

# Zooming into earthquake ruptures: from kinematics to dynamics

A DISSERTATION PRESENTED

BY

JIUXUN YIN

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

HARVARD UNIVERSITY

CAMBRIDGE, MASSACHUSETTS

DECEMBER 2021

©2021 – JIUXUN YIN  
ALL RIGHTS RESERVED.

## Zooming into earthquake ruptures: from kinematics to dynamics

### ABSTRACT

Earthquakes, especially the large ones, cause huge hazards and threaten people living nearby. It is important to improve our understanding of the earthquake source process. I do so by combining seismic data, observations, and simulation of dynamic ruptures. I first made the striking observation that the megathrust earthquakes present a ubiquitous pattern of coseismic rupture with updip low-frequency radiation and downdip high-frequency radiation, based on the backprojection and spectral analysis of teleseismic P waves. I tied this observation to the unique kinematics and dynamics of megathrust earthquakes. To relate the backprojection images, I used synthetic seismograms from theoretical and kinematic sources to explain that the backprojection images are proportional to the slip history, albeit a spatial smoothing operator that I derived. To further illustrate the observations of depth-frequency relation during megathrust earthquakes, I build a suite of 2D dynamic rupture models of megathrust earthquakes in realistic Earth structures. I find that systematic variations in the slip rate functions along-dip of the fault exhibit a systematic variation of frequency content.

Given the variety in model settings, I attribute the ubiquitous depth-frequency radiation observation to the interaction of the earthquake with the Earth's free surface, which is my preferred first-order explanation. Diving more into the observations of earthquake source time functions, we also find that heterogeneous representation of fault properties is necessary to explain the complexity in the source time function, and earthquake rupture is a dynamical process that is self-organized. These systematic studies on the seismic signature of earthquakes have shown the power of combining observations and simulations to decipher source physics from observations.

Many observations rely on the quality of these signals at a broad range of frequencies, which leads me to my last chapter. I develop a machine-learning multi-task model to separate the earthquake signals from the ambient seismic vibrations in complex seismic data. I have tested this method on an island station in Hawaii contaminated by tectonic, volcanic earthquake signals, and a strong microseismic source around the island. I show the promises of such an approach to improve earthquake and ambient noise seismology.

# Contents

1	INTRODUCTION	1
2	METHODS OF SEISMIC OBSERVATIONS FOR LARGE EARTHQUAKE RUPTURES	9
2.1	Introduction	10
2.2	Improved Compressive Sensing BackProjection (ImCS-BP)	13
2.3	Spectral source analysis	20
2.4	Conclusion	29
2.5	Data and resources	30
3	SPATIAL AND TEMPORAL EVOLUTION OF EARTHQUAKE DYNAMICS: CASE STUDY OF THE $M_w$ 8.3 ILLAPEL EARTHQUAKE, CHILE	31
3.1	Introduction	33
3.2	Application of ImCS-BP	35
3.3	Application of spectral source analysis	39
3.4	Integrated source analysis for the 2015 Illapel Earthquake	52
3.5	Rupture dynamic constraints inferred from spatial and temporal source analysis	54
3.6	Conclusion	57
3.7	Data and resources	58
4	RELATING TELESEISMIC BACKPROJECTION IMAGES TO EARTHQUAKE KINEMATICS	60
4.0	Introduction	62
4.1	Methods	65
4.2	Backprojection on kinematic sources	68
4.3	Resolvability of global earthquakes and arrays	79
4.4	Discussion	84
4.5	Conclusion	98
4.6	Data and resources	98
5	THE EARTH'S SURFACE CONTROLS THE DEPTH-DEPENDENT SEISMIC RADIATION OF MEGATHRUST EARTHQUAKES	99

5.1	Introduction . . . . .	101
5.2	Ubiquitous depth-frequency relation from back-projection analysis . . . . .	105
5.3	Dynamic rupture simulation of megathrust earthquakes . . . . .	109
5.4	Results from dynamic rupture simulations . . . . .	114
5.5	Discussion . . . . .	120
5.6	Conclusion . . . . .	129
5.7	Data and resources . . . . .	130
<b>6</b>	<b>SYSTEMATIC ANALYSIS ON EARTHQUAKE SOURCE TIME FUNCTIONS REVEALS PATTERNS IN EARTHQUAKE DYNAMICS</b>	<b>131</b>
6.1	Introduction . . . . .	133
6.2	Gaussian-subevent decomposition . . . . .	136
6.3	Dynamic time warping and clustering analysis . . . . .	140
6.4	Modeling STF complexity with dynamic rupture simulation . . . . .	152
6.5	Conclusion and Discussion . . . . .	158
6.6	Data and resources . . . . .	160
<b>7</b>	<b>A MULTI-TASK ENCODER-DECODER TO SEPARATING EARTHQUAKE AND AMBIENT NOISE SIGNAL IN SEISMOGRAMS</b>	<b>161</b>
7.1	Introduction . . . . .	163
7.2	Network design of WaveDecompNet . . . . .	166
7.3	Application to continuous seismic data . . . . .	179
7.4	Conclusion and Discussion . . . . .	190
7.5	Data and resources . . . . .	191
<b>8</b>	<b>CONCLUSION</b>	<b>192</b>
	<b>APPENDIX A SUPPLEMENTARY MATERIALS FOR CHAPTER 4, “RELATING TELESEISMIC BACKPROJECTION IMAGES TO EARTHQUAKE KINEMATICS”</b>	<b>195</b>
	<b>APPENDIX B SUPPLEMENTARY MATERIALS I FOR CHAPTER 5 “THE EARTH’S SURFACE CONTROLS THE DEPTH-DEPENDENT SEISMIC RADIATION OF MEGATHRUST EARTHQUAKES”</b>	<b>206</b>
	B.1 Text S1. Details on the dynamic rupture simulations . . . . .	207
	<b>APPENDIX C SUPPLEMENTARY MATERIALS II FOR CHAPTER 5 “THE EARTH’S SURFACE CONTROLS THE DEPTH-DEPENDENT SEISMIC RADIATION OF MEGATHRUST EARTHQUAKES”</b>	<b>225</b>
	<b>REFERENCES</b>	<b>287</b>

# Listing of figures

1.1	Overview of earthquake source study . . . . .	3
1.2	Basic idea of source observations . . . . .	4
1.3	Schematic figure of megathrust fault with several possible dynamic mechanisms of the depth-frequency relation . . . . .	6
2.1	Synthetic test from the auto-adaptive source grid refinement scheme . . . . .	16
2.2	Computational efficiency for the ImCS-BP method . . . . .	17
2.3	Schematic illustration of the reference source adjustment technique (RSAT) . . . . .	18
2.4	Synthetic test for the reference source adjustment technique. . . . .	19
2.5	Steps to retrieve the source spectra for the small events . . . . .	23
2.6	Cartoon to illustrate the method of time-varying spectral analysis . . . . .	28
3.1	Tectonic setting of southern Chile and location of the 2015 Illapel, Chile Mw 8.3 earthquake . . . . .	34
3.2	Data processing for the ImCS-BP . . . . .	36
3.3	ImCS-BP results of the 2015 Mw 8.3 Illapel Earthquake . . . . .	38
3.4	Rupture velocity estimated from ImCS-BP results . . . . .	39
3.5	Station distributions with azimuth and great circle distance for each EGF event . . . . .	41
3.6	SNR distributions of my data for the spectral analysis . . . . .	43
3.7	Source spectrum fitting for the Mw 8.3 Illapel Earthquake . . . . .	44
3.8	Spectral fitting for the total event source spectrum in each azimuth bin . . . . .	45
3.9	Source parameters estimated from the source spectrum at each station . . . . .	47
3.10	Time-varying radiated energy and falloff rate . . . . .	49
3.11	Time-varying radiated energy estimated at azimuth bin . . . . .	50
4.1	BP resolution matrix . . . . .	69
4.2	Examples of synthetic source and seismograms . . . . .	72
4.3	Takeoff angle distributions and BP resolvability of the synthetic arrays . . . . .	72
4.4	Comparison between the kinematic source model and BP results . . . . .	76
4.5	Comparison between the BP images and the kinematic source in the time domain . . . . .	77
4.6	Global map of seismic arrays . . . . .	80

4.7	Resolvability of the Hi-Net array . . . . .	81
4.8	Resolvability of all global arrays . . . . .	83
4.9	Generalized resolvability with travel time perturbation . . . . .	87
4.10	Example of array response . . . . .	94
4.11	Resolvable area as a function of frequency . . . . .	96
5.1	Ubiquitous depth-frequency relation found by back-projection observations . . . . .	103
5.2	Comparison between the LF BP centroid depth from GSN and HF BP centroid depth from regional arrays (0 - 70 km) . . . . .	108
5.3	Model setting . . . . .	110
5.4	Simulation results of small megathrust earthquakes at different depths . . . . .	116
5.5	Space-time evolution of the simulated megathrust earthquake . . . . .	118
5.6	Slip history of the simulated earthquakes . . . . .	119
5.7	Spectral properties of the source radiation . . . . .	121
5.8	Effects of the free surface and material contrasts on the dynamic behavior of megathrust earthquakes . . . . .	122
5.9	Tsunami and ground motion hazards . . . . .	127
6.1	Example of Gaussian-subevent decomposition . . . . .	137
6.2	Results of GSD for both SCARDEC STFs and simulated STFs . . . . .	139
6.3	Example of dynamic time warping clustering . . . . .	142
6.4	Source time function clustering, grouping, and conceptual interpretation . . . . .	143
6.5	Comparison between the GSD subevent number and DTW complexity groups . . . . .	145
6.6	Comparisons between prominent peak number . . . . .	146
6.7	Population distribution of four complexity groups and correlation with different source parameters . . . . .	148
6.8	Spatial distribution of DTW complexity . . . . .	149
6.9	Source parameter estimates of the SCARDEC database . . . . .	151
6.10	Earthquake distributions of different complexity groups on the vertical profiles . . . . .	153
6.11	Pre-stress distribution to generate synthetic STFs . . . . .	154
6.12	Group proportion distributions from DTW . . . . .	157
7.1	Architecture of ML network to separate earthquake and ambient noise . . . . .	167
7.2	Example 1 of waveform separation using ML network . . . . .	171
7.3	Example 2 of waveform separation using ML network . . . . .	172
7.4	Training curves of networks with different bottlenecks . . . . .	175
7.5	Comparison between networks with different bottlenecks . . . . .	176
7.6	Application to continuous seismic data from an island station IU.POHA . . . . .	181
7.7	Example of STA/LTA . . . . .	183
7.8	Single-station cross-correlation and auto-correlation functions filtered in low frequency band (0.1 - 1.0 Hz) . . . . .	186

7.9	Same as Figure 7.8 for the MF medium frequency band 1.0-2.0 Hz. . . . .	187
7.10	Same as Figure 7.8 for the HF high frequency band 2.0-4.0 Hz. . . . .	188
7.11	Coherence of the single-station cross-correlation and auto-correlation functions . . . . .	189
A.1	One example of the kinematic source and its relevant source parameters . . . . .	197
A.2	Waveforms recorded by the arrays near the nodal plane of focal mechanism . . . . .	198
A.3	Comparison between BP results and kinematic source . . . . .	199
A.4	BP image in each sub-frequency-band . . . . .	200
A.5	The variation of resolvability $\varepsilon_I$ with frequency . . . . .	201
A.6	The spectral falloff of the BP peak amplitude for the kinematic source model . . . . .	202
A.7	Calculated resolution matrices and their corresponding array response . . . . .	203
A.8	Comparison between the Linear BP image with the CSBP results . . . . .	204
A.9	Resolution matrix of the Hi-net array towards the IDN2 (Sumatra) region . . . . .	205
B.1	Domain of dynamic simulations . . . . .	208
B.2	Corresponding S wave velocity from different settings of $V_p/V_S$ ratios . . . . .	212
B.3	Regularization of slip rate functions . . . . .	213
B.4	Comparison between the LF BP centroid depth from GSN and HF BP centroid depth from regional arrays (70 - 700 km) . . . . .	216
B.5	Data and back-projection results of the Mw 9.0 2011 Tohoku-oki earthquake . . . . .	217
B.6	Data and back-projection results of the Mw 7.9 2015 Gorkha earthquake . . . . .	218
B.7	Data and back-projection results of the Mw 8.3 2015 Illapel earthquake . . . . .	219
B.8	Model settings of the five small megathrust earthquake models . . . . .	220
B.9	Space-time plot of the slip histories for all small earthquake models . . . . .	221
B.10	Results of the spectral content of the slip-rate function extracted at individual point every 10 km . . . . .	222
B.11	Results of the spectral content of the slip-rate functions averaged over 10-km subfault . . . . .	223
B.12	Results of a mode-III (anti-plane) rupture model on a vertical fault intersecting the free surface in a homogeneous medium . . . . .	224
C.1	Model settings and simulation results of model 1 . . . . .	228
C.2	Model settings and simulation results of model 2 . . . . .	229
C.3	Model settings and simulation results of model 3 . . . . .	230
C.4	Model settings and simulation results of model 4 . . . . .	231
C.5	Model settings and simulation results of model 5 . . . . .	232
C.6	Model settings and simulation results of model 6 . . . . .	233
C.7	Model settings and simulation results of model 7 . . . . .	234
C.8	Model settings and simulation results of model 8 . . . . .	235
C.9	Model settings and simulation results of model 9 . . . . .	236
C.10	Model settings and simulation results of model 10 . . . . .	237
C.11	Model settings and simulation results of model 11 . . . . .	238

C.12	Model settings and simulation results of model 12 . . . . .	239
C.13	Model settings and simulation results of model 13 . . . . .	240
C.14	Model settings and simulation results of model 14 . . . . .	241
C.15	Model settings and simulation results of model 15 . . . . .	242
C.16	Model settings and simulation results of model 16 . . . . .	243
C.17	Model settings and simulation results of model 17 . . . . .	244
C.18	Model settings and simulation results of model 18 . . . . .	245
C.19	Model settings and simulation results of model 19 . . . . .	246
C.20	Model settings and simulation results of model 20 . . . . .	247
C.21	Model settings and simulation results of model 21 . . . . .	248
C.22	Model settings and simulation results of model 22 . . . . .	249
C.23	Model settings and simulation results of model 23 . . . . .	250
C.24	Model settings and simulation results of model 24 . . . . .	251
C.25	Model settings and simulation results of model 25 . . . . .	252
C.26	Model settings and simulation results of model 26 . . . . .	253
C.27	Model settings and simulation results of model 27 . . . . .	254
C.28	Model settings and simulation results of model 28 . . . . .	255
C.29	Model settings and simulation results of model 29 . . . . .	256

# Acknowledgments

I would like to express my most sincere appreciation to my advisor Marine Denolle. When I was applying for graduate school, Marine had just started her career at Harvard. I didn't know her until I read about her research on the department website. At that time, I was about to submit my application at the last moment of the deadline. Her research deeply attracted me, so I listed her as my first advisor option and emailed her about my interest immediately AFTER I submitted my application. This was a little risky at that time, but it turns out that I would never regret my decision. Marine kindly replied to me and finally hired me as one of her first two graduate students in her research group.

I am so lucky to have Marine as my advisor on this long Ph.D. journey in my life. Her great enthusiasm for scientific research has had a significant influence on my career in academia. Marine is a very responsible advisor. No matter how busy she was, she could always make time to provide help for my study and research. It was a great pleasure to talk about science with her. She always has so many interesting new ideas that greatly enlightened me. Instead of direct supervision on specific projects, she allowed me to explore different scientific questions freely. She provided as much help as she could. Without the freedom from her, I could not have had the chance to work on different interesting projects and learn different new skills. Writing papers with Marine were also great experiences. Even when she had lots of other tasks on hand, Marine always spared time to help me revise my draft paper line by line and taught me her experience writing papers. Without her patience, it was impossible to have all my research published in different journals during my Ph.D. Marine also provided invaluable support for my career life. In the last year of my Ph.D., I planned to stay in academia and continue doing research as a postdoc. In the beginning, I was too shy to reach out to people. Marine gave me much encouragement and help during that hard time. She tried her best to advertise my research and introduce me to different people. Her efforts gave me much confidence and finally helped me find a good postdoc position. Without her generous help, this transition period of mine would not be so smooth. Marine does not only care for our research and career but also care for our non-academic life. She often organized different group activities for our group. Through those activities, all of our group members could release from intense research life, build stronger connections and better know each other. To me, Marine is more like a friend rather than an advisor. I am so happy to work with her as a graduate student and sincerely thank her for everything she has done for me.

I would also like to extend my deepest gratitude to my committee members: Brendan Meade, Jim

Rice, John Shaw, Dave Johnston. I have really learned a lot from them for these years at EPS. The completion of my dissertation would not have been possible without their help. They gave me much helpful feedback on my Ph.D. research during my qualification exam as well as the yearly progress report. Those valuable suggestions from them have greatly improved my Ph.D. research and my dissertation.

Specifically, I want first to thank Professor Brendan Meade. Brendan provided much help for my research. He was always generous to share with me the elegant research codes he wrote, and I have benefited a lot from learning his codes. I also worked as a teaching fellow in a class he taught, which was an invaluable experience for me. He showed me a great example of teaching, especially when the students were so diverse with different backgrounds. He always tried to think for each individual student, make everyone comfortable in class, and have all students really learn something from every lecture. I believe this close-distance teaching lesson from Brendan will be a great model for me to follow.

I also want to thank Professor Jim Rice. I took Jim's class "Mechanics in Earth and Environmental Science" twice. Every time I went to his class, I could always be inspired, and I enjoyed learning from and discussing with Jim. I still remember the first time I had an individual meeting with Jim to show him my work. I had just started my Ph.D. at that time, and I was so nervous about presenting in front of such a big name. But it turned out to be a delightful meeting; Jim kindly listened to my presentation and gave me many helpful suggestions, from which I benefited so much. More importantly, as a G1 graduate student at that time, this experience greatly encouraged me and gave me huge momentum to move forward.

I am also grateful to Professor John Shaw. John was the director of our department when I came to EPS, and he has done a lot for the department through those years. Although he was busy, he was always willing to listen to the voices from students and provided his help as much as possible. As Marine moved to the University of Washington in my last year of Ph.D., John was so kind to be my local advisor at Harvard. He was so responsive and could always respond to my request for help quickly and efficiently.

Thanks should also go to professor Dave Johnston. As my GSC (graduate student council) committee member, Dave also provided valuable help for me to satisfy different requirements for a Ph.D. degree.

I would like to thank Bradley Lipovsky. Brad helped me a lot on the mathematical derivations of many different projects. He also gave me many useful suggestions on the numerical simulation of earthquake rupture.

I also enjoyed working with all the current and former colleagues in the Denolle group, including graduate students: Tim Clements, Seth Olinger, Natasha Toghramadjian, Congcong Yuan; and postdocs: Chris Van Houtte, Loic Viens, Chengxin Jiang, Kurama Okubo, Zhitu Ma, Xiaotao Yang, Laura Ermert. Our group is such a diverse group with members from different countries all over the world. I was so glad to work with all of them. The daily communication, as well as research collaboration, has really enriched my Ph.D. life. Years of life in such a diverse group have taught me how to communicate with and listen to other people with different cultural backgrounds. Furthermore, the harmonic group environment serves as a great model for me to understand and learn

diversity better.

Thanks also to other professors in the department, especially those who offered courses that I took: Eli Tziperman, Jerry Mitrovica, Miaki Ishii, Peter Huybers, Zhiming Kuang. I've really learned lots of knowledge and skills from the great courses they provided. The training I received from their courses was significant and helped me better think about and solve the scientific questions I met in my research.

I would like to acknowledge the assistance of all EPS staff, former and present. Their efforts and support have made EPS a wonderful community for me to focus on my study and research. Special thanks to Aimee Smith, Esther James, Danielle da Cruz, Miguel Baquerizo, Summer Smith, Kathleen McCloskey, Elizabeth Busky, Annika Quick, Sarah Colgan, Amanda Legee, Katherine Wiryaman.

I also wish to thank all my GeoMuseum 200B office mates and EPS friends for making my Ph.D. life at EPS so memorable. Special thanks to my friends, Zhuo Yang, JJ Dong, David Wei, Packard Chan, Chris Chan.

Except for people at EPS, I would like to extend my sincere thanks to other people who are not at Harvard but also provided invaluable help. Particularly helpful to me during this time were Huajian Yao, who was my former advisor and still provided his generous help and support for my academic career; Qingyang Du, who was one of my best friends, and we went to the same high school, college and worked in the same city (he was a graduate student and postdoc at MIT); Jean-Paul Ampuero, who has developed so many user-friendly open source research codes that greatly helped me in different research projects, and also gave me lots of great suggestions on my papers.

Last but not least, I'm deeply indebted to my family. Life in a foreign country sometimes may be challenging; my parents always provide me with invaluable suggestions, selfless support and try their best to give me the best they can offer, although they are thousands of miles away. I'm extremely grateful to my beloved wife, Bing He. Bing has made several significant decisions in her life for me: coming to the US with me and marrying me. We have been through every aspect of life together, and she really changed my life and made it so colorful. With her support, every dream would have been possible. I feel that I am the luckiest person in the world to have Bing with me.

# 1

## Introduction

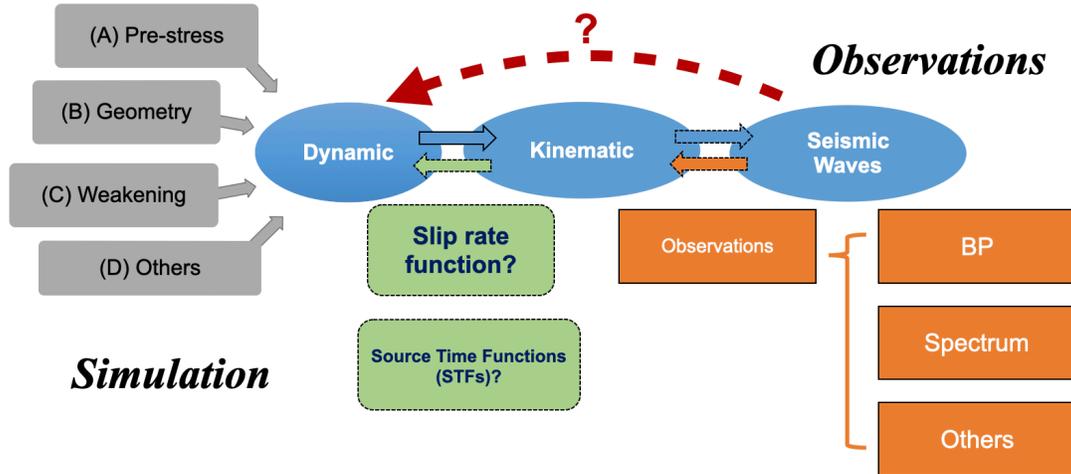
The largest earthquakes in recorded history, for instance the 2004 Sumatra  $M_w$  9.2, the 1960 Great Chilean  $M_w$  9.4, the 1964 Alaskan  $M_w$  9.3, and the 2011 Tohoku  $M_w$  9.0 megathrust earthquakes, have killed many, caused great damage, and brought tremendous financial losses. These earthquakes induce not only strong ground motion but also large tsunami waves. As populations grow, especially in the coastal regions where large earthquakes usually occur, mitigating the risk of earthquake hazards is critical. During large earthquakes, the building collapse caused by ground motion, and

the coastal destruction caused by huge tsunami waves are deadliest to humans. Both of them are directly controlled by the earthquake source, and their mitigation requires sufficient knowledge about the physics of the earthquake rupture. Different processes in the rupture, such as its direction of propagation, may incur locally different ground motion<sup>145,290</sup>. Therefore, better understanding the rupture during a large earthquake is of scientific significance and essential to mitigating the risk of seismic hazards.

The recent surge of large earthquakes<sup>173</sup> and the improvements in the observational techniques and numerical approaches provide an excellent opportunity to push forward our understandings of the rupture process of large earthquakes. For example, the recent observed depth-frequency relation of megathrust earthquakes that the low-frequency seismic wave comes from the shallow part while high-frequency seismic wave comes from the deeper part in most subduction zones provides good constraints on physics of the seismic regions in the subduction zones. However, today's observational approaches are not necessarily well posed to validate physical models. For instance, mechanisms that affect the source, the pre-stress heterogeneity (A), fault geometry (B), frictional weakening mechanisms (C), and medium inelastic response (D) all interact in the dynamic evolution.

In order to bridge seismic observations with earthquake physics, I need to tackle both aspects of seismology simultaneously (Fig. 1.1): to develop new **observational** techniques and metrics that are directly relevant to **physical models** of the source and lab experiments. This is the main topic of this Ph.D. dissertation.

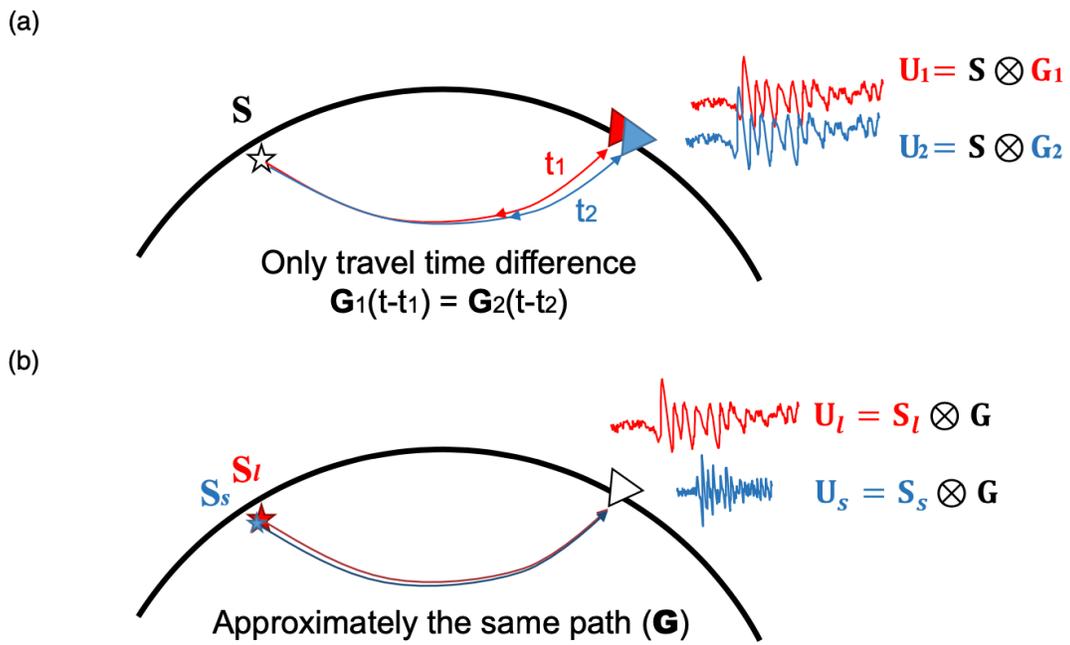
Starting from seismic waves, I mainly use the back-projection (BP) and spectral source analysis methods to observe the source from the seismic waves. The step critical to all source observation is to properly remove the path effects from the source to the station, which is parameterized by the Green's function (eGf). The back-projection and spectral source analysis approaches reflect two typical ways to remove the Green's function. BP is usually applied to dense seismic array data. For the array data, the Green's functions from the source to the array of stations can usually



**Figure 1.1:** General overview of my earthquake source study: source processes (gray boxes), observational techniques (orange boxes), investigation type (blue boxes), slip histories (green box) represent the link between dynamics and kinematics

be approximated to be identical except their travel time (Fig. 1.2). With this basic assumption, the seismic waveforms can be directly projected back to the source. Then, the stacking of the shifted waveforms at every potential location in the source region is coherent for the correct location (high amplitude) and incoherent (low amplitude) for the incorrect source location. This technique gives a time-dependent image of the source region and provides information about the earthquake rupture propagation. In the source spectra analysis (SP), I use empirical an Green's function, that is a nearby small event that share the same path with the main large event, to remove the Green's function in the frequency domain (Fig. 1.2 (b)).

As I will show in Chapter 2, these two methods require few assumptions and can provide direct information about the source process. Supplemented with information from other techniques (source inversion, directivity analysis, etc.) when possible, I can recover the kinematic source process based on these observations and try to infer the mechanisms that control earthquake dynamics. A case study on the 2015 Illapel, Chile  $M_w$  8.3 earthquake applying those techniques will be shown in



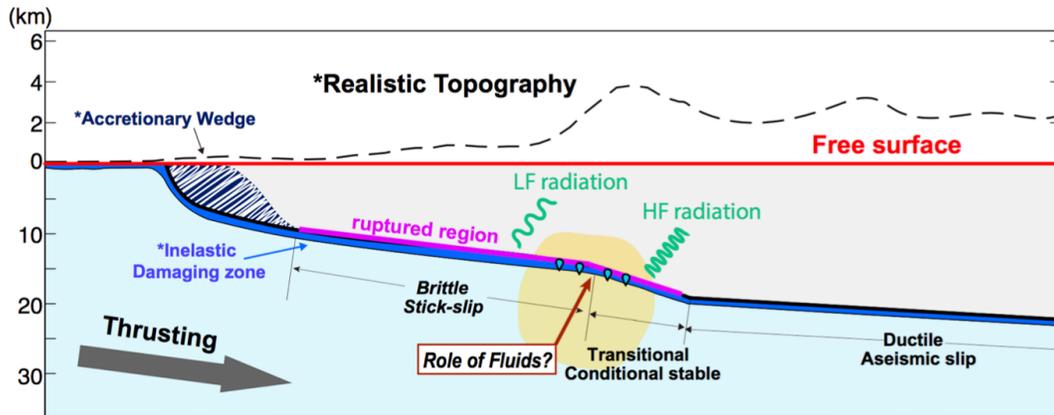
**Figure 1.2:** Basic idea of (a) back-projection methods and (b) spectral source analysis. These two methods attempt to remove the path effect (Green's function) in different ways. (a) BP is based on the assumption that the path effects can be simplified as travel time differences; (b) SP uses a nearby small event sharing the same wave path to remove the path effects for the main event.

### Chapter 3.

A rigorous understanding of their relationship is critical to interpret the BP results in terms of the earthquake source process. Theoretical studies indicate that the high-frequency seismic waves, which is usually the input data of BP, can be excited during abrupt changes in rupture velocity<sup>32,193,194,291</sup> caused either by the arrest of the rupture<sup>193</sup> or by kinks of the fault geometry<sup>195</sup>. Huang et al.<sup>129, 126</sup> introduce pre-stress heterogeneity on the fault surface to reproduce the excitation of high-frequency radiators at depth. I will show in Chapter 4 that the BP image is indeed proportional to the distribution of slip histories after a spatial smoothing. BP provides observational constraints and information about the earthquake dynamics, in complement to other independent techniques.

To bridge the seismic observations and earthquake source dynamics, I also need to find the seismic signature of various dynamic effects. It is necessary to either differentiate the dynamic mechanisms from seismic observations, such as through direct observations of seismic waveforms, their spectral content and BP images, or address the ambiguity of these mechanisms in seismic waves. To solve this question, I focus on the observation on the depth-frequency relation during megathrust events and focus on the rupture dynamics behind (Fig.1.3).

The first mechanism that can explain the depth-frequency relation of seismic radiation is the stress conditions on the fault before the rupture. Pre-stress and its heterogeneity are particularly difficult to establish, as multi-earthquake cycle simulations require both long-term fault healing and short-term dynamics<sup>84,85</sup>. Besides, the fault geometry (ramps or kinks, etc.) also plays an important role in the pre-stress heterogeneity<sup>330</sup>. A second mechanism that can explain the depth-frequency relation is the frictional properties on the fault interface. The phase transformation of the minerals constituting the near-fault fabric alters their frictional properties<sup>277</sup>. Parametrization of friction, either through velocity-<sup>78,77,259,264</sup> or slip-weakening<sup>16,132,242</sup>, is often invoked through a characteristic or critical length scale that corresponds to the frequency variation. A third mechanism is that



**Figure 1.3:** Schematic figure of megathrust fault with several possible dynamic mechanisms of the depth-frequency relation in seismic radiation.

the heterogeneous geological structure in subduction zones can exaggerate the depth dependence to seismic radiation by reducing the high-frequency radiation at the trench <sup>47,167,188,231,270</sup>. In Chapter 5, I build two-dimensional dynamic rupture models for megathrust earthquakes and explore different mechanisms for this depth-frequency relation. My results show that the Earth's free surface is necessary and sufficient to explain the observed megathrust earthquake radiation style. The realistic structure of the subduction zone is necessary to predict more accurate earthquake ground motion duration and tsunami potential.

The earthquake dynamics (stress-driven) impacts the earthquake kinematics (displacement-driven) through the moment-rate function, which is approximated from the source time function (STF). In Chapter 6, we study systematically the earthquake source time functions (STFs) and their relation to other observables of earthquake dynamics. We developed different techniques to extract physical information from a large population of earthquake STFs. We explore STF databases and find that their shapes have essential information about the earthquake dynamics. Our results imply that there is critical information hidden in the earthquake variability. This information helps constrain rupture dynamics from the analysis of population behaviors of large samples of earthquake

STFs.

In the last Chapter 7, I develop a machine learning (ML) method to separate earthquake and ambient noise signals. Separating ambient noise from seismic data can increase the signal-to-noise ratio (SNR) for the earthquake signals over a broader range of seismic frequencies. It can significantly help detect more earthquakes and further constrain the earthquake rupture process. I develop a new encoder-decoder ML network that can be directly applied to the 3-component raw seismic data in the time domain. I explore different settings of network architectures to find that the long-short-term-memory (LSTM) bottleneck outperforms other networks. I then apply the trained model to separate earthquake and noise signals for the continuous seismic raw data recorded at the IU.POHA station in Hawaii. My results show that the quality of both separated earthquake and noise signals has been improved significantly. The promising results show that my developed encoder-decoder network for the separation of earthquake and noise signals can dramatically help to improve the quality of seismic data, especially for those stations installed in noisy environments such as ocean islands or ocean bottom.

The following authors contributed to Chapter 2: Jiuxun Yin, Marine A. Denolle, Huajian Yao

Yin, J., Denolle, M.A., & Yao, H. (2018). Spatial and Temporal Evolution of Earthquake Dynamics: Case Study of the Mw 8.3 Illapel Earthquake, Chile. *Journal of Geophysical Research: Solid Earth*, 123(1), 344–367.

The following authors contributed to Chapter 3: Jiuxun Yin, Marine A. Denolle, Huajian Yao

Yin, J., Denolle, M.A., & Yao, H. (2018). Spatial and Temporal Evolution of Earthquake Dynamics: Case Study of the Mw 8.3 Illapel Earthquake, Chile. *Journal of Geophysical Research: Solid Earth*, 123(1), 344–367.

The following authors contributed to Chapter 4: Jiuxun Yin, Marine A. Denolle

Yin, J. & Denolle, M.A. (2019). Relating teleseismic backprojection images to earthquake kinematics. *Geophysical Journal International*, 217(2), 729–747.

The following authors contributed to Chapter 5: Jiuxun Yin, Marine A. Denolle

Yin, J. & Denolle, M. A. (2021). The earth's surface controls the depth-dependent seismic radiation of megathrust earthquakes. *AGU Advances*, 2(3), e2021AV000413.

The following authors contributed to Chapter 6: Jiuxun Yin, Zefeng Li, Philippe Danré, Marine A. Denolle, Bradley P. Lipovsky

Danré, P., Yin, J., Lipovsky, B. P., & Denolle, M. A. (2019). Earthquakes Within Earthquakes: Patterns in Rupture Complexity. *Geophysical Research Letters*, 46(13), 7352–7360.

Yin, J., Li, Z., & Denolle, M. A. (2021). Source time function clustering reveals patterns in earthquake dynamics. *Seismological Society of America*, 92(4), 2343–2353

The following authors contributed to Chapter 7: Jiuxun Yin, Marine A. Denolle, Bing He

# 2

## Methods of seismic observations for large earthquake ruptures

## SUMMARY

I develop a new methodology that combines compressive sensing back-projection (CS-BP) and source spectral analysis of teleseismic P waves to provide metrics relevant to earthquake dynamics of large events. I improve the CS-BP method by an auto-adaptive source grid refinement as well as a reference source adjustment technique (RSAT) to gain better spatial and temporal resolution of the locations of the radiated bursts. I also use a two-step source spectral analysis based on simple theoretical that includes depth phases and water reverberations and based on empirical P-wave Green's functions. Furthermore, I propose a time-varying source spectral analysis method that provide the time evolution of dynamic parameters such as radiated energy and falloff rates. Bridging back-projection and time-dependent spectral analysis provides me a spatio-temporal evolution of these dynamic source parameters. In this chapter, I present a detailed description on the earthquake source observational methods that I have developed.

### 2.1 INTRODUCTION

Earthquakes cause large ground motions, damaging the man-made structures and greatly threatening human lives. In order to mitigate risk of seismic hazards, it is important to understand the physics of earthquakes such as how they nucleate, propagate, terminate, and interact with and impact the integrity of Earth materials. Unfortunately, due to the remoteness of the earthquake source region, earthquake science community highly relies on the far-field, teleseismic data to infer the earthquake source process, especially for those occurring in the off-shore area. Robust and innovative observational techniques, combined with seismological theory, can significantly improve my understanding of the source process.

Inversions for slip on the fault are a classic approach to provide static and/or kinematic description of the source <sup>140,141,151,152,153</sup>. Finite slip inversions yield the total slip that occurred during the

earthquake and may be obtained from low-frequency seismic waves and/or geodetic measurements. Kinematic slip inversions solve for the time evolution of slip on the fault. They often require higher frequency seismic waves and/or high-rate GPS and often necessitate assumptions of the earthquake dynamics by imposing the slip-rate functional shape. Furthermore, these inversions are especially non-unique<sup>198</sup>, require regularization, and thus lead to different estimations of peak value and location of maximum slip. Finally, as slip models can provide important estimates of static stress drop, the regularization in the inversion directly affects the values<sup>44,359</sup>.

Backprojection of teleseismic P waves is another widely-used method to study the evolution of earthquake radiation and is particularly effective for large earthquakes. With the development of dense seismic array (e.g. Hi-net in Japan<sup>239</sup>; US Array of the United States (Earthscope program)), I can harness key information from waveform similarity or coherency through the approximation that the Green's function is similar within the seismic array except for phase delays. By shifting the travel-time differences and stacking the waveforms to enhance the coherency, the backprojection can directly provide the distribution of seismic energy bursts<sup>138,139</sup>. Its application to the recent large earthquakes ( $M_w > 8$ ) have succeeded in characterizing the spatiotemporal evolution of seismic radiation<sup>95,138,139,156,209,344,347,349,350,359</sup>. Backprojection methods hold fewer assumptions than slip inversions. Therefore, preliminary information about the earthquake rupture process can be rapidly obtained from the waveform data and in a very shortly after the waves arrive at the seismic array, for instance, IRIS DMC<sup>137</sup>. However, understanding earthquake dynamics from backprojection results, for example, what physics explains the high-frequency radiation, remains to be addressed.

Because it is difficult to infer earthquake dynamics from slip inversions and backprojection, I turn to source spectral analysis and quantify the amount of seismic energy radiated at the source. This approach does not account for the phase information as it removes path effects from the deconvolution of the observed displacement amplitude spectrum with that of the Green's function. Through model fitting<sup>6,39,45,46,92,113,192</sup>, source spectral analysis has been a popular method to

provide estimates of static and dynamic source parameters such as static stress drop (with measure of the corner frequency interpreted as source duration and assumption of uniform stress drop), radiated energy (with measures of velocity spectral shapes). Radiation efficiency, a measure of energy partitioning, can be obtained from the ratio of radiated energy with static stress drop <sup>131,161,251,315,355</sup>. Additional information such as measures of high-frequency falloff rate  $n$  brings insight to the efficiency or deficiency in high-frequency excitation. Based on the circular crack model with uniform stress drop, Kaneko & Shearer <sup>148</sup> find that the high-frequency falloff rate varies between 1.5 and 2.5 for far-field P waves with azimuth and with takeoff angles relevant to teleseismic P waves ( $0^\circ$ – $40^\circ$ ). Their results imply that great variations in  $n$  can come from simple and known dynamics of earthquake rupture.

In this study, I propose to combine two independent observational methods, backprojection and spectral analysis, to evaluate the evolution in time and in space of dynamic source parameters.

My first approach is an Improved Compressive-Sensing BackProjection method, which I refer to as ImCS-BP herein. The original CS-BP method <sup>349,360</sup> uses a frequency-domain formulation of the conventional back-projection method by extracting coherence between velocity seismograms recorded at seismic arrays through phase measurements. It can locate the seismic energy bursts with high spatial resolution <sup>349,360</sup>. Here, I increase the spatial resolution and improve the computational efficiency to provide better constraints on the spatiotemporal evolution of seismic energy bursts.

Second, to complement the phase information provided by the backprojection, I apply a two-step spectral analysis method that constructs the far-field P-wave source spectrum both for the whole event as well as throughout the rupture. My approach allows me to estimate P-wave radiated energy and to parameterize the spectral shapes with high-frequency falloff rates and corner frequencies through model fitting for the whole event. I also proceed with this method on spectrograms of the P-wave displacement waveforms and calculate interesting dynamic source parameters during the earthquake. The combination these two independent methods provides me a comprehensive and

well-consistent set of observations, with independent measures of phase and of amplitude.

## 2.2 IMPROVED COMPRESSIVE SENSING BACKPROJECTION (IMCS-BP)

The idea of Compressive Sensing (CS) for inverse problems was first proposed in the applied math and signal processing communities<sup>53,80</sup> to allow sparsity of the models through minimization of the  $L_1$  norm. Yao et al.<sup>349,351</sup> applied this inversion technique to the backprojection method in order to study large subduction zone earthquakes ( $M_w > 8$ ). Yin & Yao<sup>360</sup> carefully tested and improved the CS-BP method to be more resistant against the potential data outliers, which may carry abnormal amplitudes or phases that will affect the locating of energy bursts. The sparsity constraint of the CS scheme increases the resolution in the location of the seismic energy bursts during the rupture, an advantage to the conventional backprojection methods<sup>138</sup> as well as to beamforming methods<sup>360</sup>.

