

Observing Seismic Variations by Earth and Lab Fluids and Fractures

A DISSERTATION PRESENTED
BY
CONGCONG YUAN
TO
THE DEPARTMENT OF EARTH AND PLANETARY SCIENCES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
GEOPHYSICS AND SEISMOLOGY
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
APRIL 2024

©2024 – CONGCONG YUAN
ALL RIGHTS RESERVED.

Observing Seismic Variations by Earth and Lab Fluids and Fractures

ABSTRACT

Understanding the mechanical dynamics within the Earth's crust is crucial for both environmental and energy sustainability. These dynamics are often tied to mechanical deformations triggered by variations in fluid pressure and stress levels. The primary method of investigation involves seismic monitoring to detect changes in material properties and identify fractures or faults, making the study of the relationship between seismic properties, fluids, and fractures essential.

My research utilizes a combination of laboratory experiments, numerical simulations, and field observations to enhance our understanding of the mechanisms behind seismic variations, including changes in wavespeeds, attenuation, and the behavior of laboratory and natural earthquakes. My thesis is divided into two interrelated parts: (1) the influence of fluids and/or deformation on seismoacoustic wavespeed and attenuation, and (2) the effect of fluid properties and pressure on fracture behavior and seismoacoustic signals.

In the first part, I develop wavelet-domain techniques for analyzing frequency-dependent velocity changes through coda wave interferometry and delve into the depth sensitivity of these changes using wavefield simulations. I also employ controlled acoustic monitoring and 3D-printed media to examine how various physical conditions, such as consolidation, saturation, and strain deformation, affect frequency-dependent velocity changes and attenuation. These works aim to use these velocity change and attenuation spectra to either determine the depth of perturbations or to understand the underlying physical mechanisms. Moreover, I explore the potential of probing deep volcanic activity through inter-source interferometry, using repeating earthquakes from Mount St. Helens. This part proposes novel methods and insights to tackle the challenges in imaging and understanding of seismic property changes, with applications ranging from subsurface exploration to volcano monitoring.

In the second part, I focus on hydrofracturing dynamics, employing advanced imaging and acoustics to study fluid-induced fracturing patterns and their correlation with seismic activities, using the Cascadia region's tectonic tremor swarms as a case study. I also investigate hydromechanics within artificial fault-valve media, specifically fluid migration and its effects on fault instability and permeability. I further explore high-performance seismic processing, developing a new ensemble learning framework to enhance the generalizability of seismic phase pickers. This part aims to advance understanding of fluid-induced deformations and their implications for seismic hazard assessment and resource optimization.

Contents

1	INTRODUCTION AND MOTIVATION	1
1.1	General introduction	1
1.2	Fluid and/or deformation influence the seismoacoustic properties	4
1.3	Fluid modulates deformation, fracturing, and seismoacoustic behaviors	6
2	NUMERICAL COMPARISON OF TIME-, FREQUENCY-, AND WAVELET-DOMAIN METHODS FOR CODA WAVE INTERFEROMETRY	9
2.1	Introduction	11
2.2	Methodologies	15
2.3	Validation using full waveform modeling	21
2.4	Sensitivity with the depth of the velocity change	30
2.5	Further Discussion	33
2.6	Conclusions	40
3	LABORATORY ANALOG TO MONITORING UNCONFINED AQUIFERS WITH PASSIVE SEISMOLOGY	44
3.1	Introduction	45
3.2	Method	47
3.3	Results	51
3.4	Discussion	56
3.5	Conclusion	57
4	SEISMIC MONITORING OF MAGMATIC ACTIVITY BENEATH MOUNT ST. HELENS	59
4.1	Introduction	60
4.2	Data	62
4.3	Methods	62
4.4	Results	67
4.5	Conclusion	73

5	LABORATORY HYDROFRACTURES AS ANALOGS TO TECTONIC TREMORS	75
5.1	Introduction	76
5.2	Method	78
5.3	Results	82
5.4	Comparison with Tectonic Tremors	87
5.5	Conclusions	94
6	COMPACTION-ENHANCED INSTABILITY IN LABORATORY FAULT-VALVE MEDIA	97
6.1	Introduction	98
6.2	Methodology & Measurements	99
6.3	Results	101
6.4	Discussions	104
6.5	Conclusion	106
7	BETTER TOGETHER: ENSEMBLE LEARNING FOR EARTHQUAKE DETECTION AND PHASE PICKING	108
7.1	Introduction	109
7.2	Methodology	112
7.3	Performance of ELEP framework	120
7.4	Deployment on continuous data	134
7.5	Discussion and Conclusion	139
8	CONCLUSION AND OUTLOOK	142
	APPENDIX A	147
A.1	Demonstration that the real part of the CWT is a narrow bandpassed version of the original time series	147
A.2	Parameters for the structural model, the numerical simulations, the dv/v calculations using all approaches, and performance summary.	149
A.3	Supplementary figures - additional figures for Chapter 2	151
	APPENDIX B	171
B.1	Supplementary figures - additional figures for Chapter 3	172
	APPENDIX C	178
C.1	Supplementary figures - additional figures for Chapter 4	178
	APPENDIX D	181
D.1	Relationship between Radiated Energy and Fracture Area	181
D.2	AE Energy Calibration	183
D.3	Supplementary figures - additional figures for Chapter 5	186

Author List

The following authors contributed to Chapter 2: *Congcong Yuan, Jared Bryan, Marine Denolle.*

Portion of Chapter 3 will be published in a forthcoming paper by: *Congcong Yuan, Marine Denolle, David Weitz.*

Portion of Chapter 4 will be published in a forthcoming paper by: *Congcong Yuan, Marine Denolle.*

The following authors contributed to Chapter 5: *Congcong Yuan, Thomas Cochard, Marine Denolle, Joan Gomberg, Aaron Wech, Lizhi Xiao, David Weitz.*

Portion of Chapter 6 will be published in a forthcoming paper by: *Congcong Yuan, Yujing Du, Lluís Saló-Salgado, Marine Denolle, David Weitz.*

The following authors contributed to Chapter 7: *Congcong Yuan, Yiyu Ni, Youzuo Lin, Marine Denolle.*

Listing of figures

1.1	Overview of research themes. (A) Schematic representations of target research contexts. (B) Seismic monitoring of volcanic activities. (C) Seismo-acoustic monitoring of reservoir mechanical properties. (D) Laboratory hydrofracturing within rigid media. (E) Laboratory hydromechanics in fault-valve media.	4
2.1	Synthetic experiment for the code verification exercise. (a) Reference synthetic waveform (blue) with the stretched current waveform (red). (b) Their power spectral densities.	20
2.2	Velocity perturbations dv/v from the stretched synthetic waveform shown in Fig.2.1 measured with each method. The black dashed line represents the “true” value of dv/v . (a) Time-domain group of methods of the raw waveforms: the star indicates the result from WCC, the square indicates the result from TS, and the circle indicates the result from DTW. (b) Fourier- and wavelet-domain groups of methods for waveforms filtered at multiple narrow frequency bands: the yellow stars indicate the results from MWCS, the blue triangles indicate the results from WCS, the green squares indicate the results from WTS, and the orange circles indicate the results from WTDTW. (c) Wavelet-domain group of methods for all frequencies: the blue triangles indicate the results from WCS, the green squares indicate the results from WTS, the orange circles indicate the results from WTDTW.	22
2.3	(a) Heterogeneous V_P velocity structure is shown in the colormap, the source is the red star, and the receiver is the gray triangle. (b) Waveforms simulated with SPECFEM2D in the reference medium and medium perturbed by +0.1% (current). (c) Power spectral densities corresponding to the waveforms in (b).	24
2.4	Predictions of dv/v from coda-wave interferometry with various techniques for a uniform change in velocity. (a) dv/v results of three time-domain methods WCC (star), TS (square), and DTW (circle). (b) dv/v results at various frequency bands for MWCS (yellow stars), WCS (blue triangles), WTS (green squares), and WTDTW (orange circles). (c) The spectrum of dv/v results of the wavelet-domain methods for WCS (blue triangles), WTS (green squares), and WTDTW (orange circles).	26

2.5	Spectra of dv/v obtained from three wavelet-domain methods for three experiments in the zero-offset source-receiver case: a uniform change (a), a change in the shallow layer (b), and a change in the deep layer (c).	28
2.6	Same as Fig.2.5, but for the case of the distant-offset source and receiver configuration.	29
2.7	Apparent velocity perturbation versus the depth of the perturbed layer for TS (a, b) and WTS (c, d), both in the zero-offset source-receiver configuration (a, c) and distant-offset source-receiver configuration (b, d). The color scale denotes the depth of the perturbed layer. The vertical dashed line denotes the source frequency of 5.0 Hz for (c) and 0.3 Hz for (d).	33
2.8	Errors in velocity-perturbation estimations over 20 realizations of homogeneous velocity changes. In all three panels, the black dashed line represents zero error. (a) Errors of time-domain methods: black stars indicate the errors from using WCC, black squares indicate the errors from using TS, and black circles indicate the errors from using DTW. (b) Errors of four approaches in the bandpassed filtered Fourier and wavelet methods (second group). (c) Errors of three wavelet-domain approaches in the third group.	35
2.9	Mean velocity-perturbation computational time over 20 realizations of homogeneous velocity changes for all methods. Note that the blue, orange, and green ones represent the first, second, and third group of methods, respectively.	38
2.10	Apparent velocity perturbations for the distant-offset source-receiver configuration. The dv/v spectra measured for direct/ballistic surface waves (a) and direct/ballistic body waves (b) as a function of the depth of the perturbed layer. (c) The dv/v spectra measured versus the depth of the perturbed layer for scattered body waves at late lapse times (185 s). (d) The dv/v spectra measured with both scattered body and surface waves at early (55 s), middle (115 s), and late (185 s) lag times for a perturbed layer at 1-km depth. All results are measured using WTS. The vertical dashed line denotes the source frequency of 0.3 Hz.	41
3.1	Experimental apparatus and measurements. (a) Schematic acoustic measurement in a three-layered medium by sandwiching a granular soft medium into two rigid blocks. (b) Examples of acoustic waveforms are collected during dry (black waveforms) and partially filled conditions (red waveforms). (c) Frequency-dependent wavespeed and amplitude changes for different levels of the water table (color-coded).	52
3.2	Acoustic wavespeed and amplitude changes under strain deformations on the fully-bonded granular media in (a) unsaturated or dry state, (b) partially saturated, and (c) fully saturated states.	54
3.3	Acoustic wavespeed and amplitude changes under strain deformations on the unbonded granular media in (a) unsaturated or dry state, (b) partially saturated, and (c) fully saturated states.	55

4.1	Study region and earthquake detections. (a) Borehole and surface stations at Mount St. Helens from the UW, CC, and PB networks; (b) Repeating waveforms (in black) of one example cluster. The stacked waveform is in red. The blue and green lines represent the arrival times of P and S waves. Correlation coefficients between individual waveforms and stacked waveforms are shown top right of waveforms; (c) Temporal evolution of each family: the black and green ones are detected by borehole and surface stations, respectively. The red dots are the example with waveforms shown in (b).	63
4.2	RedPy Event Families location in horizontal and two cross-section views. (a-c) Non-LinLoc initial locations of REDPy earthquakes are in yellow. (d-f) Relocations with corrections of source-specific station terms in red squares. (g-i) 20 groups in different colors are obtained from K-means clustering with a gradual colormap. Note: background images are the P-wave velocity (V_p) perturbations with respect to the average V_p at each depth?	68
4.3	Selection from repeated signals. Two events with the highest repetitions from the shallow-depth group and intermediate-depth group. Their waveforms with ML picks for 16 individual stations are displayed below the geographic maps.	69
4.4	(a) Example waveforms of repeating events recorded at station <i>CC.STG</i> . The cross-correlation value between individual waveforms and stacked waveforms (at the top) is coded in color. Red color represents closer to 1, and blue color represents closer to 0. (b) Inter-source cross-correlation functions between reference or stack waveform in Fig.4.3d of one cluster and repeating waveforms of another cluster. Locations of two clusters are displayed in Fig.4.3a-c.	71
4.5	(a) velocity changes for each receiver; (b) velocity changes for inter-source interferometry.	72
5.1	Experimental apparatus, subtracted images, and acoustic signals for the high-viscous fluid experiment ($\mu=800$ cP). (a) Schematic of the experimental apparatus with a high-speed camera and four AE sensors placed at each sample quadrant (SW, NW, NE, SE). The filled triangles represent the AE sensor locations. The dashed red and solid cyan lines are the fracture and fluid fronts, respectively. The cyan cone is the injected fluid. See additional experimental details in Figs. D.1 and D.2. (b) Three subtracted images from the first image after $T=2.4$ ms of the first burst show nucleation, propagation, and arrest of a single AE event. The red patch and arrow represent the transverse fracture area and direction, respectively. (c) AE signals associated with the fracture event in (b), the black arrows denote the picks of the first arrivals that we use for locating the AE source, which is marked as the orange star in (b).	84

- 5.2 Kymograph, pixel intensity variation rate, and AE signals. Top panels: kymograph of the subtracted images taken in a NE direction of the specimens. The dashed red and solid cyan lines represent the fracture and fluid fronts in the upper panels, respectively. Middle panels: time series of the image intensity rate, which is measured as the averaged pixels of the differential images. Lower panels: acoustic signals recorded by the four AE sensors. The amplitude unit “a.u.” represents an “arbitrary unit” for their normalization to the peak amplitude. The back lines are the radiated energy rates of acoustic signals. (a) For a high fluid viscosity (diluted glycol, $\mu=800$ cP) experiment and (b) for a low viscosity fluid (water, $\mu=0$ cP) experiment. The full-length measurements are displayed in Figs. D.4 and D.5. Note: the time series is trimmed to focus on the hydrofracture propagation instead of its beginnings, which suffer from signal saturation, and endings, which are affected by sample edges effects. 86
- 5.3 Laboratory relationship between cumulative fractured area and cumulative radiated energy. The colored dots represent the measurements of sensors at different azimuths. The black dots are the mean estimate of four sensor-specific values. The cumulative radiated energy is proportional to the fracture area for individual fracture events. The gray line is the best-fit linear regression using the averaged measurements (black dots). (a) High fluid viscosity experiment: the first three data points come from three clear events denoted as pink stars shown in the schematic diagram. They are recognized as the first three events in the kymograph panel in Fig. 5.2a. The late two measurements represent the signals after the third event, which are indistinguishable and divided into two groups arbitrarily. (b) Low fluid viscosity experiment: the AE cumulative energy is calculated directly on the continuous vibrations. 88
- 5.4 Illustration of tremor generation mechanisms in an overpressurized fault zone in subduction. Two primary mechanisms may influence tremor dynamics. The first, the single-force mechanism^{??}, proposes that tremors occur during transient fluid pressures as barriers unclog and fluid flows, which are indicated by blue arrows in the lower left panel. The second mechanism proposed in this study suggests a mixed mode of shear and tensile fracture resulting from increased fluid pressure. The black arrows represent seismic source body forces in the lower right panel. The tensile-shear model is also described by Sibson[?]. Evidence of these crack-induced tremors can be seen in many exhumed rock samples^{??}. Alternative models also include jamming of granular media^{??}. 90

5.5	Example Tremor Swarm in Cascadia. (a) Individual tremor locations during one slow slip event that started March 10th, 2019, color-coded by occurrence time since origin. (b) Bursty tremors. The upper panel shows the southward propagation of centroid location along the latitude over time since its origin. Each dot denotes the relative distance in kilometers of centroid latitude of tremors within 2-hour intervals relative to the original location. The error bar indicates the latitude range (in kilometers) of each interval. The blue curve represents the averaged or smooth propagation over time. The lower panel shows the curves of the number and summed tremor energy in 2-hour intervals binned in the same way as in the upper panel. The yellow patches highlight the quiescence periods, in which no tremor is observed. We find the episodicity of tremors similar to laboratory observations in Fig. 5.2. 91	91
5.6	Field relationship between cumulative tremor area and cumulative radiated energy. The relationship between cumulative tremor area and cumulative radiated energy for 500 inferred tremor swarms that took place in the Cascadia region from 2017 to 2023 (adapted from ²). The color of the circles indicates the centroid latitude of each swarm, with northern to southern events shaded from red to blue, and highlighting the more energetic tremors in the north ² . (a) The black line highlights the example event in Fig. 5.5a. Swarms systematically radiate greater energies from north to south. (b) The slopes of the curves in (a), measured as the linear regression coefficients, are displayed versus the radiated energy. We find a median regression coefficient, n , as the power for $E A^n$ to be 1.16 for the 500 swarms studied, which lies between the laboratory values of $n = 1.19$ and $n = 1.07$ for the high- and low-viscosity experiments, respectively. The power exponent is more variable for the small swarm relative to larger ones. We refer to Fig. D.15 for more details. 93	93
6.1	Diagram of experimental apparatus and measurements. (a) Schematic fault valve. (b) Simplified schematic laboratory specimen. (c) Experimental apparatus design. 100	100
6.2	(a) Recorded injection pressure and volume history at the injection point, with an injection rate of 2 ml/min. Attention is drawn to the highlighted segments. The varied colors represent distinct fluid pressure (P_f) fluctuations and the corresponding physical phenomena, elaborated in the subsequent image. Three repeated patterns are observed, each marked by notable increases and decreases in P_f . (b) Kymograph detailing pixel variations correlated with movements of fluid and glass beads. This visualization is created by extracting a central vertical slice from video frames, after subtracting a background image, which is displayed on the right for reference. Differentiating the intertwined effects of fluid dynamics and bead displacement on light intensity and pixel distribution is complex. However, through manual review of the animation, we have identified and labeled the specific physical processes causing changes in pixel intensity and position. The associated video is accessible here. Note that due to the video's large size, it is segmented, analyzed in parts, and then reassembled. The background image used for reference is rotated by 90 degrees. 103	103

6.3	During the hydromechanical process, we captured continuous acoustic data (depicted in black) at various injection rates, which are displayed across three panels: 2 ml/min (upper), 3 ml/min (middle), and 4 ml/min (lower). Superimposed on these recordings are the instantaneous envelopes (in red) and the cumulative envelopes (in green). Arrows are used to indicate the occurrence of different physical processes.	105
7.1	An ensemble estimation-based framework for earthquake detection and phase picking. The main components include base predictions at broadband or multiple frequency bands (e.g., filtered data) and ensemble estimation by either statistics-, coherence-, or a learning-based approach. Note that only EqT-based pretrained models are tested.	112
7.2	Example of ELEP workflow. (a) Broadband prediction scheme. (b) Multiband prediction scheme. (c) Example three-components seismogram. (d) Time-frequency signals decomposed by a recursive multiband filter. (e) The predictions of these filtered signals are obtained and compared between broadband and multiband predictions, denoted as solid and dashed lines, respectively. Detection, P-phase, and S-phase picking are represented as orange, blue, and green lines, respectively.	114
7.3	Example of ensemble estimation for three aggregation methods. (a) P-phase probabilities and (b) S-phase probability functions predicted by 18 pretrained models. Deep colors denote high probabilities. (c) Maximum-based ensemble estimation from six EqT predictions (thin lines) for P-phase and S-phase probability functions (thick lines). (d) Semblance-based ensemble estimation from six EqT predictions for P-phase and S-phase probability functions. (e) Learning-based ensemble estimation for P- and S-phase probability functions. Note that only six 20-seconds EqT-based predictions are used as the input here.	116
7.4	Within-domain broadband and multiband predictions on the INSTANCE test dataset using the eqt-instance model for (a) detection, (b) P-phase picks, and (c) S-phase picks. Detection and picking rates are gray (broadband predictions) and orange (multiband predictions) markers. P and S pick mean errors (e.g., bias) are shown in blue (P broadband), green (S broadband), and red (P and S multiband). The error bars are the standard deviations of picking residuals. Note that the shifts of pick-error markers in (b, c) are deliberately made to avoid overlapping and for clear visualization.	123
7.5	Cross-domain broadband and multiband predictions on the INSTANCE test dataset for six EqT-based pretrained models for (a) detection, (b) P-phase picking, and (c) S-phase picking. The pretrained models are denoted by different symbols. Gray symbols illustrate detection and pick rates from broadband predictions. The orange symbols display the corresponding detection and pick rates from multiband predictions. Blue and green symbols indicate standard picking errors of broadband P- and S-phase picking residuals, respectively. The red symbols represent standard multiband-predicted errors for comparison.	124

7.6	Effect of probability thresholds on (a-c) broadband and (d-f) multiband predictions on the INSTANCE. The symbols stand for different pretrained models. The orange and gray symbols represent the detection rate and P- or S-phase pick rate, respectively. The blue and green symbols denote the mean P- and S-phase pick errors in seconds, respectively. The error bars indicate the standard deviations associated with the mean pick errors.	125
7.7	Example predictions with pretrained models and ensemble estimations on four multi-channel seismograms from the INSTANCE dataset. The dashed lines mark the labeled (or true) arrival times of the P- and S-phases. The solid lines denote the arrival times estimated by the learnable ensemble, which are only shown on the E components but calculated over the three channels. The lower plots show the probability distributions annotated with their respective base models. The upper probability distributions are annotated from the estimations by two simple ensemble methods (maximum and semblance). The P-phase and S-phase probabilities are plotted in blue and green, respectively.	127
7.8	Estimations of three ensemble methods on the INSTANCE dataset. (a, b) Broadband predictions. (c, d) Multiband predictions. Circles and squares represent P- and S-phases, respectively. Gray symbols indicate pick rates, while colored symbols show the standard deviation of the residuals for different methods. The predictions of the learnable ensemble are performed on the results of the maximum ensemble and semblance ensemble.	129
7.9	Transfer against ensemble learning in the PNW. Results from three ensemble estimation methods against test data SNR of the PNW data. (a, b) broadband workflows, (c, d) multiband workflows. Transfer learning of the eqt-pnw model is compared to (a, c) maximum ensemble and maximum + learnable ensemble and (b, d) semblance ensemble and semblance + learnable ensemble.	131
7.10	Effects of training data size on the performance of learnable ensemble predictions. (a, b) Broadband and multiband learnable ensemble predictions for INSTANCE dataset. (c, d) Broadband and multiband learnable ensemble predictions for PNW dataset.	133
7.11	Ensemble learning on OBS benchmark data. Histograms of P and S pick errors between the manual picks and estimated picks from maximum ensemble (a, c) and semblance ensemble (b, d). The ensemble estimations are performed upon multiband predictions. The outlier fraction is the fraction of outliers with pick errors larger than 1s. MAE, MAD, and RMSE denote mean absolute error, median absolute deviation, and root mean square error, respectively. The vertical gray and red dashed lines represent the median and mean pick errors, respectively. Both median and mean values are influenced by uneven errors.	135

7.12	Prediction results on 4-hour continuous seismograms recorded by borehole broadband seismometer at PB.B204. Prediction on the waveforms shown on the data itself. The first panel is continuous time series for the three channels. The second panel shows the probability results predicted by six pretrained models individually. The third panel shows the probability results predicted by the transfer-learned model (eqt-pnw). The fourth and fifth panels show the results estimated from the maximum ensemble and semblance ensemble, respectively. The picking results of the P and S-phases are marked in blue and green. The dashed red lines represent the probability or coherence threshold for the picking identification.	136
7.13	Computational efficiency on 4-hour continuous seismograms recorded by borehole broadband seismometer at PB.B204. Computation cost in milliseconds (ms) of all steps processing the continuous waveforms in (a). These steps include preprocessing (left side), six successive pretrained predictions, one transfer-learned prediction, and two ensemble estimations (right side). The comparison between CPU and GPU is also performed for prediction steps. The preprocessing and ensemble steps are only conducted on the CPU.	137
A.1	Workflow diagram of WTS and WTDTW methods. <i>CWT</i> refers to the continuous wavelet transform and <i>R</i> refers to its real part. The steps are: 1. Apply the CWT to both reference $X(t)$ and current $Y(t)$ time series; 2. at each frequency f_j , extract the real part of the CWT, $R_X(t)$ and $R_Y(t)$; 3. implement TS or DTW on $R_X(t)$ and $R_Y(t)$ to extract $dv/v(f_j)$; 3. repeat step 2 at all frequencies.	152
A.2	Absolute measurement uncertainties of the velocity perturbation results shown in Fig.2.2 for (a) time-domain methods, (b) frequency- and wavelet-domain methods with narrow-frequency bands, and (c) wavelet-domain methods at all frequencies.	153
A.3	Waveforms simulated in a layer-over-halfspace heterogeneous medium using SPEC2D solver under the zero-offset setting as in Fig.2.5. (a) Reference waveform (blue) and perturbed waveform (red) of the whole space (top), the shallow layer (middle) and the deep layer (bottom). The correlation coefficients between the perturbed and reference are 0.9901, 0.9971, and 0.9971, respectively. (b) Fourier amplitude spectra of (a) bottom panel.	154
A.4	dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.5 of the main text, zero-offset geometry, heterogeneous layer-over-halfspace medium, and uniform perturbation of the medium.	155
A.5	dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.5 of the main text, zero-offset geometry, heterogeneous layer-over-halfspace medium, and a perturbation in the shallow layer of the medium.	156
A.6	dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.5 of the main text, zero-offset geometry, heterogeneous layer-over-halfspace medium, and a perturbation in the deep layer of the medium.	157

A.7	Waveforms from the layer-over-halfspace heterogeneous medium with distant-offset source-receiver geometry to support Fig.2.6 of the main manuscript. (a) Reference waveform (blue) and perturbed waveform (red) of the whole space (top), the shallow layer (middle) and the deep layer (bottom). The correlation coefficients for three pairs of waveforms are 0.9635, 0.9692, and 0.9996, respectively. (b) Fourier amplitude spectra of (a) bottom panel.	158
A.8	dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.6 of the main text, distant-offset geometry, heterogeneous layer-over-halfspace medium, and uniform perturbation of the medium.	159
A.9	dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.6 of the main text, distant-offset geometry, heterogeneous layer-over-halfspace medium, and perturbation in the shallow layer of the medium.	160
A.10	dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.6 of the main text, distant-offset geometry, heterogeneous layer-over-halfspace medium, and perturbation in the deep layer of the medium.	161
A.11	dv/v measured for each depth of the perturbation layer in the zero-offset source-receiver geometry for all methods. (a) Time-domain WCC, TS, DTW, at various frequency bands (b) MWCS, (c) WCS, (d) WTS, (e) WTDTW, and over the spectrum of frequencies for (f) WCS, (g) WTS, and (h) WTDTW. Results are repeated in Fig.2.7(a, c) of the main manuscript.	162
A.12	dv/v measured for each depth of the perturbation layer in the distant-offset source-receiver geometry for all methods. (a) Time-domain WCC, TS, DTW, at various frequency bands (b) MWCS, (c) WCS, (d) WTS, (e) WTDTW, and over the spectrum of frequencies for (f) WCS, (g) WTS, and (h) WTDTW. Results are repeated in Fig.2.7(b, d) of the main manuscript.	163
A.13	Comparison between dv/v spectra obtained with WTS and narrow bandpassed TS (Gaussian filter), including (a) perturbation in the whole space, (b) perturbation only in the shallow layer, and (c) perturbation only in the deep layer as the layered half-space distant-offset case, which is shown as Fig.2.5 in Section 3.2. The medium and perturbation are detailed in Table B1. Briefly, the layer is perturbed at 10-km depth and by 0.1%. The black dashed line is a reference perturbation value rather than true perturbations at individual frequencies.	164
A.14	Effect of dissimilar dominant frequency between the reference and current state on the estimates of dv/v . We perform simulations in the same medium as in Section 3.1 (scattering halfspace – uniform change in velocity – zero-offset response). The dominant frequency of the reference simulation is 1 Hz. The figure organization is similar to Fig.A.12. The black dashed line is the true velocity perturbation. We perform 31 simulations in the perturbed medium with various source spectra by changing the dominant frequency of current source spectra from 0.5 Hz to 2.0 Hz with an interval of 0.05 Hz.	165

A.15	Same as Fig.A.14 but varying the lapse time (start time of the window) of the coda waveform between 5 s and 45 s.	166
A.16	Same as Fig.A.14 but varying the window length of the selected coda window between 15 s and 55 s.	167
A.17	Same as Fig.A.14 but varying the noise level. Gaussian noise is added on top of the simulated wavefields with levels varying between 0% and 30% of the maximum amplitude of the coda waveform.	168
A.18	Depth sensitivities of scattered body and surface waves at three lapse times (i.e. 55 s, 115 s, and 185 s) for the distant-offset setting. Left three panels (a, c, e) are the dv/v results from TS; right three panels (b, d, f) are the dv/v results from WTS. The experiment is the same to the distant-offset case of Section 4 but at different time lags. Note that the spurious results at frequencies around 0.1 Hz are attributed to low energy levels (spectral trough) at this frequency.	169
A.19	Depth sensitivities of scattered body (bulk) waves at three lapse times (i.e. 55 s, 115 s, and 185 s) for the distant-offset setting. Left three panels (a, c, e) are the dv/v results from TS; right three panels (b, d, f) are the dv/v results from WTS. The experiment is the same as Fig.A.18 but replacing the free surface with an absorbing boundary in order to only simulate scattered body waves.	170
B.1	Experimental apparatus and specimen.	173
B.2	(a) Calibration of wavespeeds of P- and S-waves in the hard resin. (b) Dispersion analysis of surface waves (outlined by purple lines in (a)).	173
B.3	Acoustic emission waveforms collected in the soft resin. (a) An active measurement with a fixed distance of 30 mm, (b) a shot gather over different distances.	174
B.4	(a) Spectra of AE waveform for the first case with only water-level changes within a static granular medium. (b) smoothed spectra of the same waveforms.	174
B.5	(a) Spectra of AE waveform for the cases with stress loading on the well-bonded specimen under various saturation states: un-, partially-, and fully-saturation. (b) smoothed version of same waveforms.	175
B.6	(a) Spectra of AE waveform for the cases with stress loading on the unbonded specimen under various saturation states: un-, partially-, and fully-saturation. (b) smoothed version of same waveforms.	176
B.7	Apparent shear-wave velocity (V_s) and its change for individual water heights, which correspond to the lab measurements as observed in Fig.3.1. Apparent V_s change is directly calculated from the perturbation of apparent V_s compared to the dry case. To determine the average V_s change for the whole space, we adjust the apparent V_s change by a spatial averaging factor, which are obtained from the volume and velocity of soft and rigid components, as Eq.3.4.	177
C.1	Histograms in amount of earthquake detections by (a) borehole stations and (b) surface stations.	179

C.2	(a-c) Comparison between NonLinLoc earthquake locations (yellow squares) in 3D models and PNSN catalogued earthquake locations (gray squares) in 1D. (d) Histograms of P-phase arrival-time residuals between machine learning picks and manual picks. (e) Histograms of NonLinLoc traveltime misfits of earthquake locations. (f) Histograms of earthquake location difference between NonLinLoc and PNSN.	179
C.3	(a-c) NonLinLoc locations of around 25,000 earthquakes detected by ELEP between 2007-2022. (d-f) PNSN catalogued locations of around 17,500 earthquakes since 1980, among them approximately 10,000 earthquakes were detected between 2004-2007 and around 7,500 earthquakes occurred between 2007-2022. Note: NonLinLoc locations are performed in 3D velocity models, the PNSN locations are obtained by using HypoInverse in 1D velocity models. ELEP detections are three times more than PNSN catalog.	180
D.1	Experimental apparatus. A high-pressure pump drives the fracturing fluid into the sample. A pressure transducer measures the fluid pressure as the experiment proceeds. AE sensors are placed on each of four sample quadrants and record the AE signal generated by the fractures. A high-speed camera is filming the growing fracture from the bottom of the sample, when triggered by the first AE signal. The details of the experimental apparatus components are shown in the photos in Fig.D.2.	187
D.2	Main components of the experimental apparatus. a) A high-speed camera (Phantom TMX5010) films the fracture propagation. A blue LED ring is placed around the sample to highlight the contrast between dyed fluids and fracture fronts. b) Four AE Glaser-type sensors (KRNBB-PC) listen to the fracture propagation. c) A high-pressure pump (Teledyne Isco 65D) injects the fluid at a constant flow rate of $Q=0.3$ ml/min. . . .	188
D.3	Pressure and injected volume during the experiment. The fluid injected volume (blue curves) increases the fluid pressure (red curves) until the sample ruptures (shaded blue region). Both fluid viscosity experiments are shown: 1cP in dashed lines and 800 cP in solid lines.	189
D.4	Kymograph and AE signals. (a-top) kymograph of the fracture propagation during the entire experiment in the high-viscosity case. The stick-break events are clearly visible by the increase and stagnation in the fracture radius. (a-bottom) Overall, AE signals are recorded during the entire experiment, occurring as bursts and related to breaking events in the kymograph and the pauses related to the stick events. (b) similar to (a) but for the low-viscosity experiment. In this case, the stick-break events are not clearly resolvable in the kymograph. Still, the AE signal shows interesting bursts in their signals, indicating very small amplitude stick-break instability events. Note: partial data have been shown in Fig. 5.2.	190

- D.5 Comparison between AE raw signals and intensity rates, which include raw (black), filtered (blue), and enveloped (red) ones for (a) high-viscosity and (b) low-viscosity experiments. Note: AE signals are exactly the same as the ones in Fig. D.4. Partial data have been shown in Fig. 5.2. We refer to the intensity measurements in the video description. 191
- D.6 Spectrograms of AE signals. We obtain them by performing the short-time Fourier transform on individual raw AE signals of the water experiment as the ones in Fig.D.4b. Note: the signals' unit is voltage and is nearly proportional to displacement for an almost flat instrumental response function and a constant signal amplification factor. 192
- D.7 Glass capillary test for Green's function calibration. (a) Calibration setup in which a glass capillary breaks at the impact point and generates an impact. Four AE sensors (A, B, C, and D) are deployed in a line at intervals of 15 mm from each other, recording the same single event resulting from the capillary break. The block used is made of the same 3D printed PMMA material as for the fracturing experiment and is dimensioned (140x140x50 mm) to avoid arrivals of any P-wave signal reflected from the boundary before the first arrival from the capillary break. (b) AE signals are recorded at the four sensors (successively from the impact point A, B, C, D) resulting from the capillary break. With wave types labeled, The signals are labeled with the arrivals of different waves, the measured P-wave speed is approximately 2300 m/s and estimated S- or Rayleigh wave speeds are both about 1000 m/s, because the Poisson's ratio of the material is estimated to 0.34 and as the distance between the source and successive receivers is very small compared to the overall wave speed. 192

- D.8 Fracture event spatial location by AE triangulation. (a) Arrival times are visually picked and annotated with downward arrows on the AE signals. Nucleation begins at T_n , seen as the red dashed line. The arrival time difference is compared to the nucleation times, allowing us to triangulate and determine the spatial location of the nucleation event (b) Differential image, the locations of the sensors (thick colored arrows), fracture front (red dash line), fluid front (solid blue line), and the observed location of the fracture nucleation (red oval). (c) Misfit between theoretical and observed differential travel times between sensors given the radial and azimuthal distances in a polar coordinate. The warm and cold colors denote the small and large misfits, respectively. The red oval location represents the location with the minimum misfit. The location is almost consistent with the nucleation location observed in (b). The velocity calibrated in Fig. D.7 is used for creating a travel timetable for the location search. The misfit function used for the grid search method can be seen as Eq.5.2.2 in the main text. (d) Four snapshots of differential images exhibit the propagation and arrest stages during a single tangential fracture. The third image is observed as the secondary strong fracturing stage, which radiates strong signals seen as the secondary waveforms outlined by the dashed box in (a). After estimating the travel time ($tt=56.6\mu s$) after $T_n+70\mu s$, we postulate that the dominant Rayleigh-wave signal recorded by the SW sensor is associated with the third image. In contrast, the clear arrivals at the NW sensor are unrecognized due to the short distance. This indicates the entire fracturing process has not uniformly occurred, which probably results in variations in the fluid pressure. 193
- D.9 Low fluid-viscosity (water) experiment images and AE signals. (a) Three images after at $T=2.4$ ms, at $T+50\mu s$, and at $T+80\mu s$. In the low fluid viscosity case, multiple nucleation events nucleate nearly simultaneously (tracked by the colored points from differential images) and show interesting propagating behaviors (shown by the arrows of multiple colors corresponding to their sources). (b) The four raw AE signals don't seem to give a straightforward indication of the nature of the fracture behavior. Indeed, as many events occur very close in time, their signals overlap. The change in the fracture area behind the fracture front (white or gray pixels) can be easily calculated over time. 194
- D.10 AE energy and rupture area. (a) The scaling of AE energy rate and fast transverse fracture area (small red areas in Fig.5.1b) using raw AE signals. The radiated energy is uncorrected for geometric spreading and attenuation (see Equation 2 in the main text). (b) Sensitivity analysis of attenuation factors with respect to linear regression coefficients. For both high- and low-viscosity experiments, the attenuation coefficient is close to 20 and has negligible effects on the scaling relation. Therefore, the scaling relation is strongly influenced by the effect of geometric spreading. 194

D.11	Drop Ball tests: AE combined with high-speed imaging. (a) Experimental setup of the drop ball test (inner upper picture) and zoom-in image of falling steel ball ($r=2\text{mm}$) captured by the high-speed camera. We place four AE sensors in a line to record the AE signals from the ball impacts similar to the capillary break tests (Fig.D.7). The ball falls and hits the sample at an incoming velocity and bounces back up with an outgoing velocity. (b) Incoming and outgoing velocities can be measured using high-speed imaging using a Kymograph (distance, time representation). The red and yellow lines represent the incoming and outgoing trajectories, and their slope equals the velocities (c). We perform experiments at different heights to observe the fluctuation of the incoming and outgoing velocities. The observations are shown and analyzed for the energy calibration in Fig. D.10.	195
D.12	AE Energy calibration using the drop-ball test. (a) The theoretical velocities from potential energy are almost equal to the observed incoming velocity from a high-speed camera (Fig. D.11) as shown by the black line. The coefficient of restitution, R_e , is not linearly decreasing with the height, as shown by the red line. (b) Relation between the potential energy and kinetic energy transfer during ball collision. (c) Relation between the potential energy and AE energy calculated from the integration of squared voltages over time. (d) Relation between the kinetic energy transfer and AE energy. The detailed explanation and process can be found in the supplementary Text 2.	196
D.13	Relations between fluid pressure, fracture-front propagation, and AE signals. (a, c) Interactions between absolute fluid pressure (originally measured at the injection site with a sampling rate of 3.125 MHz) and radial fracture fronts in four directions (SE, NE, NW, and SW). (b, d) Relation between the filtered fluid pressure (20 kHz) and the four AE signals. Note that the upper two panels are obtained from another high-viscosity (800 cP) fluid-injection experiment, and the lower two panels are from another low-viscosity (1 cP) fluid injection experiment. Both experiments are similar to the experiments used in the main text. However, the fluid pressure is, unfortunately, not measured by a high-resolution (i.e., high sampling rate) pressure transducer.	197
D.14	Spatio-temporal Tectonic Tremors. (a) Contours of tremor source area in each of the first 10 days of the 2007 tremor swarm occurred in the Northern Cascadia. Each area is obtained by contouring the epicenters of the daily occurring tremors. (b) The cumulative radiated energy is proportional to the cumulative tremor area for each of the four Northern Cascadia tremor swarms. Each square denotes the daily accumulated tremor area and accumulated radiated seismic energy?	198

D.15 Analysis of the regression coefficients between logarithmic (base 10) cumulative radiated energy of tremors and area in Fig.5.6. (a) The spatial distribution of these regression coefficients for 500 tremor swarms in the Cascadia region from 2017 to 2023. Each coefficient is obtained through the linear regression analysis between the logarithmic (base 10) cumulative tremor energy and area for each tremor swarm. The resulting coefficient is then employed to uniformly color-code all tremors within the specific tremor swarm. These tremor swarms exhibit varying logarithmic powers associated with tremor areas. We also display them in histograms for statistical analysis, as shown in b. The mean, median, and standard deviation of these logarithmic powers, or regression coefficients, are 1.16, 1.15, and 0.32, respectively. It is evident that most tremor swarms display a near-linear relationship between tremor radiated energy and tremor area. 199

THIS IS THE DEDICATION.

Acknowledgments

THE pursuit of a Ph.D. is fundamentally an exercise in cultivating patience, resilience, and specialized expertise. This practice could not have been completed without the guidance, encouragement, companionship, and support received along the way.

I've been incredibly lucky to have Marine Denolle as my primary advisor. Marine's dynamic personality and scholarly brilliance have profoundly impacted me. Her enthusiasm, positivity, and attention to detail have significantly shaped my approach to research and communication. The thrill and passion I feel for research have been nurtured by her insightful guidance. I am also fortunate to have John Shaw and David Weitz as my mentors. After the marine's move to UW, John and Dave took care of me a lot. They are always responsive to my inquiries. John, with his kindness, patience, and constant support, has always been available for discussions about my research progress, fostering a nurturing environment for my growth. Dave is always critical, supportive, and generous in my experimental studies. His keen insights have fundamentally shaped my approach to experimentation. Besides, I extend my heartfelt gratitude to Jie Zhang, Youzuo Lin, Xiaofeng Jia, Jinfeng Ma, Wenhuan Kuang, and Yiqiao Song for their unwavering support and encouragement. Moreover, I am thankful to my committee members, Brendan Meade, Nori Nakata, and Eli Tziperman, for their insightful feedback and constructive suggestions.

The communities at EPS and SEAS have been a source of endless inspiration, filled with incredible individuals I've had the fortune to collaborate with and learn from. Special thanks to fellow members in Harvard and UW seismology groups: Chengxin Jiang, Tim Clements, Jiuxun Yin, Zhuo Yang, Stephiane Olinger, Zhitu Ma, Xiaotao Yang, Qibin Shi, Yiyu Ni, Akash Kharita, Manuela Koepfli, Ethan Williams, and Kuan-Fu Feng. Also, many thanks to Dave's group members: Thomas Cochard, Yujing Du, Tianlei He, Zhencheng (Max) Jiang, Hong Li, Jianping Xu, Adel Djellouli, Ilya Svetlizky, Gabriele Albertini, Lizhi Xiao, Yang Wang, Yi Yang, Wenyun Wang, Jie Wang, Sijie Sun, and Negar Nazari. I would also like to thank non-academia staff: Aimee Smith, Campbell Halligan, Esther James, Kathleen McCloskey, Maryorie Grande, Matthew Zahnzinger, and Leah Burrows, for efficient assistance and inclusive environment.

To my peers, teachers, and the many friends I've made along the way, your kindness, support, and encouragement have enriched my experience immeasurably. Many thanks to Mingzhen Yu, Hairuo Fu, Zeyuan Hu, Zhongxing Chen, Yaowei Li, Hongwei Li, Boer Zhang, Jack Bruno, Kristen Chiana, Ju Chulakadabba, Yichen Geng, Jesse Gu, Junjie Dong, Brook Runyon, Qiyu Song, Zheng

Gong, Chris Parendo, Li Zeng, Lois Tang, Natasha Toghradjian, Robert Welch, Nathanael Wong, Lydia Bailey, Lluís Saló-Salgado, Andreas Plesch, Ansar Lemon, Marisa Borreggine, and Ethan Manninen, for all the uplifting pep talks, support, and encouragement received in the past. My further thanks extend to Ruben Juanes, William Frank, Camilla Cattania, Jared Bryan, Hillary Chang, Douglas Miller, Yudong Sun, Qingyu Wang, Hongrui Qiu, Lei Qin, Xin Cui, Enrico Milanese, Joan Gomberg, Qiuming Chen, Jiashun Hu, Xiaotong Zhang, Zhizhen Zhong, Yuqin (Sophia) Duan, Liang Yang, Haoxue Fan, Dongyuan Li, Connie Wang, Xin Tang, and Yawei Huang.

Above all, my deepest appreciation goes to my family—my parents, grandparents, uncles, aunts—for their endless love and emotional support. Their understanding has been the bedrock of my commitment to this journey. And to Jingyi (Rena) Ren, who came into my life during challenging times, your love, support, and warmth have been the light in my darkest days. With you by my side, this journey has not only been possible but filled with joy and fulfillment.

1

Introduction and Motivation

1.1 GENERAL INTRODUCTION

The shallow Earth is a dynamic membrane of the Earth's crust. It is subject to external forces such as transient tectonic stresses, continuously fluctuating hydrological and magmatic conditions, thermoelastic expansion and contraction, and tidal stresses, among other forcing mechanisms. These forces act across a wide range of length and time scales. Earthquake ruptures induce material dam-

age, and in turn, the structural changes affect the evolution of earthquake ruptures: Earth properties and earthquake hazards are a tightly coupled dynamical system. This coupling is often ignored in both observational and theoretical studies of earthquake phenomena. Therefore, studying the shallow Earth's materials' dynamical properties at short- and long-term time scales will improve our understanding of earthquake initiation and damage, volcanic magma activities, natural resource management, for example, tracking the groundwater, as shown in Fig. 1.1A.

The rheology of materials is a theoretical framework that relates stresses and strains that materials undergo, and it depends on material composition and deformation mechanisms[?]. Seismic properties, such as elastic wavespeed and anelastic attenuation, depend upon the material rheology[?]. Laboratory experiments^{???} and seismological observations^{???} have shown that perturbations in acoustic and elastic velocities are correlated with surrounding stress perturbations or with observed material damage, suggesting that the monitoring of seismic wavespeeds allows for the monitoring of the properties of elastic rheology. Therefore, measuring changes in time and space of these properties would enable a rapid assessment of near-surface dynamics. Time-lapse seismic velocity has become a popular tool to monitor volcanic activities[?], earthquake co-seismic damage[?], and environmental effects, such as atmospheric conditions and subsurface hydrology^{???}. Current advances^{??} are undertaken to probe and image these changes in time and space, especially the depth. Additionally, only a few studies have attempted to monitor the seismic attenuation^{???}. Because seismic attenuation is rather difficult to measure, only a few studies have measured temporal changes in attenuation^{??}.

The underlying mechanism during the earthquake nucleation is closely related to the change of the material rheological properties. There are several lines of evidence that material properties change before, during, and after earthquake ruptures. Models that use continuum damage mechanics predict a coseismic or co-shaking reduction in rigidity and, therefore, seismic velocity[?]. Observations of a reduction in seismic velocity during or due to strong shaking are numerous^{?????????}.

The evolution of the seismic velocities as it recovers from the co-seismic damage behaves in a logarithmic fashion (see reference here-in) as observed in natural earthquakes. The only reliable precursory signal observed before a major earthquake was made by Niu et al.² at the Parkfield SAFOD drill site. Most importantly, slip events' laboratory measurements predict cyclic variations in seismic velocities and attenuation during the earthquake cycles^{2,2}. Nevertheless, this has barely been observed in the past; whether this is an issue of detectability or a misunderstanding of earthquake physics is unresolved. Similar to the tectonic stress, magmatic inflation², tidal², and atmospheric² forces influences material properties of Earth's shallow crust. This emphasizes the importance of distinguishing their contributions through understanding and monitoring of their changes in seismic properties.

The fluid has a dramatic influence on the seismic properties of the material and its mechanical behavior. Fluctuations in groundwater aquifers and water table levels have impacted near-surface seismic velocities^{2,2}. Water is injected at a high rate into the ground to fracture rocks to enhance oil and geothermal energy production. To be disposed of or leaked, they have caused significant effect by inducing seismicity in intra-continental areas^{2,2,2,2}. United States induced seismicity may provide opportunities to make such observations². The fluid-related aseismic slip or creep often accompanies the seismic slip². The slow slip plays a rate-strengthening role and interacts with the weakening parts during the fracturing nucleation^{2,2,2,2}. Laboratory experiments provide tremendous opportunities to explore these interactions and overcome the limitations of the rare observations in real Earth.

My overall research objective is mainly to explore two fundamental questions: **How does fluid and/or deformation influence the seismoacoustic properties of porous media? How does fluid modulate deformation, fracturing, and seismoacoustic behaviors?** Centered around the two questions, I develop new experimental apparatus and monitoring methods aimed at elucidating empirical relationships and fundamental phenomena. These insights are then applied to interpret

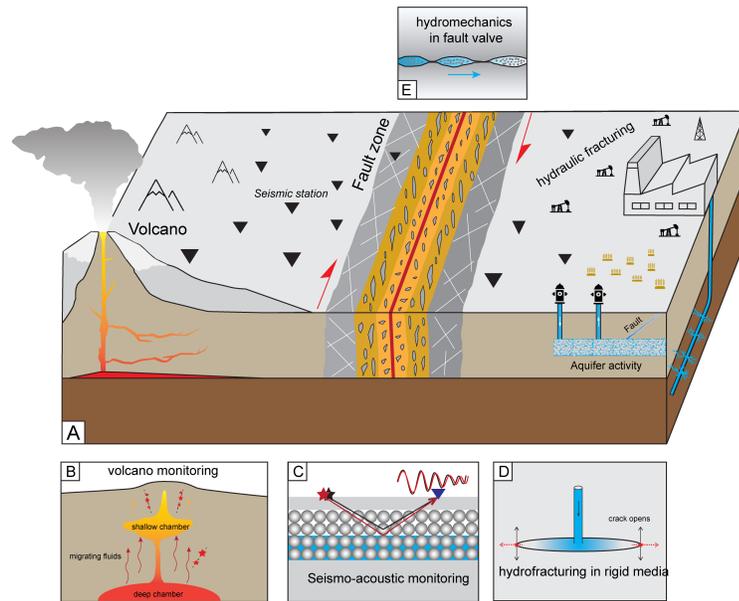


Figure 1.1: Overview of research themes. (A) Schematic representations of target research contexts. (B) Seismic monitoring of volcanic activities. (C) Seismo-acoustic monitoring of reservoir mechanical properties. (D) Laboratory hydrofracturing within rigid media. (E) Laboratory hydromechanics in fault-valve media.

observations in specific contexts such as volcanoes, fault zones, aquifers, and hydraulic fracturing scenarios, as illustrated in Fig. 1.1A.

1.2 FLUID AND/OR DEFORMATION INFLUENCE THE SEISMOACOUSTIC PROPERTIES

The seismic properties introduced earlier, such as wavespeed and attenuation, are influenced by factors like fluid presence, stress or strain, and temperature. The extent of this influence varies depending on the material and its environmental context. Numerous studies have explored these properties in diverse settings. The current efforts are on achieving 4D mapping (3D space combined with time) of seismic properties through the use of advanced technologies such as dense seismic arrays^{??} and Distributed Acoustic Sensing (DAS) systems^{??}, along with sophisticated monitoring techniques^{??}. Except for accurate estimation in spatial map of velocity changes, other challenges

include: 1) The depth of velocity changes remain challenging to resolve. Frequency-dependent data offer a potential solution for better understanding depth-related variations.^{??} ; 2) there is a notable gap in experimental research that specifically addresses the interpretation of frequency-dependent effects, particularly in terms of attenuation. My work aims to develop methodologies and provide insights to address these complex issues.

In Chapter 2, I delve into the depth sensitivity of velocity changes through numerical simulations and innovative frequency-dependent measurements. In these simulations, velocity changes serve as proxies for alterations in the mechanical properties of underground media, facilitating a concentrated examination of wavefield dynamics and analysis. I introduce two new approaches within the wavelet domain: Wavelet Transform Stretching (WTS) and Wavelet Transform Dynamic Time Warping (WTDTW), which I then evaluate and integrate with traditional approaches across the time, frequency, and wavelet domains to assess velocity changes. The frequency-dependent velocity changes we calculate are both precise and robust, enabling us to map out velocity-change spectra and their depth-related variations. The continuous measurement of frequency-based velocity changes significantly enhances our ability to deduce the depths at which these changes occur.

In Chapter 3, the numerical simulations discussed previously do not account for physical conditions like strain deformation and fluid saturation. To address this, we employ controlled active acoustic monitoring and 3D-printed granular media to study changes in wave speed and amplitudes under varying states of strain deformation and fluid saturation. Initially, I investigate how the wave speed and amplitude spectra vary with changes in water-table height within dry granular media, as illustrated in Fig. 1.1C. Subsequently, I explore the spectral responses of granular media with firmly bonded grains under dry, partially-saturated, and fully-saturated conditions. Furthermore, I examine the spectral behavior of unbonded granular media across the same moisture conditions. By simulating field-like scenarios in the laboratory, our goal is to comprehend how wave speed and attenuation respond to different static physical processes, thereby enhancing our ability to interpret

these mechanisms and estimate the depth of changes within field contexts.

In Chapter 4, I aim to explore the possibility of probing deep volcano's magmatic activity, as displayed in Fig. 1.1B. Volcanoes present a unique challenge due to internal phenomena such as magma inflation and external environmental changes. Traditional seismic interferometry techniques, whether receiver-based or inter-receiver, have shown limitations in discerning these internal changes due to interference from superficial perturbations. Here, I introduce the use of inter-source interferometry to derive the empirical Green's function between pairs of repeating earthquakes. Utilizing the catalog of repeating earthquakes identified at Mount St. Helens, I apply the phase picking method developed in Chapter 7 to pinpoint seismic phase arrivals, locate repeating earthquake sequences, and establish a workflow for inter-source interferometry. I present a case study where velocity changes between a pair of repeating earthquakes are successfully detected, showcasing the method's potential applicability to all repeating earthquake pairs. Given the distribution of these repeaters across various depths, this approach holds promise for shedding light on the dynamics of the volcanic magma plumbing system.

1.3 FLUID MODULATES DEFORMATION, FRACTURING, AND SEISMOACOUSTIC BEHAVIORS

Fluids are almost everywhere, from shallow crust to deep subduction zone. They play a pivotal role in geomechanics through their dehydration or injection and extraction or eruption processes. Overpressurized fluids can significantly interact with the material matrix and fractures/faults, potentially triggering induced seismicity and tectonic earthquakes. Understanding and monitoring these fluid-induced deformations and fractures are crucial for assessing seismic hazards and enhancing production efficiency. This subject has garnered considerable attention over several decades. Building upon the foundation laid by previous research, I aim to contribute to this field at the following aspects.

In Chapter 5, I am intrigued by the hydrofracturing dynamics due to its critical role in deci-

phering fluid-induced fracturing behaviors across a spectrum of scenarios, from deep slow slip and tremor zones to the processes of hydraulic fracturing for shale gas production. As the hydrofracturing scenario displayed in Fig. 1.1D, one novel experiment is developed and performed by utilizing advanced imaging and acoustic techniques. This experiment facilitates the study of fracturing patterns driven by fluid viscosity and pressure. By injecting pressurized fluid into a 3D-printed rigid medium, I observe the initial rupture and subsequent rapid transverse expansion of the fracture front, reaching velocities comparable to Rayleigh-wave speeds. This process results in a distinctive stick-break pattern in the fracture's forward or radial progression, characterized by the bursty energy of acoustic signals. Drawing inspiration from these minute episodic events, I delve into the tectonic tremor swarms observed in the Cascadia region between 2017 and 2023. I compare the laboratory and field observations in terms of seismic/acoustic energy evolution and earthquake/fracture migration. I aim to highlight the role of hydrofractures in regulating slow earthquakes.

In Chapter 6, I am further intrigued by the hydromechanics within porous media, where fluid movement plays a crucial role in fault zone instability and the dynamics of fluid entrapment and leakage. The fault zone is often conceptualized as a fault-valve system characterized by variable permeability, illustrated Fig. 1.1E. Within this context, compaction and dilatancy emerge as pivotal interactions between fluid pressure and the surrounding solid matrix, yet these dynamic processes are predominantly explored through numerical simulations with limited direct empirical evidence. To bridge this gap, I have devised and conducted a novel experiment to scrutinize the hydromechanical behavior in an artificial fault-valve setup during fluid injection. Utilizing high-speed imaging, I capture the intricate fluid-solid interaction and correlate these observations with injection pressure and acoustic data. This comprehensive approach aims to shed light on the underlying mechanisms of fluid migration and its contribution to triggered or induced seismic phenomena.

In Chapter 7, I am also interested in the high-performance seismic processing. Since the fluid-fracture (or fault) interactions can be inferred through seismic signals, fast and accurate seismic

phase picking is highly demanded for earthquake location and fracture identification and characterization. Current deep-learning based seismic phase pickers are incredibly efficient in such task but often limited to specific regions. To overcome this limitation, I have developed an ensemble learning framework, named ELEP, designed to improve the generalizability of these tools, allowing them to be applied to new regions without the need for additional training. I demonstrate the effectiveness of the ELEP in various public datasets and continuous data processing.

This chapter was published as Yuan C, Bryan J, Denolle M. Numerical comparison of time-, frequency- and wavelet-domain methods for coda wave interferometry. Geophysical Journal International. 2021 Aug;226(2):828-46. <https://doi.org/10.1093/gji/ggab140>

2

Numerical comparison of time-, frequency-, and wavelet-domain methods for coda wave interferometry

SUMMARY

Temporal changes in subsurface properties, such as seismic wavespeeds, can be monitored by measuring phase shifts in the coda of two seismic waveforms that share a similar source-receiver path but are recorded at different times. These nearly identical seismic waveforms are usually obtained either from repeated earthquake waveforms or from repeated ambient noise cross-correlations. The five algorithms that are the most popular to measure phase shifts in the coda waves are the Windowed Cross Correlation (WCC), Trace Stretching (TS), Dynamic Time Warping (DTW), Moving Window Cross Spectrum (MWCS), and Wavelet Cross Spectrum (WCS). The seismic wavespeed perturbation is then obtained from the linear regression of phase shifts with their respective lag times under the assumption that the velocity perturbation is homogeneous between (virtual or active) source and receiver. We categorize these methods into the time domain (WCC, TS, DTW), frequency domain (MWCS), and wavelet domain (WCS). This study complements this suite of algorithms with two additional wavelet-domain methods, which we call Wavelet Transform Stretching (WTS) and Wavelet Transform Dynamic Time Warping (WTDTW), wherein we apply traditional stretching and dynamic time warping techniques to the wavelet transform. This work aims to verify, validate, and test the accuracy and performance of all methods by performing numerical experiments, in which the elastic wavefields are solved for in various 2D heterogeneous halfspace geometries. Through this work, we validate the assumption of a linear increase in phase shifts with respect to phase lags as a valid argument for fully homogeneous and laterally homogeneous velocity changes. Additionally, we investigate the sensitivity of coda waves at various seismic frequencies to the depth of the velocity perturbation. Overall, we conclude that seismic wavefields generated and recorded at the surface lose sensitivity rapidly with increasing depth of the velocity change for all source-receiver offsets. However, measurements made over a spectrum of seismic frequencies exhibit a pattern such that wavelet methods, and especially WTS, provide useful information to infer the depth of the velocity changes.

