Supplementary Information for

Top-down and bottom-up fluid diffusion jointly induce seismicity in the Weiyuan shale gas field, China

Weixin Qi[1], Yuyang Tan*[2], Haijiang Zhang*[1,3], Maomao Wang[4], Qiangqiang Zheng[5], Jiawei Qian[6] & Jun Hu[7]

[1]School of Earth and Space Sciences, Key Laboratory of Precision Geodesy, University of Science and Technology of China, Hefei 230026, China, University of Science and Technology of China, Hefei, China.

[2]Frontiers Science Center for Deep Ocean Multispheres and Earth System, State Key Laboratory of Physical Oceanography, Key Lab of Submarine Geosciences and Prospecting Techniques, MOE and College of Marine Geosciences, Ocean University of China, Qingdao 266100, China

[3]Mengcheng National Geophysical Observatory, University of Science and Technology of China, Hefei, Anhui 230026, China

[4]Department of Marine Geology and Geophysics, College of Oceanography, Hohai University, Nanjing, China

[5]School of Civil and Architecture Engineering, Anhui University of Science and Technology, Huainan Anhui 232001, China

[6]School of Carbon Neutrality Science and Engineering, Anhui University of Science and Technology, Hefei, Anhui 231131, China

[7]State Key Laboratory of Geohazard Prevention and Geoenviroment Protection, Chengdu University of Technology, Chengdu, Sichuan, China

Corresponding authors: Yuyang Tan and Haijiang Zhang (tanyuyang@ouc.edu.cn; zhang11@ustc.edu.cn)

The supplementary information includes:

Supplementary Method 1. Earthquake location uncertainty analysis

We applied the bootstrapping method[1,2] to evaluate earthquake location uncertainties from hypoDD. In each bootstrap run, we randomly resampled differential traveltimes while keeping the total number of differential traveltimes equal to that of the original dataset. Using the resampled differential traveltimes, we then relocated the earthquakes by hypoDD. This process was repeated 100 times, yielding 100 relocation estimates for each event. The location uncertainty for each event was determined by the standard deviation of its 100 relocations. Finally, we found that the average location uncertainties of all events were 0.25 km in the E-W direction, 0.14 km in the N-S direction, and 0.46 km in the vertical direction (Supplementary Fig. 3).

To evaluate location uncertainties arising from the choice of velocity model, we compiled 10 1D velocity models of the Sichuan Basin from different studies (Supplementary Fig. 2)[3-12]. Because the model of Meng, et al.[4] based on vertical sonic logging covers only a limited depth range (0-2.8 km), we excluded it and used the remaining 9 models to evaluate location uncertainties following the same routine as described in the main text. The location uncertainty for each event was determined by the standard deviation of the 9 relocations. The average uncertainties of all events in the E-W, N-S and vertical directions are 0.54, 0.22, and 1.30 km, respectively (Supplementary Fig. 4).

Supplementary Method 2. Model resolution test

In order to evaluate the resolution of the inverted velocity models, we employed both the restoration test and the checkerboard resolution test (CRT)[13]. The restoration test was done by first calculating synthetic traveltimes based on the actual distribution of seismic events and stations, as well as the velocity models inverted by the Vp/Vs model consistency-constrained double-difference seismic tomography method (tomoDDMC)[14]. These synthetic traveltimes are used as the input for the tomoDDMC inversion with the starting models and inversion parameters same as the real data inversion. The restored models are then compared with the real data inversion models (Supplementary Fig. 9). Generally, the velocity models from the real data inversion are

accurately restored, therefore, buttressing the reliability of the inverted models.

The checkerboard velocity models were generated by adding +5% and −5% velocity perturbations to the initial models. Because the checkerboard patterns of the Vp and Vs models are set to be opposite at the same grid nodes, the velocity perturbations of the Vp/Vs model vary between +10.5% and −9.5%. These checkerboard models were then utilized to calculate synthetic traveltimes based on the actual distribution of seismic events and stations. Finally, we applied tomoDDMC method[14] to the synthetic data using the same starting models and inversion parameters (with the exception of reduced smoothing and damping parameters of 1 and 40, respectively) in the real data inversion. Supplementary Fig. 9 shows the inverted velocity models. It can be seen that the checkerboard patterns are most evident in the source areas, indicating that the velocity structures in these regions can be well resolved.

Supplementary Method 3. In-situ Vp/Vs estimation

To verify the reliability of the Vp/Vs model obtained by tomoDDMC, we estimated the Vp/Vs values for some tomography grid node using an in-situ Vp/Vs analysis method[15] based on high-precision cross-correlation differential traveltimes. First, the seismic events within 1 km of each tomography grid node is regarded as a cluster. For each pair of seismic events within a cluster, if the spacing between the two events is sufficiently small with respect to the source-station distance, the seismic velocity can be assumed to be constant locally, and the incident waveform fronts at each station can be approximated as horizontal planes. In this case, the P-wave differential traveltimes δT_p^i between the two events at the station i can be expressed as:

$$\delta T_p^i = T_{p2}^i - T_{p1}^i = \frac{\delta l_p^i}{V_p} \quad (1)$$

where T_{p1}^i and T_{p2}^i are the traveltimes of events 1 and 2, respectively, δl_p^i is the difference in raypath distance between the two events, and V_p is the local P-wave velocity. Similarly, the S-wave differential traveltimes δT_S^i between the two events at the station i can be expressed as:

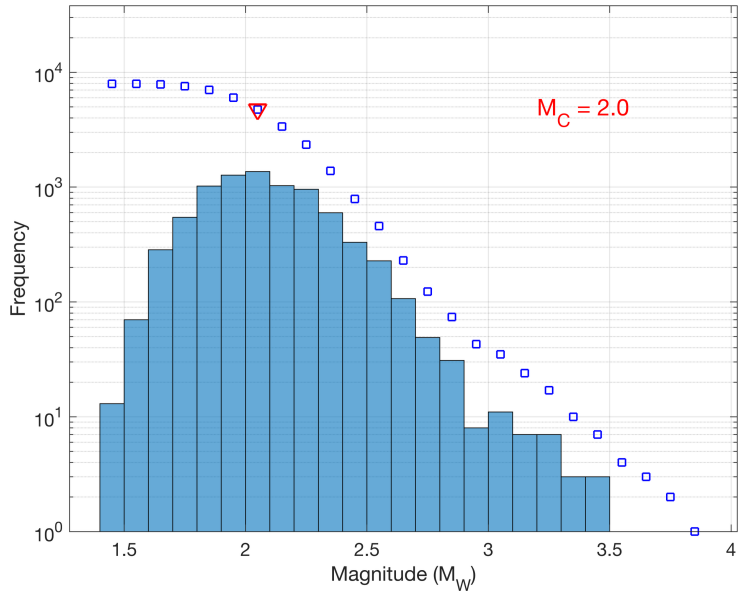
$$\delta T_s^i = T_{s2}^i - T_{s1}^i = \frac{\delta l_s^i}{V_s} \quad (2)$$

If the spacing between the two events is sufficiently small with respect to the source-station distance, it can be assumed that $\delta l_p^i = \delta l_s^i$, and therefore the in-situ Vp/Vs ratio can be estimated directly from the differential traveltimes,

