthquake data to calibrate the local 1-D P and S velocity models. We were able to relocate 5,692 explosions and other 69,229 events. Our study demonstrates the feasibility of characterizing earthquake sequences, potentially induced seismicity, and artificial sources, using machine-learning detection and timing and a high-performance sequential earthquake monitoring workflow.

Acknowledgments

The authors are grateful to Editor Xiaowei Chen, reviewer Jake Walter, and another anonymous reviewer for their helpful comments, which greatly contributed to improving the article. This study was supported by the Hong Kong Research Grants Council (No. 14303721 and N_CUHK430/16), Faculty of Science, CUHK, National Natural Science Foundation of China (Grants No. 41804015, 41661164035), and National Key R&D Program of China (2018YFC1504501-02). This work was supported also by the Stanford Center for Induced and Triggered Seismicity. PhaseNet is available at <https://github.com/wayneweiqiang/PhaseNet>. REAL and VELEST packages can be accessed via <https://github.com/Dalmzhang/REAL>. The team lead by Risheng Chu (chur@whigg.ac.cn) collected and maintained the whole dataset (Zeng Q et al., 2020). The catalogs are available upon request.

References

- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.*, 68(5), 1521–1532. <https://doi.org/10.1785/BSSA0680051521>
- Atkinson, G. M., Eaton, D. W., Ghofrani, H., Walker, D., Cheadle, B., Schultz, R., Shcherbakov, R., Tiampo, K., Gu, J., ... Kao, H. (2016). Hydraulic fracturing and seismicity in the Western Canada Sedimentary Basin. *Seismol. Res. Lett.*, 87(3), 631–647. <https://doi.org/10.1785/0220150263>
- 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>
- Bergen, K. J., and Beroza, G. C. (2018). Detecting earthquakes over a seismic network using single-station similarity measures. *Geophys. J. Int.*, 213(3), 1984–1998. <https://doi.org/10.1093/gji/ggy100>
- 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>
- Crotwell, H. P., Owens, T. J., and Ritsema, J. (1999). The TauP toolkit: flexible seismic travel-time and ray-path utilities. *Seismol. Res. Lett.*, 70(2), 154–160. <https://doi.org/10.1785/gssrl.70.2.154>
- Dong, D. Z., Shi, Z. S., Guan, Q. Z., Jiang, S., Zhang, M. Q., Zhang, C. C., Wang, S. Y., Sun, S. S., Yu, R. Z., ... Wang, S. Q. (2018). Progress, challenges and prospects of shale gas exploration in the Wufeng-Longmaxi reservoirs in the Sichuan Basin. *Nat. Gas Ind. B*, 5(5), 415–424. <https://doi.org/10.1016/j.ngib.2018.04.011>
- Ellsworth, W. L. (2013). Injection-induced earthquakes. *Science*, 341(6142), 1225942. <https://doi.org/10.1126/science.1225942>
- Ester, M., Krieger, H. P., Sander, J., and Xu, X. W. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining* (pp. 226–231). Portland, Oregon: AAAI Press.