2.1 INTRODUCTION

THE SHALLOW EARTH is a dynamic membrane of the Earth's crust. It is subject to external forces such as transient tectonic stresses, continuously fluctuating hydrologic and magmatic conditions, thermoelastic expansion and contraction, and tidal stresses, among other forcing mechanisms.

These forces act across a wide range of length and time scales and therefore are considered in most studies of subsurface rheology in various fields of the geosciences, including laboratory experimentation (10^{-3} - 10^0 m), geotechnical engineering and reservoir exploration (10^0 - 10^3 m), and regional crustal or upper mantle scale investigations (10^3 - 10^5 m). Laboratory experiments^{???} and seismological observations^{???} have shown that perturbations in acoustic and elastic velocities, which are associated with the rheology of Earth materials, are clearly correlated with surrounding changes in stresses or with observed material damage. Tracking acoustic or elastic wavespeed changes in time and space is therefore an effective approach to monitoring the rheological properties of the shallow Earth.

Measurements of time-lapse velocity changes have been widely used in seismological investigations. In passive seismology, seismic velocity changes are commonly utilized for tracking volcanic activities^{??????????????}, monitoring earthquake damage and stress changes near or far from the fault^{????????????????}, evaluating changes in subsurface hydrology due to fluid injection and/or groundwater fluctuations^{???}, extracting atmospheric effects such as pressure and temperature on subsurface rheology^{???}, monitoring hydrocarbon and geothermal reservoirs^{???}, and other applications^{???}. Besides geophysical research, temporal velocity changes often provide useful observations in rock physics and civil and geotechnical engineering, including material testing^{?????} and diagnosing infrastructure health^{???}.[?] provided a substantial overview of the applications of velocity change monitoring.

Time-lapse seismic velocity changes can be measured using repeated observations of the seismic wavefield along identical source-receiver paths. The observed seismic waves can be generated from many types of sources. Passive sources mainly include repeated earthquakes, as first proposed by [?], and ambient seismic field interferometry, as first proposed by [?]. Active sources are usually artificial sources from explosions or vibrators chosen for the target area of investigation, repeatable by design, and thus tend to provide high spatial and temporal resolution data for time-lapse imaging [?]. A hybrid approach is that of [?], who proposed using heavy freight trains as repetitive semi-passive/semi-active sources of seismic signals.

Changes in material properties may affect both elastic (wavespeed) and anelastic (attenuation) seismic properties, yielding changes in both phase and amplitude of the repeated seismograms. In this study, we focus only on the measurement of velocity perturbations. When the velocity perturbations are large, they can be detected with relative travel-time measurements of ballistic or direct waves [?]. When the velocity perturbations are small, they can be detected in the late coda waves: the later the scattered wave phases arrive at the receiver, the longer the propagation paths of these waves are, and the more sensitive the waves are to small perturbations in the medium. Whether coda waves are extracted from repeated sources (earthquake or artificial) or from ambient seismic noise cross-correlation functions, relative perturbations as small as 10^{-4} can be retrieved [?].

The relation between phase shifts, dt , and the perturbation in seismic velocity, dv/v , is often assumed as follows. The seismic rays travel a distance L over the travel time (or lag) t at a seismic velocity v . Provided that the change in seismic velocity is homogeneous, the ray path length does not change ($dL = 0$). This yields a relation between the change of seismic velocity, dv , and the phase shift, dt : $dL = 0 = t dv + v dt$, such that $dv/v = -dt/t$. There are two general approaches to measuring dt/t . The first is a *global measurement* performed by stretching two coda waves, where a linear stretching coefficient ε can be found by maximizing the correlation coefficient between the two coda waveforms such that $\varepsilon = -dv/v$. The second is a *local measurement* of phase shifts dt per-

formed at a range of phase lags t , where the phase shifts are often measured within a moving window in the time, frequency, or wavelet domain, or by using dynamic time warping^{??}. When the velocity change is uniform, it can safely be assumed that dt increases linearly with t such that the perturbed velocity dv/v is the opposite of the slope, $dv/v = -dt/t^{??}$. Heterogeneous velocity changes can induce nonlinear changes in phase shifts with lag time, but local measurements of these phase shifts are only useful when the properties of the scattered waves are perfectly known (*e.g.*, the scatterer location and expected arrival times of scattered waves). Otherwise, we are left to assume and fit a linear increase of phase shifts with lag times. This study will verify this assumption is correct in the case of fully homogeneous velocity changes and show that it is a good approximation in the case of realistic media with laterally homogeneous velocity changes.

Observed phase shifts in the coda waves depend on both the true heterogeneity of the background velocity, its perturbation, and the volumetric frequency-dependent sensitivity of seismic waves. Efforts towards locating velocity perturbations have focused on measuring dv/v over different seismic frequencies^{???} and lateral mapping via geometric regionalization^{????}. Lateral mapping of dv/v is often managed by assuming straight ray paths, or volumes around the ray, between (virtual) source and receivers and increasing the density of source-receiver path coverage. The task of extracting depth information from these measurements remains. It is commonly assumed that coda waves are mostly composed of fundamental-mode surface waves, and as such, low frequencies are sensitive to deep structures, while high frequencies are sensitive to shallow structures. However, in scattering media, without systematically depth-varying properties, body waves may dominate later in the coda[?]. [?] suggested that the depth sensitivity of coda waves could instead be related to a linear combination of body and surface wave sensitivities, with the relative importance of body waves increasing with lag time. This relation was applied to lunar and terrestrial data to discriminate velocity changes at depth from near-surface changes^{??}. Regardless of the relative contributions of body and surface waves, frequency-dependent estimates of dv/v may help constrain and infer the depth of

velocity changes simply due to scale-sensitivity of different wavelengths. To increase the frequency resolution of the dv/v measurements, ? proposed constructing a “spectrum of dv/v ”.

There are at least five methods to measure phase shifts dt or velocity perturbations dv/v that work either in the time, Fourier, or wavelet domains. In the time domain, there is the Windowed Cross-Correlation (WCC, ?), Trace Stretching (TS, ?), and Dynamic Time Warping (DTW, ?). In the Fourier domain, there is the Moving-Window Cross Spectrum (MWCS, ?). In the wavelet domain, ? recently proposed the Wavelet Cross Spectrum (WCS), which improved the time-frequency resolution of dt measurements, from which we can further estimate dv/v over a spectrum of seismic frequencies. At present, TS and MWCS are the most commonly used algorithms to monitor velocity changes. Technical comparisons between TS and MWCS have been extensively explored^{?????}. ? performed an exhaustive comparison between WCC, TS, and DTW. ? performed a comparison of MWCS and WCS. In this study, we propose two new wavelet methods, Wavelet Transform Stretching (WTS) and Wavelet Transform Dynamic Time Warping (WTDTW), and compare all time-, Fourier-, and wavelet-domain methods to assess their accuracy, robustness, and computational efficiency.

The second section of this article tackles an algorithm verification exercise by systematically comparing approaches against each other in the case of a synthetic, ground truth experiment. In the third section, we validate the assumption of linear stretching in the case of homogeneous velocity change with a numerical experiment of a scattering halfspace medium. We develop several model configurations relevant to ambient seismic noise configurations with a surface source-receiver pair at two different offsets. Finally, we explore the sensitivity of our measurements to the depth of the velocity perturbation.

2.2 METHODOLOGIES

In this section, we introduce and test published stretching and phase-shift-based methods, present two new methods, and verify them all using a simply stretched waveform.

2.2.1 WINDOWED CROSS-CORRELATION - WCC

The first approach to measuring phase shifts in the time domain is the Windowed Cross Correlation (WCC). The maximum of the cross-correlation function between two time series occurs at a phase shift dt that maximizes the similarity between the two time series windowed around phase lag t . Therefore, WCC is used to measure phase shifts dt at increasing lag times t through sliding windows in the waveforms, which makes this measurement local. It has been widely implemented to measure dv/v [1]. The algorithm requires several parameters, such as the sliding window length, a sliding step size, a start time (minimal lag), and an end time (maximum lag). The WCC is a classic and robust approach. Its limitations are that it assumes that the spectral content of the waveforms is similar (something that can be handled by narrow bandpassing of the waveforms), and it also reports that it suffers from cycle skipping issues.

2.2.2 TRACE STRETCHING - TS

TS relies on the assumption that phase shifts dt linearly increase with lag times t . The algorithm estimates dv/v by linearly stretching the lag time axis of one waveform such that it maximizes its correlation with another waveform [2]. The time axis of the stretched waveform is $t' = t(1 + \varepsilon)$, where the stretching factor is often chosen on physical grounds to be less than 1. ε is found in practice by a simple grid search to maximize the correlation coefficient between the reference and stretched waveforms and bounded by a priori knowledge of the maximum magnitude of the

changes in velocity. When the trace of interest is stretched and matches the reference, no regression is required, and the velocity perturbation is immediately given by $dv/v = -\varepsilon$.

Because TS is a global measurement over the entire trace, it does not localize perturbations within the coda. Therefore, it cannot discriminate between distinct phases in the coda that could be associated with specific scatterers, which limits us in locating the perturbation in space. Another possible challenge is to apply TS on broadband signals that have a different spectral content because coda waves of different frequencies are likely sensitive to different volumes[?]. However, the advantages of using TS are that it is capable of handling large dv/v [?] and that it is robust against noise[?].

2.2.3 DYNAMIC TIME WARPING - DTW

? introduced DTW for use in coda-wave interferometry. DTW was first proposed in speech recognition^{??} to match voices with varying speech rates. It has been widely utilized in exploration seismology to estimate nonlinear phase shifts and suppress cycle skipping^{???}. It follows a concept similar to TS, but instead of having a constant stretching of the lag time axis, it allows for a variable stretching factor at each time lag t . t' becomes $t(1 + s(t))$ and $\varepsilon = s(t)$ varies across time lags. The phase shifts dt are then found by taking the shortest warping path. The parameters to tune are the maximum warping distance allowed and a value that bounds the maximum absolute velocity perturbation. The readers are referred to ? and ? for more details on the algorithm and parameter selection.

DTW and WCC are both more general and flexible than the global stretching methods like TS in that they measure local phase shifts. However, the space-time localization of scatterers is difficult in crustal coda waves. In seismic monitoring that seeks an average dv/v , one still performs a linear fit such that $s(t) \sim \text{constant}$. ? found DTW to be robust against cycle skipping and noise, but computationally intensive.

2.2.4 MOVING WINDOW CROSS SPECTRUM - MWCS

MWCS, also called the doublet method, was first proposed by [?] for measuring velocity perturbations from a pair of repeated earthquake waveforms, or earthquake doublets. The approach is theoretically an extension of WCC to the Fourier domain, so it inherits many of the advantages and limitations of WCC. It measures the time shifts in specified frequency bands, which has the added value relative to time-domain approaches that can refine the frequency resolution. As with WCC, the free parameters are the duration of the sliding window and the sliding step size. In contrast to WCC, each phase shift dt is calculated in the frequency domain rather than the time domain. Others have further discussed its implementation [?]. Since MWCS is a frequency-domain extension of the WCC, it has the requirement that the waveform is a similar but locally phase-shifted version of the reference waveform. It is also subject to cycle-skipping issues, especially when confronting large velocity variations [?]. In part due to the implementation of the Fourier transform, it requires the specification of a frequency band and thus suffers from trade-offs between accuracy, resolution, and stability of the phase shift measurements.

2.2.5 WAVELET CROSS SPECTRUM - WCS

As stated above, estimating the velocity perturbation dv/v in various frequency bands may be desirable for locating the perturbation at depth. Because the Fourier transform suffers from a poor time-frequency resolution, [?] proposed to use the wavelet transform, which permits both time and frequency localization, to extract phase delays using the wavelet cross spectrum [?]. The wavelet cross spectrum is a complex-valued time-frequency field whose phase $\varphi(f, t)$ depends on both frequency f and phase lag t . [?] retrieved the phase shifts $dt(f, t) = \varphi(f, t)/(2\pi f)$. Similar to other methods, the frequency-dependent velocity perturbation measurement, $dv/v(f)$, is then obtained by the linear regression of $-dt(f, t)/t$.

As discussed in §, the main advantage of the method is that dv/v can be estimated over a discrete frequency spectrum. The main disadvantages are that i) the phase measurements are only reliable for the time-frequency space with sufficient power and ii) measurements may suffer from cycle skipping issues². To maintain reliability in noisy data (i.e. local changes in the power of the cross spectrum), § suggested the use of a weighting scheme based on the normalized cross-spectrum power. In our case, we find that a weighting scheme that uses the cross-coherence power yields more stable results.

2.2.6 PROPOSED METHODS: WAVELET TRANSFORM STRETCHING (WTS) AND WAVELET TRANSFORM DYNAMIC TIME WARPING (WTDTW)

Here, we propose two new methods that combine the Continuous Wavelet Transform (CWT) with TS and DTW. We call these two approaches Wavelet Transform Stretching (WTS) and Wavelet Transform Dynamic Time Warping (WTDTW), respectively. These methods rely on the proportionality between the real part of the wavelet transform and the original signal, except that each frequency element of the wavelet transform is a time series with well-resolved frequency (or a perfectly narrow bandpassed time series). Appendix A discusses the relation between a time series and the real part of its CWT. The methods follow the same strategies as the previously detailed algorithms. First, we calculate the wavelet transforms of the reference and the current time series. Second, we either apply TS or DTW to the real part of the wavelet transforms as in the time-domain approaches. A detailed workflow of both WTS and WTDTW is given in Fig.A.1. WTS and WTDTW are effectively identical to their time-domain counterparts but performed on accurate estimations of each frequency component of the signal. We test WTS against a conventional TS approach with a narrow bandpass strategy (Gaussian filter) in Section 5.1. Given the direct extension of TS or DTW to the wavelet domain, WTS and WTDTW both inherit the advantages and disadvantages of the time-domain TS or DTW, respectively. We show and analyze the results of WTS and WTDTW in detail

in the following sections.

2.2.7 VERIFICATION OF THE ALGORITHMS

Our first test is on a simple stretched synthetic seismogram. We group the methods into three categories: 1) the time-domain methods (WCC, TS, and DTW) applied to the raw and unfiltered waveforms, 2) the frequency-domain method (MWCS) and the wavelet-domain methods (WCS, WTS, WTDTW) applied in a set of frequency bands, and 3) the wavelet-domain methods applied over a vector of discrete frequencies, which forms a more complete “spectrum of dv/v ”.

The reference waveform is generated similarly to ? by convolving a zero-phase Ricker wavelet with a central frequency of 1 Hz with a time series of random values. We create a perturbed (or current) waveform by linearly stretching the reference waveform. We choose $\varepsilon = -0.1\%$ to represent a “true” homogeneous velocity perturbation of $+0.1\%$. Fig.2.1(a) shows the two truncated “coda windows” of the waveforms between lag times of 45 s and 75 s. For dilated waveforms ($dv/v < 0$ and $\varepsilon > 0$), the time series are expected to have lower frequency content than the reference waveform, and conversely, a compressed waveform ($dv/v > 0$ and $\varepsilon < 0$) will have higher frequency content. However, small values of $|dv/v|$ have limited effect in altering the amplitude spectrum (Fig.2.1(b)). Because stretching is applied to the raw waveform, the full range of frequencies is equally stretched reversely, simply due to the scaling theorem of the Fourier transform.

We now discuss and show the accuracy of the dv/v measurements retrieved by all methods (Fig.2.2). In the first group of methods, we observe that both TS and DTW accurately recover the velocity perturbation of $+0.1\%$, as shown in Fig.2.2(a). However, WCC underestimates the velocity perturbation by 2% of the true value, a bias that we attribute to waveform distortion and stretching within each sliding window. In the second group of methods, we apply MWCS and the three wavelet-based methods to signals in five narrow frequency bands (Fig.2.2(b)). We note that WTS and WTDTW are the most reliable because they provide estimates of dv/v that are closest to the true value.

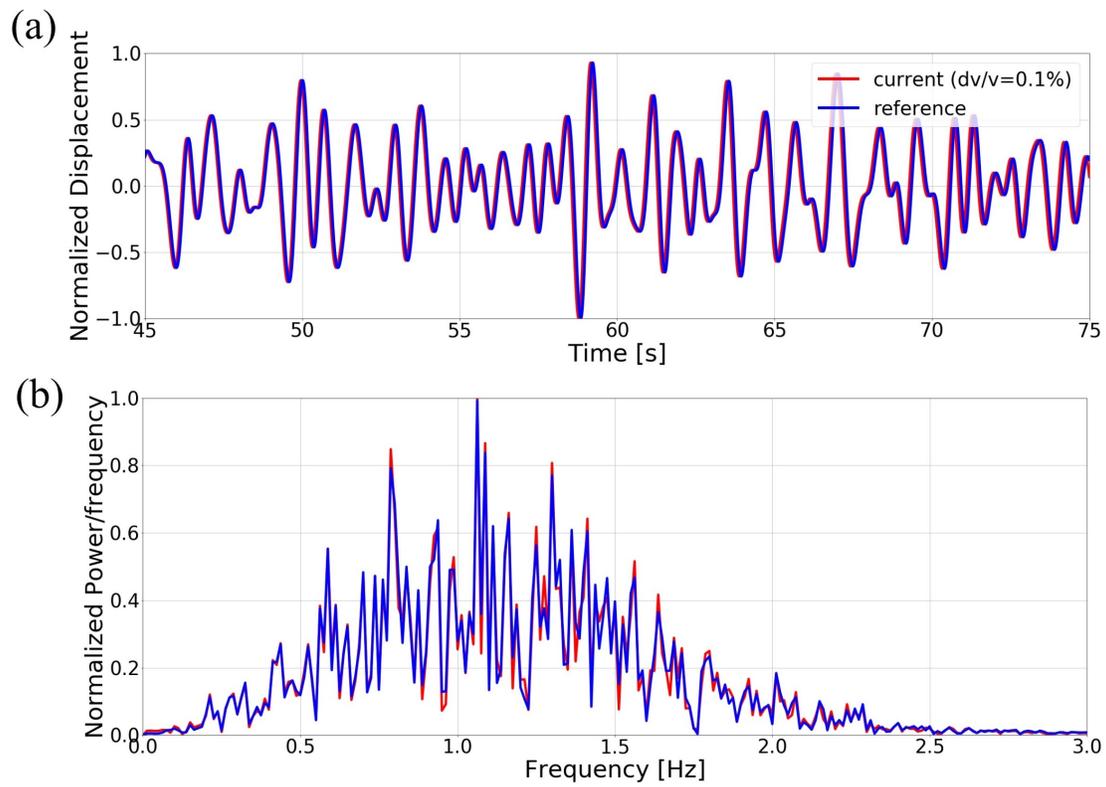


Figure 2.1: Synthetic experiment for the code verification exercise. (a) Reference synthetic waveform (blue) with the stretched current waveform (red). (b) Their power spectral densities.

MWCS and WCS both yield small but noticeable biases at frequencies with lower power (or signal). We find that the biases in WCS and MWCS measurements are correlated with energy variations in the spectral content of the signal: spectral peaks coincide with small errors, but spectral troughs coincide with large errors. Fig.2.2(c) shows the dv/v spectrum between 0.5 and 3.0 Hz obtained with the three wavelet-domain methods. Both WTS and WTDTW exhibit stable and correct results over the frequencies considered, but the errors from WCS increase with increasing frequency, especially after 2.0 Hz. The phases measured by WCS are contaminated by low-frequency signals and influenced by the large fluctuations in signal energy, a spectral leakage also highlighted by ?. Thus, WTS and WTDTW, which use both phase and amplitude of the signal, perform better than WCS, especially at high frequencies where the signal levels are low. The measurement uncertainties were too small (less than 10^{-5}) to be displayed here (see details in Fig.A.2).

To conclude, all of the methods have been verified. WCS is the only method that appears to have a systematic bias with increasing frequency; however, the errors are reasonable across a wide range of frequencies that would be used in realistic examples.

2.3 VALIDATION USING FULL WAVEFORM MODELING

This section serves several purposes. First, we test whether a homogeneous change in velocity yields a linear stretch of the coda waves. Second, we test the methods in two simple velocity structures, one halfspace model and one “layer-over-halfspace” model, to which we add three different perturbations in seismic velocities. The latter test allows us to explore whether measurements of dv/v are sensitive to the background velocity structure. Third, we investigate the first-order effects of the depth of the velocity perturbation on the spectrum of dv/v . Fourth, we explore two source-receiver configurations that are typical of ambient-noise monitoring studies: the case of a zero-offset surface source and receiver that yields the reflectivity response (in this work) as a proxy of the ambient

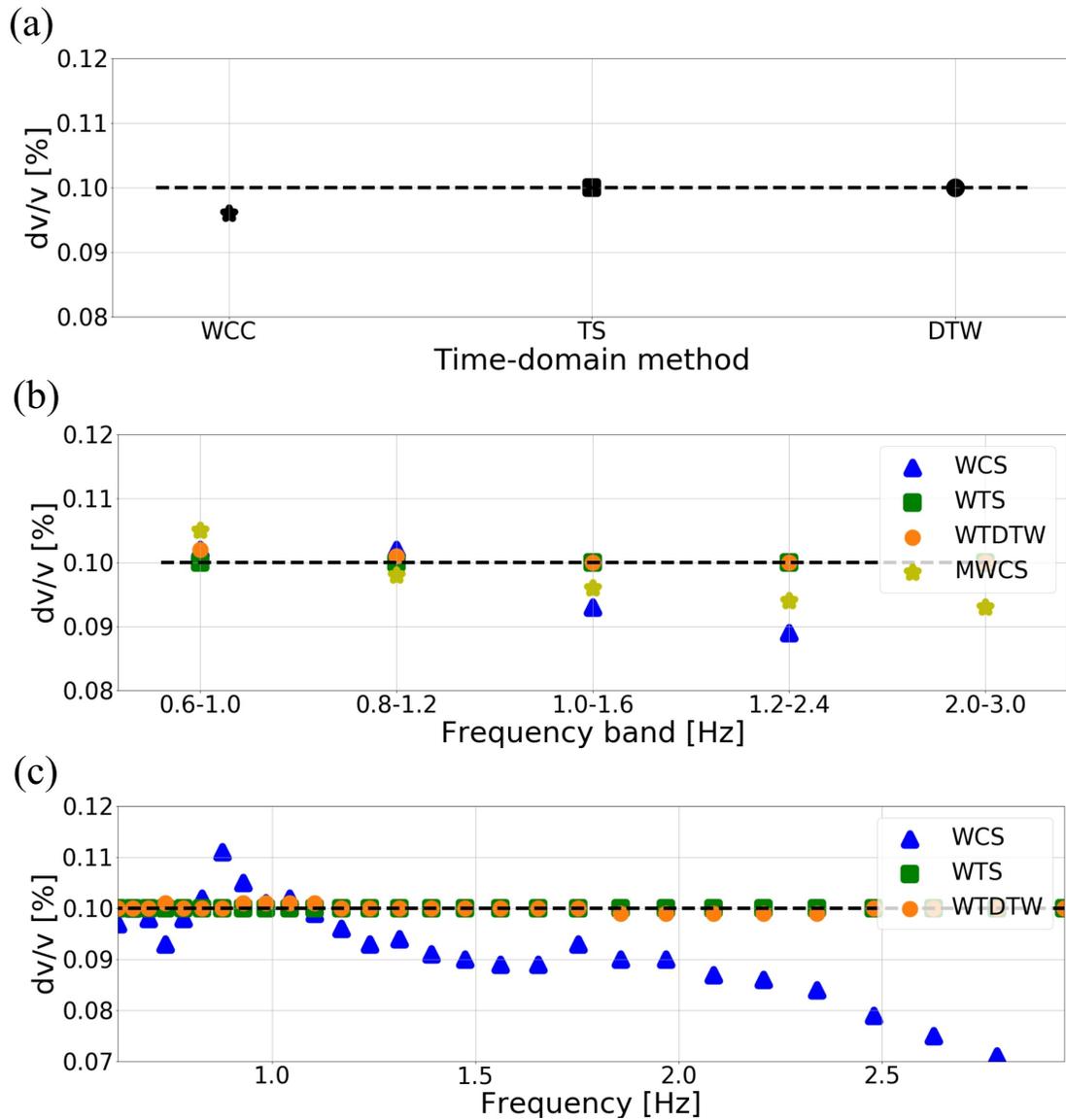


Figure 2.2: Velocity perturbations dv/v from the stretched synthetic waveform shown in Fig.2.1 measured with each method. The black dashed line represents the "true" value of dv/v . (a) Time-domain group of methods of the raw waveforms: the star indicates the result from WCC, the square indicates the result from TS, and the circle indicates the result from DTW. (b) Fourier- and wavelet-domain groups of methods for waveforms filtered at multiple narrow frequency bands: the yellow stars indicate the results from MWCS, the blue triangles indicate the results from WCS, the green squares indicate the results from WTS, and the orange circles indicate the results from WTDTW. (c) Wavelet-domain group of methods for all frequencies: the blue triangles indicate the results from WCS, the green squares indicate the results from WTS, the orange circles indicate the results from WTDTW.

noise auto-correlation^{???}, and the case of distant-offset surface source and receiver that yields the Green’s function and that approximates the ambient noise cross-correlation function[?]. This numerical exercise does not address the problem of whether the “noise cross-correlation” is a suitable approximation to the Green’s function because i) observational studies perform a wide variety of pre-processing steps on the raw data and cross-correlations that merit a separate investigation^{???}, ii) the measurements we make in the coda waves are less sensitive to an uneven noise source distribution^{??}, and iii) the computational expense would limit the number of tests we can practically perform. Additionally, using the Green’s function allows for the generalization of these experiments to active-source and repeating earthquake studies.

We perform this exercise through full waveform modeling with the software SPECFEM2D^{??}. Table A.1 details all of the parameters used to set up the medium, source, and receiver locations, and wavefield simulation. These parameters mainly include a) model size, grid spacing, and elastic properties of the background medium (i.e., P-wave velocity (V_P), S-wave velocity (V_S), and density (ρ)); b) the velocity fluctuation and correlation distance parameters of the von-Karman auto-correlation function, designed to generate scattered surface and bulk waves from ballistic waves[?]; c) source time functions, the direction of the source single force, source dominant frequency, and the sampling rate for the waveform simulations. The level of V_P perturbation is set to +0.1% for each experiment, and hence V_S perturbation is also +0.1% due to the constant V_P/V_S of $\sqrt{3}$. We use and illustrate the simulated vertical displacement seismograms. Throughout the numerical experiments, we perform all of the methods mentioned above using the parameters that are listed in Table A.2.

2.3.1 HALF-SPACE CONFIGURATION

The simulation with a zero-offset source and receiver configuration is shown in Fig.2.3(a), and the resulting seismograms in Fig.2.3(b). The waveform within the selected coda window is dominated by scattered waves as there are no distinct phases. We also show their power spectral densities in

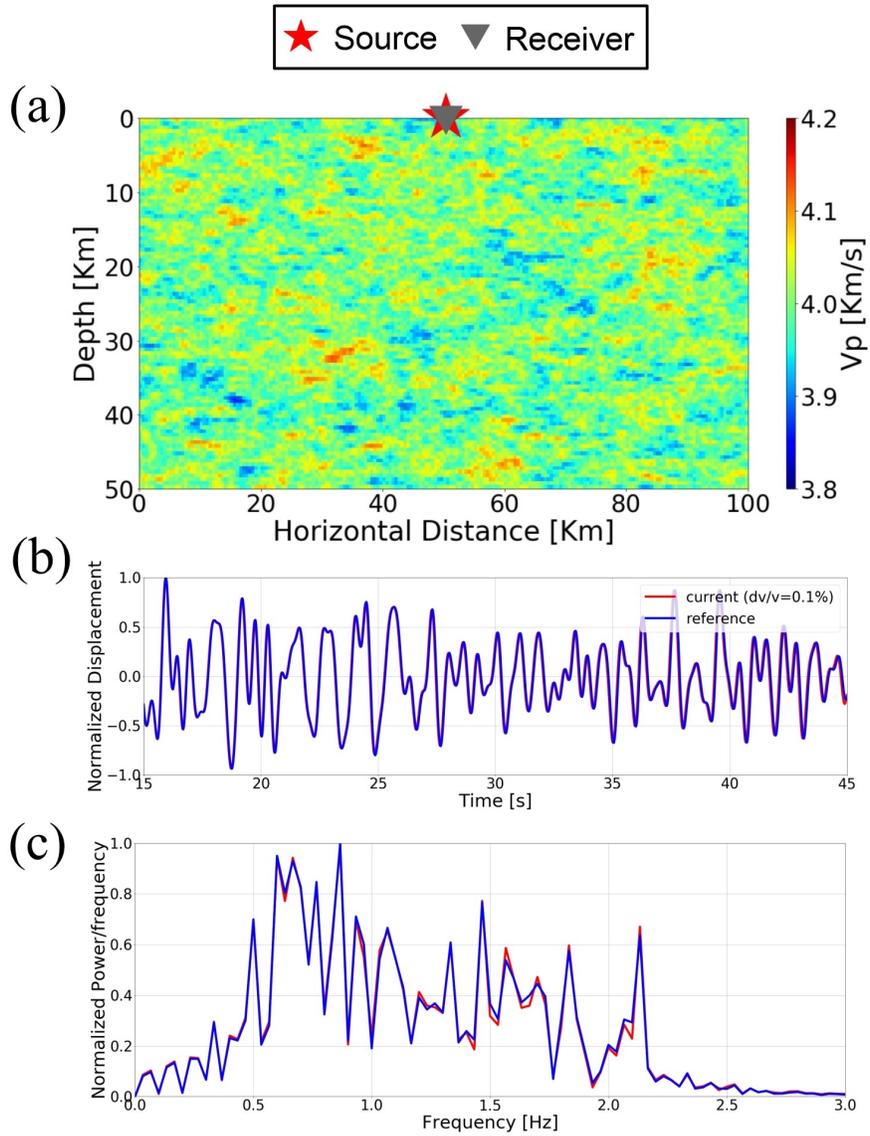


Figure 2.3: (a) Heterogeneous V_p velocity structure is shown in the colormap, the source is the red star, and the receiver is the gray triangle. (b) Waveforms simulated with SPECFEM2D in the reference medium and medium perturbed by +0.1% (current). (c) Power spectral densities corresponding to the waveforms in (b).

Fig.2.3(c) to highlight the spectral peaks between 0.5 and 2.2 Hz.

Similar to the previous experiment, we apply all methods to estimate dv/v and show their results in Fig.2.4. The time-domain group of methods yields accurate results (see Fig.2.4(a)), demonstrating that both TS and DTW accurately recover true homogeneous dv/v , with WCC performing only slightly less accurately. The second group of methods based on the Fourier and wavelet transforms also performs reasonably well for the first four frequency bands where the signal is the strongest, as shown in Fig.2.4(b)). At higher frequencies, we find that both WTS and WTDTW still yield good results but that the results of both WCS and MWCS worsen with increasing frequency. We attribute this to the lower energy of the signals at high frequencies.

The results from the wavelet-domain group of methods accentuate these patterns (Fig.2.4(c)). WCS is unstable at frequencies with low signal levels, which is also discussed in ?. Inaccuracies in the dv/v estimated from WTS and WTDTW are also noticeable (e.g., 0.005% positive bias for WTDTW at the spectral peak around 1 Hz), though both methods yield more stable results than WCS.

2.3.2 LAYER-OVER-HALF-SPACE CONFIGURATION

In this section, we explore the impact of a more realistic velocity structure (layer-over-halfspace) on the estimates of dv/v with all methods. The upper 1-km thick layer has a P-wave velocity of $V_P = 4$ km/s, and the lower half-space has a P-wave velocity of $V_P = 6$ km/s. S-wave velocities are proportional to the P-wave velocities, with $V_S = V_P/\sqrt{3}$, and density remains constant at $\rho = 2000$ kg/m³. A perturbation in P-wave velocity of +0.1% is imposed, with a corresponding S-wave velocity perturbation due to the constant Poisson ratio. The perturbation is imposed either on the whole space, on the shallow layer, or on the half-space. The time-domain signals exhibit little visual difference between the reference and perturbed waveforms (Fig.A.3). We only show the results for WCS, WTS, and WTDTW below on the zero-offset and then the distant-offset source-receiver geometry. However, we report that the same conclusions as in the previous section are drawn from the other meth-

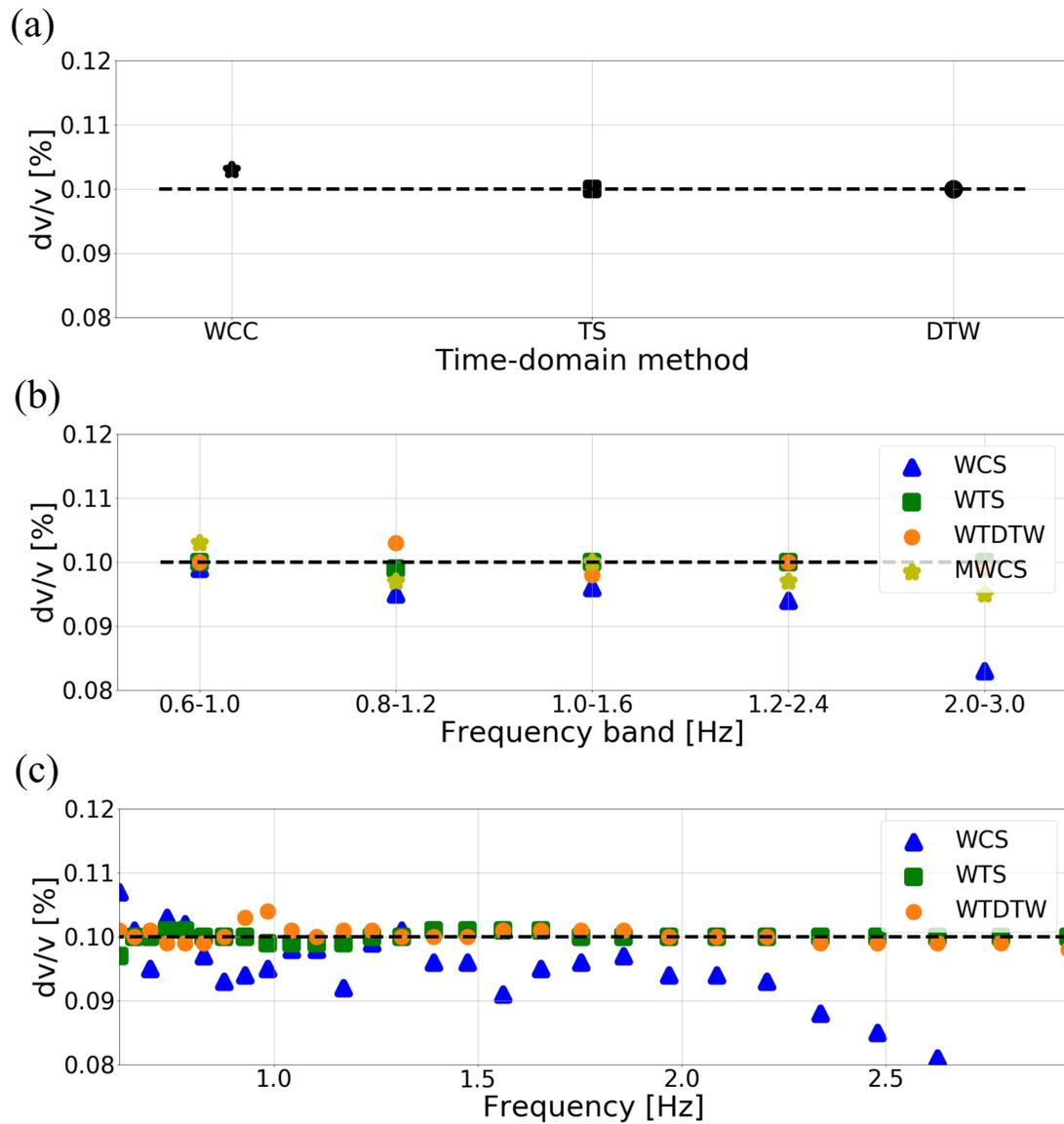


Figure 2.4: Predictions of dv/v from coda-wave interferometry with various techniques for a uniform change in velocity. (a) dv/v results of three time-domain methods WCC (star), TS (square), and DTW (circle). (b) dv/v results at various frequency bands for MWCS (yellow stars), WCS (blue triangles), WTS (green squares), and WTDTW (orange circles). (c) The spectrum of dv/v results of the wavelet-domain methods for WCS (blue triangles), WTS (green squares), and WTDTW (orange circles).

ods.

The results for a *uniform change* and with the zero-offset configuration are similar to those found in the previous two cases (compare Fig.2.5(a) to both Fig.2.2 and Fig.2.4(c)). This validates the assumption that a linear stretch indeed reflects a homogeneous change in seismic velocities. It also suggests that measuring dv/v due to a uniform change with this approach is independent of the background velocity medium. It further confirms the level of accuracy of WTS and WTDTW among the methods over a large range of frequencies.

In the case of the *shallow change*, dv/v approaches the true value of dv/v with increasing frequency (Fig.2.5(b)). Since the source and the receiver are both at the surface, near-source and near-receiver scattered waves are directly sampling the perturbed medium. Given the shorter wavelengths of high-frequency waves, one might expect high-frequency scattered waves to sample a smaller volume near the source and near the receiver than the one sampled by low-frequency scattered waves. Therefore, low-frequency scattered waves may sample or propagate through the deeper and unperturbed part of the medium.

In the case of the *deep change* in velocity and in the zero-offset configuration, the dv/v decreases with seismic frequency (Fig.2.5(c)). Here, all methods yield similar results. The interpretation of the spectrum shape is similar to that of the previous case: high-frequency seismic waves propagate in the shallow, unperturbed medium near the source and the receiver, and the low-frequency seismic waves propagate through and are sensitive to a larger volume, including the perturbed layer. We also report similar results and conclusions with all other methods (Fig.A.4-A.6).

We now perform the same tests for the distant-offset source and receiver configuration to illustrate the case of ambient noise cross-correlation monitoring. In this case, the source and receiver are separated by 100 km and the upper lower-velocity layer is 10-km thick. The simulations run for 300 s with a grid spacing of 1 km and time step of 0.01 s. The simulated waveforms only show small differences (Fig.A.7). Here again, we perturb the entire medium, the upper layer, and the lower

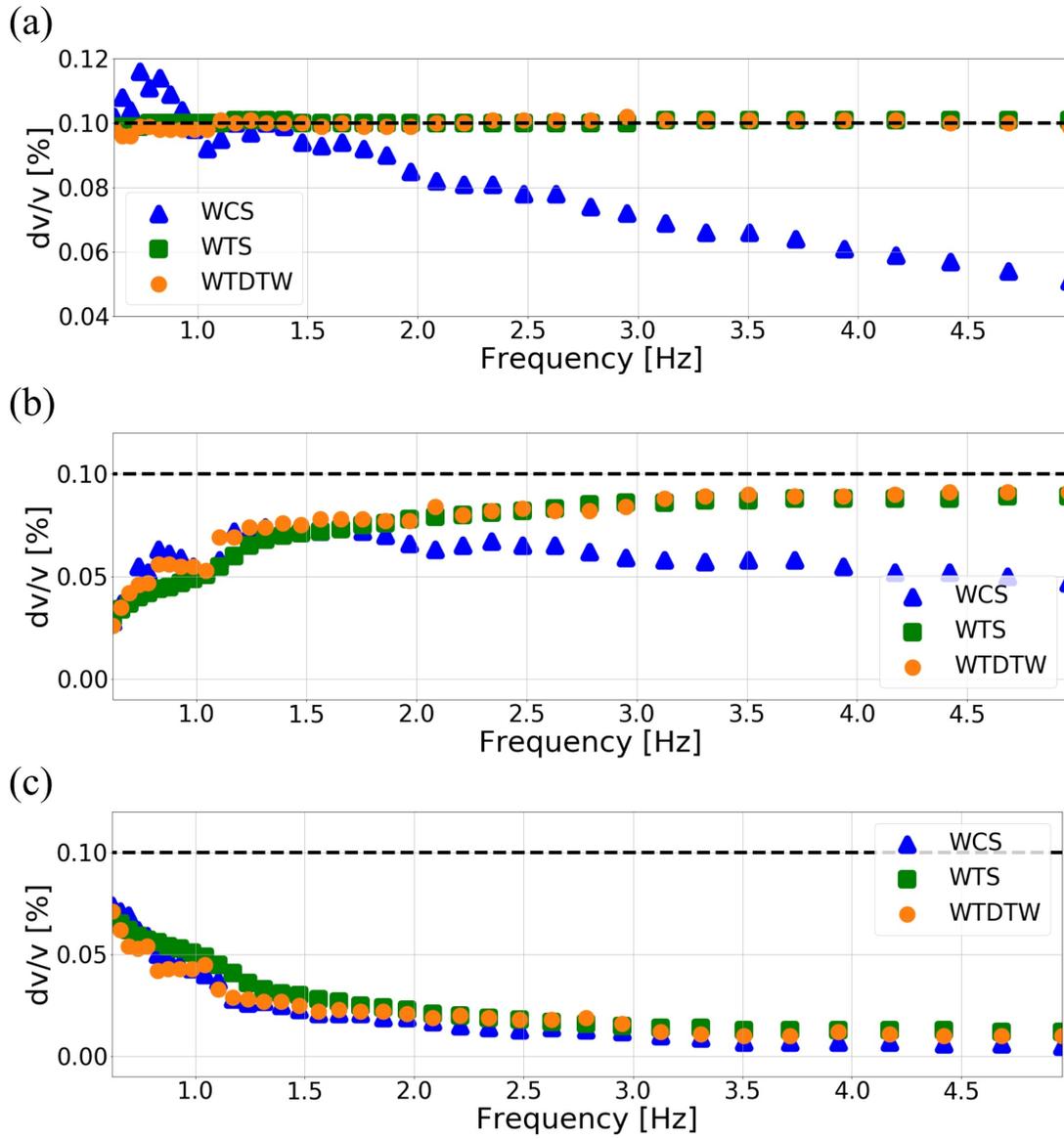


Figure 2.5: Spectra of dv/v obtained from three wavelet-domain methods for three experiments in the zero-offset source-receiver case: a uniform change (a), a change in the shallow layer (b), and a change in the deep layer (c).

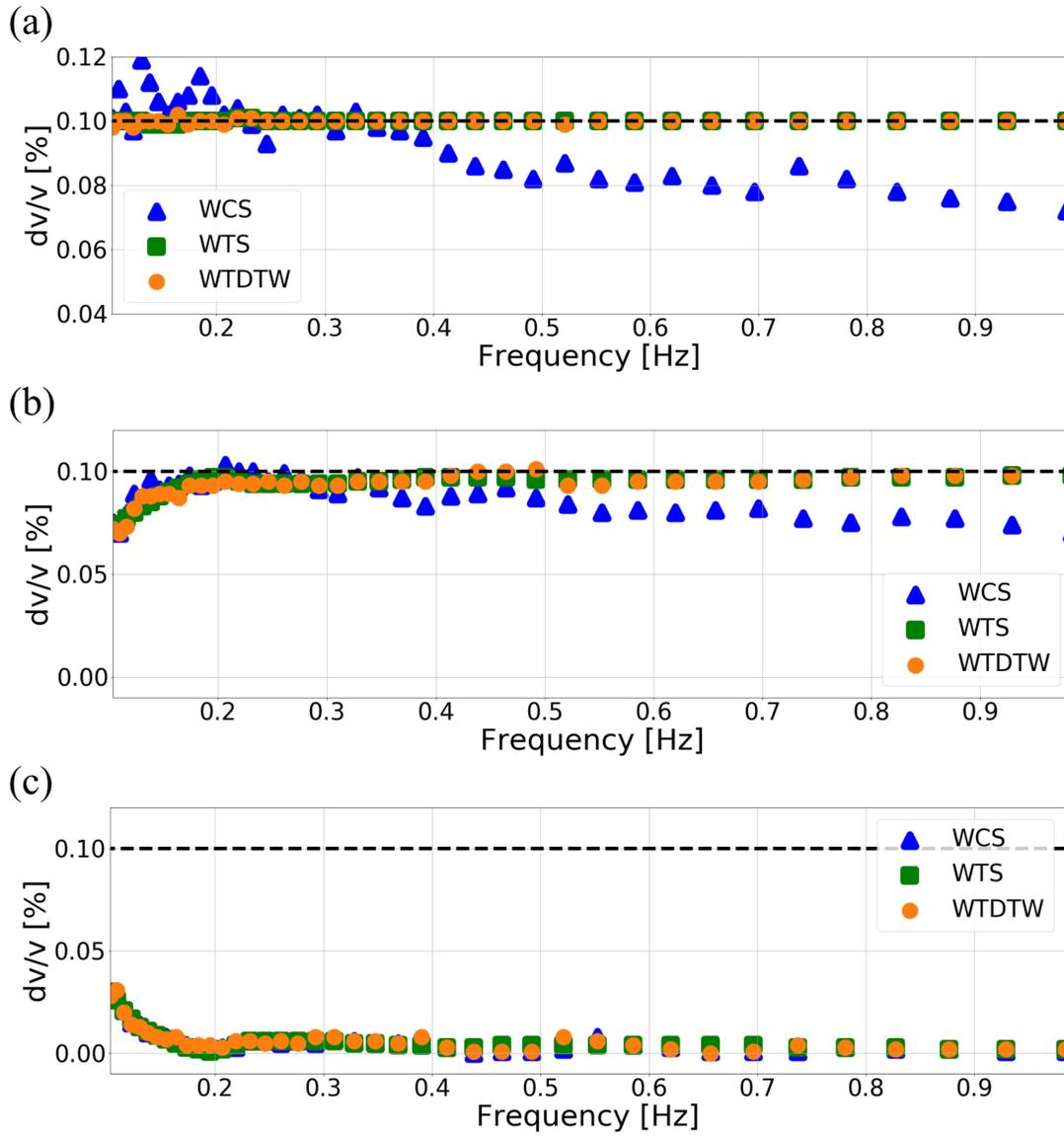


Figure 2.6: Same as Fig.2.5, but for the case of the distant-offset source and receiver configuration.

halfspace. We only show the results from the wavelet-based methods in the main article (Fig.2.6), and show similar results and conclusions for all other methods in the supplement (Fig.A.8-A.10), including deviation from the true dv/v value for WCS and MWCS due to variations in signal energy. The spectrum of dv/v behaves similarly to the zero-offset source-receiver case, albeit with dv/v measured at lower seismic frequencies.

It is interesting to note the contrasting shape of the spectrum of dv/v for deep vs shallow perturbations. This finding is one of the main contributions of this study: there is a clear seismic signature of the depth of the perturbation in the spectrum of dv/v , which promotes wavelet methods as promising tools for estimating the depth of velocity perturbations. In this narrow frequency bandwidth, the dv/v increases with frequencies for shallow velocity perturbations and decreases with frequency for deep velocity perturbations. This is unique to the wavelet-domain methods because the frequency bands of other methods are too broad to capture the characteristic shapes of dv/v spectra for different perturbation depths (Fig.A.9-A.10). We also note that the values of dv/v obtained through any of the techniques may largely underestimate the true value of dv/v in instances where the velocity change is localized in-depth, a fact that was also noted in ?. This is likely because scattered waves propagate well beyond the thin perturbed layer.

2.4 SENSITIVITY WITH THE DEPTH OF THE VELOCITY CHANGE

The previous section highlights that the depth of the perturbation has a clear effect on the spectrum of dv/v . In this section, we systematically explore this effect in a similar exercise to one taken by ?: we perturb a thin layer within the background velocity model by adding 0.1% to the velocity and then we “slide” this layer from top to bottom. Here, the medium is a heterogeneous halfspace (similar to Section 3.1).

The depth sensitivity of the coda waves relies on many factors: level of heterogeneity of the

medium, depth of the velocity perturbation, and relative contributions of surface waves (early coda) and body waves (late coda)². Here, we only consider the effect of the depth of the velocity change to dv/v measured with TS and to the spectrum of dv/v measured with WTS. We show in Fig.A.11 and Fig.A.12 the results for all other methods for the zero- and distant-offset source and receiver geometries, respectively.

First, we consider the case of the zero-offset source and receiver configuration. Each layer is 200-m thick and slides at an interval of 100 m down from the free surface to 2 km depth. We thus generate one reference model and 20 models with perturbations. Grid spacing is 100 m horizontally and 20 m vertically. We simulate all seismograms with a sampling rate of 0.0004 s, and with a 5 Hz Ricker wavelet as the source time function.

We first apply TS over the raw coda waves and plot the estimated values of dv/v for each model (Fig.2.7(a)). The recovered velocity perturbation dv/v rapidly decreases from the true value at the surface down to zero at depth. This is similar to what was established. Since the depth sensitivity vanishes completely after 1 km, we only show the dv/v results above 1 km. Because of this loss of sensitivity, it appears difficult to use the measured dv/v over the broadband coda waves and even more difficult to use TS to infer the depth of the perturbation. We now explore the effects of perturbation depth on the spectrum of dv/v , and show in Fig.2.7(c) the results for WTS. All other methods yield the same results (Fig.A.11).

Similar to the earlier analysis (Fig.2.5 and Fig.2.6), we now obtain dv/v spectra from WTS that exhibit clear and systematic variations with the depth of the perturbation (Fig.2.7(c)). The spectrum of dv/v obtained from a perturbation contained in the upper 200 m is similar to that shown in Fig.2.5(b): the high frequencies seem sensitive to a volume contained within the perturbed layer and can yield the appropriate value of dv/v .

At greater depths, the dv/v spectra have a global maximum at a frequency that decreases with increasing depths of perturbation. The global maximum of each dv/v spectrum also decreases with

the depth of the perturbation. The loss of sensitivity of the high frequencies for deep perturbations may be attributed to the fact that high-frequency waves mainly travel in the upper structure, near surface, nearest to the sources and receivers, and where there is no velocity change. The loss of sensitivity of the low frequencies for the deep perturbations is probably due to an averaging of the wavefield over the unperturbed medium, as discussed earlier.

The specific frequency of 1.38 Hz appears to be a “limit frequency” where the global maxima of the dv/v spectra converge. After testing various parameters, for both the simulation parameterization (i.e. dominant frequency, increment of layer thickness, and grid spacing) and the dv/v algorithms (i.e. lapse or lag times), we find that the selection of lag times primarily affects the value of this limit frequency. Late lag times result in a lower limit frequency, since the later coda waves incorporate more low-frequency waves that have traveled through the deeper structure. We also find that there is no limit frequency for the case of a dv/v spectrum extracted from direct surface wave (see later Section 5.4). To date, we do not find any physical explanation for such convergence.

Next, we investigate the distant-offset source-receiver geometry. In this case, the perturbed layer is 4-km thick and slides from the surface to a depth of 10 km at an interval of 1 km. We use a Ricker wavelet with a dominant frequency of 0.3 Hz. We use both TS and WTS to construct the spectrum of dv/v with varying perturbation depths. The results are shown in Fig.2.7(b) and (d), respectively. We report that lower thickness values of the perturbation layer yield a lower estimate of dv/v , likely due to the fact that seismic waves at these frequencies sample a much broader volume than the perturbed layer.

Overall, the results are quite similar to those obtained from the zero-offset source-receiver configuration (Fig.2.7(a) and (c)). The spectra have a global maximum that decreases in amplitude and frequency with increasing depth of perturbation. We only show the range of frequencies where we see the most variation in the spectra, but report that the asymptotic behavior pertains at the low and high frequency ends. Fig.A.12 illustrates that our conclusions remain with all other methods.

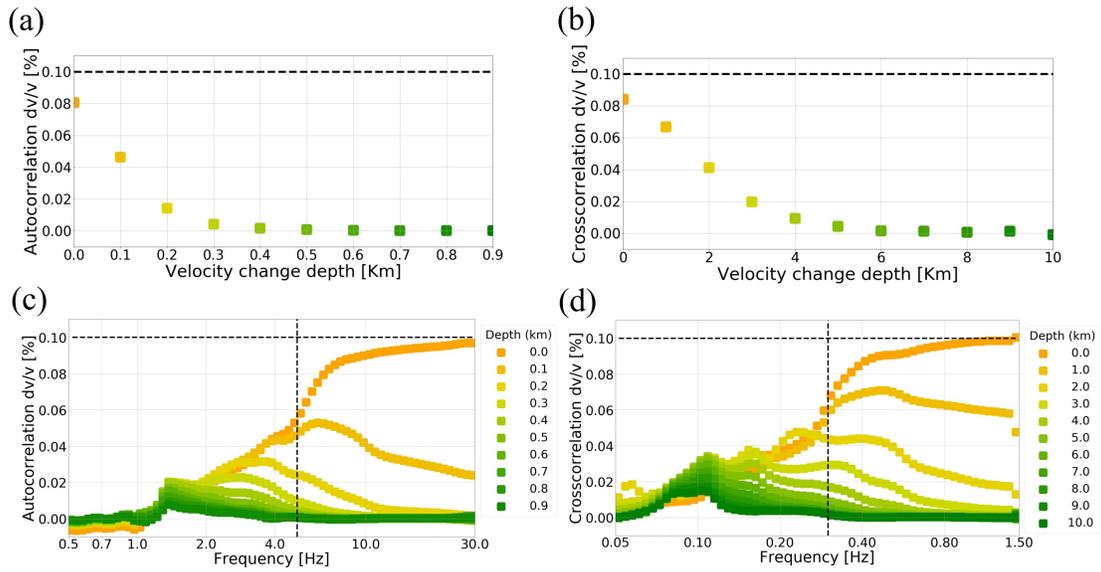


Figure 2.7: Apparent velocity perturbation versus the depth of the perturbed layer for TS (a, b) and WTS (c, d), both in the zero-offset source-receiver configuration (a, c) and distant-offset source-receiver configuration (b, d). The color scale denotes the depth of the perturbed layer. The vertical dashed line denotes the source frequency of 5.0 Hz for (c) and 0.3 Hz for (d).

2.5 FURTHER DISCUSSION

In this section, we summarize the advantages and limitations of the approaches by discussing i) accuracy, ii) stability against additional factors, and iii) computational efficiency of all the methods we tested in this work. We then discuss the influence of coda wavefield composition on the dv/v measurements.

2.5.1 ACCURACY COMPARISON

In order to investigate the accuracy of all methods, we statistically perturb the uniform background velocity model (Section 3.1) from -0.5% to 0.5% with an interval of 0.05%. This is effectively a bootstrap on the velocity structure and we perform the same analysis as in Section 3.1. The residuals between estimated and true values of velocity change are shown in Fig.2.8.

In the first group of measurements in time domain, the errors are on the order of about 10^{-5} , which are much smaller than other (wavelet and Fourier) methods. TS and DTW have greater accuracy than WCC. In the second group of methods (in frequency bands), we observe that both MWCS and WCS's errors are greater than that from WTS and WTDTW, especially in the high-frequency bands, where the errors become greater than those of WCC. We attribute the larger errors from MWCS and WCS to their lower accuracy in estimating phase lags at frequencies with low signal energy. Their accuracy would probably be comparable if the signals had similar frequency bands.

Among all approaches and in the tests presented above, WCS is the most sensitive (i.e. high error) to the low-energy signals, especially at high frequencies (see Fig.2.8(c)). WTS performs better than WTDTW, but both have the best accuracy among the suite of methods we tested. Similar results can be achieved in the third group for all individual frequencies. Small biases around 1 Hz may result from the relatively low energy levels at this frequency in the time series, as shown in Fig.2.3(c). Because WTDTW is more sensitive to relative signal amplitudes than WTS, the WTDTW errors are slightly larger than those of WTS.

Additionally, we revisit the statement that WTS is more accurate than the combination of a time-domain narrow bandpassing and TS approach. We test with a Gaussian filter and find that WTS yields higher accuracy and better computational efficiency than bandpassing (Fig.A.13).

2.5.2 ADDITIONAL FACTORS THAT CAN AFFECT THE RESULTS

In practice, the measurements of velocity perturbations may be impacted by many factors: variation in earthquake or noise source spectrum, window length, lag time, and noise levels. As we discuss in this section, the two factors that most influence the measurements are the time series spectrum and the noise level. We explore the influence of these factors on the dv/v based on the models of Section 3.1.

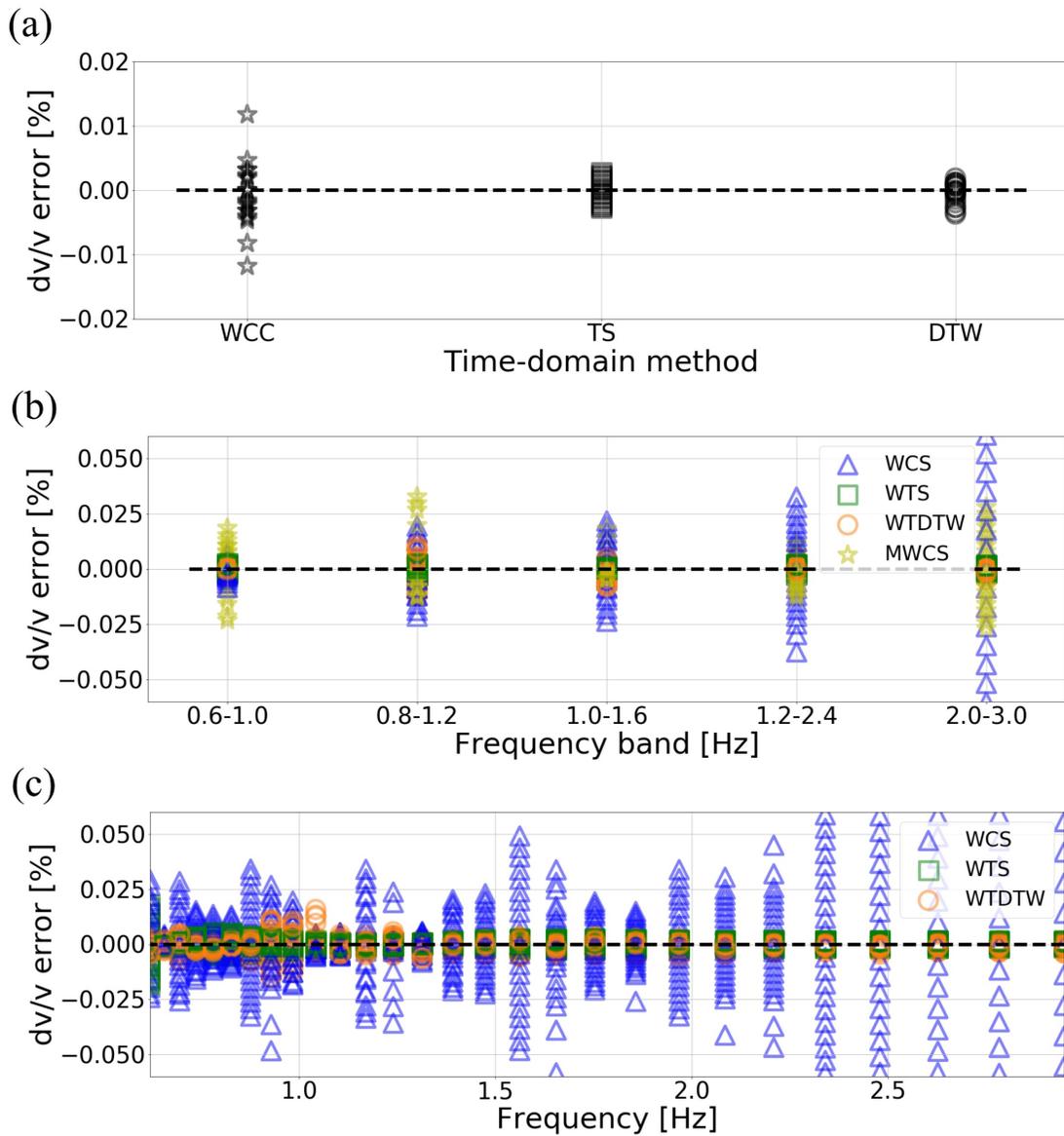


Figure 2.8: Errors in velocity-perturbation estimations over 20 realizations of homogeneous velocity changes. In all three panels, the black dashed line represents zero error. (a) Errors of time-domain methods: black stars indicate the errors from using WCC, black squares indicate the errors from using TS, and black circles indicate the errors from using DTW. (b) Errors of four approaches in the bandpassed filtered Fourier and wavelet methods (second group). (c) Errors of three wavelet-domain approaches in the third group.