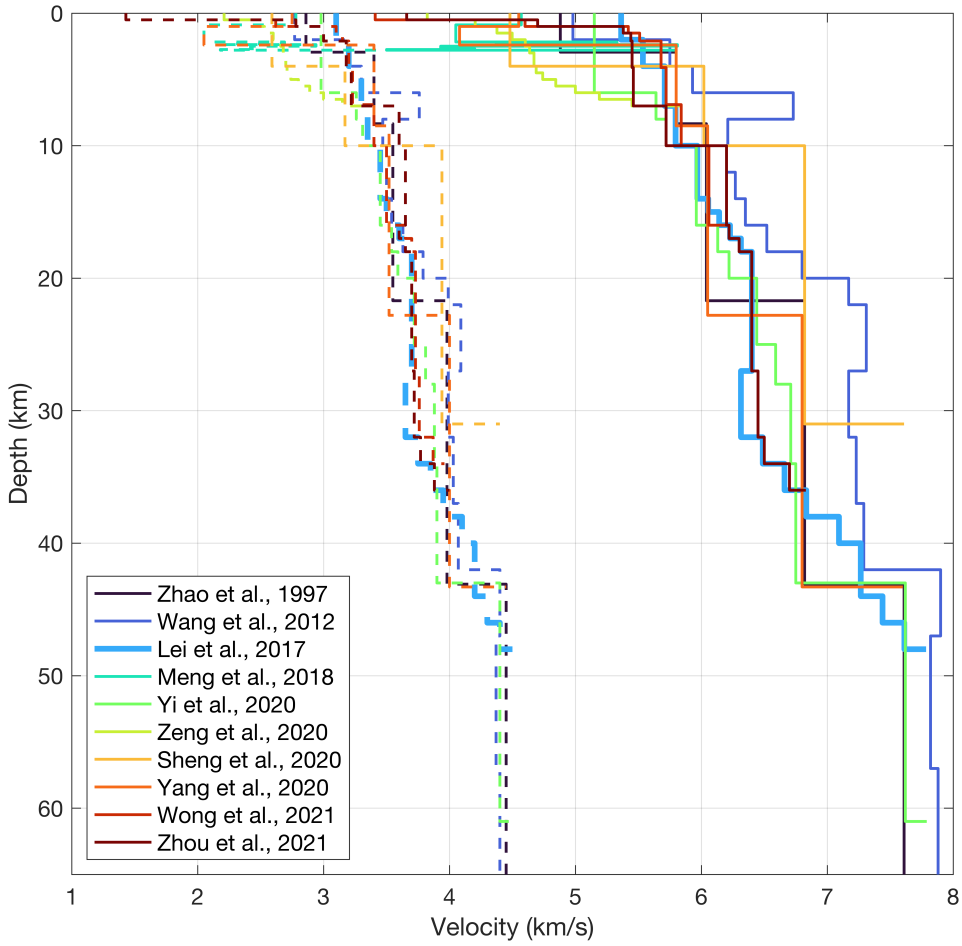
$$\frac{V_p}{V_s} = \frac{\delta T_s^i}{\delta T_p^i} \quad (3)$$

The in-situ Vp/Vs ratio for each seismic cluster can be estimated by fitting the slope of the differential traveltimes δT_p^i-δT_s^i for each event pair. For cross-correlation-based differential traveltimes of event pairs with cross-correlation coefficients cc $\geqslant$ 0.6, we first removed influential points based on Cook's distance[16]. If three or more differential traveltimes remain after removing the influential points, these differential traveltimes are retained and demeaned to minimize the effect of varying y-axis intercepts for different event pairs due to uncertainty in the seismic origin time. If the total number of retained differential traveltimes for all event pairs in a cluster exceeds 100, a least-squares fit is performed after removing high leverage and influential points[16,17] and removing points with large residuals (greater than 3 times the standard deviation of the residuals). Finally, a least-squares fit is performed using the remaining differential traveltimes and the slope of the resulting line represents the in-situ Vp/Vs ratio for the cluster.

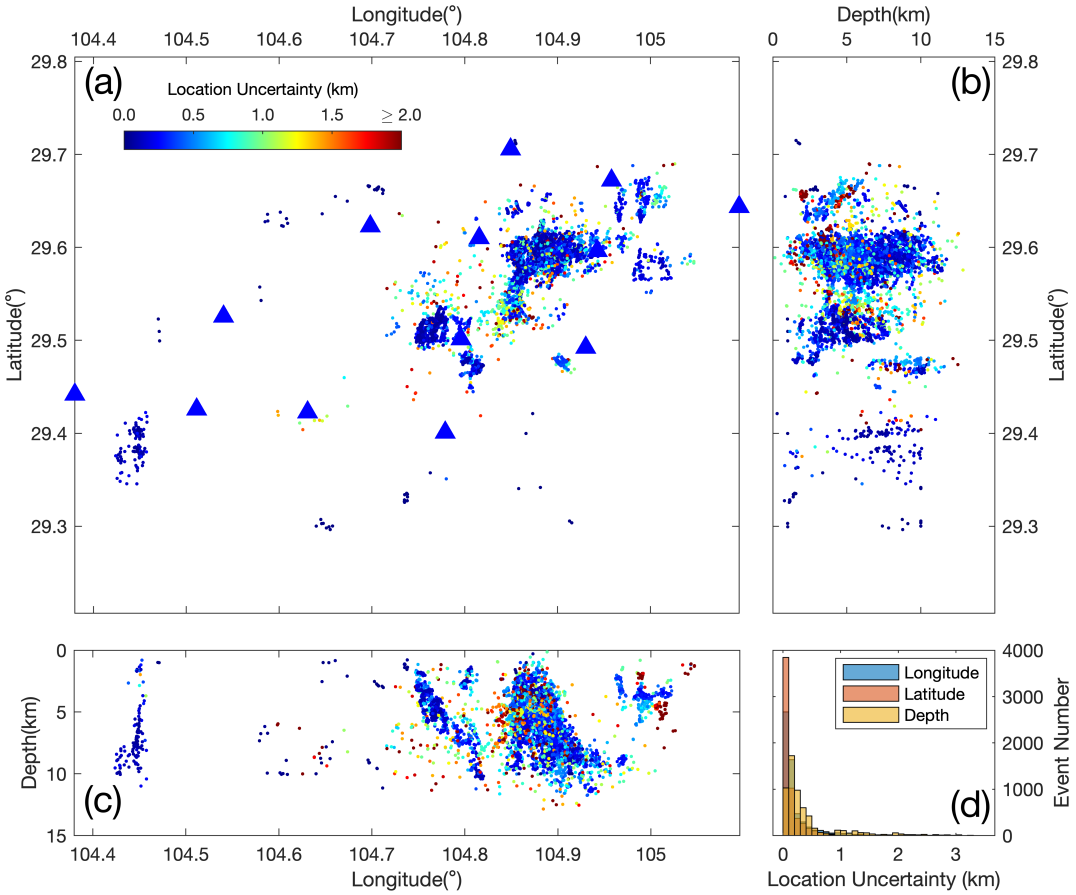
To determine the uncertainties of the in-situ Vp/Vs estimates, we applied the bootstrapping method[1,15]. For each 3D grid node, we fit the slope of the line using the randomly resampled differential traveltimes δT_p^i-δT_s^i. This process was repeated 100 times, and the uncertainty of the in-situ Vp/Vs value at each grid node was determined by the standard deviation of the 100 Vp/Vs estimates.