- Gibbons, S. J., and Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophys. J. Int.*, 165(1), 149–166. <https://doi.org/10.1111/j.1365-246X.2006.02865.x>
- 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>
- Grigoli, F., Scarabello, L., Böse, M., Weber, B., Wiemer, S., and Clinton, J. F. (2018). Pick- and waveform-based techniques for real-time detection of induced seismicity. *Geophys. J. Int.*, 213(2), 868–884. <https://doi.org/10.1093/gji/ggy019>
- Holland, A. A. (2013). Earthquakes triggered by hydraulic fracturing in south-central Oklahoma. *Bull. Seismol. Soc. Am.*, 103(3), 1784–1792. <https://doi.org/10.1785/0120120109>
- INFOPETRO. (2016). Sinopec Weiyuan-Rongxian 3D seismic exploration project completed. <http://www.cnpcjob.com/article/info.php?id=1889>
- Johnson, C. E., Lindh, A. G., and Hirshorn, B. F. (1997). Robust regional phase association. *U.S. Geol. Surv. Open-File Rept.* 94–621. <https://doi.org/doi:10.3133/ofr94621>
- Johnson, C. E. (2020). Beyond earthworm: keeping the promise. *Seismol. Res. Lett.*, 91(2A), 581–584. <https://doi.org/10.1785/0220190198>
- Kao, H., and Shan, S. J. (2007). Rapid identification of earthquake rupture plane using source-scanning algorithm. *Geophys. J. Int.*, 168(3), 1011–1020. <https://doi.org/10.1111/j.1365-246X.2006.03271.x>
- Kissling, E., Ellsworth, W. L., Eberhart-Phillips, D., and Kradolfer, U. (1994). Initial reference models in local earthquake tomography. *J. Geophys. Res.: Solid Earth*, 99(B10), 19635–19646. <https://doi.org/10.1029/93JB03138>
- 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>
- Langenbruch, C., Ellsworth, W. L., Woo, J. U., and Wald, D. J. (2020). Value at induced risk: injection-induced seismic risk from low-probability, high-impact events. *Geophys. Res. Lett.*, 47(2), e2019GL085878. <https://doi.org/10.1029/2019GL085878>
- Lee, K. K., Ellsworth, W. L., Giardini, D., Townend, J., Ge, S., Shimamoto, T., Yeo, I. W., Kang, T. S., Rhie, J., ... Langenbruch, C. (2019). Managing injection-induced seismic risks. *Science*, 364(6442), 730–732. <https://doi.org/10.1126/science.aax1878>
- Lee, W. H. K., and Stewart, S. W. (1981). *Principles and Applications of Microearthquake Networks*. New York: Academic Press.
- 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 $M_w 4.7$ induced by shale-gas hydraulic fracturing in Sichuan Basin, China. *Sci. Rep.*, 7(1), 7971. <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>
- 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. *Journal of Earth Science*, 30(2), 348–366. <https://doi.org/10.1007/s12583-017-0962-y>
- Liu, M., Zhang, M., Zhu, W. Q., Ellsworth, W. L., and Li, H. Y. (2020). Rapid characterization of the July 2019 Ridgecrest, California, earthquake sequence from raw seismic data using machine learning phase picker. *Geophys. Res. Lett.*, 47, e2019GL086189. <https://doi.org/10.1029/2019GL086189>
- Meng, L. Y., McGarr, A., Zhou, L. Q., and Zang, Y. (2019). An investigation of seismicity induced by hydraulic fracturing in the sichuan basin of China based on data from a temporary seismic network. *Bull. Seismol. Soc. Am.*, 109(1), 348–357. <https://doi.org/10.1785/0120180310>
- 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>
- Mignan, A., and Woessner, J. (2012). Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resour. Stat. Seism. Anal.*, 1–45. <https://doi.org/10.5078/corssa-00180805>
- Mignan, A., Danciu, L., and Giardini, D. (2018). Considering large earthquake clustering in seismic risk analysis. *Nat. Hazards*, 91(1), 149–172. <https://doi.org/10.1007/s11069-016-2549-9>
- Mousavi, S. M., Sheng, Y. X., Zhu, W. Q., and Beroza, G. C. (2019a). Stanford EArthquake dataset (STEAD): a global data set of seismic signals for AI. *IEEE Access*, 7, 179464–179476. <https://doi.org/10.1109/ACCESS.2019.2947848>
- Mousavi, S. M., Zhu, W. Q., Sheng, Y. X., and Beroza, G. C. (2019b). CRED: a deep residual network of convolutional and recurrent units for earthquake signal detection. *Sci. Rep.*, 9(1), 10267. <https://doi.org/10.1038/s41598-019-45748-1>
- Park, Y., Mousavi, S. M., Zhu, W. Q., Ellsworth, W. L., and Beroza, G. C. (2020). Machine-learning-based analysis of the guy-greenbrier, arkansas earthquakes: a tale of two sequences. *Geophys. Res. Lett.*, 47(6), e2020GL087032. <https://doi.org/10.1029/2020GL087032>
- Peng, Z. G., and Zhao, P. (2009). Migration of early aftershocks following the 2004 Parkfield earthquake. *Nat. Geosci.*, 2(12), 877–881. <https://doi.org/10.1038/ngeo0697>
- Richter, C. F. (1958). *Elementary Seismology*. San Francisco: W. H. Freeman.