My CS-BP method uses a sliding time windows technique<sup>349,360</sup>. I first pre-process the P-wave velocity seismogram at each station of a seismic array (see Yin & Yao<sup>360</sup> for details). Then for each sliding time-window, I take their Fourier transform and construct a vector  $\mathbf{b}(f)$  that depends on frequency and that has a dimension of the number of stations  $N$ . My model is a vector  $\mathbf{x}(f)$  containing  $M$  grid points of the potential locations of the source generating the pulse in that particular P-wave window, projected on a flat plane at the centroid depth. The reference time  $t_{n0}$  is the travel time from the epicenter to the station  $n$ , and the travel time between the location each grid point  $m$  to each station  $n$ ,  $\tau_{nm}$ . The theoretical global P-wave arrival times are based on the IASP91 velocity model<sup>150</sup>. Wave propagation is simplified to a time shift in the frequency domain and I construct the phase spectrum matrix  $\mathbf{A}(f)$ , of size  $N \times M$ , and whose entries are:

$$A_{nm}(f) = e^{-2i\pi f(\tau_{nm} - t_{n0})} \quad (2.1)$$

The basic idea of CS is to invert the system  $\mathbf{b}(f) = \mathbf{A}(f)\mathbf{x}(f)$  by minimizing the following misfit

function:

$$\|\mathbf{A}(f)\mathbf{x}(f) - \mathbf{b}(f)\|_1 + \lambda\|\mathbf{x}(f)\|_1 \quad (2.2)$$

where  $\lambda$  is a damping factor that attunes data (first term in Eq.2.2) and model (second term in Eq.2.2)  $L_1$  norms. I refer to Yao et al. <sup>349</sup>, Yin et al. <sup>359</sup> for more discussion on the choice of  $\lambda$ . I solve the sparse inversion problem using the CVX package (<http://cvxr.com>, last accessed 2012/11/03) based on convex optimization and the interior point method<sup>41</sup>. Two important specificities of my method affect the inversion: 1) the meshing of the model  $\mathbf{x}(f)$  and 2) the truncation of the time-window data  $\mathbf{b}(f)$ . This study focuses on these two factors to improve the conventional CS-BP.

### 2.2.1 AUTO-ADAPTIVE REFINEMENT OF THE SOURCE GRID

In my CS problem, the number of model parameters in  $\mathbf{x}(f)$  equals the number grid points in the source region. One might want to densify the source region, i.e. increase the number of grid points, to enhance the details of the image. However, a uniform and denser mesh of the source region (or in the model space) will greatly decrease the numerical efficiency of the inversion due to the larger dimension of the inversion (Eq.2.2) and unnecessary computation spent on the points with zero amplitude in my sparse description. Therefore, I develop an auto-adaptive grid refinement method to improve both numerical efficiency and spatial resolution of my method.

I use a synthetic test to construct the iterative refinement scheme. The synthetic source is simplified to a two-spike source with 10% random noise added to the data  $\mathbf{b}(f)$  at a frequency  $f$  chosen to be 0.5 Hz (Fig.2.1). I first start with a uniform and coarse mesh (40 km spaced grid points) on a 360 km  $\times$  320 km source region, that is, totally  $9 \times 8 = 72$  unknown model parameters. After initial results, I select the few grid points that have values exceeding  $10^{-6}$  times the maximum (Fig.2.1 (b)). Then, I refine those selected grid points by a factor of two and perform the inversion again. I repeat

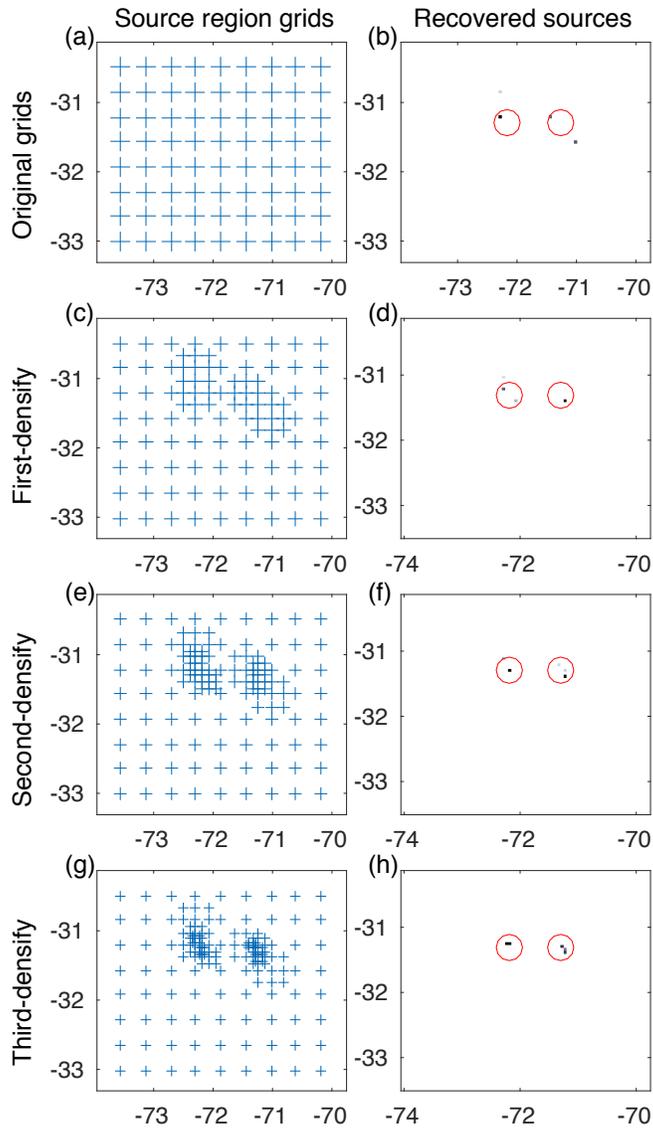
these steps until a grid size of  $5 \text{ km} \times 5 \text{ km}$ , which is close to the smallest resolvable horizontal distance  $L = 6 \text{ km}$  using a P wave with wave speed of  $c = 6 \text{ km/s}$  and takeoff angle  $i = 30^\circ$  (typical for teleseismic P waves) at a frequency  $f = 0.5 \text{ Hz}$ , i.e.,  $L = c/f\sin(i)$ . If I uniformly mesh the source region with grid size of  $5 \text{ km}$ , I need to solve a total of 4745 unknown model parameters, most of them being null in the sparse representation. However, with the auto-adaptive source grid method, I only need to solve 4 equations with only about 100 model parameters at each step. With the unnecessary and redundant computation discarded, I can get the robust CS results in both high spatial resolution as well as in good computational efficiency while remaining within reasonable physical constraints. The computational runtimes for this test case are shown in Fig.2.2.

### 2.2.2 REFERENCE SOURCE ADJUSTMENT TECHNIQUE (RSAT)

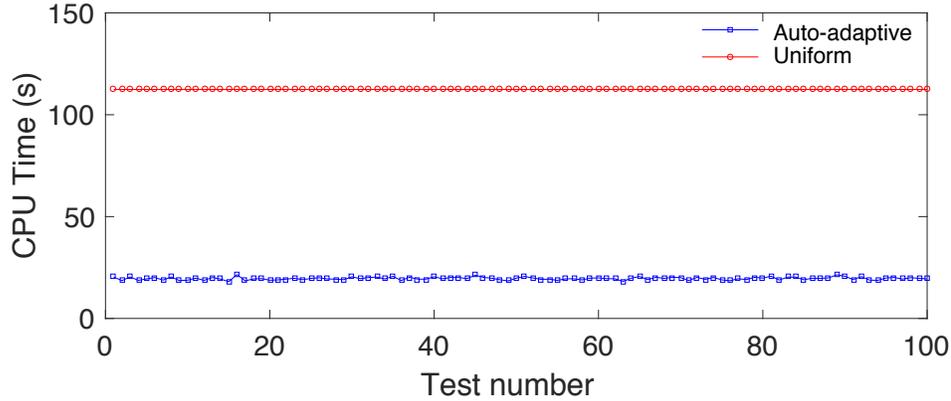
Most backprojection-based imaging methods require a reference location to align the velocity<sup>139,209,344,349,360</sup> that is usually set to be the event epicenter. The location of seismic energy bursts can then be obtained relative to that reference. Because my approach operates in the frequency domain, the alignment is performed at the beginning of the direct P wave in the first time window such that, for each sliding time window, the alignment remains unchanged throughout the waveform<sup>349,360</sup>.

For large earthquakes ( $M_w > 8$ ), as the rupture propagates, the later P waves originate from different source locations than that of the epicenter. This can bias the alignment in each sliding time window, introducing an apparent shift that increases with distance and that results in a loss of coherency in the truncated time window towards the end of the waveform. Fig.2.3 (c) illustrates this effect and shows that the effect leads to a partial sampling (truncation) of the coherent signal the sliding the window, in particular for the low frequency waveforms, and I refer to this issue as the “truncation effect”. It brings systematical errors to the data vector  $\mathbf{b}(f)$  and blurring artifacts in between real sources of energy bursts (Fig.2.4 (b)).

Instead of aligning the entire waveforms solely based on the epicenter (Fig.2.3 (a)), I propose



**Figure 2.1:** Synthetic test from the auto-adaptive source grid refinement scheme. (left) The blue crosses show the grid points used in the inversion. (right) The dots show the results recovered from the corresponding grids and the red circles indicate the positions of my synthetic sources (circle centers).



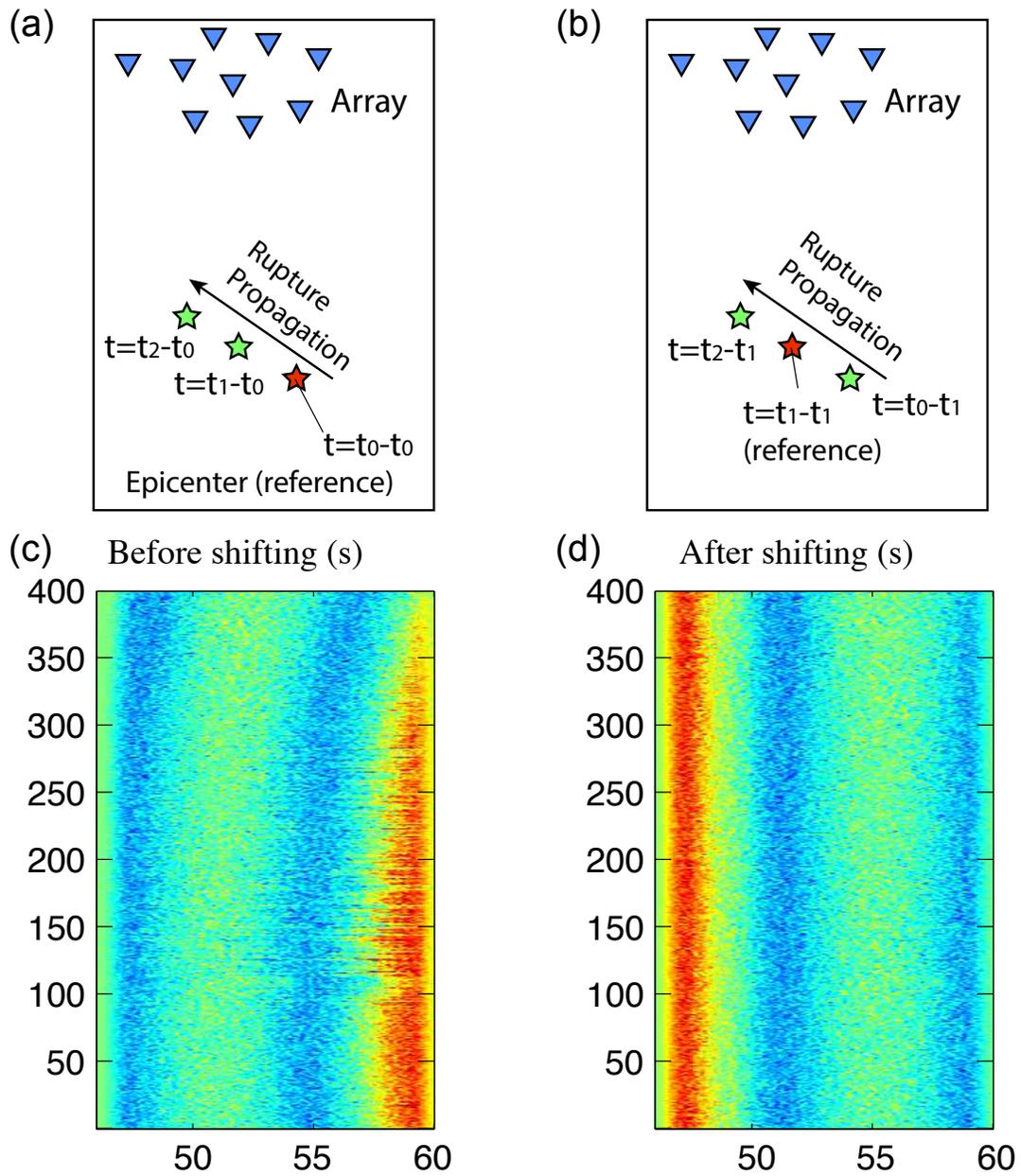
**Figure 2.2:** Computational efficiency for the ImCS-BP method. I run the synthetic test mentioned in Section 2.1 (Fig.2) on a laptop with 2.9 GHz Intel Core i5 processor for 100 times with different random noises (10% of maximum amplitude of the input data, random phase) added to the input data. Red circles show the CPU time for conventional CS-BP method with uniform grid (grid size of 5 km, totally 4745 grid points) while the blue squares show the CPU time for the ImCS-BP method with auto- adaptive grid size.

to use a sliding window adjustment technique. The first window is aligned based on travel times between the epicenter and the receivers. I solve for the position of the highest radiated energy burst  $r$ , calculate the new relative travel-time shift  $t_{nr}$  between the new reference location and station  $n$ , and use it to realign the waveforms in the current time window (Fig.2.3 (d)). This procedure leads to the new data vector  $\mathbf{b}'(f)$  and updates the phase spectrum matrix  $\mathbf{A}'(f)$ , whose values are:

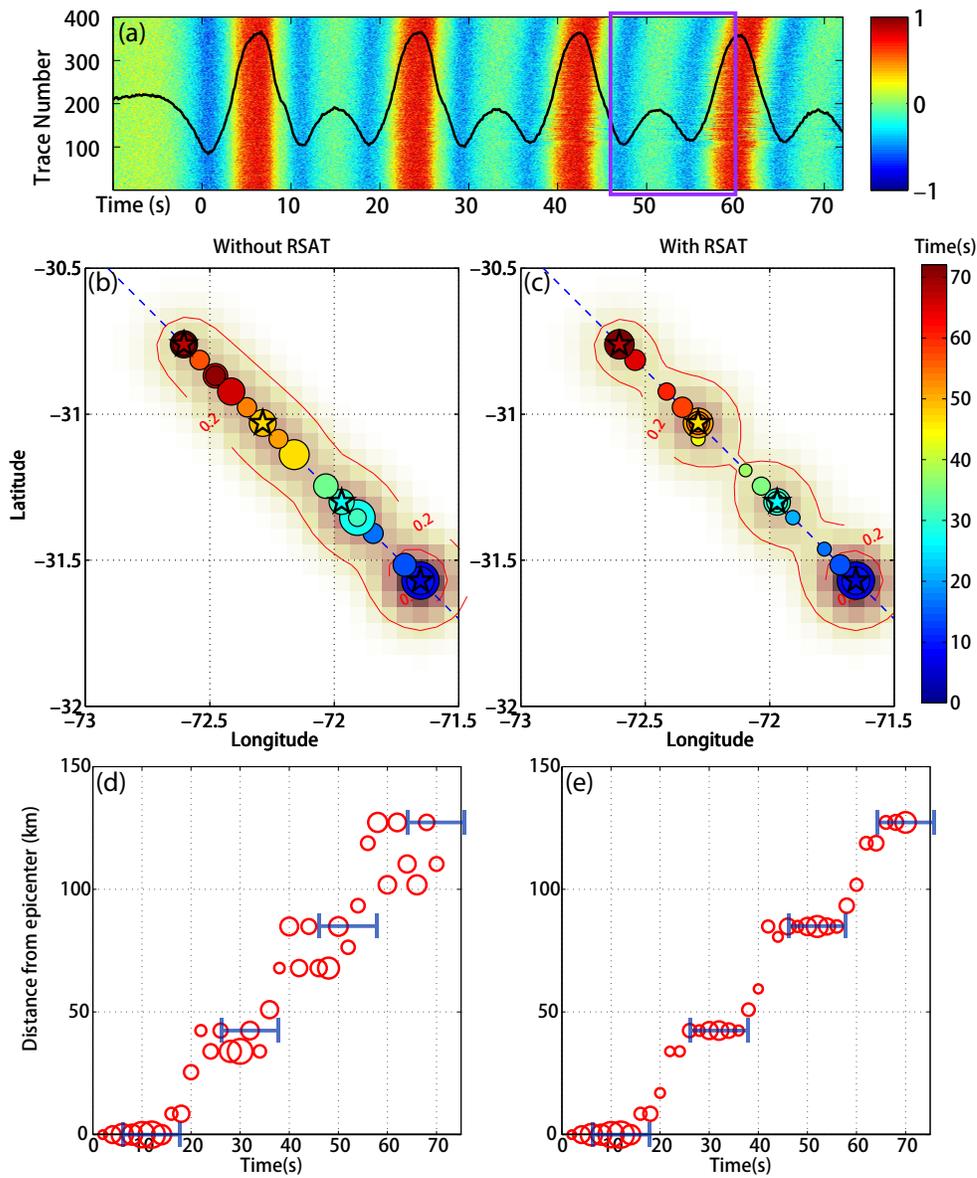
$$A_{nm}(f) = e^{-2i\pi f(\tau_{nm} - t_{nr})}. \quad (2.3)$$

The new system of equations becomes  $\mathbf{b}'(f) = \mathbf{A}'(f)\mathbf{x}(f)$  and the procedure is repeated when sliding through the entire waveform. The waveform truncation effect can be greatly suppressed as illustrated in Fig.2.3. Since I only have one reference position, this technique is most efficient for unilateral ruptures.

In this section, I again construct synthetic time series to verify my technique and measure its improvements. I set 4 point sources, each separated in space by 42 km and in time by 20 s, to mimic a



**Figure 2.3:** Schematic illustration of the reference source adjustment technique (RSAT) (a, b) and one example of waveform “truncation effect” before (c) and after (d) time adjustment.



**Figure 2.4:** Synthetic test for the reference source adjustment technique. Top panels: (a) synthetic time-domain waveforms recorded at each station where the color scale indicates the waveforms amplitude, the centered black waveform is the stacked seismogram, and the purple box shows the waveform within 40-60 s (similar to the waveform in Fig.3). Middle panels: results from conventional (b) and improved (c) CS-BP method where the stars represent my synthetic sources, the circles are the recovered sources colored by source (or rupture) time, the background copper color scale corresponds to the total energy, the red contours indicate 20% of the energy maxima, and the dashed lines show the synthetic rupture direction. Bottom panels: location of the energy bursts against time from my synthetic source (blue segments) and recovered from conventional (d) and improved (e) CS-BP (red circles, left and right respectively).

northwestward unilateral rupture with a propagation speed of 2.1 km/s. The synthetic source plane is set at the depth of 20 km to represent that of the 2015 Illapel Earthquake (details about this earthquake are covered in the following chapter) and the 400 synthetic receivers are at the location of the US Array stations. Similar to my previous work<sup>359,360</sup>, I extract and taper the first 12 s of the  $M_w$  8.3 mainshock seismograms recorded by each station and use these wavelets as synthetic waveform. I compute the travel times in IASP91 velocity model, insert the wavelets to their corresponding arrival times, and add noise (built in time domain with random amplitude kept lower than 10% of the peak signal amplitude and random (uniform distribution) phase). Fig.2.4 (a) shows that the waveforms are aligned to the epicenter location and the later phases (around 40 s and 60 s) are subject to the apparent shift due to the northwestern rupture propagation. I solve for Eq.2.2 using  $\mathbf{A}(f)$  of Eq.2.1 (without the time-adjustment technique, Fig.2.4 (b) and (d)) and  $\mathbf{A}'(f)$  of Eq.2.3 (with time-adjustment technique, Fig.2.4 (c) and (e)). I find that without the time-adjustment technique, while the overall pattern of rupture propagation can be recovered (Fig.2.4 (b)), there are many artifacts in between the synthetic point sources (Fig.2.4 (b) and (d)) that blur the location of the sources. In contrast, the sliding-window RSAT is able to recover well each point source (Fig.2.4 (c) and (e)), with limited blurring in between.

By combining both the auto-adaptive refining source grid method and the sliding-window RSAT, my improved CS-BP is able to provide more accurate positions of radiated energy bursts, which should help to constrain the spatiotemporal evolution of the dynamics of large earthquakes.

### 2.3 SPECTRAL SOURCE ANALYSIS

My ImCS-BP method can provide relevant observational information about the location and timing of high-frequency bursts. However, as any backprojection-based result, the location of the bursts found by ImCS-BP really show P-waveform coherency and their physical interpretation re-

mains unclear (I will address this question in Chapter 4). While some features of the earthquake source (e.g. minimum source dimension, rupture propagation direction, average rupture velocity, ...) can be assessed with ImCS-BP, they may not be sufficient to estimate dynamic source parameters such as radiated energy. In order to provide additional observational constraints on the rupture dynamics of large megathrust earthquake, I complement my back-projection analysis with a spectral analysis method<sup>75,76</sup>.

### 2.3.1 REMOVING PATH EFFECTS

The relation between the far-field P-wave vertical displacement waveform  $u^n(t)$  recorded at station  $n$  and the moment-rate function  $s(t)$  can be written as:

$$u^n(t) = s(t) \otimes g^n(t), \quad (2.4)$$

where  $\otimes$  denotes convolution in the time domain and  $g^n(t)$  is a far-field elasto-dynamic Green's function (or spatial derivatives of the displacement Green's function if the source is a double couple) from the source to the  $n$ th station that comprises the wave propagation along the ray path.

Transforming Eq.2.4 into the frequency domain, the displacement of the far-field P wave becomes:

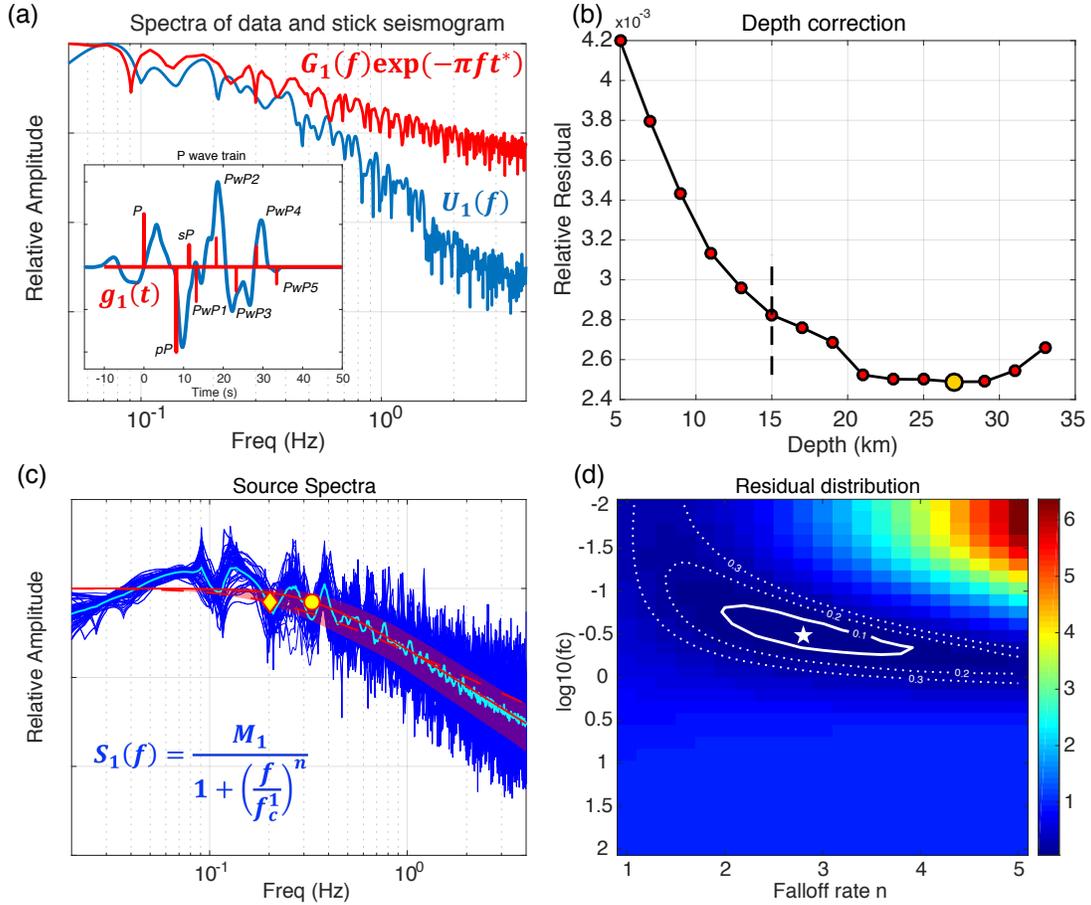
$$U^n(f) = S(f)G^n(f), \quad (2.5)$$

where  $U^n(f)$ ,  $S(f)$ , and  $G^n(f)$  are the Fourier transforms of the displacement seismogram, source moment-rate function, and Green's function, respectively. To recover the earthquake source term  $S(f)$  from the observed seismogram spectrum  $U^n(f)$ , it is essential to properly remove the effects from the wave propagation, i.e., the Green's function term  $G^n(f)$ . I use a two-step strategy to remove path effects that was proposed in Denolle & Shearer<sup>76</sup> and that combines theoretical simple Green's functions with empirical Green's functions (eGfs).

The principle of the eGf approach is to use small events nearby the large earthquake to calibrate for the true 3D path effects, but provided that I can construct their own source term. I do so by removing their path terms from a simplified and theoretical Green's function. The small earthquakes should share the same ray paths (i.e. Green's function) and the same focal mechanism (i.e. radiation pattern) as the large event. Without knowing the source spectrum of the small events, I am left with performing a spectral ratio<sup>19,214</sup>. However, it has limited value when trying to explore the high-frequency falloff rates as it only provides the relative falloff rate between the large and small events. One alternate strategy is to apply a theoretical model with global average parameters such as taking a uniform static stress drop of 3 MPa, high frequency falloff rate of 2, and a single-corner frequency parametric model such as that in Aki<sup>6</sup>, Baltay et al.<sup>25</sup>, Brune<sup>45</sup>. In the eGf method, the source spectra of both the small and large events are tightly related since the falloff of the small events directly controls the estimate of falloff for the large event. If the eGf quake is small enough (common practice is to use 3 magnitudes lower than the target event magnitude), the potential bias would be reduced. With teleseismic data, it is difficult to retrieve data with high SNR for low magnitude events. Solving for the small quake source spectrum is an attempt to reduce this effect. My spectral analysis method is summarized in Fig.2.5.

#### GETTING THE PATH EFFECTS FROM SMALL EVENTS ( $M_w$ 6)

The first step is to obtain the source spectra of the small events. The P wave train of shallow earthquakes contains not only the direct P, but also the P-wave reflections to the free surface (pP and sP) that greatly bias the displacement spectra<sup>75,76,123,333</sup>. In addition to the free surface effects, the direct P waves of shallow offshore earthquakes reverberate in the water column, potentially also interfering with the P wave train and altering their spectral shapes. In this study, I apply both the depth-phase approach of Denolle et al.<sup>75</sup>, Denolle & Shearer<sup>76</sup> and the water-phase filter proposed by Akuhara & Mochizuki<sup>8</sup> to correct my P-wave train displacement spectra. I regard the small events



**Figure 2.5:** Steps to retrieve the source spectra for the small events, using the 2015/09/18 Mw 6.1 event as example. (a) Spectra of a synthetic Green's function (red line) and the data (blue line) recorded at station AZ.FRD. With the corrected source depth, the synthetic Green's function with P wave train (stick seismogram in the inner panel, compared with the recorded P waveform) can be constructed and used to estimate a source spectrum. (b) Search of the source depth for the small event by fitting the spectral shape in (a). The dashed line indicates the GCMT centroid depth and the yellow circle shows the best-fit source depth. (c) The cyan spectrum is stacked over individual stations (blue spectra) and fit to a single-corner frequency model (red solid curves, equation (7)), with the corner frequency indicated by the yellow circle and an equivalent model with a stress drop of 3 MPa and falloff rate of 2 (dash red line and yellow diamond marker). (d) Misfit distribution of the spectra fitting for the two parameters.

as point sources and directly calculate their synthetic seismogram  $g_1(t)$  in a fluid-over-elastic half-space medium. Effectively,  $g_1(t)$  is a stick seismogram constructed from convolution of the Green's function with the moment tensor of the point source and a delta source time function:

$$g_1(t) = A_P \delta(t) + A_{pP} \hat{R}_P \delta(t - t_{pP}) + A_{sP} \hat{R}_S \delta(t - t_{sP}) + \sum_{l=1}^{\infty} A_l (-1)^l \hat{T} \hat{T}^{l-1} \delta(t - t_l). \quad (2.6)$$

The right-hand side terms of this equation represent the direct P, pP, sP, and water phases with arrival time of  $t_P = 0$  (waveforms are aligned to the direct P),  $t_{pP}$ ,  $t_{sP}$ , and  $t_l$ , respectively.  $\hat{R}_P$ ,  $\hat{R}_S$  are respectively the P- and S-wave reflection coefficients at the water-solid interface.  $\hat{T}$  and  $\hat{T}$  are transmission coefficients of the upgoing and downgoing P wave at the water-solid interface.  $\hat{r}$  is the reflection coefficient at the water-air interface.  $A_P$ ,  $A_{pP}$ ,  $A_{sP}$  and  $A_l$  are the amplitude terms that contain geometrical spreading  $1/4\pi\rho c^3 R$  ( $c$  seismic wavespeed,  $R$  source-receiver ray path length), radiation pattern, and Earth attenuation using a global  $t^*$  attenuation modeled in frequency domain<sup>331</sup> as  $e^{-\pi f t^*}$ .  $l$  is the order of water phases, i.e. the number of times the P wave reverberates in the water column, taken high enough (20 or 30) to retrieve the phases within the time window. Eq.2.6 provides a stick seismogram of P wave train containing the direct P, pP, sP, and water-phase arrival times and amplitudes. Because the source depth controls the relative arrival times and radiation pattern terms, I can find the optimal depth that minimizes the difference between the synthetic and the observed spectra<sup>75,76,333</sup> (Fig.2.5 (a)).

Ignoring directivity effects in the small events, I can recover a single reliable far-field P-wave source spectrum for each small event by deconvolving (division in the frequency domain) the observed P wave train displacement spectra with  $g_1(t)$  and stack over all stations. I fit the source spectrum to a single-corner frequency model,

$$S_1(f) = \frac{M_1}{1 + \left(\frac{f}{f_c^A}\right)^n}, \quad (2.7)$$

where  $M_1$  is the seismic moment of each small earthquake,  $f_c^A$  is the corner frequency, and  $n$  is the high frequency falloff rate. Because the corner frequency of those events is higher than my low-frequency cutoff, I can normalize the source spectra  $S_1(f)$  to the corresponding seismic moment.

To find the two source parameters (corner frequency and high-frequency falloff rate) of Eq.2.7, I minimize the L2 norm misfit in log-space between the shapes of measured and modeled spectra. I proceed by grid-search and show that the shape of the misfit function indicates a correlation between the two parameters (Fig.2.5 (d)). In addition to this correlation, high-frequency falloff rate directly trades off with my choice of attenuation. I also implemented a spatially variable attenuation model<sup>332</sup> but found no difference in the corner frequencies and only a slightly larger falloff rates in my estimates (Table 2.1). For simplicity, I use the globally averaged  $t^*$  model<sup>331</sup> in the  $g_1(t)$ . Interpretations on the small event source parameters are beyond the scope of this study. However, I validate that my results are reasonable through estimates of static stress drop, assuming a simple circular crack model with uniform stress drop,

$$\Delta\sigma = \frac{7}{16} \left(\frac{f_c}{k\beta}\right)^3 M_1, \quad (2.8)$$

where  $\beta$  is the S-wave velocity and  $k = 0.32$ <sup>92,192</sup>. I find the values of stress drop are comparable to global averages<sup>10,76</sup> (see results of a specific case study in Table 3.1 of Chapter 3). With the best-fit parameters, I use the spectral model of the small events (Eq.2.7) to recover the spectra of the large event.

Event (GCMT ID)	$t^*$		$t^* + dt^*$	
	$f_c$	n	$f_c$	n
200610121805A	0.33	2.8	0.33	3.0
201509170410A	0.13	2.2	0.17	2.6
201509180910A	0.33	2.8	0.33	3.2
201509191252A	0.33	2.8	0.33	2.8
201509210539A	0.45	3.0	0.45	3.2
201509211739A	0.33	3.0	0.33	3.2
201509220712A	0.62	3.2	0.62	3.4
201509260251A	0.45	3.0	0.45	3.4
201511070731A	0.33	3.0	0.33	3.2

**Table 2.1:** Comparison between source parameters (corner frequency and high frequency falloff rate) of EGF source models using global  $t^*$  attenuation model<sup>331</sup> as well as spatially variable attenuation model<sup>332</sup>  $t^* + dt^*$ .

## SOURCE SPECTRUM OF THE MAIN EVENT

The eGf method relies on the common Green's function  $G(f)$  between two earthquakes of interest. Given  $U_2(f)$  the observed P-wave train displacement spectrum of the large shock, and  $U_1(f)$  that of the small shock:

$$U_1(f) = S_1(f)G(f), \quad (2.9)$$

$$U_2(f) = S_2(f)G(f), \quad (2.10)$$

with  $S_2(f)$  is the target source spectrum,  $S_1(f)$  is the small even source spectrum (found in the first step). I combine Eqs.2.9 - 2.10 and find that:

$$S_2(f) = \frac{U_2(f)}{U_1(f)}S_1(f). \quad (2.11)$$

I stack all the source spectra for all small (eGf) events at all stations by averaging the log of the amplitudes and weighting by the number of stations within each  $10^\circ$  azimuth bin. Then I can calculate the P-wave radiated energy by integrating the squared velocity spectrum within a frequency band  $[f_1$

$f_2]$ :

$$E_P = \frac{8\pi}{15\rho\alpha^5} \int_{f_1}^{f_2} |fS_2(f)|^2 df, \quad (2.12)$$

in which  $\rho = 2.92 \text{ g/cm}^3$  is the density from the PREM model<sup>90</sup> and  $\alpha = 6.5 \text{ km/s}$  is the P-wave velocity from IASP91 model<sup>150</sup>. If I assume that both P and S waves share the same spectral shape, then the ratio of S-to-P radiated energy is  $3\alpha^5/2\beta^5$  and the total radiated energy becomes:

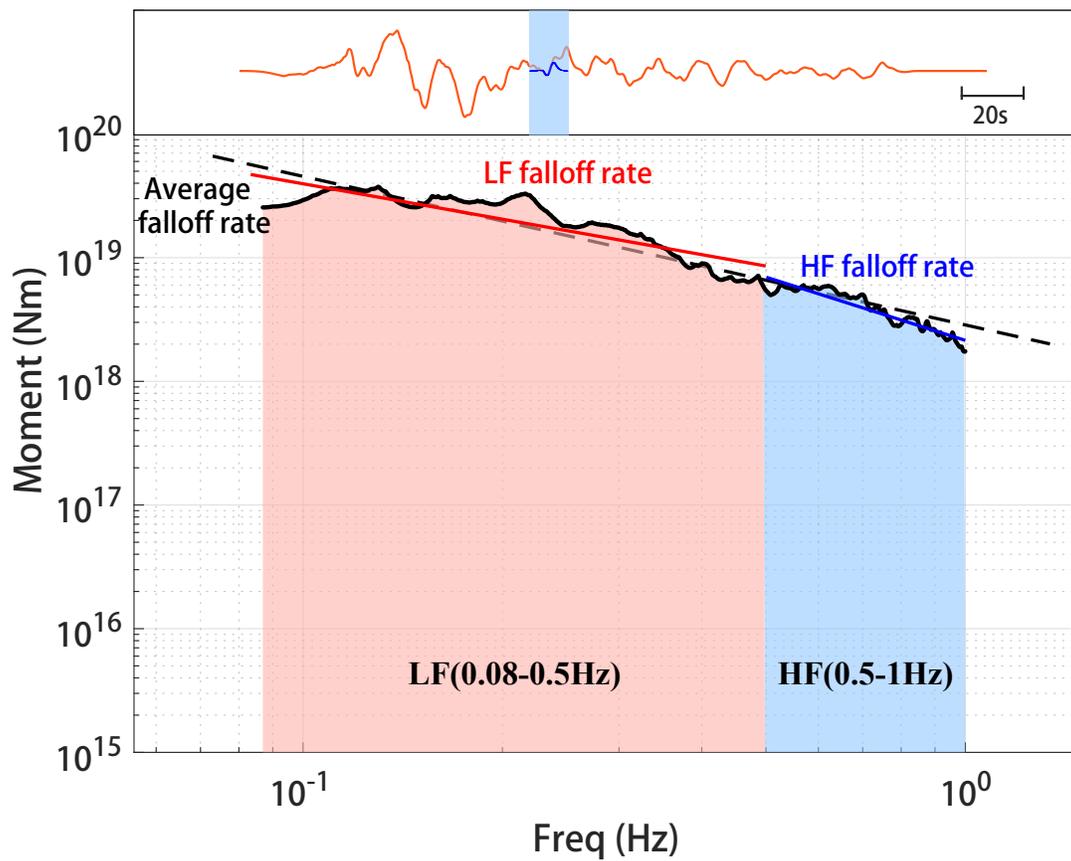
$$E_{total} = (1 + 3\alpha^5/2\beta^5)E_P. \quad (2.13)$$

I use  $\beta = 3.75 \text{ km/s}$ , the shear wavespeed from IASP91 model at the centroid depth. The total radiated energy must be estimated by integrating over the frequency interval  $[0 \infty]$  Hz. However, I have a limited frequency band within  $[0.02 \text{ } 2]$  Hz. I can use a parametric description of the source model to estimate the proportion of radiated energy missing from my limited frequency band.

### 2.3.2 TIME VARYING SPECTRAL ANALYSIS

The previous parts of spectral analysis used the whole P wave train to provide the overall source information on large earthquake event. To obtain more detailed information during the rupture itself, I perform a time varying analysis with robust estimations of radiated energy and parameterization of the spectrograms. Because I retrieve time-dependent dynamic source parameter estimate (radiated energy and falloff rates), I can directly interpret these metrics with the ImCS-BP results.

I slide a 12 seconds-long Hanning taper through the waveform of the mainshock with a 10-seconds overlap and a 2 seconds time step (Fig.2.6 upper panel). I remove the path effects simply from the eGfs, since I assume that the Green's function does not vary with time. Therefore, for each time window, I can measure a source spectrum within the resolvable frequency band (Fig.2.6 lower panel). I first average the time-dependent source spectra within each azimuth bin in order to



**Figure 2.6:** Cartoon to illustrate the method of time-varying spectral analysis. The top panel shows a waveform in red and the tapered running short time window waveform in blue. The bottom panel shows the recovered source spectra (black line) within the time window. Blue and red shaded areas indicate the two LF and HF frequency bands. Linear regressions are the best fit line for the spectrum from linear regression (blue: LF; red: HF; black dashed: overall).

investigate the azimuthal and temporal variations. Finally, I stack the azimuth-averaged spectra to construct an overall spectrogram for this event. These source spectra can represent snapshots during the dynamic source process.

The minimum frequency of the spectrograms is 0.08 Hz and I study the two different frequency bands of 0.08 - 0.5 Hz (LF) and 0.5 - 1.0 Hz (HF) to directly compare with the results from ImCS-BP. I estimate the radiated energy evolution during the earthquake by applying Eqs. 2.12 - 2.13 to each snapshot of the source spectrum and within the LF, HF, and the complete frequency band (0.08-1 Hz). Furthermore, I also fit in log-space a simple linear regression to the spectrogram and obtain the evolution of the falloff rate the LF, HF, and the complete frequency band (Fig. 2.6 lower panel). Although the physical interpretations of falloff rate remain uncertain<sup>6,148,192</sup>, the time variation of falloff rate provides a unique observation on the excitation of seismic wave throughout the rupture.

## 2.4 CONCLUSION

In this chapter, I present new methodologies that provide observations relevant to the dynamics of megathrust earthquakes. They are effective to analyze remote and offshore shallow earthquake by accounting for phase and amplitude information of the complete P wave train (direct P, depth phases, and water reverberations). Both methods provide observational constraints on the evolution of the seismic radiation during the rupture and complement each other by combining reliable phase and amplitude information. My multi-dimensional source analysis can provide the spatiotemporal evolution of seismic radiation during the earthquake, with the estimated values of radiated energy and varying falloff rate. The application of those methods to a specific case study will be covered in the following chapter.

## 2.5 DATA AND RESOURCES

I have uploaded the developed ImCS-BP to github (<https://github.com/yinjiuxun/CSBP>).

3

# Spatial and Temporal Evolution of Earthquake Dynamics: Case Study of the $M_w$ 8.3 Illapel Earthquake, Chile

## SUMMARY

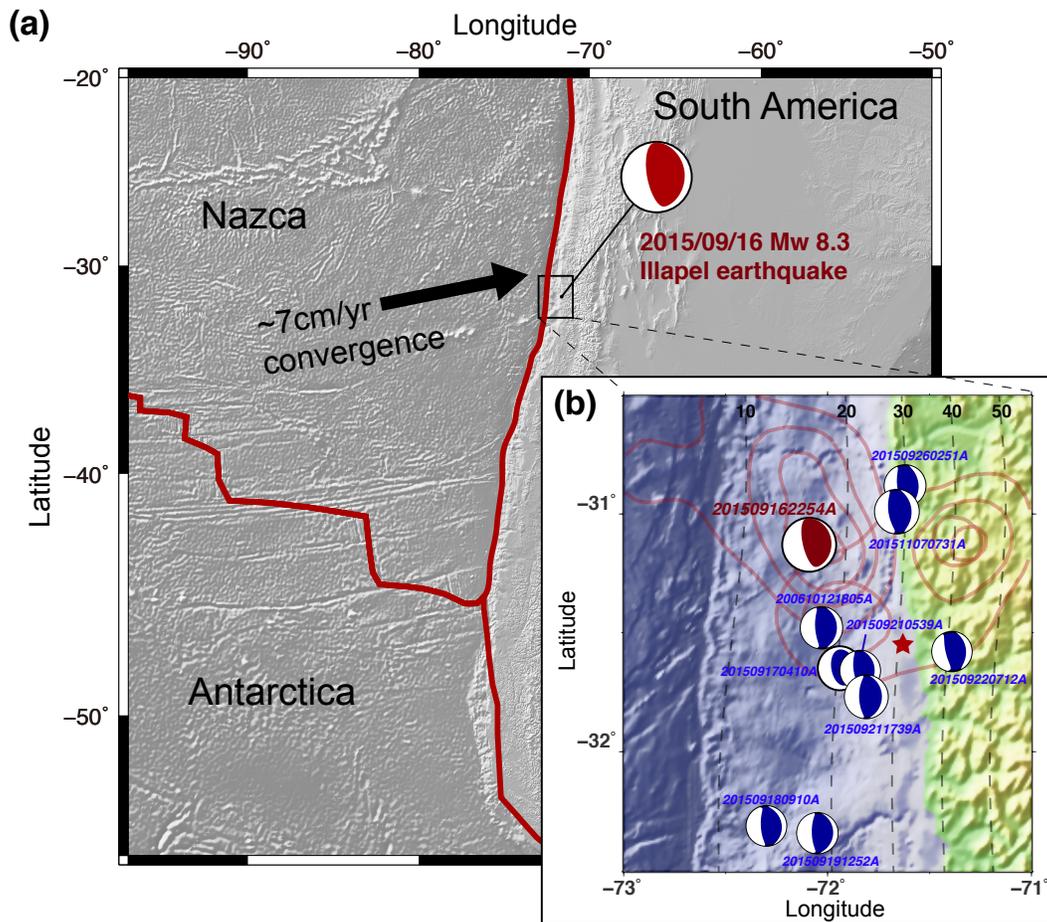
In this chapter, I apply the source analysis to the recent 2015  $M_w$  8.3 megathrust Illapel earthquake (Chile). The results from both techniques (ImCS-BP and spectral source analysis) are consistent and reveal a depth-varying rupture process, which I call the depth-frequency relation and is also found in other megathrust earthquakes. The low frequency content of the seismic radiation is located in the shallow part of the megathrust, propagating unilaterally from the epicenter towards the trench while most high frequency content comes from the downdip. Interpretation of the multiple stages of the radiation is also supported by the variations of both radiated energy and falloff rate during this earthquake. Finally, I suggest several possible mechanisms, either from pre-stress, fault geometry, and/or friction to explain my observables. My methodologies are an attempt to bridge kinematic observations with earthquake dynamics.