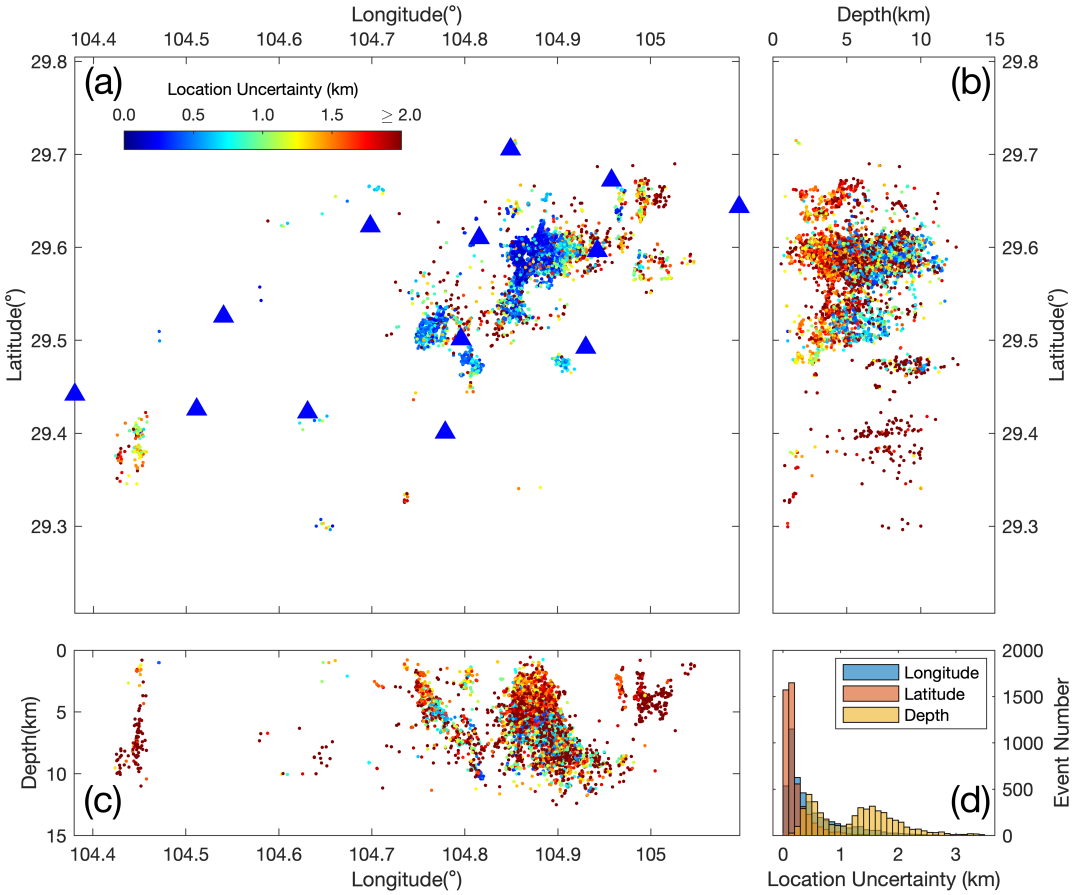
Supplementary Fig. 1. Earthquake magnitude-frequency distribution. The blue bars represent the histogram of seismic moment magnitude M_W. The open squares represent the cumulative earthquake number. The triangle represents the completeness magnitude M_C.



Supplementary Fig. 2. 1D velocity models of the Sichuan Basin assembled from different studies[3-12]. The solid and dashed lines represent the Vp and Vs models, respectively.

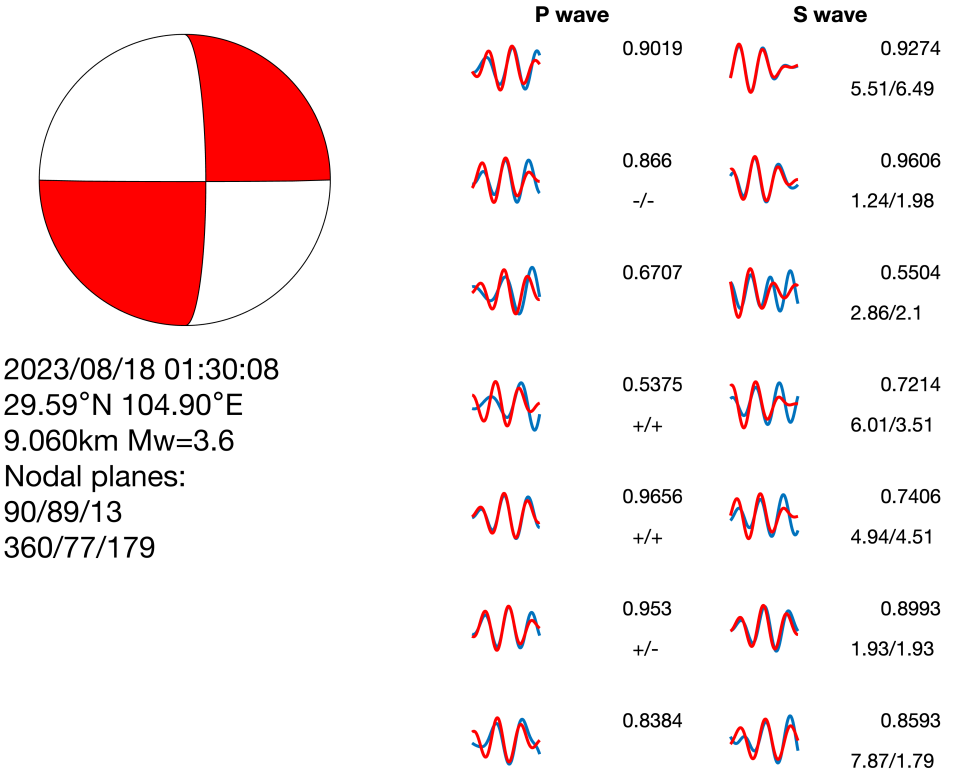


Supplementary Fig. 3. Distribution of the earthquake location uncertainties determined using the bootstrapping method. The location uncertainties are indicated by the colors of the dots. **(a)** shows the location uncertainties in the map view, while **(b)** and **(c)** show the uncertainties in the N-S and W-E directions, respectively. **(d)** shows the histograms of the location uncertainties in the longitude (blue bars), latitude (red bars) and depth (yellow bars) directions, respectively.

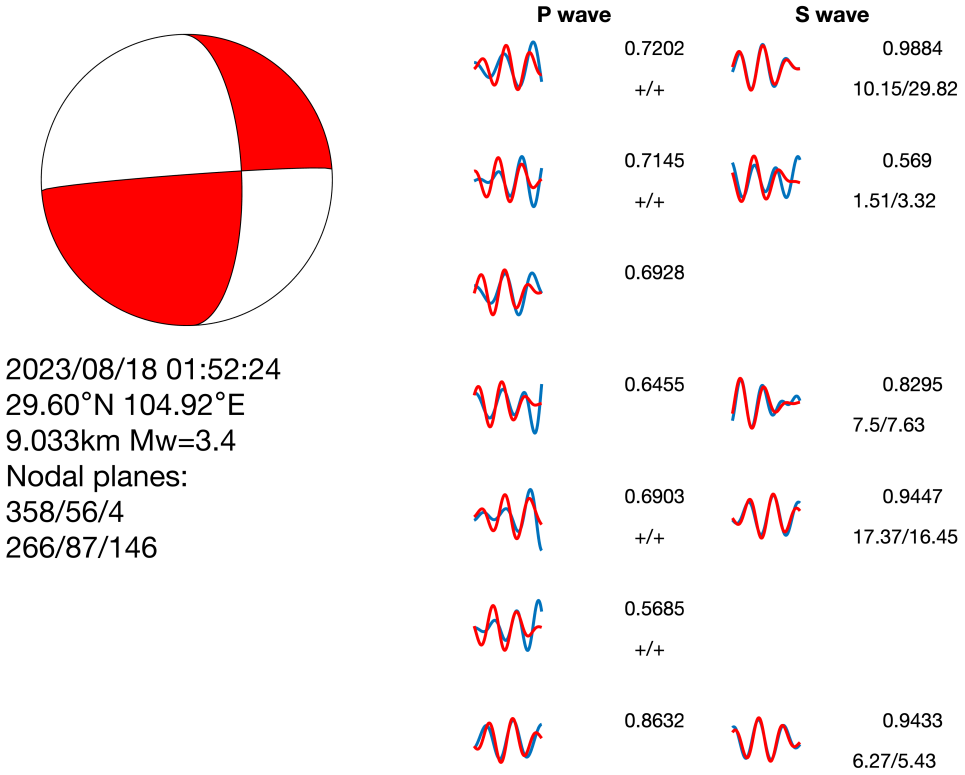


129

130 **Supplementary Fig. 4. Distribution of the earthquake location uncertainties related to**
131 **different 1D velocity models in Supplementary Fig. 2.** The symbols are the same as those in
132 Supplementary Fig. 3.