- Ross, Z. E., Meier, M. A., and Hauksson, E. (2018). P wave arrival picking and first-motion polarity determination with deep learning. *J. Geophys. Res.:Solid Earth*, 123(6), 5120–5129. <https://doi.org/10.1029/2017JB015251>
- Ross, Z. E., Yue, Y. S., Meier, M. A., Hauksson, E., and Heaton, T. H. (2019). PhaseLink: a deep learning approach to seismic phase association. *J. Geophys. Res.:Solid Earth*, 124(1), 856–869. <https://doi.org/10.1029/2018JB016674>
- Rubin, A. M., Gillard, D., and Got, J. L. (1999). Streaks of microearthquakes along creeping faults. *Nature*, 400(6745), 635–641. <https://doi.org/10.1038/23196>
- Safety and Environmental Protection Technology Research Institute of CNPC, and Beijing CNPC Construction Project Labor Safety and Health Pre Evaluation Co., Ltd. (2015). Environmental impact statement of CNPC Changning, Weiyuan, and Zhaotong shale gas construction project (in Chinese). Beijing: Safety and Environmental Protection Technology Research Institute of CNPC, Beijing CNPC Construction Project Labor Safety and Health Pre Evaluation Co., Ltd.
- Schoenball, M., and Ellsworth, W. L. (2017). A systematic assessment of the spatiotemporal evolution of fault activation through induced seismicity in Oklahoma and Southern Kansas. *J. Geophys. Res.:Solid Earth*, 122(12), 10189–10206. <https://doi.org/10.1002/2017JB014850>
- Schoenball, M., Walsh, F. R., Weingarten, M., and Ellsworth, W. L. (2018). How faults wake up: the Guthrie-Langston, Oklahoma earthquakes. *Lead. Edge*, 37(2), 100–106. <https://doi.org/10.1190/le37020100.1>
- Schultz, R., Beroza, G., Ellsworth, W., and Baker, J. (2020). Risk-informed recommendations for managing hydraulic fracturing-induced seismicity via traffic light protocols. *Bull. Seismol. Soc. Am.*, 110(5), 2411–2422. <https://doi.org/10.1785/0120200016>
- Shelly, D. R., Ellsworth, W. L., and Hill, D. P. (2016). Fluid-faulting evolution in high definition: connecting fault structure and frequency-magnitude variations during the 2014 Long Valley Caldera, California, earthquake swarm. *J. Geophys. Res.:Solid Earth*, 121(3), 1776–1795. <https://doi.org/10.1002/2015JB012719>
- 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.:Solid Earth*, 125(10), e2020JB019932. <https://doi.org/10.1029/2020JB019932>
- Skoumal, R. J., Brudzinski, M. R., and Currie, B. S. (2015). Earthquakes induced by hydraulic fracturing in Poland Township, Ohio. *Bull. Seismol. Soc. Am.*, 105(1), 189–197. <https://doi.org/10.1785/0120140168>
- Stewart, S. W. (1977). Real-time detection and location of local seismic events in central California. *Bull. Seismol. Soc. Am.*, 67(2), 433–452. <https://doi.org/10.1785/BSSA0670020433>
- Tan, Y. J., Waldhauser, F., Ellsworth, W. L., Zhang, M., Zhu, W. Q., Michele, M., Chiaraluca, L., Beroza, G. C., and Segou, M. (2021). Machine-learning-based high-resolution earthquake catalog reveals how complex fault structures were activated during the 2016–2017 central Italy sequence. *Seismol. Res. Lett.*, 92(1), 11–19. <https://doi.org/10.1785/0320210001>
- Waldhauser, F., and Ellsworth, W. L. (2002). Fault structure and mechanics of the Hayward Fault, California, from double-difference earthquake locations. *J. Geophys. Res.:Solid Earth*, 107(B3), ESE3–1–ESE3–15. <https://doi.org/10.1029/2000JB000084>
- Walsh III, F. R., and Zoback, M. D. (2016). Probabilistic assessment of potential fault slip related to injection-induced earthquakes: application to north-central Oklahoma, USA. *Geology*, 44(12), 991–994. <https://doi.org/10.1130/G38275.1>