### 3.1 INTRODUCTION

On 16 September 2015, an  $M_w$  8.3 great earthquake occurred to the west of Illapel, Chile (Fig. 3.1). This earthquake resulted from the subduction of the Nazca Plate underneath the South American Plate with a convergence rate of 70 mm/yr (<https://earthquake.usgs.gov/earthquakes/eventpage/us20003k7a#region-info>, last accessed 2017/02/17), and caused locally a tsunami wave as high as 11 m<sup>205</sup>. The fault dimension was probably 200 km  $\times$  200 km and the maximum slip was found to be about 10 m by finite slip inversion of seismic data<sup>177,353</sup> and by joint inversion of seismic and tsunami waveforms<sup>117,178</sup>. For this earthquake, Yin et al.<sup>359</sup> used CS-BP of teleseismic P waves in two frequency bands and observed a clear dependence between the frequency and the locations of the energy bursts, which are interpreted to be caused by stress changes on the fault surface.

The southern end of the rupture area coincides with the subducting Juan Fernandez Ridge<sup>359</sup>, which probably prevented the rupture from propagating farther toward the south<sup>348</sup>. The rupture region of the 2015 Illapel event locates 50 km to the north of the 2010  $M_w$  8.8 Maule event, whose rupture area is adjacent to the 1960  $M_w$  9.5 earthquake, the largest earthquake ever recorded in human history. It has been suggested that the postseismic slip of 1960 event has increased the Coulomb stress of 1.3 MPa in the southern rupture region of the 2010 Maule earthquake, which probably triggered the 2010 Maule event<sup>79</sup>. Such spatial clustering of megathrust earthquakes in Chile poses a classical example of static triggering, similar to the observations along the Sumatra megathrust<sup>348</sup>.

Although the 2010  $M_w$  8.8 Maule and the 2007  $M_w$  8.0 Peru earthquakes have shown frequency-dependent coseismic radiation<sup>175,293,327,351</sup>, whether other parts of the South America subduction zone share the similar properties remains unknown. The occurrence of the  $M_w$  8.3 Illapel earthquake provides an opportunity to evaluate along-strike variation of properties in this subduction



**Figure 3.1:** (a) Tectonic setting of southern Chile and location of the 2015 Illapel, Chile Mw 8.3 earthquake. Red bold lines indicate the plate boundaries and the black arrow shows the relative motion between the Nazca Plate and the South American Plate. The rectangle highlights the location of the insert and the earthquake source region. (b) Zoom-in of the source region: focal mechanisms (from GCMT) of the Mw 8.3 mainshock (red) and other Mw>6 shocks (blue) selected as empirical Green's functions for this study. Red contours show the scaled backprojection energy distribution that I use to represent the source region. The red star shows location of the Mw 8.3 epicenter. Black dashed lines indicate the plate interface depth from the slab 1.0 model<sup>115</sup>.

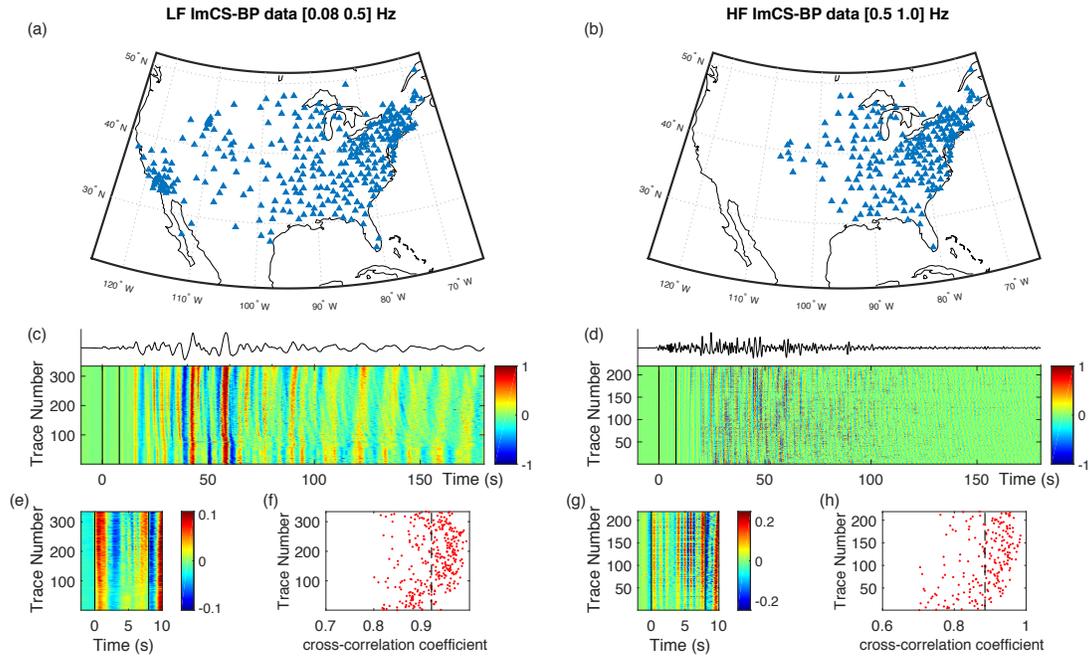
zone. In this study, I apply the source analysis methods, which are introduced in the previous chapter, to this  $M_w$  8.3 megathrust earthquake. I combine the results from two independent methods, that is, the Improved Compressive Sensing Back-Projection (ImCS-BP) and spectral source analysis, and illustrate a comprehensive rupture process during this megathrust earthquake. Finally, I suggest several possible mechanisms, either from pre-stress, fault geometry, and/or friction to explain my observables. Via this specific case study, I attempt to bridge the kinematic observations with earthquake dynamics.

### 3.2 APPLICATION OF IMCS-BP

I first apply the ImCS-BP method to the  $M_w$  8.3 Illapel earthquake. The ImCS-BP can recover the spatiotemporal evolution of the rupture process during this megathrust earthquake. The details about the methodology of this ImCS-BP technique is presented in the previous chapter. Here I focus on how to apply ImCS-BP to a specific earthquakes, and how to interpret the results from ImCS-BP.

#### 3.2.1 DATA FOR THE IMCS-BP

I apply my ImCS-BP method to the teleseismic P-wave velocity seismograms of the 2015  $M_w$  8.3 Illapel Earthquake recorded by the US Array stations in North America (TA array, data available using Wibler 3 of the Incorporated Research Institutions for Seismology Data Management Center, IRIS-DMC, [http://ds.iris.edu/wilber3/find\\_event](http://ds.iris.edu/wilber3/find_event)), selected for their high waveform coherency. I first pre-process the raw seismic data by removing means, trends and instrumental responses. Then I filter the data into two different frequency ranges: 0.05–0.5 Hz and 0.5–1 Hz, and align the waveforms via cross correlation for the first 8 s of the P waves, respectively (Fig. 3.2). Here I refer low frequency (LF) to 0.08–0.5 Hz and high frequency (HF) to 0.5–1 Hz for my CS re-



**Figure 3.2:** Data processing for the ImCS-BP. I used the data recorded by the USarray stations in North America (TA array shown in the maps (a) and (b), data available at the Incorporated Research Institutions for Seismology Data Management Center, IRIS-DMC). The aligned waveform data is shown in (c) and (d), the top waveforms are the stacked waveforms for comparison. Two solid lines in both (c) and (d) indicate the window within which the waveforms are used to do the cross-correlation. (e) – (h) show the zoom-ins of these cross-correlation windows as well as the final cross-correlation coefficients distribution. Dashed lines are the average cross-correlation coefficients.

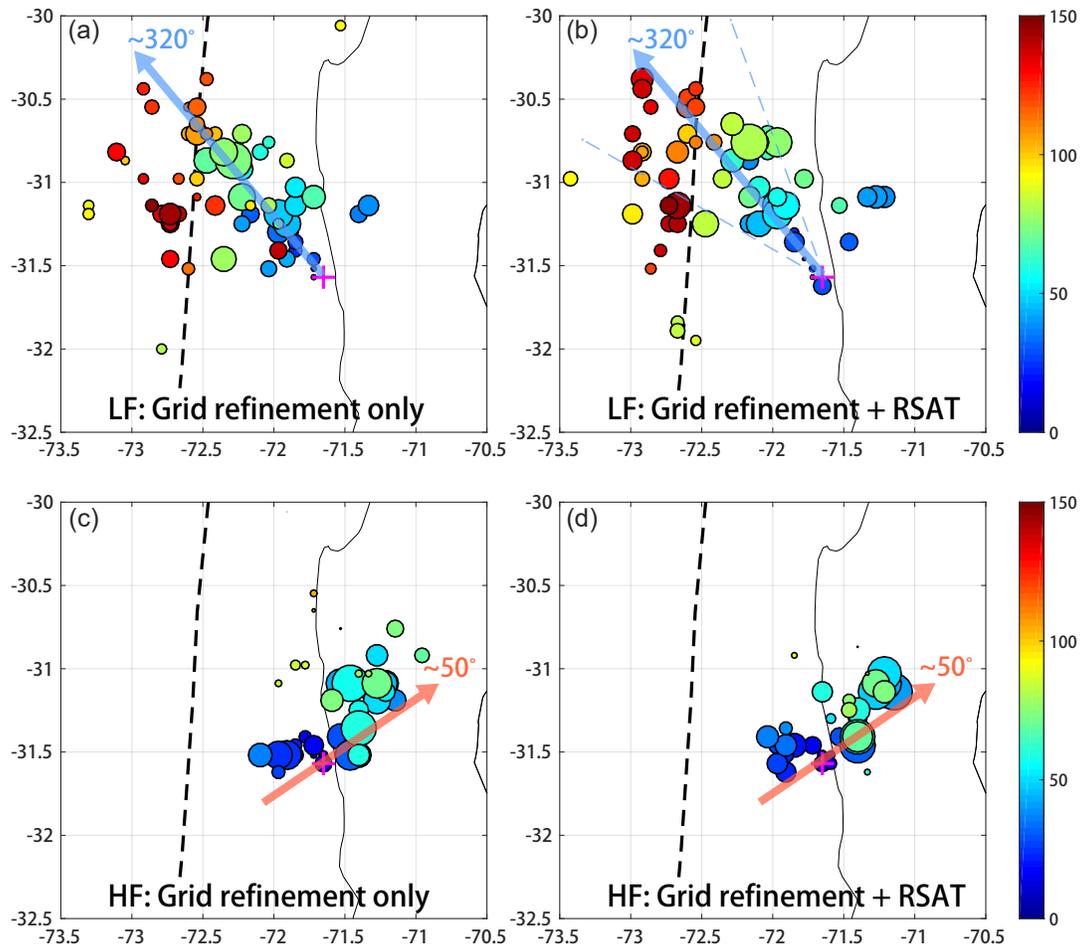
sults. Furthermore, I emphasize the meshing scheme and time-window adjustment improvements. I choose to solve for a potential source region of dimensions  $432 \text{ km} \times 432 \text{ km}$ . To apply the auto-adaptive source grid, I set the preliminary grid size to be  $48 \text{ km} \times 48 \text{ km}$  ( $9 \times 9 = 81$  grid points) and iteratively refine until reaching a grid size of  $6 \text{ km} \times 6 \text{ km}$ . I choose window lengths of  $14 \text{ s}$  for the  $0.08\text{--}0.5 \text{ Hz}$  low frequency (LF) band and of  $8 \text{ s}$  for the  $0.5\text{--}1 \text{ Hz}$  high frequency (HF) band. Sliding-window time adjustment technique is applied for each time window to suppress the waveform truncation effect.

### 3.2.2 EVOLUTION OF THE SEISMIC ENERGY BURSTS

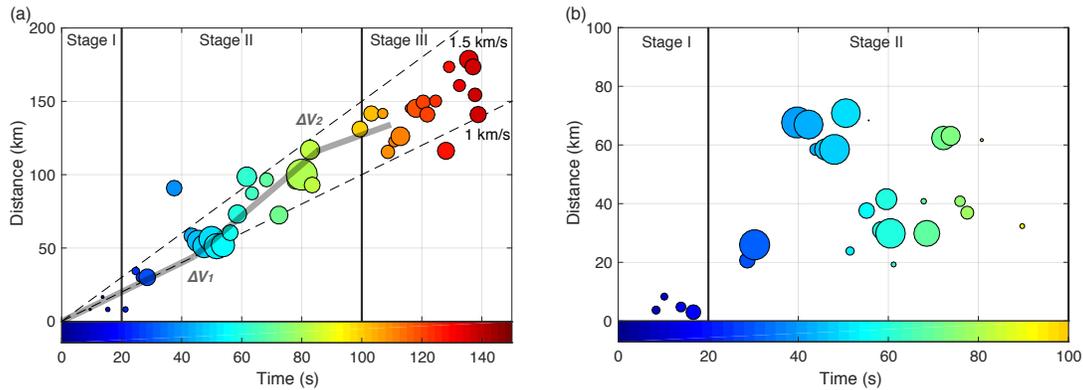
My ImCS-BP method provides the spatial and temporal evolution of the seismic energy bursts during the Illapel Earthquake. I show the results with: 1) auto-adaptive source grid only (Fig. 3.3 (a), (c)) and 2) auto-adaptive source grid and time adjustment techniques (Fig. 3.3 (b), (d)). While the overall patterns of the seismic energy distribution are similar in both cases, the results with time adjustment (Fig. 3.3 (b) and (d)) are less scattered, which I attribute to appropriate correction of the truncation effect. The most obvious improvement of the time adjustment technique appears in the later part of the rupture (from 80 s to 150 s) in the low-frequency band (Fig. 3.3 (a) and (b)). This is expected as the waveform truncation effect is most severe for the late stage of LF data (see Supplement 1, Fig. 3.2).

Given the evolution of the bursts with time, I refer to stage I as the first 20 seconds of the rupture, stage II as the time between 20 s and 100 s, and as stage III the time beyond that. The source locations of the low frequency (LF, 0.08-0.5 Hz) and the high frequency (HF, 0.5-1 Hz) bands appear to express different behaviors of the rupture (Fig. 3.3).

Most of the LF energy (Fig. 3.3 (a) and (b)) propagates unilaterally from the epicenter to a North-West direction (towards the trench) during stage I and II. Weak LF energy bursts also radiate in the North-East direction from the epicenter in stage I. In stage III, the LF energetic bursts are located well beyond the trench in the outer-rise region. In contrast, the HF bursts first propagate from the epicenter to the West for about 20 seconds (stage I) and then come to a pause. From 30 s to 70 s (early stage II), secondary HF pulses propagate to the downdip part of the megathrust, about 90 km North-East of the epicenter (Fig. 3.3 (c) and (d)). Inspecting the evolution of these two frequency bands may present a quite different signature of the evolution of the source: the LF radiation migrates updip (azimuth  $320^\circ$ , Fig. 3.3 (a) and (b)) while the HF radiation migrates downdip (azimuth  $50^\circ$ , Fig. 3.3 (c) and (d)). Results from both frequency bands suggest no southward rupture prop-



**Figure 3.3:** ImCS-BP results of the 2015 Mw 8.3 Illapel Earthquake. The black dashed lines show the trench. The arrows represent the assumed direction of propagation of the energy bursts. The solid circles indicate the energy bursts, their colors correspond to the time of the burst since the onset of the earthquake, their size corresponds to the amplitude power of energy bursts ( $|x(f)|^2$ ). Low frequency (0.08 - 0.5 Hz) results without (a) and with (b) time-adjustment. The azimuths  $300^\circ$  and  $340^\circ$  drawn in blue thin dashed lines are used to estimate a propagation velocity in Fig.3.4 (a). High frequency (0.5 - 1 Hz) results without (c) and with (d) time-adjustment.



**Figure 3.4:** Distance (y-axis) of energy bursts against time (x-axis) for (a) LF (Fig.3.3 (b)) and (b) HF (Fig.3.3 (d)). (a) I select the energy bursts between  $300^\circ$  and  $340^\circ$ , project on the axis defined from the epicenter to an azimuth of  $320^\circ$  (see arrow in Fig.3.3), and estimate the range between bursts and epicenter. (b) I select all bursts for HF results and show their range from the epicenter. The circles are sized proportionally to the amplitude of energy bursts, colored with time using the same scheme as in Fig.3.3.

agation. If the migration of the energy bursts represents a rupture propagation, the LF ImCS-BP results can also provide a rupture velocity estimate of this earthquake. Most of the energy bursts are within the azimuth range  $300^\circ$ – $340^\circ$  and I calculate a propagation speed from the distance against time plot shown in Fig. 3.4 (a). The average propagation speed is around  $1.0$ – $1.5$  km/s, which is much slower than a globally averaged shear wave speed of  $3.75$  km/s at the centroid depth (IASP91 model<sup>150</sup>). As for the HF energy bursts (Fig. 3.4 (b)), they are generally clustered along a  $50^\circ$  azimuth, but no estimate of propagation speed seems reliable.

### 3.3 APPLICATION OF SPECTRAL SOURCE ANALYSIS

My ImCS-BP method can provide relevant observational information about the location and timing of high-frequency bursts. However, as any backprojection-based result, the location of the bursts found by ImCS-BP really show P-waveform coherency. While some features of the earthquake source (e.g. minimum source dimension, rupture propagation direction, average rupture velocity, ...) can be assessed with ImCS-BP, they may not be sufficient to estimate dynamic source

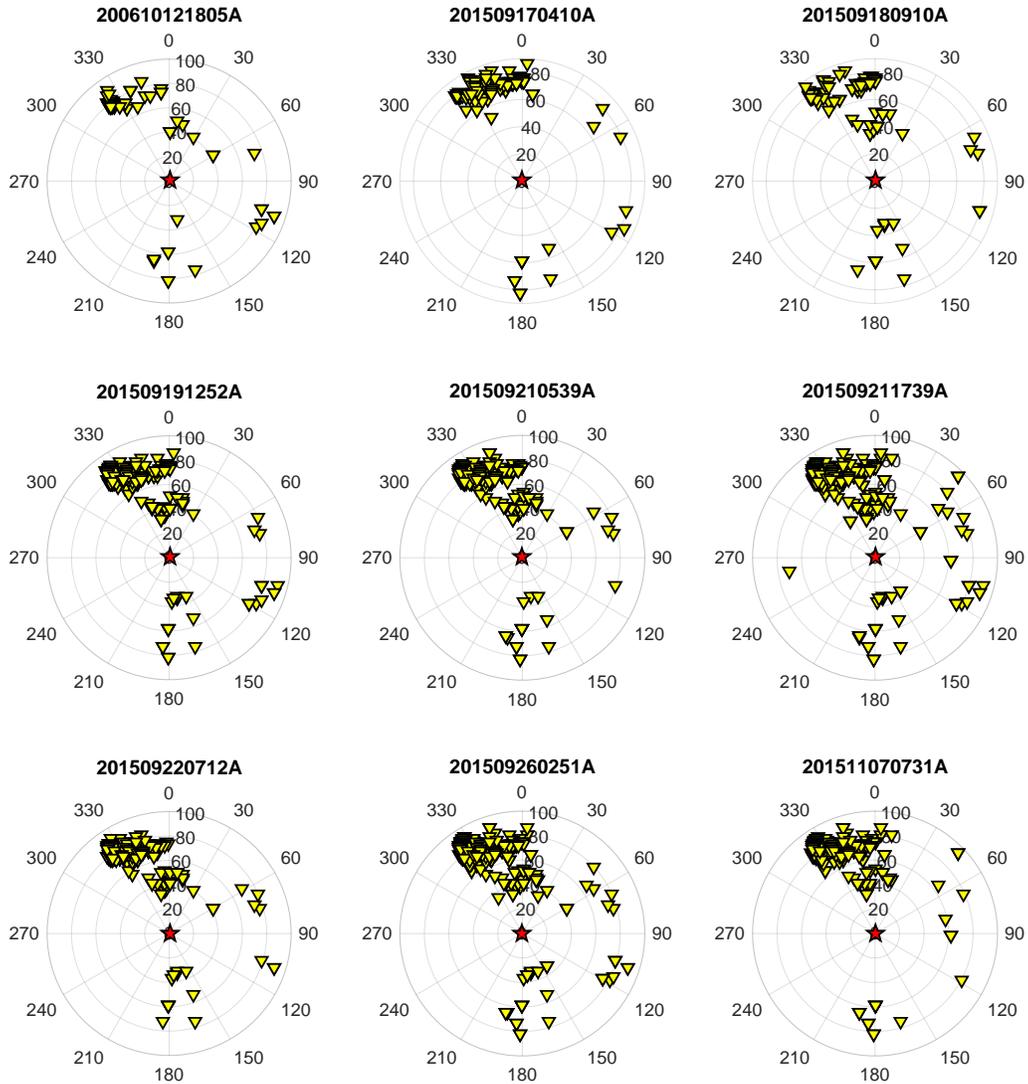
Event (GCMT ID)	Moment Magnitude	Seismic moment ( $10^{18}$ Nm)	Number of stations	Centroid depth (km)	Corrected depth (km)	Corner frequency (Hz)	Falloff rate n	Madariaga stress drop (MPa)
200610121805A	6.4	4.28	92	36.8	32.8	0.33	2.8	31.8
201509170410A	6.8	17.4	117	36	38	0.13	2.2	7.4
201509180910A	6.1	1.77	124	15	27	0.33	2.8	13.1
201509191252A	6.2	2.67	236	20	34	0.33	2.8	19.8
201509210539A	6.1	1.78	243	37	33	0.45	3.0	34.3
201509211739A	6.6	9.62	297	39	45	0.33	3.0	50.2
201509220712A	6.1	1.51	215	50	72	0.62	3.2	53.0
201509260251A	6.3	3.21	286	46	58	0.45	3.0	43.4
201511070731A	6.7	14.8	201	43.4	55.4	0.33	3.0	77.2

**Table 3.1:** Source Information of the Selected Mw 6 Earthquakes (the eGf 1Events). Gray-shading rows indicate the events without overlaying water layer.

parameters such as radiated energy. In order to provide additional observational constraints on the rupture dynamics of the Illapel Earthquake, I complement my back-projection analysis with a spectral analysis method<sup>75,76</sup>.

### 3.3.1 DATA FOR THE SOURCE SPECTRAL ANALYSIS

I choose 9 events of magnitude  $M_w > 6$  that occurred nearby the mainshock (Fig. 3.1 (b) and Table 3.1). I download teleseismic P-wave displacement data recorded at all stations located with  $30^\circ$  and  $90^\circ$  of angular distance available from the Federal Digital Seismographs Network (FDSN) (Fig. 3.5) using Obspy<sup>35</sup> (available at <https://github.com/obspy/obspy/wiki>). The pre-processing of the data consists in removing means, trends, instrumental responses, and integrating to displacement waveforms. I cut 300-seconds long waveform windows for the Illapel Earthquake, starting 15 s before and ending 285 s after theoretical direct P-wave arrival time (calculated by IASP91 model<sup>150</sup>) and I use a magnitude-dependent window length for the small shocks to capture mostly their source time function  $T_d \approx 10^{(M_w - 3.3)/2}$  s, e.g. 30 seconds for  $M_w$  6.1-6.3; 40 seconds for  $M_w$  6.4-6.6; 50 seconds for  $M_w$  6.7-6.8). I apply a Tukey window taper on 10 seconds on either end of the windows.



**Figure 3.5:** Station (yellow triangles) distributions (with azimuth and great circle distance) for each EGF event used in the source spectral analysis. Red stars show the location of Mw 8.3 main epicenter with respect to the stations.

In order to only select high-quality data, I first perform visual inspection, then calculate a signal-to-noise ratio (SNR), and decide on a SNR threshold. My noise estimate is taken in the window before the P arrival (using the same window length as that of the signal), I then Fourier transform both signal and noise windows, and define the SNR as the log of the ratio between the mean amplitude of the data and the noise in two frequency bands (0.02-0.5 Hz and 0.5-2 Hz). I select stations where the SNR exceeds 2 dB in both frequency bands (Fig. 3.6). The remaining data consist in a total of 1811 waveforms that provide a relatively good azimuthal coverage given the location of the main shock (southern Hemisphere) with respect to the distributions of global seismic stations (Fig. 3.5).

### 3.3.2 RESULTS OF TOTAL EVENT SOURCE SPECTRUM

Table 3.1 summarizes the best fit source depths and parameters necessary to construct  $S_1(f)$  for each eGf. I use these source models of small events (shown in Fig. 3.7) to apply the eGf method and retrieve the  $M_w$  8.3 mainshock source spectrum  $S_2(f)$ . Despite the differences in the small events source models (Fig. 3.7), the mainshock recovered source spectra all have very similar shape at various azimuths (Fig. 3.8). Because the corner frequency of large earthquakes ( $M_w > 8$ ) is lower than the low-frequency cutoff of my data, I only recover part of the source spectrum and cannot normalize the spectrum to the seismic moment (Fig. 3.7 and Fig. 3.8). However, unlike for small events,  $S_2(f)$  has the absolute level of amplitudes, i.e. it is scaled to the true seismic moment of the main shock through the eGf approach. Furthermore, I can use the independently estimate of seismic moment (GCMT catalog) as an additional constraint to calibrate the amplitude level of spectra at  $f \rightarrow 0$  Hz<sup>76</sup>.

At all azimuths, the recovered source spectra of the  $M_w$  8.3 event are complex at low frequencies 0.02-0.08 Hz. They are characterized by a steep drop in amplitude at 0.03-0.04 Hz (20-30 s), indicating a loss of low frequency energy (Eq. 2.12), pattern that I will address in the discussion section. At frequencies greater than 0.08 Hz, I test two parametric source models to fit the spectral shape.

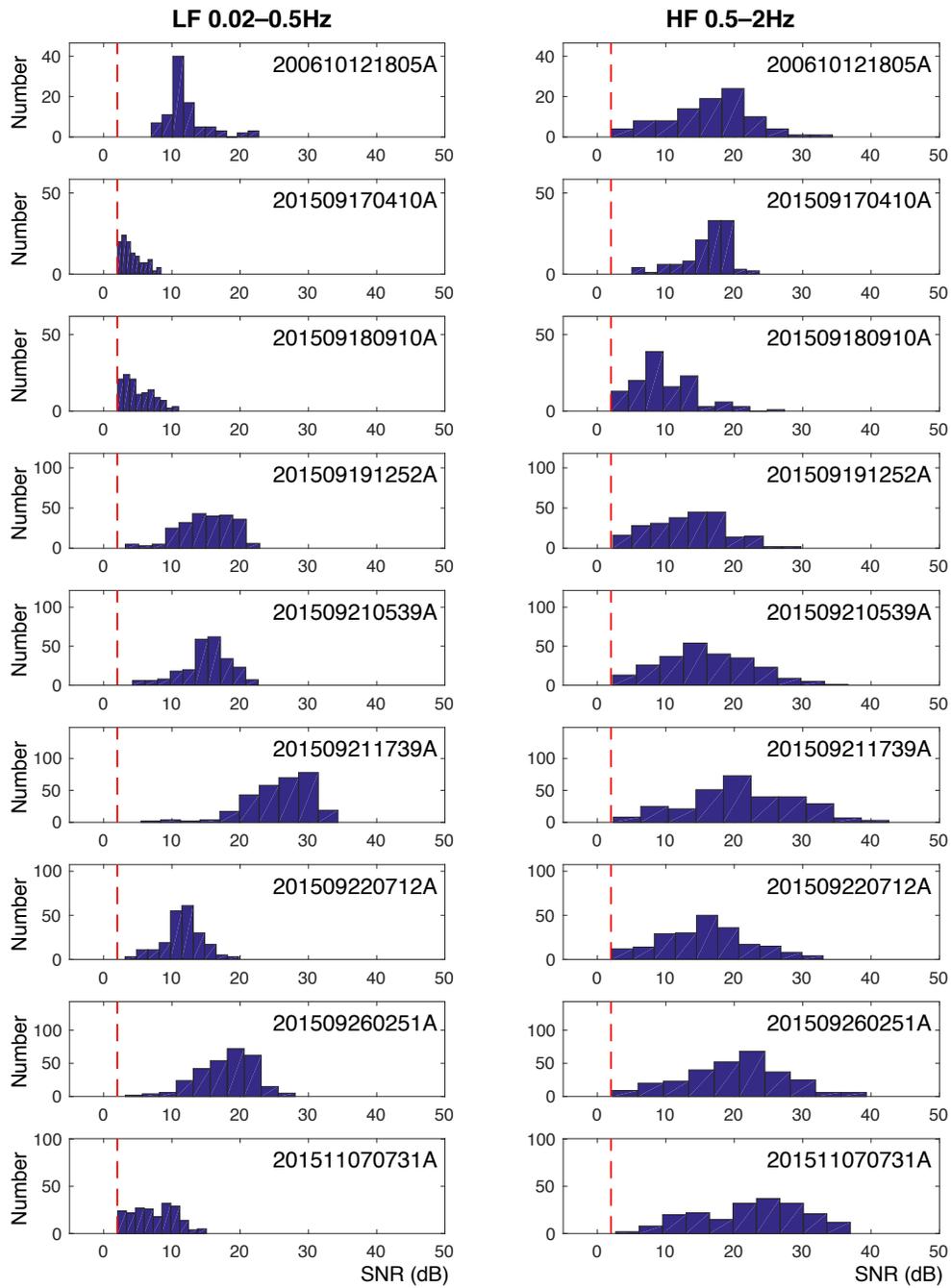
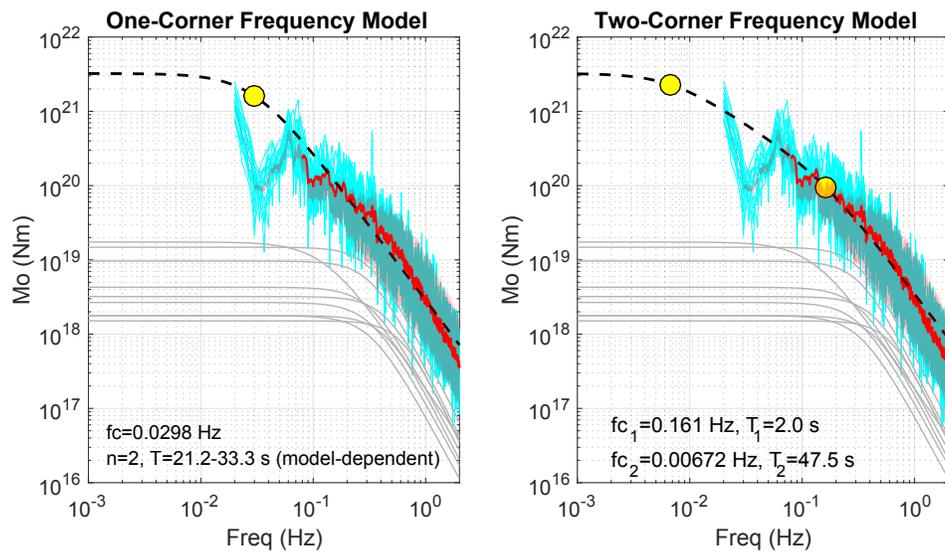
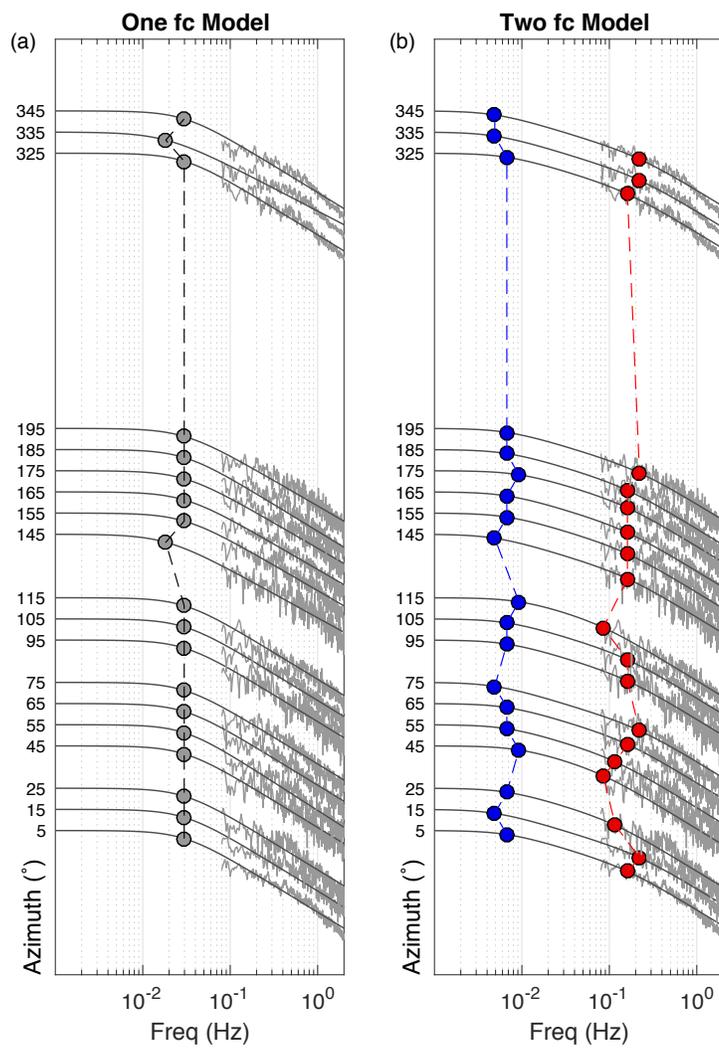


Figure 3.6: SNR distributions of my data for the spectral analysis. Red dashed lines indicate my SNR criterion of 2 dB. SNRs are estimated for each station and each event in low frequency (LF) and high frequency (HF) bands, respectively.



**Figure 3.7:** Source spectrum fitting for the Mw 8.3 Illapel Earthquake. Fitting with (a) a single-corner frequency model and (b) a double-corner frequency model. Cyan lines are the source spectra measured and averaged in each azimuth bin of 10°. Red lines are the stacked source spectra measurement in 0.08 Hz - 2 Hz, bandwidth that is used to fit with the spectral source models. Gray thin lines are the best-fit source spectral models of the small events. Black dashed lines indicate the best-fitted model with yellow circles to show their corner frequencies. Optimal source parameters (corner frequency, their associated duration, and high-frequency falloff rates) are also shown in the bottom of the panels.

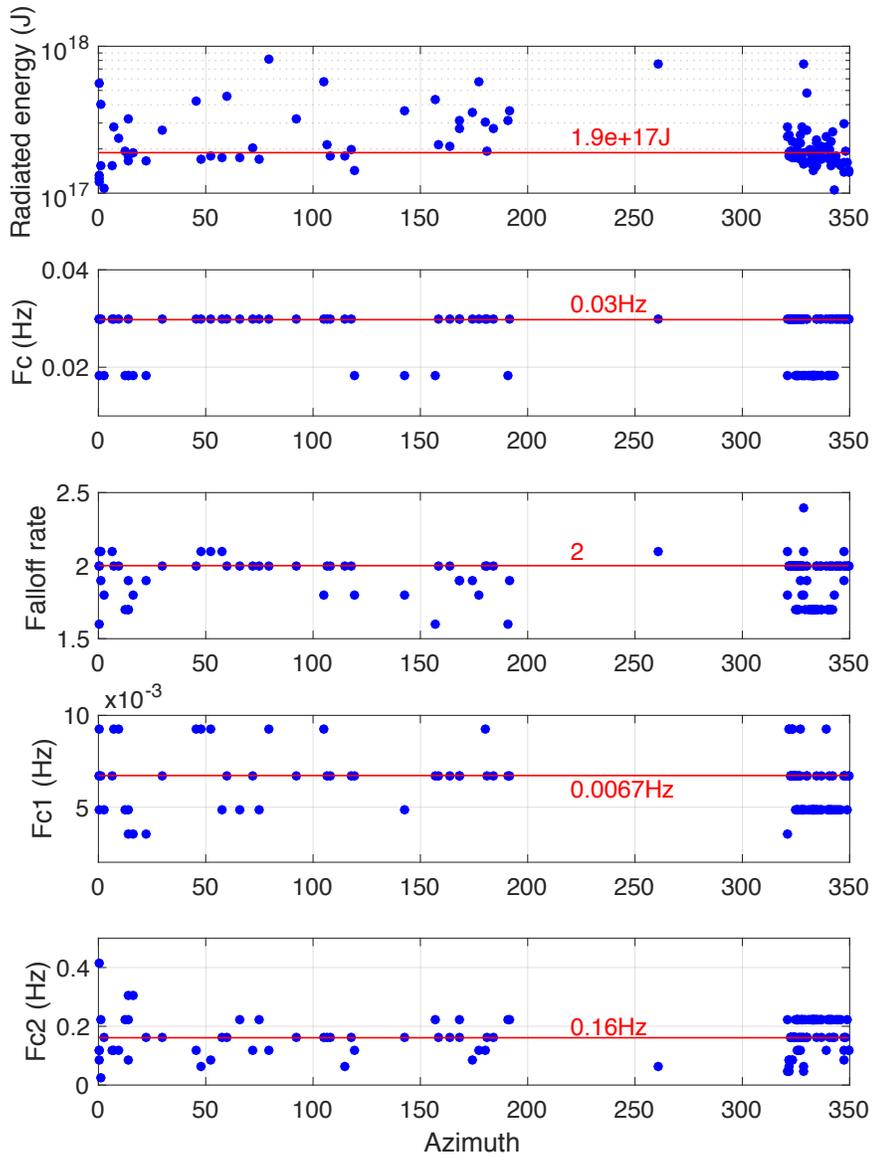


**Figure 3.8:** Spectral fitting for the total event source spectrum in each azimuth bin using single corner frequency model (a) as well as double corner frequency model (b). Circles indicate the best-fitting corner frequency of each spectrum line.

The first model is the same as Eq.2.7 with the seismic moment  $M_0$  of the main event. It is a “classic” single-corner frequency model<sup>6,45</sup> with variable falloff rate  $n$ . I find a best-fit corner frequency from this model is about 0.03 Hz and a falloff rate  $n$  of 2. The second model is a double-corner frequency model with two time scales proposed by Denolle & Shearer<sup>76</sup>:

$$S_2^{rec2}(f) = \frac{M_0}{\sqrt{[1 + (f/f_{c1})^2][1 + (f/f_{c2})^2]}} \quad (3.1)$$

The two corner frequencies from this double-corner frequency model are estimated to be 0.0067 Hz and 0.16 Hz. The duration estimate from the single-corner frequency model (Eq.2.7) likely underpredicts the true event duration. While I find no clear break of slope in the source spectra that would justify using the double-corner frequency model (Eq.3.1, Fig.3.7), it brings a more realistic value of source duration. I also show the spectra averaged over all eGfs and stations at each azimuth bin, their best fit models and source parameters in Fig.3.8, and summarize the azimuthal variations of the source parameters in Fig.3.9. Finally, by integrating the stacked spectrum using Eqs.2.12 - 2.13, I calculate the radiated energy of this  $M_w$  8.3 earthquake to be  $1.33 \times 10^{17}$  J. If I use Eq.3.1 to find the proportion of radiated energy below 0.08 Hz and that above 1 Hz, I find that my estimate is 74.3% of the total radiated energy and thus my value likely underestimates the one I would find if I had a broader frequency bandwidth. Nevertheless, my value is larger than that found by Ye et al.<sup>355</sup> ( $2.2 - 3.15 \times 10^{16}$  J) and by the IRIS products ( $3.21 \times 10^{16}$  J). My choice of elastic structure (wavespeed model) does not explain this difference. Therefore, I argue that the shape of the source spectrum, probably controlled by my choice of 3D Green’s function, can explain my overall higher estimate of radiated energy.



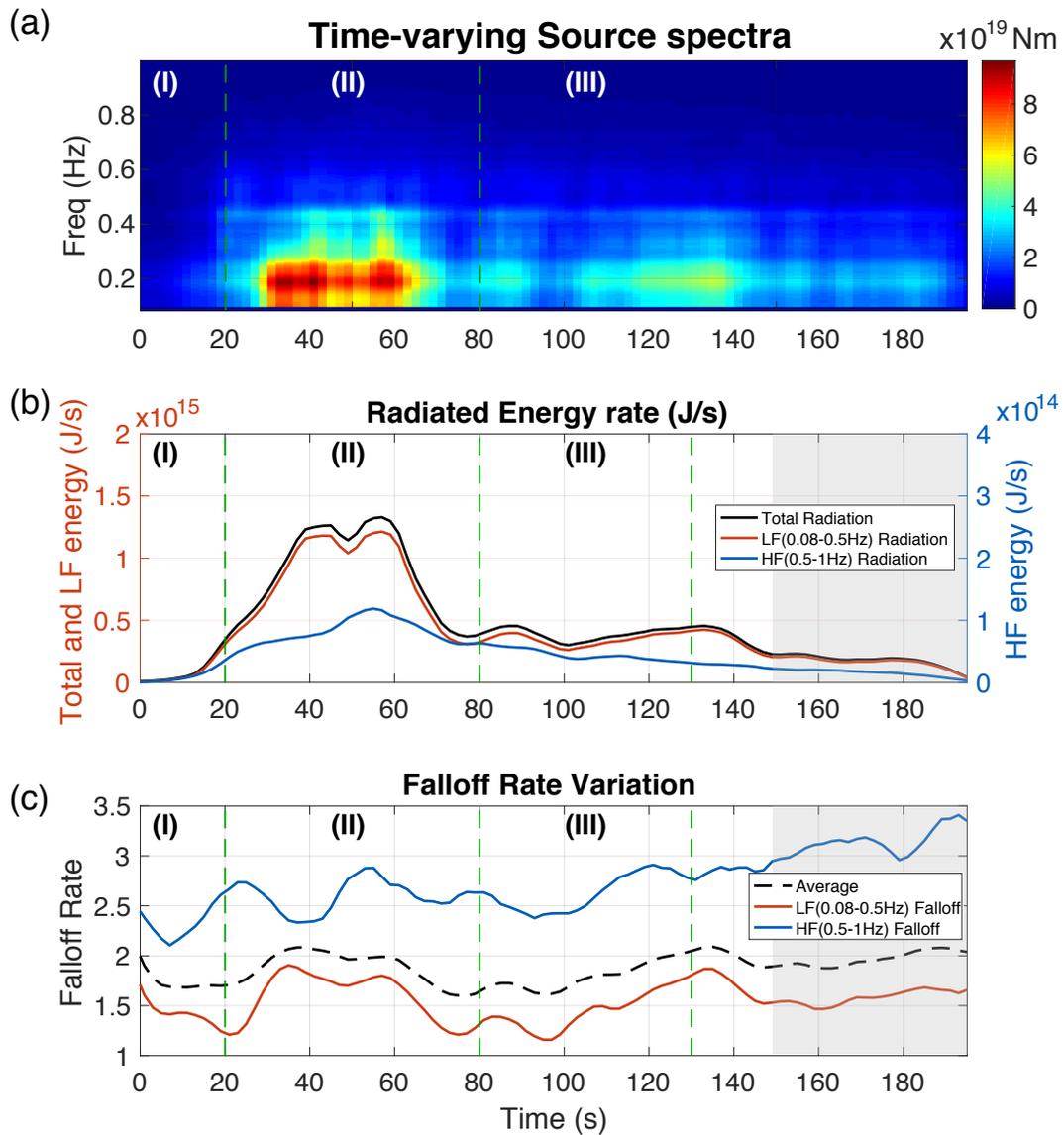
**Figure 3.9:** Source parameters estimated from the source spectrum at each station. Red lines show the median value for all the measurements. Radiated energy is calculated based on Eqs.2.12 - 2.13.  $f_c$  and falloff rate are from the single-corner frequency model Eq.2.7 fitting.  $f_{c1}$  and  $f_{c2}$  are the two corner frequencies from the double corner frequency model Eq.3.1 fitting.