Temporal changes in the ambient noise or earthquake source spectrum in coda-wave interferometry, which is also discussed in §, affect the spectral content of the coda waves. We thus examine the effects of varying the source time function spectrum. We vary the dominant frequency of the Ricker wavelet of waveform in the “current state” and compare to a reference dominant frequency of 1.0 Hz. We find that changing the dominant frequency of the Ricker wavelet may cause systematic biases that mainly fall between -0.05% and 0.05%. This effect is common to all approaches, with a slight under-estimation when the dominant frequency of the current signal is lower than the dominant frequency of the reference signal, and vice versa (Fig.A.14). The measures of uncertainty from WCS and WTS are lower ($< 0.02\%$) than for other methods. The stretching-based methods are the most stable, whereas the DTW-based methods are the least stable. This is because the DTW-based methods are sensitive to the small differences in the phase of the waveforms.

Next, we investigate the dependence of dv/v measurement stability on lag time. Here, we choose a dominant source frequency of 1 Hz, simulate the wavefield through both reference and perturbed media, and perform the dv/v analysis by varying the start time of the coda window (lag time). Overall, TS and WTS perform best over all lags (Fig.A.15). Since the amplitude spectrum of the signal varies with lag time and remains somewhat unstable for all lag times, the MWCS and WCS results have increased deviation and instability with lag time, especially at low frequencies. This test also demonstrates that WTDTW has large deviations at low frequencies.

We test the influence of the window length by varying it between 15 and 55 s (Fig.A.16) but find no significant impact on the stability of the measurements. To test the influence of the noise level on the dv/v estimates, we select one segment of coda waves and add a Gaussian noise by varying its peak absolute amplitude between 0% and 30% of the maximum absolute amplitude of the original signals. Unsurprisingly, high noise levels yield poorer results (Fig.A.17). Overall, WTS and MWCS are the most robust methods against noise.

2.5.3 COMPARISON OF THE METHODS COMPUTATIONAL EFFICIENCY

We also test the computational efficiency of all approaches in a similar way to ?. We execute each approach twenty times with the example shown in Section 3.1 on a single CPU, and then compute their mean computation time (see Fig.2.9). The time-domain methods are much faster than the other methods, likely due to the added computational cost of the Fourier and wavelet transforms and the number of frequencies over which we perform the analysis. Linear stretching is faster than dynamic time warping. Because the Fourier transform is faster than the wavelet transform, MWCS is less expensive than the three wavelet-based methods. We conclude that WCS is the most computationally efficient among the three wavelet-domain methods, because phase differences can be directly obtained from the wavelet cross spectrum. WTS and WTDTW become the most computationally expensive. We note that compute times may vary by a factor of 100 and thus their implementation will likely be weighted against accuracy.

2.5.4 ON THE SCATTERING OF BODY AND SURFACE WAVES

The relative contribution between scattered body and surface waves in coda waves is complex and changes with lag times due to multi-pathing and mutual wave-type conversions. Because surface and body waves are both emitted and received at the surface, both have sensitivities that vary with depth. Knowing the relative contribution of wave types in the coda is critical for a proper interpretation of the dv/v spectra in terms of perturbation depth. ?? first studied this using numerical simulations of wave propagation in a heterogeneous half space. They found that the “early” coda was dominated by surface waves and the “late” coda was dominated by body waves. We proceed similarly in this study using a distant-offset (100 km) source-receiver geometry with three different tests:

1. First, we start with a uniform halfspace medium with no perturbation in velocity (i.e. no scattering). We window 10 s before and 30 s after the surface-wave arrival (at about 92% of

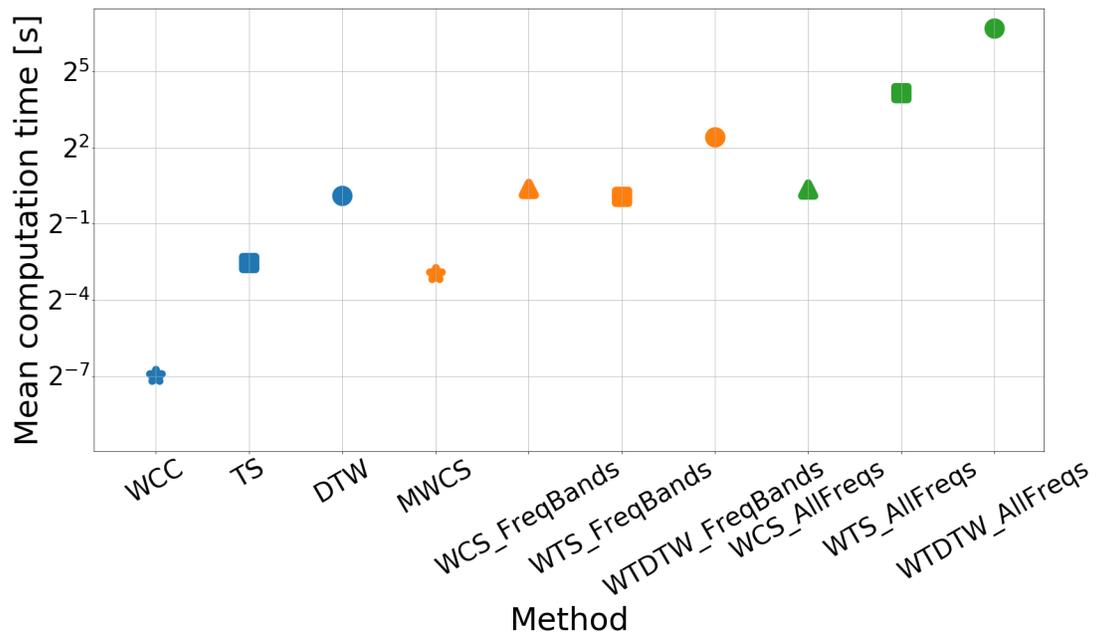


Figure 2.9: Mean velocity-perturbation computational time over 20 realizations of homogeneous velocity changes for all methods. Note that the blue, orange, and green ones represent the first, second, and third group of methods, respectively.

the shear-wave velocity) (WTS results shown in Fig.2.10(a)). The windowed signal is clearly dominated by surface wave amplitudes, thus we neglect the contribution of S waves in these results.

2. Second, we add a halfspace above the source-receiver axis and add an absorbing condition to all sides of the model, thus only exciting body waves. We window 10 s before and 30 s after the S wave (WTS results shown in Fig.2.10(b)). S waves have larger energies than P waves and thus are likely the main contributor to our measurements. We check that the window does not contain P waves and is free from artifacts of the absorbing boundary conditions. Note that the concept of “depth” here is simply a distance from the source-receiver axis.
3. Third, we simulate scattering by adding back the velocity heterogeneity in the whole space model to simulate only scattered body waves. We window from the lapse time of 185 s to 225 s (WTS results shown in Fig.2.10(c)). Similar to the previous case, the concept of “depth” here is simply the distance from the source-receiver axis.
4. Fourth, we add back the traction-free boundary condition (same as the model configuration of Section 4). We select three 40-s long windows, respectively starting from the early (55 s), middle (115 s), and late (185 s) lag times, and compute three dv/v spectra at the velocity-perturbation depth of 1 km (Fig.2.10(d)).

The first two experiments aim to provide the seismic signature of direct/ballistic surface waves and of direct body waves. The first two results show that ballistic surface waves are about twice as sensitive as ballistic body waves to the shallow perturbation (Fig.2.10(a,b)), a finding that resonates with that of?. The depth-frequency dependence seen in Fig.2.10(a) is likely a signature of the sensitivity of surface waves with depth. The depth-frequency dependence seen in Fig.2.10(b) is likely a signature of finite-frequency sensitivity of body waves.

We then investigate the seismic signature of scattered waves. The third exercise excites scattered body waves, and our results suggest that scatterings drastically reduce the sensitivity of body waves (Fig.2.10(c)). This contrasts with the results from scattered waves in a half space (Fig.2.7(d)). We deduce that the dv/v spectra in the heterogeneous half space are dominated by surface waves.

Given the findings of ? with respect to lag time, we evaluate the effect of relative wave-type contributions in the dv/v spectra in Fig.2.7(d). We find that the shape of these spectra with early (55 s) and late (185 s) lag times are similar, likely dominated by surface waves. The shape and amplitudes of the intermediate lag time (115 s) dv/v spectra is reduced, which may indicate a greater contribution from the scattered body waves.

We further show the complete dv/v spectra of scattered body and surface waves (Fig.A.18) and only scattered body waves (Fig.A.19) at the same three lag times. The evolution of partitioning between scattered surface waves and scattered body waves follows three stages: 1) scattered surface waves dominate in early coda, 2) scattered body waves may contribute about equally to scattered surface waves in the mid-coda, which fits with the findings of ? ?, and 3) scattered surface waves dominate again in the late coda, which is supported by ? and likely due to the smaller geometrical spreading of surface waves. Note that this interpretation depends on our attenuation-free simulation, and thus may be different in areas where surface waves and shear waves are attenuated. We also find that the depth sensitivities for the zero-offset setting are more stable than the distant-offset setting, and have small changes across coda lag times.

2.6 CONCLUSIONS

In this study, we compare the current techniques for measuring dv/v in coda-wave interferometry. We also develop two new algorithms, namely WTS and WTDTW, by extending the TS and DTW methods to the wavelet domain. The goal of developing these new methods is to construct a stable

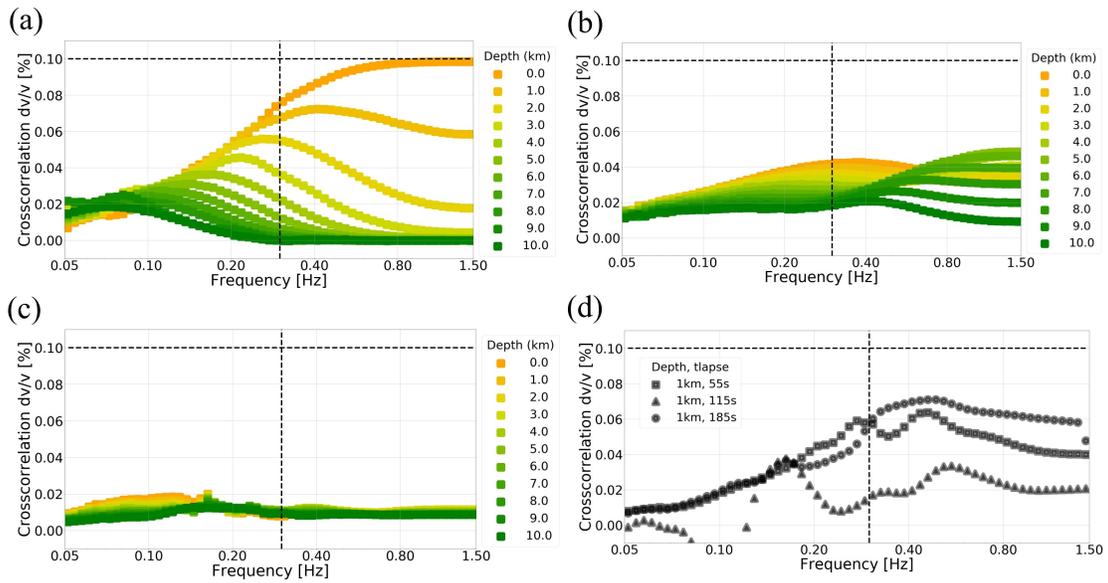


Figure 2.10: Apparent velocity perturbations for the distant-offset source-receiver configuration. The dv/v spectra measured for direct/ballistic surface waves (a) and direct/ballistic body waves (b) as a function of the depth of the perturbed layer. (c) The dv/v spectra measured versus the depth of the perturbed layer for scattered body waves at late lapse times (185 s). (d) The dv/v spectra measured with both scattered body and surface waves at early (55 s), middle (115 s), and late (185 s) lag times for a perturbed layer at 1-km depth. All results are measured using WTS. The vertical dashed line denotes the source frequency of 0.3 Hz.

“spectrum of dv/v ”, where dv/v is calculated over a dense vector of seismic frequencies.

We analyze the accuracy and stability of these techniques using full waveform modeling. We set up the numerical experiments to reflect two classic ambient noise monitoring configurations: the reflectivity response (to represent the zero-offset surface source and receiver cross-correlations) and the distant Green’s function (to represent the distant-offset surface source and receiver cross correlation). Our experiments also allow us to explore the impacts of the depth of the velocity change on the spectrum of dv/v . We have also attempted to explain the relative contribution between direct and scattered body and surface waves into our measurements of dv/v spectra.

We summarize accuracy, stability, and computational efficiency of all methods in Table A.3. The wavelet-domain methods enable high frequency resolution in the dv/v measurements (i.e. a spectrum), and are more accurate but computationally intensive than time-domain methods.

Details of the limitations and performances are also discussed in ? and ?. Overall, the main conclusions are:

1. There is a clear signature of the perturbation depth in the spectrum of dv/v .
2. The measured dv/v most likely underpredicts the magnitude of localized velocity perturbations at depth.
3. The maximum value and its corresponding frequency of the measured dv/v spectra both decrease with deeper velocity changes.
4. WTS is our recommended method to best measure spectra of dv/v .

The three main previous studies on these topics are ?, ?, and ?. The added contributions to this literature are the introduction of new methods, a systematic comparison between all published methods, and the sensitivity of the “spectrum of dv/v ” to the depth of the velocity perturbation.

The systematic variation of the “spectra of dv/v ” is an encouraging step toward mapping velocity perturbations at depth.

This study lays the ground work for further analysis. In particular, we have not addressed several issues that are pertinent to ambient noise seismological applications. Our numerical experiments may hinge on exact retrieval of the Green’s functions, while many ambient seismic noise studies make the assumption that the noise correlation functions are merely an approximation to the Green’s function. An additional study would be to analyze how a non-uniform noise source distribution would affect the results across the methods. Furthermore, many studies perform extensive pre-processing of the noise^{??} and/or post-processing of the noise correlation functions^{??}. A systematic investigation of the effects of this processing on the dv/v measurements would be interesting, since numerical experiments are limited by poor knowledge of the Earth’s structure, and in particular its heterogeneity and scattering properties. The obvious next step is to apply our techniques to real data.

3

Laboratory analog to monitoring
unconfined aquifers with passive
seismology

SUMMARY

Monitoring seismic properties such as wavespeed and attenuation are diagnostic markers of strain deformation and fluid transport in the shallow crust. While seismic observations have revealed substantial variations in temporal mechanical properties, current interpretations remain limited due to the complexity of the sensitivity of scattered wavefields to changes in partially saturated media. We propose the first analog experiment to relate measurements in seismic wavespeeds with ground truth measurements in water levels in a 3D printed laboratory analog. Our results demonstrate that 1) both wavespeed and amplitude changes have a negative relation with water-level increase in granular media; 2) amplitude changes are susceptible to the thickness and impedance of the medium while wavespeed changes are robust; 3) both wavespeed and amplitude changes at high frequencies are influenced by the cohesion between grains. Our experiments provide fundamental insights into the factors of water saturation and solid deformation for seismic monitoring in near-surface sediments.

3.1 INTRODUCTION

Seismic interferometry using ambient seismic field seismology has fueled the development of environmental seismology by enabling the monitoring of shallow subsurface changes at high temporal resolution. The methodology utilizes phase shifts in repeated source-receiver travel paths to extract changes in seismic properties along the path. Seismologists have used both direct-wave measurements^{??} e.g., and coda-wave measurements[?] e.g., to infer the depth and lateral changes in seismic wavespeed, dv/v . Once measured and imaged, these changes have been associated with magmatic and hydrothermal activities at volcanoes^{??} e.g., differential loading from Earth tides and snow or rain precipitation^{???} e.g., earthquake damage^{???} e.g., and wastewater injection or geothermal fluid injection^{??}, and groundwater monitoring (*references therein*).

The focus of this study is the recent application of the methodology to monitoring groundwa-

ter using ambient field monitoring. dv/v has been anticorrelated with pore pressure expected by ground truth, co-located groundwater well or soil moisture measurements^{1,2,3,4,5,6,7,8,9,10}, or modeled by diffusion of rainwater, snowmelt, or river stream^{11,12,13,14,15}, or inferred from remote sensing¹⁶ and piezo-electric proxies¹⁷. The inference of the depth at which these changes occur has conventionally been estimated using surface wave frequency-depth sensitivity kernel, whereby deeper changes can be probed using lower frequency^{18,19}. Beyond these empirical relationships,²⁰ proposed a physics-based model that directly expressed changes in shear wavespeed due to stress induced by pore pressure changes and induced vertical loads as competing mechanisms for decreasing wavespeed (pore pressure) or increasing wavespeed (vertical loads) depending on the permeability.

Most laboratory studies target the effects of pore pressure, stress, and relative humidity or pore pressure in rock samples where the pore pressure is often assumed to be homogenized in the sample^{21,22}. In general, studies find that shear wave velocity is negatively correlated. There is a slight mismatch between the rock physics experiments and the natural setting of an unconfined aquifer where seismologists compare dv/v to groundwater levels. We propose to fill this gap using a 3D-printed analog granular material, controlled water levels, and acoustic emissions.

In this study, we attempt to elucidate the relationship between frequency-dependent variations of seismic wavespeed measurements, varying stress, and varied water level in a 3D-printed granular porous medium. We design a new experiment where active-controlled acoustic monitoring is performed for scenarios at different stress and water levels. We consider two common cases in geological contexts: partial and full saturation in the granular media.

3.2 METHOD

3.2.1 EXPERIMENTAL APPARATUS

We develop a new experimental apparatus to measure the response of effective acoustic wavespeeds and attenuation to fluid saturation and external stress in the artificial granular media. We showed the apparatus in Fig.B.1. Our specimen has three layers: rigid resin upper and lower blocks and a soft resin 3-D printed as a granular medium in the middle. All three components are 3D-printed individually by using a Stereolithography (SLA) technique. The rigid media are printed with clear resin with a tensile modulus of 2.8 GPa. The soft medium is printed with flexible resin with a tensile modulus of 3.1 MPa. Instead of rigid material, we use a soft material to print granular media to exaggerate the deformation and mimic a low-velocity zone. We also experiment with two regimes of cohesion for the granular medium: unbonded grains and well-bonded grains.

Moreover, the post-curing with ultraviolet is applied to solidify the resin and assemble it into a complete specimen. We fabricate the specimen into a “sandwich” style to create a realistic Earth analog of three-layer sediment in the field. The top block is rigid in great part to enhance the coupling of the transducer onto the surface. The granular medium is designed with uniform grains with a diameter of 1.8 mm and interconnected pores with a porosity of 28.7%. The specifications of this granular media are similar to that of typical sandstone and carbonate rocks[?].

We inject pure water from underneath the sample (see Fig. B.1), and the fluid is free to flow in a plastic tube driven by a syringe pump, putting the experiment in a drained regime. We record the water levels with a camera and fluid height is easily read due to the dyed water.

3.2.2 ACOUSTIC MEASUREMENTS

We utilize a pair of acoustic transducers for acoustic measurements: one is the source pulse transmitter, and the other is the acoustic receiver. We first perform two calibrations for P- and S-wave velocities (V_p & V_s) of the rigid and granular media by recreating the conditions for conventional active geophysical survey: the source pulse transmitter is fixed and, once emitted, is recorded at the receiver transducer. We repeat the experiment by shifting the position of the receiver transducer every 1 mm with a regular spacing to mimic a uniform spacing linear array of geophones.

A shot gather is collected and displayed in Fig. B.2a. A manual pick of the first P wave and the S wave arrival yields a crude estimate of the wavespeed of V_p and V_s , 2.4 km/s and 1.1 km/s, respectively. The velocities are associated with the upper rigid media (clear resin). We also perform a dispersion analysis on the shot gather with a phase-shift method (slant stacking)² and highlight several dispersive modes (Fig.B.2b). The dispersion is recognized from the strong surface waves coming shortly after the S waves, given the short source-receiver spacing. We also acquire active acoustic data, as shown in Fig.B.3, on the soft medium. The V_p and V_s for the soft medium are 1.55 km/s and 0.75 km/s, respectively. We summarize the material properties in Table 3.1.

Resin	V_p (km/s)	V_s (km/s)	Density (g/cm^3)	E (GPa)	Poisson Ratio
Rigid	2.4	1.1	1.396	4.62	0.367
Soft	1.55	0.7	1.182	1.59	0.372

Table 3.1: Elastic properties of printed materials and used for the upper and lower blocks (rigid) and in the granular medium (soft).

3.2.3 ESTIMATING WAVESPEED AND AMPLITUDE CHANGES

We perform acoustic measurements at a fixed distance between transducers at different times with condition changes. Robust, repeatable, and reliable amplitude measurements are required to inter-

pret amplitude changes due to structural changes properly. We stick the transducers to the specimen using printed mounting blocks and grease for tight contact (see Fig. 3.1a). We first measure acoustic waves in the absence of water (dry context) and for each level of water injected in the aquifer analog.

The P and surface waves are clearly observed for their distinct arrival times and frequency contents; even the S wave is too short to distinguish. Hence, our data acquisition is more realistic than that of straight monitoring by using opposite transducers.

For the field cases, coda wave interferometry (CWI) is frequently utilized for seismic noise interferometry and repeating earthquakes. By working on coda waves, we conduct wavelet-transform stretching[?] to measure velocity changes (i.e., dv/v) over frequencies. The spectra of velocity changes can be seen in Fig.3.1c. Additionally, we can measure amplitude changes (e.g., dA/A) as attenuation changes due to absorption (intrinsic) and scattering. Absolute intrinsic and scattering attenuation can be obtained through multiple lapse time windows analysis and full envelope inversion[?] of body waves of earthquakes. However, measuring their temporal changes is still challenging. Under assumptions of single scattering or strong scattering (i.e., diffusion model), the decay model proposed by Aki & Chouet[?] has been widely utilized to measure attenuation changes from observed or reconstructed coda waves. Current development in 3D radiative transfer models^{??} would help to better recover these changes by considering multiple scattering for different wave types. As amplitudes are direct to reveal attenuation changes (ignore intrinsic and scattering contributions) without further processing, we are not using any aforementioned theoretical models. As shown in Fig.3.1c, the logarithmic amplitude ratios could directly indicate the response of frequency-dependent attenuation to the drainage. In the next section, we will perform similar measurements for saturation levels and stress loading.

3.2.4 SHEAR VELOCITY CHANGE OF AN EFFECTIVE MEDIUM

The effective medium approximation² is often used to estimate the apparent velocity changes when the dry granular medium is saturated with water. The apparent Vs of saturated media with different water heights can be calculated by changing the medium density as,

$$V_s = \sqrt{\frac{\mu}{\rho_{sat}}}, \quad (3.1)$$

where we approximate the volume of water by the change in weight height L_w

$$L = L_w + L_{dry}, \quad (3.2)$$

and express the change in density with the fractional change in volume and the granular media porosity φ :

$$\rho_{sat} = \rho_{dry} + \frac{L_w}{L} \varphi \rho_w. \quad (3.3)$$

The apparent velocity is estimated from the effective medium density, which increases as the water table rises in the dry granular media. The effective rigidity does not change with raised water levels, but the effective density increases. This leads to a predicted decrease in shear wavespeed as the water table rises. We calculate the velocity changes as,

$$\frac{dV_{s,ave}}{V} = c \frac{dV_s}{V}, \quad (3.4)$$

where,

$$\frac{dV_s}{V} = \frac{V_s - V_{dry}}{V_{dry}}, \quad (3.5)$$

$$c = \frac{vol_{soft} * V_s}{vol_{soft} * V_s + vol_{rigid} * V_{s,rigid}}. \quad (3.6)$$

The apparent velocity changes are calculated relative to the apparent V_s of the dry medium. We also calculate an average velocity change by a fixed averaging factor. This factor is calculated based on the volume and velocity of soft and rigid components. The $V_{s,rigid}$ has been calibrated and shown in Table.3.1.

3.3 RESULTS

3.3.1 WATER LEVEL CHANGES IN LABORATORY AQUIFER

To reconstruct the effect of changing the water table in realistic geohydrological systems, we inject water into the granular medium at various levels. The experiments are set under drained conditions with fluid pore pressure remaining constant.

we record the table height and 500 repeating acoustic measurements. These acoustic measurements will be stacked for a sufficiently robust signal. Example measurements have been displayed in Fig.3.1b. As stated in the previous section, we obtain frequency-dependent dv/v and dA/A . Such experiments and analyses are performed seven times with linearly increased water heights. We show all of dv/v and dA/A measurements in Fig.3.1c.

We observed negative dv/v changes in general. The changes are highly dependent on the frequencies, particularly for the frequencies below 100 kHz. After checking waveforms at a variety of frequency bands, we report that these signals below 300 kHz are dominated by surface waves, and the signals above 300 kHz are dominated by P waves. Thus, the strong dispersion we observed at low frequencies resulted from surface waves, as also seen for the dispersion analysis in Fig.B.2b. Some instabilities are also observed due to the signal power fluctuations as seen in Fig.B.4, although the CWI gives smooth results. With the same increase in water height, we find that the dv/v at each frequency does not decrease monotonically. This is interesting but remains unexplored.

The negative dA/A is also observed for different saturation levels. The dA/A spectra generally

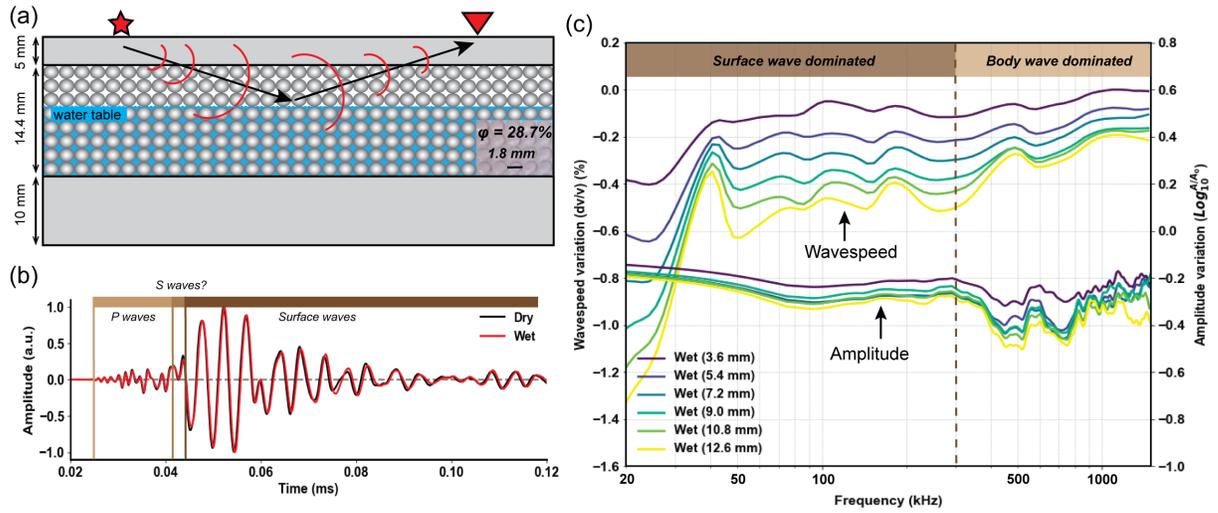


Figure 3.1: Experimental apparatus and measurements. (a) Schematic acoustic measurement in a three-layered medium by sandwiching a granular soft medium into two rigid blocks. (b) Examples of acoustic waveforms are collected during dry (black waveforms) and partially filled conditions (red waveforms). (c) Frequency-dependent wavespeed and amplitude changes for different levels of the water table (color-coded).

decrease with the water height but not monotonically, as we observed for the dv/v spectra. Similar to dv/v , the change magnitude mainly depends on the baseline velocities, where we see strong dispersion at low frequencies. The variation of dA/A at different frequencies is mainly decided by the baseline amplitudes: large-power frequencies suffer from relative small variation while small-power frequencies have relative large variations. The dA/A spectra are directly measured from the signal spectra and inherently have strong instabilities. The dA/A spectra shown in Fig. 3.1c are derived upon the smoothed signal spectra over 100 frequency points. Additionally, when the medium is fully saturated (i.e., water height is 14.4 mm), we notice the spectral power has a strong attenuation as Fig.B.4. We do not include here the dv/v for spurious results due to large change in signal phases and dA/A for the consistency. Additionally, we calculate the shear wave changes with the effective medium approximation in the previous section and show the velocity changes for different water heights in Fig.B.7. The significant velocity reduction indicates that the CWI underestimates those changes due to water percolation.

3.3.2 COHESIVE GRANULAR MEDIA: DRY, PARTIALLY-SATURATED, AND FULLY SATURATED STATES

To consider a more realistic scenario, we deform the well-bonded granular media (i.e., cohesive media) under an applied strain. With the elastic media, the strain and stress are simply converted through elastic modulus. We apply the same strain deformation of 1.5 mm (i.e., compression) onto the granular media under dry, partially-saturated, and fully-saturated states. We keep the system under a drained condition so that no over-pressure is considered. At each state, we perform and stack 500 acoustic records before and after deformation. We conduct the same measurement for dv/v and dA/A as in the previous section. The derived results are displayed in Fig. 3.2.

Comparing dv/v spectra between different saturation states, we observe that deformation can result into large reduction when the granular medium becomes more saturated. Considering the velocity changes are dominated by shear-wave speed change, the increase in density due to filled water and compressed pore size reduce shear-wave speed from the effective medium theory. Those changes are nonlinear and frequency-dependent. At low frequencies (≤ 50 kHz), the velocity change becomes less sensitive to the strain deformation, possibly because the wavelength ($\lambda \approx 20$ mm) is larger than the thickness of granular media. The waves at intermediate frequencies (100–500 kHz) are very responsive to the deformation, probably these waves mostly travel inside of the medium. When wave frequency goes higher, we observe the changes become weaker, probably because high-frequency waves cannot sense that deep. This is unseen in the dry or unsaturated case, but it becomes evident when water is filled. In particular, when fluid is fully saturated, the dominant frequency of dv/v spectra shifts to a higher frequency. This states that the water height reaches the sensing depth of most high frequencies, and the wet medium slows down the high-frequency waves.

dA/A spectra are also measured for the deformation at three saturation states. In the dry test, we find that the amplitude will be decreased by compressing the specimen. The count-intuitive behav-

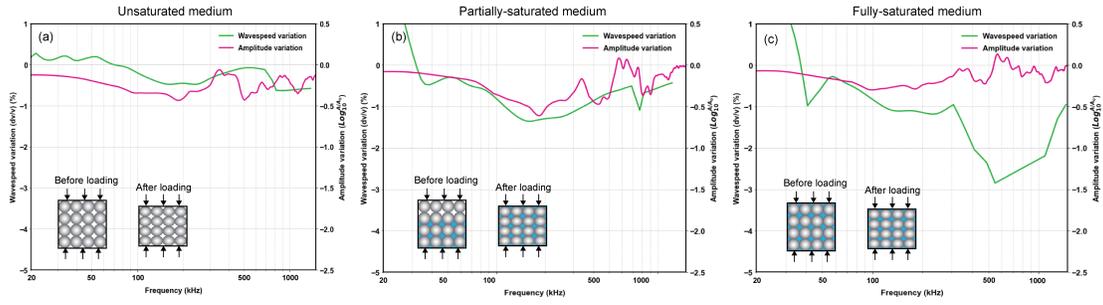


Figure 3.2: Acoustic wavespeed and amplitude changes under strain deformations on the fully-bonded granular media in (a) unsaturated or dry state, (b) partially saturated, and (c) fully saturated states.

ior has resulted from the structure of our specimen, where the soft (low-velocity) granular media is placed at the middle layer. The strain deformation that mainly occurs in the soft medium causes the decrease in thickness and pore size of the granular media, thereby decreasing the amplification intensity of surface waves². This is mainly occurred on the dominant frequencies around 150 kHz. Upon the granular medium is partially saturated, the amplitude would be attenuated for more saturation due to the compression. The amplitude generally decreased more than in the dry case, while the high-frequency amplitudes remained almost unchanged. When the granular medium is fully saturated, we find that the strain deformation has less effect on the amplitude change compared to the dry and partially saturated case. The main reason is that the amplification factor does not decrease much due to much-decreased velocity (i.e., enhanced amplification) other than the decreased thickness (i.e., reduced amplification). Although the thickness plays a dominant role, the amplification factor has been affected a lot and has less amplitude change.

3.3.3 NONCOHESIVE GRANULAR MEDIA: DRY, PARTIALLY-SATURATED, AND FULLY SATURATED STATES

Other than the cohesive granular media, we consider the noncohesive granular medium since it is also widespread in the field. We use the granular media with unbonded grains. Similarly, we apply

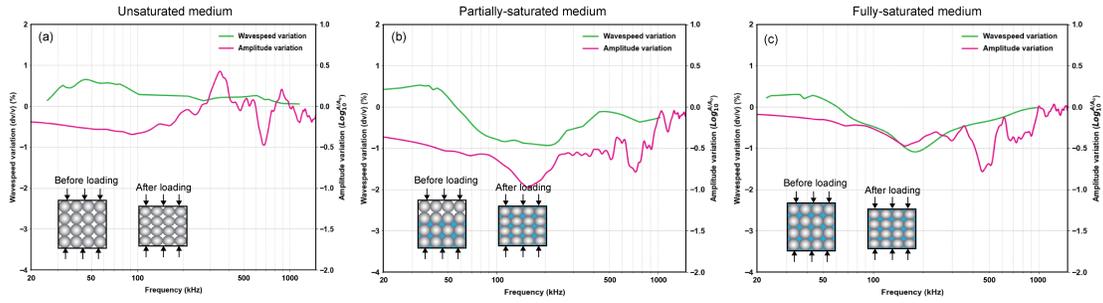


Figure 3.3: Acoustic wavespeed and amplitude changes under strain deformations on the unbonded granular media in (a) unsaturated or dry state, (b) partially saturated, and (c) fully saturated states.

the same strain deformation (1.5 mm) on the specimen at different saturation states: dry (i.e., unsaturated), partially saturated, and fully saturated. We perform the same acoustic measurements before and after strain deformation. We show the acoustic measurements for different saturation states in Fig. 3.3.

Compared with the cohesive case, we find that the compression leads to an increase in shear-wave speeds because the shear modulus increases more than the density. We also recognize that the amplitude increases at high frequencies (above 200 kHz) for the tight grain contacts, while the amplitude change stays negative below 200 kHz. This states that the effect of medium thickness dominates the amplitude change. When the state switches to the partial saturation, we observe that the main velocity changes in the noncohesive case become less compared to the cohesive case because of the compensation of the increased shear modulus. Interestingly, we observe a significant reduction in amplitude, which is unlike the one we observe in the cohesive case. This is mainly because the amplification in the noncohesive case is larger than that in the cohesive case, but the amplification is similar for both cases after the compression. Thus, the larger $\Delta 0$ in the noncohesive case causes larger negative changes in the amplitude ratio. When the state is fully saturated, we do not see a big decrease in the velocities, as seen in the cohesive case. Again, the increase in shear modulus compensates for the loss of the increased density. Similar to the partially-saturated state, we also observe

larger negative changes compared to the cohesive case. Similarly, the reason is the larger $\Delta 0$ in the noncohesive case.

3.4 DISCUSSION

Our laboratory observations can provide the first insight for frequency-dependent field observations of using seismic interferometry. Seismic changes, mainly in velocity, in shallow sediments and sedimentary rocks, have been measured extensively in various scenarios. The dispersion characteristic of dv/v spectra has been utilized to constrain the depth of velocity changes due to water levels[?]. The seismic properties of shallow sediments are easily impacted by environmental factors, such as temperature, snow, tidal, and rainfall. Although their combined effects are convoluted, the sediments are mainly influenced in terms of their saturation or moisture, pores or cracks, and elastic moduli, which are fundamental variables considered in our parametric experiments.

Stress from earthquake waves or tidal force can cause strain deformations, which open or close cracks. Earthquakes may usually give rise to cracks due to strong ground motions, while the tidal stress opens up cracks through the extension. These have been investigated through uniaxial stress experiments on rocks^{???} and ambient noise monitoring^{?????}. The generic understanding is that the opening and closing of cracks mainly result in negative and positive changes in seismic velocity. These cracks would mainly lower material stiffness and, thus, a negative relation between seismic velocity and estimated volumetric or areal strain changes. Our case with noncohesive granular medium validates this, consistent with other active-source controlled seismic experiments on rocks^{???}. However, our cohesive case shows an opposite effect, which suggests microcracks dominate the velocity changes in natural rocks.

Our results also emphasize the competing factor from the fluid. Fluid can saturate pores and potentially open cracks under high fluid pressure. Fluid flow or overpressurized processes can de-

crease velocity by enhancing density or opening cracks. Both processes may coexist as the dilatant effect, for instance, opposite velocity changes is observed due to tides[?], which is different from previous observations^{??}. This indicates that the ambiguity in explaining observed velocity changes is because of the lack of localized geological knowledge. Basic poroelastic property of Earth's materials is needed to understand better and model the dv/v behaviors^{??}. Besides, frequency-dependent measurements would become more important to constrain the depth of dv/v as shown in the experimental results as well as in the numerical studies[?]. While measuring attenuation changes is still challenging, our results demonstrate it serves a great complement to understand physical mechanisms, especially for anelastic and nonlinear deformation[?].

Our experiments provide a protocol that we even extend to consider the effect of crack shape and density, fluid overpressure, and thermal expansion. They can be involved in the consideration of specific contexts. Our cases do not involve the change in acoustic attenuation due to the interaction between wave (i.e., stress) and pores or cracks (known as squirt flow), but this has been widely studied by another experimental apparatus with low-frequency measurements^{???}. The effect of squirt flow can be ignored for low-power waves. Discussing this is already out of the scope of this study. Through a simple linear scaling, the dimension of our sample thickness (e.g., 10 mm) can become 1 km when the dominant frequency (e.g., 100 kHz) of acoustic signals corresponds to the dominant frequency (e.g., 1 Hz) of seismic signals.

3.5 CONCLUSION

We develop an experimental apparatus where active-controlled source monitoring and 3D-printed granular media are leveraged to investigate the effects of fluid saturation and strain deformation on acoustic properties, such as wavespeed and attenuation. We measure frequency-dependent wavespeeds and attenuations with coda wave interferometry. Our results have a generic agreement

with previous laboratory measurements and provide a comprehension on the frequency-dependent wavespeed and attenuation under the conditions of consolidation, deformation, and varying saturation. Although current observations mostly focus on one or several frequency bands, our continuous frequency measurements would be very helpful in understanding the frequency-dependent velocity and attenuation change in field observations.

4

Seismic Monitoring of Magmatic Activity beneath Mount St. Helens

SUMMARY

Seismic monitoring is one important approach to tracking volcanic activity and alerting eruption risk. Inter-station measurements of phase changes, such as ambient noise monitoring or coda wave interferometry, are frequently utilized for long-term monitoring of volcanoes, but the sensitivity of the measurements is typically confined to the shallow subsurface. To explore deeper changes in seismic velocity at Mount St. Helens, we propose to use instead an inter-source monitoring technique. We leverage a vast database of detected repeating earthquakes. We undertake the full earthquake location task in a 3D velocity model to locate the events at depth: this entails deep-learning aided phase picking and 3D event relocation. We recompose the measurements by selecting shallow and deep clusters. Our preliminary results of using coda wave interferometry and inter-source interferometry show a discontinuous and short-term velocity change pattern with intriguing variations. We would complete dv/v measurements of all possible waveform pairs and recompose them for a long-term and continuous change, potentially help understand magma transport or fluid pressurization.

4.1 INTRODUCTION

Probing a volcano's deep magmatic system is of critical importance to understanding the active volcanic systems and assessing the eruptive hazards. Geodetic and seismic measurements are two main approaches for long-term monitoring. While geodetic measurements, such as Global Navigation Satellite System (GNSS), interferometric synthetic aperture radar (InSAR), and ground-based strain-/tilt-meter, could provide sufficient resolution of ground surface deformations[?], they may miss the deep activity that does not cause surface deformation and has limited resolution in the depth-dependent deformations[?]. Seismic data can serve as complementary measurements for shallow- and deep-depth monitoring. Seismic monitoring of earthquakes and overall seismic power emitted at the volcano are the cornerstone of volcano monitoring[?]. Often, a seismic network and

sophisticated methods are necessary for the accurate location of earthquakes[?] .

Because the mechanical properties of rocks vary under stress conditions, crack orientation, pore fluid pressure and composition, monitoring the changes in seismic wavespeed, polarization, and attenuation is becoming a popular approach to tracking the evolution of subsurface properties of volcanoes^{????} . Recent eruptions in Iceland have demonstrated the success of integrating such monitoring into operation^{??} . Uncovering elastic property changes (e.g., seismic wavespeeds or velocities) is therefore helpful to understanding and capturing interactions between crustal deformations and magmatic plumbing systems.

Several passive seismic interferometry techniques have been frequently utilized to track temporal changes in seismic velocities. Ambient noise monitoring has recently gained much attention given its potential for continuous monitoring[?] . While at occasion successful in indicating pre-eruptive activities^{??} , whether velocities increase or decrease before eruption depends on the eruptions. Furthermore, it is difficult to distinguish the cause of these changes as they may come from fluid migrations (magmatic or hydrothermal) or rock damage. Often, surface conditions such as tidal and atmosphere stress, snow loading/melting, and rainfall or precipitation also affect these velocities and should be accounted for^{??} . These effects are particularly strong on glaciated volcanoes and in MSH^{??} . Therefore, probing deep velocity changes remains to be undertaken.

Inter-source interferometry can recover the empirical Green's function between two earthquake sources^{??} , similar to inter-source interferometry, which retrieve the approximation of Green's function between two stations. Under approximate source-receiver geometry that ensure the wavefield in the station phase zone, the body and coda wavefields between two sources can be reconstructed with negligible phase error[?] . We extend this technique to repeating earthquakes to explore the possibility of probing changes in inter-source velocities. This method would complement current receiver-based interferometry methods to probe deep structural properties^{??} and even refining inter-event source distance^{??} .

This study develops a workflow to extract deep velocity changes by leveraging the automated detection of repeated seismicity at Mount St. Helens (MSH)^{??}, the longevity of the seismic monitoring, advances in deep learning to pick phases in seismograms, and aforementioned seismic interferometric methods. Hotovec-Ellis et al.^{??} have generated a continuous record of seismic velocity changes in the shallow subsurface and demonstrated dominated by seasonal changes. This study focuses on the post-eruption phase between 2009 and 2023.

4.2 DATA

We leverage the comprehensive detections of repeated seismicity from the (RedPy catalog, last accessed 2024/01/01) at MSH. We select 1,663 families –referred to as clusters in RedPy– in total, including 468 clusters obtained from borehole stations only and 1,195 clusters detected by surface stations. The distribution of the seismic network is shown in Figure 4.1a. We find common events among the two sets of families, and we eventually recompile the event information and waveforms of 1,544 unique families. The number of individual earthquakes in each family varies from 2 to 1,536. Such irregular occurrence can be recognized in Fig. 4.1c and Fig. C.1. Overall, all of these families have 16,064 earthquakes that occurred from January 1st, 2009, to January 1st, 2024. We downloaded waveforms of individual earthquakes within each family according to the start times of the cataloged earthquakes. One example of similar waveforms for one family can be seen in Fig. 4.1b.

4.3 METHODS

4.3.1 PHASE PICKING

To locate these families, we pick the arrival times of P and S phases on the realigned and stacked waveforms (e.g., the red waveform in Fig. 4.1). Since the REDPy workflow on the Cascades volca-

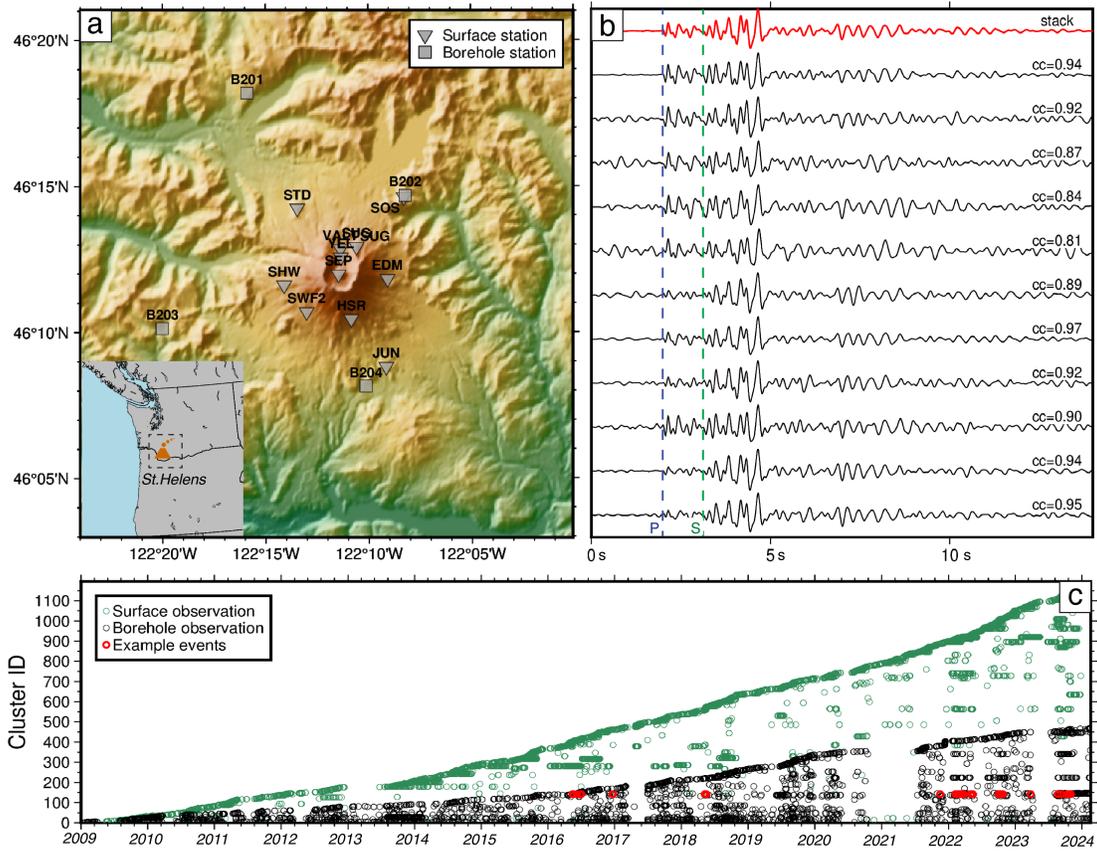


Figure 4.1: Study region and earthquake detections. (a) Borehole and surface stations at Mount St. Helens from the UW, CC, and PB networks; (b) Repeating waveforms (in black) of one example cluster. The stacked waveform is in red. The blue and green lines represent the arrival times of P and S waves. Correlation coefficients between individual waveforms and stacked waveforms are shown top right of waveforms; (c) Temporal evolution of each family: the black and green ones are detected by borehole and surface stations, respectively. The red dots are the example with waveforms shown in (b).

noes only selects earthquakes with a minimum correlation coefficient (cc) of 0.7 at least 4 stations, we assume the events in the same family are nearly co-located^{???}.

We download waveforms of 60 second duration, 15 seconds before and 45 seconds after the time of the RedPy detection. We do not remove the instrumental response, but we filter the waveforms between 1 and 10 Hz. All sensors we use as velocity meters (e.g, channel BHZ and EHZ). For each source-receiver group of waveforms in a single family, we first realign the waveforms to the first event using cross-correlation phase shift. In the second step, we re-calculate the correlation coefficient (CC) and obtain the stack waveform by stacking all shifted individual waveforms with CCs larger than 0.7. Those noisy repeating waveforms are thus discarded.

For each family, we thus gather the stacked waveform of each station, increasing the signal-to-noise ratio and enabling a more robust phase picking. We use our recently developed ensemble-learning-based phase picker (ELEP, see Chapter 7 and[?]) to pick P- and S- phases. The base models are the EqTransformer models[?] that have been trained from scratch for 6 different benchmark data sets[?] and made available through SeisBench[?]. The ensembling method used here is the *semblance*. Examples of stacked waveforms with ELEP-derived picks are shown in Fig.4.3. For ground-truthing, we also manually picked 2440 stacked waveforms. The performance of ELEP is thus evaluated against these manual picks in Fig.C.2d. We keep the P and S phase picks for the stations that are three components and only the P phase picks for stations with single vertical channels because single-channel input degrades the picking performance and S phase picks are not predicted confidently. We choose a threshold of 0.1 for P and 0.05 S for S waves to select the minimum confidence of phase picks. We obtain a total of 13652 P picks and 12052 S picks.

4.3.2 FAMILY EVENT LOCATION

With ELEP-picked phases, we employ the NonLinLoc developed by Lomax et al.[?] to globally search for locations of the 1,544 families in a 3D velocity model. The stacked waveforms represent

an averaged event for all events that belong to a single family. The synthetic P- and S-traveltime tables for all 16 stations are produced by running Eikonal equations[?] in the 3D tomographic velocity models[?].

We first calculate an initial location of 1,544 families. The equal differential-time (EDT) is utilized as the objective function in the inversion is:

$$L(\mathbf{x}) = \left[\sum_{a,b} \frac{1}{\sqrt{\sigma_a^2 + \sigma_b^2}} \exp \left(- \frac{\{ [T_a^O - T_b^O] - [TT_a^C(\mathbf{x}) - TT_b^C(\mathbf{x})] \}^2}{\sigma_a^2 + \sigma_b^2} \right) \right]^N, \quad (4.1)$$

where \mathbf{x} is the earthquake location, T_a^O and T_b^O are the observed arrival times and TT_a^C and TT_b^C are the calculated travel times for two observations a and b ; the sum is taken over all pairs of observations, and N is the total number of observations. Standard deviations σ_a and σ_b summarize the assigned uncertainties on the observed arrival times and calculated travel times. As tested by Lomax et al.[?], the EDT misfit function is more effectively compared to classic single-station based misfit.

We show the initial locations in Fig.4.2a-c with the above objective function for 3D global search. For further refine results, the source-specific, station-term (SSST^{??}) traveltime corrections is performed in 3D volume, specifically for each station and seismic phase type. Given an initial set of gridded travel-times and event locations, 3D grids of SSST corrected traveltimes for each station-phase are created iteratively[?]. The EDT location is then performed after each SSST corrections on calculated traveltimes. We perform two iterations by changing the radius of source regions from 4 km to 2 km and found it sufficient to achieve better results. Three iterations are usually needed to update locations in using 1D velocity models[?], we find that the third iteration has a strong distortion in locations and degrade location results in using 3D velocity models. Besides, as we are using 3D velocity models for locations, SSST relocations do not produce significant improvement results as seen in Fig.4.2e-f. The SSST is mainly designed for improving the initial locations of using 1D velocity model[?]. We also report that additional constraints to the source relocation, such as NLL-

coherence and double-difference, may not help refine location results much because those families are unique and dissimilar to each other.

4.3.3 CODA-WAVE INTERFEROMETRY

We have at least two events for each family. To solve for change in velocity among repeating events, we apply the coda-wave interferometry² to extract time variations in coda waves. In particular, we utilize the wavelet-domain stretching, WTS² in specific frequency bands. Currently, we focus on the frequency band of 2-6 Hz, which dominates the signal power, and on the vertical component seismograms. We first preprocess individual waveforms by realigning waveforms based on the P and S waves ($[T_p-1s, T_s+1s]$) using cross-correlation and discarding waveforms with cc values less than 0.7. An example of selected waveforms is shown in Fig.4.4a. Like time domain stretching, the optimal stretching factor is found until the highest similarity between wavelet transform is reached.

Moreover, we calculate the inter-source interferometry (ISI) between two target events. ISI technique² can generate approximations to the Green's functions (GFs) between two events provided some conditions about the location of the sensors at the surface with respect to the inter-source axis. We utilized the processing workflow from² describe to meet the conditions to recover the Green's function from the inter-source cross-correlation: 1) the pair of waveforms of two repeating events recorded at the same stations are first aligned to their respective origin time of events, 2) for each station, we cross-correlate waveforms of various repeating events, 3) the cross-correlation functions at each station are stacked, and 4) the individual cross-correlations are selected if their correlation coefficient with the stacked cross-correlation is greater than 0.5. Examples of these selected cross-correlation functions between one reference event and a family of repeated sources are shown in Fig.4.4b.

The parameters chosen to perform the coda-wave interferometry are identical to the ones that would be chosen for inter-station interferometry. The coda window of the cross-correlation is

estimated as 0.5 seconds before the S waves and 5.5 seconds after the S waves, assuming a shear wavespeed of 3.5 km/s. The coda window is 5 seconds. We eventually perform WTS for extracting velocity changes between individual functions and the stacked one, but only for the cc value (before stretching) larger than 0.7.

4.4 RESULTS

4.4.1 EARTHQUAKE LOCATIONS

As we can see in Fig.4.2, the repeating earthquakes are quite distributed, not necessarily localized on sharp features that would highlight faults. The near-surface earthquakes or those events above stations are found much scattered and located with high uncertainties because of distorted location probabilities[?], inaccurate near-surface velocities, and/or large uncertainty in observed picks.

Additional corrections or constraints need to be performed. In this study, we focus on subsurface earthquakes with depths from 1 km to 20 km instead of those near-surface events, for which we have uncertainties. The PNSN has also reported that it is not handling the depth of shallow events at Mt St Helens well (Hartog, *pers.comm.*). We performed a test to characterize the effects of 3D varying seismic structure relative to the 1D velocity structure that the PNSN uses for event location. We filter out paired events within origin-times difference within 3 seconds between PNSN catalogued events and REDPy detected multiplets. Similar to ?, we find that including a 3D varying structure shifts the best solution of the earthquake location shallower.

Compared with the PNSN catalogued earthquakes, we found that the velocity model has a large effect on the location, particularly for the source depth (Fig.C.2). In contrast, the general location pattern is similar to PNSN catalogued earthquake locations Fig.C.3b. Furthermore, the northern earthquakes seem to be more diffused than the locations in the south, and most earthquakes are located near the crater. From the depth distributions, we observe a significant bifurcation at a depth

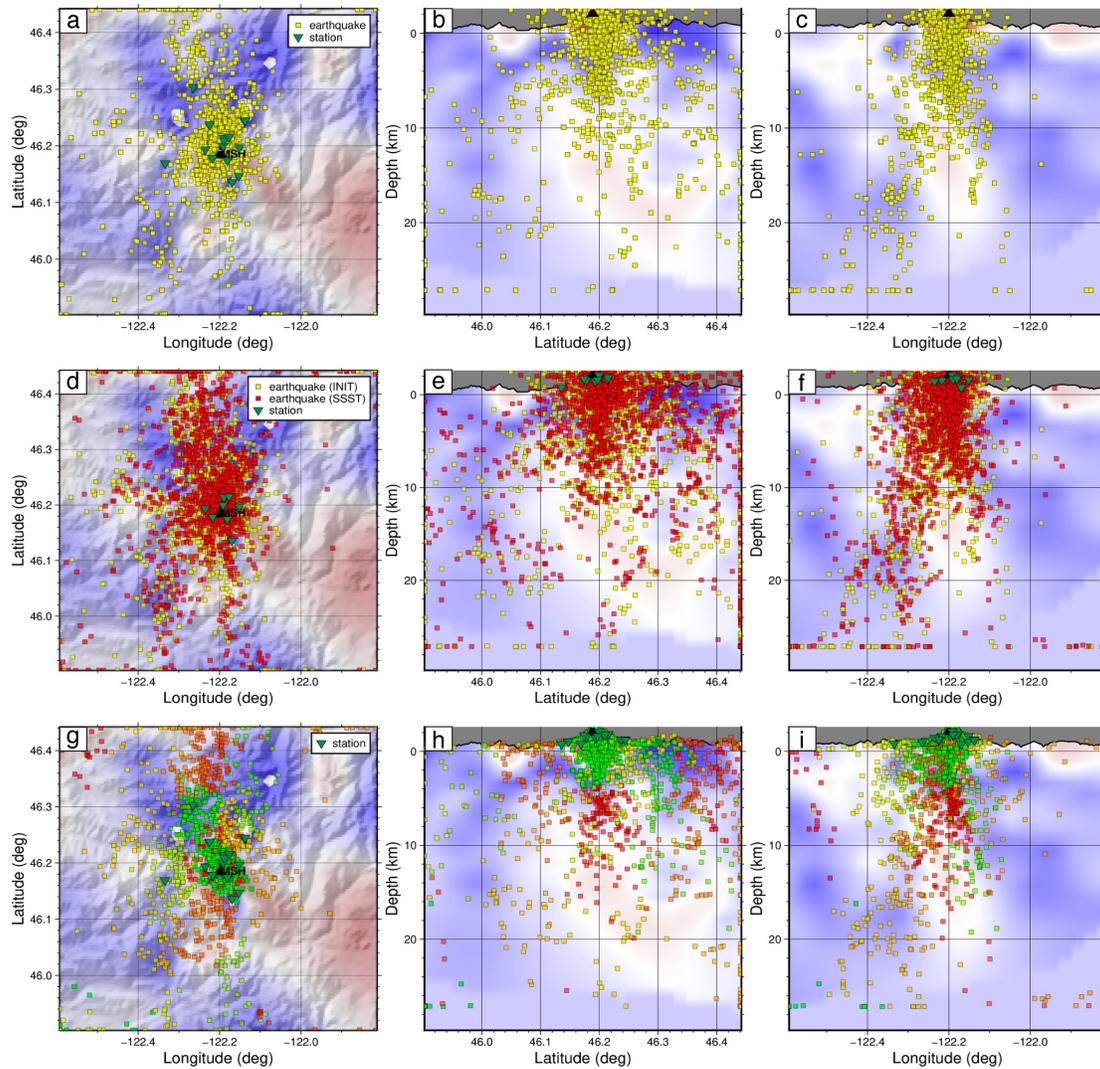


Figure 4.2: RedPy Event Families location in horizontal and two cross-section views. (a-c) NonLinLoc initial locations of REDPy earthquakes are in yellow. (d-f) Relocations with corrections of source-specific station terms in red squares. (g-i) 20 groups in different colors are obtained from K-means clustering with a gradual colormap. Note: background images are the P-wave velocity (V_p) perturbations with respect to the average V_p at each depth[?].