- Wang, J., Xiao, Z. W., Liu, C., Zhao, D. P., and Yao, Z. X. (2019). Deep learning for picking seismic arrival times. *J. Geophys. Res.:Solid Earth*, 124(7), 6612–6624. <https://doi.org/10.1029/2019JB017536>
- 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>
- Wang, X., Ma, S. L., Guo, Z., Lei, X. L., Xia, Y. J., Guo, X., Yu, G. Z., Gou, X. B., and Jiang, X. D. (2013). S-wave velocity of the crust in Three Gorges Reservoir and the adjacent region inverted from seismic ambient noise tomography. *Chin. J. Geophys.*, 56(12), 4113–4124. <https://doi.org/10.6038/cjg20131216>
- 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>
- Wiemer, S. (2001). A software package to analyze seismicity: ZMAP. *Seismol. Res. Lett.*, 72(3), 373–382. <https://doi.org/10.1785/gssrl.72.3.373>
- Wong, W. C. J., Zi, J. P., Yang, H. F., and Su, J. R. (2021). Spatial-temporal evolution of injection induced earthquakes in weiyuan area by machine-learning phase picker and waveform cross-correlation. *Earth Planet. Phys.*, 5(6), 485–500. <https://doi.org/10.26464/epp2021055>
- Wu, X., Xia, J., Guan, B. S., Yan, X. M., Zou, L., Liu, P., Yang, L. F., Hong, S., and Hu, S. (2019). Water availability assessment of shale gas production in the Weiyuan play, China. *Sustainability*, 11(3), 940. <https://doi.org/10.3390/su11030940>
- Xu, W. B., Dutta, R., and Jónsson, S. (2015). Identifying active faults by improving earthquake locations with InSAR data and Bayesian estimation: the 2004 Tabuk (Saudi Arabia) earthquake sequence. *Bull. Seismol. Soc. Am.*, 105(2A), 765–775. <https://doi.org/10.1785/0120140289>
- Yang, H. F., Zhu, L. P., and Chu, R. S. (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. (2015). Recent advances in imaging crustal fault zones: a review. *Earthq. Sci.*, 28(2), 151–162. <https://doi.org/10.1007/s11589-015-0114-3>
- 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 M_L 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>
- Yoon, C. E., O'Reilly, O., Bergen, K. J., and Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. *Sci. Adv.*, 1(11), e1501057. <https://doi.org/10.1126/sciadv.1501057>
- Yoon, C. E., Bergen, K. J., Rong, K. X., Elezabi, H., Ellsworth, W. L., Beroza, G. C., Bailis, P., and Levis, P. (2019). Unsupervised large-scale search for similar earthquake signals. *Bull. Seismol. Soc. Am.*, 109(4), 1451–1468. <https://doi.org/10.1785/0120190006>
- Yu, H., Harrington, R. M., Liu, Y., and Wang, B. (2019). Induced seismicity driven by fluid diffusion revealed by a near-field hydraulic stimulation monitoring array in the Montney Basin, British Columbia. *J. Geophys. Res.:Solid Earth*, 124(5), 4694–4709. <https://doi.org/10.1029/2018JB017039>
- Zeng, Q., Chu, R. S., Sheng, M. H., and Wei, Z. G. (2020). Seismic ambient noise

- tomography for shallow velocity structures beneath Weiyuan, Sichuan. *Chin. J. Geophys.*, 63(3), 944–955. <https://doi.org/10.6038/cjg2020N0177>
- Zhang, M., and Wen, L. X. (2015). An effective method for small event detection: match and locate (M&L). *Geophys. J. Int.*, 200(3), 1523–1537. <https://doi.org/10.1093/gji/ggu466>
- 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). Crustal structure and accurate hypocenter determination along the Longmenshan fault zone. *Acta Seismol. Sin.*, 10, 761–768. <https://doi.org/10.1007/s11589-997-0008-0>
- 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>
- Zhu, L. P., and HelMBERGER, D. V. (1996). Advancement in source estimation techniques using broadband regional seismograms. *Bull. Seismol. Soc. Am.*, 86(5), 1634–1641. <https://doi.org/10.1785/BSSA0860051634>
- 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>