### 3.3.3 RESULTS OF TIME VARYING SPECTRAL ANALYSIS

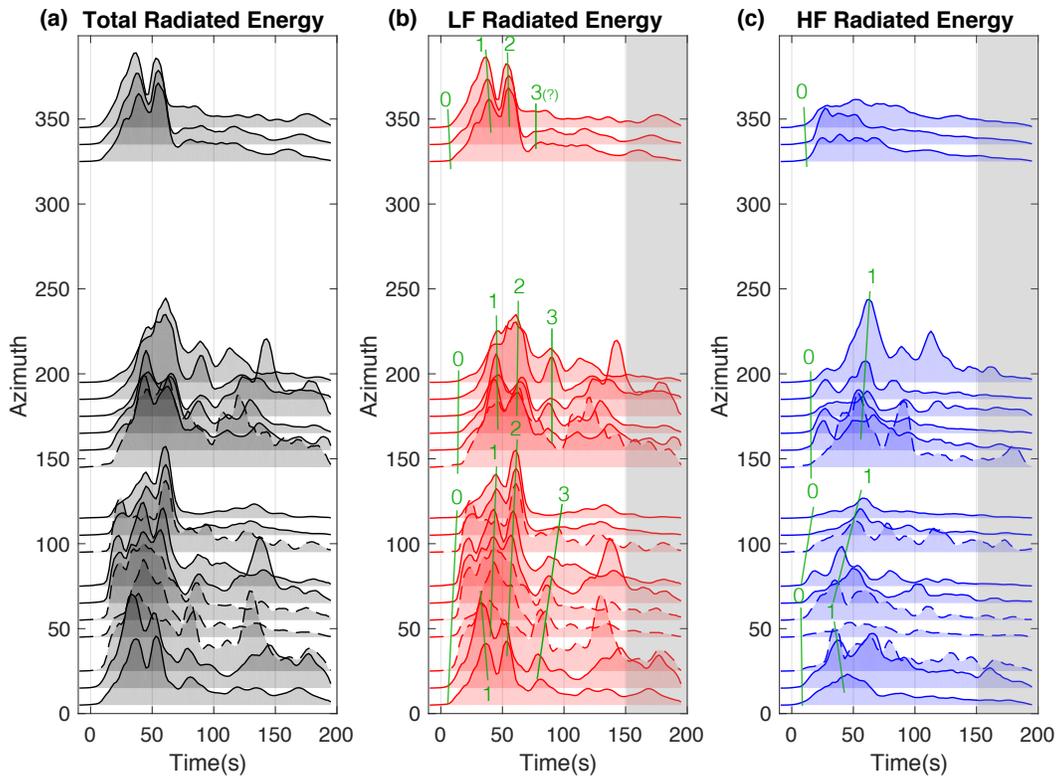
My temporal spectral analysis provides a unique view on the evolution in time of the source parameters. I show the time-varying radiated energy and the falloff rate in Fig. 3.10 within the same frequency bands of LF and HF as the previous ImCS-BP results. I limit my interpretation to the first 150 seconds, as the later phases, shown by the ImCS-BP results, are located in the outer-rise region and are probably not associated with the rupture process (Fig. 3.10 and Fig. 3.11).

The radiated energy slowly increases in the stage I from the earthquake onset to 10 s after, which probably reflects the earthquake nucleation process. Most of the seismic energy is released during the stage II between 20 s and 100 s and presents two pronounced peaks at about 40 s and 60 s (Fig. 3.10 (a)). After 100 s, the total radiated energy stabilizes and finally dies until 200 s after the onset. Variations in radiated energy in the LF band are quite abrupt. Because the LF radiation is very similar to the total radiation, the LF energy dominates (probably from being closer to the corner frequency of the main event) and the two most energetic peaks are relatively low-frequency energy. This compares well with my backprojection results (Fig. 3.3 (a) and (b)). Variations in radiated energy in the HF band are smoother than those of the LF band. The overall HF energy steeply increases at about 20 s, peaks at 50 - 60 s, and dies after that. This is very consistent with my HF backprojection results (Fig. 3.3 (c) and (d)).

The patterns of time-varying falloff rates from my observations are intriguing (Fig. 3.10 (b)). Because of my choice of regression, LF and HF falloff rates are independent measures of spectral slope. It is important to remember that a low falloff rate is equivalent to a high ratio of high-frequency over low-frequency radiation. I show the variations of falloff rate from the limited frequency bandwidths LF, HF, and from the full bandwidth (called average falloff rate here). In the first stage, a general decrease appears at all frequency bands, indicating an increase in high-frequency content, which is well timed with the nucleation inferred from radiated energy (Fig. 3.10 (a)).



**Figure 3.10:** Time-varying radiated energy and falloff rate. In both panels, green dashed lines and the roman numbers indicate the timings of the stages, 0–20 s, 20–100 s, and 100–200 s. Black lines are averaged values, blue lines are the LF, and red lines are the HF results. Results after 150 s are shaded for not being representative of the evolution of the source. (a) Evolution of radiated energy with time: left y-axis is for the total and LF radiated energy while the right y-axis is for the HF energy. (b) Evolution of falloff rates with time with similar color scheme as in (a).



**Figure 3.11:** Time-varying radiated energy estimated at  $10^\circ$  azimuth bin: (a) total radiated energy, (b) LF radiated energy, and (c) HF radiated energy. Green thin lines and numbering indicate reference and indexing for the coherent radiated energy peaks. Results after 150 s are shaded for being less reliably interpretable to source processes.

At the beginning of stage II, LF and HF falloff rates appear to be anticorrelated. The LF falloff reaches a minimum at 22 s and peaks at 35 s during the overall steep growth in LF radiated energy. In contrast, HF falloff peaks at 24 s, which also corresponds well to the rapid growth in HF energy. Both falloffs are similar at 40 s (peak of seismic radiation), favoring a smooth spectral shape with a single falloff. After 50 s and during stage III, both LF and HF falloff rates seem to vary together until the end of the rupture. Additionally, while there is no obvious correlation between HF falloff and HF radiated energy, LF falloff rate and LF radiated energy share similar variations. If there is no systematic correlation between radiated energy and falloff rates, understanding the contributions of low vs high frequencies during the rupture and to the radiation of seismic wave is a metric that can be tested with dynamic source simulations.

I further detail the evolution of my three metrics of radiated energy (total, LF, and HF) by exploring their variations with azimuth. Fig. 3.11 shows the time varying radiated energies averaged over  $10^\circ$  azimuth bins that have at least 10 stations per bin. Unfortunately, there is a clear lack of data coverage at azimuths where I expect the source directivity. The features common to all azimuths are two energy peaks 1 and 2 in the LF and broadband time series, consistent with what I find previously (Fig. 3.7 (a) and (c); Fig. 3.10 (a)). The difference in time interval between the peaks are indicators of source directivity, somewhat in a similar way to source time function analysis.

Focusing on the LF radiation, there are at least 3 coherent energy peaks (labeled 1, 2 and 3 in Fig. 3.11 (b)). The time interval between 0 and peak 1 is shorter at azimuth  $30^\circ$ , corresponding to the northeastward propagation of the rupture observed by the backprojection results (Fig. 3.7 (a) and (c)). The time interval between peak 1 and 3 around azimuths  $320^\circ - 350^\circ$  corresponds well to the northwestward rupture in the LF backprojection results. It is, however, not well constrained because of a data gap between  $200^\circ - 320^\circ$  and the energy peak 3 is not clearly observed by stations within  $320^\circ - 350^\circ$ .

In contrast to the LF radiation, the HF radiation is simple. I estimate only a single coherent peak

(labeled as 1) and the directivity towards azimuth  $50^\circ$  is obvious, which is also supported by back-projection results (Fig.3.7 (b) and (d)).

### 3.4 INTEGRATED SOURCE ANALYSIS FOR THE 2015 ILLAPEL EARTHQUAKE

In stage I of this earthquake (0 - 20 s), ImCS-BP results show that both LF and HF energy bursts cluster around the epicenter in the first 10 seconds (Fig.3.3). During this period, the time-varying spectral analysis results present a similar slow and weak increase of radiated energy (Fig.3.10 (a)) and the falloff rate decreases from 2 to 1.7, indicating a growth in the proportion of high frequencies (Fig.3.10 (b)). During this stage, the velocity curves provided by the LF energy bursts may also infer a slow growth with, if interpreted as such, earthquake rupture velocities under 1 km/s (see Fig.3.4). This stage can characterize a slow initiation or nucleation.

As mentioned in the previous section, I define stage II as the period of time between 20 s and 100 s. My integrated analysis during this stage brings up many aspects of this rupture. Based on the LF ImCS-BP results, there are two distinct energy bursts North-West to the epicenter around 45 - 60 s and 60 - 80 s after onset (Fig.3.3 (a) and (b), Fig.3.4 (a)). They correspond to the largest amplitudes of the velocity waveforms at about 40 s and 60 s (Fig.3.2) and to the two radiated energy peaks from my time-dependent spectral analysis results (Fig.3.10 (a) and Fig.3.11 (b)).

My integrated results show that most of the seismic radiated energy is released between 40 s and 80 s and their location is implied by the LF energy bursts in Fig.3.3 (b). Furthermore, the LF ImCS-BP results may indicate two changes in rupture velocity around 50 s and 80 s (Fig.3.4 (a)). These velocity changes are well timed with the peaks of the backprojection energy bursts and radiated energy from spectral analysis. This supports that changes in rupture velocity excite high-frequency radiation<sup>193</sup>. Yin et al.<sup>359</sup> compared the CS-BP results with various slip inversion models for this earthquake and highlighted that most of the LF energy bursts were located within the high slip

regions, which are common to most slip models<sup>117,353</sup>. Melgar et al.<sup>205</sup> also showed that their low frequency backprojection results consistently overlapped with the large slip regions, indicating a relation between low frequency backprojection results and coseismic slip.

There are about 10 s time difference between the peaks of ImCS-BP energy burst (peaks at about 50 and 80 s, Fig. 3.4 (a)) and the peaks of the time-varying source spectral analysis (peaks at about 40 and 60 s, Fig. 3.10 (a)). As the rupture approaches the seismic array (TA Array in this study), the seismic waves arrive earlier because of the shorter traveling distance. If left ignored, the rupture propagation will bring some errors into the timing of the source (Fig. 3.10). However, the ImCS-BP method can correct these source time shifts/errors and provide more accurate source time estimations. I refer to Yin & Yao<sup>360</sup> for the details and test on this correction.

The evolution of radiated energy (with time and with azimuth (Fig. 3.11 (b))) supports the directivity effects inferred from my backprojection results: the time interval between the onset of radiation (0) and the first energy peak (1) shortens at azimuths  $30^\circ - 50^\circ$ , implying a probable direction of the rupture propagation similar to that found by the ImCS-BP (Fig. 3.3 (a) and (b)). Although the ImCS-BP results indicate a clear northwestern propagation towards the trench (azimuth of about  $320^\circ$ ) from 30 s to 80 s (Fig. 3.3 (a) and (b)), there is no station available in the azimuth ranges of  $200^\circ - 320^\circ$  and  $120^\circ - 140^\circ$  to support directivity with the spectral analysis. This justifies well the use of backprojection techniques as a complement to spectral analysis to evaluate source directivity. In contrast to the low frequency radiation, the HF results from ImCS-BP show a propagation directed to the North-East (azimuth  $30^\circ - 50^\circ$ ), with a downdip unilateral propagation and a peak of coherence at 40-60 s. Both radiation peak time as well as likely source directivity are well consistent with the time-varying spectral analysis in the HF (Fig. 3.10 (a) and Fig. 3.11 (c)).

Supported by both backprojection and spectral analyses, the main characteristic of the seismic radiation of this earthquake is the obvious discrepancy in the evolution of the low and high frequency content during the rupture. They may represent two processes of dynamic rupture characterized

by different time scales (inverse of frequency) at the updip and at the downdip part of the megathrust. The different rupture behaviors at updip and downdip regions of this event have also been observed by Melgar et al. <sup>205</sup> for the two resolved largest slip patches. They interpret the shallow part to be similar to a tsunami earthquake <sup>143</sup>, whereas the deeper portion would behave like a “normal” source.

Finally, I ought to discuss during the stage III, beyond 100 s into the rupture. Lee et al. <sup>177</sup> suggest that this event is a very long lasting earthquake (duration  $\sim 250$  s), as a combination of a megathrust earthquake that propagates updip in the first 150 s with a tsunami earthquake that propagates along the trench with a slow rupture speed from 150 s to 250 s, which they propose to be equivalent to a  $M_w$  8.08 earthquake. However, the ImCS-BP low-frequency energy bursts are clearly located on the outer-rise region (Fig. 3.3 (a) and (b)) rather than along the trench. Moreover, the dominant period of the coherent signal is roughly 10 s, which is realistic for a two-way travel time in a 7.5-km thick water column with a sound speed of 1.5 km/s. Therefore, these bursts likely arise from reverberations in the water column, and compared with the previous 2 stages, this stage III does not seem to be a real physical stage for the rupture but rather a wave propagation effects in the outer-rise region.

### 3.5 RUPTURE DYNAMIC CONSTRAINTS INFERRED FROM SPATIAL AND TEMPORAL SOURCE ANALYSIS

#### 3.5.1 DEPTH-FREQUENCY RELATION

The spatiotemporal distribution of seismic energy bursts from ImCS-BP potentially brings insight on physical properties on the fault interface. Many backprojection studies of the large subduction-zone thrust earthquakes, such as the  $M_w$  9.3 2004 and  $M_w$  8.8 2005 Sumatra, the  $M_w$  9.0 2011 Japan, the  $M_w$  8.8 2010 and  $M_w$  8.3 2015 southern Chile <sup>175,324,349,351,359</sup>, the  $M_w$  8.0 2007 Peru <sup>293</sup>,

and in large continental subduction zone such as the  $M_w$  7.9 2015 Nepal<sup>364,361</sup> reveal clear systematic differences between the frequency content of the updip and downdip regions. My results from ImCS-BP are consistent with previous studies that identify LF radiation in the updip part and HF radiation in the downdip region of the megathrust earthquakes.

The depth-dependence of seismic radiation is also observed from small and moderate shallow (< 50 km) earthquakes on the megathrust.<sup>121</sup> found a trend in source time functions that, when adjusted to the seismic moment scaling with source duration, shorten as source depth increases, implying a higher frequency content. Denolle & Shearer<sup>76</sup> find a similar trend in Japan and in Sumatra for moderate and large events, trend also confirmed by Ko & Kuo<sup>161</sup> for a wide range of earthquake magnitudes in the Japan trench. Bilek & Lay<sup>36</sup> interpret this shortening of source pulses as a systematic variation in rigidity with depth. Furthermore, studies on the generation of high-frequency strong ground motion simplify the source of megathrust earthquakes to localized strong motion generation areas, that often locate downdip of the fault<sup>19,265</sup>.

Regarding any systematic changes in high-frequency falloff rate with depth, Ye et al.<sup>355</sup> argue for a decrease in high-frequency falloff rates with depth for  $M_w$  7+ in subduction zones, further suggesting a high ratio of high-frequency over low-frequency radiation for the deeper parts of the megathrust. However, Ko & Kuo<sup>161</sup> find in the Japan trench a relatively constant radiation efficiency with depth, and given the observations of increase stress drop with depth, the high-frequency falloff rates should at most remain constant with depth.

## EARTHQUAKE DYNAMICS: GEOMETRY VS FRICTION PROPERTIES

The first mechanism that can explain depth-dependence in seismic radiation is the stress conditions on the fault before the rupture. Pre-stress and its heterogeneity are particularly difficult to establish, as multi-cycle earthquakes simulations require both long-term fault healing and short term dynamics<sup>84,85</sup>. The frequency dependence of the seismic radiation during the Tohoku earthquake has been

explained in Huang et al. <sup>129, 126</sup> by dynamic simulations though heterogeneous distribution pre-stress on the fault surface.

A second mechanism that can explain the depth-dependence in the frequency of seismic radiation is the frictional resistance on the fault interface. The phase transformation of the minerals constituting the near fault fabric alters their frictional properties <sup>175,277,351</sup>. Parametrization of friction, either through velocity-<sup>77,78,259,264</sup> or slip-weakening <sup>16,132,242</sup> mechanisms, is often invoked through a characteristic or critical length scale that can be tuned <sup>129,126</sup> to simulate high-frequency radiation. Often the shallowest part of the plate interface is considered unfavorable to earthquake nucleation <sup>199,267</sup>, but other dynamic weakening mechanisms (flash heating, thermal pressurization <sup>234,258,318</sup> may allow propagation through those areas.

A third, but not least, mechanism that can generate a depth-varying seismic radiation during the rupture is the geometry of the fault. The hanging wall, or accretionary wedge in subduction zone, undergoes large dynamic stresses during the rupture <sup>47,167,188,231,237</sup>. The materials in the accretionary wedge may thus respond inelastically to strong dynamic stresses, an effect that Ma & Hira-kawa <sup>190</sup> proposed to exaggerate the depth-dependence to seismic radiation by reducing the high-frequency radiation at the trench. The accretionary prism, if described as inelastic, may act as an “energy sink”, reducing the directivity effects in seismic radiation and absorbing the high-frequency radiation at the trench. Wedge failure may explain my observations.

#### PULSE-LIKE VS CRACK-LIKE

In results from both backprojection and spectral analysis, the LF and HF seismic radiation differ in location of excitation and in direction of propagation, behaving as if they occurred due to two different rupture processes. My observations suggest a unilateral propagation of the seismic radiation, either updip with the LF radiation, or downdip with the HF radiation. If interpreted independently, both measures would favor interpretations of unilateral pulse-like ruptures, such as

that of the  $M_w$  7.9 2015 Nepal Earthquake<sup>21,94,102,105,361</sup>. However, I only observe a single dominant time scale in the whole event spectrum (the source duration), which favors a crack-like model, as against a pulse-like model that is dominated by the source duration and the rise time<sup>76,113</sup>. Because of this contrasting behavior of the radiation, it is difficult to assess whether the rupture had a more pulse-like behavior (such as a typical Haskell model) or a crack-like behavior. One possibility is an asymmetry in the slip-rate function at the rupture front where, for the propagation of a simple crack front, the updip slip-rate function would be smoother (deficient in high frequencies) than the downdip slip-rate function.

### 3.5.2 TSUNAMI GENESIS

My final discussion point is the spectral shape of the main event at low frequencies. My strict SNR criterion at low frequency provides me confidence in the peculiar trough in the spectrum at 0.035 Hz, which made me ignore the 0.02 – 0.08 Hz frequency band in my spectral model fitting. This type of feature is similar to that observed during tsunami earthquakes, e.g., 1994 Java<sup>2</sup>; 2006 Java<sup>13</sup>; 2011 Mentawai<sup>174</sup>; 2006 Java and 2010 Mentawai<sup>76</sup> earthquakes. While the tsunami generated by the Illapel event reached locally heights up to 11 m<sup>205</sup>, it is possible that the spectral shape reflects some features common to tsunami earthquakes. Currently and to the best of my knowledge, there exist no physics that explains such a notch at long periods in tsunami earthquake P-wave spectra.

### 3.6 CONCLUSION

I applied both methods to the  $M_w$  8.3 2015 Illapel Earthquake and found a distinct evolution of the low frequency radiation compared to the high-frequency radiation. Low frequency radiation is not surprisingly more prominent in large earthquakes and appears consistently shallower than high-frequency radiation.

I hope that my combined approach to image the seismic radiation during the earthquake provides relevant metrics to bridge kinematic observations and dynamic models. Understanding the physical mechanisms that can explain my observations is central to improving my knowledge of large megathrust earthquakes. Different mechanisms (friction, pre-stress, geometry) can produce variation with depth of the frequency content of seismic radiation. Determining whether one mechanism dominates over the others deserves to be addressed.

### 3.7 DATA AND RESOURCES

The  $t^*$  model used in this study is from Linda M. Warren and Peter M. Shearer. The USArray seismic data for the ImCS-BP results are downloaded from IRIS-DMC (Incorporated Research Institutions for Seismology Data Management Center). Networks DOIs are the following: AZ: [doi:10.7914/SN/AZ](https://doi.org/10.7914/SN/AZ); BK: [doi:10.9732/BDSN](https://doi.org/10.9732/BDSN); CI: [doi:10.7914/SN/CI](https://doi.org/10.7914/SN/CI); CN; II: [doi:10.7914/SN/II](https://doi.org/10.7914/SN/II); IM; IU: [doi:10.7914/SN/IU](https://doi.org/10.7914/SN/IU); IW: [doi:10.7914/SN/IW](https://doi.org/10.7914/SN/IW); LD; MX; N4: [doi:10.7914/SN/N4](https://doi.org/10.7914/SN/N4); NM; OK: [doi:10.7914/SN/OK](https://doi.org/10.7914/SN/OK); PB; PE: [doi:10.7914/SN/PE](https://doi.org/10.7914/SN/PE); PY: [doi:10.7914/SN/PY](https://doi.org/10.7914/SN/PY); SC; TA: [doi:10.7914/SN/TA](https://doi.org/10.7914/SN/TA); US: [doi:10.7914/SN/US](https://doi.org/10.7914/SN/US); UU: [doi:10.7914/SN/UU](https://doi.org/10.7914/SN/UU); YN: [doi:10.7914/SN/YN\\_2010](https://doi.org/10.7914/SN/YN_2010)) using Wibler 3. The global data for the spectral analysis are from the Federal Digital Seismographs Network (FDSN). Networks DOIs are as follows: AF: [doi:10.7914/SN/AF](https://doi.org/10.7914/SN/AF); AI: [doi:10.7914/SN/AI](https://doi.org/10.7914/SN/AI); AU; AZ: [doi:10.7914/SN/AZ](https://doi.org/10.7914/SN/AZ); BK: [doi:10.9732/BDSN](https://doi.org/10.9732/BDSN); CI; CC: [doi:10.7914/SN/CC](https://doi.org/10.7914/SN/CC); CI: [doi:10.7914/SN/CI](https://doi.org/10.7914/SN/CI); CM; CN; CU: [doi:10.7914/SN/CU](https://doi.org/10.7914/SN/CU); G: [doi:10.18715/GEOSCOPE.G](https://doi.org/10.18715/GEOSCOPE.G); GE: [doi:10.14470/TR560404](https://doi.org/10.14470/TR560404); GT: [doi:10.7914/SN/GT](https://doi.org/10.7914/SN/GT); II: [doi:10.7914/SN/II](https://doi.org/10.7914/SN/II); IM; IU: [doi:10.7914/SN/IU](https://doi.org/10.7914/SN/IU); IW: [doi:10.7914/SN/IW](https://doi.org/10.7914/SN/IW); MB: [doi:10.7914/SN/MB](https://doi.org/10.7914/SN/MB); MX; TA: [doi:10.7914/SN/TA](https://doi.org/10.7914/SN/TA); TR) using Obspy (available at <https://github.com/obspy/obspy/wiki>). Most of the data processing, calculations, and figure plottings are carried out using Matlab. Some figures in the manuscript are also plotted using GMT (the Generic Mapping Tools, <http://gmt.soest.hawaii.edu/projects/gmt/>

[wiki/Download](#)).

# 4

## Relating teleseismic backprojection images to earthquake kinematics

## SUMMARY

Backprojection (BP) of teleseismic P waves is a powerful tool to study the evolution of seismic radiation of large earthquakes. The common interpretations on the BP results are qualitative comparisons with earthquake kinematic observations, such as the evolution of slip on the fault and rupture velocity. However, the direct relation between the BP images and physical properties of the earthquake rupture process remains unclear and is needed for further application of this technique. In this study, I start from a theoretical formulation of the BP images, which is linear in the frequency domain, and carry on a synthetic exercise with kinematic source representations and virtual receivers embedded in a homogeneous fullspace. I find that the fundamental linear formulation of the BP method is most correlated with the true kinematic source properties: in frequency domain the BP images are proportional to the images of slip motion through a scaling matrix  $\mathbf{F}(\omega)$  that accounts for radiation pattern and source-receiver geometry and that acts as a spatial smoothing operator. Overall, the synthetic BP images match relatively well the kinematic models and my exercise validates that the BP image can be directly used to track the spatio-temporal propagation of rupture front. However, because  $\mathbf{F}(\omega)$  is not strictly an identity matrix due to limited station coverage in space (azimuth and distance) and to the limited frequency bands of the seismograms, it remains difficult to recover the details in the rupture fronts from BP images. I define a resolvability parameter  $\varepsilon_I(\omega)$  built from  $\mathbf{F}(\omega)$  that incorporates fault geometry, radiation pattern, and wave propagation (source-array geometry) to quantify the ability of the BP method to resolve details of the rupture on the fault.  $\varepsilon_I(\omega)$  successfully captures the similarity between BP images and kinematic source. I analyze the resolvability of most tectonically active regions and the most commonly used seismic arrays. Based on this global resolvability analysis, I propose an empirical relation between the seismic frequency, resolvable area, and earthquake magnitude. It provides general guidelines to choose the lowest frequency in seismic waveform (for example, about 0.3 Hz for  $M_w$  8 and 1 Hz for  $M_w$  7 earthquakes) and to

interpret the BP image in terms of the source kinematics. In general, this work attempts to provide a clear interpretation of the BP images in light of the real earthquake rupture process and give a systematic evaluation of seismic data limitations.

#### 4.0 INTRODUCTION

With the development of dense seismic arrays (e.g., Hi-net in Japan<sup>236,239</sup>; USArray (Earthscope program)), seismologists are able to harness key information of earthquake sources from seismic waveform coherency. The backprojection (BP) of high-frequency teleseismic P waves (usually from 0.1 to 4 Hz) is a method widely used to study the evolution of earthquake rupture and has been particularly effective for the study of large earthquakes. It provides relative location of the seismic radiation coherency on the projection of the fault plane at the hypocentral depth. Its application to the recent large earthquakes ( $M_w > 8$ ) has succeeded in characterizing a spatio-temporal evolution of seismic radiation of earthquakes<sup>94,138,139,156,209,326,344,346,349,350,361,358</sup>. BP in general requires fewer assumptions than kinematic slip inversions that necessitates, for example, fault geometry, slip-rate function shapes<sup>140,141</sup>, and rupture velocity in some cases<sup>151</sup>. In addition, the simplicity of the method allows for rapid calculations. Therefore, preliminary information about earthquake rupture processes can be rapidly obtained from waveform data, soon after the seismic waves arrive at the array of receivers (e.g., Incorporated Research Institutions for Seismology Data Management Center, IRIS DMC<sup>137</sup>). Despite the success of the BP approach, the physical interpretation of the images in terms of rupture properties is yet to be verified.

While the BP images are the spatial and temporal distributions of high frequency waveform coherency, they are often referred to as relative radiated energy<sup>139</sup> and/or energy burst<sup>350</sup>. Qualitative comparisons between BP results and independent kinematic inversions for the recent large earthquake events exhibit some spatial and temporal correlation between the BP images and the source

kinematic evolution<sup>22,164,175,205,308,324,347,359,361</sup>. In particular, the BP results constructed from low frequency waves (about 0.1-0.5 Hz) are mostly collocated with large coseismic slip and thus to negative coseismic shear stress change (stress drop)<sup>205,359,361</sup>. In contrast, the BP results constructed from the high frequency seismic waves (0.5-1 Hz) are consistent with the edges of large slip areas, and thus with the positive stress change (stress loading).

The temporal evolution of the earthquake seismic radiated energy, or seismic power, may be related to specific locations on the fault with the help of BP images: Denolle et al.<sup>75</sup> and Yin et al.<sup>358</sup> apply a time-varying spectral analysis to calculate the time history of earthquake radiated energy and directly compare it with BP results, showing the correlation between high coherency and high radiated energy. However, these comparisons remain qualitative, and the interpretation of BP images with respect to seismic energy or excitation is yet to be investigated.

A first element to discuss is the physical dimensions of the BP image. The BP algorithm involves the alignment and stacking of observed seismic waveforms. Therefore, the BP approach is essentially a manipulation of the seismic data, and the BP images carry the physical units of the data. Fukahata et al.<sup>99</sup> present a theoretical framework on the relationship between the BP results and a classical linear inversion solution. They focus on the conventional BP<sup>138,139</sup> and Hybrid BP<sup>346</sup> methods with linear stacking. They suggest that these BP images represent the slip motion on a fault, thereby approximately equal to a kinematic slip inversion, provided that the Green's function is sufficiently close to a shifted delta function. These conclusions are enlightening to understand the dimension of the BP images. However, their deductions rely on the assumption of strong decorrelation between source locations other than the true source. That is, whether the correlation between Green's functions from multiple sources to a single receiver is delta function in space. This assumption on decorrelation may not hold as it is widely used in seismic interferometry analysis<sup>51</sup>.

A second element to discuss is the ability of BP methods to resolve small wavelengths features in source radiation in various frequency bands. One of the conventional approaches to quantify

resolution is to perform the seismic array response<sup>262</sup>. The array response carries important information about the limitations in spatial resolution of a seismic array toward specific region; it represents the BP image given a delta source in time and space. Another method to test the resolution of a BP method is to setup a series of synthetic point sources with different locations and/or source times, then to apply the BP method and see whether these point sources can be correctly recovered<sup>209,326,349,360,358</sup>. These synthetic tests are popular to establish the spatial resolution limits of BP. Another example of such exercise is how Wang et al.<sup>326</sup> integrate these two approaches. They express the BP images as the convolution of an array response and a series of point sources, and then attempt to solve for the high-frequency radiators (source series) through an inversion scheme. However, and in general, an instantaneous point-source representation of the on-fault radiation may not be appropriate and the process zone (zone of active slip) is likely distributed in realistic earthquakes.

Addressing these two elements is necessary to interpret source physics from the BP images and to better apply the BP methodology to study earthquake ruptures. Specifically, the physical unit of the BP image determines whether I can interpret the BP images as snapshot of slip motions; the BP resolution controls whether, and how well, I can use the BP images to map rupture propagation (i.e. for appropriate estimate of rupture velocity).

Realistic kinematic source generators provide great opportunities to investigate the relation between BP images and kinematic properties. This study attempts to address the elements mentioned above using synthetic waveforms. I restrict my discussion to idealistic wave propagation in a homogeneous full space in order to focus on the relation between source and seismic waveforms and ignore the effects of 3D elastic structure (and Green's function) that might alter the results<sup>139,210,362</sup>. In a homogeneous full space there are (i) analytical formulations of the far-field body waves<sup>7</sup> and (ii) reliable kinematic source representations (in this study, I use the kinematic source generator developed by Crempien & Archuleta<sup>68</sup>, Liu et al.<sup>183</sup>, Schmedes et al.<sup>276</sup>). Moreover, I consider the simplest approach to backprojection, that is, the linear stacking in the Fourier domain. This frame-

work enables a direct reading of the BP image in light of the source slip-rate functions. Given this linear formulation, I propose a simple scalar metric to quantify the BP resolution solely based on the source-receiver geometry and for a given seismic frequency. Then, I test the linear BP images against realistic and heterogeneous kinematic sources. Finally, I extend these theoretical formulations to explore realistic limitations of BP techniques given the distribution of global seismicity and of globally available seismic arrays.

## 4.1 METHODS

### 4.1.1 SYNTHETIC SEISMOGRAMS FOR KINEMATIC SOURCES

In the homogeneous full space, the direct teleseismic P-wave displacement seismograms  $d_k(t)$  recorded by the  $k^{th}$  station can be regarded as the summation over the fault plane (or source) of individual slip-rate functions  $u_n(t)$  (subfault  $n$ ) with terms of radiation pattern  $R_{kn}^P$ , geometrical spreading, and travel-time delay  $t_{kn}$ <sup>7</sup>:

$$d_k(t) = \sum_{n=1}^N \frac{R_{kn}^P}{4\pi\rho\alpha^3} \frac{\mu\Delta S}{r_{kn}} \dot{u}_n(t - t_{kn}), \quad (4.1)$$

where  $r_{kn}$  is the distance from the  $n^{th}$  subfault to the  $k^{th}$  station;  $\rho$ ,  $\alpha$ , and  $\mu$  are the density, P-wave velocity and shear modulus in the source region, respectively.  $\Delta S$  is the area of the subfault.

This is a discretized formulation of the representation theorem<sup>48</sup> applied in the far field for a source with known slip history. After a Fourier transform, the travel-time delays become phase shifts  $e^{-i\omega t_{kn}}$  at the angular frequency  $\omega$ ,

$$D_k(\omega) = \sum_{n=1}^N \frac{R_{kn}^P}{4\pi\rho\alpha^3} \frac{\mu\Delta S}{r_{kn}} e^{-i\omega t_{kn}} \dot{U}_n(\omega). \quad (4.2)$$

Given the linearity of the formulation in the frequency domain, I form an vectorial representation to incorporate seismograms from an array of stations (seismogram spectra):

$$\begin{bmatrix} D_1(\omega) \\ D_2(\omega) \\ \vdots \\ D_K(\omega) \end{bmatrix} = \mathbf{A}(\omega) \begin{bmatrix} \dot{U}_1(\omega) \\ \dot{U}_2(\omega) \\ \vdots \\ \dot{U}_N(\omega) \end{bmatrix}, \quad (4.3)$$

where the wave propagation matrix  $\mathbf{A}(\omega)$  is:

$$\mathbf{A}(\omega) = \frac{\mu\Delta S}{4\pi\rho\alpha^3} \times \begin{bmatrix} \frac{R_{11}^p}{r_{11}} e^{-i\omega t_{11}} & \dots & \frac{R_{1N}^p}{r_{1N}} e^{-i\omega t_{1N}} \\ \frac{R_{21}^p}{r_{21}} e^{-i\omega t_{21}} & \dots & \frac{R_{2N}^p}{r_{2N}} e^{-i\omega t_{2N}} \\ \vdots & \ddots & \vdots \\ \frac{R_{K1}^p}{r_{K1}} e^{-i\omega t_{K1}} & \dots & \frac{R_{KN}^p}{r_{KN}} e^{-i\omega t_{KN}} \end{bmatrix}_{K \times N}. \quad (4.4)$$

Being a linear operator in the frequency domain, the vectorial formulation of Eq.4.4 is convenient to separate the two main variables that constitute a seismogram: the source term with the slip-rate function  $\dot{U}_n(\omega)$  and the wave-propagation term  $\mathbf{A}(\omega)$ . The latter can be revised to accommodate radiation pattern, geometrical spreading, and travel-time elements calculated in a 3D Earth model. The linear BP in the frequency domain is similar to beamforming<sup>262,326,360</sup>. In the practical application of frequency-domain BP, the waveform data are windowed and Fourier transformed to construct Eq.4.3. This provides the temporal dependence of the BP images.

#### 4.1.2 FORMULATION OF LINEAR BP IN THE FREQUENCY DOMAIN

The two key ingredients of BP are waveform alignment and stacking<sup>138</sup>. The literature is rich in method development to improve both ingredients<sup>139,206,210,320,344,346,350,366</sup>. The alignment in my

synthetic exercise is known and trivial. The linear stacking scheme is chosen in order to relate source kinematics to BP images, which differs from other studies that may favor nonlinear  $n^{th}$ -root stacking scheme to enhance resolution.

The alignment and linear stacking are carried out by multiplying a phase-shift matrix  $\tilde{\mathbf{A}}(\omega)$  to the left hand side of  $\mathbf{D}(\omega)$  in Eq.4.3:

$$\tilde{\mathbf{A}}(\omega) = \begin{bmatrix} e^{j\omega t_{11}} & e^{j\omega t_{21}} & \dots & e^{j\omega t_{K1}} \\ e^{j\omega t_{12}} & e^{j\omega t_{22}} & \dots & e^{j\omega t_{K2}} \\ \vdots & \ddots & \ddots & \vdots \\ e^{j\omega t_{1N}} & e^{j\omega t_{2N}} & \dots & e^{j\omega t_{KN}} \end{bmatrix}_{N \times K} \mathbf{W}, \quad (4.5)$$

where the matrix  $\mathbf{W}$  is a  $K \times K$  diagonal matrix that is used in linear weighted stack to balance the contributions of seismograms. The weighting matrix  $\mathbf{W}$  is usually applied to normalize the different amplitude of waveforms or adjust the uneven distribution of stations in a seismic array<sup>320,321,350</sup>. In this example, I apply uniform averaging by choosing  $\mathbf{W} = \frac{1}{K}\mathbf{I}$  for the evenly distributed synthetic array and omit it in the following discussion. In practice, the travel-time terms in the  $\tilde{\mathbf{A}}(\omega)$  are theoretically calculated based on a specific Earth velocity model. Therefore, I can obtain the BP results, or I call BP image at frequency  $\omega$ :

$$\mathbf{U}^{BP}(\omega) = \tilde{\mathbf{A}}(\omega)\mathbf{D}(\omega) = \tilde{\mathbf{A}}(\omega)\mathbf{A}(\omega)\dot{\mathbf{U}}(\omega) = \mathbf{F}(\omega)\dot{\mathbf{U}}(\omega). \quad (4.6)$$

This simple form provides a linear relation in the frequency domain between the BP image constructed from displacement seismograms  $\mathbf{U}^{BP}(\omega)$  and slip-rate field on the fault surface  $\dot{\mathbf{U}}(\omega)$ . Specifically, the BP image should be proportional to the band-pass filtered slip-rate field. The scaling factor is the matrix  $\mathbf{F}(\omega) = \tilde{\mathbf{A}}(\omega)\mathbf{A}(\omega)$ , which is a frequency dependent function of the source-array geometry, wave propagation effects, and radiation pattern (Fig.4.1 (a) - (c)). Note that  $\tilde{\mathbf{A}}(\omega)$

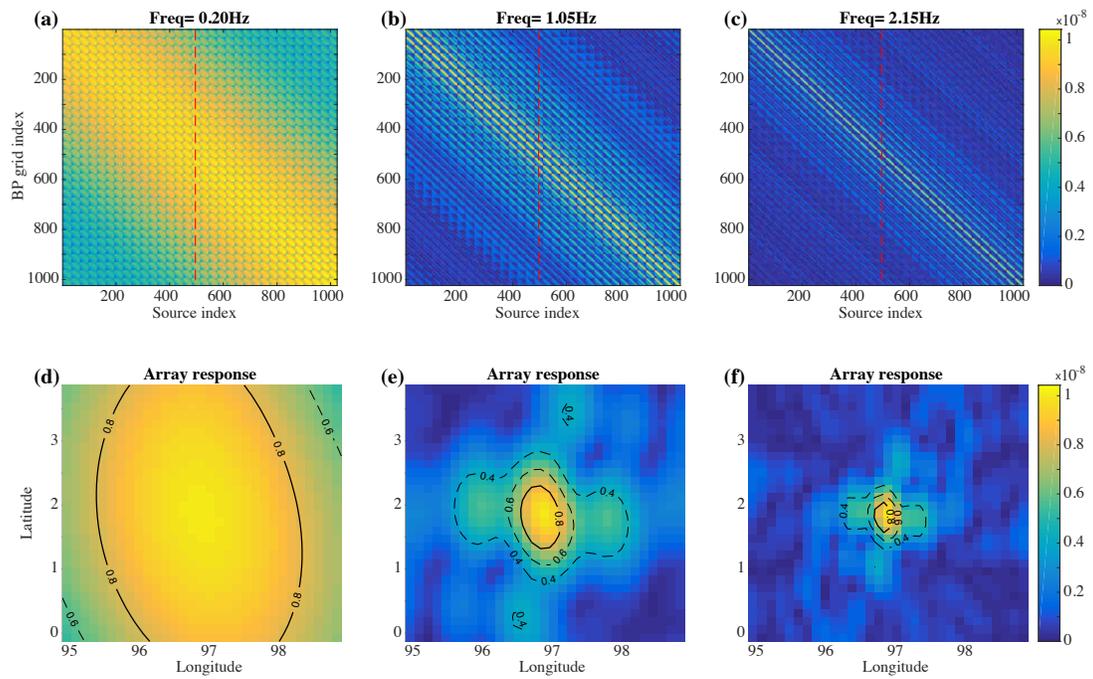
is not  $\mathbf{A}(\omega)^{-1}$ , which would turn the problem into a kinematic slip inversion. Instead,  $\mathbf{F}(\omega)$  bears great similarity with the array response: each column is the array response to an impulse source at a particular source location with specific radiation pattern terms (Fig.4.1 (d) - (f)). I refer to  $\mathbf{F}(\omega)$  as the resolution matrix because of its spatial smoothing effects on the slip-rate field at a specific frequency  $\omega$ . The closer  $\mathbf{F}(\omega)$  is to identity, the greater the similarity between BP image and slip-rate field. I thus proceed to investigate the impact of  $\mathbf{F}(\omega)$  onto interpreting the slip-rate distribution from BP images. Fig.4.1 (a) - (c) shows an example of  $\mathbf{F}(\omega)$ : it varies in shape as it converges to diagonal with increasing seismic frequency. To quantify the similarity between the resolution and identity matrices, I define the resolvability parameter  $\varepsilon_I$  as the 2D correlation coefficient between the  $\mathbf{F}(\omega)$  and an identity matrix with same size:

$$\varepsilon_I(\omega) = |\text{corr2}(\mathbf{F}, \mathbf{I})| = \frac{|\sum_m \sum_n (F_{mn} - \bar{F})(I_{mn} - \bar{I})|}{\sqrt{[\sum_m \sum_n (F_{mn} - \bar{F})^2][\sum_m \sum_n (I_{mn} - \bar{I})^2]}}, \quad (4.7)$$

$m, n$  being the elements of the matrices.  $\varepsilon_I(\omega)$  varies between 0 and 1 and provides a compact form to quantify the resolution of linear BP for specific array settings and the deterioration effects of the source-receiver geometry on the BP image. I refer to  $\varepsilon_I(\omega)$  as measure of resolvability. It does not carry the units of spatial resolution, instead it encapsulates multiple parameters relevant to BP processing. This choice bears some similarity with other metrics, such as the Goodness-Of-Fit criteria that combines multiple ground motion metrics to quantify broadband waveform fitting<sup>240</sup>.