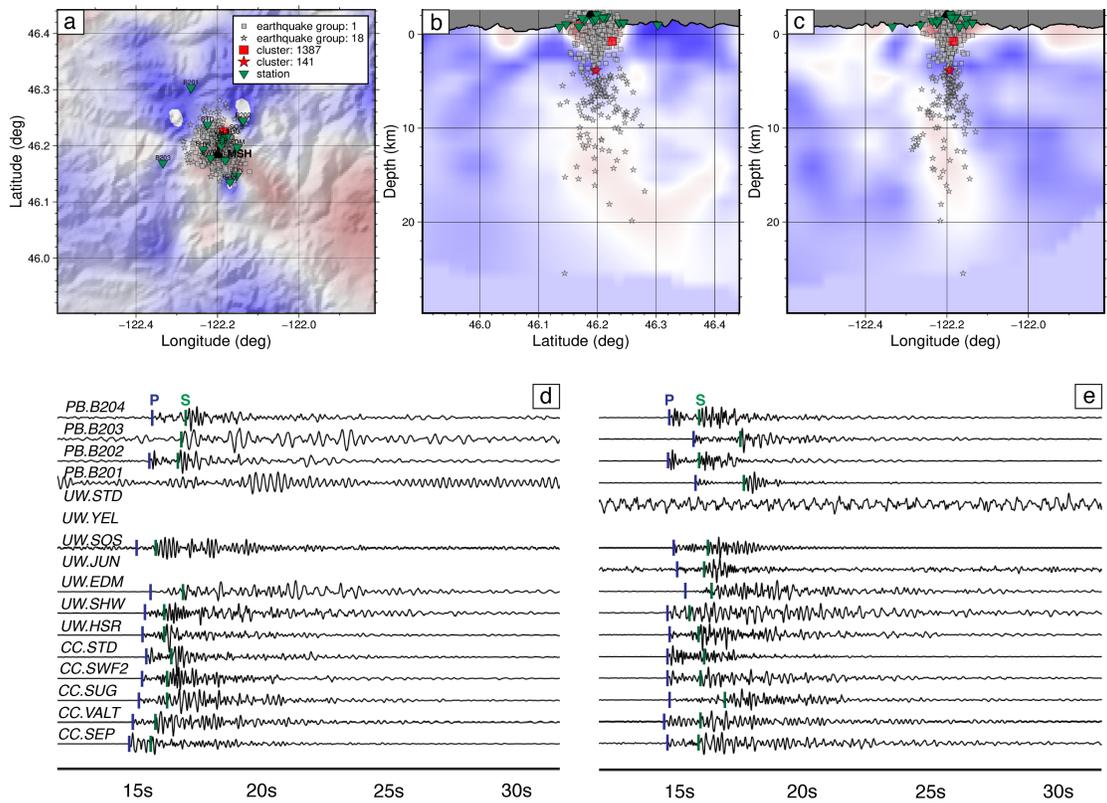


Figure 4.3: Selection from repeated signals. Two events with the highest repetitions from the shallow-depth group and intermediate-depth group. Their waveforms with ML picks for 16 individual stations are displayed below the geographic maps.

of 10 km in Fig.4.2c and f, mainly around the negative P-wave velocity perturbation zone (i.e., red region). As interpreted from² and multiple seismic imaging investigations², the negative anomaly indicates the magma storage zone due to partial melting. This magma chamber is active and may replenish the shallow chamber around the depth of 2 km.

These repetitive events can be leveraged to understand activities associated with the plumbing system. We categorize the whole dataset into 20 groups by using K-means clustering² and a defined number of groups set to 20. The number of groups is referenced from other clustering algorithms, such as DBSCAN² and self-organized maps², which automatically generate the number of groups. Those grouped events are shown in Fig.4.2g-i color-coded in a colormap. In the next section, we focus the analysis on extracting data from two clusters: a shallow and one at intermediate depth above the chambers (< 10 km).

4.4.2 CROSS-CORRELATIONS

We perform the cross-correlation by selecting the stacked waveforms of a given redpy family from either shallow or deep cluster and cross correlating it with the individual earthquake waveforms of a family in the alternate cluster. We show an example of waveforms and their cross-correlations in Fig.4.4. To highlight the similarity between waveforms, we calculate the correlation coefficient of the coda window to be cross-correlated between 2 and 6 seconds after the S wave. We find that some waveforms have a low correlation coefficient. In fact, these windows seem to exhibit a difference in relative P-to-S amplitudes yet a similar P-to-S travel time, which indicates that the focal mechanism of these events might be different than the composite or averaged mechanism. These windows are rejected in the selection for the cross-correlation. The resulting cross-correlation shows a mostly anticausal signal with an expected arrival at the predicted S-wave arrival time. Groups of waveforms that are clustered in time exhibit high similarity among themselves. However, the inter-group similarity is low.

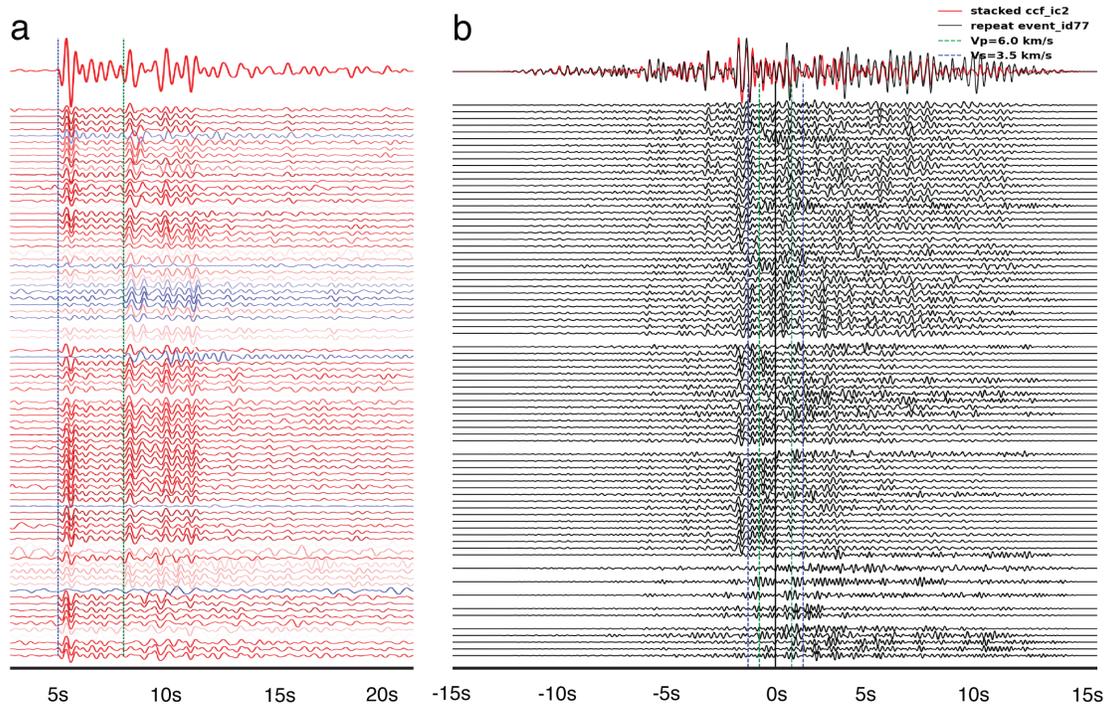


Figure 4.4: (a) Example waveforms of repeating events recorded at station *CC.STG*. The cross-correlation value between individual waveforms and stacked waveforms (at the top) is coded in color. Red color represents closer to 1, and blue color represents closer to 0. (b) Inter-source cross-correlation functions between reference or stack waveform in Fig.4.3d of one cluster and repeating waveforms of another cluster. Locations of two clusters are displayed in Fig.4.3a-c.

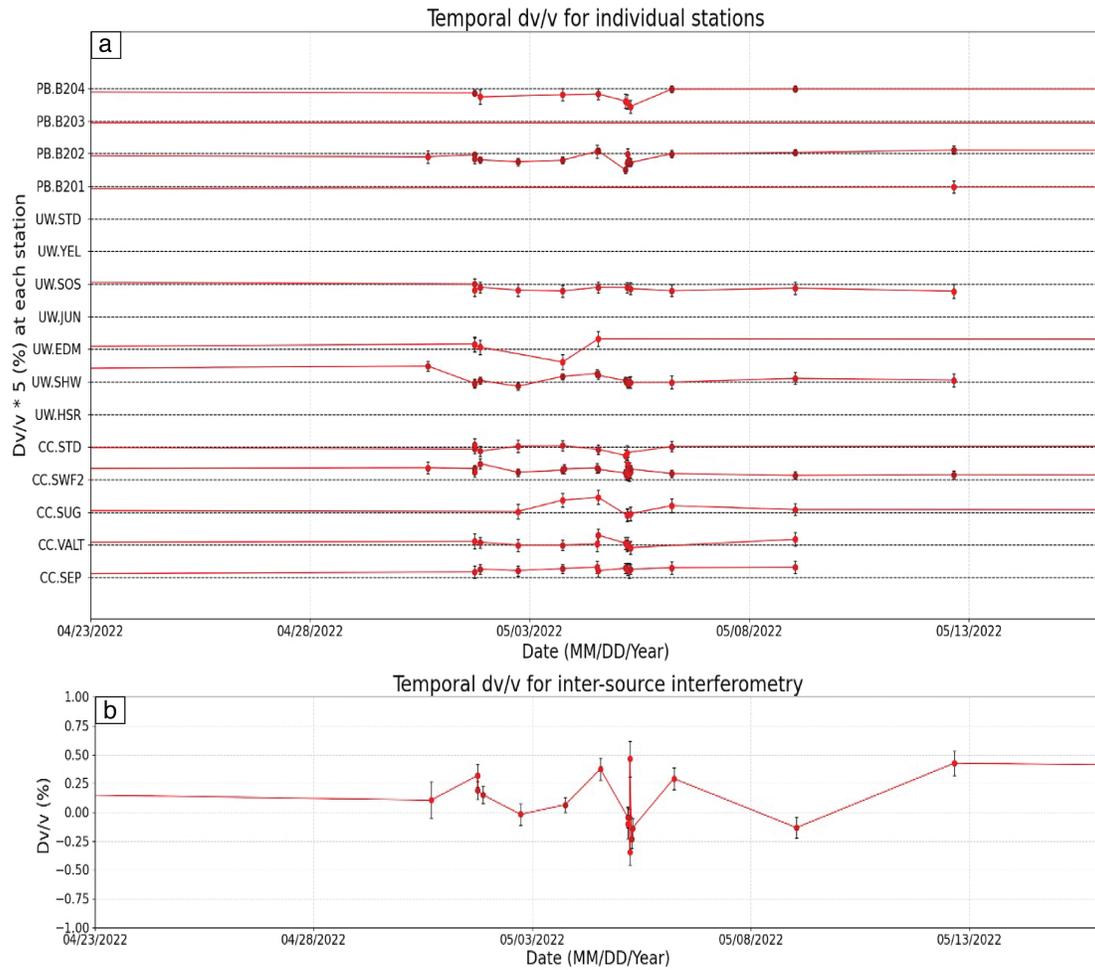


Figure 4.5: (a) velocity changes for each receiver; (b) velocity changes for inter-source interferometry.

4.4.3 VELOCITY CHANGES

We generate the dv/v time series by stretching cross-correlation to the averaged cross-correlation. Given the intermittency of the reply cluster, we obtain velocity changes irregularly sampled in time for each source-receiver path in Fig.4.5a and inter-source region in Fig.4.5b.

Despite sparse sampling, we could recognize that overall variations are consistent across different measurements at stations and inter-source. One striking observation is a strong reduction around May 5th, 2022, when a swarm was frequently (every hour) occurring. The uneven distribution hinders us from observing gradual variations in velocities. These observation gaps can be filled by using all members in one specific group and assuming they experience the same velocity changes.

4.5 CONCLUSION

Deep volcano's magmatic unrest patterns are challenged to be deciphered. Leveraging the detection of repeating earthquakes at Mount St. Helens between 2009 and 2023, we explore a novel method to extract big changes in seismic velocities that are typically more difficult to extract given the near-surface strong environmental variations and measurement sensitivity.

We have developed a workflow that integrates various machine learning strategies and conventional but state-of-the-art methodologies. After the semi-supervised detection from the RedPy workflow, we pick the arrival time of P and S waves using an ensemble deep learning algorithm of the family centroid waveform event, locate this centroid using nonlinear location in a heterogeneous 3D velocity structure. We then group the families of events by their location using K-means clustering. Once grouped, we select near vertical paths to extract changes in seismic velocities observed between sources.

Our preliminary result of estimated velocity changes suggests some interesting changes and needs additional work to compile a near continuous time series using all pairs of families, as done in??.

After obtaining those changes at different depths, we hope to interpret the deep changes in relation to deep seismic activities to characterize the evolution of the magma chamber in this post-eruption phase.

*Yuan C, Cochard T, Denolle M, Gomberg
J, Wech A, Xiao L, Weitz D. Laboratory hydrofractures as analogs to tectonic tremors. AGU
Advances. 2024 Feb;5(1):e2023AV001002.
<https://doi.org/10.1029/2023AV001002>*

5

Laboratory Hydrofractures as Analogs to Tectonic Tremors

SUMMARY

The fracture of Earth materials occurs over a wide range of time and length scales. Physical conditions, particularly the stress field and Earth material properties, may condition rupture in a specific fracture regime. In nature, fast and slow fractures occur concurrently: tectonic tremor events are fast enough to emit seismic waves and frequently accompany slow earthquakes, which are too slow to emit seismic waves and are referred to as aseismic slip events. In this study, we generate simultaneous seismic and aseismic processes in a laboratory setting by driving a penny-shaped crack in a transparent sample with pressurized fluid. We leverage synchronized high-speed imaging and high-frequency acoustic emission (AE) sensing to visualize and listen to the various sequences of propagation (breaks) and arrest (sticks) of a fracture undergoing stick-break instabilities. Slow radial crack propagation is facilitated by fast tangential fractures. Fluid viscosity and pressure regulate the fracture dynamics of slow and fast events, and control the inter-event time and the energy released during individual fast events. These AE signals share behaviors with observations of episodic tremors in Cascadia, United States; these include: 1) bursty or intermittent slow propagation, and 2) nearly linear scaling of radiated energy with area. Our laboratory experiments provide a plausible model of tectonic tremor as an indicative of hydraulic fracturing facilitating shear slip during slow earthquakes.

5.1 INTRODUCTION

EARTHQUAKES can happen within seconds for fast or regular events to years for slow earthquakes. The events that rupture fast at the Rayleigh wave velocity excite seismic waves that can result in colossal damage to urban infrastructures. Conversely, earthquakes with slow rupture and slip velocities, three to four orders of magnitude lower than the Rayleigh wave speed, do not cause significant

damage and are only detectable using remote sensing techniques. Slow earthquakes appear as aseismic slips on shear faults, but they often co-evolve with weak coherent ground motions known as tectonic tremors^{???}. Slow slip and accompanying tremor events are frequently observed with a wide range of sizes and velocities^{???}. Unlike the seismic signals of fast earthquakes, tectonic tremor is emergent, prolonged in duration, often depleted in high frequencies^{??}, and may exhibit behaviors of intermittency and burstiness^{??}. Tectonic tremor was discovered using seismic arrays[?] and can be located using the coherence of the envelope of the signals^{??}. It is detected worldwide, mainly in subduction zones^{??????}, but also on strike-slip faults^{??}, and other plate boundaries^{??}.

The coexistence of slow and fast earthquakes also appears in lab experiments^{??}. Lab-generated slow-slip events are observed acoustically as low-amplitude tremors, while fast-slip events have impulsive and energetic acoustic signals^{??}. Tremors are observed during the acceleration of pre-slip before large earthquakes[?]. Studies have found that slow earthquakes may promote or inhibit large, fast megathrust earthquakes in subduction zones^{??}. Investigating slow earthquakes may improve our understanding of the triggering mechanism and potential area of fast earthquakes[?].

The observed coexistence of slow slip and tectonic tremor has led seismologists to use tremor, or its constituent low-frequency earthquakes (LFEs), as markers of slow earthquakes. Their timing can help detect and locate slow-slip events^{???}, estimate their size^{??}, and infer the mechanical and stress state on the fault^{??}. Despite the ample observations of tectonic tremors, their source mechanisms and relation to slow slip remain uncertain. Studies of radiation patterns of stacked LFEs suggest that the LFEs that constitute tectonic tremor may coincide in space with shear slip on the megathrust, where fluid pressures are thought to be high^{??}. Recently, ? and ? suggested an alternative model whereby tremors may be markers of hydrofracture unclamping a fault that moves in shear motion, generating some component of tensile motions. While the physical model of a single force for the seismic source may be unrealistic[?], the concept of tensile motions, particularly those facilitating shear slip, corroborates an alternative model to LFE generating processes proposed

previously (e.g.,² and references therein). The subduction of Earth materials generates free fluid from the dehydration of minerals during their phase transformation²². Such fluid generation and storage released are thought to play a major role in the genesis of tremors and slow slip involving fault dilatancy and compaction²³. Fluid diffusion, whether it is constant or intermittent, may be prevalent in subduction zones where brittle fractures are expected²⁴.

Although slow earthquakes have been induced in dry experiments by regulating slip rates, there have been limited investigations on the hydrofracturing mechanism to generate tectonic tremors. This study examines whether stick-break instabilities of tensile cracks may contribute to the processes by which tremors are observed during slow slip events. We inject overpressured viscous fluid in an intact transparent rigid material and observe the nucleation and propagation of fractures that generate seismic signals similar to tectonic tremors. We then investigate the interaction between fluids and fractures to understand better the evolution of slow slip rupture source area and radiated seismic energy in a laboratory setting. Finally, we observe and discuss similarities between the laboratory and natural cases, particularly the near proportionality between cumulative radiated seismic energy and tremor area and the intermittent-burst behavior of stick-break instabilities.

5.2 METHOD

5.2.1 EXPERIMENTAL APPROACH

The experimental apparatus is schematically shown in Fig. 5.1a. More details, including a depiction of the experimental apparatus, are provided in Figs. D.1 and D.2. To build a stiff and transparent sample with a notch included in the design to inject the fluid, we apply stereolithography (STL) 3D printing of polymethylmethacrylate (PMMA) using a FormLabs Form 3 printer. The sample consists of a transparent cylinder of 100 mm in diameter and 32 mm in height. We use a high-pressure pump (Teledyne Isco 65D) to inject the fluid into the PMMA sample at a constant flow rate of 0.3

ml/min. A pressure transducer measures the inline fluid pressure and indicates that the sample starts breaking at a pressure of about 30 MPa (FigD.3). At this instant, the compressed fluid expands suddenly and drives a fracture. The fluid is dyed with fluorescein, allowing us to visualize the fracture and the fluid independently. We use a high-speed camera (Vision Research Phantom TMX 6410) that records 500×500 pixel images ($200 \mu\text{m}/\text{pixel}$) at 100,000 frames per second ($10 \mu\text{s}/\text{image}$). With the images from the high-speed camera, we create subtracted images, taking the difference of each pixel between each frame and the first one, to track the cumulative fracture area. We also create differential images, taking the difference of each pixel between two consecutive images, to visualize detailed tangential fracturing. In addition, we use four Glaser-type broadband sensors (KRNBB-PC) to record acoustic signals associated with the fracturing process. Those acoustic-emission (AE) receivers have an exceptional frequency range of 20-1000 kHz with nearly flat instrumental response in this frequency range². These four AE sensors are placed at each quadrant of the specimen and record acoustic signals at a frequency of 2 MHz. The experimental schematic is displayed in Fig. D.1a and detailed in Figs. D.1-D.3. We perform two fracturing experiments varying the injected fluid viscosity to be $\mu=1$ cP and 800 cP in each. The complete fracture front and AE recordings are displayed in Figs. D.4 and D.5. We also display the spectrograms of these AE signals in Fig. D.6.

5.2.2 AE NUCLEATION LOCATION

The difference in first arrival times of the AE signals allows us to estimate the nucleation location of the AE sources. Often, the AE signals interfere with each other (Fig. D.5), especially due to waves reverberating in the sample, so we do not use all of them. In the high-viscosity experiment, we pick the arrival times of acoustic waves from discernable stick-break events. We utilize the relative arrival time between receivers that are sufficient to locate the events, and subsequently increase the location accuracy using waveform cross-correlations. We pre-calculate all theoretical travel times using the compressional (P-) wavespeed found by performing calibration tests (Fig. D.7) and all possible

source locations in a polar coordinate system (radial distance, r , from the center of the specimen and, θ , the azimuth angle taken from the East direction). We then perform a grid search as a global inversion to find the original location of the AE signal. The misfit function L is defined as:

$$L(r, \theta) = \sum_{ij} (tt_{ij} - TT_{ij})^2. \quad (5.1)$$

where i and j denote the indices of the receivers. TT is the theoretical travel-time difference between receivers i and j . tt is the observed arrival time difference between receivers i and j (origin time cancels). The neighboring receivers are paired as four groups for the inversion. We demonstrate the example process of inverting the location of the first fracture event from their four arrival picks in Fig. D.8. All fracture event locations are processed identically.

5.2.3 ESTIMATING RADIAL AND TRANSVERSE FRACTURE AREA

By image processing, we track the contour of the slow radial fracture front to compute the radial fracture area over frame time. Since we are imaging at a constant frame rate of 100,000 frames per second, we can calculate the area of the crack surface every 10 μ s. As observed in Figs. D.4 and D.8, measurements include both the radius of the slowly expanding aseismic radial crack and the area of fractures that propagate tangentially to this front at seismic speeds. The latter manifests as illuminated pixels in sequential differential images and as AE signals. The details can be seen in Movies S1 and S2 and are described in the later results. We focus on the propagation phases to avoid signal saturation at the beginning of the experiment and boundary effects as the radial crack reaches out of bounds in both low-viscosity and high-viscosity fracturing experiments. There are some differences in area measurement for the two experiments. In the high-viscosity experiment, we calculate the transverse fracture area of discrete AE events, which lasts longer than the 10 μ s measurement interval, as the differential area between the start and end times of the AE events. These individual

fracture areas are then summed up over time and because there is no radial propagation in between AE events, this equals the cumulative radial fracture area. In the low-viscosity experiment, due to continuous and weak AE signals, we simply define the start and end times of a moving window of a fixed length (10 μ s). An example of estimating the radial fracture area as a function of time is shown in Fig. D.9. The radial fracture area consists of multiple transverse fracture areas.

5.2.4 ESTIMATING THE AE ENERGY

To calculate the radiated energy of a single AE event, we need to correct the acoustic signals for path effects as the amplitude decays due to geometrical spreading and attenuation with distance. Because surface waves dominate the amplitudes of the wavefield that propagate in the traction-free sample, we use the square root of the distance as the geometrical spreading term. We model attenuation as an exponential decay with an attenuation parameter for the material absorption and scattering[?]. Therefore, the following equation is utilized to correct the signal amplitude loss:

$$A = A_0 \sqrt{x} e^{-\alpha x}. \quad (5.2)$$

where A_0 is the raw AE signal in voltage, x is the distance between the transverse fracture front and receiver, which is the receiver- or azimuth-specific (e.g., the radius is measured in the northwest direction to correct the signals recorded by the receiver located in the same corner). As suggested by[?], the attenuation coefficient $\alpha = 21.5$ for PMMA using a peak acoustic frequency of 75 kHz. The spectrograms of recorded acoustic signals demonstrate the peaks of the dominant acoustic power around 40 kHz, as shown in Fig. D.6. We then perform a sensitivity analysis over a range of attenuation coefficients between 0.01 and 100. Here we choose $\alpha = 20$ to correct for attenuation. We report that an attenuation coefficient of less than 50 does not impact our main results. We show the results for these correction experiments in Fig. D.10.

After the signal correction, we calculate the total AE radiated energy for four individual receivers using:

$$E = \int_{t_0}^{t_1} A^2 dt. \quad (5.3)$$

where E is the radiated energy calculated from voltage (AE) data, which is shown to be proportional to the kinetic energy in a drop-ball test and, therefore, proportional to the elastic strain energy (Figs. D.11 and D.12). We acknowledge that A is proportional to displacements and in voltage units, but our empirical calibration demonstrates that it is also a valid measure of radiated energy. Other studies have also approximated the AE radiated energy using voltage data to approximate the released elastic strain energy^{??}. Similar processing for tectonic tremor signals, such as attenuation correction and integration of velocity squared, is performed to calibrate and calculate signal energy^{??}. In equation 5.3, t_0 and t_1 are selected for each experiment for the energy integration. For the high-viscosity fluid experiment, we manually pick t_0 and t_1 to be the window for individual discernable AE events after high-pass filtering signals (50-200 kHz). The time window varies from the start of the fracture AE event until the start of the next event. For the low-viscosity fluid experiment, we use an interval of 10 μ s. We then compare the cumulative AE energy with the evolving fracture surface area.

5.3 RESULTS

5.3.1 STICK-BREAK INSTABILITIES

We conduct two experiments with varying fracturing fluids: the first with water (low viscosity, $\mu = 1cP$) and the second with a mixture of glycerol and water (high viscosity, $\mu = 800cP$). As studied by ?, a stick-break instability occurs in all experiments. The fracture propagates slowly in

the radial dimension overall (thousands of times smaller than the Rayleigh wave speed), but the slow radial fracture is accommodated by seismic transverse fractures, which propagate rapidly (close to Rayleigh wave speed). The break amplitude (i.e., when the radial fracture front advances and transverse fracturing occurs) and the stick time (i.e., time between breaks) increase as the fluid viscosity increases, which were also shown by ?. Here, we focus on a single small but fast seismic fracture.

In the high fluid viscosity experiment, the local nucleation of the fracture occurs mainly in the northeastern quadrant of the sample. Following the initial nucleation, the fracture expands transversely very rapidly, as shown by red patches in Fig. 5.1b. A single transverse fracture takes approximately 90 μs to revolve around the entire perimeter of the fracture. Given the fracture length, we estimate the fracture speed of 1000 m/s, very close to the experimentally measured Rayleigh-wave velocity in the 3D printed material (see glass-capillary break calibration in Fig. D.7). The AE signals of a single event recorded by the four sensors at each quadrant of the sample allow us to confirm the dynamics of these events. The first AE signal arrives at each sensor with a delay based on the location of its initiation point (Fig. 5.1c). Thus, we can triangulate the signals to locate the position at which the elastic waves originated with high accuracy (see Methods and Fig. D.8 for further details). We find that the position of the source of the signals coincides with the exact location of the fracture nucleation point identified visually, as shown in Fig. 5.1b. We conclude that the radiated elastic energy is released in the form of elastic waves due to the fast anti-plane propagation of a tensile crack. Furthermore, we reveal that the whole course of the transverse fracture propagation radiates seismic signals since, fortunately, we capture the surface wave associated with another fracture event in the later propagation stage (Fig. D.8).

To provide a basis for comparison, we conduct a water-induced fracturing experiment. In contrast with the high viscosity fluid experiment, we observe numerous individual events occurring with smaller amplitudes but much more frequently. The very small time interval between the nucleation of stick-break events in this case causes strong interference in their signals, which resemble

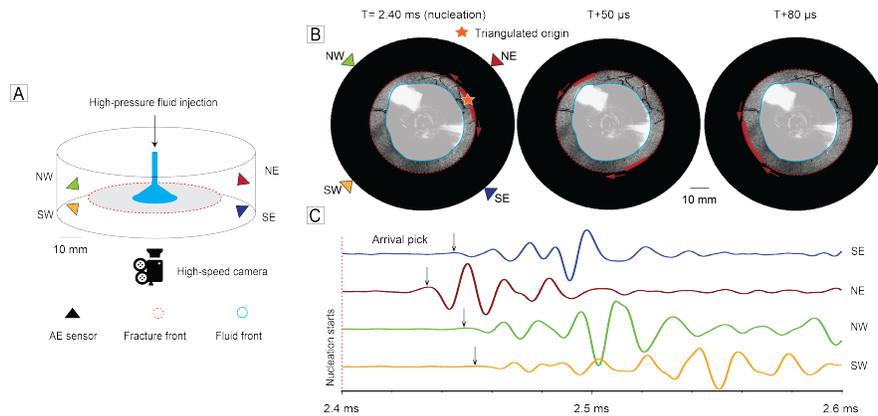


Figure 5.1: Experimental apparatus, subtracted images, and acoustic signals for the high-viscous fluid experiment ($\mu=800$ cP). (a) Schematic of the experimental apparatus with a high-speed camera and four AE sensors placed at each sample quadrant (SW, NW, NE, SE). The filled triangles represent the AE sensor locations. The dashed red and solid cyan lines are the fracture and fluid fronts, respectively. The cyan cone is the injected fluid. See additional experimental details in Figs. D.1 and D.2. (b) Three subtracted images from the first image after $T=2.4$ ms of the first burst show nucleation, propagation, and arrest of a single AE event. The red patch and arrow represent the transverse fracture area and direction, respectively. (c) AE signals associated with the fracture event in (b), the black arrows denote the picks of the first arrivals that we use for locating the AE source, which is marked as the orange star in (b).

seismic tremors or swarms. Despite the difference in frequency and amplitude, every event follows the same growth cycle from nucleation to arrest, as shown by tracking each individual fracture event in Fig. D.9.

5.3.2 CONSISTENCY BETWEEN OBSERVED FRACTURE DISCONTINUITIES AND AE SIGNALS

To further analyze the causal relationship between stick-break events and AE signals, we compare and associate transient AE signals with the radial fracture radii in the subtracted images and intensity variations from differential images (Fig. 5.2). We plot the kymographs, time representations of radial fracture radii taken in the northeast direction, for both fracturing cases with different fluid viscosities. In these representations, we highlight the fracture front with a dashed red line and the fluid front with a solid cyan line. The episodic displacement and stop pattern over the radial frac-

turing process is clearly observed on the top panel of Fig. 5.2a and b. In the high-viscosity fracturing fluid experiment, although the fracture speed is fast (≥ 180 m/s) for a single transverse event, the average radial propagation speed is slow (≈ 4 m/s), resulting from periods of pauses (i.e., stick events). The velocity of the fluid expansion is close to the average radial fracture propagation speed. The kymographs at other radial directions reveal similar fracture propagations as the contour of the fracture front is approximately symmetrical (Fig. 5.1b and Movie S1). Besides visualizing the optical images, four AE sensors record the acoustic signals emitted by the fracture nucleation and propagation. We can identify at least three clear events with recognizable associations with fracture discontinuities and peak intensity rates. The AE wave packet may include a single event or multiple events. As we observe, several smaller events that follow the third event occur in short periods, which results in strong interference between the signals. The density of concentrated signals thus suggests the frequent occurrence of transverse fractures. This gives rise to a faster average fracture propagation than earlier propagation with longer stops.

In contrast, for the low fluid viscosity experiment the kymography indicates a rather continuous slow fracture propagation relative to the first observation. The fluid front is very close to the fracture front, making them hardly differentiable. The pressure fluctuations during the fluid expansion induce frequent and small fractures. The average fluid expansion equals the average slow radial fracture velocity (≈ 2 m/s) and is close to the one estimated in the high viscous case. The intensity-rate curve with multiple peaks discloses a fast-paced fracture propagation. Accordingly, we observe continuous acoustic signals filled with concurrent events with signals that overlap with each other, which thus makes individual fracture events difficult to distinguish, as shown in Fig. 5.2b. Unlike the high-viscosity experiment (Fig. 5.2a), uncovering the origin of these signals is a greater challenge. The acoustic signals from either low-viscosity or high-viscosity experiments display very similar waveforms to tectonic tremors consisting of either clear or unclear discrete seismic events, as often observed in Japan and Cascadia.

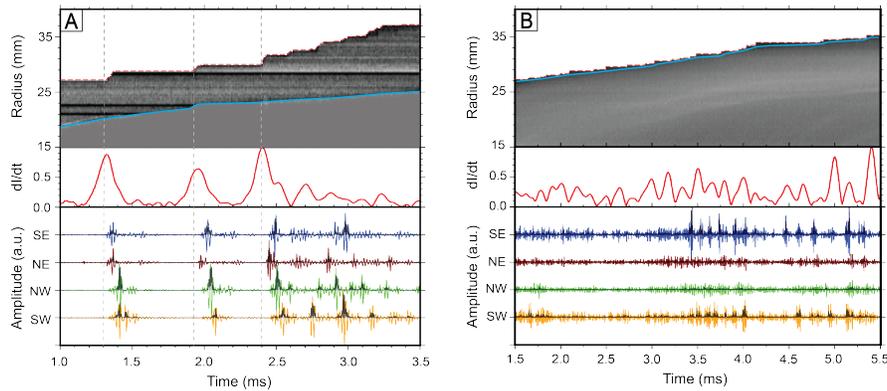


Figure 5.2: Kymograph, pixel intensity variation rate, and AE signals. Top panels: kymograph of the subtracted images taken in a NE direction of the specimens. The dashed red and solid cyan lines represent the fracture and fluid fronts in the upper panels, respectively. Middle panels: time series of the image intensity rate, which is measured as the averaged pixels of the differential images. Lower panels: acoustic signals recorded by the four AE sensors. The amplitude unit “a.u.” represents an “arbitrary unit” for their normalization to the peak amplitude. The back lines are the radiated energy rates of acoustic signals. (a) For a high fluid viscosity (diluted glycol, $\mu= 800$ cP) experiment and (b) for a low viscosity fluid (water, $\mu= 0$ cP) experiment. The full-length measurements are displayed in Figs. D.4 and D.5. Note: the time series is trimmed to focus on the hydrofracture propagation instead of its beginnings, which suffer from signal saturation, and endings, which are affected by sample edges effects.

5.3.3 SCALING RELATIONSHIP BETWEEN FRACTURE AREA AND AE ENERGY

We investigate the relationship between fracture area and AE energy and thus the fracture size can be characterized with AE energy. In the case of a fracturing fluid of 800 cP, there are multiple successive and very discernable tensile fracture events. By measuring the start and end times of each of these clear fracture events, we can calculate the fracture area by subtraction of images between the end and the beginning of the event and measuring the area (pixels) highlighted with high image intensity. We correct the effects of geometrical spreading and attenuation of the acoustic signals corresponding to each event and calculate their radiated energy as described above. We plot the cumulative AE energy against the cumulative fracture area and highlight a nearly linear relation for the first three events in Fig. 5.3a. To better support the interpretation of a near-linear relation, we split the acoustic signals arbitrarily after the third fracture event into two parts. After calculating the fracture surface area changes and AE energy based on the split time, we consider them as two individual

points. The larger variance at later times is also observed compared to the earlier times because of the uneven distribution of the events with respect to the different sensors. Averaging the energy over four sensors can eliminate the influence of azimuthal variations to a large extent. We confirm that the near-linear relation between cumulative radiated energy and cumulative fracture area remains valid using sensor-averaged measurements.

In the water (low viscosity) fracturing experiment, we cannot calculate fracture area and AE energy for individual fracture events, which are inseparable in time (Fig. 5.2b). Instead, we calculate fracture area and AE energy for each frame interval (10 μ s). The acoustic signals are corrected for geometrical spreading and attenuation using the transverse fracture front as location, calculate the distances between the transverse fracture fronts and the sensors, and calculate their respective energy, similarly to the previous case. We show the cumulative energy against the cumulative area in Fig. 5.3b again and observe the near-linear relationship between the two again. There is an exception for the NW sensor, which we interpret as a deviation from the circularity of the rupture front (e.g., Movie S2). Overall, the consistent observations in both high- and low-viscosity cases suggest that the fracture energy increases with the fracture area almost linearly, regardless of the fluid viscosity. This relation is further theoretically explained in supplementary Text 1.

5.4 COMPARISON WITH TECTONIC TREMORS

We observe two types of fracturing behaviors in our laboratory experiments: 1) the slow average radial fracture velocity (2-4 m/s) actually occurs as intermittent short-duration rapid radial advances (i.e., breaks) separated by long nearly stationary periods (i.e., stick events); 2) the fast advances nucleate at the radial fracture front and propagate transversely much faster (1000 m/s), at the Rayleigh-wave speed of the material. The radial fracture advancement is driven by the stress concentration applied at the crack tip. As shown in ?, the increase in the fluid viscosity leads to larger stick-break

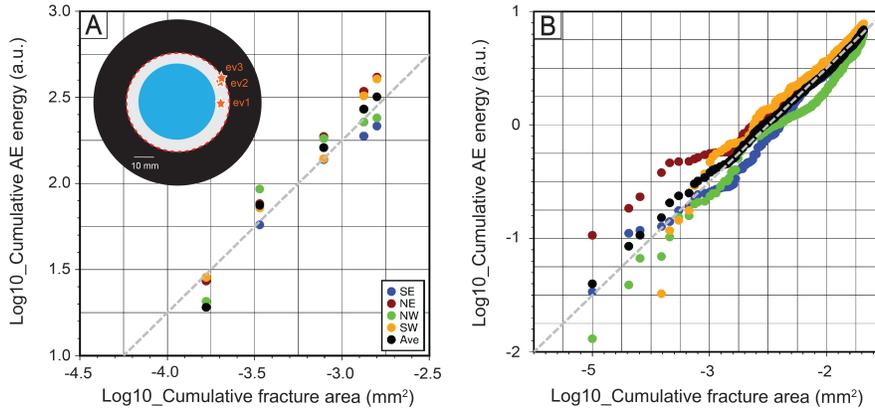


Figure 5.3: Laboratory relationship between cumulative fractured area and cumulative radiated energy. The colored dots represent the measurements of sensors at different azimuths. The black dots are the mean estimate of four sensor-specific values. The cumulative radiated energy is proportional to the fracture area for individual fracture events. The gray line is the best-fit linear regression using the averaged measurements (black dots). (a) High fluid viscosity experiment: the first three data points come from three clear events denoted as pink stars shown in the schematic diagram. They are recognized as the first three events in the kymograph panel in Fig. 5.2a. The late two measurements represent the signals after the third event, which are indistinguishable and divided into two groups arbitrarily. (b) Low fluid viscosity experiment: the AE cumulative energy is calculated directly on the continuous vibrations.

events, though does not alter much the overall radial fracture speed. In such a tensile-mode fracture, the stick-break instability results from extended fracture propagation because the pressure front is heterogeneous (i.e., is not strictly uniform in Fig. 5.1) despite the fluid being injected at a uniform rate, and seems ubiquitous to all-natural fractures[?]. Although in nature the slow slip often plausibly is too small to be measured, tectonic tremor is thought to consist of low frequency earthquakes and may be broadband^{???}, suggestive of a universal slow earthquake model^{??}. The tremors generated in our laboratory experiments also exhibit broadband characteristics (Fig. D.6).

5.4.1 PLAUSIBLE MECHANISMS FOR TREMOR GENERATION

The role of fluid migration in the generation of tectonic tremor has been proposed in multiple studies. First, where tremor occurs fluids likely come from the dehydration of minerals as they transform

with increasing temperatures and pressure[?]. Their pore-fluid pressure is inferred to be high, if not close to lithostatic^{??}. ? suggested that the rapid tremor migration events that occur backward relative to the slow, aseismic rupture front are induced by the slow propagation (10 m/s) of a wave of high pore pressure, which then reduces the effective normal stress and allows slip. Such pore pressure diffusion is evoked in the more general case of slow-slip and tectonic tremor by ?. Tremor fronts migration described as diffusive has been observed and inferred to reflect fluid processes in numerous other studies^{???????}.

The rupture mechanisms for low-frequency earthquakes that compose tremors are observed to be consistent with shear slip along the megathrust^{?????} that can be activated by overpressurized pore fluid pressure^{??}. Recent investigations by ? and ? argued similar measurements of seismic radiation pattern can be generated by a dipole or single-force mechanism, owing to insufficient station coverage. While shear deformations are evident in geological observations, the associated tensile veins generated at the depth of tectonic tremors also suggest that tensile fracture plays an important role in accommodating the deformation in the context of high fluid pressure^{???}. Here, we propose that tectonic tremor may not solely be shear dislocation alongside shear slow slip but could also facilitate shear slip through hydraulic fracturing that promotes the propagation of pore fluids and overall shear deformation (Fig. 5.4). We discuss two observations that coincide in the laboratory and nature.

5.4.2 STICK-BREAK INSTABILITIES AS A MODEL FOR INTERMITTENT TECTONIC TREMOR

Our laboratory experiments have demonstrated that the slow propagation of hydrofractures proceeds in a stick-break fashion, with inter-event times and fracture sizes driven by pressurized fluids that increase with fluid viscosity[?]. In nature, tectonic tremors also happen in bursts and exhibit temporal and spatial clustering^{???}. We present tremor observations during an example slow slip event in Cascadia in Fig. 5.5. Despite a general trend of tremors migrating southward in that spe-

Fault-valving model

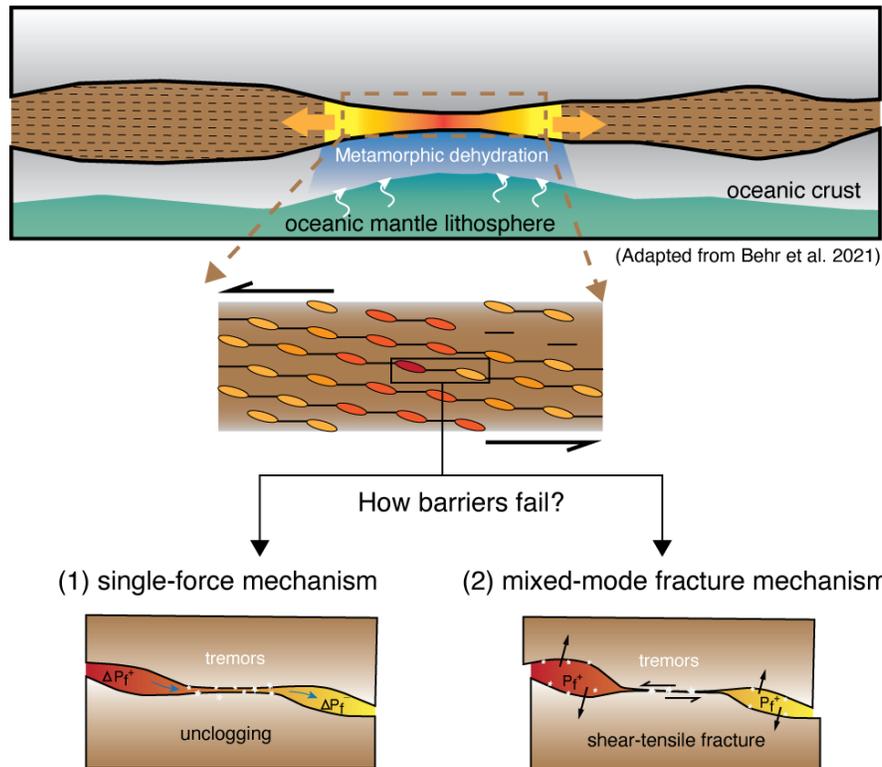


Figure 5.4: Illustration of tremor generation mechanisms in an overpressurized fault zone in subduction. Two primary mechanisms may influence tremor dynamics. The first, the single-force mechanism^{??}, proposes that tremors occur during transient fluid pressures as barriers unclog and fluid flows, which are indicated by blue arrows in the lower left panel. The second mechanism proposed in this study suggests a mixed mode of shear and tensile fracture resulting from increased fluid pressure. The black arrows represent seismic source body forces in the lower right panel. The tensile-shear model is also described by Sibson[?]. Evidence of these crack-induced tremors can be seen in many exhumed rock samples^{??}. Alternative models also include jamming of granular media^{??}.

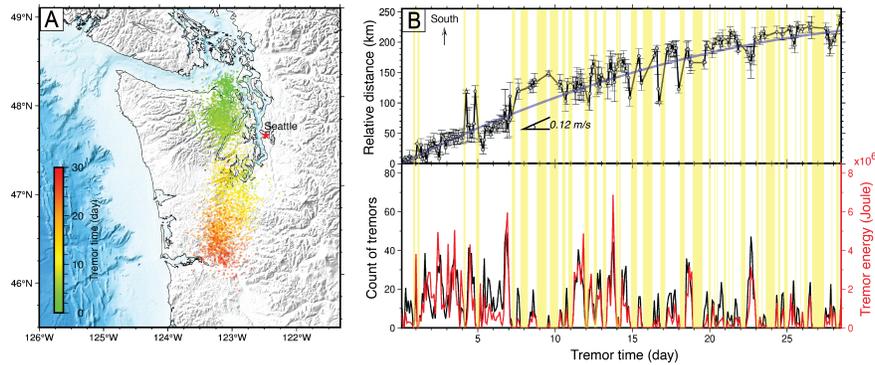


Figure 5.5: Example Tremor Swarm in Cascadia. (a) Individual tremor locations during one slow slip event that started March 10th, 2019, color-coded by occurrence time since origin. (b) Bursty tremors. The upper panel shows the southward propagation of centroid location along the latitude over time since its origin. Each dot denotes the relative distance in kilometers of centroid latitude of tremors within 2-hour intervals relative to the original location. The error bar indicates the latitude range (in kilometers) of each interval. The blue curve represents the averaged or smooth propagation over time. The lower panel shows the curves of the number and summed tremor energy in 2-hour intervals binned in the same way as in the upper panel. The yellow patches highlight the quiescence periods, in which no tremor is observed. We find the episodicity of tremors similar to laboratory observations in Fig. 5.2.

cific event, we also observe intermittent occurrence of tremor, punctuated by pauses with no radiation interpreted as temporary stalling of tremor-initiating slow slip^{????}, and large variability in centroid locations of the tremors occurring in two-hour intervals. Aspects of these observations mirror our lab-based findings, albeit with greater complexity, and may suggest the potential reopening of sealed or healed fault valves (e.g., backward ruptures). Unfortunately, our lab experiment is unconfined and does not include healing/sealing mechanisms, therefore, we do not observe all these behaviors in our laboratory experiments. Moreover, the intermittent and bursty characteristics of natural tremors are noticeable when considering the radiated energy of the tremors. Comparing this with the AE energies observed in Fig. 5.2 reveals that the stick-break instabilities could be an effective proxy model for simulating tectonic tremor-like events; we further corroborate this proposition in the next section.

5.4.3 SCALING RELATION BETWEEN TECTONIC TREMOR SOURCE AREA AND RADIATED SEISMIC ENERGY

Since the tremors usually present a low signal-noise ratio and lack resolvable low-frequency signals, the seismic moment (i.e., the total size and slip) is difficult to measure. Radiated seismic energy has been utilized as an alternative measure of size to define the tremor magnitude^{??}. Our scaling relation between radiated energy and fracture area provides the first experimental evidence supporting use of radiated seismic energy as a measure of tremor source area or size. We use an enhanced catalog of[?] with 1056 tremor swarms in the Cascadia region from 2017 to 2023, each assumed to be driven by a slow slip event. We utilize the 500 swarms with the greatest number of tremors for measuring their tremor areas and energies. We measured the tremor area and energy by employing a grid-based approach that discretizes the entire Cascadia transition zone into many grid cells with a grid size of 7.5x7.5 km². For each swarm (slow slip event), in a similar manner to summing pixels illuminated after fracturing in the laboratory, we sum up the areas of grid cells that include tremors and the corresponding tremor energies within 2-hour intervals over the whole course of the event. In Fig. 5.6, we display the relations between cumulative tremor area and energy for all 500 swarms. Similar results can be found when using data from[?] serves as an independent verification (Fig. D.14). We notice that the northern tremor swarms are systematically louder than the southern ones, which may be controlled by the geographic variations in fault strength.

The area-energy relation demonstrates that the daily radiated seismic energy varies quasi-linearly with the tremor source area, as shown in Fig. 5.6. The seismic signals recorded during these swarms show very similar behavior to the AE signals recorded during the stick-break instability experiments. In addition, the observed spread of energy-area scaling coefficients may indicate the variations in apparent stress states on the subducted slab among events. The variability could be attributed further to heterogeneities in material properties and stress loading of the subducted slab^{??}. In the labora-

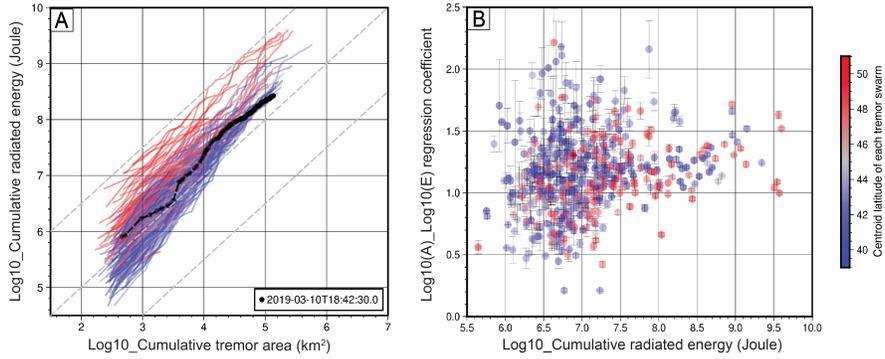


Figure 5.6: Field relationship between cumulative tremor area and cumulative radiated energy. The relationship between cumulative tremor area and cumulative radiated energy for 500 inferred tremor swarms that took place in the Cascadia region from 2017 to 2023 (adapted from ²). The color of the circles indicates the centroid latitude of each swarm, with northern to southern events shaded from red to blue, and highlighting the more energetic tremors in the north ². (a) The black line highlights the example event in Fig. 5.5a. Swarms systematically radiate greater energies from north to south. (b) The slopes of the curves in (a), measured as the linear regression coefficients, are displayed versus the radiated energy. We find a median regression coefficient, n , as the power for $E A^n$ to be 1.16 for the 500 swarms studied, which lies between the laboratory values of $n = 1.19$ and $n = 1.07$ for the high- and low-viscosity experiments, respectively. The power exponent is more variable for the small swarm relative to larger ones. We refer to Fig. D.15 for more details.

tory experiment, we also observe a slight asymmetry, which shows preferred nucleation and propagation direction due to the heterogeneity of the loading from the fluid pressure on the fracture tip. Therefore, the pre-existing stress condition and heterogeneity of material properties accentuate the complexity of rupturing process in the field scenarios.

Since the range of individual event sizes was not large enough, we only focus on the relation between cumulative values for energy and area for the comparison. Regardless, relating radiated energy and fracture area may be interpreted in several ways. ?? proposed a Brownian model for shear events for tectonic tremor that predicts a random distribution of single moment rates functions and a proportionality, or near proportionality between tremor area and radiated energy. A second consideration might be to think about seismic moment from these tensile fractures. ? concluded that the seismic moment is proportional to fluid pressure and cubed source radius. Our finding that radiated energy scales nearly, or quasi-linearly with area suggests that the event slip is nearly constant

with tremor area. Slip-opening invariance is quite plausible given the aspect ratios of veins². The detailed analysis on the energy-area relationship is referred to supplementary Text 1.

5.5 CONCLUSIONS

The intermittency and migration of tectonic tremors are frequently observed in subduction zones^{2,22,23}. Tremors are often hypothesized as occurring passively as shear events initiated by slow slip, from seismogeodetic observations^{2,22,23}. However, recent geological investigations^{2,22,23} and numerical simulations²² reveal that fluid propagation can potentially generate seismic behaviors similar to tectonic tremor via a combined mechanism of shear and tensile deformations. While we focus only on the tensile mechanisms due to our experimental limitations, our results suggest previous studies of tremor intermittency that attribute it exclusively to sporadic fault locking and stalled tremor-initiating slow shear slip^{2,22,23} should also consider a role for variable pore pressures that may cause tensile fracturing and sealing. Additionally, in the model we propose in Fig. 5.4, the cracking and sealing of fractures actually modify the potential for shear slip and thus tremor cannot be considered strictly as a result of the overall driving slow slip event, but instead as a facilitator of shear slip.

We perform a laboratory experiment in ambient pressure and temperature conditions of a pure slow tensile hydrofracturing event driven by a radially spreading fluid pressure front, accompanied by fast fracture propagating transversely to the front. Unprecedented high spatial and temporal resolution of the fracture dynamics reveal a stick-break instability², whereby AEs from the fast break events behave like seismic tectonic tremor. The fluid decompression controls the slow radial fracture propagation pattern (Fig. D.13). The observations of decelerated radial propagations, which are due to decreases in transient fluid pressure, explain that the migration of tremors slows down, as observed in Fig. 5.5b. This temporal evolution has been observed in numerous other studies^{2,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100}, describing it as diffusive and thus in some cases attributing it to pore pres-

sure migration. The laboratory stick-break stabilities also are similar to the intermittencies of tremor rupture observed in Cascadia. Our study further suggests a near-linear scaling of cumulative energy with active breaking area for both natural observation of tectonic tremor in Cascadia and the laboratory experiments. The idealized geometry of our experiment likely only represents elements of a rather complex system of veins and existing cracks.

The rate of fast transverse fracture and size vary with fluid viscosity. While the overall behavior of the slower rupture is not affected much by fluid viscosity, we observe pauses and breaks with various durations during the slow radial fracture propagation (Figs. 5.2 and D.4). We observe that in the high-viscosity fluid experiments, the slow radial fracture propagates with large break amplitudes followed by long pauses. In contrast, in the low-viscosity fluid, the experiment proceeds with small break amplitudes and short pauses. The tremor intermittency might be an indication of fluid viscosity, as in our experiment. Despite these differences, the overall propagation of the slow radial fracture (and fluid diffusion) occurs at a similar rate and coincides with fluid propagation. Simple models accounting for the different fracture styles and fluid viscosity and with an intact, simple internal stress structure showed that the fluid viscosity impacted the stress loading conditions (i.e., the lag between the fluid and the fracture front;[?]. In nature, the viscosity of geofluids may vary tremendously[?], as well as pre-existing permeability[?]. These factors may control the migration of fluid pressure and therefore tremor and slow slip propagation patterns, which deserves further exploration.

Our experimental results suggest a plausible model whereby fluid propagation generates seismic events with characteristics that are comparable to field observations. The transverse breaks propagate at much higher speeds (≥ 180 m/s) than the overall propagation speed (2-4 m/s), and are accompanied by fast fractures that propagate transversely to the radial rupture front and radiate AEs. This highlights the importance of imaging with a high temporal resolution to observe the fracture dynamics in detail. ? recently observed short-time tremor bursts that may migrate with a faster speed (3-25 m/s) than long-time tremor events. We suspect that even shorter bursts that comprise those

individual bursts might reach higher velocities, but that temporal resolution limits the observations. Furthermore, local nucleation observed at some locations along the fracture front, followed by transverse expansion of the fracture, suggests that the transient fluid pressure may rapidly vary differential stress and cause complex deformations^{??}.

To conclude, our experimental approach permits the measurement of a fracturing transparent sample in a nearly circular fracture geometry with unprecedented high spatiotemporal resolution, providing novel insights into the generation and characterization of tectonic tremors and slow slip. Further work should include differential stress conditions to evaluate the contribution of shear stress on the propagation of these fractures. Alternative models to tremor generations remain at depth. One of them describes slow earthquakes and tremor as deformation of shear zones within granular media, where particle jamming from compression and expansion builds force chains^{??} that may explain magnitude-frequency distributions of LFEs. While our study suggests a role for tensile cracking (i.e., hydrofracture) in tremor genesis, it is rather likely that the reality is a combination of various mechanisms. On rare occasions have, slow slip earthquakes not been accompanied by tremor[?], which suggests that hydrofracture mechanisms for tectonic tremor are not required for slow slip and additional physics ought to be understood.

6

Compaction-enhanced Instability in Laboratory Fault-valve Media

SUMMARY

The underlying mechanism of fluid-solid interactions remains elusive in the fault-valve systems. Here, we design a new experimental apparatus that enable us to both visualize and listen to the fluid transport and its interplay with the solids within artificial fault-valve media. This setup, featuring regions of alternating permeability, allows us to explore the complex behaviors that emerge under a steady fluid injection flow rate. Our research seeks to unravel two primary questions: 1) how does fluid flow interact with porous media and barriers? 2) how to reveal those interactions from acoustic and pressure measurements? Through our investigations, we discover that compaction and dilatancy play pivotal roles in controlling fluid movement and the integrity of barriers. Furthermore, we find that acoustic and pressure data can, under certain conditions, provide insights into these physical phenomena. Our direct observations offer a deeper understanding of fluid dynamics in fault zones and their potential to trigger seismic activities.

6.1 INTRODUCTION

Fault valves, ranging from minor veins to significant megathrusts, exhibit complex mechanical characteristics, comprising both permeable pathways and impermeable barriers with very low permeability. Overpressurized fluids, resulting from processes such as fluid injection^{??} or mineral dehydration^{??}, tend to navigate through these fault channels, driven by pressure diffusion. The advancement of fluid pressure is intricately linked with its interactions with the surrounding media and the strength of the fault^{??}, influencing fault stability and potentially triggering earthquakes^{??}. Additionally, these interactions can lead to the reactivation of fluid pathways, resulting in fluid leakage^{??}.

Fault dilatancy and compaction are pivotal mechanisms influencing the weakening or strengthening of fault zones due to fluctuations in fluid pressure, a phenomenon well-documented through

laboratory experiments^{???}, numerical simulations^{???}, and field observations^{???}. The interplay between compaction, dilatancy, and fault movement is intricately linked to fluid pore pressure, which is often utilized to deduce the states of dilatancy and friction. In addition, the viscous flow of the matrix and fluid-mediated mass transfer significantly contribute to both compaction and dilatancy, particularly in deeper regions subject to high temperatures^{???}. Despite extensive research, the intricate dynamics between fluid pore pressure and the processes of compaction and dilatancy have not been fully captured in laboratory settings, highlighting the need for further exploration of these complex hydromechanical processes.