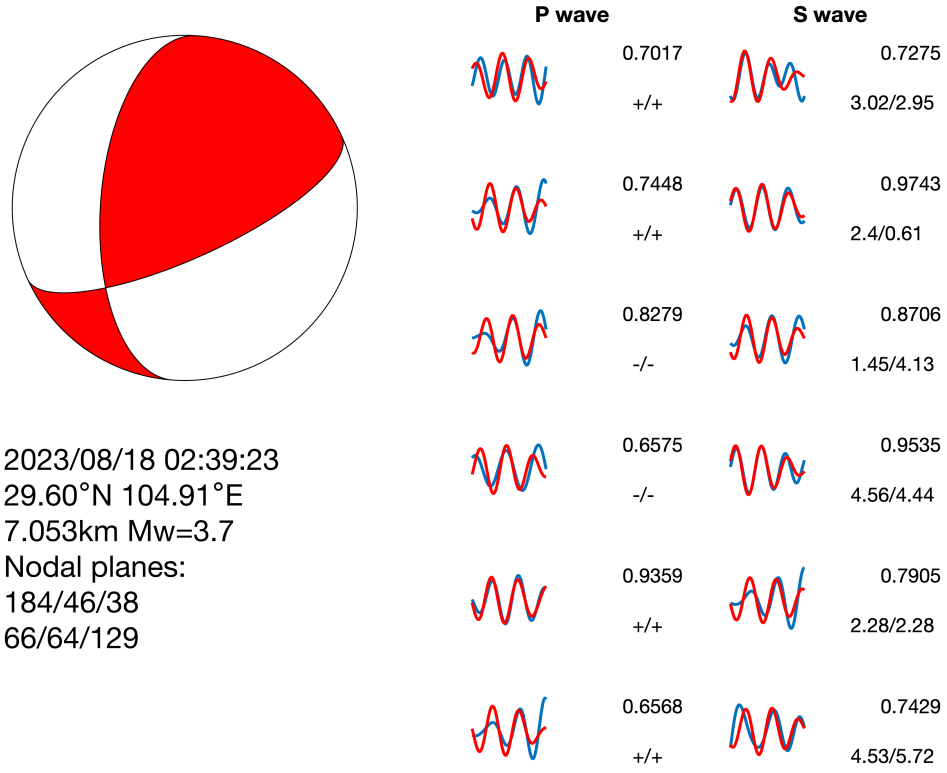


Supplementary Fig. 5. Focal mechanism solution (represented by the beachball) of the 2023/08/18 Mw 3.6 earthquake. The source parameters of this event are given below the beachball. The blue and red lines in the right panel represent the observed and modeled seismograms, respectively. The number at the upper right corner of each waveform pair represents the maximum cross-correlation coefficient between observed and modeled seismograms. The signs at the lower right corner of each waveform pair in the left column represent the comparison between observed (left) and modeled (right) P wave first motion polarities. The numbers at the lower right corner of each waveform pair in the right column represent the comparison between observed (left) and modeled (right) S/P amplitude ratios.



143

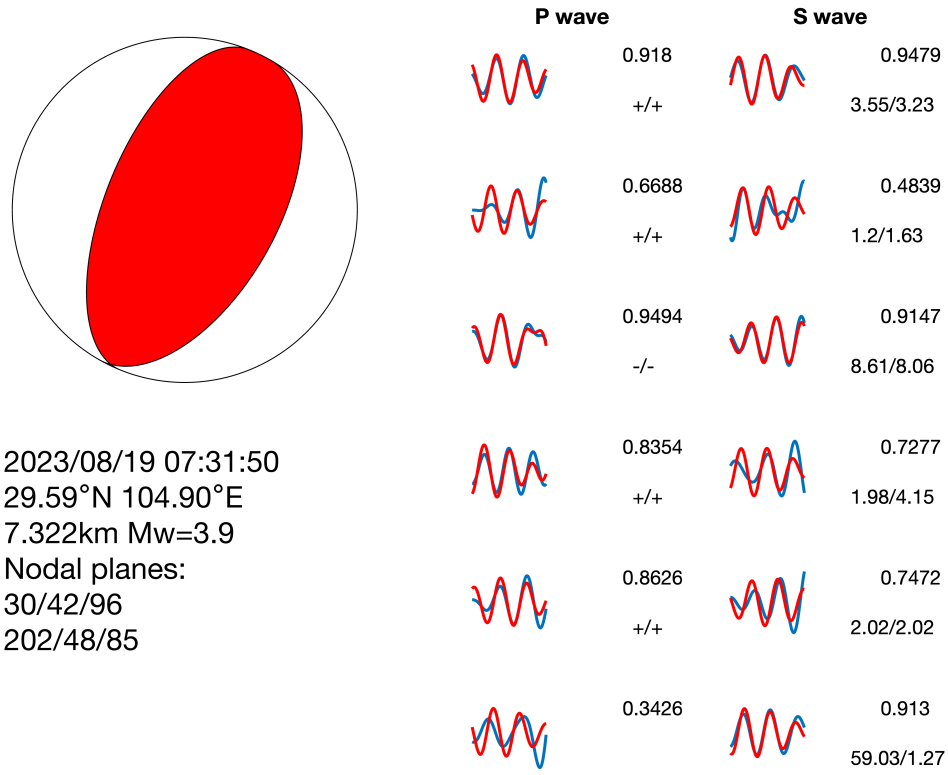
144 **Supplementary Fig. 6. Focal mechanism solution (represented by the beachball) of the**
145 **2023/08/18 Mw3.4 earthquake.** The symbols are the same as those in Supplementary Fig. 5.
146



147

148 **Supplementary Fig. 7. Focal mechanism solution (represented by the beachball) of the**
149 **2023/08/18 Mw3.7 earthquake.** The symbols are the same as those in Supplementary Fig. 5.