## 4.2 BACKPROJECTION ON KINEMATIC SOURCES

I test the linear BP method using the theoretical formulation of Eq.4.6 and its usefulness in interpreting kinematic properties on synthetic sources that has kinematic complexity.



**Figure 4.1:** Absolute value of the resolution matrices of the Hi-Net array toward IDN2 region (see location in Figure 7) at (a) 0.2 Hz; (b) 1.05 Hz and (c) 2.15 Hz. (d) - (e) show the corresponding array response at the source location indicated by the red dashed lines in the top panels. Contours indicate the 0.4, 0.6 and 0.8 of the maximum value.

#### 4.2.1 SYNTHETIC EXAMPLE SET UP

A pseudo-dynamic source model is a statistical representation of the source built upon the correlations among kinematic parameters found in earthquake dynamic models<sup>196,275</sup>. I use a kinematic source generator developed by Liu et al.<sup>183</sup> and Crempien & Archuleta<sup>68</sup>. The kinematic source parameters are local slip, rise time, rupture velocity, peak time. After a spatial discretization of the fault plane, I obtain a series of correlated distributions of seismic moment, rupture velocity, and rise time (Appendix A Fig.A.1). Onset time, which is the time when each subfault begins to slip, is calculated using the wave equation on the rupture velocity field<sup>96</sup>. I use the moment-rate function defined in Liu et al.<sup>183</sup>:

$$\dot{u}(t) = \begin{cases} C_N[0.7 - 0.7 \cos(\pi t/\tau_1) + 0.6 \sin(0.5\pi t/\tau_1)] & (0 \leq t < \tau_1) \\ C_N[1.0 - 0.7 \cos(\pi t/\tau_1) + 0.3 \cos(\pi(t - \tau_1)/\tau_2)] & (\tau_1 \leq t < 2\tau_1), \\ C_N[0.3 + 0.3 \cos(\pi(t - \tau_1)/\tau_2)] & (2\tau_1 \leq t < \tau) \end{cases} \quad (4.8)$$

where  $C_N = \pi/(1.4\pi\tau_1 + 1.2\tau_1 + 0.3\pi\tau_2)$  is a normalization constant,  $\tau$  is the rise time,  $\tau_1 = 0.3\tau$  is the peak time and  $\tau_2 = \tau - \tau_1 = 0.7\tau$ . Therefore, the  $n^{th}$  subfault patch on the fault surface has the corresponding slip-rate function:

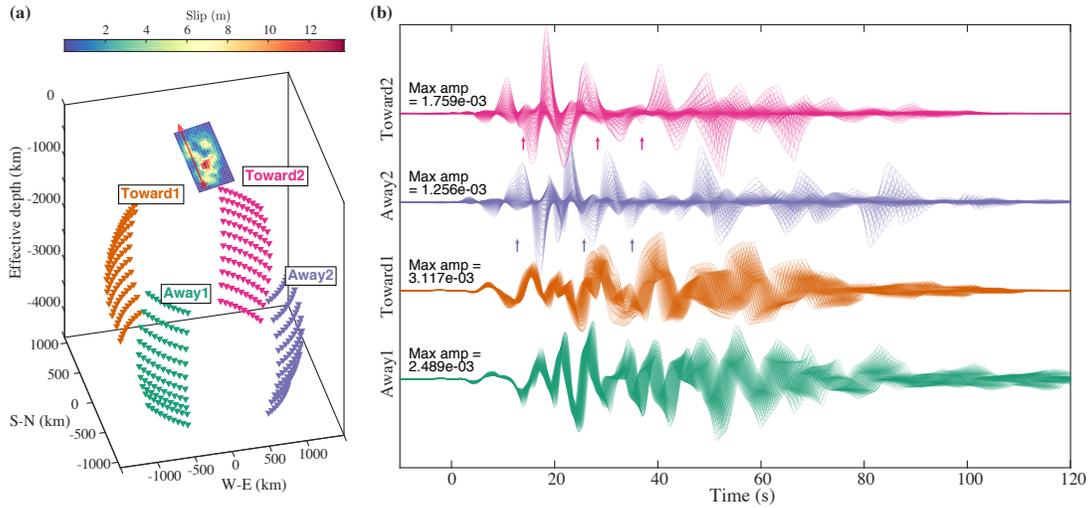
$$\dot{u}_n(t) = \dot{u}(t - t_0^n)M_0^n/(\mu\Delta S), \quad (4.9)$$

where  $t_0^n$  and  $M_0^n$  are the onset time and seismic moment at the  $n^{th}$  subfault, respectively. The total seismic moment of these pseudo-dynamic sources corresponds to  $M_w$  8, above which magnitude the BP methods is generally applied and seem to work best (more details are discussed in the section

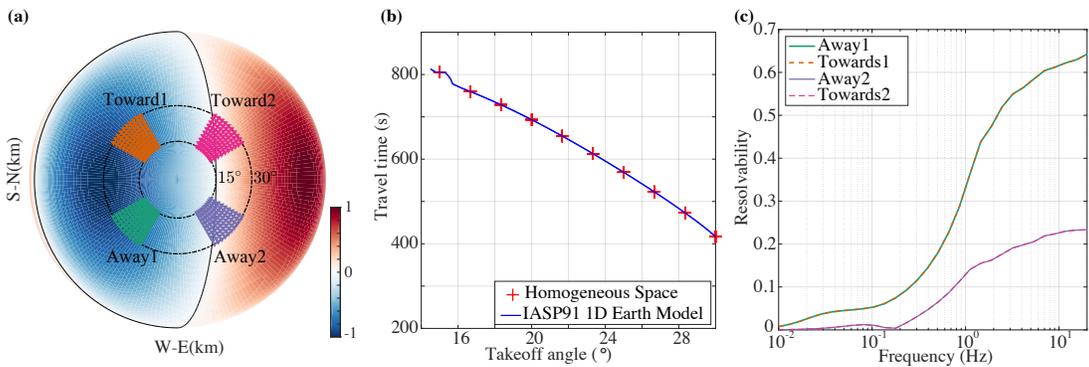
2.4 and section 2.5.4).

The synthetic seismograms are constructed from the slip history of each source model. To focus on source rather than wave propagation effects, I keep wave propagation simple and embed the source in a homogeneous full space (Fig.4.2 (a)) of elastic properties density  $\rho = 2,700 \text{ kg/m}^3$ ; shear modulus  $\mu = 2.43 \text{ GPa}$ ; P and S wave velocity are  $V_P = 5.2 \text{ km/s}$  and  $V_S = 3.0 \text{ km/s}$ , respectively. Eq.4.1 then allows me to numerically compute the direct P-wave seismic waveforms for a specific array of receivers. The focal mechanism at each subfault is a pure shear double couple with  $15^\circ$  dip angle,  $0^\circ$  strike angle, and  $90^\circ$  rake angle to render the typical slip direction of megathrust earthquakes. I strategize to place the synthetic receivers beneath the synthetic source to mimic the steep takeoff angles of teleseismic P waves ( $15^\circ - 30^\circ$ , see Fig.4.2 (a) and Fig.4.3 (a)). Rupture velocity information can be inferred from directivity effects. Therefore, I apply the linear BP method for the synthetic source with two types of seismic arrays: (i) arrays Toward1 and Toward2, both located ahead of the rupture direction; (ii) arrays Away1 and Away2, both located behind the direction of rupture. In each type of arrays, I also design the locations of two arrays to sample different parts of the radiation patterns: Set1 with those labeled 1 (Toward1 and Away1) have rays that sample the same quadrant of the P-wave radiation pattern (i.e. identical polarity) while the Set2 arrays labeled 2 (Toward2 and Away2) mostly sample the P-wave nodal plane (Fig.4.3 (a)). I adjust the distance to the kinematic source with the known takeoff and azimuth angles of each virtual station (Fig.4.3 (a)) and make the travel time identical to those calculated from the IASP91 1D Earth velocity model (Fig.4.3 (b), the velocity model is from Kennett & Engdahl <sup>150</sup>). All these settings aim to keep the synthetic BP tests resembling the real applications. It is intuitive that Set2, which samples the nodal plane, is greatly impaired by waveform de-coherence (Fig.4.2 (b)) among virtual receivers and thus produce a low resolvability  $\varepsilon_I(\omega)$ .

I apply the basic linear BP method described in Section 2.2 to these synthetic waveforms, which I filter in several narrow frequency bands within 0.1 to 1 Hz. I slide through the waveforms with a



**Figure 4.2:** Synthetic example: (a) Final slip distribution of a kinematic source model (colorscale) as well as locations of the 4 seismic arrays (colored triangles) embedded in the homogeneous full space. The spatial dimensions of the source are exaggerated 5 times for better view. The red star and arrow indicate the location of the hypocenter and the overall direction of rupture propagation, respectively. (b) Synthetic waveforms filtered [0.1 1] Hz recorded by each array. The numbers indicate the maximum waveform amplitudes in each array. Arrows indicate some unsystematic polarity shifting of array waveforms due to rupture propagation (Also see Appendix A Fig.A.2).



**Figure 4.3:** Takeoff angle distributions and BP resolvability of the four synthetic arrays: (a) Focal mechanism (lower hemisphere) of the synthetic source as well as the projection of takeoff ray path of each stations in bird's-eye view. The focal mechanism is color-coded by the radiation pattern and the nodal planes are also indicated by the black thin lines. The two dashed circles show the 10° and 30° takeoff angles, respectively. (b) Blue curve shows the P-wave takeoff angle against travel time based on the IASP91 1D Earth velocity model<sup>150</sup>. Red crosses indicate the same setup for the synthetic arrays in the homogeneous full space (Fig.4.2 (a)). (c) BP resolvability  $\epsilon_I(\omega)$  calculated for each array. Array colors are the same as in Fig.4.2

20% Tukey window taper (20% total window length for the cosine taper) every time step of 0.5 s. The length of time window is chosen as 4 times of longest period ( $4/f_{min}$  seconds) of the bandpass filters (40 s: 0.1-0.2 Hz; 20 s: 0.2-0.4 Hz; 10 s: 0.4-0.7 Hz; 6 s: 0.7-1 Hz) to capture enough periods in the waveforms. Then, I transform the windowed waveforms to frequency domain, obtain the synthetic data spectra  $\mathbf{D}(\omega)$ , and calculate the corresponding phase-shift matrix  $\tilde{\mathbf{A}}(\omega)$  for the pre-defined source location. Therefore, I can obtain the BP images at each frequency  $\omega$  and for each time window (Eq.4.6). It is common in frequency-domain backprojection to correct the window time to the appropriate source time: the motion of a source stretches of the seismic signal that distorts the windowing time axis (similar to Doppler effects, see the directivity effects in waveforms in Fig.4.2 (b)) and thus requires a time calibration. I apply the same calibration method as introduced by Yin & Yao<sup>360</sup> (see their Eq.(11)) and use the location of highest BP amplitude to calibrate the window time for the correct source time.

In this controlled experiment, I can directly compare the BP results with the ground truth parametrization of the rupture. Since the relation between BP results and source kinematics is built in the frequency domain (Eq.4.6), it is necessary to combine the BP images at various frequencies and compare with the slip motions in a continuous frequency band. However, I cannot equate the time series of broad-band BP results (i.e. inverse Fourier transform of the BP value at each subfault  $\int \mathbf{U}^{BP}(\omega)e^{i\omega t} d\omega$ ) and slip-rate field (i.e. inverse Fourier transform of  $\dot{\mathbf{U}}(\omega)$ ) simply from Eq.4.6 because the resolution matrix  $\mathbf{F}(\omega)$  is frequency dependent and is not identity (Fig.4.1). Instead, I focus on the spatial similarities between the BP images and slip motions distribution of the kinematic sources. I compare the averaged the BP results with all central frequencies (13 discrete frequency values in total: 0.125 Hz, 0.15 Hz, 0.175 Hz, 0.20 Hz, 0.25 Hz, 0.30 Hz, 0.35 Hz, 0.40 Hz, 0.50 Hz, 0.60 Hz, 0.70 Hz, 0.83 Hz, 1.00 Hz) and the filtered slip-rate field within the broader frequency band of 0.1 to 1 Hz. I normalize the BP images at each frequency due to the large differences in the absolute amplitude of these BP results. The frequency-dependent normalization factor is taken as

the peak amplitude of the image over the entire source duration. By averaging the normalized BP results over all frequencies, I can obtain the average BP image in the corresponding frequency band. To quantify the similarity between the images, I measure the 2D correlation coefficient (CC) also defined in Eq.4.7 between snapshots of the averaged BP image and of the bandpass filtered slip-rate field.

#### 4.2.2 RESULTS OF SYNTHETIC BACKPROJECTION

##### RESOLVABILITY

First, I estimate the resolvability for all four synthetic arrays in the way that was introduced in Section 2.2 Eq.4.7 (Fig.4.2). The resolvability  $\varepsilon_I(\omega)$  increases with seismic frequency (Fig.4.3 (c)). Because of the symmetry of the array distributions with respect to the radiation pattern, the resolvability curve of Toward<sub>1</sub> and Toward<sub>2</sub> overlap with those of Away<sub>1</sub> and Away<sub>2</sub>, respectively. Moreover, the resolvability of Toward<sub>1</sub> and Away<sub>1</sub> is systematically higher than Toward<sub>2</sub> and Away<sub>2</sub> due to better coherency of the waveforms (Fig.4.2 (b) and Appendix A Fig.A.2).

Precaution ought to be given to arrays that sample the nodal plane of the focal sphere. The lower resolvability  $\varepsilon_I(\omega)$  of Set<sub>2</sub> indicates the lower BP resolution of seismic arrays near the nodal plane of focal mechanisms due to the source-receiver geometry. Although the early waveform polarity can be manually adjusted by changing the signs of elements in the weighting matrix  $\mathbf{W}$ , it is difficult to track the later polarity flips due to the propagation of rupture (see arrows in Fig.4.2 (b) as well as in the Appendix A Fig.A.2). In addition, moving ruptures induce two effects that might dominate near the nodal planes: i) the moving rupture changes the source-receiver geometry and ii) the radiation pattern is likely to vary due to non-planar fault geometry (for example, the 2002 Denali  $M_w$  7.9 earthquake<sup>91</sup>; the 2012 Sumatra  $M_w$  8.6 earthquake<sup>207</sup>; and the 2016 Kaikoura  $M_w$  7.8 earthquakes<sup>88</sup>). Therefore, delayed polarity flipping can greatly impair waveform coherence and yield

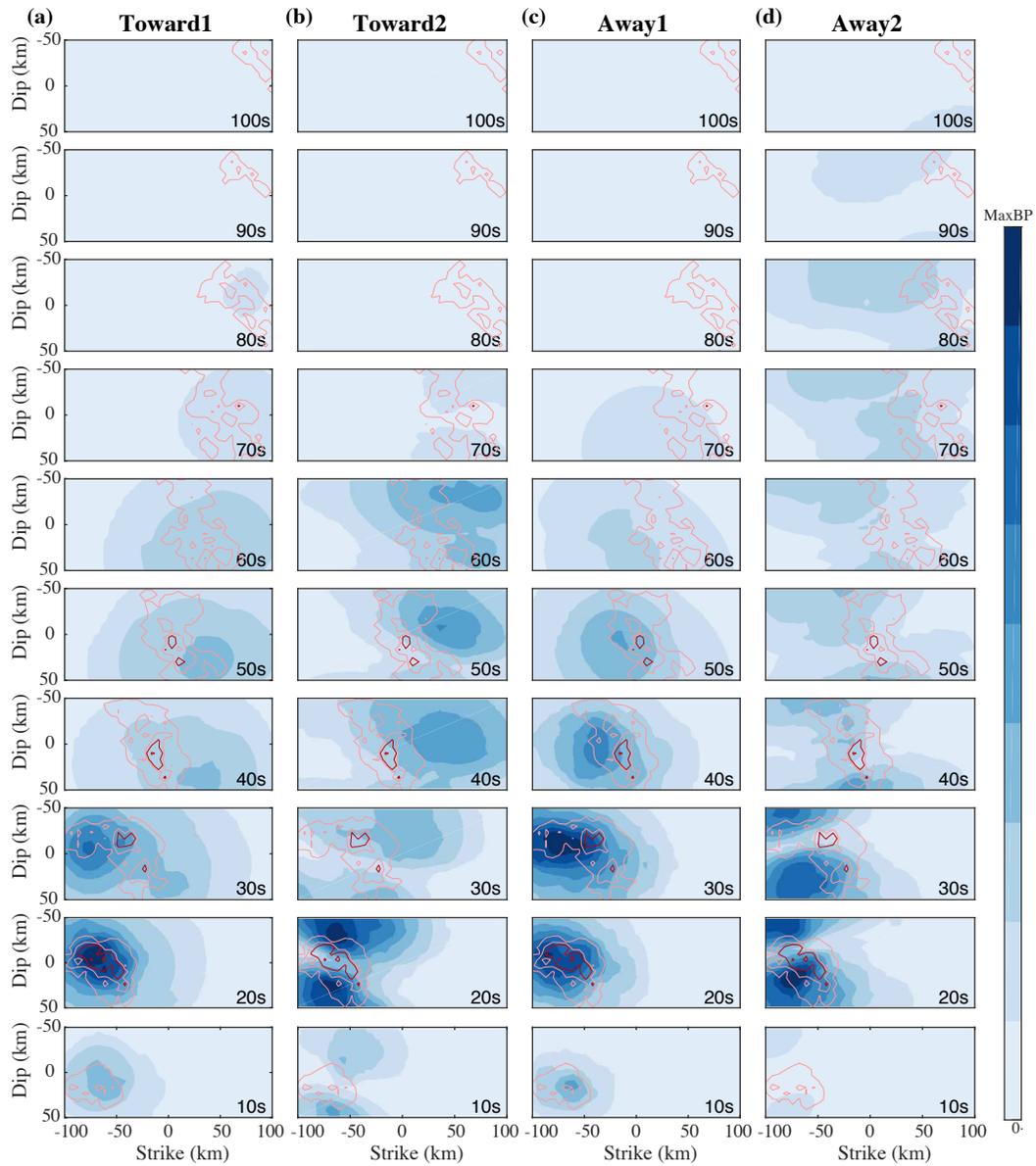
poor BP resolution and significant bias in the results. In general, arrays with rays taking off in the vicinity of the nodal planes will be subject to uncertain BP results.

To conclude, the resolvability parameter provides a metric to select array location and confidence in the BP resolution. It incorporates source-receiver geometry and radiation pattern effects present in the resolution matrix  $\mathbf{F}(\omega)$  and thus in  $\epsilon_I(\omega)$ . The resolvability can be easily extended to more complex station distributions like realistic seismic arrays (see later Section 4).

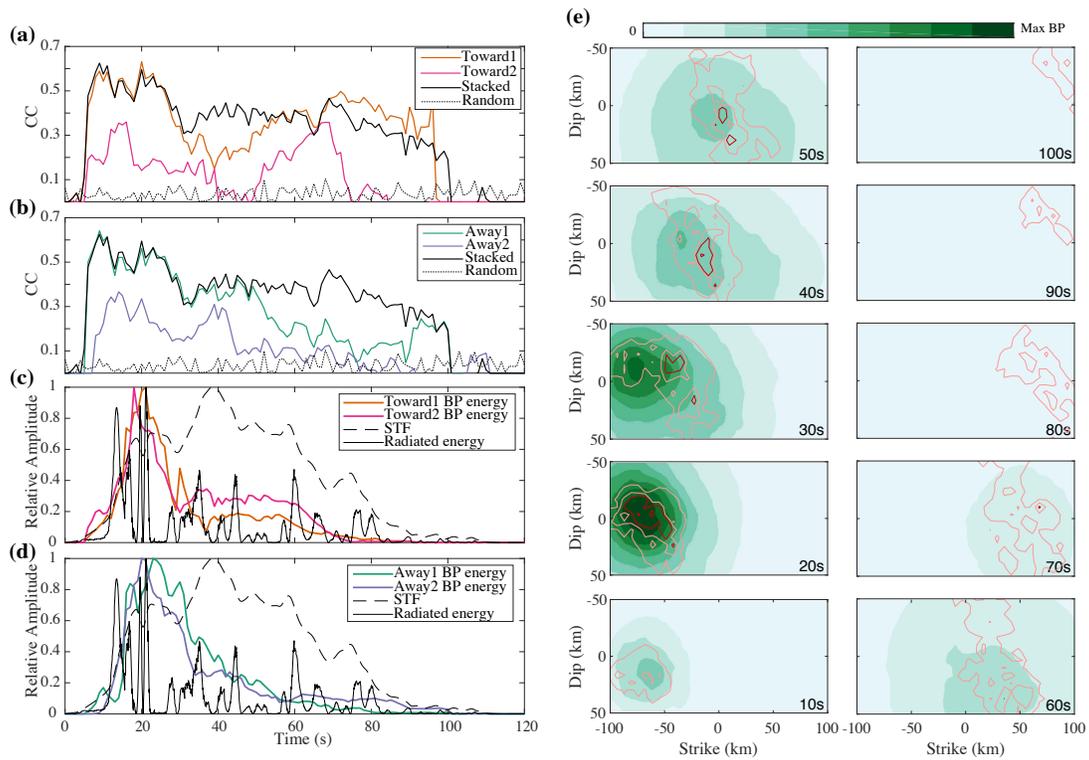
#### BP IMAGES VS SLIP-RATE IMAGES

The absolute amplitude of the BP images is controlled by the geometrical spreading and attenuation, which is in general poorly constrained. Resolvability is better at higher frequency, but the displacement and velocity seismograms are dominated by low frequencies due to the long source duration. Thus, I normalize the BP images at each frequency between 0.1 and 1 Hz and average them for each array. Fig.4.4 shows these images against the known band-passed filtered slip-rate field. Overall, the general features of the BP images are consistent with the evolution of high slip rates (Fig.4.4). The CC values generally vary between 0.1 to 0.6 (Fig.4.5 (a) and (b)), which indicates that each array is able to capture relatively well the propagation of rupture on the fault surface, even with lower resolvability. I now discuss the second order disparities among the BP images.

The CCs from Set1 (range 0.2 - 0.6) are systematically higher than those obtained with Set2 (about 0.1 - 0.4), especially during the major stage of moment release in the first 80 s (Fig.4.5 (a) and (b)). It is expected to occur from the higher resolvability values of Set1. Taking the 10 and 20 s snapshots for example, Set 2 arrays produce 2 peaks instead of the single peak of the slip-rate distribution (Fig.4.4 (b) and (d)). Therefore, these two peaks are likely artifacts due to the improper source-receiver geometry, i.e. the sampling of the nodal planes on the focal sphere. Because the source directivity effects are expected to occur at equal strength in both Sets, such as those seen in the raw waveforms (Fig.4.2 (b)), I attribute these first-order differences to the source-receiver geometry, ra-



**Figure 4.4:** Comparison between the kinematic source model and BP results from (a) Toward1; (b) Toward2 and (c) Away1 and (d) Away2 arrays. The blue colorscale corresponds to the BP image (averaged over frequencies) at each time step. The pink thin contours correspond to 1% and 10% while red bold contours correspond to 20%, 50% and 80% of the maximum amplitude of filtered slip rates.



**Figure 4.5:** Comparison between the BP images and the kinematic source in the time domain: (a) - (b) the time-varying correlation coefficients (CC) between moment-rate distribution and BP images from all 4 different arrays (thin gray lines). The Toward1 and Toward2 arrays are color-highlighted in (a) while Away1 and Away2 arrays are highlighted in colors in (b). Bold black lines show the time-varying CC curve between moment-rate distribution and stacked BP results (also shown in (e)). The gray dashed line shows the correlation between moment-rate distribution and random images produced from uniform distribution. (c) - (d) show the normalized peak BP energy burst evolution from each array (Colored curves: Toward1 and Toward2 in (c) while Away1 and Away2 in (d)) comparing with the normalized source time function (STF, in gray dashed lines) and radiated energy evolution (squared time derivative of STF, in black lines). (e) Green images show the stacked BP images compared with slip rate distribution. Other symbols are the same as in Fig.4.4.

diation pattern effects, which are captured in  $\varepsilon_I$  (Fig.4.3 (c)). Therefore, the higher resolvability of Set1 confirms that Set1 is able to better image the slip-rate evolution.

The BP images are also affected by rupture directivity effects. The BP images from the Toward arrays (Fig.4.4 (a) and (b)) capture the beginning (0 - 30 s) as well as the end (50 - 100 s), but have lower quality results in between (30 - 50 s); the BP results from Away1 and Away2 arrays (Fig.4.4 (c) and (d)) are slightly more consistent with the slip-rate distribution within 30 - 50 s but give poorer constraint on the later stage of rupture after 50 s. During the first 30 s Toward and Away arrays exhibit quite similar results and have approximately the same level of CC values within both Sets (Fig.4.5 (a) and (b)).

The complementary results obtained from the Toward and Away arrays imply that I can attempt to improve the BP results through stacking of seismic arrays. This stacking strategy has been successfully employed in previous studies<sup>252,366</sup>. Based on Eq.4.6, the stacking over various arrays is effectively a stack of their resolution matrix  $\mathbf{F}(\omega)$  for the same source term  $\mathbf{U}(\omega)$  and thus improves the resolvability. I perform the stacking on the BP images from single array (Fig.4.4) to obtain the stacked results in Fig.4.5 (e). In practice, the stacking over different seismic arrays may require some weighting of the contributions of different arrays<sup>366</sup>. But in my synthetic test on stacking, the absolute amplitudes of BP images from each single array are preserved without extra weighting when stacking over arrays. This is reasonable because the aperture and scale of four synthetic arrays are similar but amplitude of waveforms varies a lot (Fig.4.2 (b)). Therefore, the direct stacking naturally allows the BP images from Set1 arrays with higher resolvability to dominate. As expected, the stacking can provide a sharper image and a better fit with stable CC from 0.4 to 0.6 (Fig.4.5 (a) and (b)) for the entire rupture duration.

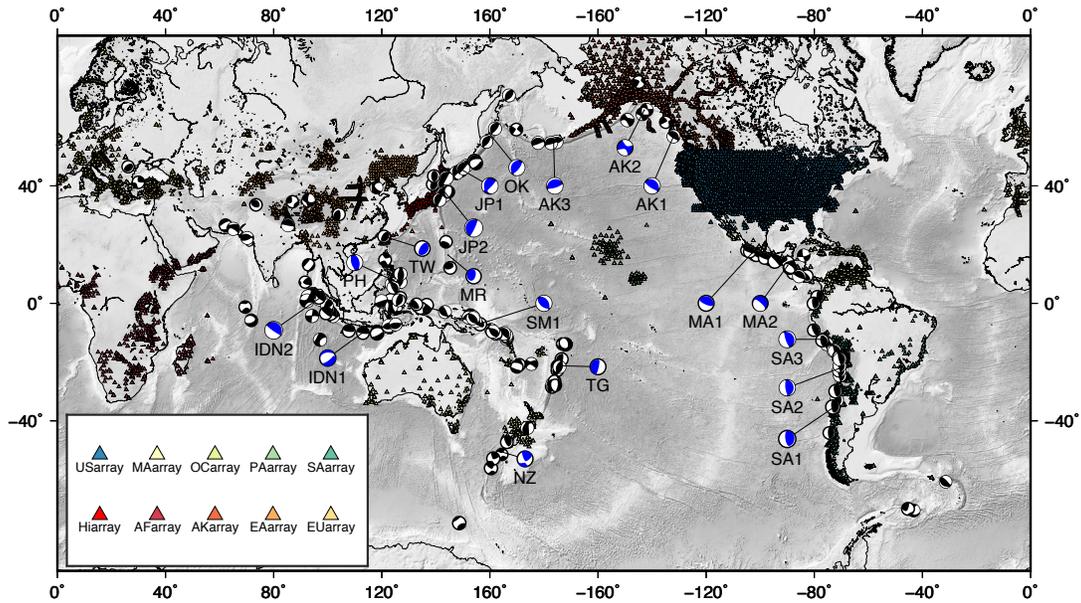
Finally, it is common to analyze the temporal and spatial evolution of the peaks of the BP images. I can either look at (i) the squared peak BP amplitudes, which is usually called relative energy radiation<sup>139</sup>, or (ii) track the spatial variation of BP peaks to estimate the rupture velocity.

(i) I compute the temporal evolution of the peak squared BP amplitude, that is, the relative energy radiation for each array (Fig.4.5 (c) and (d)). I also compare them with the squared moment acceleration (time derivative of source time function), which is proportional to the radiated energy (black lines in Fig.4.5 (c) and (d)). The BP peak amplitude from all arrays captures quite well the onset of the moment-rate and moment-acceleration functions, as also captured by the high CC values. Furthermore, the time series of BP energy resembles that of the squared moment acceleration. One possible explanation is the whitening of the BP spectrum during the stacking over frequency, which effectively brings up the level of the high frequencies. However, their strict similarity is hindered by methodological limitations such as off diagonal terms in the resolution matrix  $\mathbf{F}(\omega)$ , rupture directivity, even structural effects for the real BP applications.

(ii) Since the BP peaks are consistent with the peak locations of slip motion on the fault (Fig.A.3 (a) in the Appendix A), I can estimate the average rupture velocity from propagation of BP peaks. I use the BP results from the Away<sub>1</sub> array (Fig.4.4 (c)) as an example. Similar to many BP studies<sup>209,325,350,359,358</sup>, I estimate the average rupture velocity through a linear fit between the distance from epicenter to BP peaks and time (Fig.A.3 (c) in the Appendix A). I find that the rupture velocity estimated from the slip-rate peaks is  $1.75 \pm 0.03$  km/s while the rupture velocity from BP peaks is  $1.55 \pm 0.06$  km/s. The rupture velocity estimated from other arrays is generally consistent with slight difference (Toward<sub>1</sub>:  $1.69 \pm 0.09$  km/s; Away<sub>2</sub>:  $1.53 \pm 0.08$  km/s; Toward<sub>2</sub>:  $1.62 \pm 0.10$  km/s).

### 4.3 RESOLVABILITY OF GLOBAL EARTHQUAKES AND ARRAYS

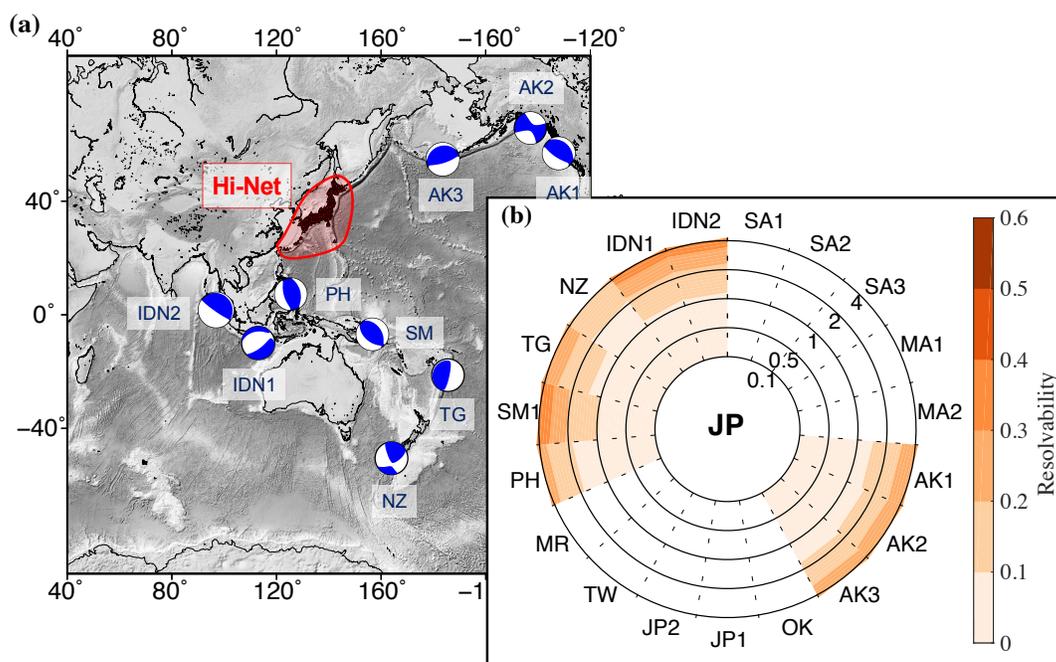
In addition to the synthetic exercise, my study aims to provide recommendations for BP studies through the evaluation of resolvability  $\epsilon_I(\omega)$  given the global seismicity and accessible seismic networks. Based on Eqs.(4.4) and (4.5), I simply calculate  $\mathbf{F}(\omega)$  with the radiation pattern terms  $R_{kn}^P$



**Figure 4.6:** Global map of seismic arrays and focal mechanisms in this study. Colored triangles indicate various seismic arrays available since 2004 to 2018 on IRIS SeismicQuery website (<https://ds.iris.edu/SeismiQuery/station.htm>) and NIED Hi-net websites (<http://www.hinet.bosai.go.jp/>). Many of these stations may not be available/deployed during the same period of time. Black focal mechanisms are those of shallow (depth < 100 km) earthquake with magnitude  $M_w > 7.5$  from the Global Centroid Moment Tensor (GCMT) solution (<http://www.globalcmt.org/>). The average focal mechanism in each region is indicated by blue beach balls.

and the relative position between global seismic stations and global source regions. I then use the Global Centroid Moment Tensor data base (GCMT, <http://www.globalcmt.org/>) to estimate the global seismicity radiation pattern (Fig.4.6)

In practice, the compilation of regional focal mechanisms of past moderate and large magnitude earthquakes allows me to construct an effective radiation pattern through averaging of strikes, dips, and rakes. I choose 19 regions in the world where the occurrence of large earthquakes ( $M_w > 7.5$ ) is frequent (Fig.4.6). For each region I only select focal mechanisms from the  $M_w > 7.5$  earthquakes with depth < 100 km and then directly average their source parameters: the 6 components of their moment tensor, longitude, latitude, and depth. I naturally weight the averages based on their seismic moment and let the focal mechanisms of the largest ( $M_w$  8-9) dominate. Finally, I scale the



**Figure 4.7:** Resolvability of the Hi-Net array toward source regions within  $30^{\circ}$ - $90^{\circ}$  teleseismic distances. (a) Map view of the Hi-Net array and the averaged source focal mechanisms in each regions (blue beach balls). (b) The frequency-varying resolvability of Hi-Net array toward different regions. The concentric circles correspond to frequency from 0.1 Hz to 4 Hz in log scale. The resolvability is color-scaled in orange.

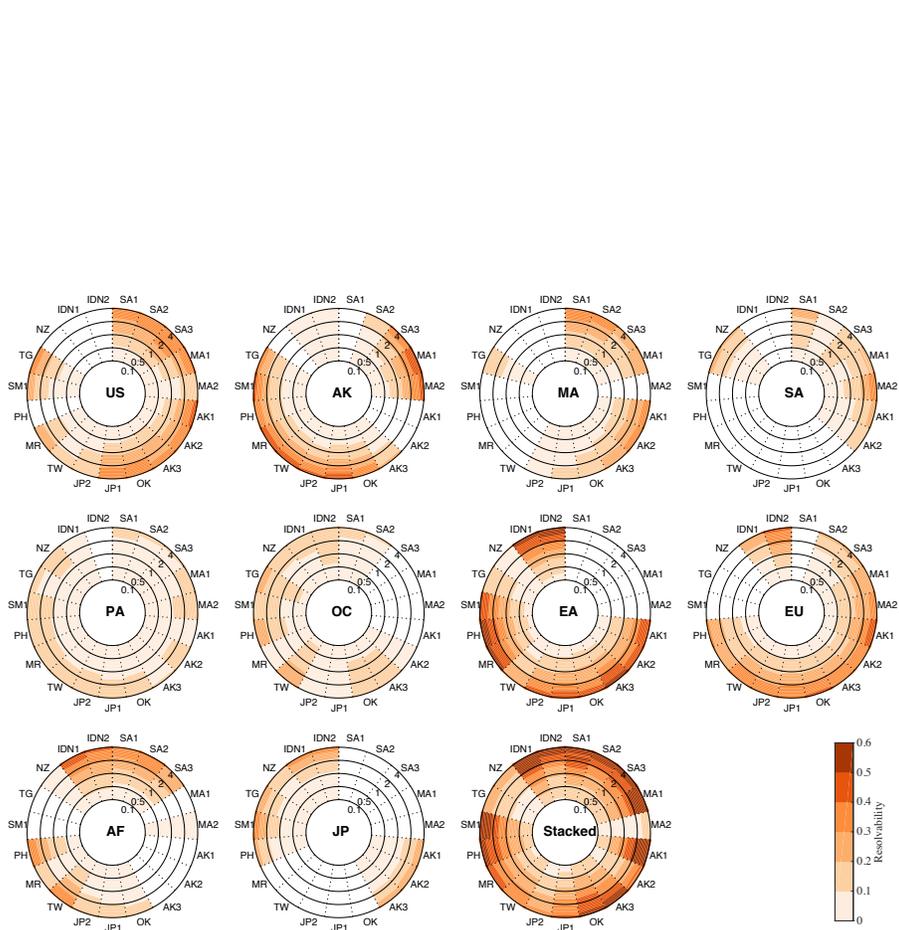
seismic moment magnitude of these average earthquakes to be of  $M_w$  8, above which BP method using teleseismic data works best (further discussion on this part in the later section).

As for the distributions of the stations, I download the locations of all available stations from IRIS SeismicQuery website (<https://ds.iris.edu/SeismicQuery/station.htm>) and NIED Hi-net websites (<http://www.hinet.bosai.go.jp/>). Then, I cluster all these stations into large arrays. These arrays, including all temporary array stations, provide the ideal data coverage to apply the BP methods (Fig.4.6).

This study aims to provide an informed recommendation on the resolvability of the BP images given the source-receiver location. I first take the example of the Hi-net seismic array, a high quality dense seismic array<sup>239,236,138,320</sup> and then provide a global perspective.

Fig.4.7 (a) shows an example of the seismic active regions within teleseismic distances ( $30^\circ$ - $90^\circ$ ) of the Hi-net array in Japan. The Hi-Net array can cover many major subduction zones including Indonesia (IDN<sub>1</sub>-Java and IDN<sub>2</sub>-Sumatra), Philippine (PH), Solomon (SM), Tonga (TG), and Alaska (AK<sub>1</sub>-Aleutian and AK<sub>3</sub>) subduction zones. In addition, there are also two transform plate boundaries in New Zealand (NZ) and Alaska (AK<sub>2</sub>). The average focal mechanisms shown in Fig.4.7 (a) are consistent with the geometry of the plate boundaries. I set the size of the potential source regions to be horizontal  $4^\circ \times 4^\circ$  planes discretized with  $32 \times 32$  grid points and choose the average depths of the  $M_w$  7.5+ earthquake sources. Travel times from each grid point source to each station are computed using the IASP91 model<sup>150</sup>. I then use Eqs.(4.4) and (4.5) to calculate the resolution matrix  $\mathbf{F}(\omega)$  and the corresponding resolvability from Eq.4.7. I focus on the frequency band from 0.1 to 4 Hz that is often used in backprojection studies. Fig.4.7 (b) shows the resolvability of the Hi-Net array toward all source regions. The resolvability is quite low below 1 Hz but rapidly improves at higher frequencies. Hi-net array can well resolve sources in Sumatra, Solomon, and Alaskan subduction zones. But it does not work well for the New Zealand (NZ) region because it is located too close to the nodal plane, which is the similar case as shown in my synthetic test results for the Set2 arrays.

I then show the resolvability distributions of all global seismic arrays in Fig.4.8. The systematic increase in resolvability with frequency is notable at all arrays and for all sources. Most of the large scale and dense arrays (USA (US), Eurasia (EA), Europe (EU), and Africa (AF)) have good resolvability to most source regions.



**Figure 4.8:** The same resolvability as Fig.4.7 (b) but for all global arrays. The location of arrays as well as the source regions can be found in Fig.4.6. The resolvability of all stacked array is shown in the bottom right.

## 4.4 DISCUSSION

### 4.4.1 USING THE LINEAR BP IMAGE RESULTS TO EXPLAIN EARTHQUAKE RUPTURE

The theoretical formulation as well as the synthetic tests on complex kinematic sources help me to better interpret the BP images in light of earthquake kinematics. Since the displacement seismograms are determined mainly by integrating the slip-rate functions over the fault plane from Eqs.(4.1) and (4.2), the linear BP results constructed from the synthetic seismograms correspond well to the slip motions, i.e. the slip rates for displacement seismograms (this study) or slip accelerations for velocity seismograms. In the frequency domain, the BP image at each narrow frequency is actually consistent with the slip motion distribution filtered around that frequency (see Fig.A.4 in the Appendix A), consistent with my theoretical formulation Eq.4.6. However, in frequency domain BP, the displacement BP image and velocity BP image at the same frequency  $\omega$  ought to be proportional  $i\omega$ .

As indicated in Fig.4.5 (a) and (b), for the four single arrays and the composite one, the average correlation coefficients between the average BP image and filtered slip-rate distribution generally varies between 0.1 to 0.6. This range of CC indicates that the BP method can recover relatively well the first order features of slip motion such as the slip peaks and spatial extent of rupture. The direct comparison between BP peaks and peak slip rates in Fig.A.3 (a) can validate this consistency.

Since the peak slip rate always occurs slightly behind the true rupture front, my theoretical formulation and synthetic tests indicate that the BP image can give a good estimation on, at least, the lower limit of the average rupture velocity. In some specific cases, even the detailed changes of rupture velocity during an earthquake rupture can be possibly observed<sup>325,358</sup> given the good resolvability (Fig.4.8). Given the variations in rupture velocity that are estimated with the source-receiver geometry, I suggest that the rupture velocity obtained from BP studies is a robust lower limit estimation of the earthquake rupture velocity. On the other hand, the large variability of CC values

(0.1-0.6) and lack of perfect value ( $CC=1$ ) imply that the BP results cannot recover the exact slip history. I attribute this due to the shape of the resolution matrix  $F(\omega)$  that is not proportional to identity. A critical element of conventional BP is whether the waveforms can constructively or destructively interfere in the stacking. Low frequency waveforms have a wider sensitivity zone and are likely to interfere within a large source region (e.g. Fig.4.1 (a) and (d)), which further lowers the resolvability. On the other hand, the observed high-frequency data is limited due to attenuation and the non-stationary station coverage. This can be clearly quantified by the spectrum of BP resolvability. For instance, the resolvability of BP images constructed from the Hi-net stations of an earthquake in the Indonesia region (IDN1 or IDN2) increases from 0.1 at 0.1 Hz to 0.75 at 12 Hz (Appendix A Fig.A.5). However, seismic attenuation in the mantle constrains the upper observable teleseismic frequency to a maximum of 4 Hz, above which the signal-to-noise ratio of teleseismic seismograms is very low<sup>331,332</sup>.