Geological studies[?] have revealed that fault zones are composed of fragmented rocks, encompassing pores, cracks, and joints, with permeability that varies with depth. This raises the question of how the geological properties of fault gouge affect the propagation of fluid pore pressure. Understanding this dynamic is crucial for grasping how fluid pore pressure spreads and potentially initiates earthquake ruptures. In this study, we introduce a new experimental setup utilizing artificial porous media coupled with high-frequency measurement tools, such as cameras and acoustic transducers, allowing us to observe, for the first time, the intricate hydromechanical interactions within a simulated fault-valve system.

6.2 METHODOLOGY & MEASUREMENTS

APPARATUS AND SAMPLE

As depicted in Fig. 6.1, we design a new experimental setup using 3D-printed materials and highly sensitive detection methods, including a high-speed camera and a high-frequency acoustic transducer. Our experiment features a sample with five chambers separated by four barriers, each 1.5 mm thick, designed to simulate the impermeable or low-permeability barriers found in fault valve systems. These barriers serve as a sort of 'toggle switch', as described by[?], where their permeability is

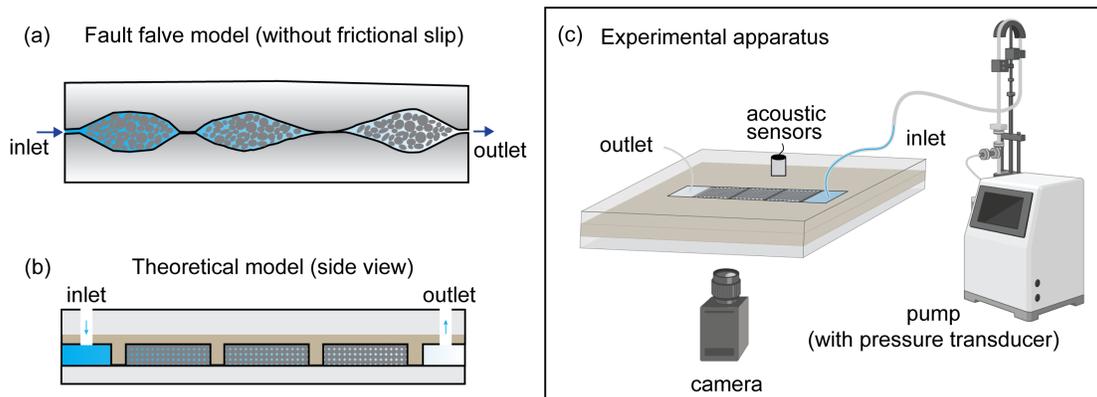


Figure 6.1: Diagram of experimental apparatus and measurements. (a) Schematic fault valve. (b) Simplified schematic laboratory specimen. (c) Experimental apparatus design.

negligible until they fail, at which point it becomes significantly high. To simulate permeable zones, the central three chambers are filled with glass beads. Alternatively, samples can be created with porous media integrated directly into the chambers. While our experiments utilize both types of samples, we primarily focus on those filled with glass beads here. For the experiments, the bead-filled sample is placed between two sturdy, clear blocks and then clamped to compress the beads (without bonds) and prevent fluid leakage during fluid injection.

INSTRUMENTAL MEASUREMENT

In Fig.6.1c, fluid injection is facilitated by a pump that introduces fluid through an inlet on one side and allows it to exit from an outlet on the opposite side. The entire procedure is captured from beneath by a high-speed camera, while a high-frequency acoustic transducer positioned adjacent to the chamber monitors the process. The high-speed camera is activated by the acoustic emissions generated as the first porous chamber compacts under increasing fluid pressure. With the recording at a rate of 5,000 frames per second, the camera provides detailed visual data, whereas the acoustic transducer captures continuous audio signals for up to 20 seconds at once. Concurrently, the pump records the injection pressure and volume, with data points collected every 0.3 seconds, ensuring a

comprehensive understanding of the fluid dynamics during the experiment.

6.3 RESULTS

6.3.1 INJECTION PRESSURE AND HYDROMECHANICAL VARIATIONS

As depicted in Fig. 6.2b, we introduce water into the chamber at a steady rate of 2 ml/min. The pump logs both the injection pressure and the volume of water, with these measurements presented in Fig. 6.2a. By integrating kymographs, video recordings, and pressure data, we categorize the entire hydromechanical process into distinct stages.

In the initial stage, the fluid pressure change exhibits a sequence of accelerated increase, slowing increase, stabilization, and then accelerated decrease. As water enters the first chamber, there is an exponential rise in fluid pressure, leading to the compaction of the second chamber by compressing the first barrier, notably around 80 seconds. This compaction is evident from the increased pixel intensities in Fig.6.2b. The rise in pressure then slows until the first barrier fails, followed by a rapid decrease in pressure until a higher pressure is necessary for the fluid to enter the next porous chamber.

The second stage presents a different fluid pressure change pattern: a nearly linear increase, followed by a slow decrease, a brief transitional phase, and then a rapid decrease. This stage highlights fluid drainage under injection pressure. Before complete saturation of the porous chamber, backward saturation begins around 100 seconds, causing a slight drop in pressure. As saturation progresses, compaction induced by fluid pressure accelerates this process, visible in both kymograph and video footage. The video reveals notably more compaction around the 65 mm mark. This compaction leads to a surge in fluid pressure, followed by significant dilatancy observed as compaction at both ends with dilation in the center. This central dilation reduces pressure by increasing porosity and drawing in water. The high-pressure wave moving forwards leads to further compaction and

eventually breaches the second barrier, while the wave moving backwards reduces pressure at the injection site, demonstrating how compaction and dilatancy can drive diffusive pressure surges along the fault². Similar to the first stage, barrier failure results in a drop in fluid pressure.

The third stage mirrors the second in its pattern of fluid pressure changes: a steady increase, then a slowing decrease, a prolonged transitional phase, and finally an accelerated decrease. The initial increase is less pronounced than in the first stage, owing to weaker compaction in this chamber and the reduced pressure required for fluid drainage. Approximately 40% of the chamber shows clear compaction. Post-drainage and saturation, dilation at the front of the compaction zone occurs due to a surge in fluid pressure, with simultaneous forward and backward compaction as seen in the previous stage. As dilatancy weakens, the decrease in observed fluid pressure slows. The failure of the third barrier, taking longer than in the second stage, results in an extended transitional phase before another decrease in observed fluid pressure.

In the fourth stage, the pattern of fluid pressure change simplifies to a brief increase followed by a nearly constant fluid pressure level. This initial rise is akin to that seen in earlier stages. However, due to the loose nature of the glass beads, neither dilation nor compaction is observed following the increase in fluid pressure. To sustain fluid flow under these undrained conditions, a consistent injection pressure is maintained.

6.3.2 ACOUSTIC MONITORING

Using a similar experimental setup, we delve into the temporal characteristics of acoustic signals throughout the hydromechanical process at varying fluid injection rates. In Fig.6.3, we present the acoustic data and their corresponding envelopes for injection rates ranging from 2 ml/min to 4 ml/min. Video analysis reveals that the acoustic emissions primarily result from slips between glass beads. Similar to previous observations, processes such as pressure-driven compaction, dilatancy, and barrier failure occur, each leaving distinct acoustic signatures, as highlighted by the cumulative

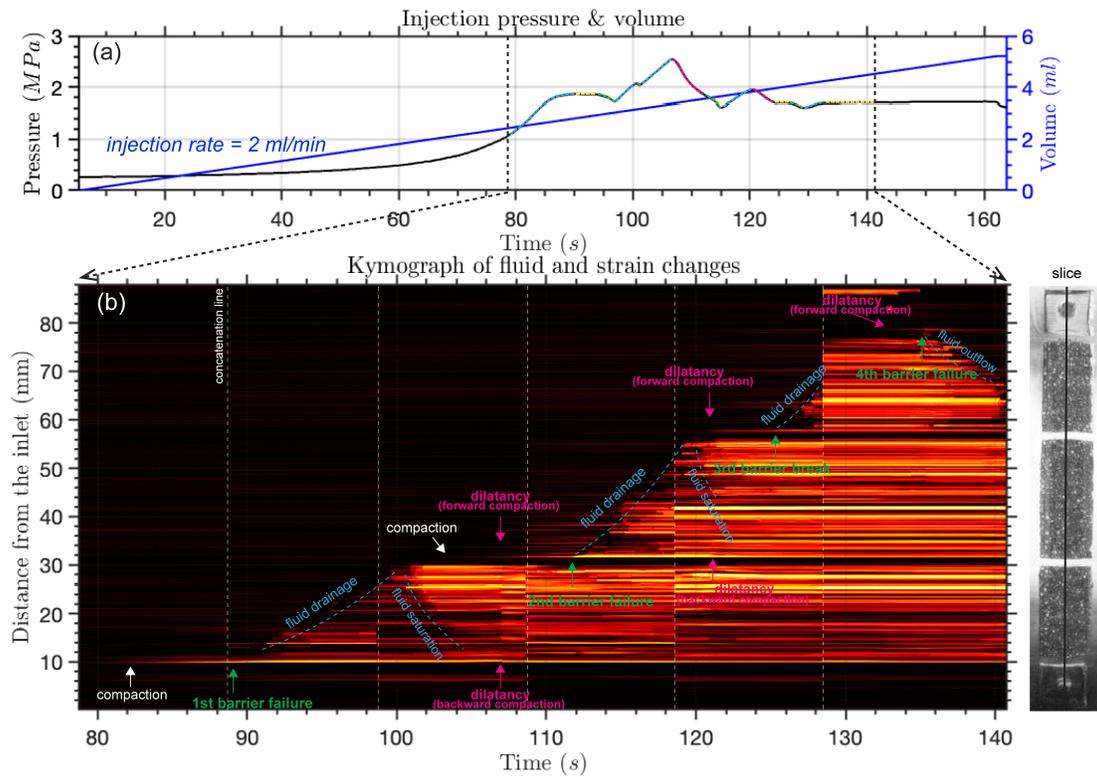


Figure 6.2: (a) Recorded injection pressure and volume history at the injection point, with an injection rate of 2 ml/min. Attention is drawn to the highlighted segments. The varied colors represent distinct fluid pressure (P_f) fluctuations and the corresponding physical phenomena, elaborated in the subsequent image. Three repeated patterns are observed, each marked by notable increases and decreases in P_f . (b) Kymograph detailing pixel variations correlated with movements of fluid and glass beads. This visualization is created by extracting a central vertical slice from video frames, after subtracting a background image, which is displayed on the right for reference. Differentiating the intertwined effects of fluid dynamics and bead displacement on light intensity and pixel distribution is complex. However, through manual review of the animation, we have identified and labeled the specific physical processes causing changes in pixel intensity and position. The associated video is accessible [here](#). Note that due to the video's large size, it is segmented, analyzed in parts, and then reassembled. The background image used for reference is rotated by 90 degrees.

envelopes (green lines). Notably, the dilatancy phase, characterized by numerous slips and acoustic activity, often culminates in barrier rupture, which is marked by a notable increase in the cumulative envelope.

The overall duration of these processes is influenced by the injection rate or the rate of pressure migration. Higher injection rates lead to more significant barrier failures but afford little time for precursor signals to emerge. In the raw acoustic data, pinpointing the exact moment of barrier failure can be challenging due to subsequent acoustic emissions. Moreover, variations in fluid pore pressure contribute to isolated, relatively large slips and pronounced acoustic signals. Despite these complexities, a thorough analysis of the complete acoustic dataset allows for the identification and understanding of the primary hydromechanical processes at play.

6.4 DISCUSSIONS

We have no shear stress placed on the sample. The frictional slip along the interface is not explicitly involved. The mechanical deformations observed are solely driven by variations in fluid pore pressure. These deformations encompass the stretching and breaking of barriers between granular chambers as well as slips among glass beads and their subsequent movements. Consequently, porosity and permeability emerge as crucial factors influencing fluid pressure dynamics.

Although the porosity can be changed in a variety of ways, such as fracture, slips, granular flow, viscous flow of the matrix and chemical dissolution-precipitation processes, we shed some lights into the fracture and slips influenced by the compaction and dilatancy during mass transfer as observed in our experiments.

Fluid pressure migration and stress changes due to rupture propagation predominantly determine slip behavior along faults. Elevated fluid pore pressures facilitate aseismic slips, which are strengthened by dilatancy[?]. Such aseismic slips can redistribute stress, potentially inducing earth-

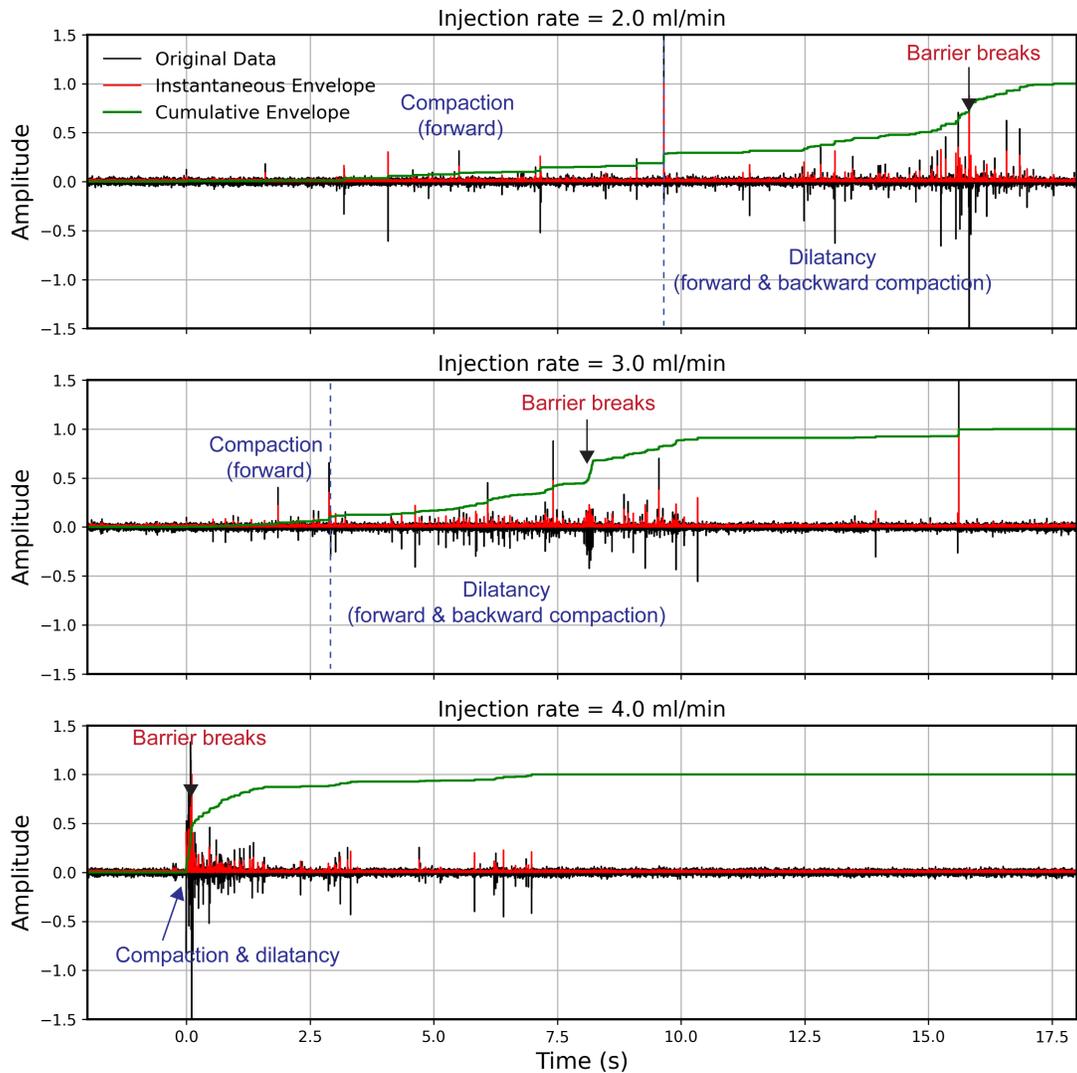


Figure 6.3: During the hydromechanical process, we captured continuous acoustic data (depicted in black) at various injection rates, which are displayed across three panels: 2 ml/min (upper), 3 ml/min (middle), and 4 ml/min (lower). Superimposed on these recordings are the instantaneous envelopes (in red) and the cumulative envelopes (in green). Arrows are used to indicate the occurrence of different physical processes.

quakes[?] and slow earthquakes[?]. Moreover, slips can reciprocally affect porosity and permeability, thereby influencing fluid pore pressure[?]. This interplay indicates that fault slip phenomena are intricately interconnected with the processes of dilatancy and compaction within the fault zone.

Our findings illustrate that barriers or zones with a contrast in permeability compared to their surroundings can induce compaction and dilatancy. Such variability in porosity and permeability is likely a common feature within fault zones.

Firstly, the rate of compaction is influenced by the fluid pressure (P_f), which in turn can impact the rate of aseismic slip, essentially affecting the strengthening rate of the fault.

Secondly, the combined effects of compaction and fluid injection can lead to the formation of high P_f anomalies, potentially contributing to the initiation of rupture. The compaction resulting from aseismic slips accelerates this process, potentially transitioning to seismic slip.

Thirdly, an increase in P_f is followed by dilatancy, characterized by forward compaction and backward compaction. Forward compaction near a barrier can amplify P_f in that area, eventually leading to its failure. If seismic slip occurs concurrently with compaction, the dilatancy effect may be diminished as P_f is expended in facilitating the seismic slip. Conversely, if seismic slip precedes dilatancy, the ensuing increase in porosity can amplify the dilatancy effect. Additionally, dilatancy on its own can slow down seismic slip or reduce the rate of fault weakening.

6.5 CONCLUSION

In our newly developed experimental setup, we employ both visualization and acoustic monitoring to observe hydromechanical behaviors under consistent fluid injection rates. Our findings distinctly highlight phases of compaction and dilatancy preceding barrier failure. Particularly, we observe that surges in fluid pressure, stemming from compaction, manifest not immediately around the barrier but rather at the leading edge of the compaction zone, often within the chamber's central region.

The breakdown of the barrier is precipitated by forward compaction coupled with pressure migration occurring during dilatancy phases. An intriguing observation is that backward compaction leads to a marked reduction in pressure at the point of fluid injection, suggesting that fluid backflow might be more common in natural settings than previously assumed. Moreover, the pattern of pressure decrease offers a means to differentiate between dilatancy, characterized by a slowing decrease, and barrier failure, marked by a rapid drop in pressure. Additionally, the accumulation of acoustic energy reflects ongoing deformation and highlights the significance of the spatiotemporal distribution of laboratory-induced or natural earthquakes in understanding underlying physical processes. These laboratory insights are invaluable for enhancing our comprehension and surveillance of fault-zone hydrodynamics and seismic activity.

Yuan C, Ni Y, Lin Y, Denolle M. Better Together: Ensemble Learning for Earthquake Detection and Phase Picking. IEEE Transactions on Geoscience and Remote Sensing. 2023 Sep 28. 10.1109/TGRS.2023.3320148

7

Better Together: Ensemble Learning for Earthquake Detection and Phase Picking

SUMMARY

The detection and picking of seismic waves is the first step toward earthquake catalog building, earthquake monitoring, and seismic hazard management. Recent advances in deep learning have leveraged the amount of labeled seismic data to improve the capability of detecting and picking earthquake signals. While these deep learning methods have shown great promise, their success remains hindered by low generalizability and poor performance in low signal-to-noise ratios (SNRs) data. Here, we propose a new processing workflow that integrates pretrained deep learning models, multi-frequency band predictions, and ensemble estimations to enhance the generalization of these algorithms. We test the performance of the ensemble model using three benchmark datasets, one of which is within-domain and has been used for training the deep learning models, the other two being cross-domain test datasets. We explore the performance given data and model characteristics. We also compare an ensemble approach with a transfer-learning approach and discuss the benefits and drawbacks of these two approaches when deploying on continuous data. Our experiments demonstrate that ensemble learning can drastically improve generalization ability and hence alleviate the need for transfer learning in the case where no labeled datasets exist.

7.1 INTRODUCTION

EARTHQUAKE detection and seismic phase picking are the most fundamental steps for cataloging, monitoring, and alerting earthquake hazards. The detection of earthquakes is often done automatically at seismic networks using features of earthquake waveform as filters (e.g., impulsivity). The location of earthquakes requires the knowledge of seismic wave travel time between potential sources and receivers at seismic stations, especially the arrival time of P and S waves. Algorithms exist for automated picking, though seismic analysts tend to manually pick the arrival phases to improve

the quality, particularly in noisy waveforms. From these arrival times, algorithms are designed to rapidly locate an earthquake, associate its magnitude, and broadcast public warnings if the shaking is expected to damage human infrastructure[?].

Seismology is becoming a big-data field as the volume of seismic data acquired is vastly expanding[?], and as sensing technology is evolving[?]. Recently, the design and implementation of convolutional neural networks in deep learning (DL)[?] for seismic data has dramatically transformed seismological research. In particular, the contribution of DL in seismology has been in large part for feature extraction in seismograms to automate the characterization of earthquake sources from continuous seismic data^{?????}.

A particular success in DL for seismological research is the picking of seismic wave arrival times, a task referred to as “phase picking”. Phase picking is among the top uses of DL in seismology and is mainly designed as a supervised learning problem that requires large labeled datasets to train from scratch. The typical workflow is to develop a new architecture of a deep neural network associated and trained with a specific dataset (either a “standard benchmark” dataset or a user-specific labeled dataset).

There exist today several popular phase pickers, such as GPD[?], PhaseNet[?], EqTransformer[?], Siamese EqTransformer[?], DPP[?], and ARRU[?], which are well-used by the community. There also exist a number of benchmark datasets to train these models from, SCEDC[?], STEAD[?], Iquique[?], LENDB[?], NEIC[?], INSTANCE[?], DiTing[?], PNW[?], and DeepBlue[?]. Woollam et al.[?] designed a toolbox SeisBench for interoperable machine learning models and benchmark datasets and user-friendly access to these various models, which enables systematic investigations of the performances of model designs and their weights obtained after training on multiple datasets. New users of these models may refer to Münchmeyer et al.[?] to evaluate the appropriateness of specific pretrained models on the specific use case. Cross-domain performance varies a lot either from different benchmark datasets or different DL models[?], suggesting that generalization capability is rather limited and

model design may impact the performance of the models on a given dataset. Although the DL methods for detection and picking have demonstrated their high performance in specific applications, they still face the challenge of poor generalizability, high noise sensitivity, and overproducing false results. Researchers have mitigated these limitations with elaborate workflows of multiple-sequential DL models^{2, 3, 4}. Recently, Park et al.⁵ documented the varied prediction consistency of the DL models and provided a strategy to mitigate some of their degraded performance that aggregate models applied on a small stride running window.

Transfer learning, the step of continuing the training of a machine-learning model on a new labeled data, is needed to fine-tune the model weights to improve their performance (i.e., picking rate and accuracy) to new data^{6, 7, 8}. Transfer learning typically needs a labeled dataset, making it impractical to apply to datasets without prior knowledge about phase arrival times. However, as with any training of DL models, limitations in training performance remain due to the quality and diversity of the training data. In those cases, the transfer-learned model performance may be thus insignificantly improved for cases without a large and diverse dataset.

The difficulty remains in selecting a specific DL model for target study data. For supervised learning algorithms, the literature only reflects the development of DL models specific to particular datasets. The DL models may vary their prediction performance with the training on different datasets. One can achieve better prediction performance by leveraging and integrating the performance of different base models^{9, 10}. In seismology, random forest-based ensemble learning has been successfully utilized for laboratory earthquake prediction¹¹ and earthquake wave discrimination¹². There remain opportunities to bring ensemble learning to DL frameworks.

This study proposes to leverage the diversity of the training datasets to improve the DL performance and the generalization for earthquake detection and phase picking. We augment bagging (one aggregation method,¹³) by making predictions over multiple narrow-bandpass versions of the seismic data. We explore various ensemble techniques, such as maximum, semblance, and meta-

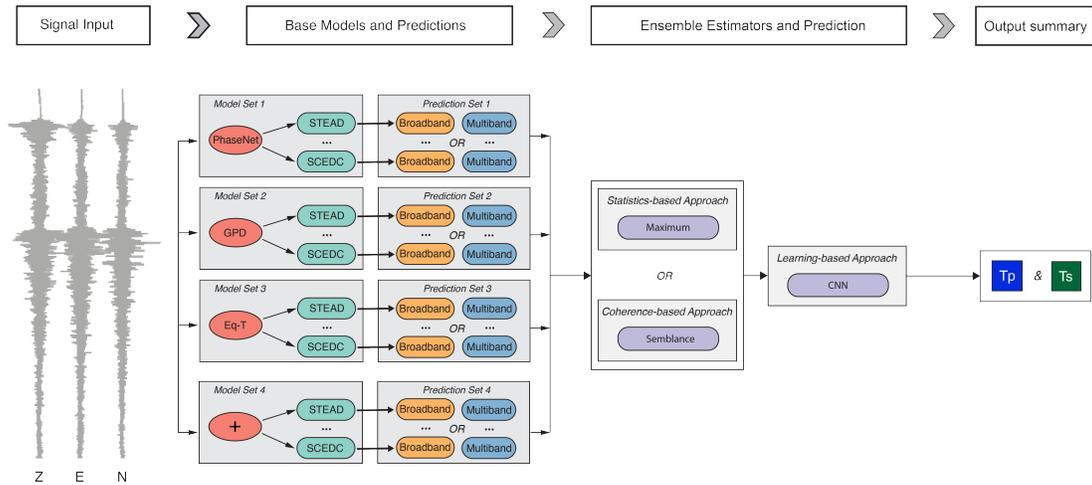


Figure 7.1: An ensemble estimation-based framework for earthquake detection and phase picking. The main components include base predictions at broadband or multiple frequency bands (e.g., filtered data) and ensemble estimation by either statistics-, coherence-, or a learning-based approach. Note that only EqT-based pretrained models are tested.

learner, to compare the performance of phase picking. We test on within-domain datasets and cross-domain data with a comparison with a transfer-learned model, which tests the model’s performance. Finally, we deploy the phase picker on continuous data and find a much-reduced rate of spurious detections.

7.2 METHODOLOGY

We develop a framework comprising two essential workflows that we illustrate in Fig.7.1: 1) multiple individual model predictions and 2) aggregation (ensembling). We refer to this workflow as Ensemble Learning Earthquake Prediction (ELEP).

The first component utilizes DL-based models to predict the input three-component seismograms trained on multiple datasets. The framework may include multiple model architectures, each of which can be trained on different datasets for pretrained models with different neural weights. Here, we call models that are DL models with either different architectures or pretrained weights as

“base models”. We apply individual base models either on the raw data (broadband) or on filtered data (multiband) to generate multiple predictions on the same input seismograms. The predictions of interest are two time series of probabilities of the P- or S-phase in the input seismic data. The second component involves sequential ensemble estimations that aggregate all predictions to obtain the final arrival times of the P- and S-phases.

7.2.1 PART I: BROADBAND AND MULTIBAND PREDICTION WORKFLOWS

The base models typically solve classification problems: they take input data as three-component seismograms of length 6000 (i.e., 60 s at 100 Hz), and they output three channels, respectively, for detection, P-phase, and S-phase time series probabilities. These probability time series share the input size, i.e., identical time sample length.

The base models have been trained with datasets that have various spectral content. For example, the original PhaseNet was trained on raw data from southern California that the authors compiled², while the EqTransformer (EqT) we will use was retrained by² using 1-45 Hz filtered data from STEAD that compiles local and regional data from around the world^{2, 3}. The noise present in the original dataset influences the model performance. For instance, Yin et al.² found that using the benchmark dataset STEAD but adding seismic noise local to the station improved the ability of deep neural networks to characterize earthquake signals.

We propose two workflows to obtain predictions of those probabilities: a broadband (BB) and a multiband (MB) prediction, as shown in Fig.7.2a and b. The first one is straightforward and uses broadband (1-45 Hz) seismograms as the input to the base models. However, this approach is noise-sensitive, particularly for low-amplitude signals or low-magnitude earthquakes². To mitigate this, we develop the second prediction workflow based on multiple frequency bands, summarized in Fig.7.2b. We add two steps for the MB prediction compared to the BB one. We first decompose the signal into multiple frequency bands, as illustrated in Fig.7.2d for the vertical component of the

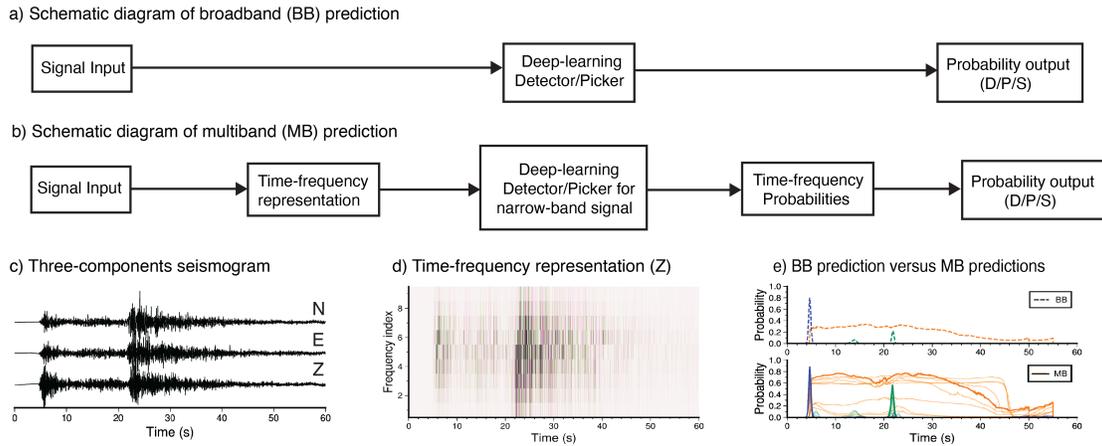


Figure 7.2: Example of ELEP workflow. (a) Broadband prediction scheme. (b) Multiband prediction scheme. (c) Example three-components seismogram. (d) Time-frequency signals decomposed by a recursive multiband filter. (e) The predictions of these filtered signals are obtained and compared between broadband and multiband predictions, denoted as solid and dashed lines, respectively. Detection, P-phase, and S-phase picking are represented as orange, blue, and green lines, respectively.

seismograms. Time-frequency representations can be realized by either the short-time Fourier transform, a wavelet-based transform, or the S-transform²², but these transforms are computationally prohibitive and considerably limit the deployment of these workflows on continuous data. We use the recursive multiband filter (MBF) performed in the time domain proposed by Poiata et al.²³, a filter adapted from Lomax et al.²⁴ as an efficient alternative for picking seismic phases in real-time earthquake monitoring. We filter the data into ten frequency bands to balance computational efficiency and prediction effectiveness in design choice. We illustrate an example of such decomposition in Fig.7.2d. Each filtered or narrow-band signal is put into the DL model for individual predictions. The DL model utilized in the BB prediction is directly adaptive to the MB prediction without designing new DL models. We thus perform ten BB-like predictions for each MB prediction. Each of the predictions from the original EqT model² has three-channel time series outputs: detection, P-phase picking, and S-phase picking probabilities. We show predicted results for all ten frequency bands as solid lines in Fig.7.2e (lower panel). In the MBF workflow, each forward model has ten

predictions, and we perform the first level of aggregation by selecting the given prediction with the highest probability (thick lines) in each of the three outputs (detection, P-pick, S-pick), which we call the “optimal output”.

Comparison between the BB and MB predictions are shown in Fig.7.2e. Even with the simple and obvious example shown in Fig.7.2c, the optimal MB predictions (thick lines) have much higher probabilities and retain the prediction accuracy of individual predictions as inferred from the temporal width of these time series of probability. The MBF provides more representations of the data, some of which may be more familiar to the trained DL base model. We conclude from this experiment that the MBF enhances the confidence of the DL model upon the predictions on multiple signal transformations.

7.2.2 PART II: ENSEMBLE ESTIMATIONS AND PREDICTIONS

This step aims to aggregate the prediction from multiple base models, treat them as an ensemble, and combine them for an improved prediction. In the following, we ignore the detection branch that some models may output (e.g., the EqT) and focus on P and S arrival-time picks.

We propose and develop three distinctive methods for ensemble estimation: statistical-, coherence-, and learning-based approaches. Each of these methods is independently capable of generating the final prediction. We illustrate the concept in Fig.7.3 using BB predictions. We gather the individual base predictions from eighteen pretrained models: PhaseNet, GPD, and EqT models, each trained on six different training data. Fig.7.3a and b illustrate the predictions for the P-phase and S-phase for a given example waveform. In this study, only six EqT-based pretrained models are utilized for computational efficiency.

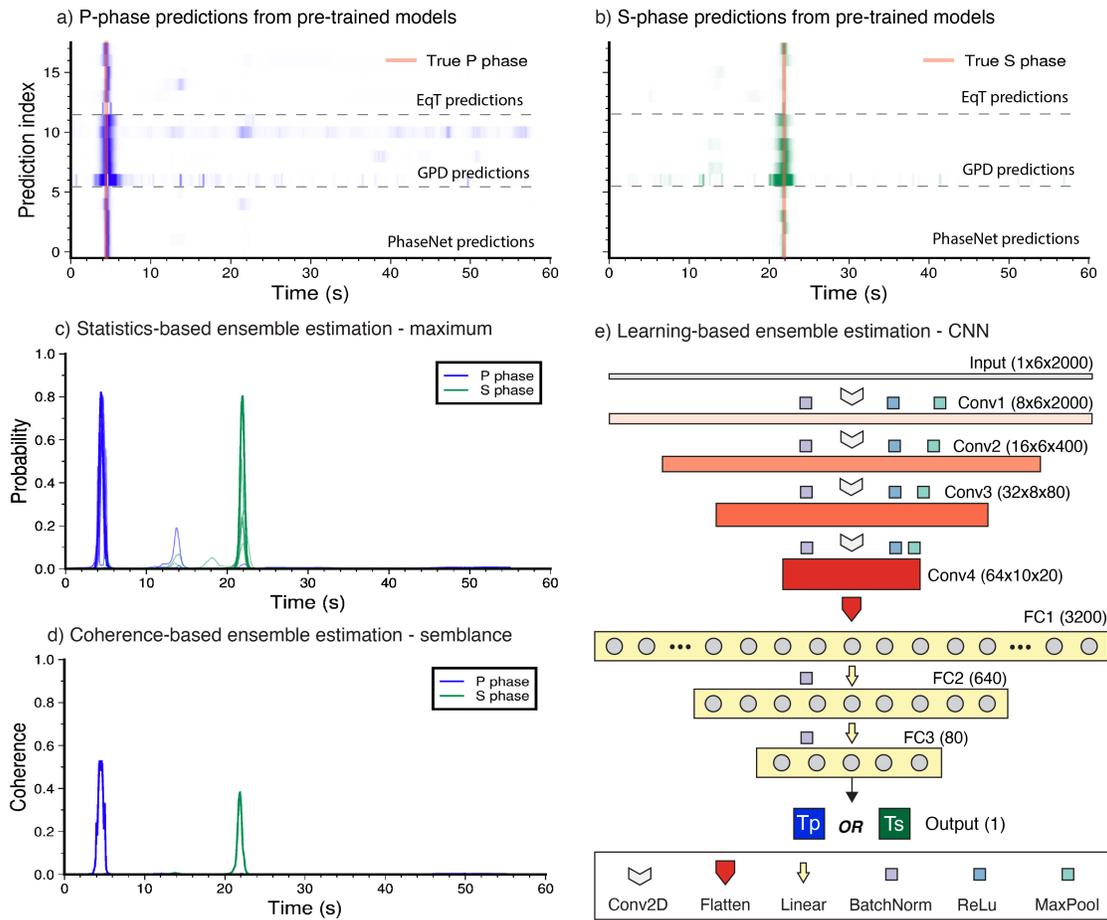


Figure 7.3: Example of ensemble estimation for three aggregation methods. (a) P-phase probabilities and (b) S-phase probability functions predicted by 18 pretrained models. Deep colors denote high probabilities. (c) Maximum-based ensemble estimation from six EqT predictions (thin lines) for P-phase and S-phase probability functions (thick lines). (d) Semblance-based ensemble estimation from six EqT predictions for P-phase and S-phase probability functions. (e) Learning-based ensemble estimation for P- and S-phase probability functions. Note that only six 20-seconds EqT-based predictions are used as the input here.

MAXIMUM-BASED ENSEMBLER

For the statistics-based ensemble estimation, we identify the prediction with the greatest probability by applying the maximum function. This process is independently conducted for both the P- and S-phases, as articulated in the following equation:

$$Pb^{max} = MAX(Pb^1, Pb^2, \dots, Pb^m), \quad (7.1)$$

where Pb is the predicted probability trace of P or S phase, m is the index of the DL model for probability prediction, and MAX denotes the maximum operator.

Fig.7.3c illustrates with an example the results of the maximum ensembling: the peaks of these predictions (thick lines) are the probabilities of the arrival times of the P- and S-phases.

SEMBLANCE-BASED ENSEMBLER

The second method is a coherence-based ensemble estimation method to predict the final probabilities, which applies semblance analysis to coherence estimation, a practice widely adopted due to its demonstrated robustness and computational efficiency⁷. The original semblance C^0 (for either P or S wave) can be determined with each model m probability Pb^m as follows:

$$C^0(t_i) = \frac{\sum_{t_i-\delta t/2}^{t_i+\delta t/2} (\sum_{m=1}^M Pb^m)^2}{N \sum_{t_i-\delta t/2}^{t_i+\delta t/2} (\sum_{m=1}^M (Pb^m)^2)}, \quad (7.2)$$

where t_i stands for the i^{th} time sample, δt is the length in seconds of the time window. This could be considered through the shape length of the phase probability label. Here, we choose an optimal time length of 0.5 s after multiple tests. M is the number of DL models used for predictions. The coherence of probabilities is measured at each t_i centered in the time window. This original semblance analysis effectively measures coherence across probability traces but may enhance noise or

artifacts. Therefore, we modify the original version to weight the coherence according to the following weights W :

$$W(t_i) = \text{MAX}(Pb^1(t_i), Pb^2(t_i), \dots, Pb^m(t_i)), \quad (7.3)$$

where the weight $W(t_i)$ is defined as the highest one among all probabilities at the time point t_i .

We finally write the semblance:

$$C(t_i) = W(t_i)C^0(t_i)^\nu. \quad (7.4)$$

where the exponent $\nu = 2$ is set to balance the suppression of noise and signal coherence. The semblance analysis is performed individually for the P- and S-phases. Fig.7.3d shows the final coherence traces, in which the peaks point to the phase arrivals accurately.

While both previously mentioned methods can be utilized to predict probability traces, each exhibits limitations. First, the ensemble *maximum* is only picking the maximum probability and, therefore, is sensitive to the bias of the highest probability. In experimenting with the method, we find that the method achieves a high picking rate but is also susceptible to larger picking errors. Second, the ensemble *semblance* method, while prioritizing probability consistency, often compromises on picking rate to output with low coherence. Finally, both methods still output time series that require an additional post-processing step of peak detection and thresholding to decide on a positive outcome. We, therefore, propose a third method that overcomes these limitations.

META-LEARNER ENSEMBLER

The base models are treated for solving classification problems: they output three channels with time series of probability where every timestamp is a class. For the EqT model, the second and third time series are outputs for the probability of seismic phase arrivals in the window. We skip the detection probability as we focus on phase picking. Users then decide on thresholds to pick phases, which

adds a hyperparameter to choose the appropriate threshold. By construction, the two previously mentioned ensemble methods also need a hyperparameter for such a final decision.

The seismic phase picking is intrinsically a regression problem. Therefore, there remain opportunities to adapt the base models and predict scalar values of the phase arrival times. Regarding the base models as weak learners, we add a meta-learner to transform these classification models into regression.

We propose such a meta-learner to be a convolutional neural network (CNN)-based learning approach, leveraging its robust capabilities in pattern recognition and semantic comprehension². We use PyTorch and its taxonomy for neural networks². Fig.7.3e illustrates the primary architecture of the CNN, comprising four convolutional neural network blocks and three fully connected neural networks (Linear). Each convolutional block incorporates a convolutional layer (Conv2D)², batch normalization (BatchNorm)², a rectified linear unit (ReLU)², and max pooling (MaxPool)². The input data is a 3-dimensional tensor, in which the dimensions sequentially refer to the size of batch prediction, six base probability predictions for each trace, and 20-second (2000 time samples) probability prediction truncated from previous base predictions. The dimension of each layer is described in Fig.7.3e. We create two independent models to treat respective P and S and avoid the model to learn the relative time between P and S (e.g., earthquake location). The output of the meta learner is a single scalar value of travel time, either T_p or T_s , representing the time since the trace starts.

In this concept, any combination of base models that provide the same output size will be possible. In practice, we found that loading many models in memory was a high computational cost and restricted our base models to the EqT trained on six different datasets.

7.2.3 MODEL TRAINING

The meta-learner is a regression problem. Therefore, we use a mean-square-error (MSE) loss function. We choose the Adam optimizer for weight updates, and a learning rate of 0.2². The model is

small, and the training is efficient since these probabilities are relatively simple compared to learning this directly from waveforms. Additionally, the CNN model has robust performance, as later seen in its weak dependence on data size. However, the ensemble learner struggles to eliminate the categorically spurious probabilities, rendering it prone to estimation errors on any given input. Despite its ability to yield reasonable picking results under low and inconsistent probabilities, the error rate is currently unfit to function independently. The erratic behavior of this simple learning leads us to add a constraint to the predictions.

Therefore, we integrate the ensemble learner with either ensemble maximum or ensemble semblance, as depicted in Fig.7.1. These preliminary ensemble estimations can discern signals from noise and provide reference phase picks. We truncate the output predictions for 20 seconds around the maximum or semblance ensemble picks and use these truncated probability time series as input to the meta-learner. The ensemble learner is subsequently deployed as the final prediction, with the aim of enhancing prediction accuracy. In essence, drawing parallels with the widespread bagging and stacking ensemble learning methods[?], we conclude that our framework is primarily driven by ensemble estimations.

7.3 PERFORMANCE OF ELEP FRAMEWORK

We propose to test the performance of the ELEP framework in three separate contexts. First, we take the approach of ML developers in seismology, who typically use benchmark datasets. Second, we take the approach of ML users in seismology with plenty of labels that might continue training (transfer learning) the base models to a specific region. Third, we take the approach of seismologists who want to deploy these models on waveform data in a completely different seismic environment on the sea floor.

We perform both broadband and multiband predictions from six pretrained models. These mod-

els are called eqt-original, eqt-ethz, eqt-instance, eqt-scedc, eqt-stead, and eqt-neic, named with the model architecture hyphenated with the training dataset. Further details of these pretrained models are referred to in the SeisBench documentation[?]. The base model predictions for each seismogram are the probabilities for detection, P-, and S-phase. A detection, P-, and S-phase picking are successful for their respective probabilities above threshold values of 0.1, 0.1, and 0.05. For maximum-based ensembler, we set the probability threshold values of 0.1 and 0.05 for P- and S-phase picks. For coherence-based ensembler, we set the corresponding probability thresholds to 0.05 and 0.05 as the coherence is compromised across predictions.

Following the EqT model training protocol[?], each seismogram is pre-processed by truncating the time series to 60 seconds and filtering the data between 1 and 45 Hz using a Butterworth filter of a 4th-degree pole. The processing uses the ObsPy python package[?]. Some data are three-channel seismograms. Some data may be single-channel components with two channels filled with zeros.

We use several performance metrics in the following tests. First, the detection and pick rates depend on the threshold values discussed below. More than one event can be processed and returned through probabilities that exceed thresholds. Second, the pick errors are found by measuring the arithmetic mean (bias) and standard deviation (variance, standard picking error) of the residuals between the phase pick predictions and their ground truth labels. Unlike other studies that ignore many outliers with residuals larger than around 0.5s[?], we use the residuals under 2s for calculating the arithmetic mean and 10s for calculating the standard deviation to keep at least 80% low-quality phases and 90% high-quality phases.

7.3.1 PERFORMANCE ANALYSIS ON THE INSTANCE DATASET

In this section, we take the approach of an ML developer for seismology. We test the performance of our workflow against benchmark datasets. We use INSTANCE as a test dataset. The INSTANCE seismological benchmark dataset contains over 1.3 million waveforms from local earthquakes in

Italy[?]. We arrange the seismograms into 15 groups based on their respective Signal-to-Noise Ratios (SNRs). We randomly select 15,000 seismograms from each SNR group to ensure an equitable comparison and optimized computation. Each waveform has “ground truth” labels for the P and S picks. Some of the INSTANCE data may also be contained in other benchmark datasets, such as STEAD and NEIC. We randomly select 3,000 waveforms out of each SNR group of the INSTANCE data to generate a training dataset for the ensemble learner. We leave 12,000 other waveforms for testing. This simulates the case of limited training data.

BROADBAND AND MULTIBAND PREDICTIONS

We first evaluate the performance of the multiband workflow using *within-domain* predictions with the eqt-instance model trained on the INSTANCE data. We apply the prediction workflows separately on the fifteen SNR groups and show the results in Fig.7.4. Because we make 15,000 predictions in each SNR group, we show the results in the form of arithmetic means (biases) and standard deviations (standard picking errors) over these distributions for broadband and multiband predictions in Fig.7.4a-c for the detection rate, the P, and the S pick errors. As expected for within-domain predictions, the eqt-instance performs well because it has recognized most of the test data. The broadband waveforms tested are likely the training data for eqt-instance, therefore yielding high accuracy in predictions.

We now discuss the differences between broadband and multiband predictions. For both sets of predictions, the performance increases with SNR levels. The performance is generally lower for broadband seismograms than for multiband predictions. The detection rates are lower by about 5% across all SNR groups. The P-wave picking rates and uncertainties are somewhat similar, with a much-improved performance with the SNR level. S waves are generally poorly picked, which is a general problem in seismology that yield much lower picking rates and larger uncertainties. Multiband predictions greatly improve the picking rate, while broadband predictions only go from 80%

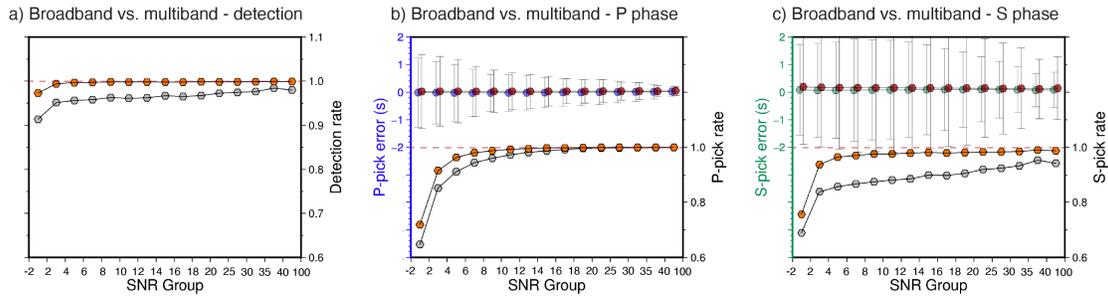


Figure 7.4: Within-domain broadband and multiband predictions on the INSTANCE test dataset using the eqt-instance model for (a) detection, (b) P-phase picks, and (c) S-phase picks. Detection and picking rates are gray (broadband predictions) and orange (multiband predictions) markers. P and S pick mean errors (e.g., bias) are shown in blue (P broadband), green (S broadband), and red (P and S multiband). The error bars are the standard deviations of picking residuals. Note that the shifts of pick-error markers in (b, c) are deliberately made to avoid overlapping and for clear visualization.

to 95% with the SNR.

However, the increased detection and picking rates of the multiband predictions bring greater uncertainties and biases. The broadband and multiband both have a bias of about 0.1s in their predictions. We attribute this to inherit from the base model performance issues, which were also apparent in ².

The next exercise tests the generalization of these models (e.g., ²) and the performance of the multiband variant. We now contrast the *within-domain* predictions with the *cross-domain* predictions. We gather all measures of performance (detection and picking rates and standard phase pick error) and show them in Fig.7.5. We ignore the mean picking errors (i.e., the biases) because they are small similar to the ones in Fig.7.4.

For the event detection results, the eqt-ethz and eqt-neic perform comparably well with the eqt-instance, which is reasonable because the former two pretrained datasets include data similar to INSTANCE. This is not strictly cross-domain predictions due to the data leakage. The other three base models (eqt-original, eqt-scedc, and eqt-stead) perform worse than the eqt-instance on waveforms with low SNR but perform comparatively with eqt-instance on the data with SNR over 10. This suggests that the generalized capability of pretrained models is reasonably well for high SNR

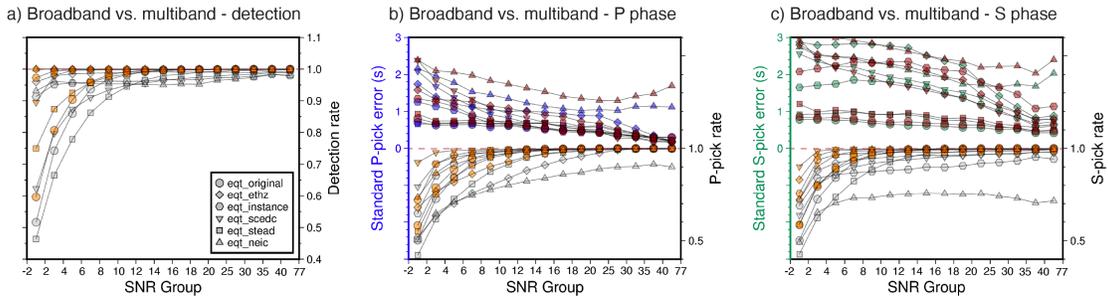


Figure 7.5: Cross-domain broadband and multiband predictions on the INSTANCE test dataset for six EqT-based pretrained models for (a) detection, (b) P-phase picking, and (c) S-phase picking. The pretrained models are denoted by different symbols. Gray symbols illustrate detection and pick rates from broadband predictions. The orange symbols display the corresponding detection and pick rates from multiband predictions. Blue and green symbols indicate standard picking errors of broadband P- and S-phase picking residuals, respectively. The red symbols represent standard multiband-predicted errors for comparison.

data but will fail at low SNR data. We attribute this to the noise of the data, which may be region-specific, but the seismic characteristics of earthquakes are global. The multiband predictions alleviate the noise effect from frequency bands and give rise to higher detection probabilities. Detection rates are thus increased compared to broadband predictions.

As previously shown for the within-domain prediction, the multiband prediction workflow also improves the picking rate, especially for S-wave picks (see Fig.7.5c and f). Similarly to our previous within-domain prediction, the additional picks have more errors. This is particularly true from the NEIC dataset, which includes local, regional, and teleseismic earthquake signals, and may not be well labeled². Interestingly, we find that the eqt-scedc performs even better than the eqt-instance in terms of pick rates, but it also has a larger variance in pick errors. We postulate that the reason behind the high picking rate of the eqt-scedc is that the scedc dataset includes much more low-SNR data for training, though we can only guess since DL models tend to be trained on augmented data². This suggests that the involvement of more well-labeled low-SNR data may improve the model performance. Overall, the multiband prediction performs better than the broadband prediction. S-phase picking performs much worse than the P-phase, which further states the accurate

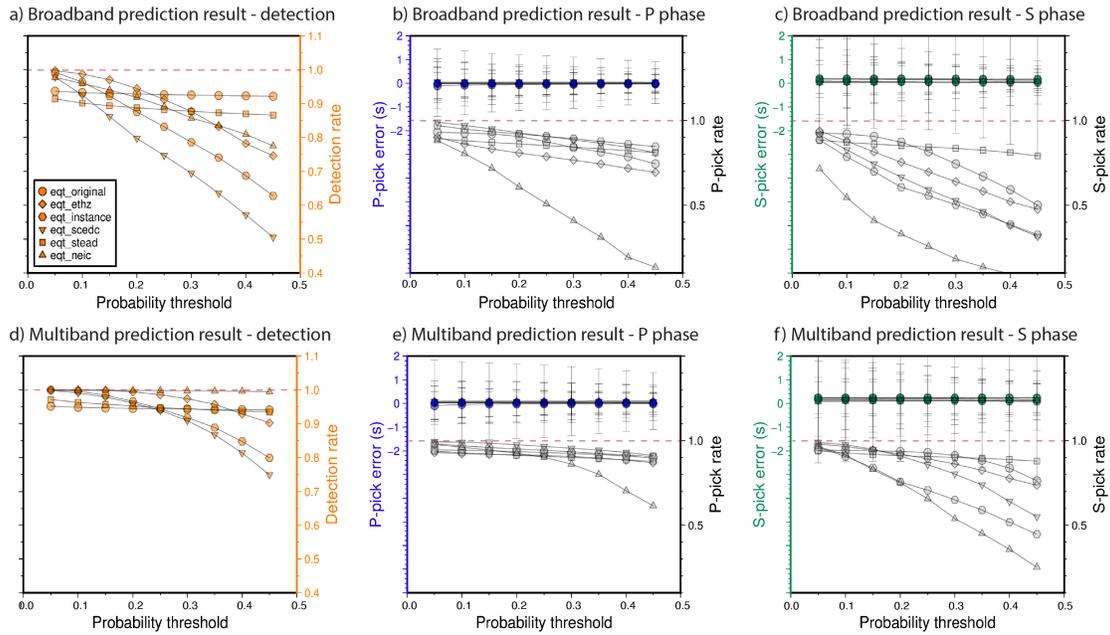


Figure 7.6: Effect of probability thresholds on (a-c) broadband and (d-f) multiband predictions on the INSTANCE. The symbols stand for different pretrained models. The orange and gray symbols represent the detection rate and P- or S-phase pick rate, respectively. The blue and green symbols denote the mean P- and S-phase pick errors in seconds, respectively. The error bars indicate the standard deviations associated with the mean pick errors.

picking of S-phases is still challenging.

The selection of probability thresholds influences the decision. Low threshold values imply higher detection and picking rates but lower quality (greater error) in the picking values. We show the variability of prediction performance in terms of detection rate, P-phase pick rate and error, and S-phase pick rate and error when varying probability thresholds from 0 to 0.5 at the interval of 0.05. We show the results in Fig.7.6. Detection and pick rates decrease with increased thresholds but at different rates with different models.

Broadband predictions are much more sensitive to thresholds (Fig.7.6a-c) than multiband predictions (Fig.7.6d-f). This means that choosing a multiband workflow is less sensitive to hyperparameter tuning. Naturally, the multiband predictions have much-improved performance because the workflow increases the predicted probabilities, and more predictions meet the threshold. We con-

clude that the multiband workflow increases the predicted probability, therefore, the detection and pick rates. However, it also gives more errors in the picks of phase arrival times.

PERFORMANCE OF ENSEMBLE MODELS

With the predicted results from all pretrained models, we use three ensemble estimation methods to obtain the final picks: maximum, semblance, and learnable. The estimation methods have been described in section II of this chapter. We now discuss the ensemble performance with examples that include the picking predictions from six pretrained models and three ensemble estimations from high-SNR to low-SNR data in Fig.7.7. The first two models output a time series of probabilities with the same dimension as the base models. The meta-learner transforms the multi-class classifiers (base models) into regression and is only shown as tick marks on the seismic waveforms in Fig.7.7.

The examples shown in Fig.7.7 have relatively low SNR values. S waves are difficult to pick, yet the DL models predict a reasonable value around the ground truth label. The examples illustrated here highlight the variability among the model predictions for the same waveforms and how the maximum and semblance models predict.

We first compare the two first ensemble learning methods. The ensemble maximum mainly depends on the probabilities with the highest values, regardless of their accuracy. The multiband workflow is effectively a maximum-ensemble method, and we have demonstrated that it increased the detection and pick rate but also increased the errors. Therefore, we can expect the maximum ensemble model also to bring more detections but less accurate picks.

In contrast, we expect the semblance-ensemble model to depend rather on the consistencies of the values. The examples shown in Fig.7.7 illustrate that the resulting probabilities are lower by nature of the product of the individual models. A user must choose a lower threshold value for the same confidence in the picks.

We now compare them over a broad test dataset. The test dataset contains the 12,000 waveforms

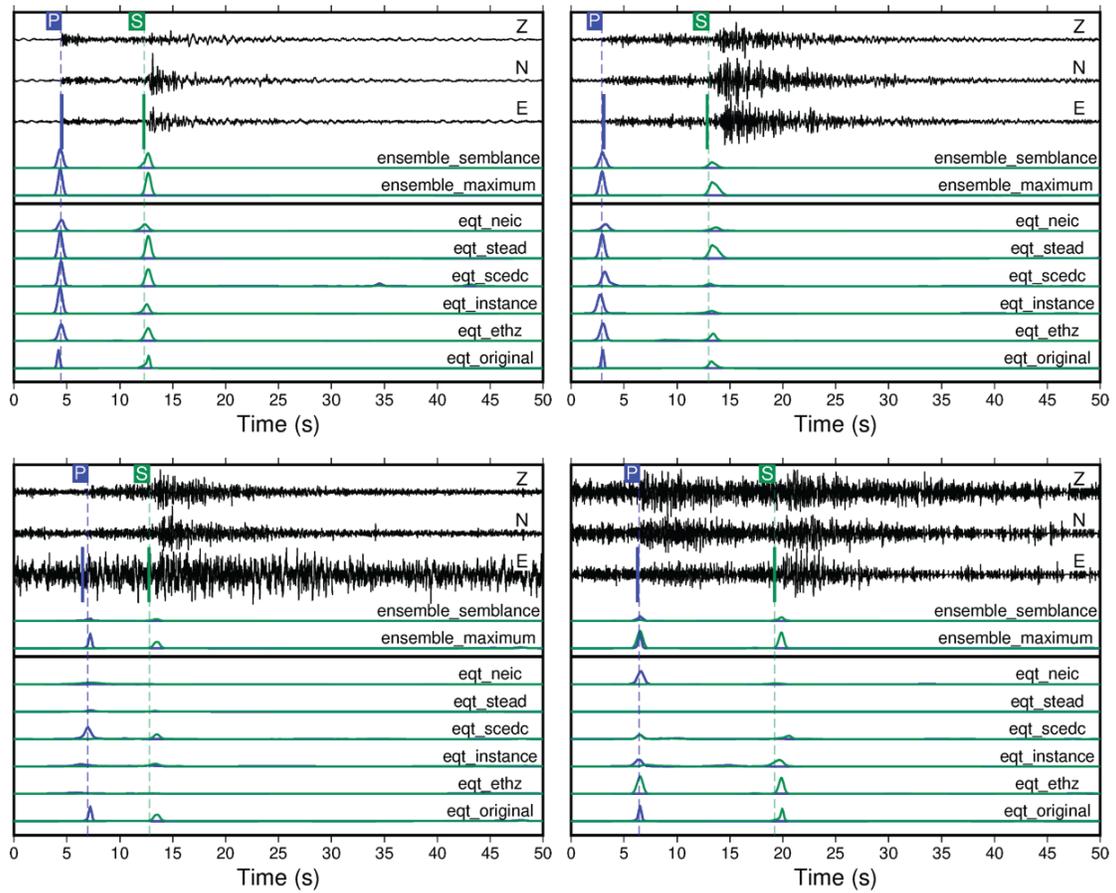


Figure 7.7: Example predictions with pretrained models and ensemble estimations on four multi-channel seismograms from the INSTANCE dataset. The dashed lines mark the labeled (or true) arrival times of the P- and S-phases. The solid lines denote the arrival times estimated by the learnable ensemble, which are only shown on the E components but calculated over the three channels. The lower plots show the probability distributions annotated with their respective base models. The upper probability distributions are annotated from the estimations by two simple ensemble methods (maximum and semblance). The P-phase and S-phase probabilities are plotted in blue and green, respectively.

of each SNR group not used for training the meta-learner (see previous section for details). All models improve performance (i.e., pick rates and errors). We show results in Fig.7.8. Similarly to the comparison between the multiband and the broadband methods, the maximum-ensemble method detects more events (e.g., higher P- and S-phase pick rates) but with larger errors.