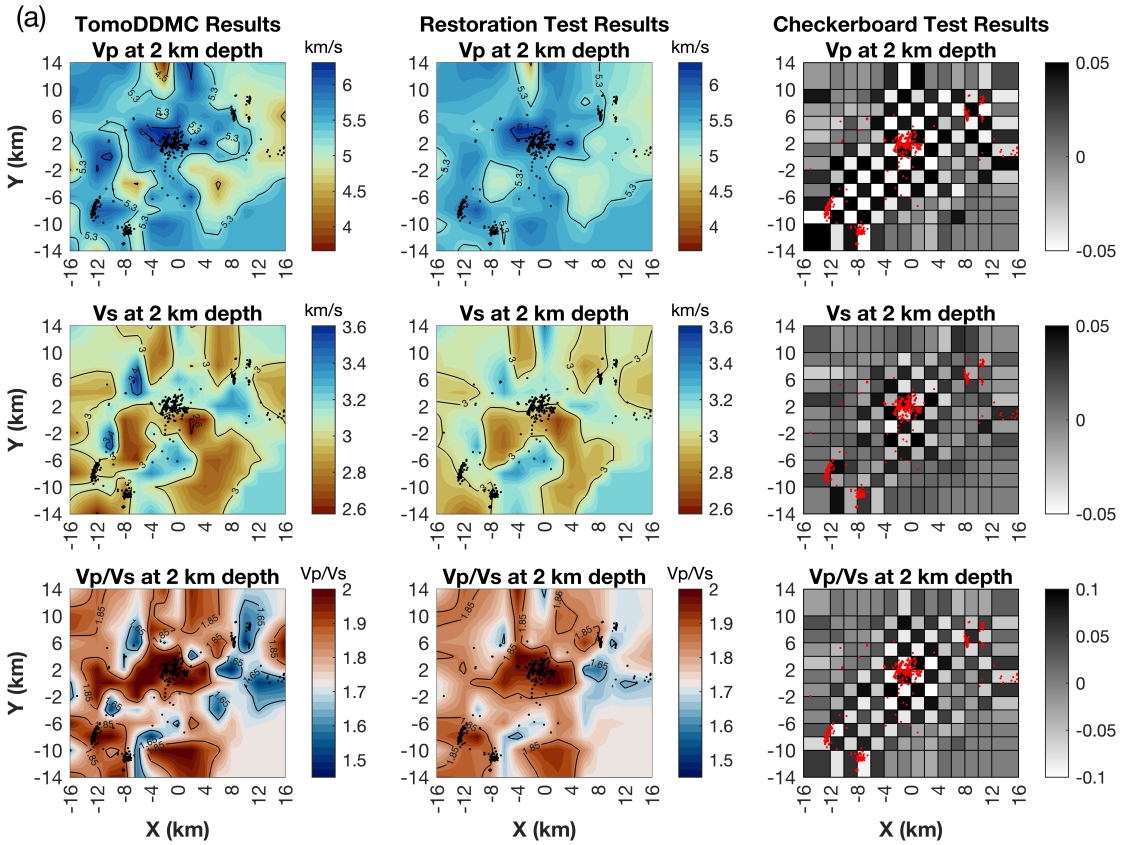
150

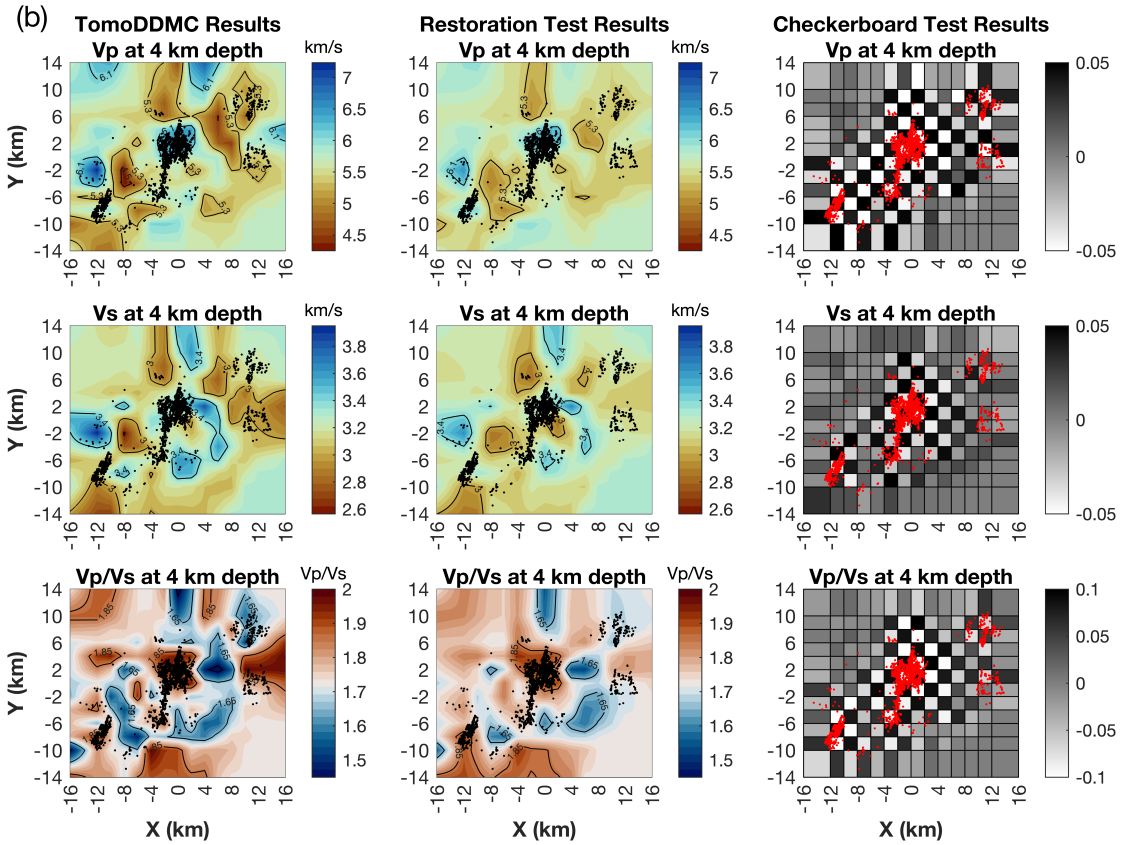


151

152 **Supplementary Fig. 8. Focal mechanism solution (represented by the beachball) of the**
153 **2023/08/19 Mw3.9 earthquake.** The symbols are the same as those in Supplementary Fig. 5.