Therefore, I conclude that the BP images derived from raw seismic data, if corrected for attenuation, are proportional to the slip-rate field after a spatial smoothing, which can be parameterized by the resolution matrix  $F(\omega)$  (Fig.4.1). This is similar to the conclusions of Fukahata et al.<sup>99</sup> that the BP image represents the slip motion on a fault, provided that the Green's function is sufficiently close to a delta function. My results, however, show that the Green's function cannot realistically be a delta function, but that general features of the slip motions may be recovered within limited frequency bandwidth.

The relation between the BP image and the kinematic source process provides a unique way to infer the slip behaviors in the relatively higher BP frequency band: the high frequency components of the slip history, parameterized either with slip rate or acceleration, are sensitive to the sudden change of rupture propagation<sup>193,194</sup> and thus can be used to estimate the overall pattern of rupture propagation such as the rupture extend or lower limit of rupture velocity.

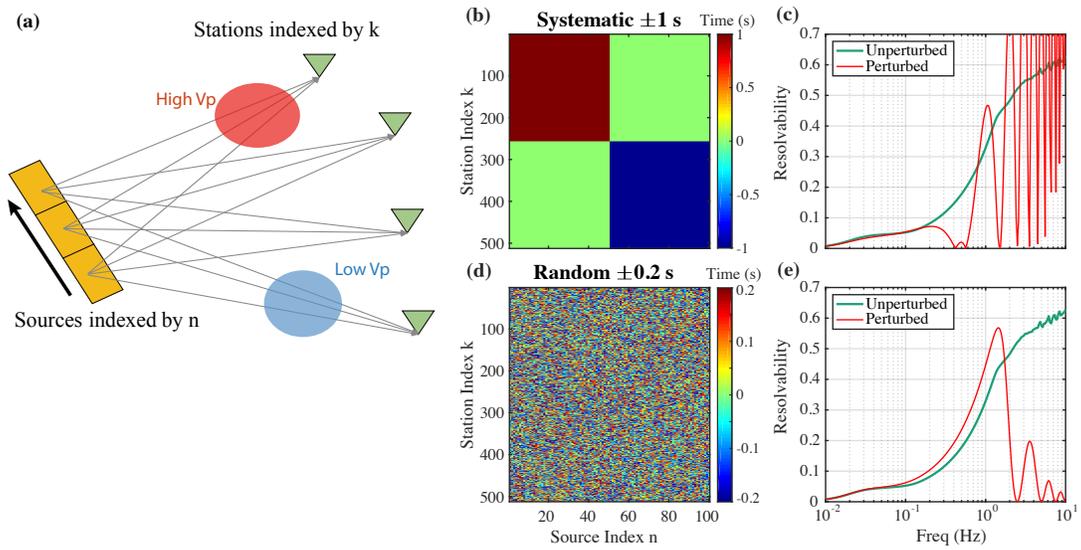
Many previous studies on the megathrust events reveal a frequency-depth relation of the seismic

radiation coherence in the BP results <sup>175,205,293,324,351,359</sup>. Interestingly, this pattern cannot be clearly observed in my synthetic kinematic sources: neither in the filtered slip motion distribution nor in the BP images (see Fig.A.4 in the Appendix A). This implies that the occurrence of frequency-dependent seismic radiation may require additional source heterogeneities that would cause systematic spatial variations of rise times or slip-rate functional forms, but that are not modeled in my kinematic source. These heterogeneities may be better modeled with realistic dynamic models that account for pre-stress <sup>129</sup>, friction <sup>257,277</sup>, fault geometry <sup>195</sup> or even inelasticity effects <sup>190</sup> along dip direction, not included in this study but worth further investigations.

Finally, I discuss the spectral decay of the BP amplitudes. At any time, the BP amplitude decays with frequency in a way that is similar to the source spectral decay (Fig.A.6 in the Appendix A). The high frequency spectral falloff rate of this BP peak amplitude spectrum from linear regression varies from 2.9 to 3.3 for all four arrays in my synthetic tests with kinematic source (Fig.A.6 in the Appendix A). Considering the spectral falloff rate of the sliding time window, which is 1 for the Tukey taper used in this study, the corrected source spectral falloff rate estimated from BP peak amplitude can be 1.9 to 2.3 in 0.1-1 Hz. The falloff rates of BP peak amplitude spectra are roughly consistent with the spectral falloff rates of the sources, 2.2 (Fig.A.1 in the Appendix A) for the kinematic model. In practice and for further interpretation of the spectral decay in terms of rupture process, a correction of the amplitude for high frequency attenuation is required and remains challenging.

#### 4.4.2 DISCUSSION ON 3D STRUCTURAL EFFECTS ON RESOLVABILITY $\varepsilon_f(\omega)$

In this study, the theoretical relation of Eq.4.6 is described in a homogeneous full space and so I assume an ideal case that I can perfectly correct the travel time: the travel-time terms in the matrices  $\tilde{\mathbf{A}}(\omega)$  and  $\mathbf{A}(\omega)$  are equal. Under this assumption, I have ignored the uncertainty of the travel-time corrections that may present in practice. While both source and path complexity affect the seismograms, my primary motivation of this study is to map the source complexity with idealized path



**Figure 4.9:** Generalized resolvability with travel time perturbation. (a) A schematic cartoon shows the structural heterogeneities along the ray path. The yellow rectangle is a source with 3 grids indexed by  $n$ . Black arrow shows the rupture propagation. Green triangles are the stations indexed by  $k$ . The two circle patches show the velocity anomalies along the ray paths from each source grid to stations. (b) - (c) The systematic travel-time perturbation matrix  $\Delta t_{kn}^1$  and its corresponding resolvability, respectively. (d) - (e) The random travel-time perturbation matrix  $\Delta t_{kn}^2$  and its corresponding resolvability, respectively. In (c) and (e) the green lines are the Away1 array resolvability, the same as shown in Fig.4.3 (c) and the red lines are the corresponding resolvability from travel time perturbations.

terms. In this section, I briefly address the impact of path complexity on resolvability. The concept of BP resolvability  $\varepsilon_I(\omega)$  is to propose an upper bound of my confidence in the BP images, i.e., to what extent I can recover the source kinematics from BP images.

A first element I can incorporate is a variable contributions of stations and arrays. For example, the relation of Eq.4.6 ignores the weighting matrix  $\mathbf{W}$ . I can generalize the resolution matrix  $\mathbf{F}^g(\omega) = \tilde{\mathbf{A}}(\omega)\mathbf{W}\mathbf{A}(\omega)$  to account for the waveform normalization, different array contributions and polarity reversal.

A second element I can incorporate is travel-time uncertainty due to the unknown 3D structure. In realistic situation, the travel-time terms in the wave propagation matrix  $\mathbf{A}(\omega)$  and the BP phase shift matrix  $\tilde{\mathbf{A}}(\omega)$  are different: in the former the  $t_{kn}$  is the true travel time while in the latter  $t_{kn}$  is a theoretical estimate. To account for this difference, I note  $\tilde{\mathbf{A}}(\omega)$  to be  $t'_{kn}$  as the theoretical travel time and regard the  $t_{kn}$  in  $\mathbf{A}(\omega)$  as the true travel time. For example, the diagonal phase-shift terms now become  $F_{nn}(\omega) = \sum_k \frac{R_{kn}^p}{r_{kn}} e^{j\omega(t'_{kn} - t_{kn})}$ . Then, I can model uncertainties in travel time due to my limited knowledge of the Earth structure, in particular for small length-scale anomalies rays travel through.

To simulate these effects on the BP resolution, I design two different kinds of travel time perturbations, one that is far-field systematic shift, one that is typical of local site effects (Fig.4.9 (a)). I use the synthetic setting of Away1 array as an example. I add the travel-time perturbations as  $\Delta t_{kn} = t'_{kn} - t_{kn}$ , re-construct  $\mathbf{F}(\omega)$  as well as the resolvability. The first uncertainty  $\Delta t_{kn}^1$  is a systematic travel-time shift of  $\pm 1$  s added to half of the source-receiver pair (-1s for 1/4 and +1s for the other 1/4, see Fig.4.9 (b)). The second kind of perturbation  $\Delta t_{kn}^2$  is a simple random shift taken from a uniform distribution with maximum amplitude of 0.2 s (Fig.4.9 (d)).

Both types of uncertainty impact the resolvability. The systematic perturbation causes significant fluctuations in the resolvability (Fig.4.9 (c)): the resolvability drops at specific frequencies. Because these time shifts act as waveform re-alignment, it is likely that the alignment and stacking produce

spurious arrivals, shifted by the uncertainty that interfere constructively or destructively at the specific frequencies harmonic to the inverse of the uncertainty phase shift. Intuitively, it is similar to taking the Fourier transform of a time series with two pulses<sup>75</sup>. This large effect in the resolvability yields a systematic location bias in the BP images (Appendix A Fig.A.7 (b) and (e)). On the other hand, the random perturbation has little effect on the resolvability at low frequency and even provides even a higher resolvability (Fig.4.9 (e)). This is because the incoherent part of waveforms can be better destructively stacked after adding this random perturbation. The random perturbation becomes rough but also slightly “sharpens” the edge of BP images (Appendix A Fig.A.7 (c) and (f)), thus leads to relatively higher resolvability. However, it causes a steep decrease of the resolvability at the high frequency, indicating a severe loss of waveform coherency and poor resolution on the short-wavelength features.

My tests confirm that travel-time uncertainty can greatly influence the resolution in BP images. Besides, these tests also suggest a high frequency cutoff of applicability of the BP techniques of 2 Hz in this test, given a 0.2 s travel-time uncertainty. This factor, together with the structural attenuation, poses an upper limit on the frequency of BP technique. In real applications, many efforts have been devoted to better corrections on structural effects, using theoretical or empirical methods<sup>139,210,208</sup>.

A third element present in 3D structure are the near-source body-wave reflections such as depth phases<sup>172,333,75,358</sup> and water reverberation<sup>60,8,362</sup> that are particularly visible in megathrust events. It is possible to include these phases in a more generalized wave propagation matrix  $\mathbf{A}^g(\omega)$  as a linear summation of the phases (e.g., see Eq.(6) in Yin et al.<sup>358</sup>):

$$\mathbf{A}^g(\omega) = \mathbf{A}^P(\omega) + \mathbf{A}^{pP}(\omega) + \mathbf{A}^{sP}(\omega) + \dots \quad (4.10)$$

Then the corresponding BP phase-shift matrix would be:

$$\tilde{\mathbf{A}}^g(\omega) = \tilde{\mathbf{A}}^P(\omega) + \tilde{\mathbf{A}}^{pP}(\omega) + \tilde{\mathbf{A}}^{*P}(\omega) + \dots \quad (4.11)$$

Interferences and coherence among depth phases will appear in the generalized resolution matrix as the product of these summed matrices. The arrival times of depth phases and water reverberation are source-specific and a rather systematic parameter space study of these effects are left for future work. Nonetheless, this scheme is theoretically simple and may be useful in the future to better evaluate how BP can work under the more realistic conditions.

#### 4.4.3 RELATION TO OTHER IMPROVED BP TECHNIQUES

The relation shown in Eq.4.6 provides a fundamental framework between BP images and slip-rate field, provided that the stacking scheme is linear. This well motivated my work and also has been emphasized in previous studies<sup>155</sup>. Sophistication of the data processing that loses the linearity in Eq.4.6 is attempting to improve image resolution: for instance, the use of sparsity regularization in Compressive Sensing BackProjection (CSBP)<sup>349,358,360</sup>, hybrid backprojection (HyPB)<sup>99,346</sup> and the  $n^{th}$  root stacking processing<sup>209,262,344</sup>.

Inspired by techniques developed in signal processing and applied mathematics communities, Yao et al.<sup>349</sup> develops a compressive sensing BP method, CSBP, to invert for a sparse distribution of the source  $\dot{\mathbf{U}}(\omega)$  (or  $\ddot{\mathbf{U}}(\omega)$  from velocity seismograms) in Eq.4.3. Since this system is under-determined ( $K \ll N$ ), I cannot get a unique solution without smoothing constraints. The basic assumption of CSBP is that the source distribution is sparse in space so the problem is solved via optimization,

$$\mathbf{U}^{CS}(\omega) = \underset{\mathbf{U}}{\operatorname{argmin}}\{\|\mathbf{D}(\omega) - \mathbf{A}(\omega)\mathbf{U}(\omega)\|_{l_2} + \lambda\|\mathbf{U}(\omega)\|_1\}, \quad (4.12)$$

where  $\lambda$  is a damping factor chosen to balance the contributions of data misfit (first term in right hand side) and model constraint (second term in right hand side). Instead of directly aligning and stacking in a sense of “grid-search” like conventional BP, CSBP is based on an inversion scheme that attempts to directly solve for the source  $\mathbf{U}(\omega)$  with the specific constraint of sparsity. The advantage of sparsity constraint is its relatively high spatial resolution. The sparsity constraint helps to accurately locate the sub-events, especially when limited by lower seismic frequencies. Ignoring the damping required to balance data and model misfit, the CSBP is equivalent to the sparse solution of  $\mathbf{U}^{BP}(\omega)$  in Eq.4.6, constrained by the data. The sparse representation inevitably eliminates details about the source but can provide more robust locations of the dominant sources. This latter effect is practical when the spatial resolution of conventional BP method is relatively poor (see Fig.A.8). Similarly, I can also look at the CSBP peaks and estimate the average rupture velocity (see Fig.A.3 (b) and (d) in the Appendix A for instance). Overall, CSBP provides a sparse fit to the slip-rate field.

The Hybrid BP technique<sup>346</sup> is another improved BP technique that can be clearly discussed under the framework in this study. In my study, the alignment of the waveform,  $\mathbf{A}(\omega)$ , is carried by simplifying the Green’s function to a shifted delta function, that is, directly time/phase shifting without changing waveforms. The HyBP, however, incorporates the full Green’s function in  $\mathbf{A}^{thG}(\omega)$ , as a combination of slip inversion with conventional BP techniques. The basic assumptions are that the cross-correlation between the theoretical Green’s function and real Green’s function can be approximated to the auto-correlation of real Green’s function, and that it is sufficiently close to a delta function<sup>99</sup>. If these assumptions are satisfied, the cross-correlation function can directly reflect the slip motion occurring at the source thus I can use the HyBP to directly recover the slip motion. The calculation of the cross-correlation function is equivalent to multiplying a  $N \times K$

cross-correlation matrix  $\mathbf{A}^{tbG}(\omega)$  to the left-hand side of Eq.4.3:

$$\mathbf{A}^{tbG}(\omega)\mathbf{D}(\omega) = \mathbf{A}^{tbG}(\omega)\mathbf{A}(\omega)\mathbf{U}(\omega). \quad (4.13)$$

The elements of  $\mathbf{A}^{tbG}(\omega)$  are:

$$(\mathbf{A}^{tbG}(\omega))_{nk} = C_{nk} \frac{R_{kn}^P}{r_{kn}} e^{+i\omega t_{kn}}, \quad (4.14)$$

where  $C_{nk}$  is a normalization constant for the cross correlation and the subscripts  $k = 1, 2, \dots, K$  and  $n = 1, 2, \dots, N$  correspond to the station and source index for the theoretical Green's function, respectively; other terms have the same notation as in Eqs.4.1 - 4.4. Compared with Eqs.(4.4) and (4.5), the HyBP method effectively consists of the multiplication  $\mathbf{A}^{tbG}(\omega) = \mathbf{A}^H(\omega)$ , the conjugate transpose of  $\mathbf{A}(\omega)$  to the spectral data, and forms a new resolution matrix  $\mathbf{F}^{HyBP}(\omega) = \mathbf{A}^H(\omega)\mathbf{A}(\omega)$ . Based on my theoretical formulation Eqs.4.4 - 4.5 and 4.13 - 4.14, the HyBP method is the same as Linear BP in the frequency domain, except for its resolution matrix  $\mathbf{F}^{HyBP}(\omega)$ . Both methods can be interpreted as cross-correlation: linear BP is the result from cross-correlation with a phase-shifted delta function  $\delta(t - t_{kn})$  while the HyBP is the outcome from cross-correlation with the theoretical Green's function. The difference in the BP results due to their respective resolution matrices is negligible (Appendix A Fig.A.9). In practice, the cross-correlation with an accurate Green's function can potentially suppress incoherent noise and thus enhance the signal levels of the source waveforms. However, basic assumptions of HyBP are difficult to satisfy: (i) accurate theoretical Green's functions are difficult to compute due to limited knowledge of structure and computation cost of high frequency wave propagation; (ii) even the theoretical Green's function is equal to real Green's function, the auto-correlation of a Green's function is not exactly a delta function due to finite-frequency effects. Therefore, the BP images from HyBP are still not the perfect match to slip motion on the fault surface.

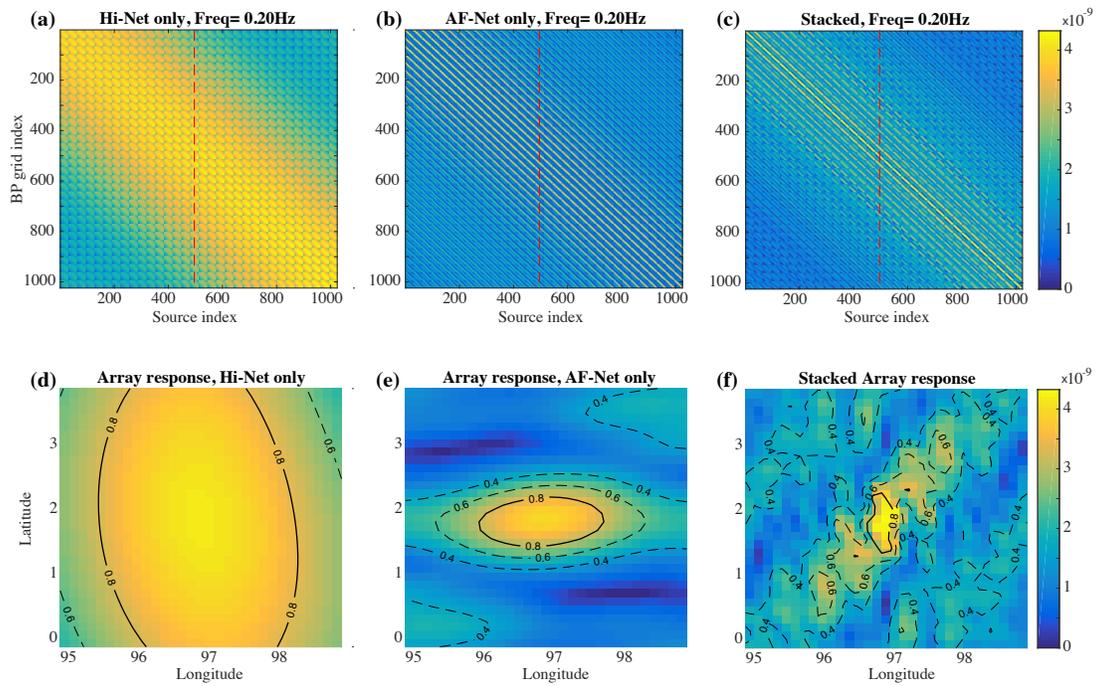
Finally, I briefly discuss the popular non-linear stacking schemes.  $n^{th}$  root stacking<sup>209,262,344</sup> is another classical beamforming technique. It first calculates the  $n^{th}$  root ( $n=2,3,4,\dots$ ) of the seismogram in Eq.4.1 before stacking. This power-law processing removes the linearity between slip-rate and displacement waveforms, and thus I have already lost the information about the slip motion in the data. However, it is practical to enhance phase coherency<sup>262</sup> and thus to provide better resolution of radiation locations. To sum up, the  $n^{th}$  root stacking can definitely improve the resolution of BP image but in order to keep the slip information about the source (dimension of slip motions), linear stacking is necessary.

#### 4.4.4 GLOBAL ARRAY STACKING AND FREQUENCY RESOLUTION

Nowadays, there are several available seismic arrays within the teleseismic distance of a given earthquake. This allows me to combine multiple arrays and improve the array response and resolution of BP method. The BP stacking over multiple arrays has been applied in various recent studies<sup>366,252</sup>. Here, I relate the multiple arrays stacking to my theoretical formulation and indicate how well it improves the BP results.

The shape of resolution matrix itself carries information about the data resolution given a source-receiver geometry. Each column of  $F(\omega)$  corresponds to the array response<sup>262,344</sup> of a seismic array toward a single grid point source at a specific seismic frequency. The array response is determined by both the azimuth and distance coverage<sup>155</sup>, and a wide azimuth-distance coverage lead to the different distributions of array response.

For example, Fig.4.10 shows the resolution matrices of Hi-Net and AF-Net arrays as well as their array responses at 0.2 Hz for the IDN2 region, where the 2004 Sumatra earthquake occurred. Ishii et al.<sup>138</sup> use the Hi-net array to recover the rupture process of this event. However, the array response of Hi-net array shows a north-south distributed patch (Fig.4.10 (d)) and the size of this patch is very large due to the limited coverage of Hi-net array. On the other hand, if there had been



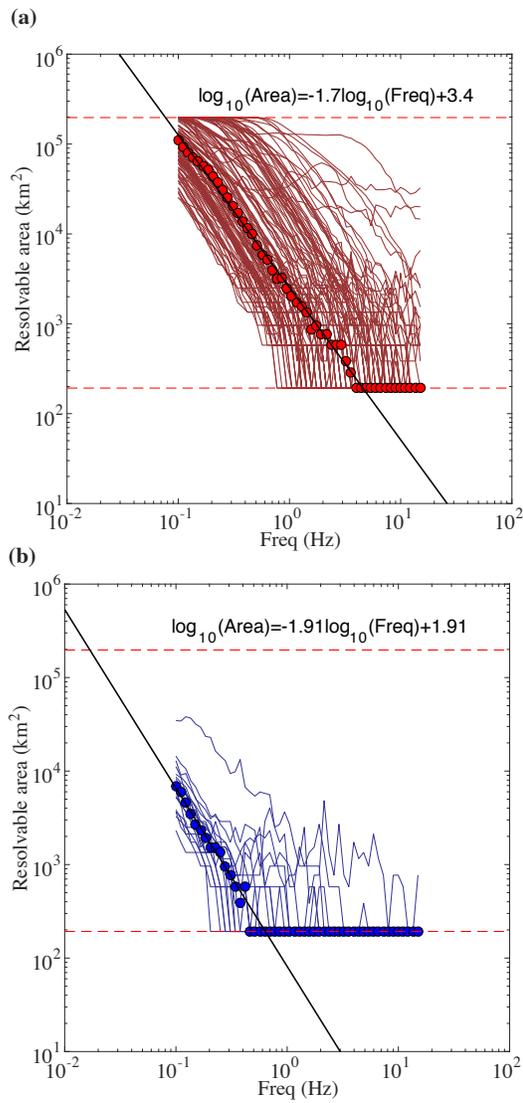
**Figure 4.10:** Example of array response varying with array locations and the improvement from array stacking at a given seismic frequency. Absolute value of resolution matrices of Hi-Net (a) and AF arrays (b) toward the region IDN2 at 0.2 Hz. Array responses at a point source location corresponding to Hi-Net (c) and AF (d) arrays. Absolute value of stacked resolution matrix from all available arrays within teleseismic distance to the region IDN2 (AK, OC, EA, EU, AF, JP) (e) and the corresponding array response at the same source location (f). Areas within the 0.8 contours of array response distribution will be used to estimate the resolvable areas of Fig.4.11.

enough high quality stations in Africa, the corresponding array response at the same point is east-westward distributed with smaller size (Fig.4.10 (e)) due to better spatial coverage. Moreover, the resolution matrices of these two arrays are different at most locations but both have peak values at the diagonal parts of the resolution matrix (Fig.4.10 (a) and (b)). This is actually the basis of the multiple-array stacking that can improve the convergence of the resolution matrix to a diagonal matrix. For the IDN<sub>2</sub> region, I stack the resolution matrices of all the available arrays (JP (Hi-Net), AF, OC, AK, EA, EU arrays). The resolvability is greatly improved (Fig.4.10 (c) and (f)), even at low frequency. For all other regions, the improvements of resolvability are all obvious (last sub-figure in Fig.4.8).

For each array response matrix, i.e. for each source-array configuration, and at each seismic frequency, I attribute as resolvable area as the integrated area within 80% of the peak array response function (Fig.4.1 (d) - (f) and Fig.4.10 (d) - (f)). For all source-array configurations and at all frequencies, I construct an empirical relation between the spatial resolution and the frequency of the data. All available configurations are shown in Fig.4.11 (a) and exhibit unique levels of resolvability, whereby the resolvable area decreases with seismic frequency, and equivalently, spatial resolution increases. By taking the median of individual area measurement at each frequency, I construct an empirical relation between the BP resolvable area  $S_0^{\text{BP}}$  (in km<sup>2</sup>) and seismic frequency  $f$  (in Hz) as a power law of seismic frequency:

$$S_0^{\text{BP}} \approx 10^{3.4} f^{-1.7}. \quad (4.15)$$

In the ideal case that each source region can be well recorded by all available arrays, I proceed by stacking over arrays to increase resolvability (Fig 4.11 (b)). The optimal median resolvable area  $S^{\text{BP}}$  - seismic frequency  $f$  is:



**Figure 4.11:** Resolvable area as a function of frequency. Red dashed lines corresponds to the minimum (grid size) and maximum (total) area of the source region. (a) Thin red curves show the frequency-varying resolvable area from each array-source region pair. Red circles corresponds to the median value of resolvable area at each frequency. Black line shows the relation between median resolvable area and frequency from linear regression. (b) Blue curves show the variation of the resolvable area as a function of seismic frequency from multiple-array-stacking for each source region. Blue circles corresponds to the median value of the stacked resolvable area at each frequency. The black line shows the best-fit relation from linear regression to the median values.

$$S^{\text{BP}} \approx 10^{1.91} f^{-1.91}. \quad (4.16)$$

An additional practical consideration is that of earthquake size scaling. If the fault length is  $L = (S^{\text{BP}})^{1/2}$ , then  $L \approx 10/f \approx 2V_p/f = 2\lambda_p$ . That is, my empirical relation implies a twice P wavelength resolution for the BP. Given scaling between fault length and earthquake magnitude  $M_w$  provided by Table 2A of Wells & Coppersmith <sup>338</sup>,

$$S \approx 10^{(-3.42+0.9M_w)}. \quad (4.17)$$

In order to resolve the rupture propagation, the BP resolvable area  $S^{\text{BP}}$  should be smaller than the total rupture area. For example, if  $S^{\text{BP}} \leq S/10$  is required, I can build a relation between earthquake magnitude and lowest BP frequency  $f_{min}^{\text{BP}}$  required to resolve source features:

$$f_{min}^{\text{BP}} \approx 10^{(3.31-0.47M_w)}. \quad (4.18)$$

In order to resolve the source features of a  $M_w$  8 earthquake using multiple-array BP, the lowest seismic frequency required is approximately 0.35 Hz; 1.02 Hz for a  $M_w$  7, and 3.02 Hz for a  $M_w$  6 earthquake. Because the relation Eq.4.18 between rupture area and earthquake magnitude from Wells & Coppersmith <sup>338</sup> is mostly from continental earthquakes, the actual rupture area of megathrust events in the subduction zones can be larger. Therefore, the corresponding lowest BP frequency can be smaller than the value predicted from Eq.4.18 when BP is applied to the megathrust events.

This purely empirical relation only provides crude guidelines on the lower bound of the BP frequency analysis. Further considerations such as attenuation, structure, signal levels will impact the upper bound frequency.

## 4.5 CONCLUSION

My theoretical formulation of the linear backprojection algorithm indicates that the BP image is indeed related to the slip motion on the fault, granted a spatial smoothing. A resolvability parameter, which I defined as the norm of the resolution matrix, provides a metric to evaluate the spatial resolution of backprojection method for a specific source-receiver geometry. I further test the BP method on a synthetic kinematic source to validate the theoretical formulation. The synthetic tests indicate that the BP image can provide a reliable estimation on the general pattern of rupture propagation.

In addition, I estimate the strengths and limitations of the linear BP algorithm in light of realistic source and seismic array configurations. I find that stacking arrays considerably increases the resolution thereby reducing the resolvable area. Finally, I construct a relation between resolvable area and seismic frequencies. Given the scaling of earthquake size with source length, my analysis provide simple guidelines to the lower bounds of seismic frequencies required to image details of the source provided earthquake magnitude.

## 4.6 DATA AND RESOURCES

Station information is obtained from IRIS SeismiQuery website (<https://ds.iris.edu/SeismiQuery/station.htm>) and NIED Hi-net official website (<http://www.hinet.bosai.go.jp/>). Focal mechanisms of historical earthquakes are downloaded from GCMT website (<http://www.globalcmt.org/>). Most of the data processing, calculations and figures are carried out using Matlab and some map figures are made using GMT (the Generic Mapping Tools, <http://gmt.soest.hawaii.edu/projects/gmt/wiki/Download>)

5

# The Earth's Surface Controls the Depth-Dependent Seismic Radiation of Megathrust Earthquakes

## SUMMARY

The largest earthquakes occur on the megathrusts of subduction zones and generate huge ground motions and devastating tsunami waves that threaten the coastal populations. Via seismic observations on recent megathrust earthquakes, those earthquakes exhibit a ubiquitous seismic radiation style: low-frequency (LF) seismic energy is efficiently emitted from the shallowest portion of the fault, whereas high-frequency (HF) seismic energy is efficiently emitted from the deepest part of the fault. Although this is observed in many case-specific studies, I show that it is ubiquitous in global megathrust earthquakes between 1995 and 2021. Previous studies have interpreted this as an effect of systematic depth variation in either the plate interface frictional properties<sup>175</sup> or the P wavespeeds<sup>270</sup>. This work suggests an alternative hypothesis: the interaction between waves and ruptures due to the Earth's free surface is the leading mechanism that generates this behavior. Two-dimensional dynamic rupture simulations of subduction zone earthquakes support this hypothesis. My simulations show that the interaction between the seismic waves reflected at the Earth's free surface and the updip propagating rupture results in LF radiation at the source. In contrast, the downdip propagation of rupture is less affected by the free surface and is thus dominated by HF ra-

diation typical of buried faults. To a second degree, the presence of a realistic Earth structure derived from P-wave velocity ( $V_P$ ) tomographic images and realistic  $V_P/V_S$  ratio estimated in boreholes further enhances the contrast in source radiation. I conclude that the Earth's free surface is necessary to explain the observed megathrust earthquake radiation style, and the realistic structure of subduction zone is necessary to better predict earthquake ground motion and tsunami potential.

## 5.1 INTRODUCTION

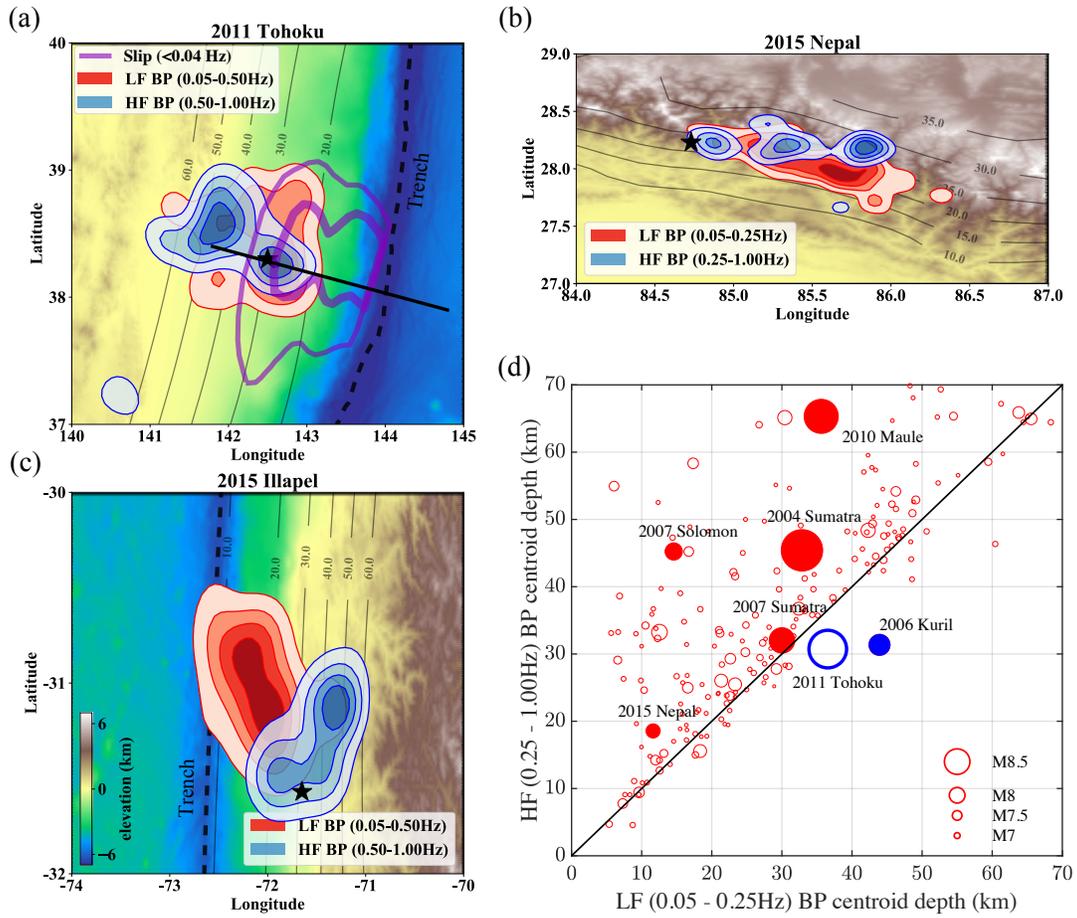
The largest and most damaging earthquakes occur offshore in subduction zones: the Mw 9.4 1960 Great Chilean earthquake, the 1964 Mw 9.3 Great Alaskan earthquake, the Mw 9.2 2004 Sumatra earthquakes, and the Mw 9.0 2011 Tohoku-oki earthquake. Because almost 1 in 10 people in the world live on the coast, understanding the rupture behavior of megathrust earthquakes is critical for seismic and tsunami risk mitigation in coastal areas. The recent occurrence of multiple of these events has coincided with a vast expansion in seismic networks, which, in turn, has led to the discovery of a multitude of processes surrounding the rupture of these large earthquakes<sup>138,175</sup>.

A remarkable observation of these earthquakes' seismic signature is that low-frequency (LF) seismic waves are mostly generated at the shallow, updip region, while high-frequency (HF) seismic waves tend to come from the deep, downdip part. I refer to this as the "depth-frequency relation" in this work. It is manifested in three ways. First, studies on earthquake source time functions highlight a shortening of the source pulse that is well explained by an increase in elastic moduli with depth<sup>37,121,309</sup> and an increase in the relative contributions of HF radiation at depth and along the megathrust<sup>58,355</sup>. Second, the strong ground motions that are responsible for damaging urban infrastructure have been observed to originate from the downdip end of the megathrust<sup>19,97,171</sup>. The third class of seismic observations is the back-projection (BP) image reconstructed from teleseismic P waves<sup>138</sup>. The BP image is effectively a blurred representation of the slip history on the

fault<sup>99,357</sup>. Consequently, the images constructed at various frequency bands relate to the slip function's whole-event spectral content on the fault. Event-specific studies have shown that high frequencies are more efficiently generated at the downdip portion of the megathrust rather than its updip end<sup>156,205,209,289,293,351,358,359,361</sup>.

Here, I show three examples of such images using an Improved Compressive Sensing BackProjection (ImCS-BP) method<sup>349,358</sup> for the Mw 9.0 2011 Tohoku-oki earthquake<sup>175,324,349</sup>, the Mw 7.9 2015 Gorkha earthquake<sup>22,361,364</sup>, and the Mw 8.3 2015 Chilean Illapel earthquake<sup>205,359</sup>. I show both the LF and HF BP images in Fig. 5.1 (a) - (c). These images clearly illustrate that HF source signals are emitted at greater depths than LF source signals. I then turn to global databases of BP images provided by The Incorporated Research Institutions for Seismology (IRIS) over all the Mw 6.5+ earthquakes since 1995<sup>137</sup>. Here, I select 461 earthquakes between 1995 and 2021 within the latitude-longitude range of the available Slab2 plate interface model<sup>114</sup>. I then project the HF and LF BP peaks of each earthquake onto the Slab2 model and calculate the corresponding HF and LF centroid depths. For most earthquakes, especially the large magnitude ones with a likelihood of better time and spatial resolution of the BP image, I find that the centroid depth of the HF BP peaks is systematically greater than that of the LF peaks (Fig. 5.1 (d) and Fig. 5.2). Two events stand out as exceptions: the Mw 9.0 2011 Tohoku-oki earthquake and the Mw 8.3 2006 Kuril Island earthquake<sup>12</sup>. For the 2011 Tohoku-oki earthquake, the exception is due to the different choice of frequency bands by the IRIS database, and I have shown that the refined BP results clearly present the depth-frequency relation (Fig. 5.1 (a), or figures in Yao et al.<sup>349</sup>).

A common interpretation for these observations is the systematic depth variation in frictional properties that result from increasing temperature and pressure with depth and associated phase transformation of the minerals that compose the downgoing oceanic lithosphere. The argument is that systematic depth variations in fault properties can explain the evolution of the seismicity rates with depth<sup>277</sup>. It has also been widely used to explain the depth-varying seismic radiation of large



**Figure 5.1:** Ubiquitous depth-frequency relation found by back-projection observations. (a)-(c) BP images of the Mw 9.0 2011 Tohoku-oki, the Mw 7.9 2015 Gorkha, and the Mw 8.3 2015 Illapel earthquakes, respectively. The BP images are reconstructed using the ImCS-BP method developed by Yin et al. <sup>358</sup>, and only the contours of 20%, 40%, 60%, and 80% maximum power are shown. The dashed black lines indicate the trench. The thin gray contours show the Slab2 model <sup>114</sup>. The purple contours in (a) show the 20 m and 50 m of coseismic slip distribution during the 2011 Tohoku earthquake from Lay et al. <sup>175</sup>, and the black solid line shows the location of the velocity profile of Miura et al. <sup>213</sup>. (d) Centroid depths of the low-frequency (0.05 - 0.25 Hz) BP images compared with the high-frequency (0.25 - 1 Hz) BP images from 245  $M > 6.5$  earthquakes.

megathrust earthquakes<sup>175,351,361</sup>. Studies that simulate the dynamic rupture have adopted this with a parameterization of pre-stress or fault strength heterogeneity in the deeper portion of the seismogenic megathrust and have successfully reproduced HF and LF's relative contributions in seismic radiation<sup>103,129</sup>. Other studies have shown that it may be explained by a depth dependence in fault rheology, whereby the transition of frictional behaviors occurs, result in HF radiation at the rupture front<sup>211,234</sup>. A recent alternative interpretation is that the systematic increase in P wavespeed ( $V_p$ ) with depth in subduction zones directly impacts the wavelength and frequency of seismic waves emitted at the source<sup>270</sup>. However, such an argument would also pertain to earthquakes in a wide depth range and from other tectonic environments. But I do not observe it for deeper earthquakes in the IRIS database (see Fig.B.4).

Another major impact on megathrust earthquake dynamics is the asymmetrical fault-surface geometry: a shallow dipping fault intersects the Earth's free surface, and the accretionary and frontal wedge materials (hanging wall) are highly compliant compared to the footwall materials. This particular structure tends to trap seismic waves within the wedge and cause significant dynamic stress perturbations<sup>47,100,108,185,186,189,231,238,295</sup>. Such high stresses can lead to material yielding<sup>190,191</sup> or unclamping and flapping of the hanging wall<sup>47,100,295</sup>.

This study evaluates the impact of realistic structures in subduction zones, including the free surface and heterogeneous velocity structure, on the rupture dynamics and seismic radiation of megathrust earthquakes. I use two-dimensional (2D) dynamic models to investigate the radiation style of these earthquakes. A similar exercise was undertaken by Lotto et al.<sup>185,186</sup>, albeit a simplification of the 2D elastic structure and a focus on fault rheology and tsunamigenesis. Instead, this contribution uses a tomography-derived elastic model, a realistic model of the shear wavespeed ( $V_s$ ), and provides a comprehensive analysis of the seismic waves generated by these ruptures.

My results show that all simulations that contain a traction-free surface can reproduce the observations: HF seismic waves are more efficiently generated at depth, LF seismic waves are more

efficiently generated near the trench. I define the free-surface effects as the dynamic interactions between the rupture and the seismic waves reflected from the surface. I propose that the free-surface effects are the first-order explanation to the observed depth-frequency relation. Furthermore, the subduction of a cold and wet slab produces a strong material contrast across the plate interface or fault, which favors the evolution of pulse-dominated rupture front and enhances high-frequency strong ground motions from the downdip region near the coast. Because such realistic velocity models exacerbate the contrast in radiation style, I propose that realistic heterogeneous Earth velocity models give a significant second-order effect on controlling the seismic radiation. I conclude that realistic Earth structures are necessary to predict tsunami and coastal ground motion hazards better.

## 5.2 UBIQUITOUS DEPTH-FREQUENCY RELATION FROM BACK-PROJECTION ANALYSIS

### 5.2.1 RECENT LARGE EARTHQUAKE BP IMAGES

I show the back-projection (BP) results of the Mw 9.0 2011 Tohoku-oki earthquake, the Mw 7.9 2015 Gorkha earthquake, and the Mw 8.3 2015 Illapel earthquake. I obtain these BP results using a high-resolution improved Compressive Sensing back-projection (ImCS-BP) method that I developed. Detailed information about this methodology can be found in Yin et al. <sup>358</sup> and also in Chapter 2.

I download the available teleseismic P wave velocity seismograms of the 2011 Tohoku earthquake recorded by the USArray stations (TA array, Fig.B.5 (a) - (b)) in North America (TA [doi: 10.7914/SN/TA](https://doi.org/10.7914/SN/TA), data is downloaded using Wilber 3 of the Incorporated Research Institutions for Seismology Data Management Center, IRIS-DMC, [http://ds.iris.edu/wilber3/find\\_event](http://ds.iris.edu/wilber3/find_event)). The raw data is first processed by removing the mean, trend, and instrumental responses. Then I filter the waveforms (Butterworth filter, order 2) into the low-frequency (LF) band (0.05 - 0.5 Hz)

and high-frequency (HF) band (0.5 - 1 Hz) and align the waveforms based on the P wave arrival time Fig.B.5 (a) - (b). The sliding time window technique is used to get the time evolution of the earthquake rupture, and I choose a window length of 14 s for the 0.05 - 0.5 Hz LF band and 8 s for the 0.5 - 1 Hz HF band. The step of the moving time window is set 2 s. Within each time window, I apply the ImCS-BP with auto-adaptive source grid refinement<sup>358</sup> (also see Chapter 2 for details) to locate the coherent peaks and finally get the back-projection images of the Tohoku earthquakes in different frequency bands (Fig.B.5 (c) - (d)). In the main text, I integrate the BP results over the entire duration to construct a total BP image for each frequency band (Fig.5.1 (a)). My BP results of the 2011 Tohoku earthquake are well consistent with the relevant previous studies<sup>175,324,349</sup>.