In this exercise of deploying these models on ground truth information, we observe that a greater picking rate from maximum-based ensembling may yield large errors: pick errors of several seconds may imply the picking of spurious arrivals that would become impractical for earthquake location techniques. Ensemble learning is a large step of improvements, with errors twice as small, relative to a pick rate reduced by 10%.

7.3.2 PERFORMANCE ANALYSIS ON THE PNW DATASET

We now explore the cross-domain performance of the ensemble methods relative to new, unknown data and compare to transfer-learning one of the base models with a new ground truth dataset in the Pacific Northwest of the United States. The PNW region has a relatively low seismicity rate compared to the active tectonics of the region. While there might be some PNW data in STEAD, we believe data leakage is not consequential. The PNW datasets[?] contain regional earthquakes with relative travel time between P and S greater than one minute (the window length of the eqt-base models). The PNW data is not used to train the base models and serves as independent data to test the generalizability of our machine learning model. We have also retrained the EqT using 70% of the PNW data, so about 130,000 waveforms in previous work[?]. We use the ComCat datasets that include roughly 180,000 earthquake and explosion seismograms. The resulting model is called eqt-pnw. In this case, we do not select based on SNR groups but exclude approximately 1,000 seismograms with too large of relative arrival time between $T_s - T_p > 60s$.

Similarly to our analysis with the INSTANCE dataset (Fig. 7.4 and 7.5), we test the individual model predictions with the PNW data. We find similar results in the cross-domain predictions.

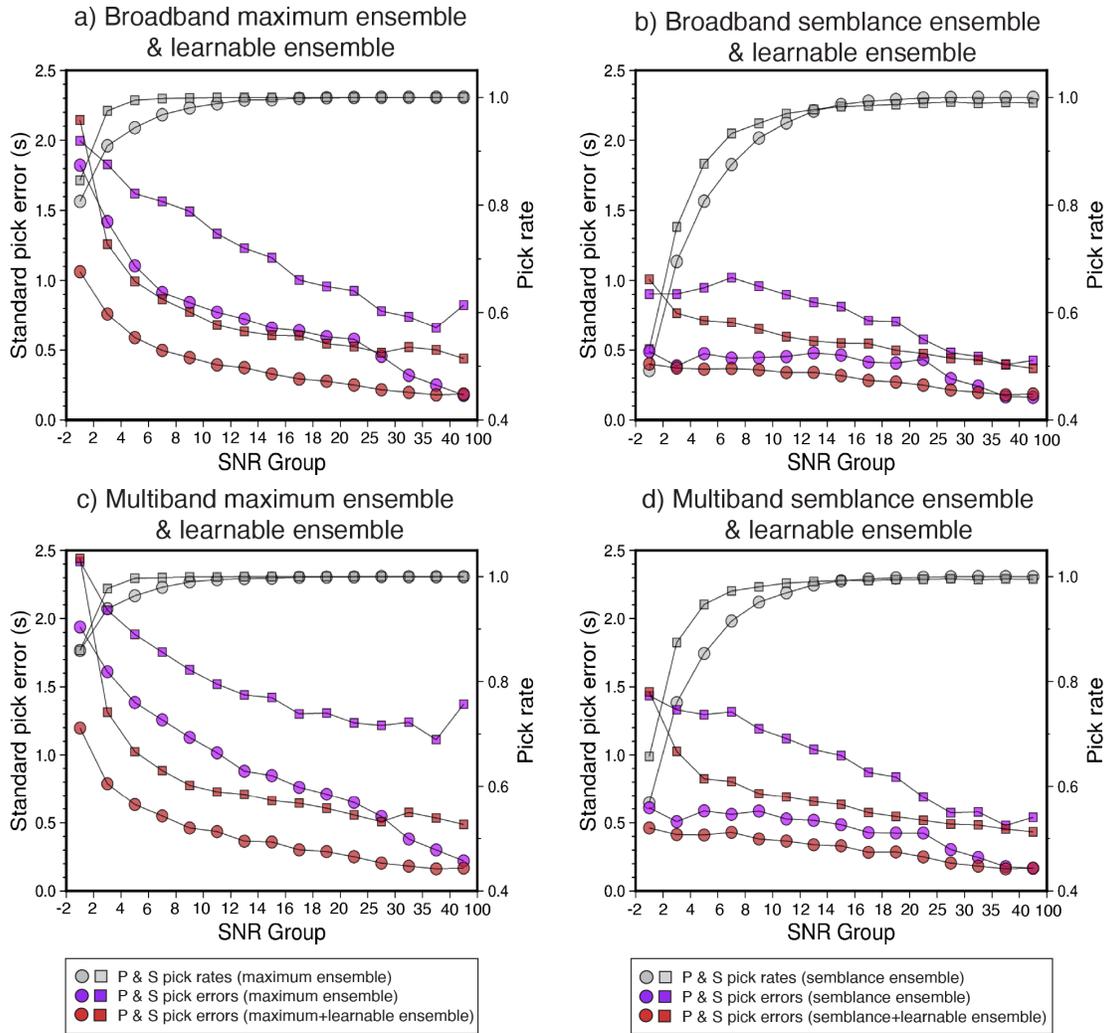


Figure 7.8: Estimations of three ensemble methods on the INSTANCE dataset. (a, b) Broadband predictions. (c, d) Multi-band predictions. Circles and squares represent P- and S-phases, respectively. Gray symbols indicate pick rates, while colored symbols show the standard deviation of the residuals for different methods. The predictions of the learnable ensemble are performed on the results of the maximum ensemble and semblance ensemble.

Overall, the performance increase with SNR is less apparent, which we attribute to the various waveforms per SNR bin. In general, the cross-domain results show that earthquakes can be detected with another model, which assesses some of the generalization power of the base models. Again, the pick errors can be significant, especially for S waves, and become impractical for continuous deployment and earthquake location. We also confirm that the multiband results exhibit an improved performance again: increased detection and pick rates but without the cost of an increased pick error compared to the INSTANCE case.

We now focus on comparing the transfer-learned model (i.e., eqt-pnw model) against the ensemble models. We train the learnable ensembling method using 3,000 waveforms from each SNR group, leaving the other waveforms in the test set. In Fig. 7.9, we show the results of pick rates and errors when comparing transfer-learned against the multiple workflows (ensemble techniques and broadband vs. multiband) on such test data.

We apply the eqt-pnw model to pick the test data used for ensemble estimations, although a big portion of the testing data has been exposed to the model for training. As such, we expect the rates and errors to be small from training. In general, the pick errors of all models are comparable, except for the maximum+learning ensemble models (Fig. 7.9). The maximum ensemble gives greater pick rates than transfer learning, especially at intermediate SNRs, but the pick errors are also much larger.

Semblance ensemble models behave similarly or better than the transfer-learned model in terms of picking accuracy. The pick rates are lower for semblance learning but also more accurate. This indicates that the semblance model generalizes relatively well to new areas or as much as what transfer-learning would do. The learnable ensemble model aided by the semblance model is comparable with the eqt-pnw, with the trade-off of picking rate versus accuracy.

Therefore, this first test on benchmark datasets demonstrates that ensemble learning is as effective as a transfer-learned model. However, it does not necessitate training a learning-based ensemble to

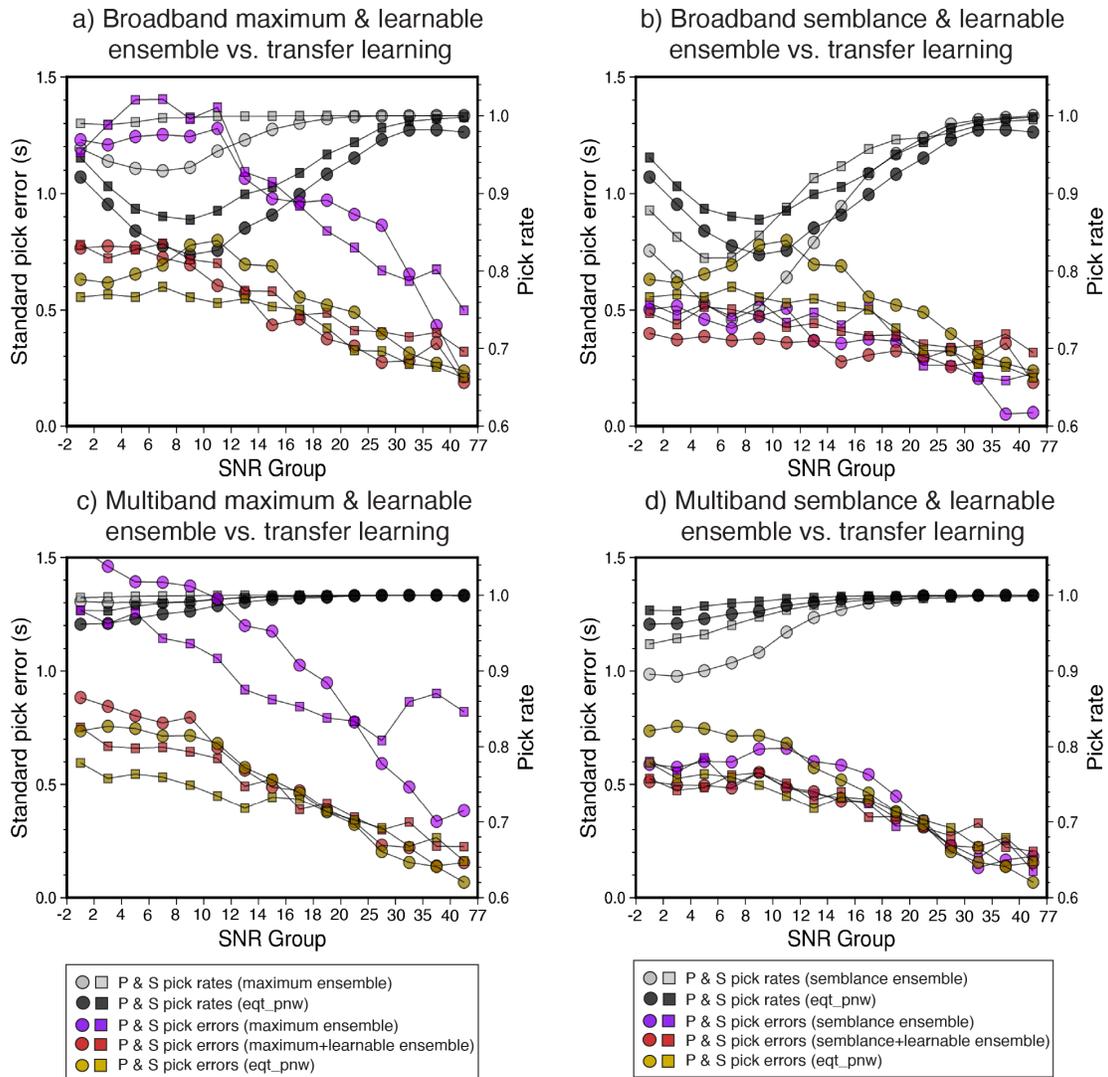


Figure 7.9: Transfer against ensemble learning in the PNW. Results from three ensemble estimation methods against test data SNR of the PNW data. (a, b) broadband workflows, (c, d) multiband workflows. Transfer learning of the eqt-pnw model is compared to (a, c) maximum ensemble and maximum + learnable ensemble and (b, d) semblance ensemble and semblance + learnable ensemble.

obtain satisfactory performance.

Given the simplicity of the learnable ensemble method, we now explore the performance of the ensemble model with the size of the training data (Fig. 7.10) for two situations: training on the INSTANCE (some within-domain) and on the PNW (nearly complete cross-domain). We select the data of equal amounts from different SNR groups. Because small biases have been observed in previous cases², we explore these as the mean absolute errors (MAE) and the standard picking errors (standard deviation). The training is done quickly over 20 epochs. We find rapid convergence of the training with the amount of training data: 1000 labeled data already provides acceptable errors and is improved until a few 1000s. This demonstrates that little data is needed to achieve good performance in ensemble learning.

7.3.3 BLIND TEST ON OCEAN BOTTOM SEISMOMETER (OBS) DATA

Seismic data collected on the sea floor have distinct seismic signals from seismological research, mostly on land seismic data. The seismic signals are largely impacted by oceanic noise that pollutes earthquake signals². Therefore, data obtained from Ocean Bottom Seismometers (OBS) present a considerable challenge to test our ensemble learning approach. We use the recently published benchmark datasets by² to test the model performance. It is typical to work on bandpassed data in offshore seismology, as whale calls and other oceanic wave signals may contaminate the seismograms².

We focus on performance tests on the two non-learnable ensemble learners (maximum and semblance) in the multiband predictions. The test is completely cross-domain without any possible data leakage, given the recent compilation of the benchmark datasets. We illustrate the performance of the two ensemble learning models on the predictions of P and S picks in Fig. 7.11. Consistent with previous tests, the ensemble semblance demonstrates superior performance over the maximum ensemble in terms of Mean Absolute Error (MAE), Median Absolute Deviation (MAD), and Root

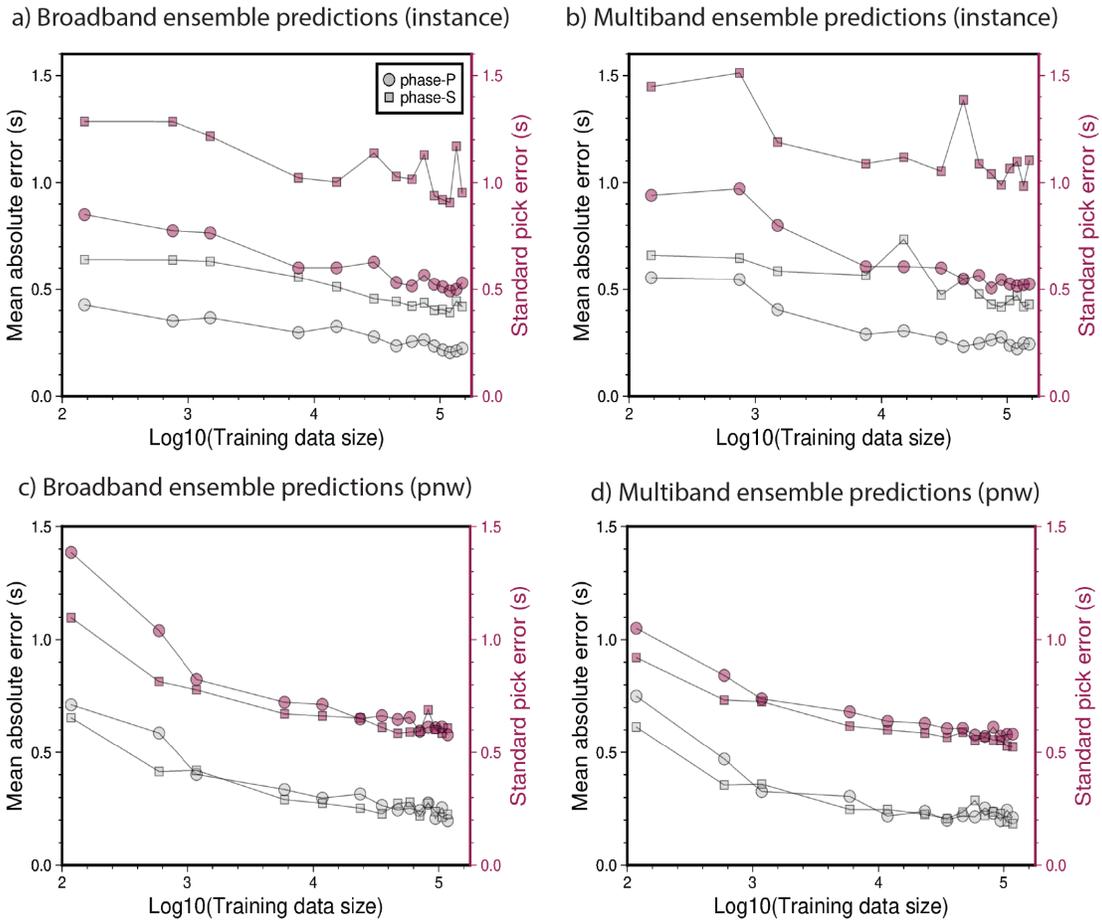


Figure 7.10: Effects of training data size on the performance of learnable ensemble predictions. (a, b) Broadband and multiband learnable ensemble predictions for INSTANCE dataset. (c, d) Broadband and multiband learnable ensemble predictions for PNW dataset.

Mean Square Error (RMSE), even though the ensemble maximum method results in more picks than the ensemble semblance. As expected, the picking results of the P-phase are superior to those of the S-phase. We report that our results are better than those obtained transfer-learned model as documented in², without retraining. For P waves, we predict an RMSE is 0.21 (they predict 0.33), MAD is 0.01 (they predict 0.07), MAE is 0.08 (they predict 0.32) with fewer outlier factors (0.06 in our case, 0.07 in their case). Our results for S waves are also improved by a factor of approximately 40%. These results are encouraging that deploying a multiband semblance ensemble will provide a good algorithm for the continuous deployment of OBS data.

We perform an additional test to train the ensemble learner on the OBS benchmark data set². We report a decreased performance for the phase picks. This could be attributed to large biases across data predictions, which is difficult for the learnable ensemble. Given that the original semblance predictions were better than the transfer model, we suggest the use of the semblance ensemble as a sufficient workflow for simplicity.

7.4 DEPLOYMENT ON CONTINUOUS DATA

One goal of these models is to detect and pick direct wave arrival times on continuous data. We demonstrate the prediction performance and computational cost using a four-hour seismic waveform captured by borehole station PB.B204, located near Mount St. Helens in the Pacific Northwest. A borehole sensor is an appropriate choice for this test because borehole instruments are a lot quieter than surface measurements.

To prepare the continuous waveforms for analysis, we first perform a few preprocessing operations: i) resampling to 100 Hz and detrending (using the ObsPy function `detrend`), ii) filtering within a fixed frequency band (1-45 Hz) using the ObsPy function `bandpass`, and iii) trimming the data into 60-second windows with a 50% overlap, resulting in a total of 479 windowed waveforms.

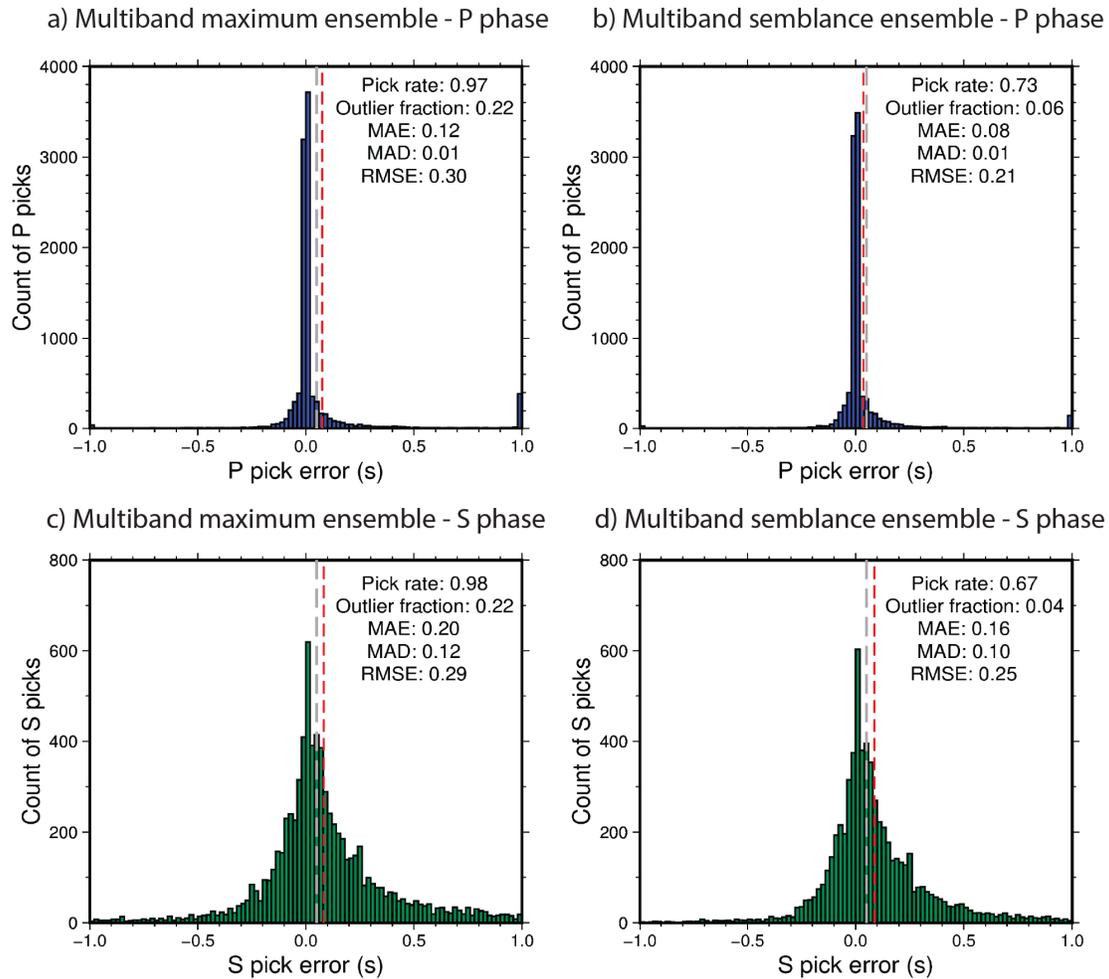


Figure 7.11: Ensemble learning on OBS benchmark data. Histograms of P and S pick errors between the manual picks and estimated picks from maximum ensemble (a, c) and semblance ensemble (b, d). The ensemble estimations are performed upon multiband predictions. The outlier fraction is the fraction of outliers with pick errors larger than 1s. MAE, MAD, and RMSE denote mean absolute error, median absolute deviation, and root mean square error, respectively. The vertical gray and red dashed lines represent the median and mean pick errors, respectively. Both median and mean values are influenced by uneven errors.

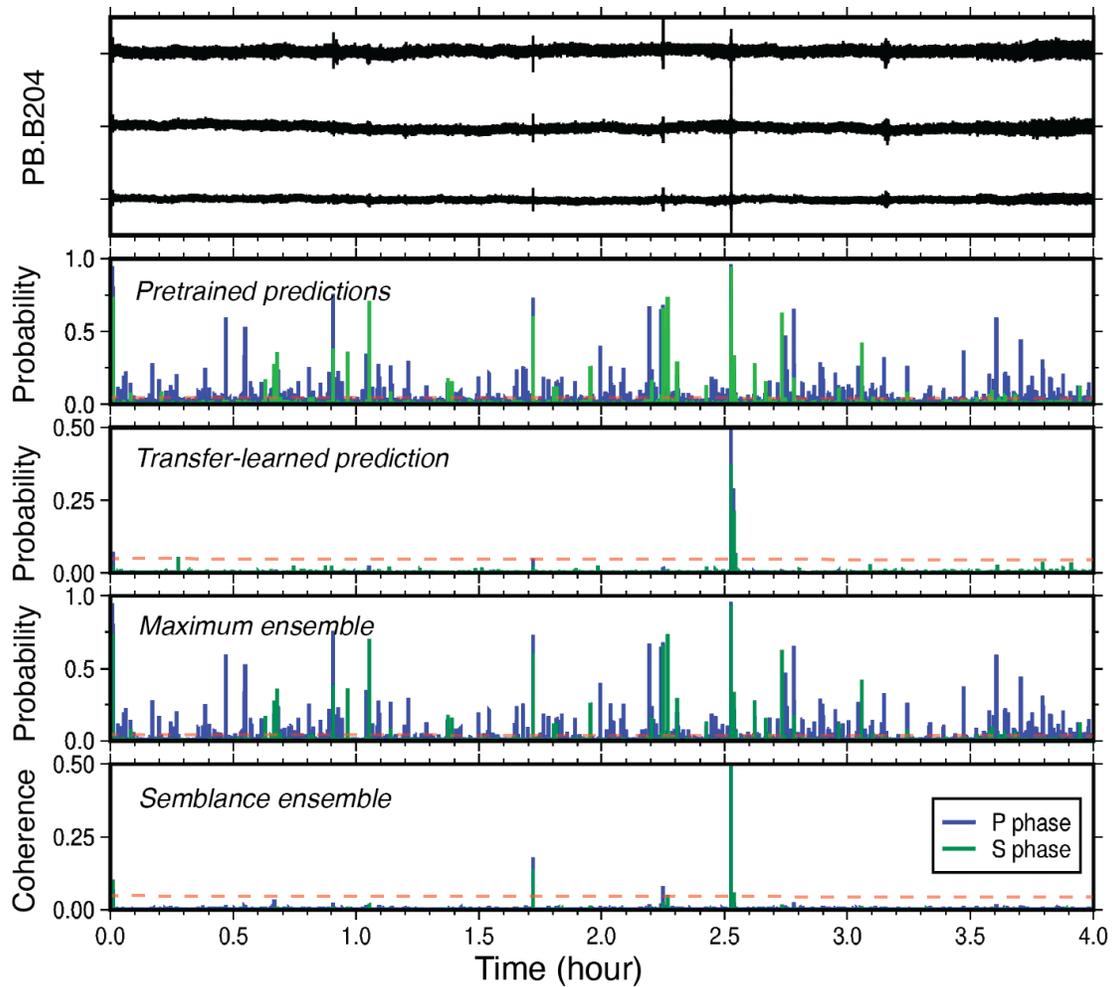


Figure 7.12: Prediction results on 4-hour continuous seismograms recorded by borehole broadband seismometer at PB.B204. Prediction on the waveforms shown on the data itself. The first panel is continuous time series for the three channels. The second panel shows the probability results predicted by six pretrained models individually. The third panel shows the probability results predicted by the transfer-learned model (eqt-pnw). The fourth and fifth panels show the results estimated from the maximum ensemble and semblance ensemble, respectively. The picking results of the P and S-phases are marked in blue and green. The dashed red lines represent the probability or coherence threshold for the picking identification.

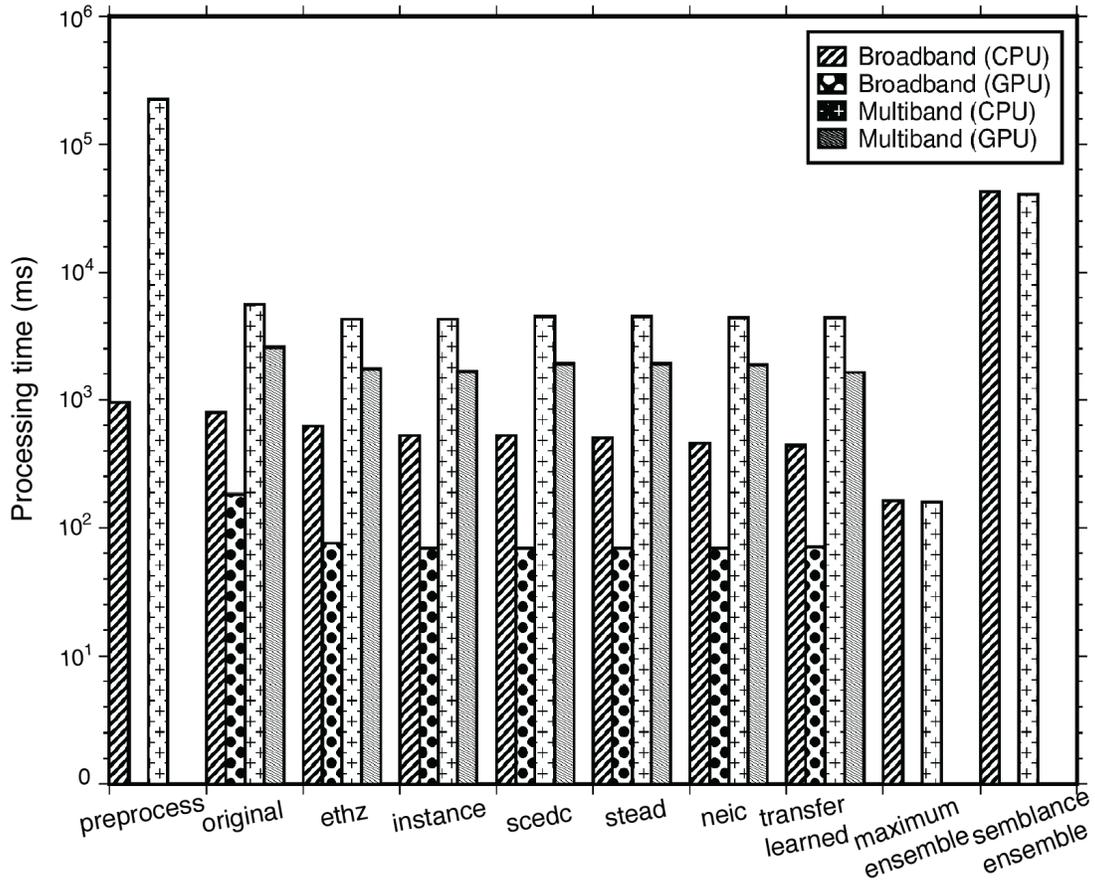


Figure 7.13: Computational efficiency on 4-hour continuous seismograms recorded by borehole broadband seismometer at PB.B204. Computation cost in milliseconds (ms) of all steps processing the continuous waveforms in (a). These steps include preprocessing (left side), six successive pretrained predictions, one transfer-learned prediction, and two ensemble estimations (right side). The comparison between CPU and GPU is also performed for prediction steps. The preprocessing and ensemble steps are only conducted on the CPU.

Next, we apply the six pretrained model predictions and the eqt-pnw (transfer-learned) prediction using batch prediction.

We predict phase picks from the maximum-based and semblance-based ensemble methods based on the six pretrained models. Fig.7.12 shows the probabilities of picks for P and S waves on the continuous time axis. The test on continuous data suggests that predictions from the individual base models (cross-domain) and the maximum ensemble model give rise to so many high-probability pre-

dictions that many must be artifacts and false positives. These artifacts are problematic in building an earthquake catalog because they need to be further analyzed with other stations to be selected or dismissed. Furthermore, errors may also occur for the picking arrival times, yielding poor estimates of earthquake locations.

The best model predictions are from the semblance ensemble and transfer-learned models. The ensemble model predicts the four events and appears to suppress the rest as noise. The transfer-learned model (eqt-pnw) picks up three events, missing a pair of P and S picks from two events. We also compare the results with the PNSN catalog of phase picks, which only includes two events. These picked events are visually relatable to small earthquakes. When using a nearby surface station (UW.SHW), we report that the transfer-learned model performs much worse than the ensemble model. Repeated tests also demonstrate that the ensemble semblance behaves more robustly by picking fewer false positives compared to the ensemble maximum and compared to the individual base models. Furthermore, it consistently detects and picks more events than the transfer-learned model does. We inspect detected events individually, which are all low-magnitude earthquakes and picked up with accurate phase arrivals.

The computational cost of deploying the proposed framework is crucial for considering the deployment “at scale” of the proposed phase picker. All the steps of the workflow are serialized and do not include obvious improvements such as parallelization. Fig. ?? shows the processing time for batch predictions in both CPU and GPU environments within Google Colaboratory (T4 instance). The figure illustrates the case for one 4-hour continuous data. We also compare the computing cost of individual batch predictions from base models. The multiband workflow requires ten predictions per windowed waveform and therefore takes ten times longer to predict. Given the lack of a CUDA version for multiband filtering, multiband processing requires significantly more time than standard processing. However, parallel processing could alleviate the computational load with multiple processors. Example tutorials on deploying ELEP on continuous data are listed in the acknowledgment

section.

7.5 DISCUSSION AND CONCLUSION

We have developed a sophisticated framework that integrates existing pretrained models with ensemble estimations for earthquake detection and seismic phase picking. The effectiveness of this framework has been substantiated through testing on two established benchmark datasets (INSTANCE and PNW), a blind test on a recent dataset (PickBlue), and continuous data. Our framework is primarily comprised of two main components.

The first is a Multiband Filter (MBF)-based prediction workflow to generate baseline detection and picking predictions for each pretrained model. This methodology, although more computationally demanding, outperforms traditional broadband predictions by enhancing event detection, phase picks, and overall picking accuracy. Multiband workflows overall outperform the broadband workflow, especially at low SNR levels. This is promising to detect robustly new small events that are masked by ambient seismic noise.

Second, we introduce three ensemble estimation methods: ensemble maximum, ensemble semblance, and ensemble learner. The first two estimations (maximum and semblance) are deployed for phase detection and to establish reference final picks. The results from these two estimations may be input in a third ensemble model to predict the scalar values of P- and S-phase picks. Remarkably, the ensemble learner's accuracy is on par with transfer-learning or within-domain predictions, with picking detection contingent upon ensemble maximum or ensemble semblance. We find, in general, that the semblance ensemble method provides accurate and conservative detection of P and S picks in seismic data. Overall, P and S picks are much improved compared to the predictions from individual base models. A blind cross-domain application on a new benchmark dataset of OBS demonstrates impressive performance in predicting the P and S picks without the need for transfer

learning.

Certain limitations remain. First, the computational expense of the multiband predictions is considerable. The computational cost scales linearly with the number of frequency bands and the number of base models chosen. Users may decide on the frequency bands appropriate for their research applications considering characteristic frequencies of the earthquake signals and seismic noise. For instance, researchers may anticipate the frequency band of the earthquake signals relative to the local noise and decide on corresponding frequency bands that would optimally separate them. The broadband prediction may suffice for common seismic monitoring. Furthermore, we find for the maximum ensemble that by increasing the true positive rate (recall) and identifying more false negatives, the ensemble methods could potentially increase false positives, thereby reducing precision. Therefore, to maintain a high Area Under the Curve (AUC, one performance measurement of classification thresholds) value, it might be beneficial to set a higher selection threshold. We refer to Münchmeyer et al.² for a detailed performance evaluation of broadband predictions. Semblance ensembling provides a good balance while the overall probability becomes lower, and users ought to adjust for this expectation. Our blind tests conducted on two sets of four-hour continuous data illustrate that the ensemble semblance method outperforms the ensemble maximum method, which often misidentifies noise as signals.

Additionally, the ensemble learner, framed as a regression problem, has the ability to estimate any given input data without a need for a threshold. Despite its proficiency in identifying low probability or inconsistent phase predictions, it currently lacks the capability to independently exclude non-phase signals or detect multiple-phase signals. An additional limitation of our framework is that it requires a fixed input data of six based models and a specific window length. The simple model architecture can be adapted efficiently to other base models and seismic window lengths.

This study demonstrated the vast range of model behaviors with respect to the diversity of training data and the usefulness of combining predictions for a much-improved performance. We also

demonstrated that ensemble learning might be as powerful as transfer learning, except that it does not necessarily require labeled data to make accurate and reliable predictions, which we assess based on previously published analysis. Our study thus emphasizes the importance of training datasets, rather than model complexity, in generalizing supervised learning algorithms.

8

Conclusion and outlook

At this stage, I would finalize this thesis with an emphasis on the main outcome, contribution, and potential impact of the works that I have completed and have been working on. My studies target and contribute to two fundamental scientific questions: **How does fluid and/or deformation influence the seismoacoustic properties of porous media? How does fluid modulate deformation, fracturing, and seismoacoustic behaviors?**

In Chapter 2, I develop two wavelet-based approaches, namely wavelet transform stretching

(WTS) and wavelet transform dynamic time wrapping (WTDTW), and compare and complement them to existing time-domain, frequency-domain, and wavelet-domain approaches for measuring velocity changes. I also explore the depth-sensitivity of velocity changes with wavelet-domain approaches. I conclude that the velocity-change spectrum varies with the depth of velocity change. This indicates that the spectrum can be utilized to constrain and infer the depth of velocity change in the field observation. Beyond that, I also conduct factor analysis of the depth sensitivity of velocity change. I find that the characteristic frequency is most sensitive to the background velocity model and time-lapse of measured coda waves, while the characteristic velocity change is most sensitive to the perturbation depth and thickness[?]. Further, I have applied this method to one field case of monitoring wastewater injection, where I utilized WTS for measuring velocity-change spectra and constraining the depth of velocity change with the depth sensitivity analysis[?].

Currently, 2D coda-wave sensitivity kernels for homogeneous media can be accurately calculated for different frequency bands and time lapses. The depth of the imaging is usually calculated through the depth-sensitivity kernels of surface waves. Our method of recovering continuously frequency-dependent changes could help the diagnosis of velocity-change depth as stated previously. Additionally, the velocity changes in depth can be better constrained with the accurate calculation of 3D coda-wave sensitivity kernels. Other alternatives would use ballistic waves instead of scattered waves for 4D velocity-change imaging. Current advances in dense seismic array and DAS acquisition are opening up new opportunities for time-lapse imaging techniques.

In Chapter 3, I develop an experiment apparatus by leveraging 3D-printed granular media and active-controlled acoustic monitoring technique to investigate frequency-dependent phase shifts (i.e., velocity changes) and amplitude change (i.e., attenuation change) with respect to saturation and/or deformation. During the water table increases in the dry media, I find a gradual change in both velocity and attenuation spectra. I also find that both velocity and attenuation spectra have different responses to consolidated and unconsolidated media. With the fluid saturation, however,

different responses disappear as the water dominates the velocity and attenuation changes. One limitation of our experiments is that the overpressure is unconsidered. As the overpressure-induced deformation and saturation may also be important mechanisms behind observed velocity changes, I will complement this mechanism investigation into this ongoing study. Additionally, my new platform could help us to investigate the strain- or stress-sensitive velocity and attenuation spectral changes. This would provide insights response of Earth's materials under different conditions to the stress perturbation.

In Chapter 4, I explore the possibility of probing deep volcano's magmatic activity by leveraging inter-source interferometry and repeating earthquakes. Our example demonstration shows the wavefield between two repeaters can be effectively constructed. I also perform the coda wave interferometry on the cross-correlations between two sources over time. The velocity changes show some interesting patterns, that ought to further be improved and compiled for other pairs of sources. Further efforts are needed for: 1. performing inter-source interferometry and coda wave interferometry for all possible source pairs; 2. inverting the measure discrete velocity changes for continuous velocity changes; 3. discovering hidden clues in the calculated velocity changes for possible magmatic activity.

In Chapter 5, I perform hydrofracturing experiments, where detailed fluid flow, fracture, and radiated acoustic emissions (AEs) can be clearly visualized and listened to. I conduct a holistic analysis of fractures and AE energy for a high-viscosity case and a low-viscosity case. Both fracturing cases demonstrate similar stick-break instabilities. I find that radial crack propagation is slow (~ 3 m/s) and facilitated by tangential fractures, which is much faster (~ 1000 m/s). Fluid viscosity and pressure The fracture dynamics of slow and fast events are controlled by fluid viscosity and pressure for the inter-event time and the energy released during individual fast events. Additionally, I find that these AE signals share behaviors with observations of episodic tremors in Cascadia, United States in terms of bursty or intermittent slow propagation and nearly linear scaling of radiated energy with

area. The laboratory experiments provide a plausible model of tectonic tremor as an indicative of hydraulic fracturing facilitating shear slip during slow earthquakes.

In Chapter 6, I develop a new experiment apparatus by injecting high-pressure water into an artificial fault-valve model. I visualize the complete compaction and dilatancy dynamics prior to a failure of the barrier. I find a fluid surge caused by the compaction followed by the bi-direction migration during the dilatancy. The forward pressure wave breaks the barrier while the backward pressure decreases the injection pressure as measured. The recorded injection pressure variation reveals the hydromechanical behaviors. In particular, both dilatancy and barrier failure can cause pressure drop but with different temporal decays. Additionally, I use cumulative acoustic energy to indicate mass deformation. The increase rate could be used to suggest the barrier failure. Inferring physical processes from acoustic energy evolution is difficult. In the field, this challenge can be overcome by the detailed analysis of earthquake source property. Further, I plan to use hydromechanical modeling to reproduce the laboratory's physical processes and scale them up to realistic contexts.

In Chapter 7, I develop an ensemble-learning framework, namely *ELEP*, for earthquake detection and phase picking. I find it has great generalizability by ensemble estimations over probability predictions from multiple deep-learning pickers. Without transfer learning or further refining, I demonstrate the effectiveness of *ELEP* in the on-land and obs datasets, triggered events, and continuous waveforms. My software package can be readily utilized for any seismic recordings and is being tested for full operation at a seismic network. The recent tests on the DAS recordings further show the robustness and effectiveness of the proposed *ELEP*. I also utilize it to scan the 2007-2022 continuous data recorded by multiple permanent stations deployed at Mount St. Helens and obtain over three times more than the PNSN catalog, which may warrant further analysis to associate the new picks and identify actual events. The result is shown in Chapter 4. Moreover, the current *ELEP* includes the module for determining first motions, which would help to solve for focal mechanisms of small earthquakes. In the near future, I aim to build up a machine-learning-powered automated

toolkit for high-performance seismic monitoring and earthquake early warning.

In summary, this thesis contributes significantly to the understanding of fluid interactions within the Earth's subsurface and their impact on seismic behaviors. It reveals how fluids and deformations alter seismic properties in porous media, affecting fracturing and seismoacoustic activities. The research develops new wavelet-based methodologies for analyzing velocity changes and their depth sensitivities, providing insights into subsurface fluid movements and their implications on seismic imaging. Experiments with 3D-printed media and high-pressure fluid injections shed light on the hydrofracturing dynamics and the intricate relationship between fluid viscosity, pressure, and fracture behaviors, drawing parallels with natural tectonic tremors. Additionally, the study delves into the hydromechanical behaviors of artificial fault-valve systems, illustrating the complex interactions between fluid pressure, compaction, and dilatancy dynamics. On the technological front, the thesis advances seismic processing through the development of an ensemble-learning framework for improved earthquake detection and phase picking, showcasing its application in enhancing seismic network operations. Collectively, these findings offer a comprehensive model of fluid-induced seismic phenomena, bridging laboratory observations with field-scale seismic activities, and paving the way for refined seismic hazard assessments and resource management strategies.



A.1 DEMONSTRATION THAT THE REAL PART OF THE CWT IS A NARROW BANDPASSED VERSION OF THE ORIGINAL TIME SERIES

The time series reconstruction from the wavelet domain is derived by ? with the following equation:

$$\bar{X}_n = \frac{\delta_j \delta t^{1/2}}{C_\delta \psi_0(0)} \sum_{j=j_1}^{j_2} \frac{\mathcal{R}\{W_n(s_j)\}}{s_j^{1/2}}, \quad (A1)$$

where \mathcal{R} denotes the real part. The factor $\psi_0(0)$ removes the energy scaling. The factor C_δ is derived by reconstructing a δ function from its wavelet transform. It is a constant for each wavelet

function and is derived in [?]. The \bar{X} is the reconstructed time series over a band from $f_1^{#b}$ to $f_2^{#b}$ scale or frequency. The reconstruction is performed at a specific scale or frequency when j_1 equals to j_2 . In this case, the above equation is reduced to

$$\bar{X}_n = \alpha \mathcal{R}\{W_n(s_j)\},$$

$$\alpha = \frac{\partial j \partial t^{1/2}}{C_\delta \psi_0(0) s_j^{1/2}}. \quad (A2)$$

Based on the reconstruction equation, we can transform the wavelet function of particular frequency or frequency bands back to the corresponding time series in the time domain. The wavelet transform and its reconstruction of specific frequency or frequency bands are equivalent to the frequency filter in signal processing. One natural way of obtaining dv/v of each frequency is to apply TS or DTW on the time series filtered at each frequency with the wavelet-domain. However, reconstruction of the time series at each frequency would be computationally costly. Based on the above equations, we find that the real part of the wavelet transform $\mathcal{R}\{W_n\}$ is used for the reconstruction, and the coefficient α is an independent constant, which allows us to directly use $\mathcal{R}\{W_n\}$ without considering the scaling constant α , since the scaling is usually ignored in extracting phase variations of two time series. In this way, we are permitted to employ TS or DTW on the $\mathcal{R}\{W_n\}$ without additional time cost for reconstruction. To measure dv/v results in specific frequency bands (as against single frequencies), one can alternatively apply the WTS and WTDTW on the filtered signals in frequency bands of interest.

A.2 PARAMETERS FOR THE STRUCTURAL MODEL, THE NUMERICAL SIMULATIONS, THE dv/v CALCULATIONS USING ALL APPROACHES, AND PERFORMANCE SUMMARY.

Table A.1 gathers information about the numerical exercise. Table A.2 gathers the explicit parameters necessary to apply the algorithms for computing dv/v either in time-, frequency-, and wavelet-domains. The parameters can be found for either different source-receiver geometry (i.e. zero or distance offset) or different media (i.e. half-space and layered medium). We summarize the performances of all three categories of methods in terms of their general practicability, accuracy, computational efficiency, stability, and noise resistance in Table A.3.

Table B1. The configuration of parameters for media and simulations in all numerical experiments.

Parameter Setup	Half-space	Layered half-space cases		Depth-sensitivity cases	
	Case	Zero-offset	Distance-offset	Zero-offset	Distant-offset
Size (l_x, l_z) (km)	(100,50)	(100,50)	(200,100)	(20,5)	(200,100)
Spacing (dx, dz) (km)	(0.5,0.5)	(0.5,0.5)	(1.0,1.0)	(0.1,0.02)	(1.0,0.5)
Background V_p (km s ⁻¹)	4.0	layer1=4.0 layer2=6.0	layer1=4.0 layer2=6.0	4.0	4.0
Medium V_p/V_s	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$
density (kg m ⁻³)	2000	2000	2000	2000	2000
Velocity fluctuation (per cent)	20	20	20	20	20
Correlation distance (a_x, a_z) (km)	(0.8,0.5)	(0.8,0.5)	(3.0,1.0)	(0.2,0.2)	(3.0,1.0)
Velocity perturbation (per cent)	0.1	0.1	0.1	0.1	0.1
Source-receiver distance/offset (km)	0	0	100	0	100
Simulation Source function	<i>Ricker</i>	<i>Ricker</i>	<i>Ricker</i>	<i>Ricker</i>	<i>Ricker</i>
Source direction	<i>Vertical</i>	<i>Vertical</i>	<i>Vertical</i>	<i>Vertical</i>	<i>Vertical</i>
Source frequency (Hz)	1.0	1.0	0.3	5.0	0.3
Sampling rate (s)	0.01	0.01	0.01	0.0004	0.006
Simulation duration (s)	80	80	300	16	240

Table A.1: Simulation configuration.

Table B2. The dominant parameters setup for all approaches in Table B3 to measuring dv/v for all numerical experiments.

Parameter setup	Half-space	Layered half-space cases		Depth-sensitivity cases	
	case	zero-offset	distant-offset	zero-offset	distant-offset
Time window (T_{min}, T_{max}) (s)	(5,35)	(20,60)	(180,240)	(2.5,6.5)	(185,225)
Frequency range (F_{min}, F_{max}) (Hz)	(0.5,3.0)	(0.5,5.0)	(0.1,1.0)	(0.5,30)	(0.05,1.5)
Moving-window duration (s) for WCC and MWCS	5.0	5.0	5.0	n/a	n/a
Moving-window step (s) for WCC and MWCS	2.5	2.5	2.5	n/a	n/a
Computing dv/v Maximum dv/v for TS and WTS	0.02	0.02	0.02	0.02	0.02
Coefficient number for TS and WTS	1000	1000	1000	1000	1000
Maximum searching points for DTW and WTDTW	50	50	50	n/a	n/a
b-value for DTW and WTDTW	1	1	1	n/a	n/a

Table A.2: Measurement configuration.

Table B3. Performance comparison within each one of three groups of approaches, which include the one with three time-domain approaches, the one with four approaches for narrow-frequency bands, and the one with three approaches for all individual frequencies. Ranking goes from low to high in the following order: Low, Intermediate, High and Advanced High.

Methods	General Practicability	Accuracy	Computational efficiency	Stability	Noise Resistance
Three time-domain approaches					
WCC	High	Intermediate	High	High	Intermediate
TS	Intermediate	High	Intermediate	High	Low
DTW	Low	High	Low	Intermediate	High
Four approaches for narrow-frequency bands					
MWCS	Low	Intermediate	High	Low	Intermediate
WCS	Intermediate	Intermediate	Intermediate	Low	Intermediate
WTS	Advanced High	Advanced High	Intermediate	High	High
WTDTW	High	High	Low	Intermediate	Low
Three approaches for all individual frequencies					
WCS	Low	Low	High	Low	Intermediate
WTS	High	High	Intermediate	High	High
WTDTW	Intermediate	Intermediate	Low	Intermediate	Low

Table A.3: Performance summary over all approaches

A.3 SUPPLEMENTARY FIGURES - ADDITIONAL FIGURES FOR CHAPTER 2

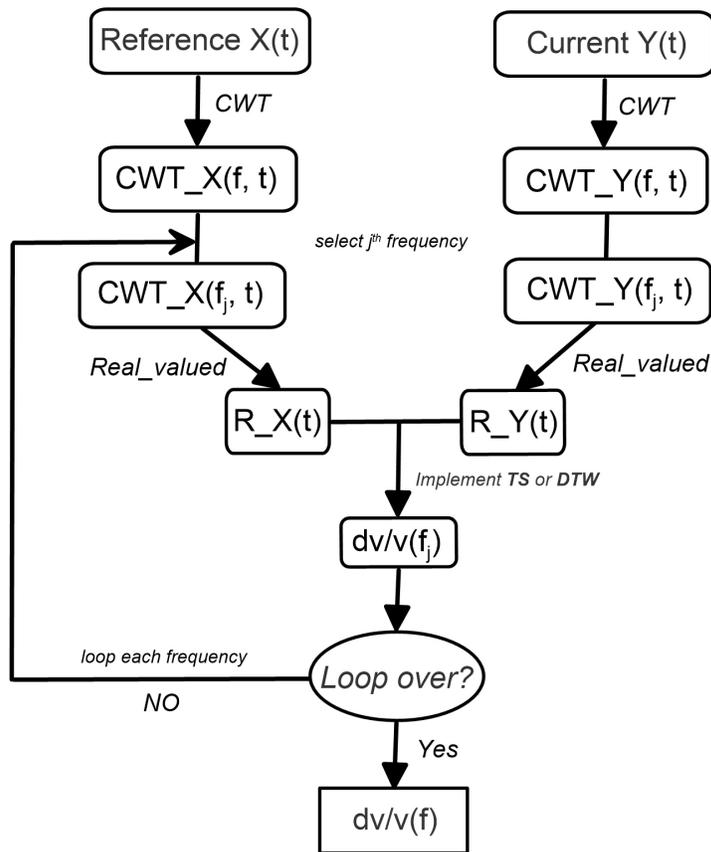


Figure A.1: Workflow diagram of WTS and WTDTW methods. *CWT* refers to the continuous wavelet transform and *R* refers to its real part. The steps are: 1. Apply the CWT to both reference $X(t)$ and current $Y(t)$ time series; 2. at each frequency f_j , extract the real part of the CWT, $R_X(t)$ and $R_Y(t)$; 3. implement TS or DTW on $R_X(t)$ and $R_Y(t)$ to extract $dv/v(f_j)$; 3. repeat step 2 at all frequencies.

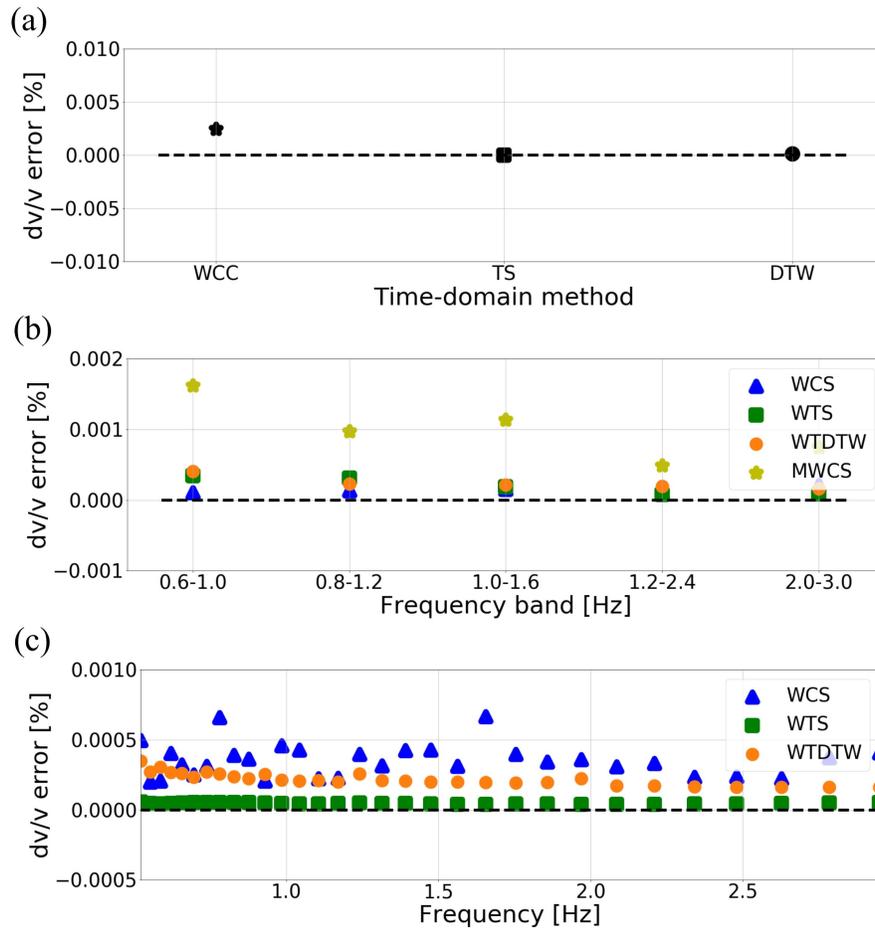


Figure A.2: Absolute measurement uncertainties of the velocity perturbation results shown in Fig.2.2 for (a) time-domain methods, (b) frequency- and wavelet-domain methods with narrow-frequency bands, and (c) wavelet-domain methods at all frequencies.

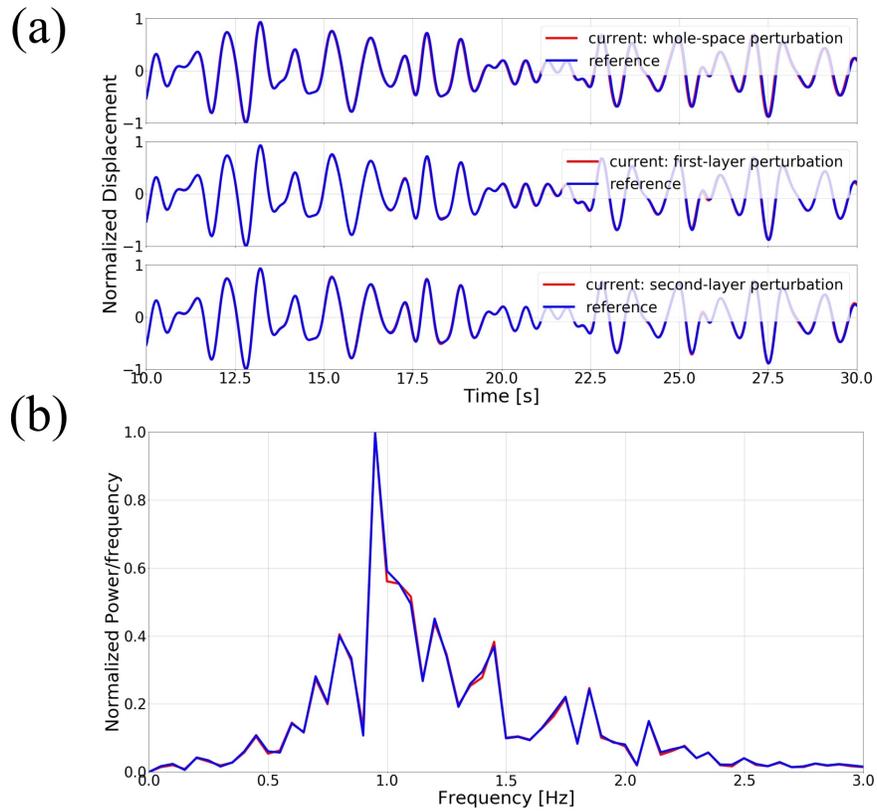


Figure A.3: Waveforms simulated in a layer-over-halfspace heterogeneous medium using SPECSEM2D solver under the zero-offset setting as in Fig.2.5. (a) Reference waveform (blue) and perturbed waveform (red) of the whole space (top), the shallow layer (middle) and the deep layer (bottom). The correlation coefficients between the perturbed and reference are 0.9901, 0.9971, and 0.9971, respectively. (b) Fourier amplitude spectra of (a) bottom panel.