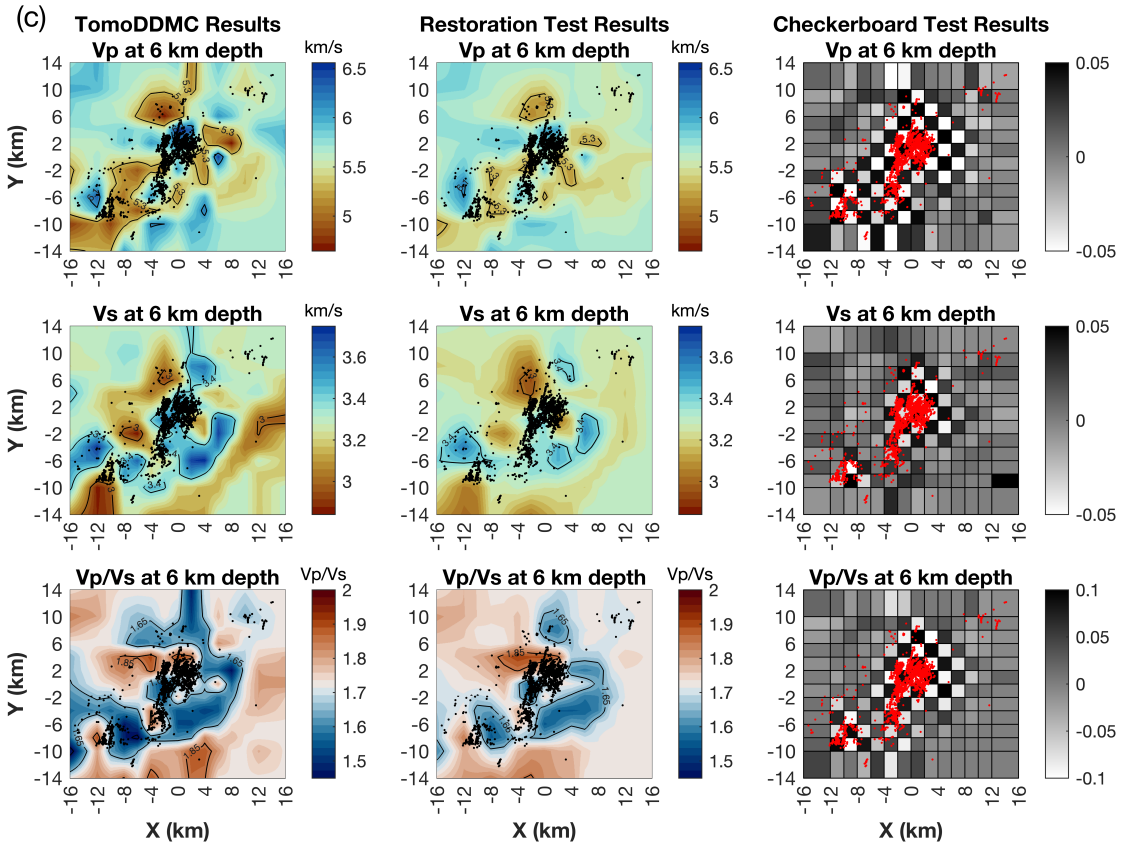
154



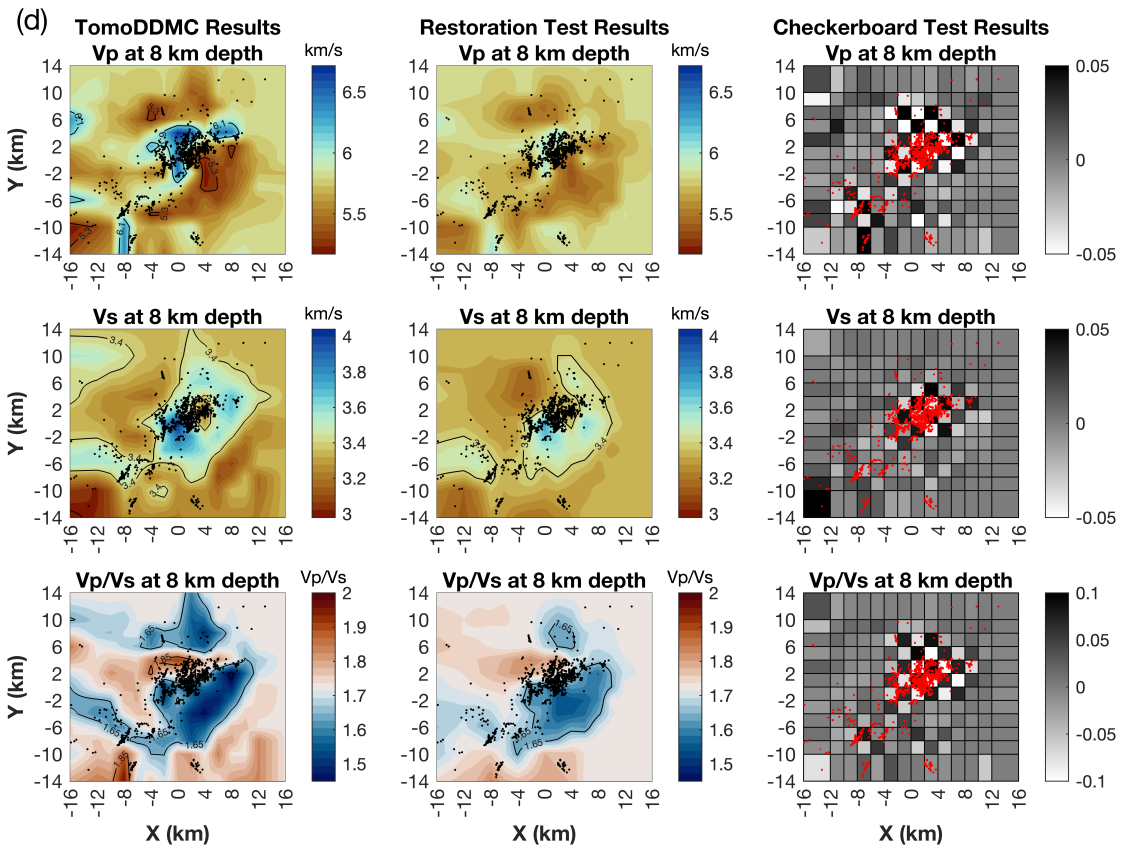


Supplementary Fig. 9. Depth slices of the inverted and restored Vp, Vs and Vp/Vs models as well as recovered checkerboard models at different depths. (a) Slices at 2 km depth. **(b)** Slices

159 at 4 km depth.