For the 2015 Nepal and Chile earthquakes, I use the same ImCS-BP technique (introduced in Chapter 2) and the same USArray data as my previous studies<sup>358,361</sup>. The waveforms of the Mw 7.9 2015 Gorkha earthquake are filtered at 0.05 - 0.25 Hz and 0.25 - 1.0 Hz frequency bands while the waveforms of the Mw 8.3 2015 Illapel earthquake are filtered at 0.05 - 0.5 Hz and 0.5 - 1.0 Hz frequency bands. The difference in the frequency band is due to handling different magnitudes of earthquakes<sup>357</sup>. Here I simply show the data and BP results of both events (Figs.B.6 - B.7) and refer to the previous publications for more details on the interpretation and reliability of the images given the source and receiver array configuration<sup>358,359,361</sup>.

### 5.2.2 ANALYSIS OF THE IRIS BP DATABASE

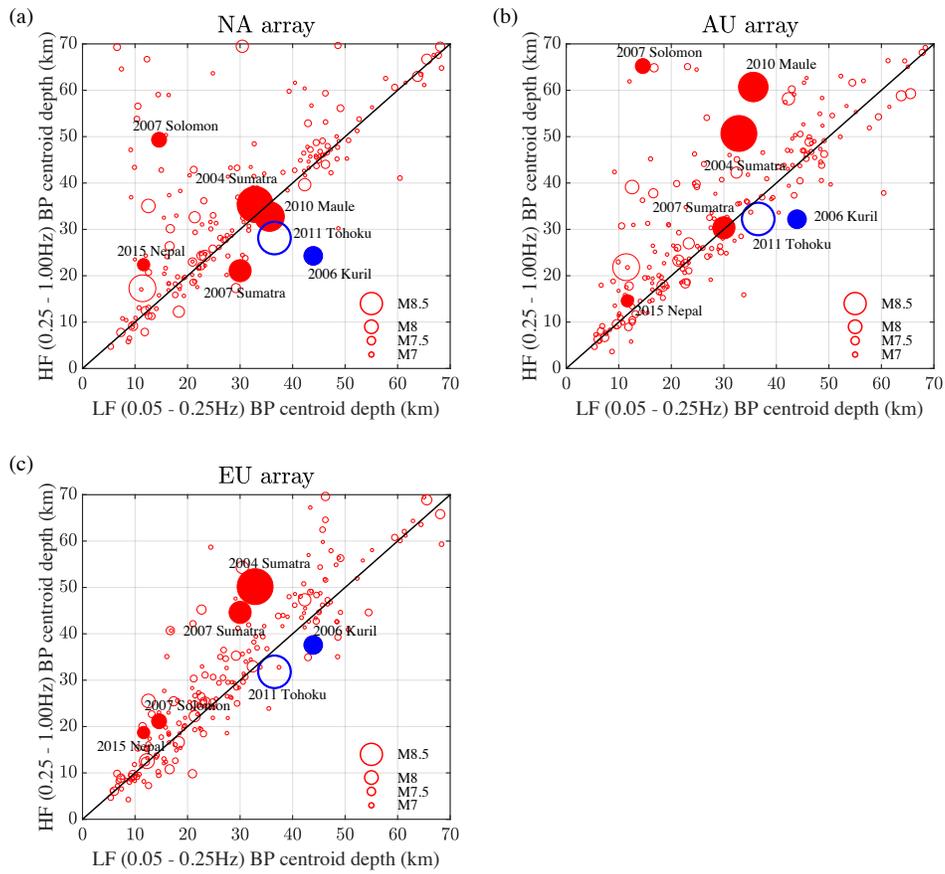
I further explore whether the depth-frequency relation exists for most megathrust earthquakes with the help of the back-projection database of the Incorporated Research Institutions for Seismology (IRIS). The IRIS back-projection database<sup>137</sup> automatically generates the BP images from three regional arrays (NA: northern America; EU: Europe; AU: Australia) and the Global Seismic Network (GSN) for all the M6.5+ earthquakes since 1995<sup>137</sup>. The three regional arrays can produce the HF (0.25 - 1.00 Hz) BP images, and the GSN can produce the LF (0.05 - 0.25 Hz) BP images. This

provides an opportunity to compare the depth-frequency relation systematically (i.e., with a single method) instead of making an inventory of results based on different methods applied to different earthquakes.

I collect the HF and LF BP peaks of all the 842 earthquakes present in the IRIS database (available at <http://ds.iris.edu/spud/back-projection>, last accessed on 02/27/2021). Among the events from the IRIS database, I only select those with BP results from all four arrays/networks. Because the BP results are recovered from the teleseismic P waves, which have poor depth resolution, I project the latitude and longitude of the BP peaks onto the corresponding Slab2 slab model<sup>114</sup> to infer the depth of the BP results. Only 461 earthquakes (mostly megathrust earthquakes) within the latitude-longitude range of the available Slab2 models are kept.

Next, I calculate the average depth of all the BP peaks weighted by the BP peak amplitude for each array. I define the average depth as the BP centroid depth of the earthquake for each specific array. In this way, I can obtain the BP centroid depth from the GSN BP results in the low-frequency band of 0.05 - 0.25 Hz and the 3 estimates of the HF BP centroid depths from the dense regional arrays NA, AU, and EU in the high-frequency band of 0.25 - 1 Hz. Because I focus on the megathrust earthquakes in this study, I only keep the 245 events with BP centroid depth less than 70 km and the comparison results of all three regional arrays are shown in Fig. 5.2. Finally, I take the mean of the HF BP centroid depths across all three arrays as the representative HF BP centroid depth and show the comparison with LF BP centroid from GSN in Fig. 5.1 (d) of the main text. I also show the same results for the deep earthquakes with depth from 70 km to 700 km in Fig. B.4 to show that the frequency-depth relation disappears for deep earthquakes.

My systematic analysis on the BP results from the IRIS Back-projection database confirms the findings from previous studies on specific earthquakes that the depth-frequency relation is ubiquitous for most megathrust earthquakes. Furthermore, my results highlight that the depth-frequency relation only shows up for the shallow earthquakes above 70 km. For those deeper earthquake be-



**Figure 5.2:** Comparison between the LF BP centroid depth from GSN and HF BP centroid depth from (a) North America NA array; (b) Australian AU array and (c) European EU array for the megathrust earthquakes in the IRIS back-projection database.

low 70 km, the depth-frequency relation disappears. This pattern implies that the depth-frequency relation should be derived from the shallow features in the subduction zones.

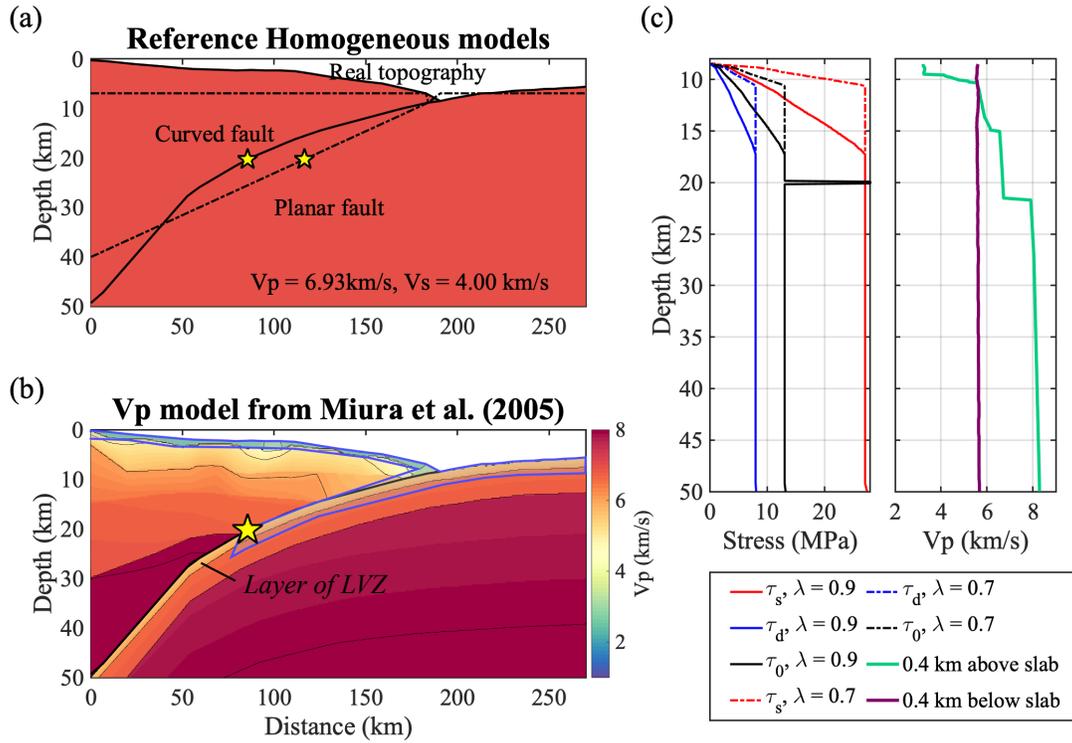
### 5.3 DYNAMIC RUPTURE SIMULATION OF MEGATHRUST EARTHQUAKES

Numerical simulation of dynamic rupture for earthquakes is a powerful tool to study the earthquake physics. To investigate the reasons for the ubiquitous depth-frequency relation observed during megathrust earthquakes, I perform a set of dynamic rupture experiments in 2D media of small and large earthquakes. Significant ingredients of dynamic rupture simulation includes initial stress, fault friction, geological structures and a numerical solver<sup>111</sup>. I set up the ingredients of dynamic rupture simulation based on the relevant observational constraints for the subduction zone earthquakes.

Firstly, I build five models that are dedicated to small earthquakes in a simple homogeneous half space on a flat fault. Spontaneous dynamic ruptures are nucleated at different depths so I can systematically investigate the depth-dependence of this depth-frequency relation.

Next, I target at the large megathrust earthquakes. I gradually increase structural complexity from homogeneous to realistic elastic structures. By building up complexity, I explore to what degree the realism in Earth models impact the rupture. Combining different parameter settings, I obtained 29 representative rupture models. For both the small and large megathrust rupture models, I analyze the spectral properties of the rupture slip history. I further investigate the nearby ground motions and the tsunami potential for those megathrust models.

The technical details about the numerical simulations are shown in the Appendix B. In this chapter, I focus on the simulation results and how those results can be used to explain the observation of depth-frequency relation.



**Figure 5.3:** Model setting. (a) Model configuration in the homogeneous structure: a flat half-space with planar slab/fault geometry and a flat topography (dashed lines), a half-space with realistic slab geometry and seafloor topography (solid lines, referred to later as REF), hypocentral locations (yellow stars). (b) Heterogeneous half-space with realistic seafloor topography and  $V_p$  structure from Miura et al.<sup>213</sup>. The shaded blue areas highlight where  $V_p/V_s$  ratio is varied. (c) Fault properties: static strength levels  $\tau_s$  (red), dynamic strength levels  $\tau_d$  (blue), initial shear stress  $\tau_0$  (black) with different values of pore-pressure ratio  $\lambda$  of 0.7 (dashed lines) and 0.9 (solid lines),  $V_p$  along with two profiles projected at 400-m above (green) and 400-m below the plate interface (purple).

### 5.3.1 REPRESENTING A REALISTIC MEGATHRUST STRUCTURE

I choose the Tohoku region in northeastern Japan as my study case. I start from a benchmark case in the homogeneous full-space medium without free surface. Then I increase the complexity of the medium from a homogeneous half-space with a planar shallow dipping fault ( $11.8^\circ$  degrees, Fig. 5.3 (a)) to a heterogeneous half-space with realistic geometry and a regional  $V_p$  structure from Miura et al.<sup>213</sup> (Fig. 5.1 (a) and Fig. 5.3 (b)). The elastic structure varies considerably along the dip of the

megathrust, especially  $V_p$  in the upper plate<sup>270</sup>. Another aspect of the structural complexity is the high compliance of the sediments that constitute the accretionary wedge (see Von Huene et al.<sup>319</sup>, and references therein). Here, I describe the megathrust fault zone into two canonical fault zone structures: 1) the updip fault zone has low-velocity properties and high  $V_p/V_S$  ratio, a nearby free surface, and a wide damaged zone, and 2) the downdip fault zone has a sharp contrast in material properties across the fault.

I focus my efforts to model a realistic updip region (above 20 km) on generating a realistic  $V_S$  structure. The compilation of  $V_p/V_S$  ratio values provided by Brocher<sup>42</sup> suggests that low  $V_p$  materials have high  $V_p/V_S$  ratios. In light of this, I discuss three regions of possibly elevated  $V_p/V_S$  ratios. The first region is the subduction channel, the thin upper layer of the downgoing slab that is composed of fluid-rich seafloor sediments<sup>118,226,268,368</sup> and hydrated minerals in a mafic fractured crust<sup>40,118,232,249,284</sup>. The second region is the slope apron, the thin layer of the seafloor sediments that covers the wedge, which is best accessed by offshore drilling and active seismic surveys<sup>98,244,307,368</sup>. The third region I consider is the frontal prism that is the tip of the accretionary wedge where dragging of high  $V_p/V_S$  ratio sediments may occur<sup>98,118,228,268</sup>. Due to the range of  $V_p/V_S$  values found in the literature, I vary the ratios between  $\sqrt{3} \sim 1.73$ , 1.83, 1.94, 2.04, 2.14, 2.24, 2.35, and 2.45 in the three specific regions discussed above (Fig .B.2). Although higher values have been reported within layers of seafloor sediments<sup>368</sup>, these are likely too thin to be resolved by my numerical exercise.

I now focus my attention on modeling material contrasts at the plate interface in the downdip region (between 20 and 50 km depth). Although the downgoing oceanic plate is denser than the overriding plate, the several-kilometer thin upper portion of the oceanic crust exhibits low seismic velocities. It is present in most subduction zones and is referred to as the Low-Velocity Zone (LVZ). To confirm this common feature of subduction zones, I compile the range of  $V_p$  in the LVZ and across the fault in the upper plate in Appendix B Table.B.1.

Finally, I embed the realistic structure in a homogeneous half-space and generate a larger simulation domain to avoid artifacts from the absorbing boundary conditions. I impose a 5-km smoothing operator to taper off velocity changes between the realistic structural model and the homogeneous half-space (Appendix B Fig.B.1).

### 5.3.2 MODELING THE DYNAMIC RUPTURE

The other ingredients necessary to model earthquake ruptures are fault properties such as the stress field, the pore pressure, and the frictional conditions (Fig. 5.3 (c)). I explore several frictional conditions. In most models, I apply linear slip weakening on the entire fault. I test for slip-neutral and slip-strengthening conditions in the upper  $\sim 10$  km of the along-dip direction, in a zone of low-grade metamorphism where neutrally stable conditions may occur<sup>129,168,185,186,234</sup>. I also test the frictional constitutive relation proposed by Murphy et al.<sup>225</sup> that is derived from laboratory experiments. In addition to increasing the  $V_p/V_s$  ratio, the fluid content also affects the stress fields by reducing overburden lithostatic pressure  $\sigma_L$  with pore fluid pressure  $p$ . I use the pore pressure ratio  $\lambda$  defined in Hubbert & Rubey<sup>130</sup> to impose a pore pressure  $p = \lambda\sigma_L$  as well as the effective normal stress  $\bar{\sigma}_n = (1 - \lambda)\sigma_L$ . Given the uncertainties in  $\lambda$ , I test two values of  $\lambda$  (0.7 and 0.9) and assume that the pore fluid pressure becomes lithostatic when  $\bar{\sigma}_n = 40$  MPa (Fig. 5.3 (d)). These conditions are similar to those discussed and imposed in previous studies<sup>186,225,256,268</sup>. The earthquake rupture naturally evolves on the fault in response to an over-stressed nucleation patch (see Fig. 5.3 (c)). A full description of all model parameters is in Appendix B (Text S1). I use the SEM2DPACK software<sup>14</sup> (available at <https://github.com/jpampuelo/sem2dpack>, last accessed on 06/08/2021) to simulate both the dynamic slip on the fault and the wavefield in the two-dimensional elastic domain.

### 5.3.3 PARAMETERIZATION OF THE SOURCE RADIATION

To understand the relative contributions between LF and HF seismic waves emitted by the rupture, I parameterize the local slip-rate function's spectrum and improve from the qualitative discussions in Figure 3 (c) of Ma & Hirakawa <sup>190</sup> and Figure 12 (d) of Galvez et al. <sup>103</sup>. In this study, I systematically measure and compare the along-dip spectral variations with two metrics.

The first approach fits the Fourier amplitude spectrum of the local slip-rate function with a flat model at low frequencies and a power-law decay at high frequencies. I apply a model commonly used in source seismology,  $S(f) = 1 / (1 + (f/f_c)^n)$ , where  $f_c$  and  $n$  are the corner frequency and spectral falloff rate, respectively. The spectral model fits the shape of far-field P-wave pulses that originate from circular crack ruptures with uniform stress drop and elliptical slip distribution <sup>45,92,192</sup>. It is common to perform spectral fitting over the spectrum of the far-field body-wave pulse of the entire event, which is the moment-rate pulse <sup>3,10,306,329</sup>. Slip rate functions and overall moment-rate functions differ because the latter is the spatial integration of the former. This leads to differences in spectral shapes. For instance the slip-rate spectral shape may be sensitive to the breakdown time <sup>128,300</sup>. I use this spectral shape solely to characterize the spectral shapes and relative HF-LF content. The corner frequency  $f_c$  is inversely proportional to the pulse duration, which is also referred to as "rise time" in the kinematic representation of the earthquake source. The spectral falloff rate  $n$  describes how fast the high-frequency component decays in amplitude. The two spectral parameters trade off each other during the spectral fitting <sup>76,306</sup>. Combining both can help to quantify the relative portions of LF and HF seismic radiation: larger  $f_c$  and smaller  $n$  correspond to relatively more HF radiation, while smaller  $f_c$  and larger  $n$  correspond to relatively more LF radiation. I apply a non-linear least-square solver to find  $f_c$  and  $n$  from fitting the  $\log_{10}$  of the amplitude spectra of the local slip-rate functions interpolated on a logspace frequency vector, a strategy similar to other observational studies (see Shearer et al. <sup>282</sup> for a recent review).

The second measure of relative contribution in frequency content estimates the seismic power generated by the local slip-acceleration function. Similar methods have been applied in previous studies to quantify the spectral power of slip rate from different frequency components<sup>129,128,211</sup>. Here I choose slip acceleration as the ground motion unit because far-field velocity seismograms are commonly used for teleseismic P-wave back-projection studies<sup>99,357</sup> and are proportional to moment accelerations. I estimate the power by bandpassing (Butterworth, four corners, zero phase) and integrating the squared time series of local slip-acceleration functions in two frequency bands below the resolvable frequency: for small earthquake rupture in the homogeneous medium, LF 0.001-0.1 Hz and HF 0.1 - 1 Hz; for megathrust rupture, LF 0.001-0.06 Hz and HF 0.06 - 0.3 Hz. The central frequencies 0.1 Hz and 0.06 Hz are arbitrarily chosen as approximately the middle of the log-scale frequency band, but other tested values did not affect the general trends in the results. Details about the frequency resolution are in the Appendix B. I then use the HF and LF seismic powers, specifically the HF/LF power ratio, to measure their relative contributions.

## 5.4 RESULTS FROM DYNAMIC RUPTURE SIMULATIONS

### 5.4.1 CASES OF SMALL SUBDUCTION ZONE EARTHQUAKES

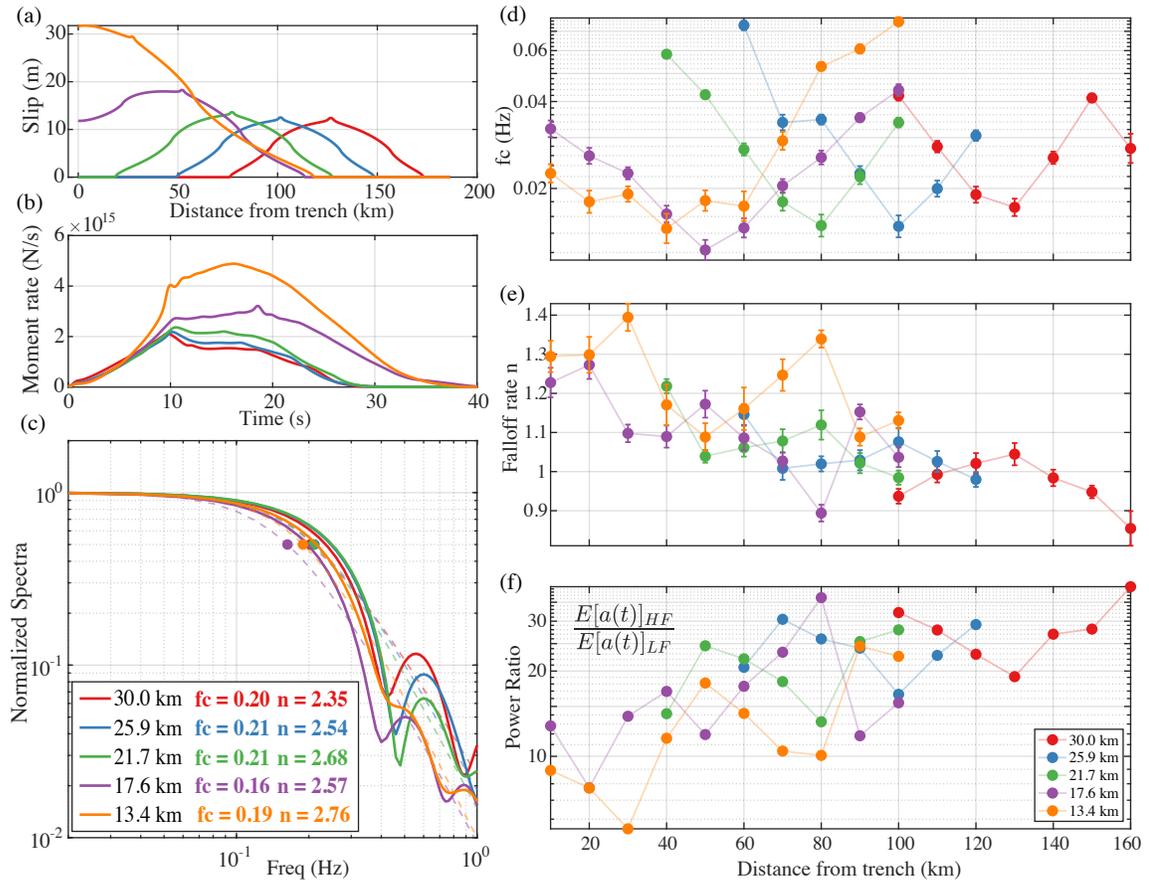
I start by inquiring whether the model setup can reproduce the differences in pulse width and fall-off rate that are reported from observations of small subduction-zone events<sup>121,355</sup>. I systematically model five small ruptures initiated at the depths of 13.4 km, 17.6 km, 21.7 km, 25.9 km, and 30.0 km in a homogeneous structure with a planar fault and flat, free surface (Fig.5.3 (a) and Fig.B.8). I impose pre-stress conditions to constrain the rupture length and keep other parameters equal in all simulations (see Fig.B.8). Finally I apply my parameterizations to quantify the contributions from LF and HF radiation for these rupture models. Any difference in rupture style may then be attributed to free-surface effects controlled by the depth (or distance from the free surface) at which

the rupture occurs.

The simulation results show that only the two shallower ruptures have reached the surface while the three deeper ones remain buried (Fig. 5.4 (a), Fig. B.9). As the shallow ruptures reach the trench, they interact with the scattered wavefield. Such wave-rupture interaction disappears in the case of a source deeper than 20 km as the rupture almost terminates before the arrival of free-surface reflections (Fig. B.9). Effectively, the deep sources are in a full-space. The shallow ruptures end up releasing about twice the moment (per unit of fault width) of the deep ruptures (Fig. 5.4 (a)) due to the “mirror effect” from free surface<sup>187</sup>.

Next, I fit the overall moment-density-rate function with the spectral model mentioned in Section 2.3 up to a resolvable frequency 1 Hz (Fig. 5.4 (c)). This is in practice very similar to the seismological studies that explore earthquake source parameters<sup>3,25,76,306</sup>. However, here I only use this model to quantify the spectral shape and avoid any dynamic implications on source parameters due to the circular-crack assumption of this spectral model<sup>45,92,192</sup>. The spectral analysis shows that the source spectra of the two shallow earthquakes have lower  $f_c = 0.19$  Hz and  $f_c = 0.16$  Hz than the deeper ones with  $f_c$  about 0.2 Hz because of the longer duration of shallow ruptures. I also find a systematic trend of the spectral falloff rate  $n$  that the value of  $n$  systematically decreases along depth (Fig. 5.4 (c)), implying that the moment-rate spectrum is more depleted in HF waves than deep earthquakes.

Moreover, I investigate how the local slip-rate functions vary with depth for each model. Details of the space-time rupture evolution can be found in Appendix B Fig. B.9. Here, I select an individual slip-rate function every 10 km along with the plate interface and measure corner frequency  $f_c$ , falloff rate  $n$ , and the corresponding HF/LF power ratio (Fig. 5.4 (d) - (f)). There is no evident systematic along-depth variation of  $f_c$ ; instead, it varies with the distance from the nucleation site as expected from crack models (rise time is longest at the nucleation patch). However, I find systematic along-depth variations of  $n$  and HF/LF power ratio:  $n$  decreases while HF/LF power ratio increases with



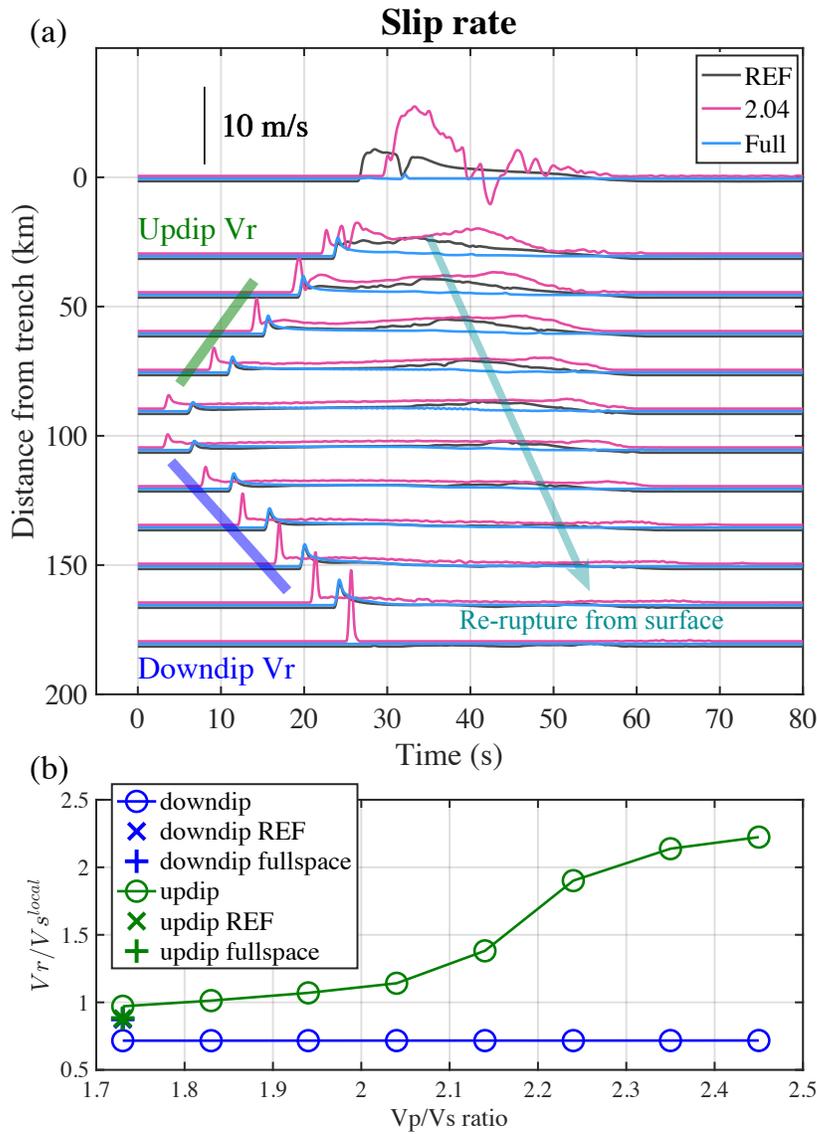
**Figure 5.4:** Simulation results of small megathrust earthquakes at different depths. (a) Final slip distributions on fault; (b) Moment-rate functions (per unit length along strike) averaged over the entire fault; (c) Amplitude-normalized source spectra (solid lines) as well as the corresponding best-fitted spectral models (dashed lines). The dots indicate the values of corner frequency  $f_c$ . (d)-(e) Best-fitted parameters of the slip rate functions at different depths for all the models: corner frequency  $f_c$ , spectral falloff rate  $n$  and HF/LF power ratio of slip acceleration, respectively.

depth for all models in general. Both  $n$  and the HF/LF ratio suggest that more HF components are radiated during the deeper ruptures. Since the only difference between the models is the source depths, i.e., the distances from the free surface, I suggest that free-surface effects are the origin of the depth-frequency relation.

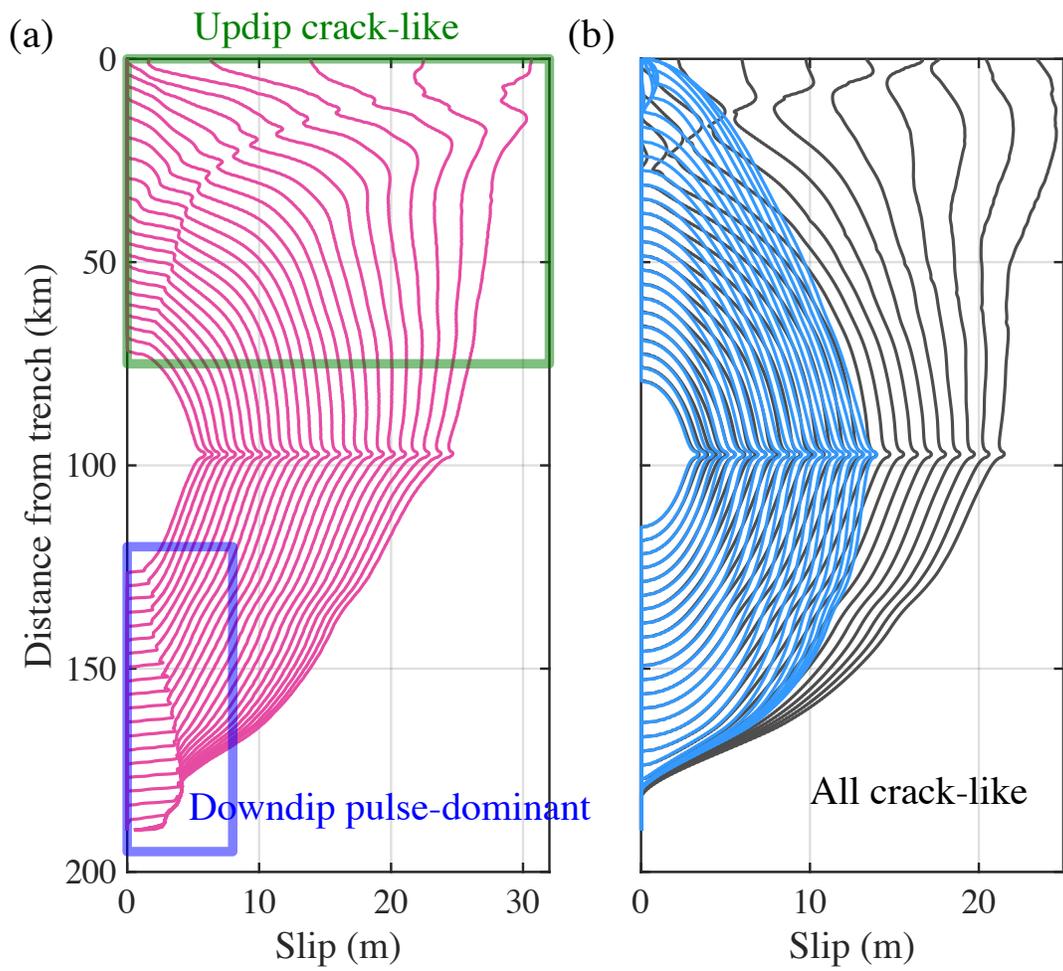
#### 5.4.2 CASES OF MEGATHRUST EARTHQUAKES

In this section, I present my simulation results of the large megathrust earthquake models. Examples of the space-time evolution from the ruptures in the homogeneous full-space model (Full), homogeneous half-space model (REF) and heterogeneous model ( $V_P/V_S = 2.04$  in the  $V_P/V_S$ -elevated regions) are shown in Fig. 5.5 (a). My half-space simulations are typical of 2D models of dynamic rupture<sup>129,168,185,253</sup>. All simulated ruptures reach the trench, last about 60 seconds, and their final slip increases from small downdip to large updip. The rupture first propagates bilaterally from its nucleation patch. The updip rupture then hits the trench with a high slip rate, and a weak re-rupture front propagates back downdip. The downdip rupture propagates with a constant rupture velocity and dies at the end of the fault. The slip profiles along the dip (Fig. 5.6) are similar to many of those inferred for the Mw 9.0 2011 Tohoku-oki earthquake<sup>328</sup>. By comparison, the simulation in the homogeneous full-space model presents symmetric rupture behavior at the updip and downdip propagating fronts. The slight asymmetry of full-space model is due to the initial stress distribution (Fig. 5.5 (a) and Fig. 5.6 (b)). I refer to Appendix C for each model's detailed results and summarize their general patterns.

To explore the depth-varying properties, I apply the same parameterization in previous sections to all megathrust rupture models (Fig. 5.7). First, I perform the spectral fitting for each slip-rate function. I find that all models with a free surface present similar along-dip (or depth) variations of the spectral properties (Fig. 5.7 (a) and Fig. B.10). The spectral falloff rate  $n$  generally decreases with depth: it is about 1.8 - 2.0 (model median) on the shallow segment from 0 - 20 km and 0.8 - 1.0 (model median) on the deep segment. Second, I calculate the HF/LF power ratio of slip accelerations in the HF (0.06 - 0.3 Hz) and LF (0.001 - 0.06 Hz) bands. Here again, I find a clear pattern that the HF/LF power ratio increases with depth (down-dip) for all those half-space models (Fig. 5.7 (b) and Fig. B.10). I also repeat the measurements for the segment-averaged slip-rate functions (on



**Figure 5.5:** Space-time evolution of the simulated megathrust earthquake. Comparisons of simulation result from the model in heterogeneous medium with  $V_p/V_S=2.04$  (Fig.5.3 (b)) and homogeneous models: REF model with real topography in Fig.5.3 (a) (dark gray) and full-space model (light blue). (a) Space-time slip-rate evolution: green and blue lines crudely mark the updip and downdip rupture front; the gray-blue arrow indicates the weak re-rupture propagating downdip from the surface. The pore pressure ratio  $\lambda=0.9$  in this comparison. (b) Rupture speeds of updip (in green, 40 km to 80 km from the trench) and downdip (in blue, 110 km to 160 km from the trench) propagation for each model.  $V_S^{local}$  is chosen based on the P wave velocity 0.4 km above slab (green line in Fig.5.3c) and the  $V_p/V_S$  ratio in each model.



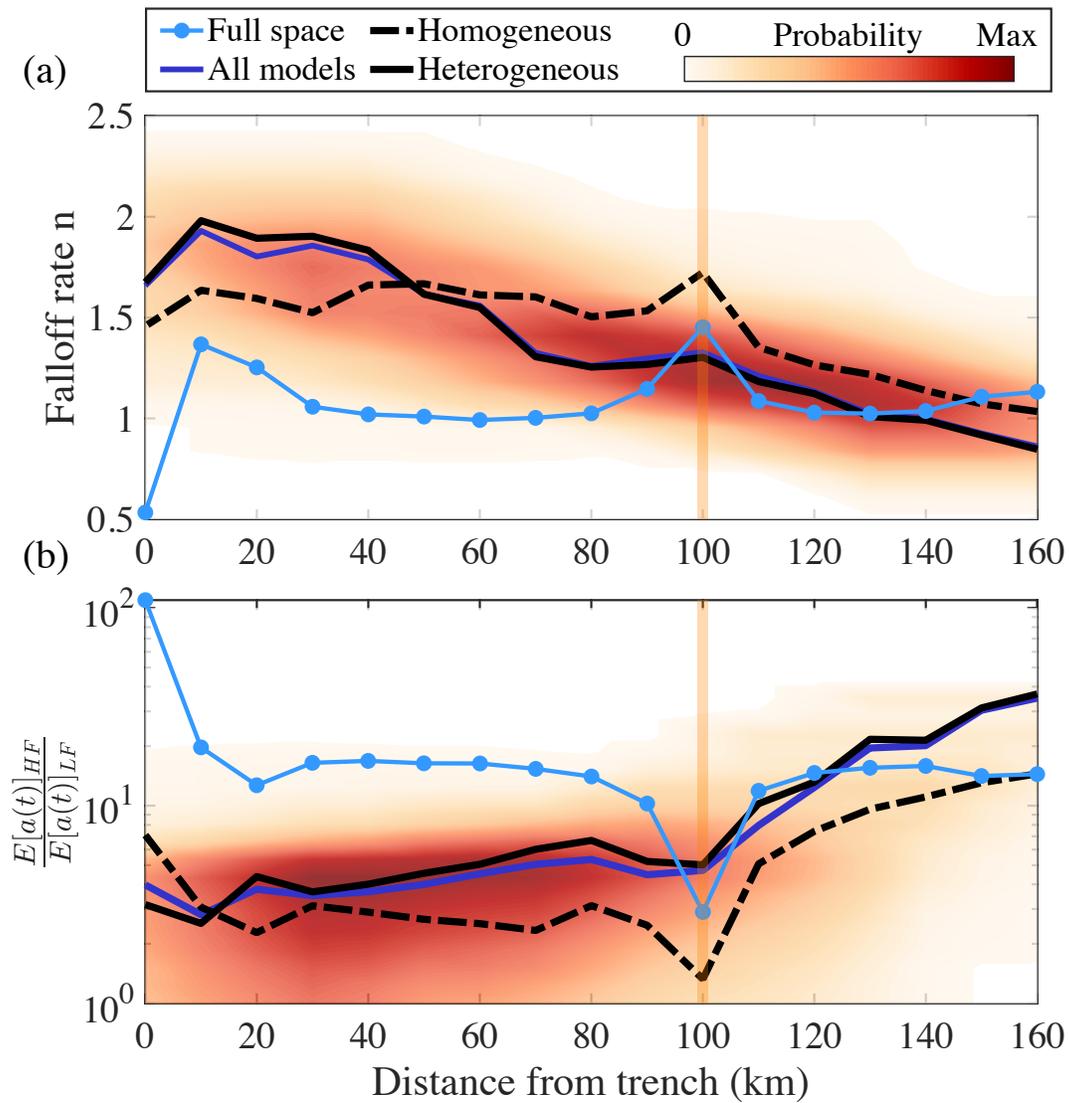
**Figure 5.6:** Slip history of the simulated earthquakes. Snapshots of slip distribution from (a) models with heterogeneous velocity structures and  $V_p/V_s=2.04$ ; (b) homogeneous models in half-space with real topography (dark gray) and in full-space (blue). Slip contours from every second between 10 to 35 s are plotted.

the 10-km subfaults), and the patterns stay the same (Fig .B.11).

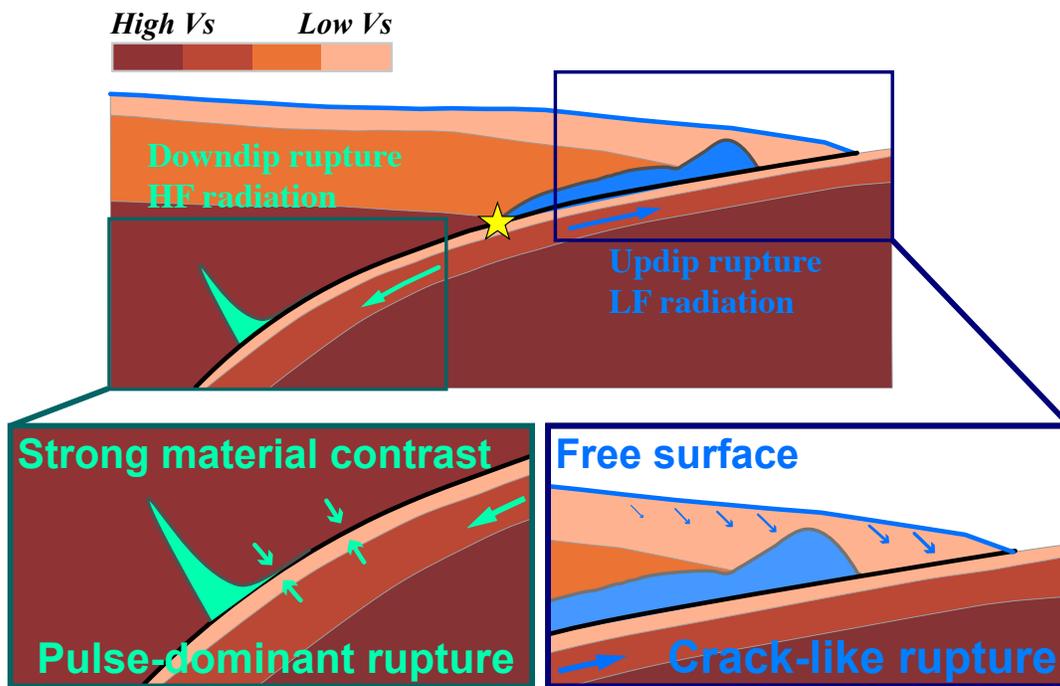
In all free-surface models, both measures of the local slip-rate functions' relative frequency content vary systematically with depth. Such systematic variation contrasts with the results obtained with the full-space model's case: both the spectral falloff rate  $n$  and the HF/LF ratio remain constant (Fig.5.7) because of the symmetry of slip history (Figs.5.5 - 5.6). This is consistent with the results from the small subduction-zone megathrust earthquakes in Section 3.1, and again suggests that free-surface effects are the first-order mechanism that explains the frequency-depth radiation during megathrust earthquakes. Furthermore, I notice that the rupture models in the realistic heterogeneous mediums present stronger contrast in radiation style, that is stronger variations of falloff rate  $n$  and HF/LF power ratio with depth than the models in the homogeneous structure. It means that the realistic velocity structure can be a second-order mechanism and further enhance the observations of depth-frequency relation.