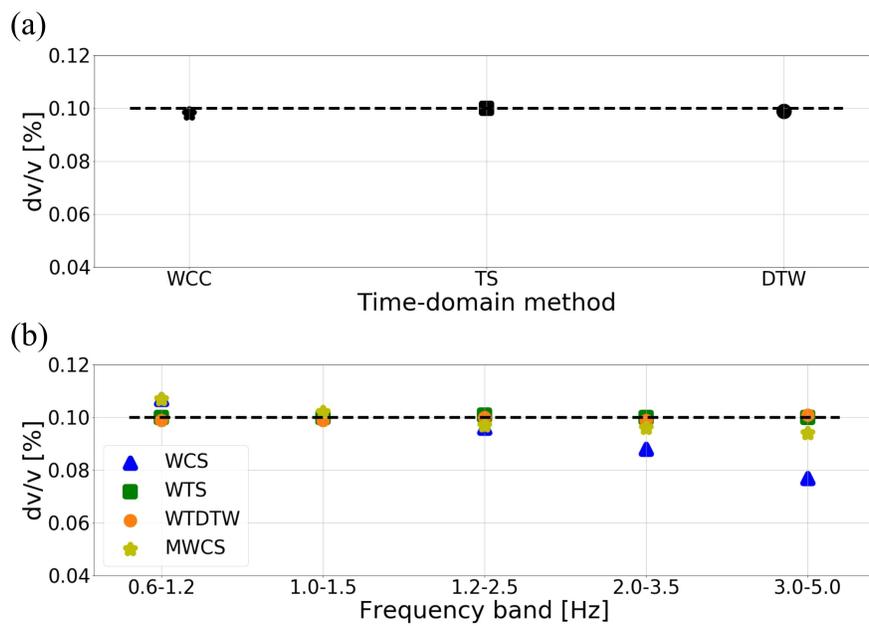


Figure A.4: dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.5 of the main text, zero-offset geometry, heterogeneous layer-over-halfspace medium, and uniform perturbation of the medium.

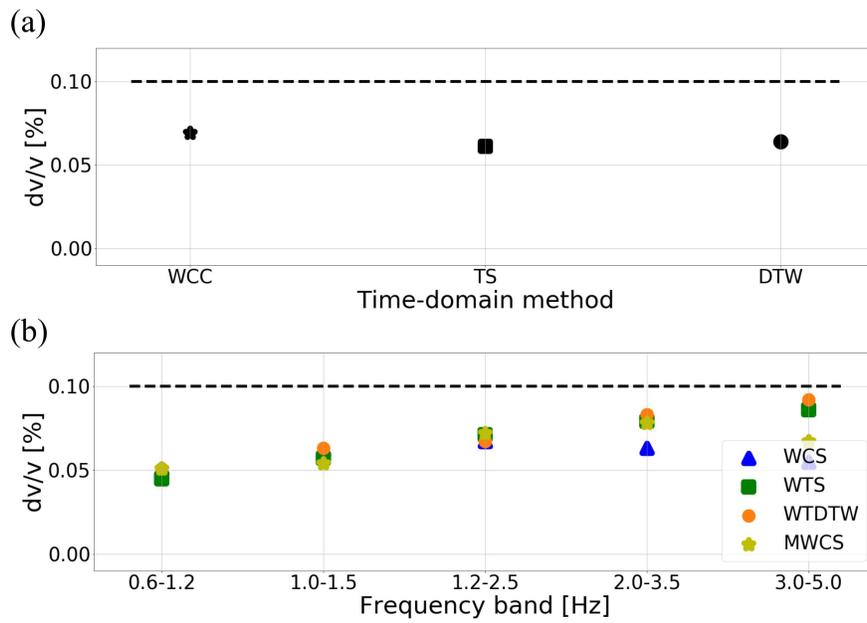


Figure A.5: dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.5 of the main text, zero-offset geometry, heterogeneous layer-over-halfspace medium, and a perturbation in the shallow layer of the medium.

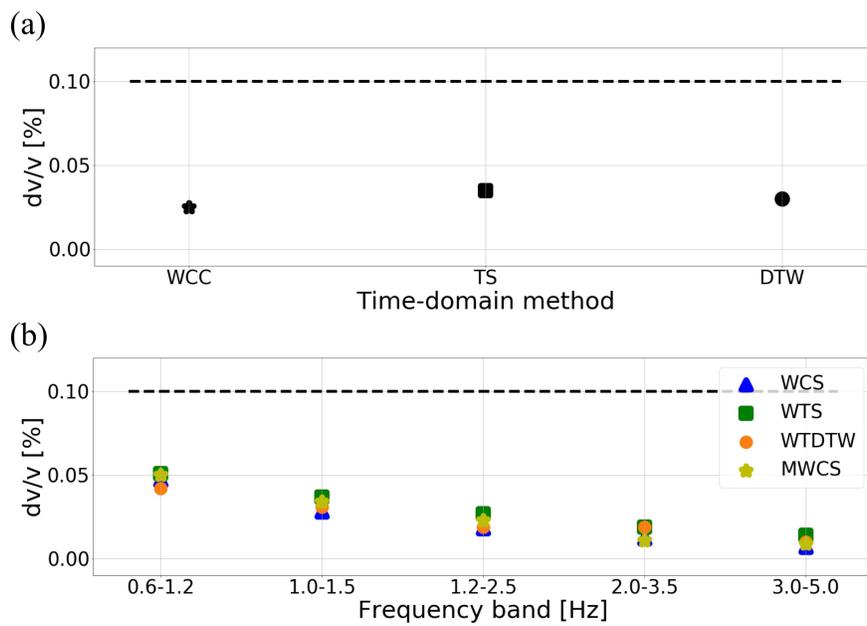


Figure A.6: dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.5 of the main text, zero-offset geometry, heterogeneous layer-over-halfspace medium, and a perturbation in the deep layer of the medium.

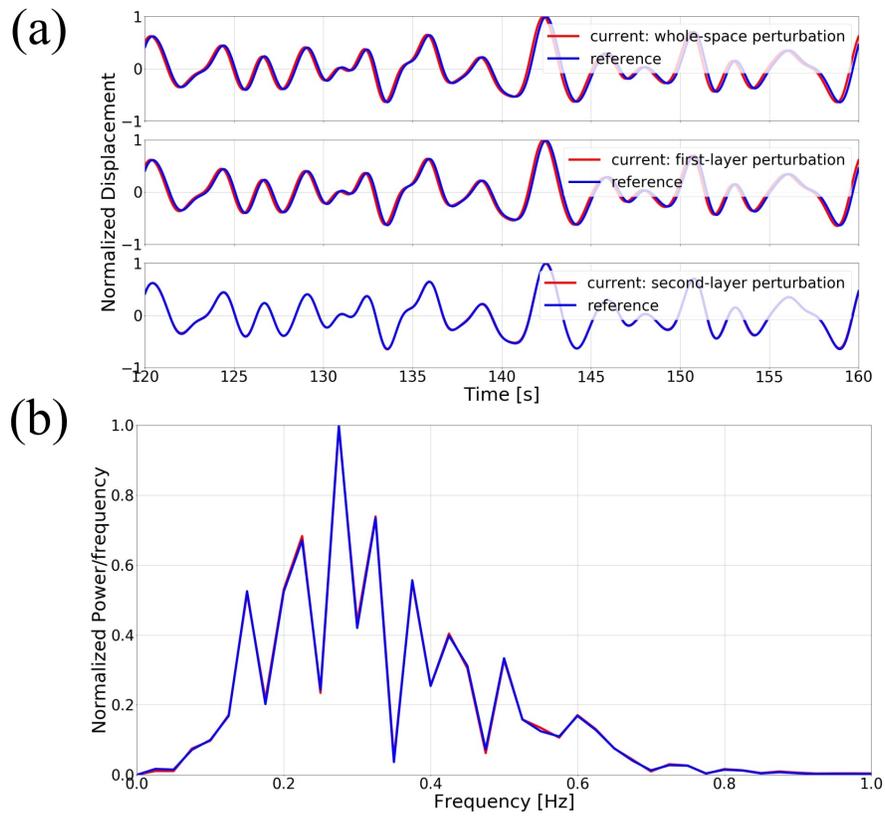


Figure A.7: Waveforms from the layer-over-halfspace heterogeneous medium with distant-offset source-receiver geometry to support Fig.2.6 of the main manuscript. (a) Reference waveform (blue) and perturbed waveform (red) of the whole space (top), the shallow layer (middle) and the deep layer (bottom). The correlation coefficients for three pairs of waveforms are 0.9635, 0.9692, and 0.9996, respectively. (b) Fourier amplitude spectra of (a) bottom panel.

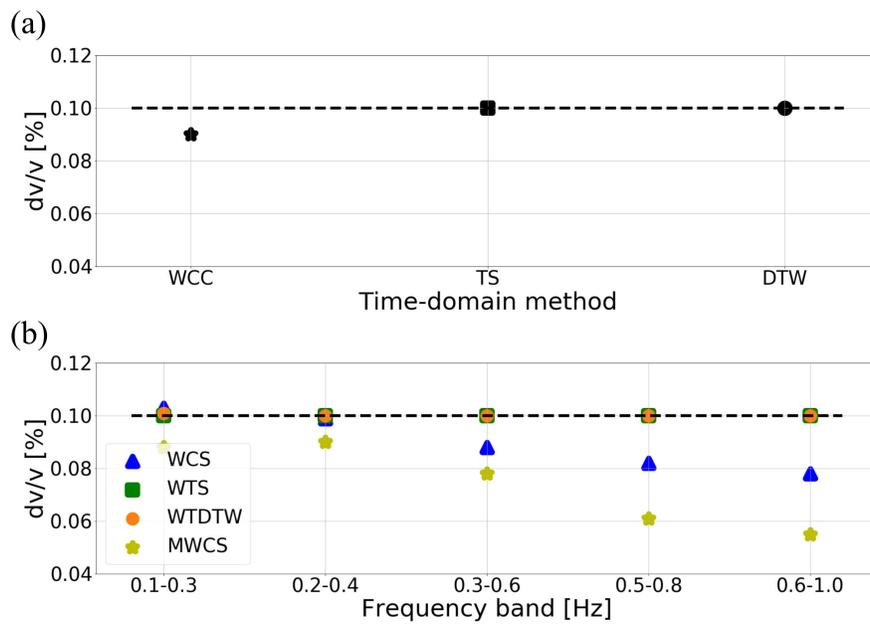


Figure A.8: dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.6 of the main text, distant-offset geometry, heterogeneous layer-over-halfspace medium, and uniform perturbation of the medium.

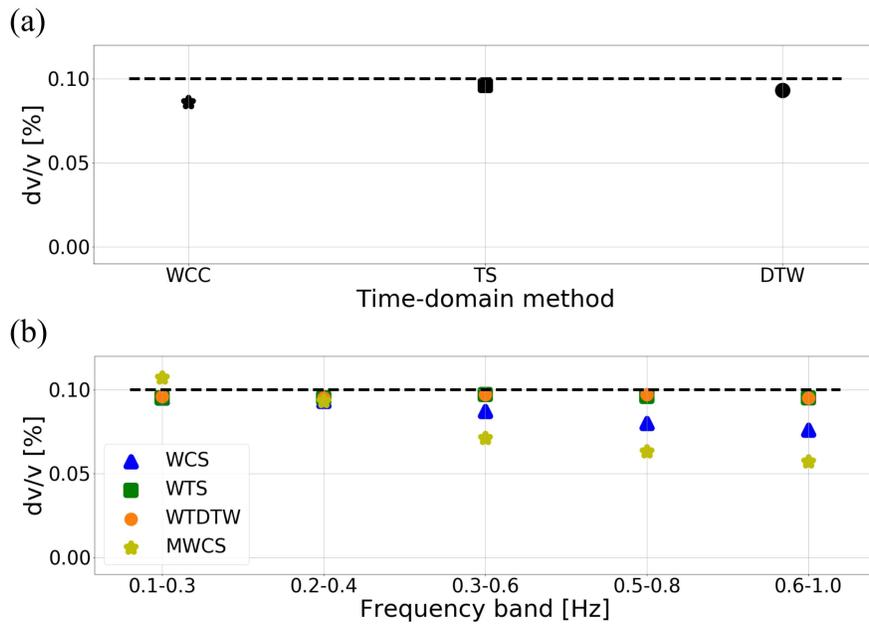


Figure A.9: dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.6 of the main text, distant-offset geometry, heterogeneous layer-over-halfspace medium, and perturbation in the shallow layer of the medium.

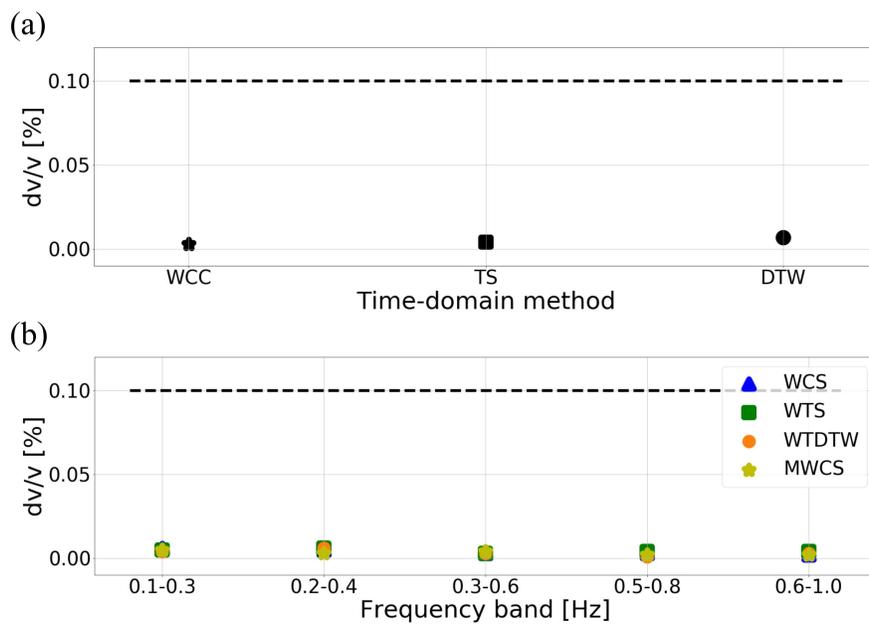


Figure A.10: dv/v estimates for time-domain (a) and frequency-band (b) methods, in support of Fig.2.6 of the main text, distant-offset geometry, heterogeneous layer-over-halfspace medium, and perturbation in the deep layer of the medium.

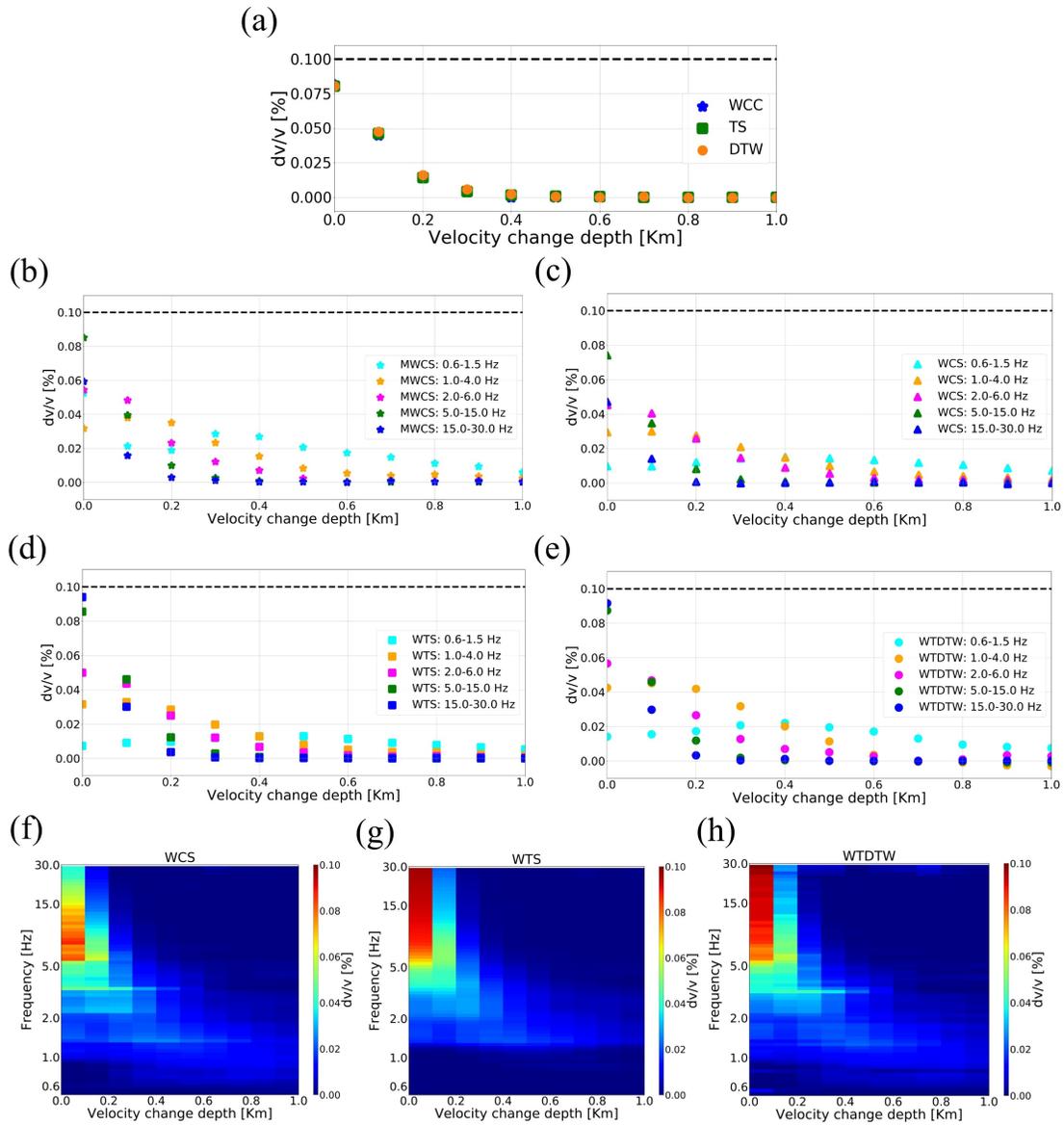


Figure A.11: dv/v measured for each depth of the perturbation layer in the zero-offset source-receiver geometry for all methods. (a) Time-domain WCC, TS, DTW, at various frequency bands (b) MWCS, (c) WCS, (d) WTS, (e) WTDTW, and over the spectrum of frequencies for (f) WCS, (g) WTS, and (h) WTDTW. Results are repeated in Fig.2.7(a, c) of the main manuscript.

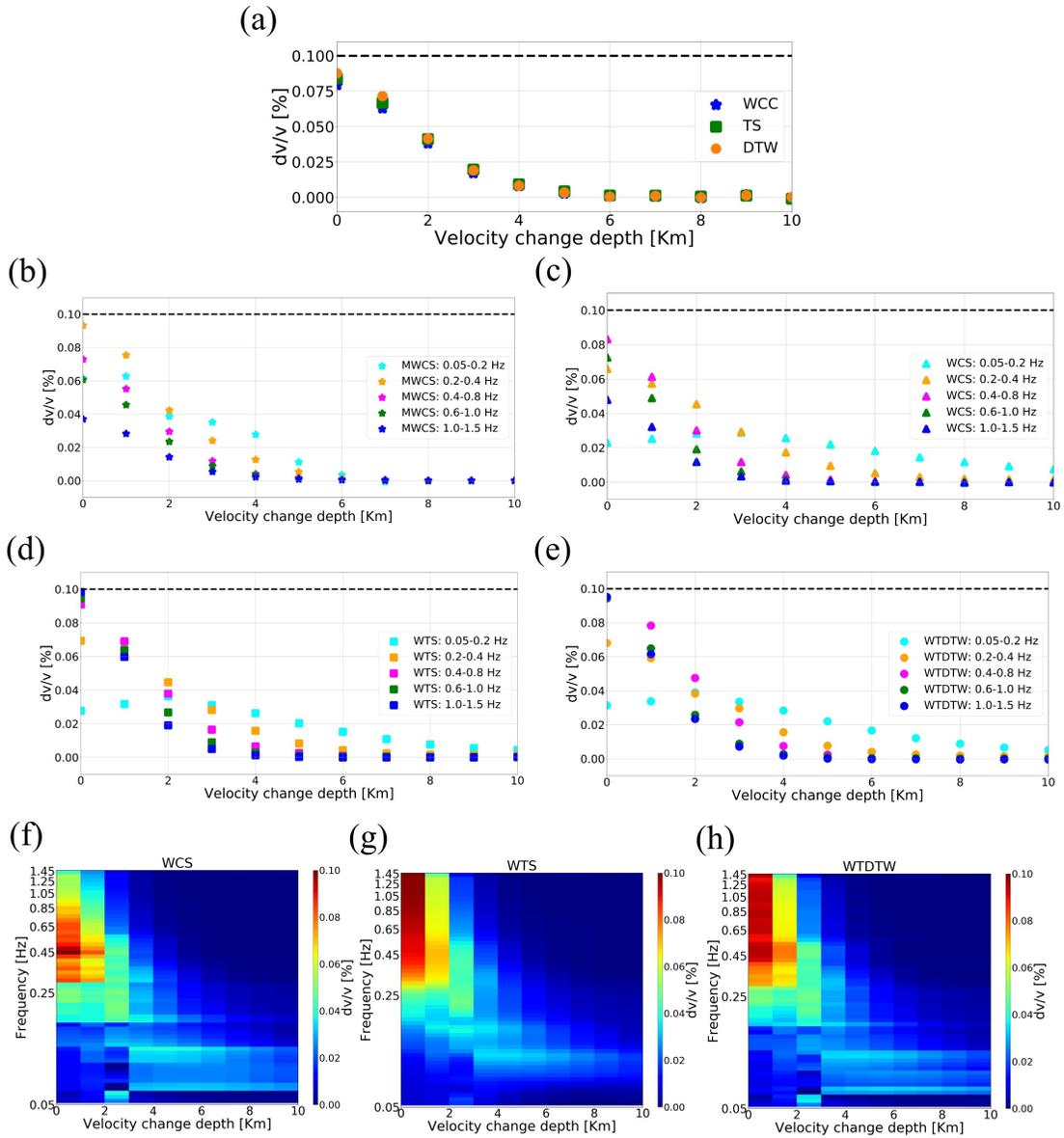


Figure A.12: dv/v measured for each depth of the perturbation layer in the distant-offset source-receiver geometry for all methods. (a) Time-domain WCC, TS, DTW, at various frequency bands (b) MWCS, (c) WCS, (d) WTS, (e) WTDTW, and over the spectrum of frequencies for (f) WCS, (g) WTS, and (h) WTDTW. Results are repeated in Fig.2.7(b, d) of the main manuscript.

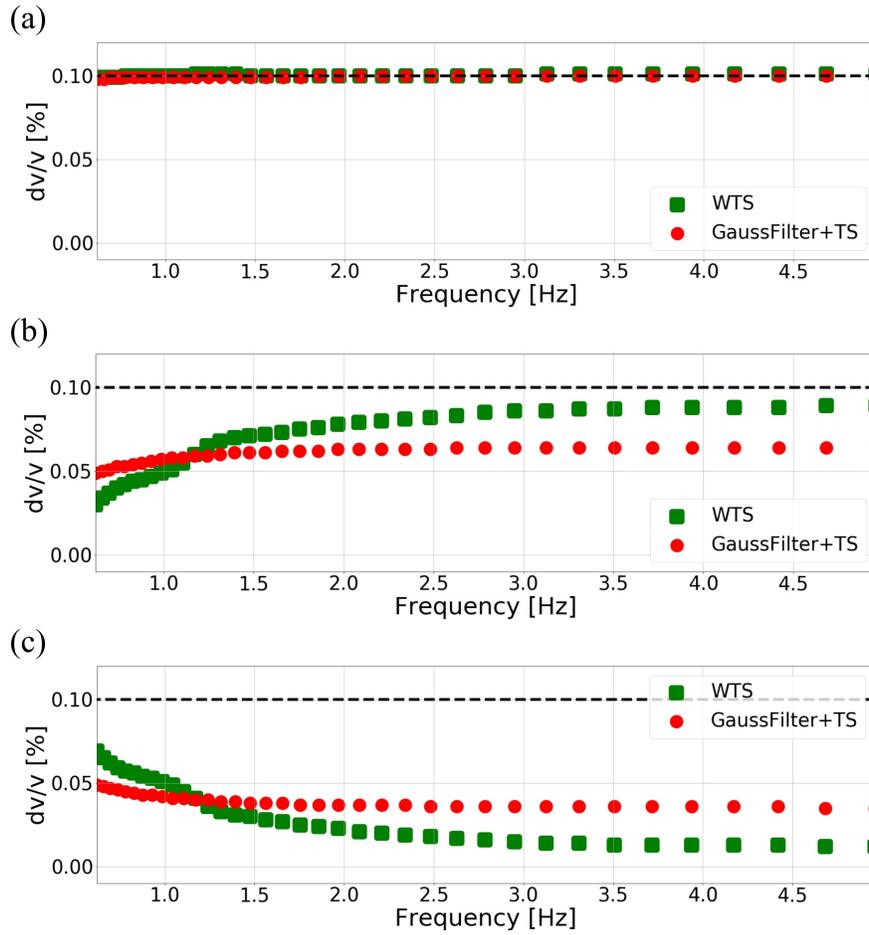


Figure A.13: Comparison between dv/v spectra obtained with WTS and narrow bandpassed TS (Gaussian filter), including (a) perturbation in the whole space, (b) perturbation only in the shallow layer, and (c) perturbation only in the deep layer as the layered half-space distant-offset case, which is shown as Fig.2.5 in Section 3.2. The medium and perturbation are detailed in Table B1. Briefly, the layer is perturbed at 10-km depth and by 0.1%. The black dashed line is a reference perturbation value rather than true perturbations at individual frequencies.

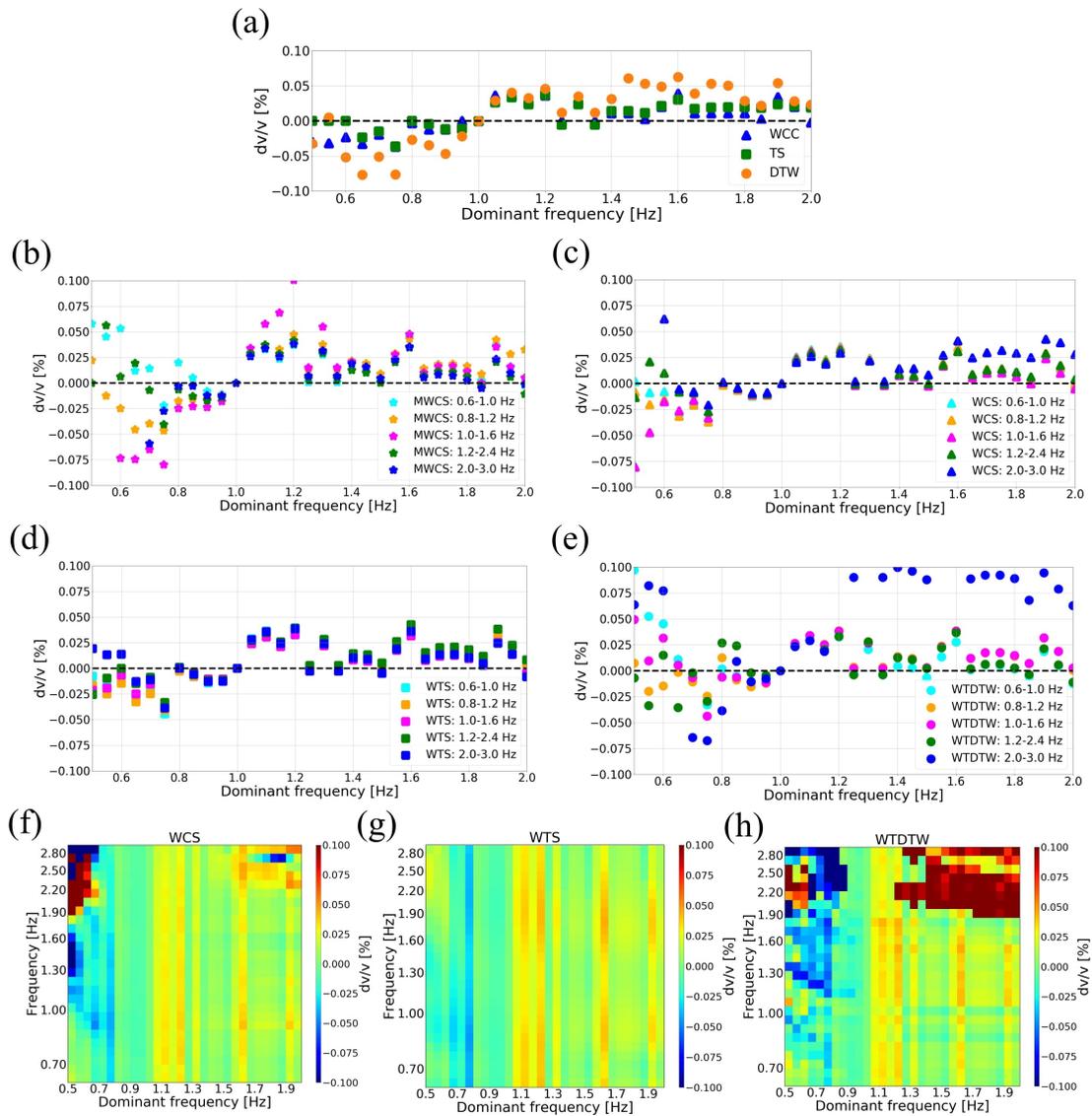


Figure A.14: Effect of dissimilar dominant frequency between the reference and current state on the estimates of dv/v . We perform simulations in the same medium as in Section 3.1 (scattering halfspace - uniform change in velocity - zero-offset response). The dominant frequency of the reference simulation is 1 Hz. The figure organization is similar to Fig.A.12. The black dashed line is the true velocity perturbation. We perform 31 simulations in the perturbed medium with various source spectra by changing the dominant frequency of current source spectra from 0.5 Hz to 2.0 Hz with an interval of 0.05 Hz.

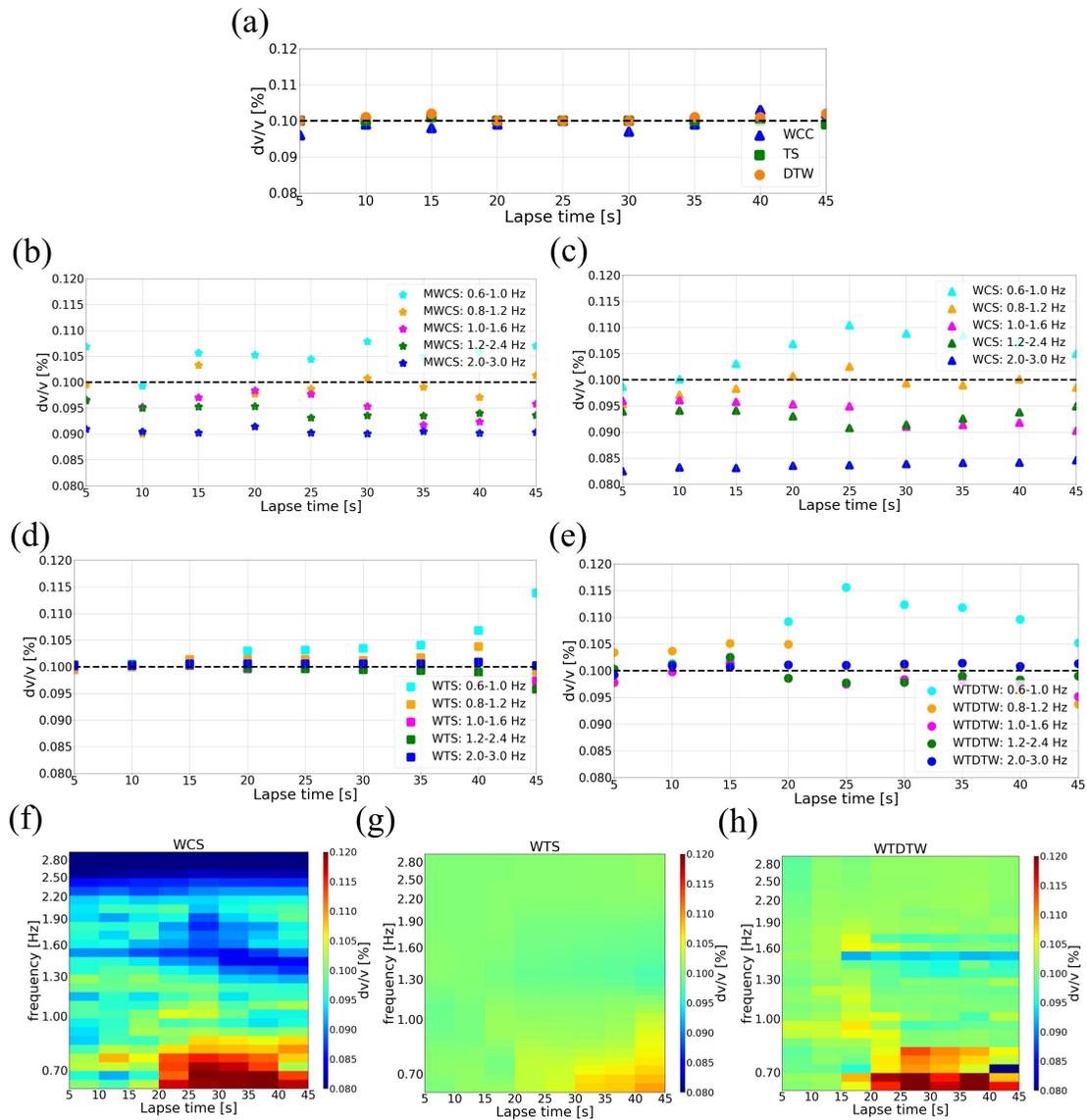


Figure A.15: Same as Fig.A.14 but varying the lapse time (start time of the window) of the coda waveform between 5 s and 45 s.

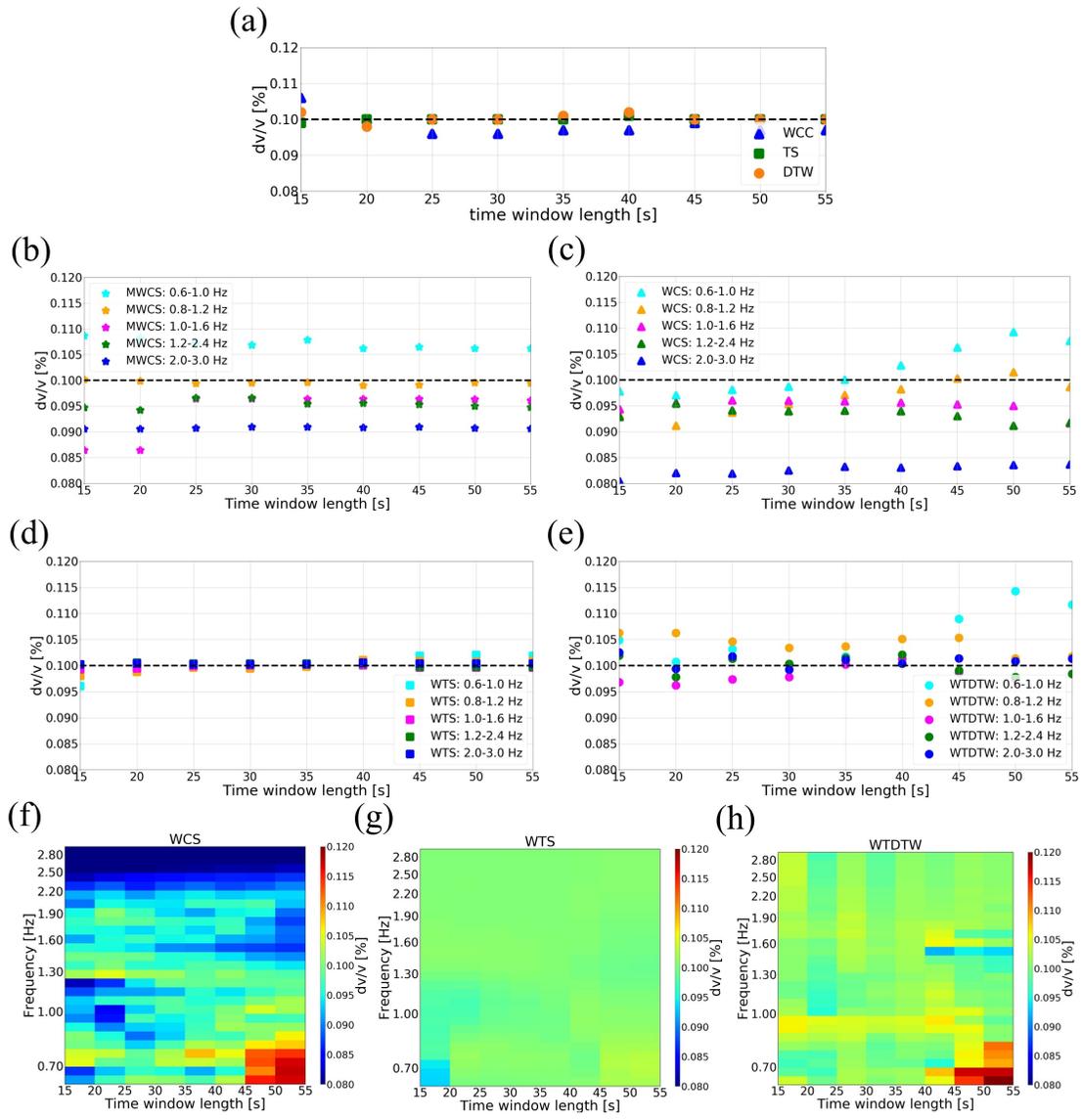


Figure A.16: Same as Fig.A.14 but varying the window length of the selected coda window between 15 s and 55 s.

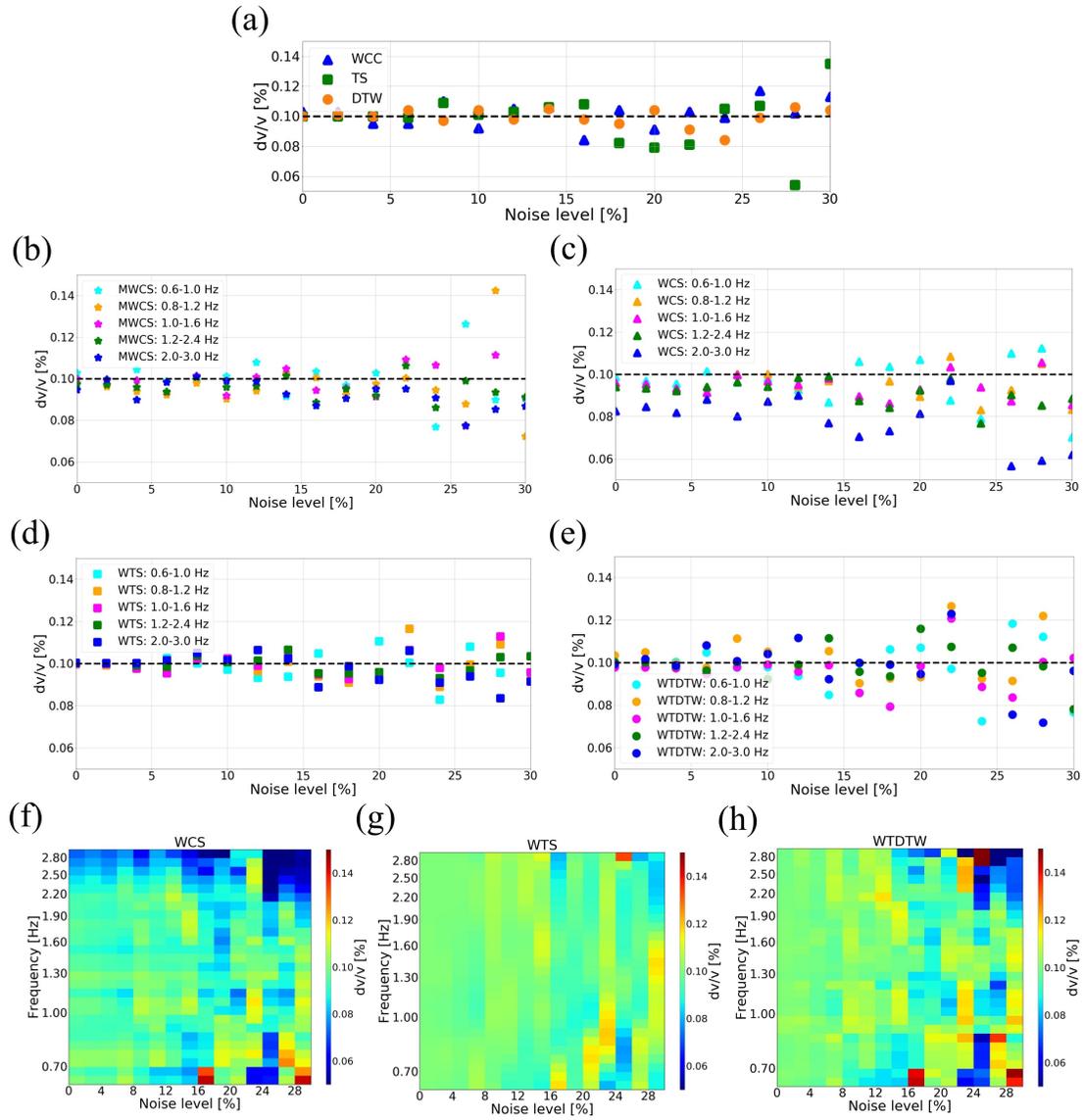


Figure A.17: Same as Fig.A.14 but varying the noise level. Gaussian noise is added on top of the simulated wavefields with levels varying between 0% and 30% of the maximum amplitude of the coda waveform.

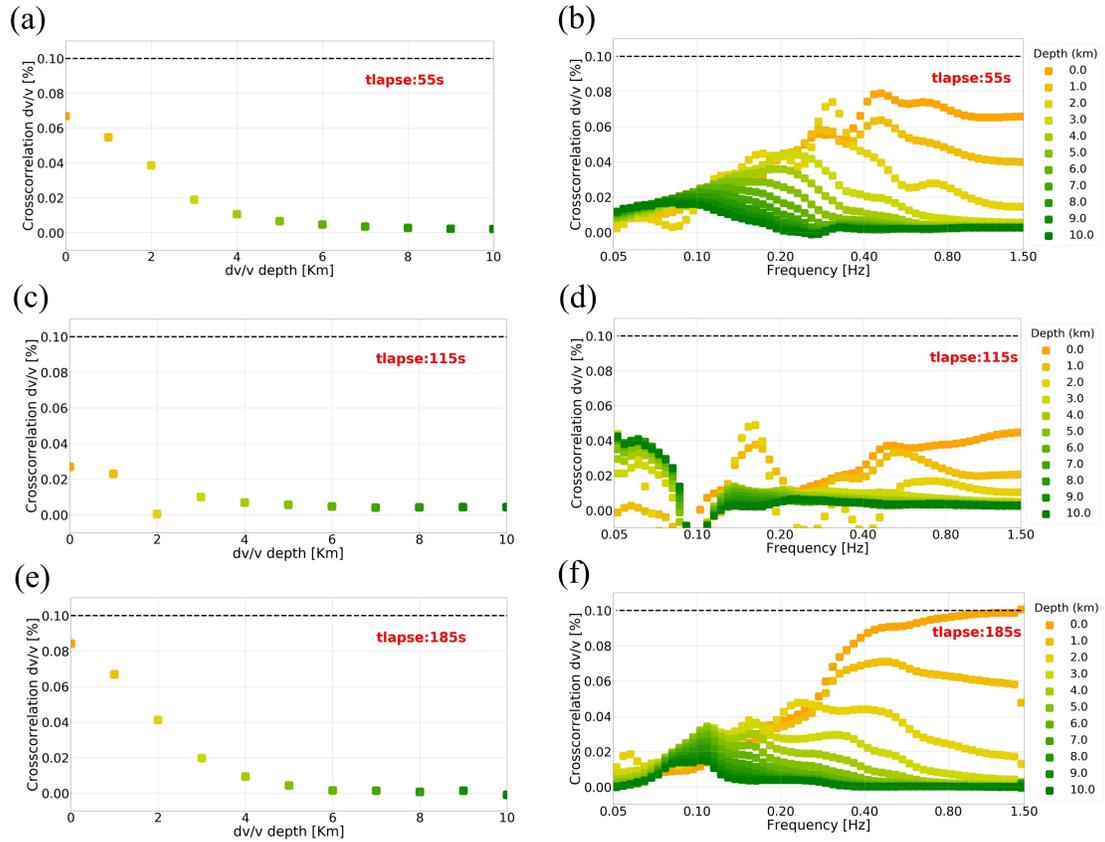


Figure A.18: Depth sensitivities of scattered body and surface waves at three lapse times (i.e. 55 s, 115 s, and 185 s) for the distant-offset setting. Left three panels (a, c, e) are the dv/v results from TS; right three panels (b, d, f) are the dv/v results from WTS. The experiment is the same to the distant-offset case of Section 4 but at different time lags. Note that the spurious results at frequencies around 0.1 Hz are attributed to low energy levels (spectral trough) at this frequency.

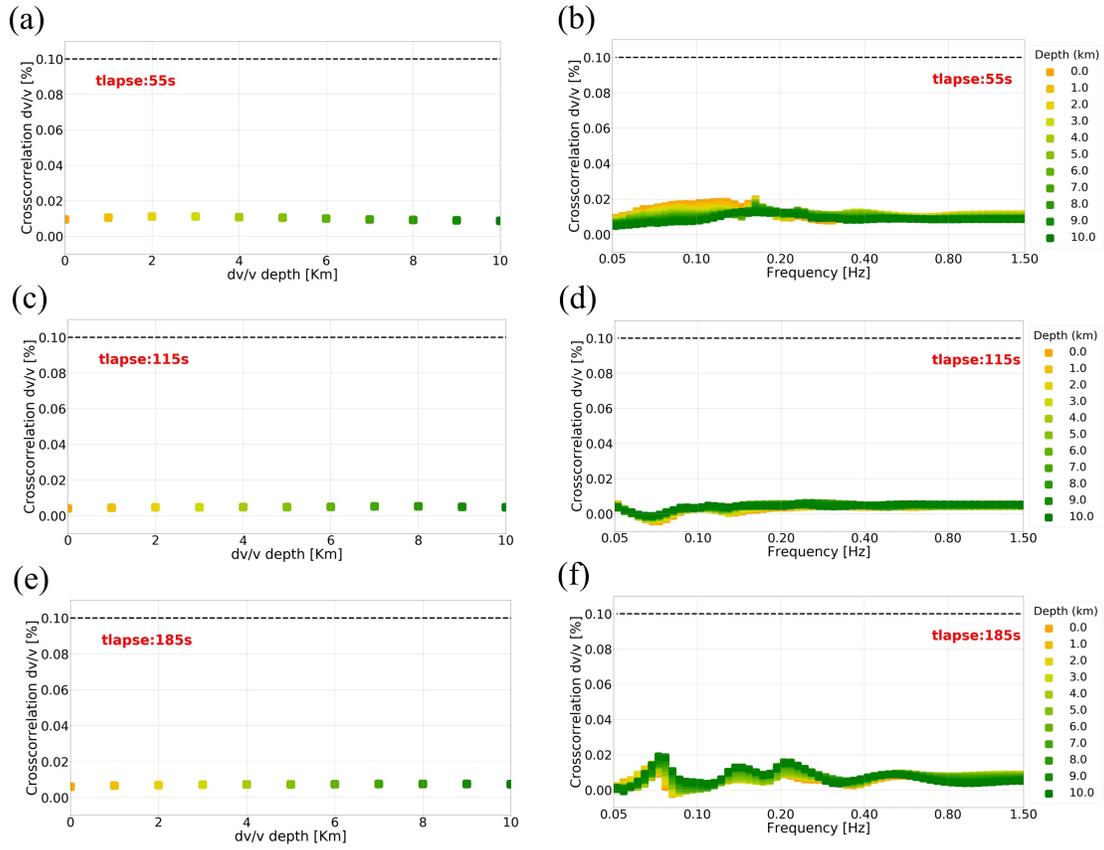


Figure A.19: Depth sensitivities of scattered body (bulk) waves at three lapse times (i.e. 55 s, 115 s, and 185 s) for the distant-offset setting. Left three panels (a, c, e) are the dv/v results from TS; right three panels (b, d, f) are the dv/v results from WTS. The experiment is the same as Fig.A.18 but replacing the free surface with an absorbing boundary in order to only simulate scattered body waves.

B

B.1 SUPPLEMENTARY FIGURES - ADDITIONAL FIGURES FOR CHAPTER 3

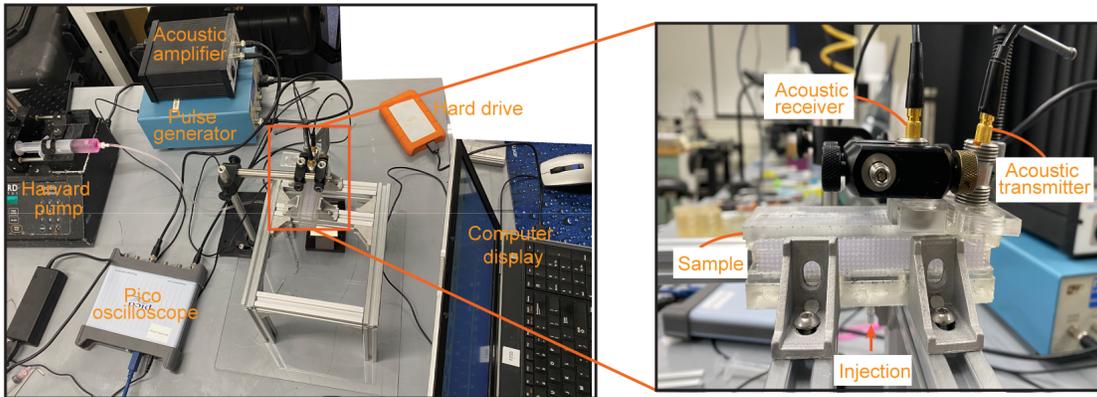


Figure B.1: Experimental apparatus and specimen.

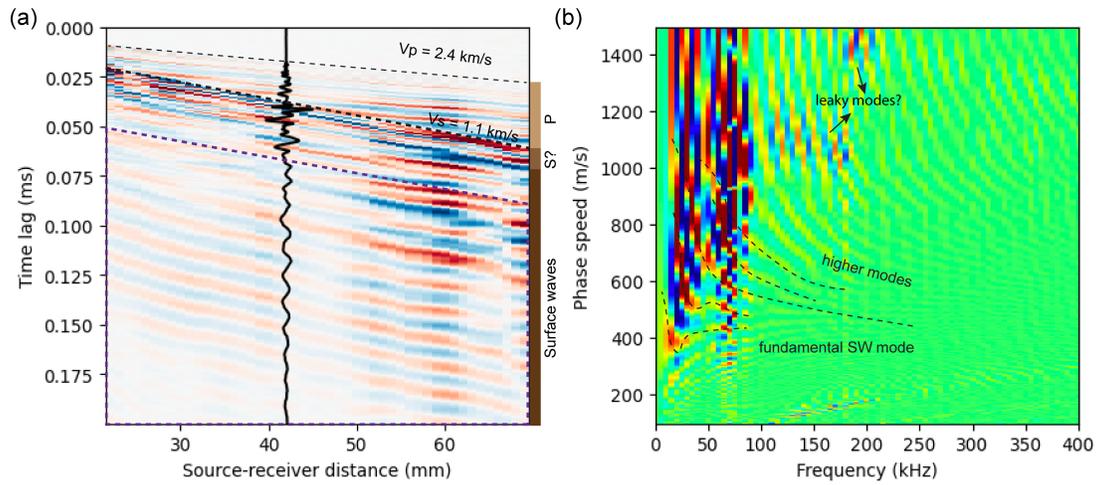


Figure B.2: (a) Calibration of wavespeeds of P- and S-waves in the hard resin. (b) Dispersion analysis of surface waves (outlined by purple lines in (a)).

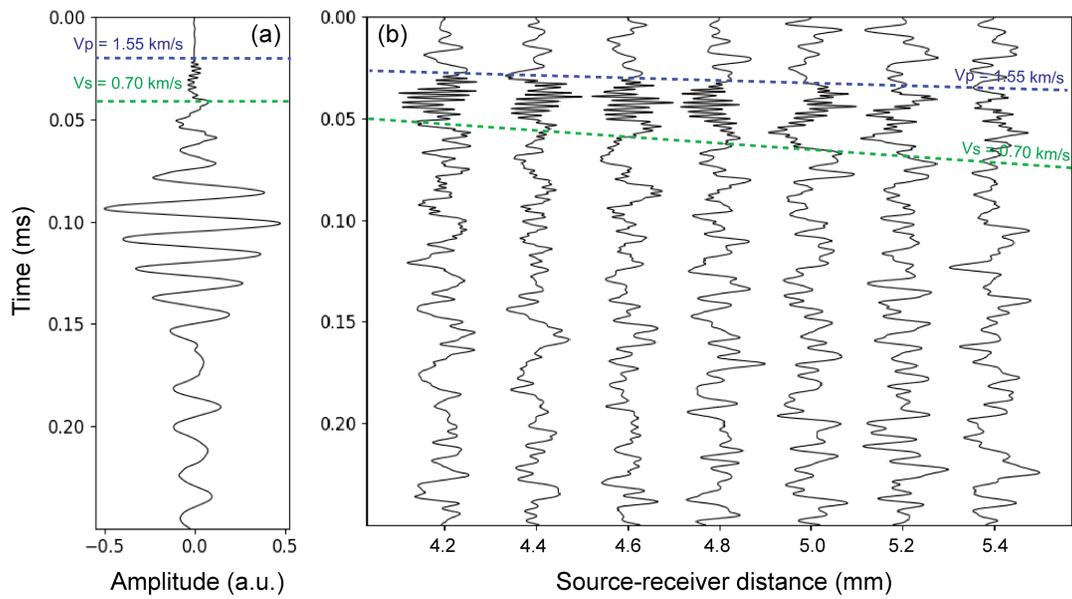


Figure B.3: Acoustic emission waveforms collected in the soft resin. (a) An active measurement with a fixed distance of 30 mm, (b) a shot gather over different distances.

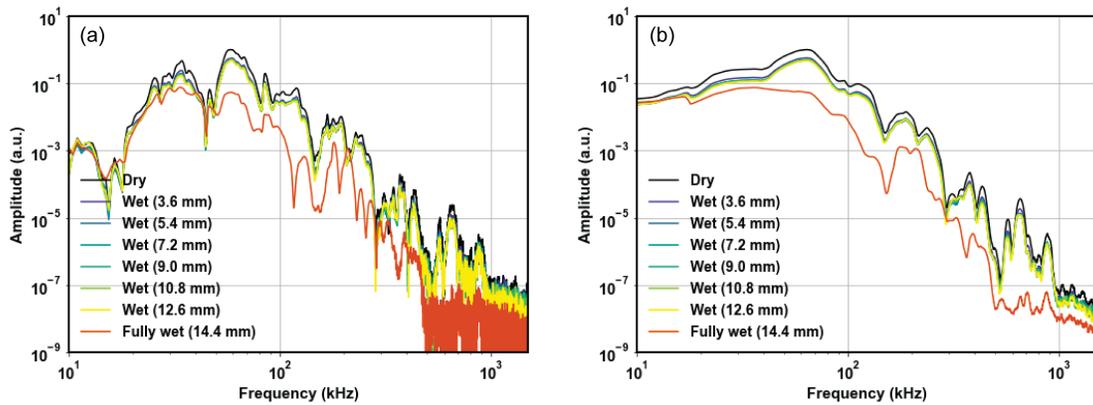


Figure B.4: (a) Spectra of AE waveform for the first case with only water-level changes within a static granular medium. (b) smoothed spectra of the same waveforms.

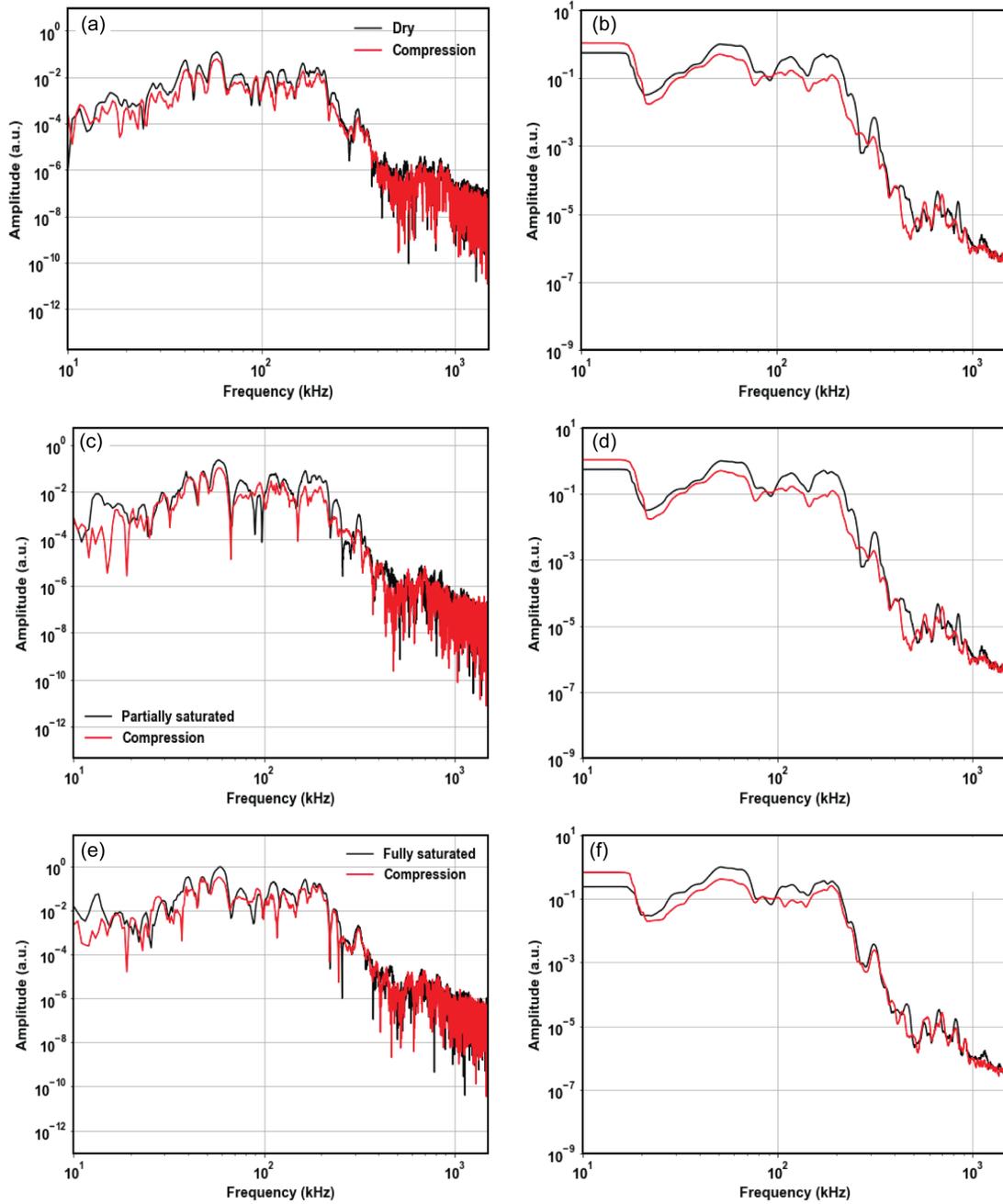


Figure B.5: (a) Spectra of AE waveform for the cases with stress loading on the well-bonded specimen under various saturation states: un-, partially-, and fully-saturation. (b) smoothed version of same waveforms.