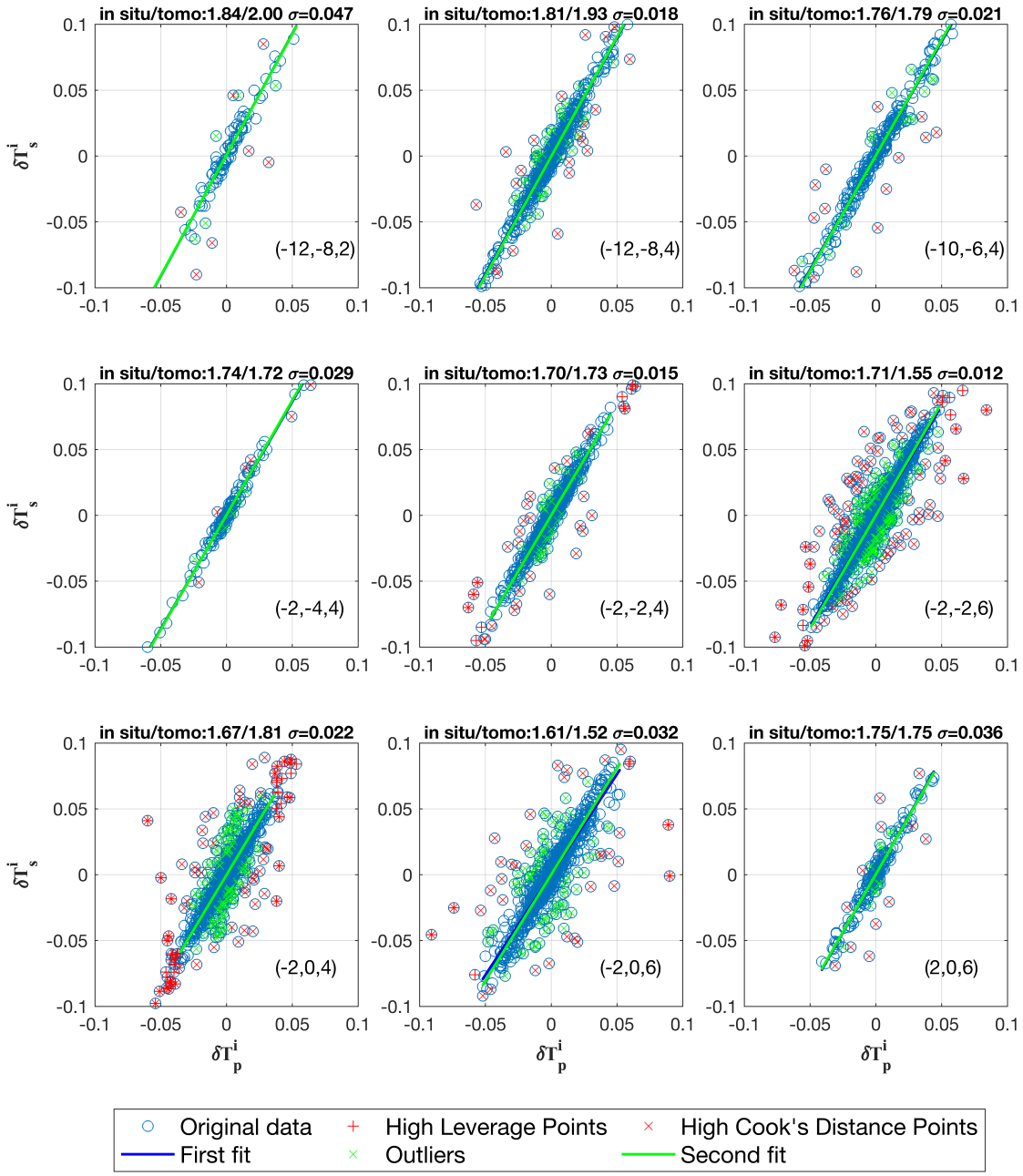


160

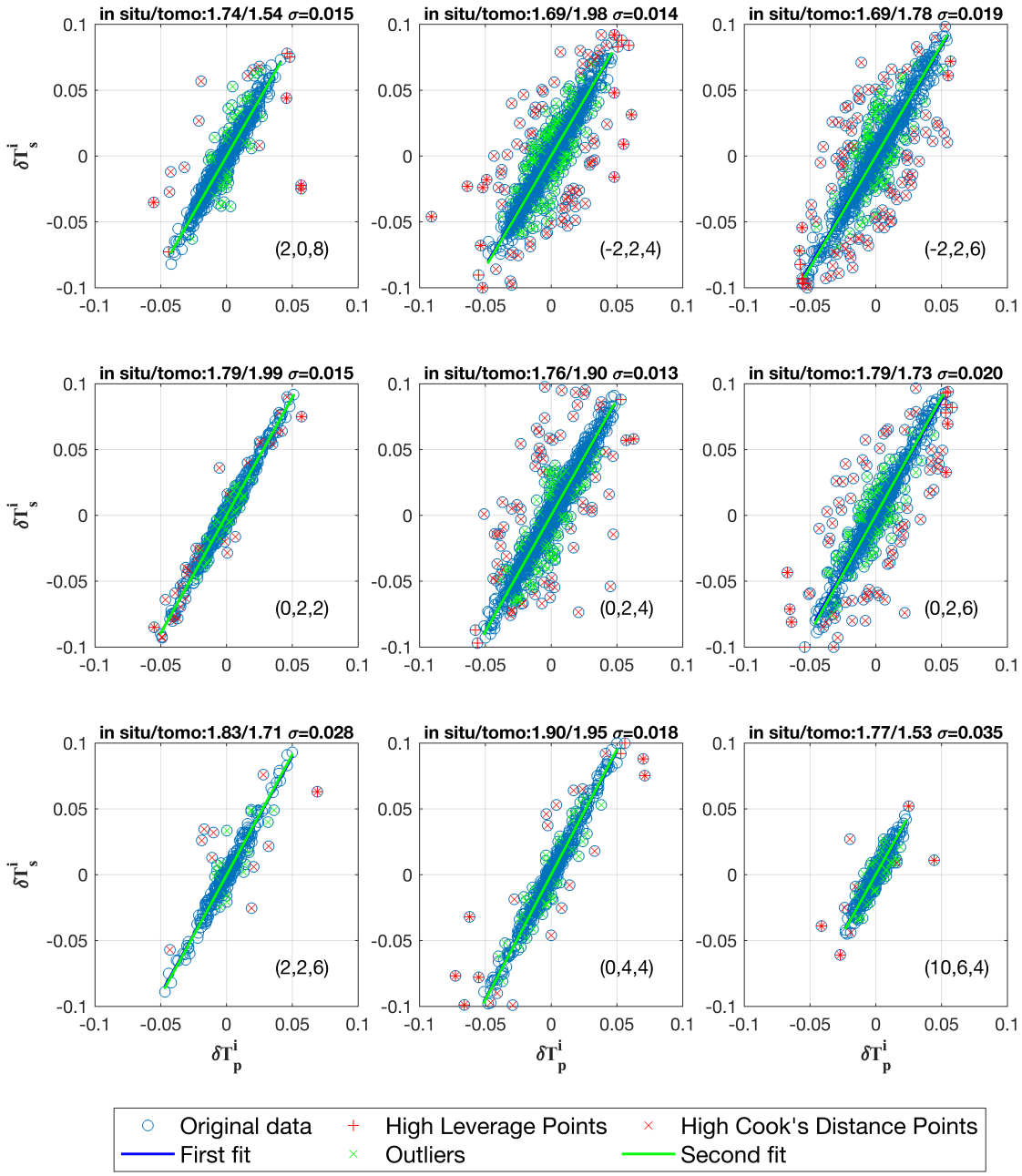


161

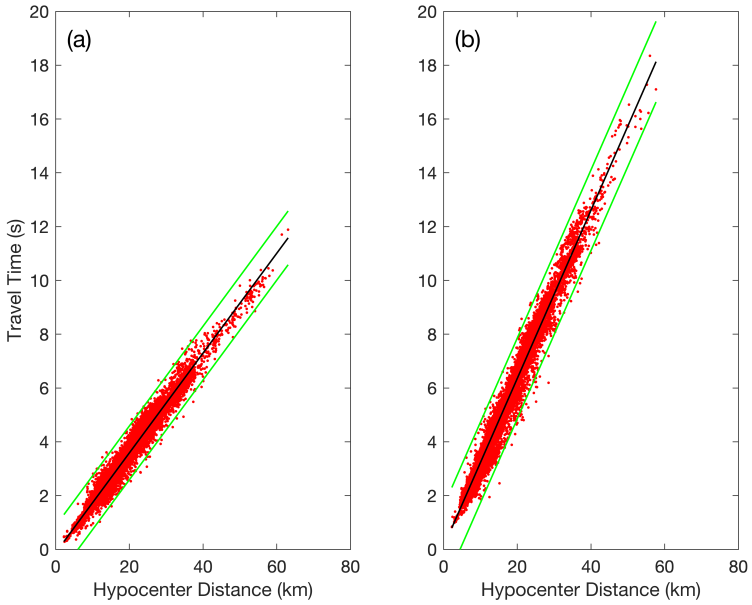
162 **Supplementary Fig. 9. Continued. (c)** Slices at 6 km depth. **(d)** Slices at 8 km depth.



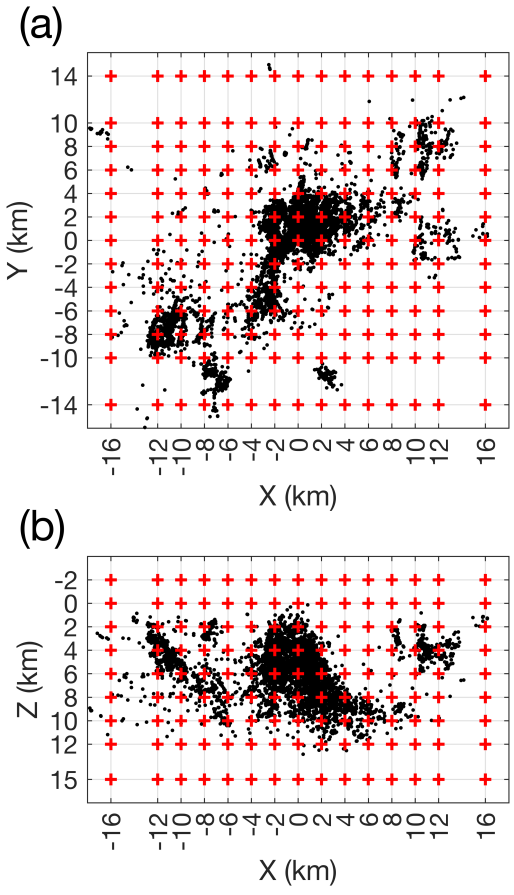
Supplementary Fig. 10. Cross plots of differential P-wave (δT_p^i) and S-wave traveltimes (δT_s^i) for each grid nodes. The Vp/Vs values determined by in-situ estimation and seismic tomography as well as the uncertainty obtained using the bootstrapping method are given in the title of each sub-figure. The coordinates of the grid nodes are given in the lower right corner of each sub-figure. The blue circles represent the original data. The red plus signs represent the high leverage points. The red crosses represent the high Cook's distance points. The blue lines represent the first fit. The green crosses represent the outliers after the first fit. The green lines represent the second fit.