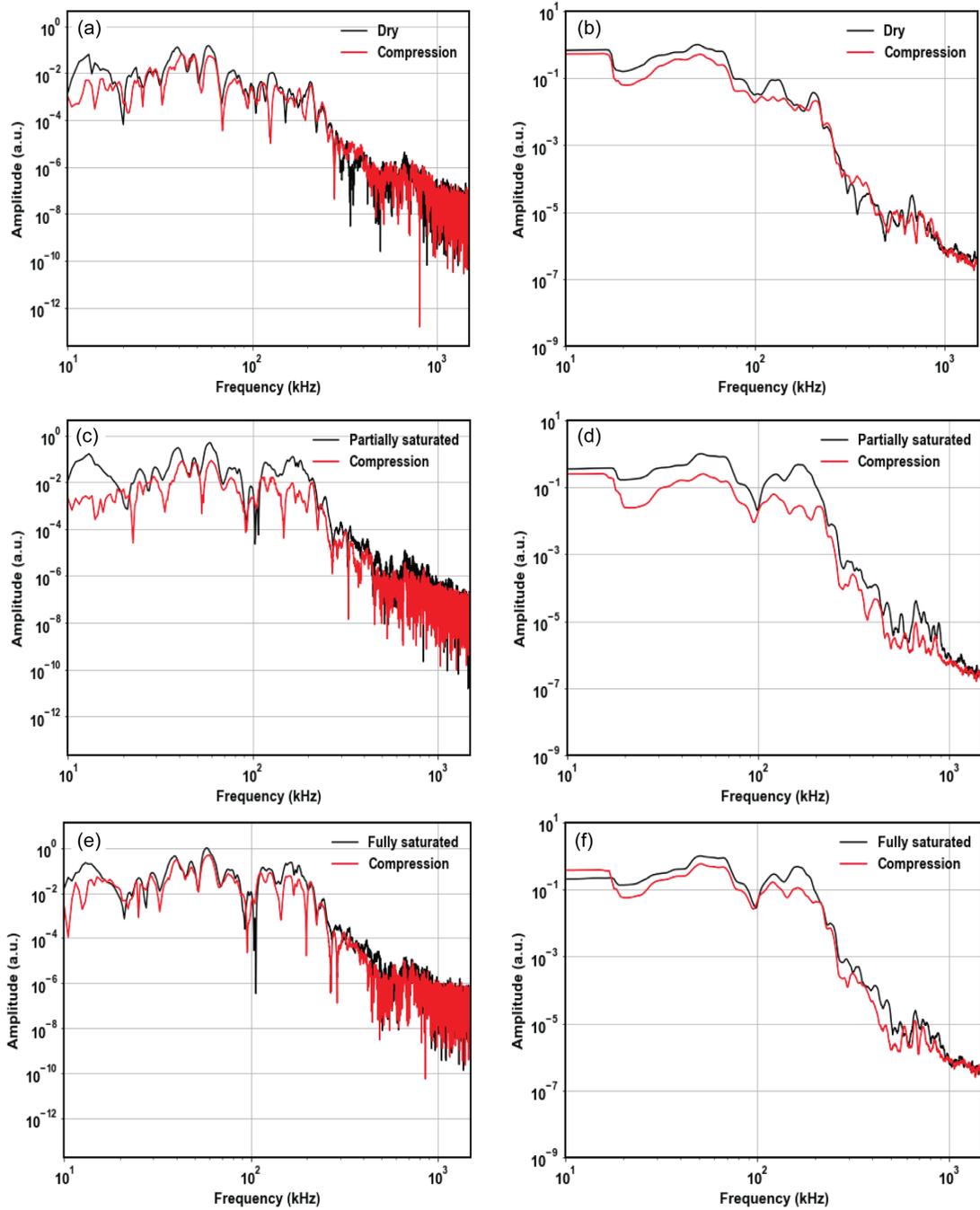


Figure B.6: (a) Spectra of AE waveform for the cases with stress loading on the unbonded specimen under various saturation states: un-, partially-, and fully-saturation. (b) smoothed version of same waveforms.

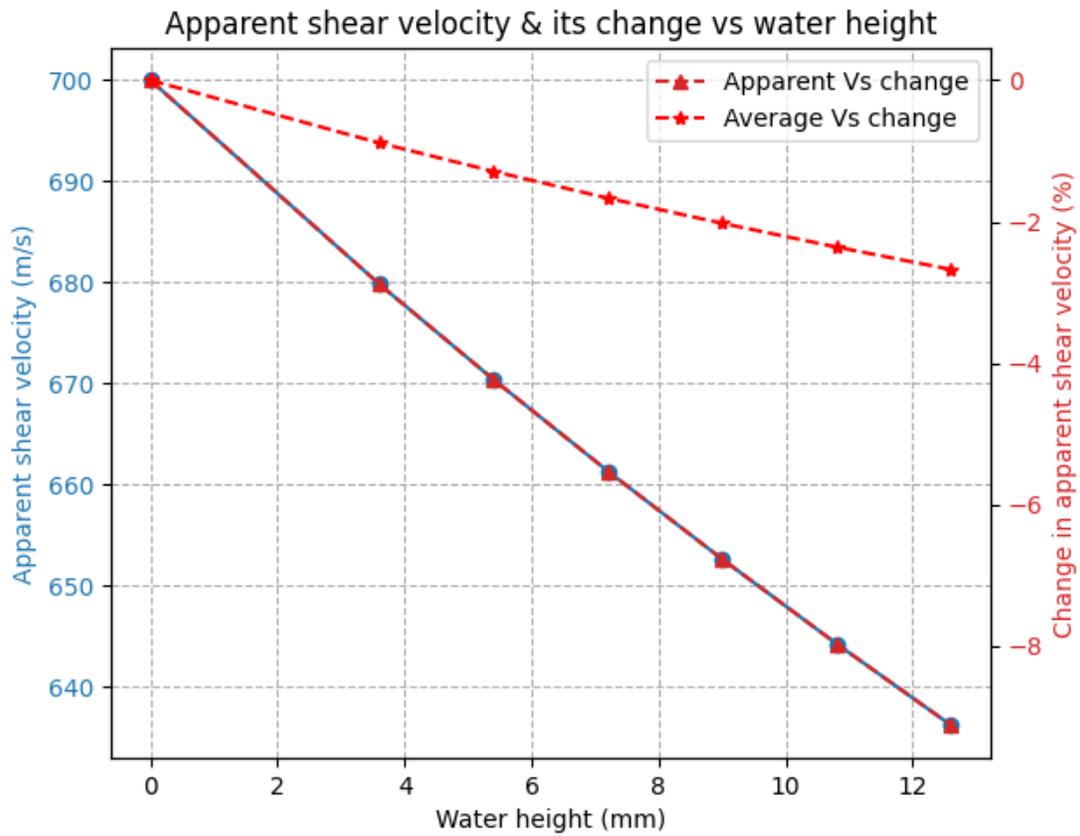


Figure B.7: Apparent shear-wave velocity (V_s) and its change for individual water heights, which correspond to the lab measurements as observed in Fig.3.1. Apparent V_s change is directly calculated from the perturbation of apparent V_s compared to the dry case. To determine the average V_s change for the whole space, we adjust the apparent V_s change by a spatial averaging factor, which are obtained from the volume and velocity of soft and rigid components, as Eq.3.4.

C

C.1 SUPPLEMENTARY FIGURES - ADDITIONAL FIGURES FOR CHAPTER 4

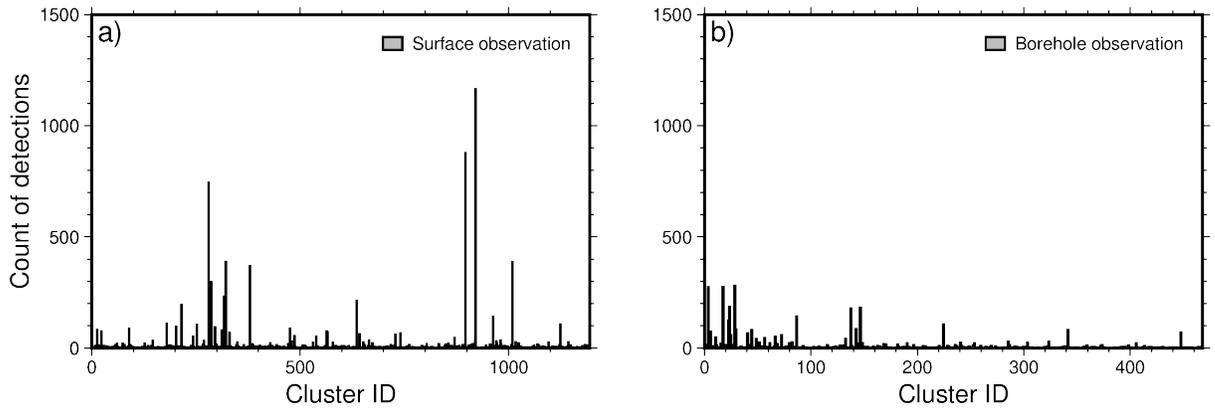


Figure C.1: Histograms in amount of earthquake detections by (a) borehole stations and (b) surface stations.

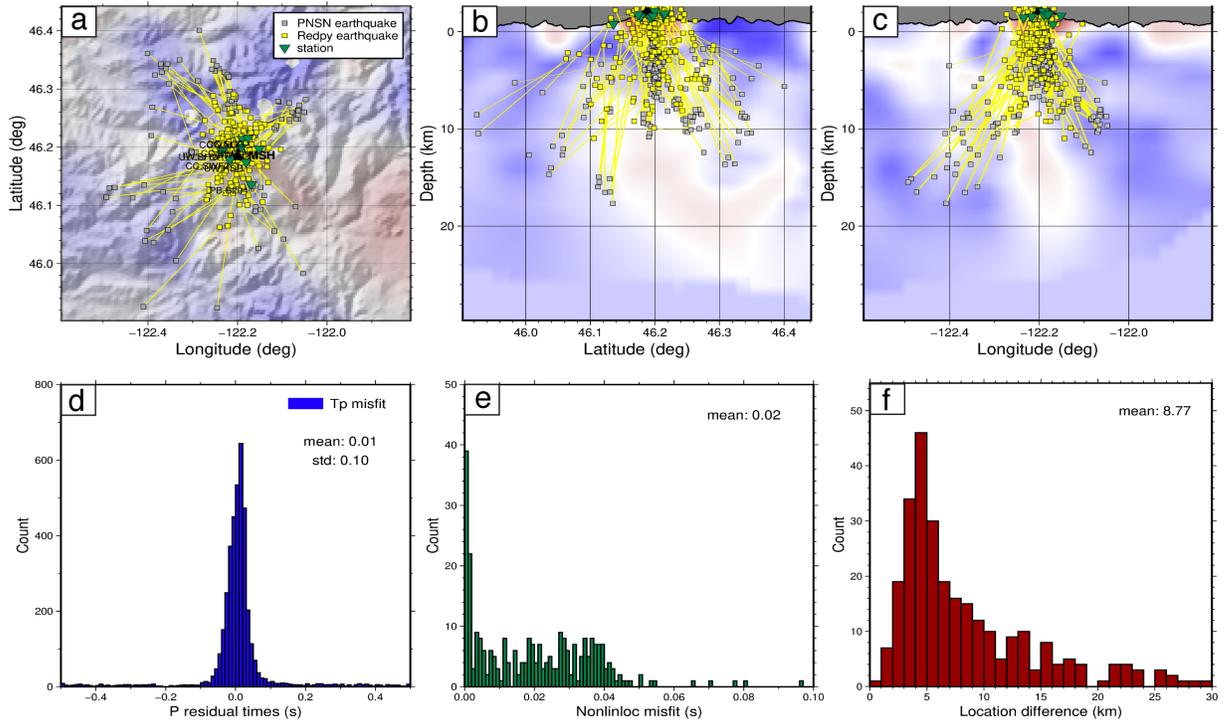


Figure C.2: (a-c) Comparison between NonLinLoc earthquake locations (yellow squares) in 3D models and PNSN catalogued earthquake locations (gray squares) in 1D. (d) Histograms of P-phase arrival-time residuals between machine learning picks and manual picks. (e) Histograms of NonLinLoc traveltime misfits of earthquake locations. (f) Histograms of earthquake location difference between NonLinLoc and PNSN.

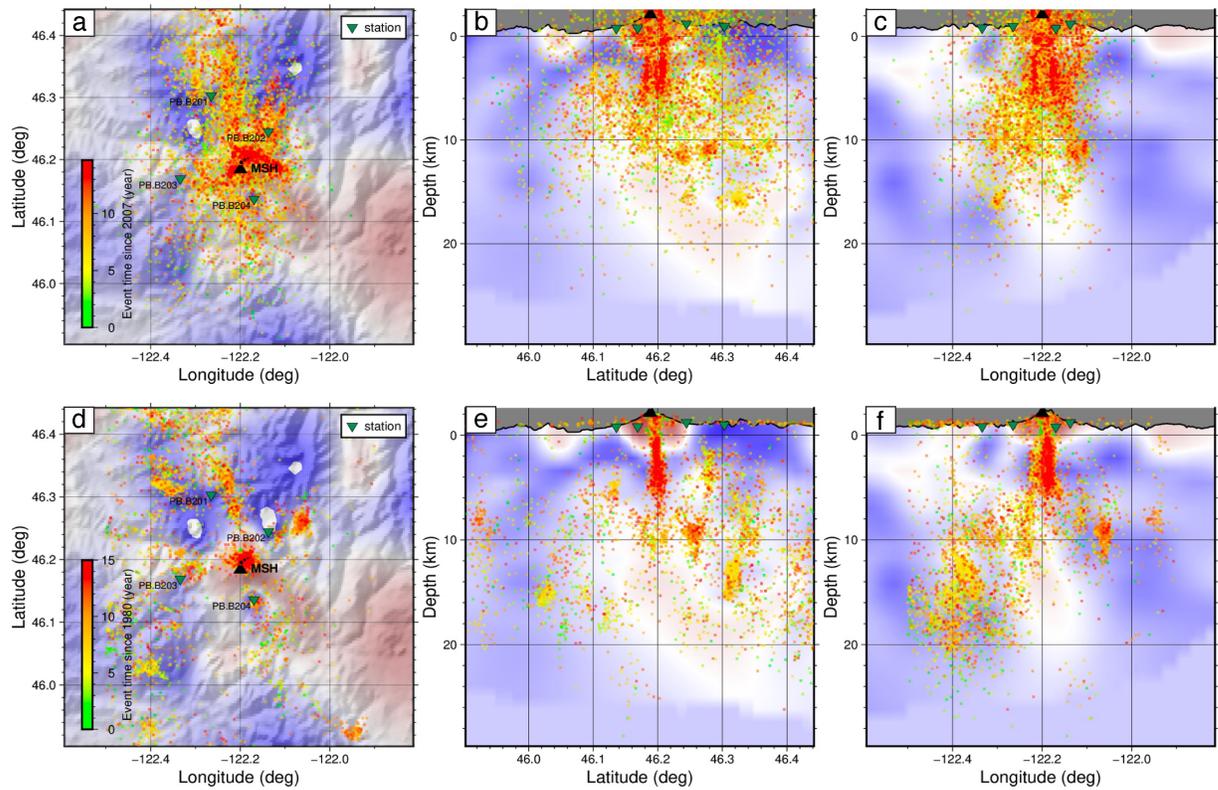


Figure C.3: (a-c) NonLinLoc locations of around 25,000 earthquakes detected by ELEP between 2007-2022. (d-f) PNSN catalogued locations of around 17,500 earthquakes since 1980, among them approximately 10,000 earthquakes were detected between 2004-2007 and around 7,500 earthquakes occurred between 2007-2022. Note: NonLinLoc locations are performed in 3D velocity models, the PNSN locations are obtained by using HypoInverse in 1D velocity models. ELEP detections are three times more than PNSN catalog.

D

D.1 RELATIONSHIP BETWEEN RADIATED ENERGY AND FRACTURE AREA

Because natural systems tend to fail more in shear mode, we compare the energy budget between mode I and mode II fractures. The potential energy (E) associated with the crack formation is consumed as fracture or surface energy (E_f) and radiated elastic energy (E_r) (heat energy loss is ignored)². Thus, the E is approximately a summation of E_f and E_r , as:

$$E \approx E_f + E_r, \tag{D.1}$$

We consider how this depends on the fracture speed (v) due to crack dynamics[?], writing this equation in terms of the strain energy density rates, $G(v)$, and fracture area $A(v)$, in which:

$$E \approx G(v)A(v) \approx [\Gamma(v) + \chi(v)]A(v), \quad (D.2)$$

$\Gamma(v)$ and $\chi(v)$ are respectively the fracture energy and radiated energy release rate, both of them replies on the fracture speed. G is related to the stress intensity factor K and the young's modulus γ . Hence, the potential energy E may be written further as K^2/γ , which is a universal representative for both crack mode-I and mode-II. In our experiments, the fracture area varies nearly linearly with the rupture duration, which has small fluctuations that may be associated with the speed of fracturing propagation. The $\Gamma(v)$ and $\chi(v)$ are minimally influenced by varying moments and almost keep constant. Similar explanation can be derived with a Brownian walk model for seismic moment of slow earthquakes^{??}. Accordingly, the radiated energy release scales with the fracture area, which explains our observations.

D.1.1 ALTERNATIVE STOCHASTIC MODEL

The near-linear variation of the radiated energy with the area was also found under the assumption of a randomly varying moment rate function of tremor[?]. For weak dependence between tremor slip and tremor length, a near-linear relation was also modeled[?].

D.1.2 SCALING RELATION WITH SEISMIC MOMENT

? and others showed that for elongated rupture, or rupture with high aspect ratios, the scaling of $Er M0$ could be explained by the scaling $Er slip * W * L$, where W is now fixed, and slip and area (WL) allowed to vary. Assuming that $Er Area (3/2)$ is $M0 a^3$, where slip $WL = a$ and area $LW a^2$. So for skinny rupture, $Er a^2$, where W is constant. We find that the observations of the

Cascadia tremor catalog and area suggests a scaling of $Er \text{ Area}^{(1.16 \pm 0.3)}$, which is slightly lower than that of regular earthquakes, but similar to skinny rupture. The laboratory experiments suggest a nearly linear scaling, and there again, each individual AE is a skinny rupture.

An alternative consideration is that of the definition of moment for tensile cracks. In particular, we consider the work from ? that writes equations to relate seismic moment and tensile crack dimension (a) and fluid pressure (P), such as $M0 \propto Pa^3$. This is similar to the shear moment $M0 \propto \Delta\sigma a^3$, where $\Delta\sigma$ is shear stress change. It is possible that a slight deviation from the regular scaling ($3/2$) can be accommodated by a scaling between fluid pressure change (P) and tremor radiated energy. Non-self similar behavior is not uncommon in recent observations of shear regular earthquakes (e.g., ? and many others).

D.2 AE ENERGY CALIBRATION

We validate accurate measurement of the radiated energy from the AE sensors data by comparing the radiated and the elastic strain energies from multiple drop ball tests (see Fig.D.11). The validation follows a four-step process as follows:

First, the potential energy, E_p , must equal the incoming kinetic energy, E_{ki} . This is confirmed by the consistency between the theoretical velocity from potential energy and observed incoming velocity from a high-speed camera (Fig.D.11), as the black line in Fig. D.11a. Both air resistance and rotational energy are negligible. Second, the coefficient of restitution, R_e , (ratio between the outgoing velocity of the bouncing ball, v_0 , and the incoming velocity of the falling ball, v_i) is the parameter that describes the potential to kinetic energy loss at each bounce. If it is a perfectly elastic collision or bounce, then $R_e = 1$, which is almost impossible. In our case, R_e is not linearly decreasing with the height, as shown by the red line in Fig.D.12a.

Third, we define the energy transferred to the material as the kinetic energy transfer, ΔE_{ki} , or the

difference between the ball's incoming and outgoing kinetic energies, E_{ki} and E_{k0} . By ignoring air resistance, rotational energy, plastic energy, and heat, the sample consumes the energy as the elastic strain energy, which produces radiated energy and is later received by the sensor. The relation, $\Delta k_{ki} = E_{ki}(1 - R_c^2) = E_p(1 - R_c^2)$, can be seen in Fig.D.12b, and will be utilized to convert the potential energy and into kinetic energy transfer, ΔE_{ki} , by multiplying it by $(1 - R_c^2)$.

Fourth, we use this relation to calibrate another group of drop ball tests, which record the AE signals. We first compare the AE energy calculated from the integration of squared voltages over time to the potential energy in Fig. D.12c. We then calculate the kinetic energy transfer ΔE_{ki} from the potential energy, E_p , by using the previously calibrated relation in Fig.D.12d. We see AE radiated energy is almost proportional to ΔE_{ki} .

We conclude that the AE energy is a good metric for ΔE_{ki} , which is mainly elastic strain energy. We also compare the AE energy calculated from the integration of squared velocities over time to further confirm the radiated energy ends up with the kinetic energy of particle motions in the sample or sample surface. The elastic strain energy is released and propagates as the radiated energy.

Movie S1. Animated demonstration of observations on the high-viscosity ($800cP$) fluid-induced fracturing experiment. Raw observations include the optical images (upper-left panel) captured by a high-speed ($fps = 100,000$) camera and acoustic-emission (AE) signals (lower-right panel) recorded by four AE sensors. Other measurements are also derived, such as the time-evolving fracture fronts (upper-right panel) in four directions (SE, NE, NW, and SW), which are extracted from individual frames; the image difference (background of the lower-left panel) relative to the first frame and thus clear expanding of fracture and fluid; the derivative image (overlaid red image in the lower-left panel) which indicates the opening cracks; the intensity of each differential image is averaged in space and normalized in time and displayed as the red line at the bottom of the lower-right panel (Note: the intensity of the differential image also means the intensity rate of raw images). The overall view of all observations clearly demonstrates the fracture generation process and produced mechanical vibrations. [[Link to video](#)]

Movie S2. Same to Movie S1 but for the low-viscosity ($1cP$) fluid-induced fracturing experiment. [[Link to video](#)]

D.3 SUPPLEMENTARY FIGURES - ADDITIONAL FIGURES FOR CHAPTER 5

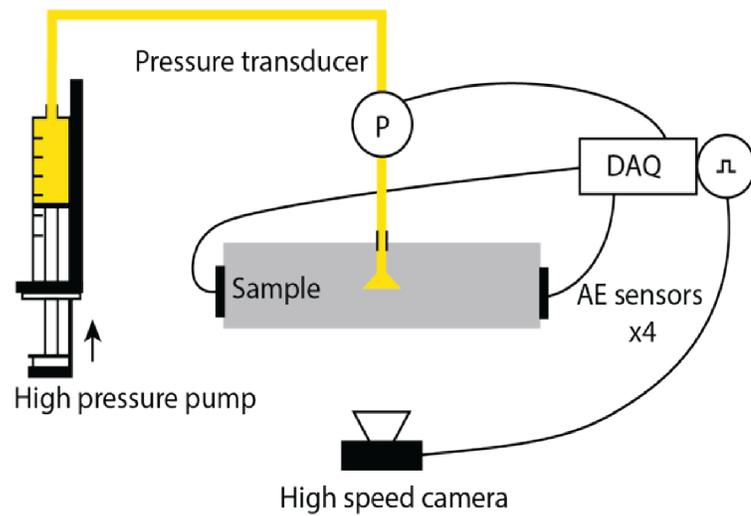


Figure D.1: Experimental apparatus. A high-pressure pump drives the fracturing fluid into the sample. A pressure transducer measures the fluid pressure as the experiment proceeds. AE sensors are placed on each of four sample quadrants and record the AE signal generated by the fractures. A high-speed camera is filming the growing fracture from the bottom of the sample, when triggered by the first AE signal. The details of the experimental apparatus components are shown in the photos in Fig.D.2.

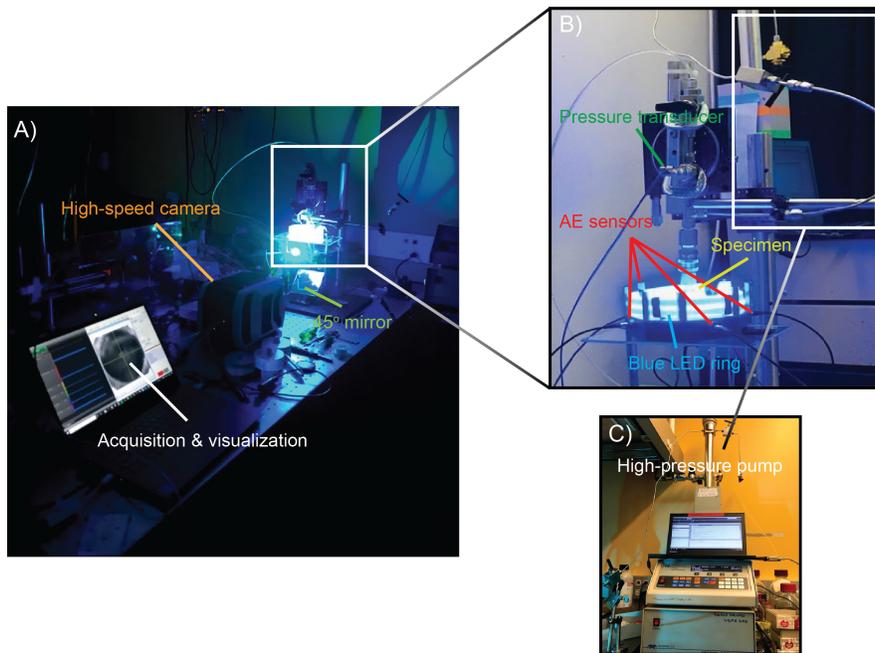


Figure D.2: Main components of the experimental apparatus. a) A high-speed camera (Phantom TMX5010) films the fracture propagation. A blue LED ring is placed around the sample to highlight the contrast between dyed fluids and fracture fronts. b) Four AE Glaser-type sensors (KRNB-PC) listen to the fracture propagation. c) A high-pressure pump (Teledyne Isco 65D) injects the fluid at a constant flow rate of $Q=0.3$ ml/min.

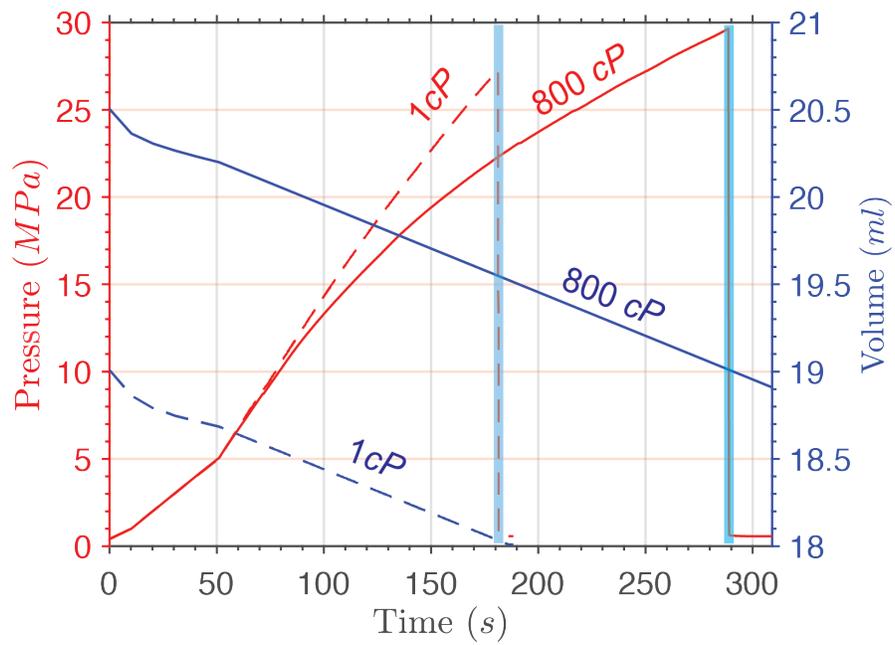


Figure D.3: Pressure and injected volume during the experiment. The fluid injected volume (blue curves) increases the fluid pressure (red curves) until the sample ruptures (shaded blue region). Both fluid viscosity experiments are shown: 1cP in dashed lines and 800 cP in solid lines.

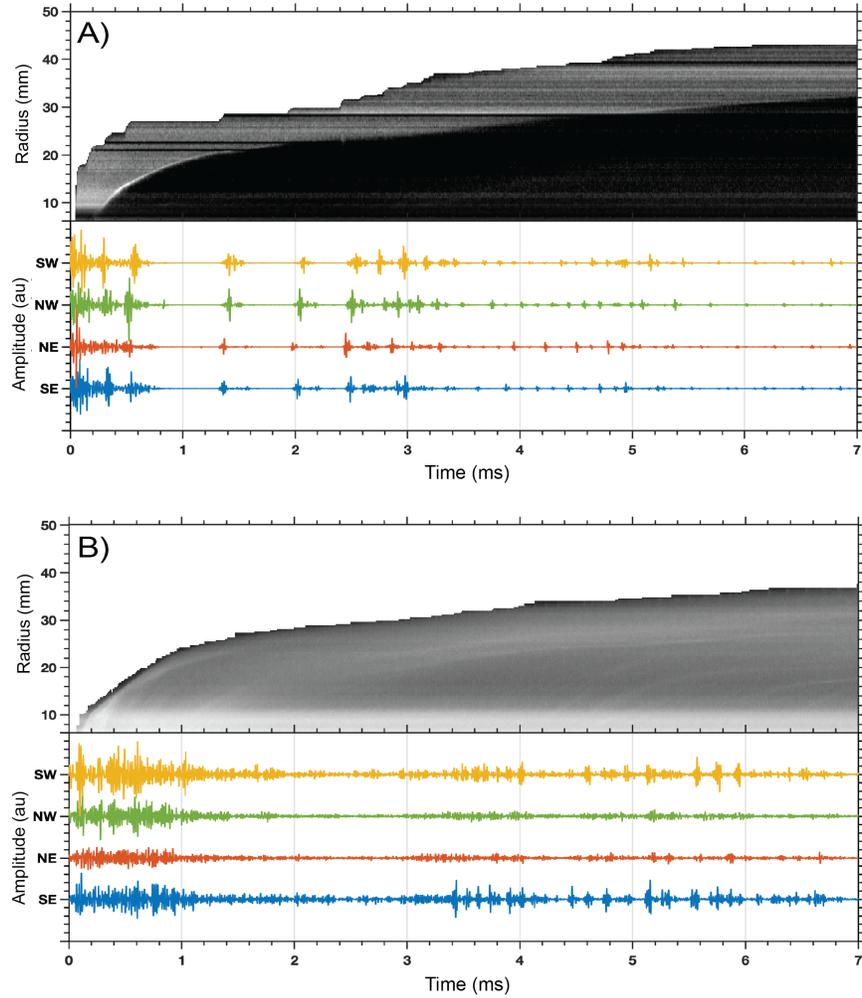


Figure D.4: Kymograph and AE signals. (a-top) kymograph of the fracture propagation during the entire experiment in the high-viscosity case. The stick-break events are clearly visible by the increase and stagnation in the fracture radius. (a-bottom) Overall, AE signals are recorded during the entire experiment, occurring as bursts and related to breaking events in the kymograph and the pauses related to the stick events. (b) similar to (a) but for the low-viscosity experiment. In this case, the stick-break events are not clearly resolvable in the kymograph. Still, the AE signal shows interesting bursts in their signals, indicating very small amplitude stick-break instability events. Note: partial data have been shown in Fig. 5.2.

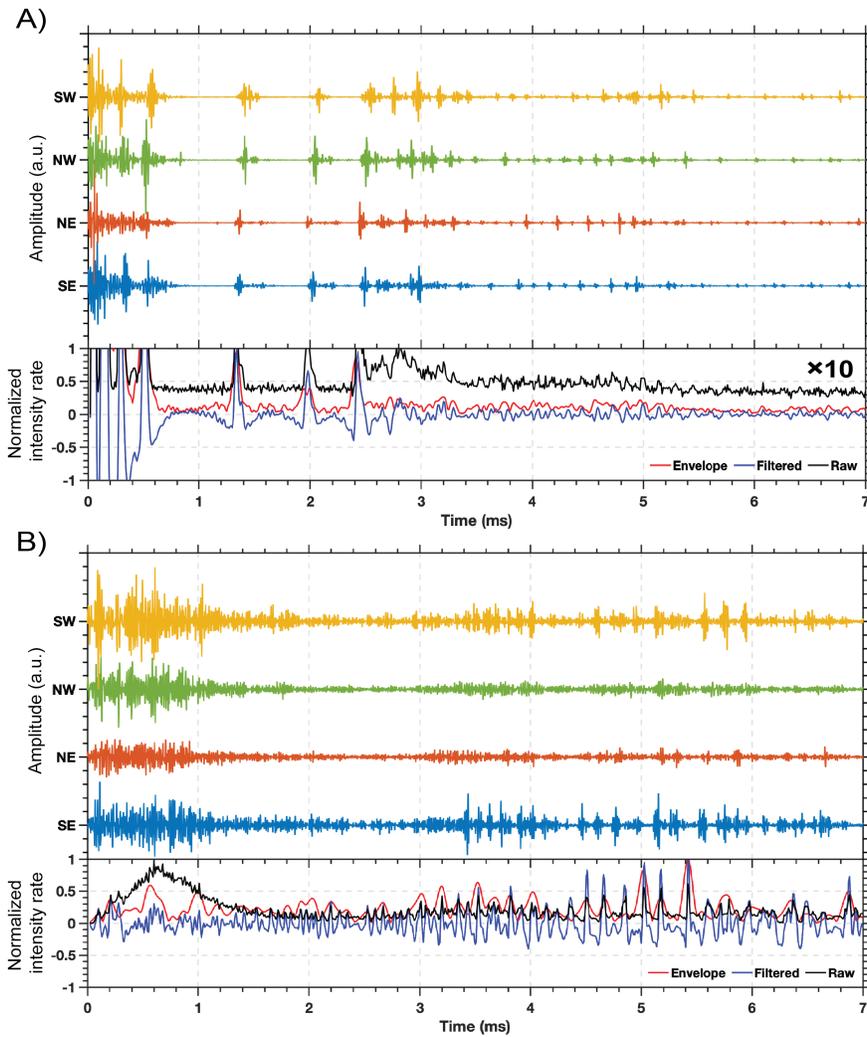


Figure D.5: Comparison between AE raw signals and intensity rates, which include raw (black), filtered (blue), and enveloped (red) ones for (a) high-viscosity and (b) low-viscosity experiments. Note: AE signals are exactly the same as the ones in Fig. D.4. Partial data have been shown in Fig. 5.2. We refer to the intensity measurements in the video description.

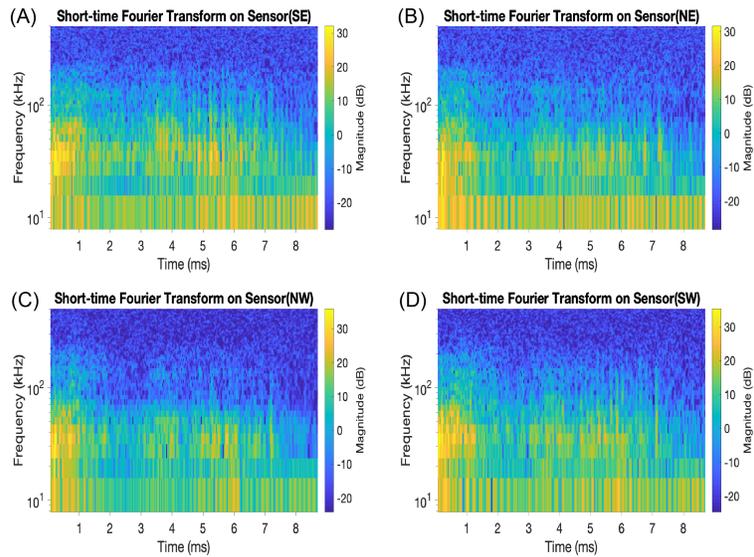


Figure D.6: Spectrograms of AE signals. We obtain them by performing the short-time Fourier transform on individual raw AE signals of the water experiment as the ones in Fig.D.4b. Note: the signals' unit is voltage and is nearly proportional to displacement for an almost flat instrumental response function and a constant signal amplification factor.

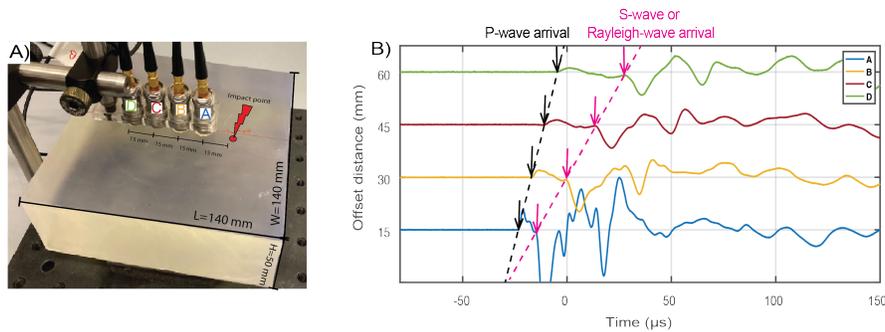


Figure D.7: Glass capillary test for Green's function calibration. (a) Calibration setup in which a glass capillary breaks at the impact point and generates an impact. Four AE sensors (A, B, C, and D) are deployed in a line at intervals of 15 mm from each other, recording the same single event resulting from the capillary break. The block used is made of the same 3D printed PMMA material as for the fracturing experiment and is dimensioned (140x140x50 mm) to avoid arrivals of any P-wave signal reflected from the boundary before the first arrival from the capillary break. (b) AE signals are recorded at the four sensors (successively from the impact point A, B, C, D) resulting from the capillary break. With wave types labeled, The signals are labeled with the arrivals of different waves, the measured P-wave speed is approximately 2300 m/s and estimated S- or Rayleigh wave speeds are both about 1000 m/s, because the Poisson's ratio of the material is estimated to 0.34 and as the distance between the source and successive receivers is very small compared to the overall wave speed.

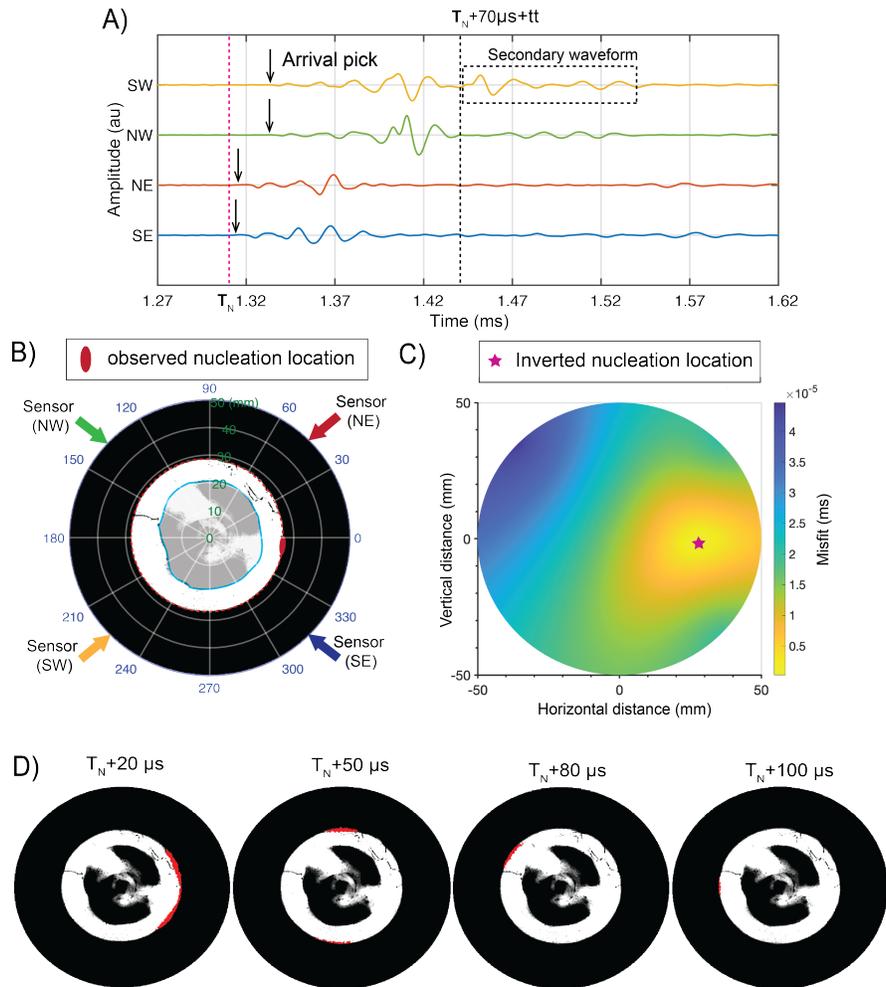


Figure D.8: Fracture event spatial location by AE triangulation. (a) Arrival times are visually picked and annotated with downward arrows on the AE signals. Nucleation begins at T_N , seen as the red dashed line. The arrival time difference is compared to the nucleation times, allowing us to triangulate and determine the spatial location of the nucleation event (b) Differential image, the locations of the sensors (thick colored arrows), fracture front (red dash line), fluid front (solid blue line), and the observed location of the fracture nucleation (red oval). (c) Misfit between theoretical and observed differential travel times between sensors given the radial and azimuthal distances in a polar coordinate. The warm and cold colors denote the small and large misfits, respectively. The red oval location represents the location with the minimum misfit. The location is almost consistent with the nucleation location observed in (b). The velocity calibrated in Fig. D.7 is used for creating a travel timetable for the location search. The misfit function used for the grid search method can be seen as Eq.5.2.2 in the main text. (d) Four snapshots of differential images exhibit the propagation and arrest stages during a single tangential fracture. The third image is observed as the secondary strong fracturing stage, which radiates strong signals seen as the secondary waveforms outlined by the dashed box in (a). After estimating the travel time ($tt=56.6\mu s$) after $T_N+70\mu s$, we postulate that the dominant Rayleigh-wave signal recorded by the SW sensor is associated with the third image. In contrast, the clear arrivals at the NW sensor are unrecognized due to the short distance. This indicates the entire fracturing process has not uniformly occurred, which probably results in variations in the fluid pressure.

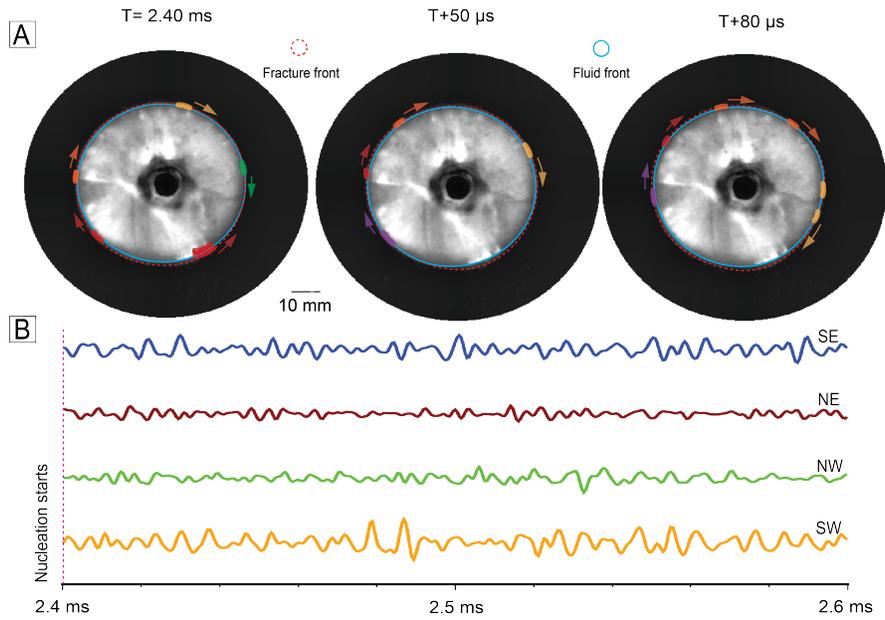


Figure D.9: Low fluid-viscosity (water) experiment images and AE signals. (a) Three images after at $T=2.4$ ms, at $T+50 \mu\text{s}$, and at $T+80 \mu\text{s}$. In the low fluid viscosity case, multiple nucleation events nucleate nearly simultaneously (tracked by the colored points from differential images) and show interesting propagating behaviors (shown by the arrows of multiple colors corresponding to their sources). (b) The four raw AE signals don't seem to give a straightforward indication of the nature of the fracture behavior. Indeed, as many events occur very close in time, their signals overlap. The change in the fracture area behind the fracture front (white or gray pixels) can be easily calculated over time.

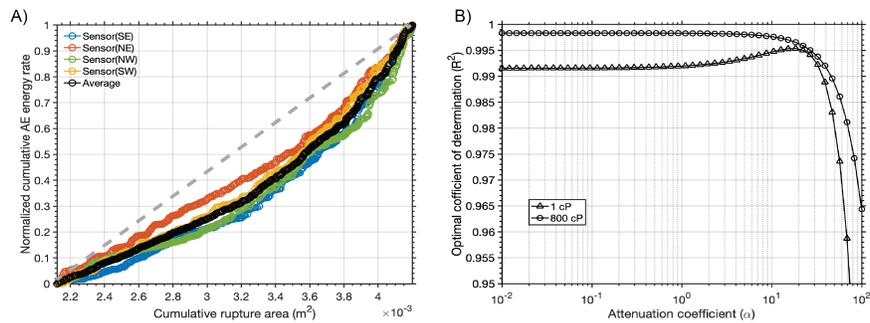


Figure D.10: AE energy and rupture area. (a) The scaling of AE energy rate and fast transverse fracture area (small red areas in Fig.5.1b) using raw AE signals. The radiated energy is uncorrected for geometric spreading and attenuation (see Equation 2 in the main text). (b) Sensitivity analysis of attenuation factors with respect to linear regression coefficients. For both high- and low-viscosity experiments, the attenuation coefficient is close to 20 and has negligible effects on the scaling relation. Therefore, the scaling relation is strongly influenced by the effect of geometric spreading.

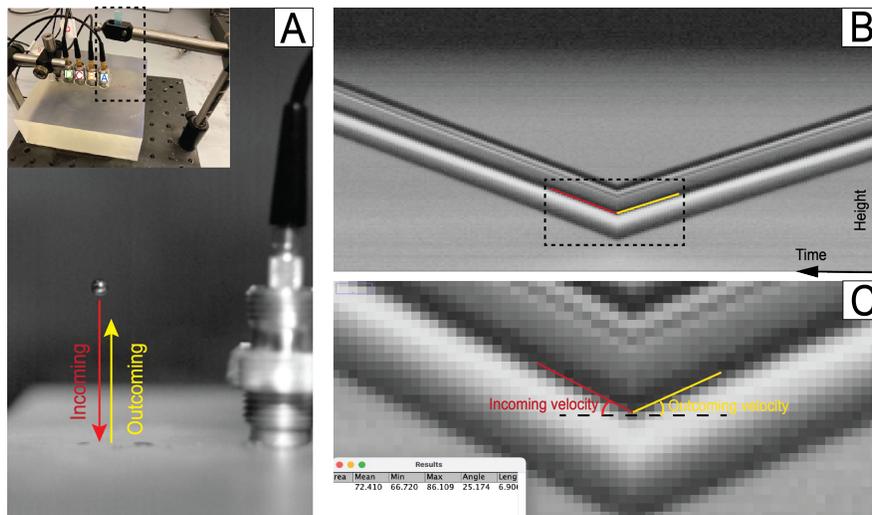


Figure D.11: Drop Ball tests: AE combined with high-speed imaging. (a) Experimental setup of the drop ball test (inner upper picture) and zoom-in image of falling steel ball ($r=2\text{mm}$) captured by the high-speed camera. We place four AE sensors in a line to record the AE signals from the ball impacts similar to the capillary break tests (Fig.D.7). The ball falls and hits the sample at an incoming velocity and bounces back up with an outgoing velocity. (b) Incoming and outgoing velocities can be measured using high-speed imaging using a Kymograph (distance, time representation). The red and yellow lines represent the incoming and outgoing trajectories, and their slope equals the velocities (c). We perform experiments at different heights to observe the fluctuation of the incoming and outgoing velocities. The observations are shown and analyzed for the energy calibration in Fig. D.10.

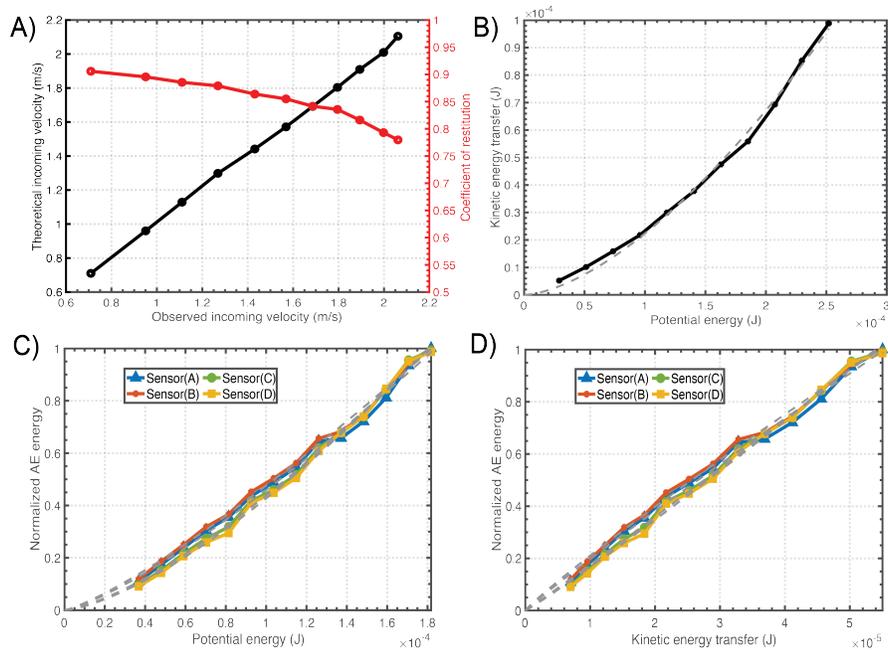


Figure D.12: AE Energy calibration using the drop-ball test. (a) The theoretical velocities from potential energy are almost equal to the observed incoming velocity from a high-speed camera (Fig. D.11) as shown by the black line. The coefficient of restitution, Re , is not linearly decreasing with the height, as shown by the red line. (b) Relation between the potential energy and kinetic energy transfer during ball collision. (c) Relation between the potential energy and AE energy calculated from the integration of squared voltages over time. (d) Relation between the kinetic energy transfer and AE energy. The detailed explanation and process can be found in the supplementary Text 2.

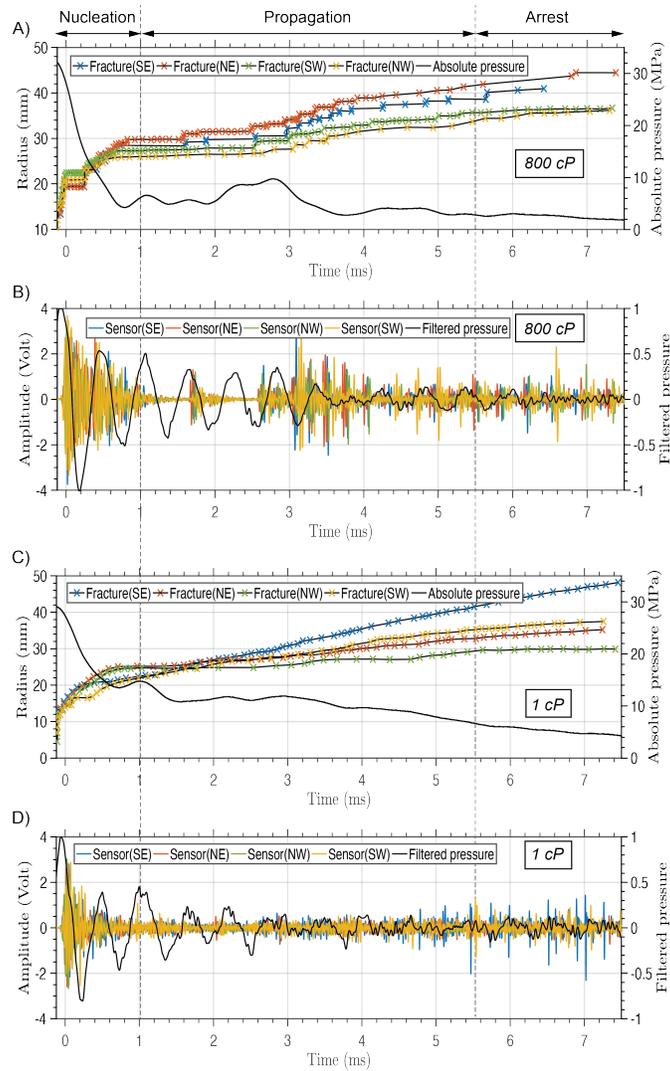


Figure D.13: Relations between fluid pressure, fracture-front propagation, and AE signals. (a, c) Interactions between absolute fluid pressure (originally measured at the injection site with a sampling rate of 3.125 MHz) and radial fracture fronts in four directions (SE, NE, NW, and SW). (b, d) Relation between the filtered fluid pressure (20 kHz) and the four AE signals. Note that the upper two panels are obtained from another high-viscosity (800 cP) fluid-injection experiment, and the lower two panels are from another low-viscosity (1 cP) fluid injection experiment. Both experiments are similar to the experiments used in the main text. However, the fluid pressure is, unfortunately, not measured by a high-resolution (i.e., high sampling rate) pressure transducer.

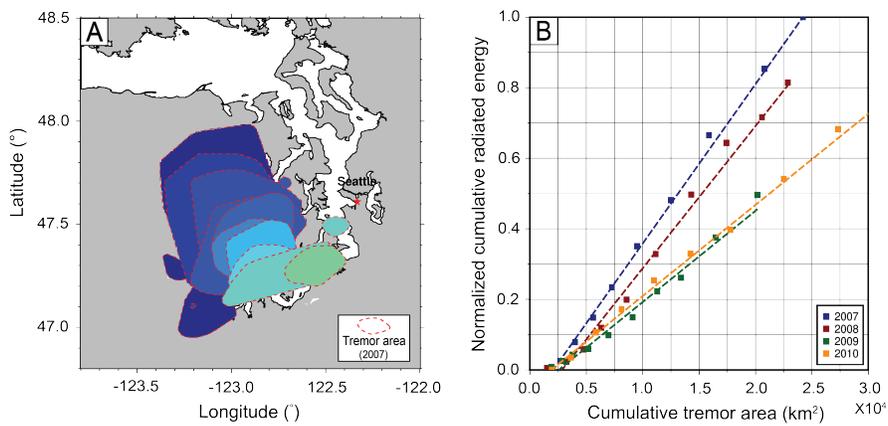


Figure D.14: Spatio-temporal Tectonic Tremors. (a) Contours of tremor source area in each of the first 10 days of the 2007 tremor swarm occurred in the Northern Cascadia. Each area is obtained by contouring the epicenters of the daily occurring tremors. (b) The cumulative radiated energy is proportional to the cumulative tremor area for each of the four Northern Cascadia tremor swarms. Each square denotes the daily accumulated tremor area and accumulated radiated seismic energy².

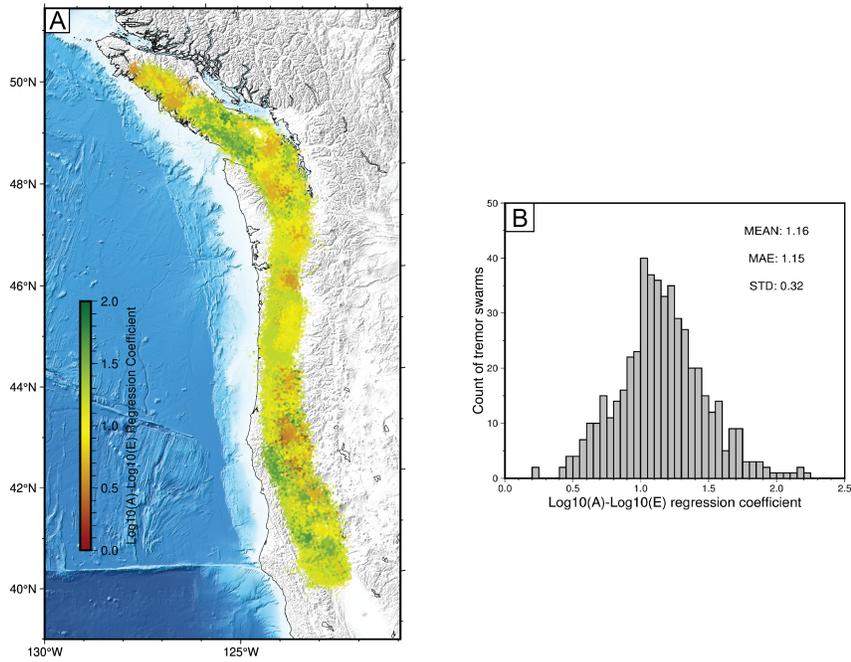


Figure D.15: Analysis of the regression coefficients between logarithmic (base 10) cumulative radiated energy of tremors and area in Fig.5.6. (a) The spatial distribution of these regression coefficients for 500 tremor swarms in the Cascadia region from 2017 to 2023. Each coefficient is obtained through the linear regression analysis between the logarithmic (base 10) cumulative tremor energy and area for each tremor swarm. The resulting coefficient is then employed to uniformly color-code all tremors within the specific tremor swarm. These tremor swarms exhibit varying logarithmic powers associated with tremor areas. We also display them in histograms for statistical analysis, as shown in b. The mean, median, and standard deviation of these logarithmic powers, or regression coefficients, are 1.16, 1.15, and 0.32, respectively. It is evident that most tremor swarms display a near-linear relationship between tremor radiated energy and tremor area.