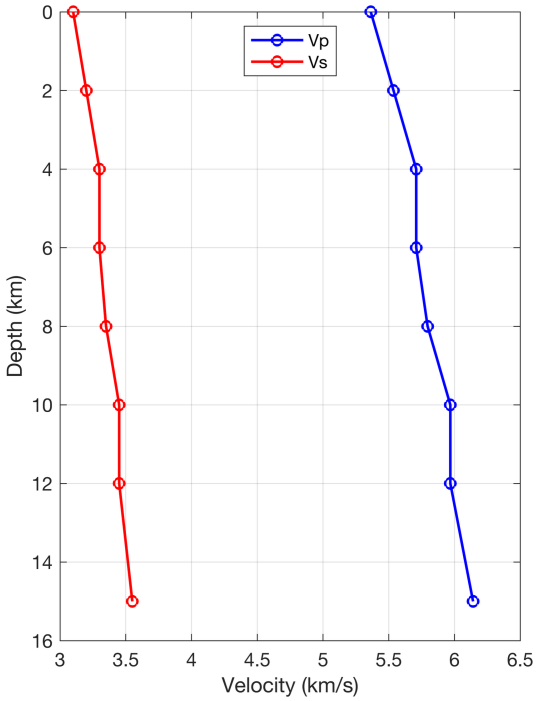
Supplementary Fig. 10. Continued.



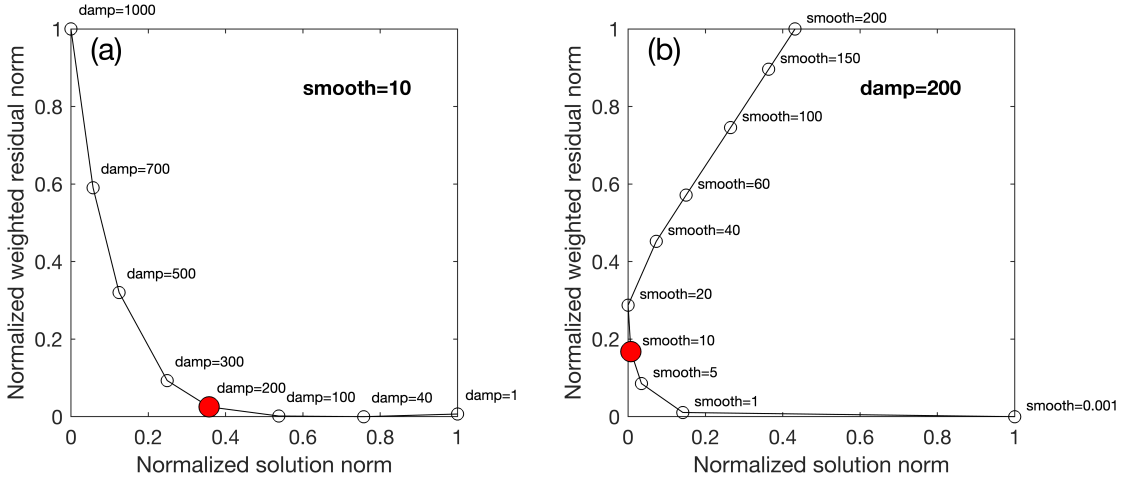
Supplementary Fig. 11. Traveltime-hypocentral distance curves for P- (a) and S-waves (b). The black lines represent the best-fit linear regression derived from the data. Data points lying outside the green threshold lines were excluded.



Supplementary Fig. 12. Inversion grid setting for double-difference seismic tomography. (a) The plane view of the grid nodes (red crosses). **(b)** Vertical cross-section of the grid nodes. The black dots represent the earthquakes.



Supplementary Fig. 13. Regional 1D velocity model with Vp/Vs ratio of 1.73 used as the starting model for the earthquake location and seismic tomography.



Supplementary Fig. 14. Regularization parameters determined using the trade-off curves. The red circles represent the optimal values of the damping **(a)** and smoothing **(b)** parameters.

Supplementary References

1 Efron, B. & Tibshirani, R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Stat. Sci.* **1**, 54-75 (1986).

2 Sheng, M. *et al.* Earthquakes triggered by fluid diffusion and boosted by fault reactivation in Weiyuan, China due to hydraulic fracturing. *J. Geophys. Res. Solid Earth* **127**, e2021JB022963 (2022).

3 Lei, X. *et al.* Fault reactivation and earthquakes with magnitudes of up to Mw4.7 induced by shale-gas hydraulic fracturing in Sichuan Basin, China. *Sci. Rep.* **7**, 7971 (2017).

4 Meng, X. *et al.* 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**, 1643-1658 (2018).

5 Sheng, M. *et al.* Source Parameters of Three Moderate Size Earthquakes in Weiyuan, China, and Their Relations to Shale Gas Hydraulic Fracturing. *J. Geophys. Res. Solid Earth* **125**, e2020JB019932 (2020).

6 Wang, X. *et al.* Fine velocity structure and relocation of the 2010 M_L5.1 earthquake sequence in Rongchang gas field. *Seismol. Geol.* **34**, 348-358 (2012).

7 Wong, W. C. J., Zi, J., Yang, H. & Su, J. 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**, 485-500 (2021).

8 Yang, H. *et al.* A shallow shock: The 25 February 2019 ML 4.9 earthquake in the Weiyuan shale gas field in Sichuan, China. *Seismol. Res. Lett.* **91**, 3182-3194 (2020).

9 Yi, G., Long, F., Liang, M., Zhao, M. & Wang, S. Geometry and tectonic deformation of seismogenic structures in the Rongxian-Weiyuan-Zizhong region, Sichuan Basin: insights from focal mechanism solutions. *Chin. J. Geophys. (Chin.)* **63**, 3275-3291 (2020).

10 Zeng, Q., Chu, R. S., Sheng, M. H. & Wei, Z. G. Seismic ambient noise tomography for shallow velocity structures beneath Weiyuan, Sichuan. *Chin. J. Geophys. (Chin.)* **63**, 944-955 (2020).

11 Zhao, Z., Fan, J., Zheng, S.-H., Hasegawa, A. & Horiuchi, S. Crustal structure and accurate hypocenter determination along the Longmenshan fault zone. *Acta Seismol. Sin.* **10**, 761-768 (1997).

12 Zhou, P. *et al.* Machine-learning-facilitated earthquake and anthropogenic source detections near the Weiyuan Shale Gas Blocks, Sichuan, China. *Earth Planet. Phys.* **5**, 501-519 (2021).

13 Humphreys, E. & Clayton, R. W. Adaptation of back projection tomography to seismic travel time problems. *J. Geophys. Res. Solid Earth* **93**, 1073-1085 (1988).

14 Guo, H., Zhang, H. & Froment, B. Structural control on earthquake behaviors revealed by high-resolution Vp/Vs imaging along the Gofar transform fault, East Pacific Rise. *Earth Planet. Sci. Lett.* **499**, 243-255 (2018).

15 Lin, G. & Shearer, P. Estimating local Vp/Vs ratios within similar earthquake clusters. *Bull. Seismol. Soc. Am.* **97**, 379-388 (2007).

16 Cook, R. D. Detection of influential observation in linear regression. *Technometrics* **19**, 15-18 (1977).

17 Chatterjee, S. & Hadi, A. S. Influential observations, high leverage points, and outliers in linear regression. *Stat. Sci.* **1**, 379-393 (1986).