## 5.5 DISCUSSION

This study focuses on the effects of free surface and realistic Earth structure on the dynamic rupture behavior of megathrust earthquakes. While I test one particular subduction zone in northeastern Japan<sup>213</sup>, the overall structure exists in many other subduction zones (Appendix B Table.B.1). Three specific structural features appear to impact the depth-frequency relation of megathrust earthquakes (Fig.5.8): 1) the free surface in the near-source region, 2) the high compliance of the sediments in the updip wedge, and 3) the low-velocity zone below the plate interface downdip. My systematic simulations show that free-surface effects are the first-order mechanism, and the heterogeneity in material compliance further enhances the radiation contrast. I illustrate this in Fig.5.8. I now discuss the varied rupture behavior, their impact on the depth-frequency relation, and further implication for ground motion and tsunami hazards.



**Figure 5.7:** Spectral properties of the source radiation, shown by the probability distribution from all models. These distributions are obtained from the Box-kernel smoother. (a) Best-fitted spectral falloff rate  $n$  along dip from the simulated megathrust earthquake with different model settings. (b) The power ratio of high frequency (HF) 0.06 - 0.3 Hz and low frequency (LF) 0.001 - 0.06 Hz slip acceleration along dip. Dark blue solid line shows the median value of all models. Black dashed line and solid line show the median values from all homogeneous models and all heterogeneous models, respectively. The light blue dotted line shows the result from the homogeneous full-space model. The orange bar indicate the location of rupture nucleation.



**Figure 5.8:** Effects of the free surface and material contrasts on the dynamic behavior of megathrust earthquakes. In the updip region, the free surface leads to crack-like rupture (slip-rate distribution is shown in blue) and enhanced low-frequency radiation. In the downdip part, the substantial material contrast at the top of the LVZ favors pulse-dominant rupture (slip-rate distribution is shown in green) and enhanced high-frequency seismic radiation.

### 5.5.1 UPDIP RUPTURE: LARGE AND FAST CRACK RUPTURE TO THE TRENCH

The rupture accelerates updip and evolves as a crack (Fig. 5.6): the shallow rupture velocities are higher than typically observed<sup>59</sup> and greater than the surrounding  $V_S$ , and slip continues until the end of rupture. My simulations shed light on two major factors that control this updip behavior: the free surface and the shallow compliant fault zone.

Previous studies have shown that the free surface can significantly change the normal stress during rupture, due to waves reflecting at the free surface and traveling back to the fault<sup>47,231,237,238,295,330</sup>. My simulation results are no different: clear surface-reflected phases cause the prolonged and persistent slip in the updip portion (Appendix C). Free-surface effects also induce acceleration of rupture propagation with supershear velocity: a secondary “daughter crack” can be triggered by the surface-reflected shear wave, which can be seen in other studies<sup>129,185</sup> and in other tectonic regimes such as strike-slip earthquakes<sup>147</sup>. The “mirror effect” of the free surface to seismic waves can also cause larger coseismic slip even with a constant stress drop value<sup>187</sup>.

The highly compliant structure of the shallow hanging wall of the megathrust acts as a seismic waveguide. The upper plate low-velocity sediments can trap seismic waves, amplify their amplitudes and extend their duration. This wave propagation effect is similar to how seismic waves amplify when traveling in sedimentary basins<sup>50</sup>. Despite differences in model settings, all simulations show that the initial wave emitted at the rupture front, the free-surface reflections, and other wedge captured and scattered waves interfere together to energize rupture propagation and further increase the final slip. In my simulations, these normal stress changes and fault-parallel slip are so extreme, with peak slip rates on the order of 10 m/s, that some models with standard  $V_P/V_S$  ratios predict co-seismic backslip (Fig. 5.5 (a)). In simulations with higher  $V_P/V_S$  ratios, much lower  $V_S$  may delay the propagation of scattered waves in a way that limits their constructive interference back to the fault. Regardless, such extreme values of slip rates generate large dynamic stresses that can cause (not

modeled) inelastic failure<sup>190,191</sup>, wedge flapping<sup>47,100</sup>. This phenomenon may be the cause for the suggested dynamic overshoot during the Mw 9.0 2011 Tohoku-oki earthquake<sup>133</sup>.

Moreover, the downgoing plate is fractured and hydrated on the foot-wall side with low velocities and elevated  $V_P/V_S$  ratios (10 - 20 km depth in Fig. 5.3 (b)). Altogether, the structure is similar to that observed in crustal damage zones<sup>29</sup>. Harris & Day<sup>112</sup> suggested that the low-velocity structure around the fault can affect the rupture speed and slip-velocity pulse shape. Furthermore, such a low-velocity structure dramatically impacts rupture propagation and termination, such as multiple slip pulses, supershear rupture velocity, and rotation of background stress<sup>15,126,124,263</sup>.

In the homogeneous case with a uniform  $V_P/V_S$  ratio of  $\sqrt{3}$  and realistic fault and seafloor geometries (REF model in Fig. 5.7 and Fig. 5.9), the rupture velocity for both updip and downdip rupture has a typical value of  $0.87V_S$ . In the models that have realistic  $V_P/V_S$  ratios, the updip rupture velocity becomes greater than the local  $V_S$ . This is typical for 2D elastic models of earthquakes on the megathrust of subduction zones<sup>186</sup> and in damaged fault zones<sup>126,127,340</sup>.

### 5.5.2 DOWNDIP RUPTURE: PULSE-DOMINANT RUPTURE ALONG WITH THE LVZ

As the rupture propagates to the downdip region, there is no impact from free-surface reflections as the rupture ends before waves travel back to the fault. All models present a sharp rupture front (Fig. 5.5 (a)). In the models with a homogeneous structure, the slip-rate functions have typical long tails<sup>166</sup>. In the models with heterogeneous structures, the slip-rate functions are characterized by a shortening of the slip pulse (stronger healing) with depth (or along dip with hypocentral distance). In both situations, my quantification on the spectrum shows that the HF energy dominates due to the impulsive slip-rate function shape.

The material contrast across the fault can explain the evolution of short and sharp slip pulses downdip of the hypocenter. Theoretical studies have predicted the slip pulse produced by the material contrast at the fault interface<sup>18,335</sup>. Moreover, Shlomag & Fineberg<sup>286</sup> perform and analyze

lab experiments with an in-plane shear of the two blocks with different compliance. They show that such a bimaterial interface can host both rupture modes: one self-healing slip pulse that moves in one direction of rupture and one slip crack that propagates in the opposite direction. The experimental configuration is similar to that of the subduction zones downdip of the seismogenic zone with the contact between the LVZ and the overhanging upper mantle material. As the rupture propagates downdip, in the moving direction of the more compliant oceanic plate, the slip-rate functions are short and sharp pulses (Fig. 5.5 (a)). The corresponding downdip rupture speed  $V_r$  is about  $0.71 V_S^{local}$  (Fig. 5.5 (b)), which is the local shear wavespeed of the continental crust near the slab (Fig. 5.3 (c)), but is about 5% higher than the shear wavespeed in the LVZ. This has also been shown by the experiments of Shlomaï & Fineberg<sup>286</sup>.

Previous theoretical and numerical studies show that the generation of a self-healing slip-pulse on bimaterial interface required specific conditions of initial stress, friction or geometry<sup>15,69,241,263,285</sup>. This study has not covered the parameterization of those conditions for my dynamic models of megathrust rupture. But I leave them as a future direction to explore in combination with theoretical studies and constraints on how fault is localized in subduction zones from geological observations.

### 5.5.3 DEPTH-FREQUENCY RELATION OF MEGATHRUST EARTHQUAKES

In this study, I have shown that all earthquakes simulated in half-spaces exhibit similar along-dip (along-depth) variations in the values of the spectral parameters and HF/LF ratios of the local slip-rate functions (Fig. 5.4, Fig. 5.7), which is consistent with the observed depth-frequency relation (Fig. 5.1). In contrast, the benchmark full-space simulation is not consistent with the observations. Therefore, I propose that free-surface effects are the first-order factor in explaining the observed depth-frequency relation of megathrust earthquakes.

The cases of the simulated small earthquake ruptures reveal that the shallower earthquakes are

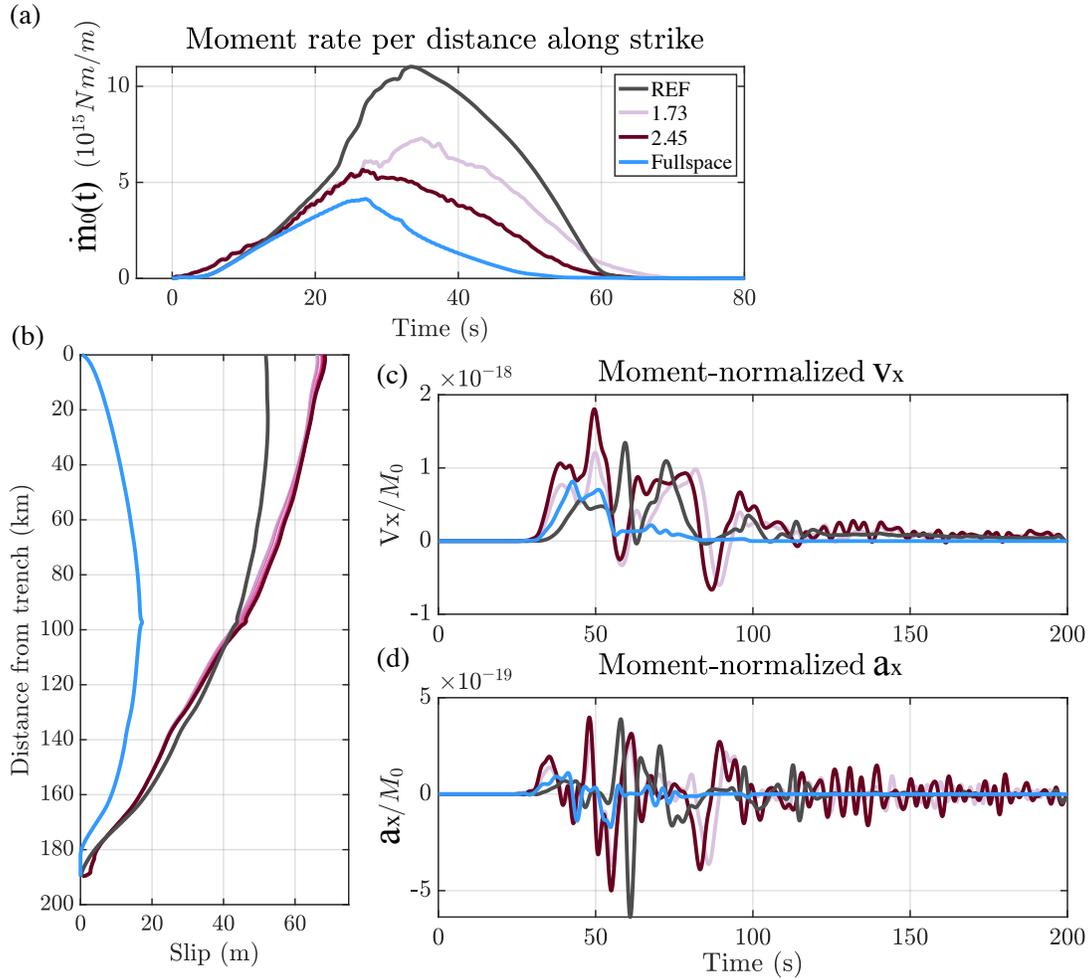
more depleted in high-frequency radiation than the deeper ones (Fig. 5.4). These patterns are consistent with the observed systematic depth variations of source parameters for small-to-moderate earthquakes<sup>76,121,161,355</sup>. The depletion in HF content is mainly caused by the interference between direct rupture and the free-surface reflection (Fig. B.9).

The cases of the simulated large earthquake ruptures further support the claim that free-surface effects are the leading factor to explain the depth-frequency relation during large megathrust earthquakes. The deep portion of the rupture has elevated HF radiation compared to the shallow portion, regardless of model setting (Fig. 5.7). My study suggests that a crack-like rupture mode exemplifies the updip rupture of megathrust earthquakes. In contrast, the sharp slip-pulses are the dominant mode of the downdip ruptures, at least as seen by seismic radiation (Fig. 5.5 (a)).

This study focuses on the Tohoku region, however my results are generalizable since the free-surface dominate the response. I also tested a shallow vertical mode-III (anti-plane) rupture in a homogeneous halfspace and found a similar patterns in the spectral content (Fig. B.12). These findings imply that the depth-frequency relation may also exist for other types of earthquakes such as strike-slip events at shallow depth. However, there is no observation of such phenomenon, which I attribute to the poor resolution with depth using teleseismic waves. Improvements in the Green's function for near surface source may help find the seismic signatures.

Moreover, models that include realistic velocity structures exhibit a stronger variations in  $n$  and HF/LF ratios with depth (Fig. 5.7). I attribute this stronger contrast to the wave effects in a realistic velocity structure. First, the shallow, compliant, high- $V_P/V_S$  accretionary wedge trap waves more effectively, slows their propagation, and increase the duration of slip on the fault, which enhances LF radiation near the trench. Second, the deep strong material contrast between the LVZ and the continental, overriding mantle can lead to more pulse-dominant slip histories with more HF radiation in the downdip region. Therefore, the realistic elastic structure in the subduction zone is another controlling factor to the depth-varying frequency-dependence of seismic radiation.

#### 5.5.4 IMPLICATIONS FOR TSUNAMI AND GROUND MOTION HAZARDS



**Figure 5.9:** Tsunami and ground motion hazards. (a) Moment-rate density function of each model with different  $V_P/V_S$  ratios. (b) Final along-dip slip distribution from different models. (c) Moment-normalized velocity seismograms (horizontal x direction) recorded by the far-field station (location shown in Fig.B.1). (d) Corresponding moment-normalized acceleration seismograms (horizontal x-direction) recorded by the same virtual station.

My simulations indicate that the final slip distribution varies considerably with the model settings. The final moment magnitude of the homogeneous half-space models is larger than the heterogeneous models, probably due to the greater shear modulus at the shallow portion (Fig.5.9 (a)).

However, the final slip is greater underneath shallow and highly compliant structures (Fig. 5.9 (b)), which was also found by Lotto et al.<sup>186</sup>. The final slip at the trench directly impacts the tsunami height. I apply a simplified relation from Tanioka & Satake<sup>298</sup> to estimate the initial tsunami height at the trench:  $\eta_{ts} = u_y - mu_x$ , where  $u_x$ ,  $u_y$ , and  $m = -0.1$  are the horizontal displacement, vertical displacement, and the horizontal gradient of the bathymetry at the trench, respectively. I find that  $\eta_{ts} = 8.6$  m for the homogeneous half-space model (REF model), 11.0 m for the heterogeneous model with  $V_p/V_S = \sqrt{3}$  and 11.3 m for the heterogeneous model with  $V_p/V_S = 2.45$ . This simple exercise reaffirms the results from previous studies that the realistic velocity structure, especially the shallow  $V_S$  structure, is necessary to estimate better the potential tsunami hazards<sup>186</sup>.

I also compare the ground motions that would be recorded at a station in the coastal region (Fig. 5.9 (c) - (d) and Fig.B.1). The strong ground motions that are responsible for damaging urban infrastructure may arrive as distinct high-frequency bursts from the downdip part of the megathrust<sup>19,97,171</sup>. Moment-normalized velocity and acceleration seismograms produced by the different models of this study have relatively similar peak amplitudes. The earliest peak amplitudes of ground motions occur when the rupture hits the trench. However, the duration of strong shaking is much greater in realistic structures. I attribute this to the wave reverberation in the wedge (wave propagation effects) and not a source effect since the source duration is comparable ( $\sim 60$  s). The presence of the LVZ naturally increases the strong ground motion hazard: it is located nearby the coastal regions and tends to produce three times more HF seismic power than in reference, uniform models (Fig. 5.7). Previous studies have illustrated the existence of distinct strong-motion generation areas (SMGAs)<sup>19,97,171</sup>. The SMGAs imply that there may be heterogeneity in the LVZ such that the spatial variations in elastic structure may control variations in slip-front healing (i.e., more or less healing of the slip pulse). These can also be modeled by heterogeneity in fault properties<sup>129</sup>.

## 5.6 CONCLUSION

Global databases of BP images show a systematic depth variation of the frequency content in source radiation. While this finding was discussed in Lay et al.<sup>175</sup> for several large events, here I show that it is a systematic pattern among most moderate-to-large subduction zone earthquakes. This study provides a simple and generalizable explanation of this observation. I find that the inclusion of Earth's free surface is sufficient to explain this ubiquitous observation. I propose that the dynamics of shallow rupture are dominated by free-surface effects that are, in turn, the first-order factor in explaining the depth-frequency relation. The second-order effect is the evolution of earthquake rupture in a realistic velocity structure that is typical of shallow subduction zones ( $< 50$  km), one that has a compliant wedge and a low-velocity zone atop the downgoing slab. The presence of anomalously low  $V_S$ , relative to  $V_P$ , also impacts the rupture behavior that further enhance the depth-dependence of seismic radiation. Furthermore, my findings resonate with previous work that realistic structures are necessary to correctly model tsunami and ground motion hazards in future subduction zone earthquakes<sup>186</sup>. Because elastic wavespeed properties are likely better constrained than frictional properties at depth, my study promotes the use of tomographic images in dynamic rupture modeling and ground motion predictions.

There are several key limitations to this work and avenues to improve upon it. Free-surface effects consist of multiple factors including fault geometry/curvature, depth of earthquake rupture, seafloor topography, bulk properties, which could be explored in a rather systematic way in future analysis. My preliminary attempts to produce a synthetic backprojection by coupling the dynamic rupture models using SPEC2D<sup>304</sup> failed due to a poor resolution of the BP peaks in the 2D dynamic modeling setting. Part of this limitation is likely due to the 2D modeling against 3D modeling, which would provide more spatial dimension to separate the BP peaks. 3D realistic structure effects may matter more for the along-strike propagation of rupture, which is not explored in this

setting. I have also ignored the water layer, though this should not affect too much the rupture process<sup>168</sup>. I also have not included inelastic rheology, which would smooth the slip evolution at the trench and further enhance the depth-frequency relation<sup>190,191</sup>. These are fantastic avenues for future work.

## 5.7 DATA AND RESOURCES

The high resolution velocity model of Japan trench is from Seiichi Miura at JAMSTEC (Japan Agency for Marine-Earth Science and Technology). The back-projection results in IRIS database is downloaded from the IRIS data services products: back-projection (<https://ds.iris.edu/ds/products/back-projection/>, last accessed on 02/27/2021). The Slab2 model is downloaded from U.S. Geological Survey (<https://www.sciencebase.gov/catalog/item/5aa1b00ee4b-0b1c392e86467>, last accessed on 02/27/2021). The code of dynamic rupture simulation is SEM2DPACK and can be downloaded from <https://github.com/jpampuerto/sem2dpack>. I use resources from Harvard FAS Research Computing (<https://www.rc.fas.harvard.edu>) to solve for the dynamic rupture models. All the down-sampled simulation results and relevant scripts are in Figshare ([https://figshare.com/projects/The\\_Earth\\_surface\\_controls\\_the\\_depth-dependent\\_seismic\\_radiation\\_of\\_megathrust\\_earthquakes/98360](https://figshare.com/projects/The_Earth_surface_controls_the_depth-dependent_seismic_radiation_of_megathrust_earthquakes/98360)) for result reproduction.

6

# Systematic analysis on earthquake source time functions reveals patterns in earthquake dynamics

## SUMMARY

Earthquake source time functions (STFs) carry information about the complexity of seismic rupture. We developed different metrics to extract physical information from a large population of earthquake STFs. We explore STF databases and find that their shapes have significant information about earthquake dynamics. Our first method, the Gaussian-subevent decomposition (GSD), shows that earthquake complexity, as represented by the number of subevents, grows with earthquake magnitude. Patterns in rupture complexity arise from a scaling between subevent moment and main event moment. Our second method, dynamic time warping (DTW) clustering, exhibits different degrees of complexity of the STF shapes and suggests an association between STF complexity and earthquake source parameters. The results from two independent methods are consistent with and complementary to each other. In order to investigate possible factors that control the STF complexity patterns, we perform a large number of 2-dimensional dynamic rupture simulations with stochastic distributions of pre-stress, and apply the same metrics to the resulting synthetic STFs. We have two main findings: 1) heterogeneity in the stress field on the fault interface is necessary to reproduce the large STF variability observed from our subevent-decomposition results; 2)

variation of frictional parameters can obviously change the group proportions of STFs from our DTW clustering. Our results imply that there is very important dynamic information hidden in the earthquake variability, and it is very helpful for constraining rupture dynamics from the analysis on population behaviors of large samples of earthquake STFs.

## 6.1 INTRODUCTION

Earthquakes are known to break in diverse manners: some events rupture on a geometrically simple fault with a relatively smooth slip distribution<sup>345</sup>, while others break a network of faults and/or have heterogeneous slip distribution<sup>11,54,179,207</sup>. Although the complexity of earthquakes can be directly observed, in some cases, from surface fault trace<sup>146,179,201</sup>, many ruptures are buried at depth so that seismic waves are the only observations available to infer the source process. Derived from seismic waves through waveform deconvolution or kinematic inversion, the earthquake Source Time Function (STF) is a foremost important seismic observation that describes the time history of moment release during a rupture. Moreover, the shape of the STF directly controls the variability and uncertainty in the strength and duration of strong ground motion.

Conventional approaches to estimating the STF fall into two categories. The first inverts for a kinematic evolution of slip on the fault from recorded body and surface waves and yields the moment rate function<sup>141,153,354,352</sup>. Due to data resolution, this approach works best for large earthquakes; published databases have few examples. The second approach makes a point source approximation to the rupture and directly deconvolves recorded seismic waves with a Green's function<sup>222,297,311</sup>. This second method assumes that the far-field pulse is proportional to the moment rate function when averaged over stations, is appropriate for small to moderate earthquakes, and results in larger ensembles of STFs.

Observations of global earthquake STFs and source spectra have shown significant inter-event

variability among earthquakes<sup>10,20,59,352,74</sup>. Such variability may partly come from differences in data processing strategy<sup>134</sup>. Therefore, large catalogs of STFs (or their spectra) obtained from a uniform approach is preferable to analyze relative differences among earthquakes<sup>10,59,67,76,311,352</sup>.

Recently, such catalogs of STFs (or of their spectra) have enabled multiple discoveries about earthquake source processes. For example, the total seismic moment  $M_0$  (the time integral of the STF) scales with source duration  $T^3$  (the duration of the STF) for most small to moderate size earthquakes, which implies that the earthquake stress drop is roughly invariant with earthquake magnitudes. At larger magnitudes, this scaling may differ (e.g.  $M_0 \sim T^2$  from Denolle & Shearer<sup>76</sup>). Their properties also have indicated that the ratio of the radiated energy  $E_R$  over the moment, also referred to as the scaled energy  $E_R/M_0$ , varies spatially and with depth but remains invariant with earthquake magnitude<sup>25,67,76</sup>.

However, both the amplitude and the source duration of the STF vary by orders of magnitude. This requires careful strategies of amplitude and time scaling for across-magnitude visualization and comparison. One approach is to scale the time axis to a duration metric and normalize the amplitude to seismic moment (i.e. the integral of the STF). However, source duration is difficult to measure because near-source and near-site scattering of seismic waves may interfere with waves radiating from the end of the seismic rupture. Therefore previous studies have proposed several metrics of duration: moment-based duration<sup>121</sup>, threshold-based duration<sup>74,309</sup>, and centroid-based duration<sup>203</sup>. Because these measures are not strictly equivalent, the shapes of the scaled and stretched STFs differ as well. For instance, Meier et al.<sup>203</sup> find that average STFs have rather a triangle shape whereas Denolle<sup>74</sup> suggests a rather skewed-Gaussian functional form.

Here, we propose to weaken the assumption of a choice in source duration metrics and instead use Gaussian-subevent decomposition (GSD) and dynamic time warping (DTW) to compare the shapes of the STFs. We apply these methods to the global SCARDEC (Seismic source ChAracteristics Retrieved from DEConvoluting teleseismic body waves) catalog of STFs<sup>311</sup> (available at <http://>

[scardec.projects.sismo.ipgp.fr/](https://scardec.projects.sismo.ipgp.fr/), last accessed 01/20/2020), which contains 3529 earthquakes of magnitude greater than 5.5 from 1/1/1992 and until 12/31/2018. The SCARDEC functions are constructed from the deconvolution of teleseismic P waves (P, pP, sP, PP, and PcP) with a theoretical Green's function that is estimated from a radially symmetric and anelastic Earth. Because the global models of attenuation are better constrained by seismic frequencies lower than 1 Hz, we do not interpret signals that are shorter than 1 s in the following analysis.

GSD decomposes the STF as a sum of “subevents” that are Gaussian pulses, similar to the pulse stripping method of Kikuchi & Kanamori<sup>151</sup> and Zhan et al.<sup>365</sup>. We iteratively perform the subevent decomposition from onset of rupture (time zero): searching for the best-fit Gaussian pulse for the first peak, then subtracting that Gaussian subevent from the STF, repeating this for the residual until no obvious peaks can be found. In this way, we can quantify the STF complexity with the number and magnitudes of those subtracted Gaussian subevents. A clear scaling pattern has been found between the main event magnitude and corresponding subevent magnitudes.

DTW<sup>1,33</sup> performs a non-uniform stretching of time and amplitude to match the shape of two time series via the optimal warping path with minimum distance. We measure the similarity between STFs with the DTW distance and cluster the STFs according to the DTW distance. We re-group the clusters based on their complexity and finally form 4 groups. We find that the association to a particular group or degree of complexity is correlated with several earthquake source parameters, such as focal mechanisms, depth, and scaled energy.

To test whether the current physical understanding of earthquake processes reproduces the observations on the STFs, we perform dynamic simulations of earthquake ruptures with linear slip-weakening friction to construct synthetic STFs. We find that heterogeneity in the stress field on the fault interface is necessary to reproduce the large STF variability observed from our GSD results. Furthermore, we find a strong correlation between the grouping distribution of STF shapes from DTW and frictional parameters, such as the characteristic slip-weakening distance  $D_c$ . The propor-

tion of groups in the SCARDEC database is most similar to those simulated STF<sub>s</sub> with small values of  $D_c$ , thus the group proportion of a large number of STF<sub>s</sub> can potentially provide observational constraints to earthquake dynamics.

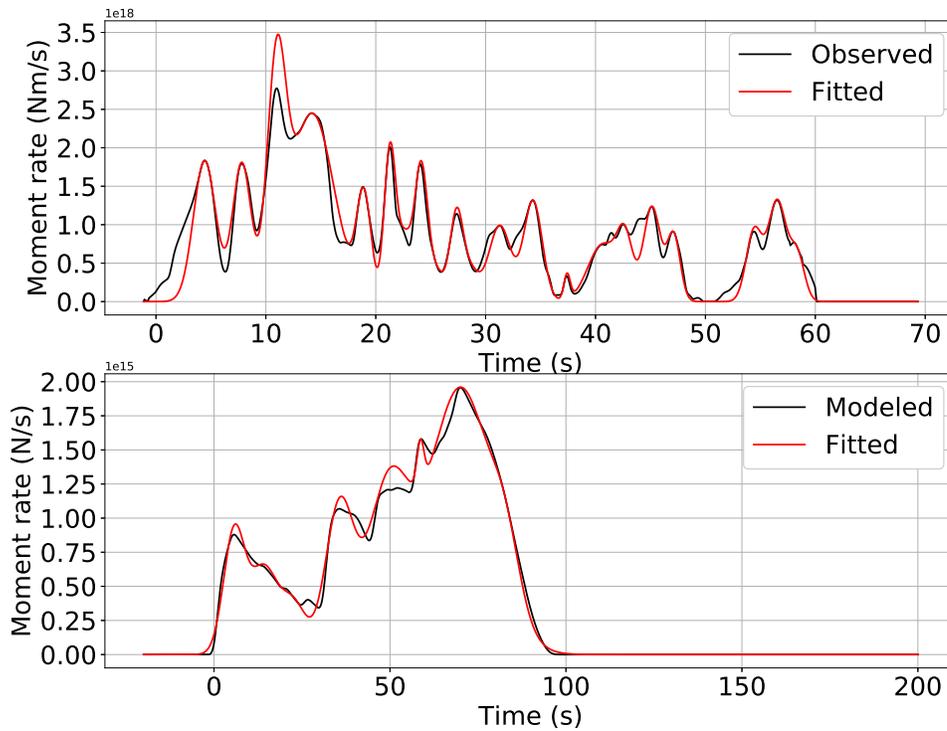
## 6.2 GAUSSIAN-SUBEVENT DECOMPOSITION

### 6.2.1 METHOD OF GSD

The roughness of STF<sub>s</sub> may be described in several ways. Several studies use the zero crossings of the moment acceleration function<sup>122,121,246,271</sup>, while others use the log-residuals between the STF and a smoothed model<sup>203</sup>. Here, we decompose the STF as a sum of “subevents” that are Gaussian pulses, similar to the pulse stripping method<sup>151,365</sup>. Gaussian subevents were found to best fit the shape of the SCARDEC STF<sub>s</sub>. We perform the subevent decomposition from onset of rupture (time zero) as follows:

1. go forward in time and detect a peak  $S_S$  (local maximum over three points or 0.21 s) at time  $t_S$  that satisfies  $S_S > 0.1 \max(\text{STF})$ ;
2. fit a Gaussian function centered around  $t_S$  to the STF with an amplitude to  $S_S$  and a width  $\sigma$  estimated using a grid search minimizing the root-mean-square residuals (effectively to get a subevent duration and moment) over 11 grid points, or 0.77 s;
3. if  $4\sigma > 1$  s, count the detection as subevent and move forward, otherwise go back to step 1;
4. subtract the fitted Gaussian function from the STF;
5. if more time remains in the STF, return to step 1.

We apply the same algorithm to all STF<sub>s</sub> (simulated and the observed SCARDEC). Examples of the reconstructed STF<sub>s</sub> are shown in Fig.6.1 for both the observed STF and the simulated STF. The



**Figure 6.1:** (top) Observed Source Time Function (STF) from the SCARDEC database from the 14 March 1994 M7.1 Central Mid-Atlantic Ridge Earthquake in black and the fitted Gaussian-built STF in red. (bottom) Modeled STF in moment rate per unit of fault width (black) and the corresponding fitted Gaussian-built STF (red). This illustrates the reliability of our fitting procedure in capturing the shape and moment of STFs.

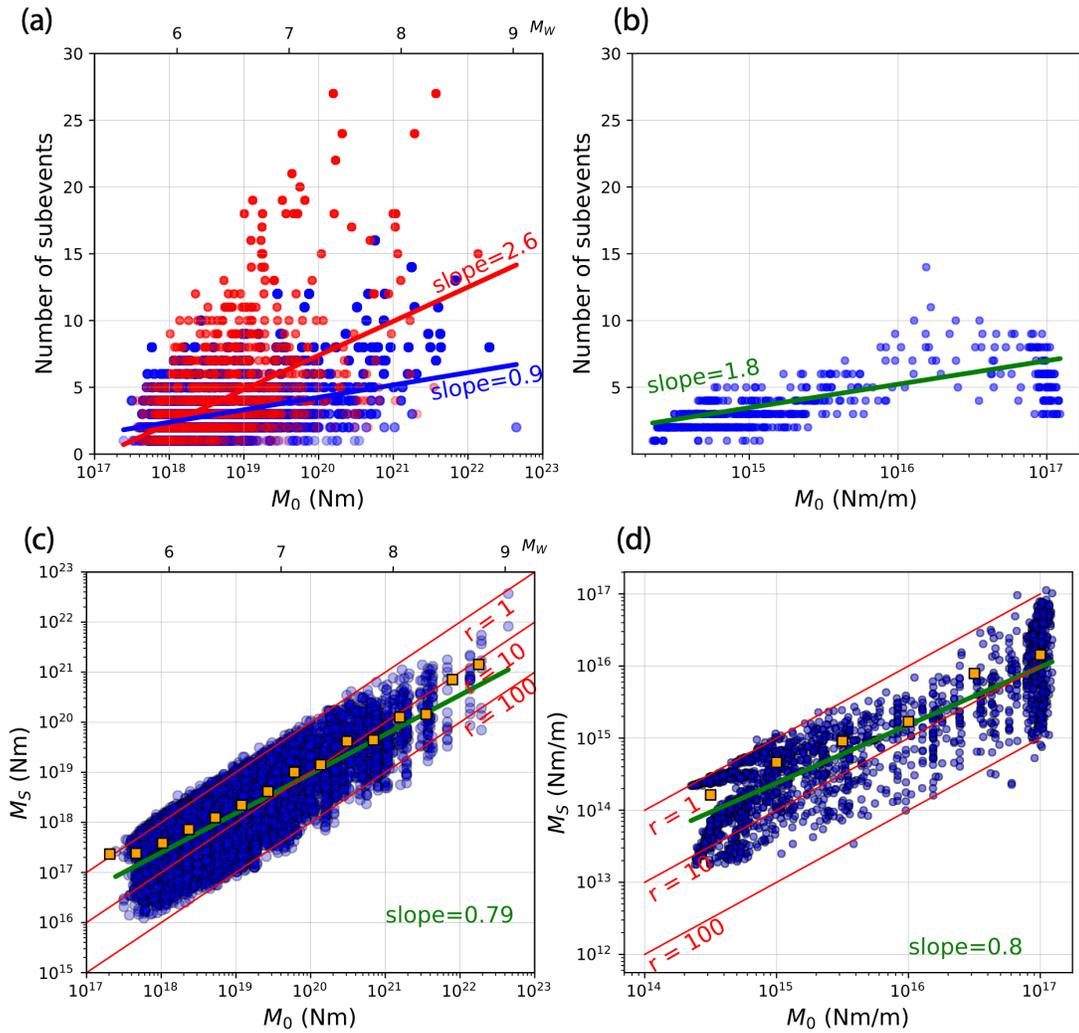
choice of a threshold to select the peak based on the maximum amplitude of the STF, here  $\sigma_{\tau}$ , is necessary to ignore spurious residuals that are not resolvable by the data (e.g., for shorter duration or lower amplitude signals). Other thresholds resulted in detecting smaller subevents that we interpret as being overfit. This interpretation is based on the low likelihood of resolving small subevents given the depletion of high-frequency energy at teleseismic distances. Regardless of the threshold, the results presented below remained unchanged.

### 6.2.2 RESULTS OF GSD

Using the number of subevents as a metric for earthquake complexity, we find our first main result: earthquake complexity increases with earthquake size (Fig.6.2 (a)). In general,  $M_0 < 4 \times 10^{19}$  (Nm) ( $\sim M_w 7$ ) earthquakes have about 1 – 4 subevents (mean of 3.08), while the  $M_0 > 4 \times 10^{19}$  earthquakes have 4 – 10 subevents (mean of 5.84).

The growth in the number of observed subevents with earthquake magnitude is more pronounced for strike-slip earthquakes with magnitudes greater than  $M_w \approx 6.75$  ( $M_0 \leq 1.68 \times 10^{19}$  Nm) as the slope is 2.5 subevents/ $\log(M_0)$  for lower magnitudes and 3.3 subevents/ $\log(M_0)$  for greater magnitude. These earthquakes have a source dimension  $\sim 15$  km, a typical seismogenic depth of crustal earthquakes (Fig.6.2 (a)). Large crustal strike-slip earthquake tend to have subvertical faults limited in width by the seismogenic depth and are known for their complex multifault geometry (2001 Kokoxili <sup>160,302</sup>, 2002 Denali <sup>91</sup>, 2012 Sumatra <sup>207,363</sup>, 2016 Kaikōura <sup>109,339</sup>). Off-shore strike-slip event complexity may be overestimated, however, by contamination of the direct seismic phases with their reverberation in the water column <sup>362</sup>. Furthermore, among the strike-slip earthquakes of magnitudes greater than  $M_w 6.5$ , the 267 crustal events (depth lower than 35 km) have a median number of 5 subevents while the 142 deeper events have a median number of 3 subevents. Large dip-slip earthquakes tend to exhibit less complexity. Their growth in the number of subevents is clearer for moments of  $M_0 \leq 6 \times 10^{20}$  ( $M_w 7.8$ ), where the slope becomes 2.5 subevents/ $\log(M_0)$  if we ignore the  $M_w 9.0$  2011 Tohoku earthquake. The  $M_w 9.0$  2011 Tohoku is an outlier and an extreme case in this regard consistent with rupture of a single, large patch <sup>133,280</sup>. Depth has less influence for  $M_w > 6.5$  dip-slip earthquakes since both crustal and deep earthquakes have a median number of subevents of 3.

Our second result from GSD is that subevent moment is correlated with main event total moment (Fig.6.2 (c)). To our knowledge, this finding has not previously been reported in the litera-



**Figure 6.2:** Results of GSD for both SCARDEC STF and simulated STF. Top row shows the number of subevents as a function of main event seismic moment for (a) the observed SCARDEC STF and (b) the modeled STF. Dots are the individual earthquakes. Dots are colored according to whether the faulting type parameter FM defined in Shearer et al. <sup>283</sup>, which is between  $-0.5$  and  $0.5$  for strike slip (red dots), and between  $-1$  and  $-0.5$  or  $0.5$  to  $1$  for dip slip (blue dots). The slopes that best explain the variance in a linear regression of both subsets are shown in colored letters. The growth in complexity is monotonic with earthquake size for the observations, except for the M9.0 2011 Tohoku earthquake, and the growth is noticeable for the simulations. Bottom row shows individual subevent moments plotted against main earthquake moment for observed SCARDEC (c) and modeled (d) STF. In both figures, blue dots represent a single measurement of subevent/main event pair, orange squares represent the medians over moment bins. The upper horizontal axis in (c) shows the equivalent moment magnitude. Red lines indicate a ratio of seismic moment of subevent to earthquake  $r$  of 1, 10, and 100, respectively. The green lines are linear regressions in a log-log space of individual measurements that yield about  $M \propto M^{0.8}$  for the observations (c) and the simulations (d).

ture. Independent of event magnitude, the ratio of the main event moment to subevent moment is between 1 and 100. The lower bound of this distribution arises from the fact that subevents are rarely found to be larger than the main events, which is true by definition, but the approximation of the Gaussian function to the local moment rate function sometimes brings errors. The upper bound of this distribution arises from the choice of threshold to detect subevents and from ignoring the subevents with duration shorter than 1 s ( $4\sigma \geq 1$  s). Of all STF from SCARDEC database, 19 are left out of the analysis with this latter criterion. The scaling remains unchanged by varying the threshold. Another interesting and supporting aspect is that we did not detect small subevents in the large  $M_w$  8+ earthquakes: subevents for such earthquakes are either buried in the signal of the largest subevents and undetectable in our decomposition, or they are absent. The scaling between subevent moment  $M_S$  and main event  $M_0$  for the observed STF is  $\log_{10}(M_S) = 0.79\log_{10}(M_0) + 3.22$ . It is likely that complexity in the small STF is underestimated, as discussed in Houston et al.<sup>122</sup>, because the teleseismic data are less reliable at frequencies greater than 1 Hz. Thus, we may not be able to see the smaller and higher frequency subevents. Furthermore, the use of a symmetric Gaussian function for the pulse shape may not be physical<sup>301</sup> as it is intrinsically smoothed. Similar limitations apply to the simulated STF given the low-pass filtering we apply and given the numerical limitations of the experiment.

## 6.3 DYNAMIC TIME WARPING AND CLUSTERING ANALYSIS

### 6.3.1 METHOD OF DTW

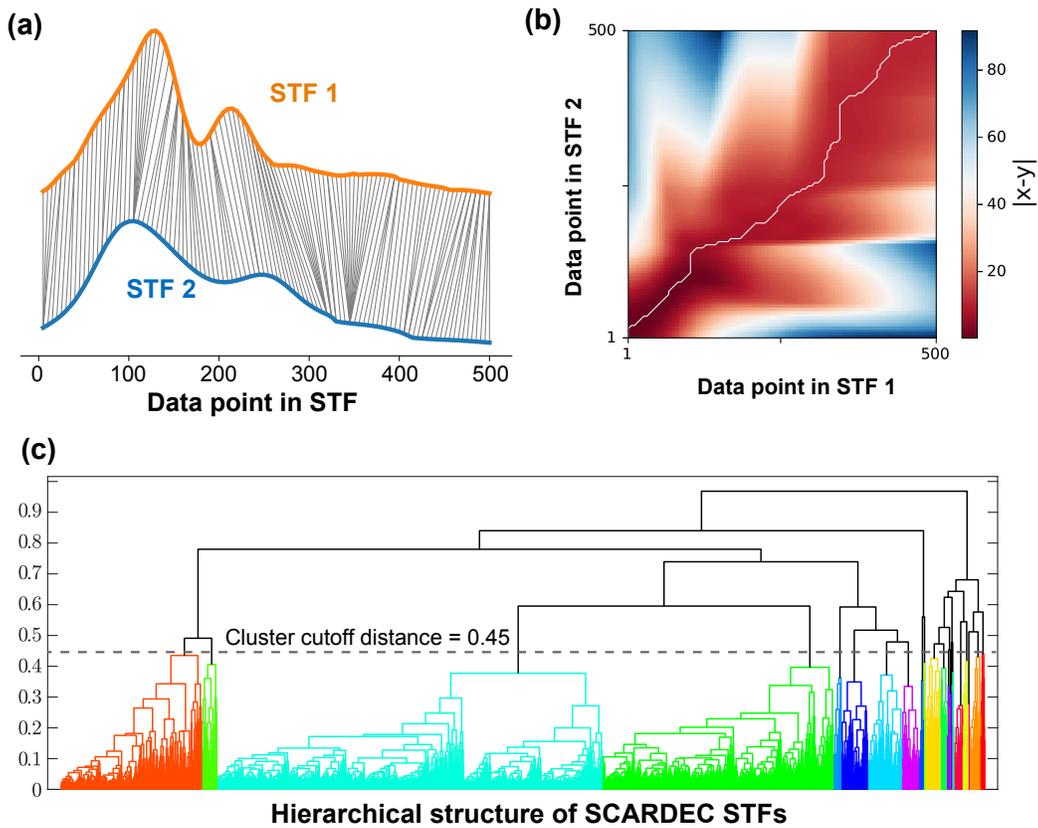
DTW measures the similarity between two time series that may not share the same frequency content or the same sampling rate. The series are “warped” (or stretched) non-uniformly in the time dimensions to optimally match two series (Fig.6.3). This algorithm is widely used in automated speech recognition in which different audio sequences may have different speaking speeds<sup>1,33</sup>. One

important advantage of DTW is its ability to preserve topological structures of the time series by assimilating their temporal elongation or compression. Once stretched, the DTW distance is taken as a new metric for STF similarity, which can be used for clustering. Our approach follows four steps:

1. Pre-conditioning of STF shapes;
2. Calculation of DTW distance;
3. Hierarchical clustering based on DTW distance ;
4. Re-grouping clusters around centroid events;

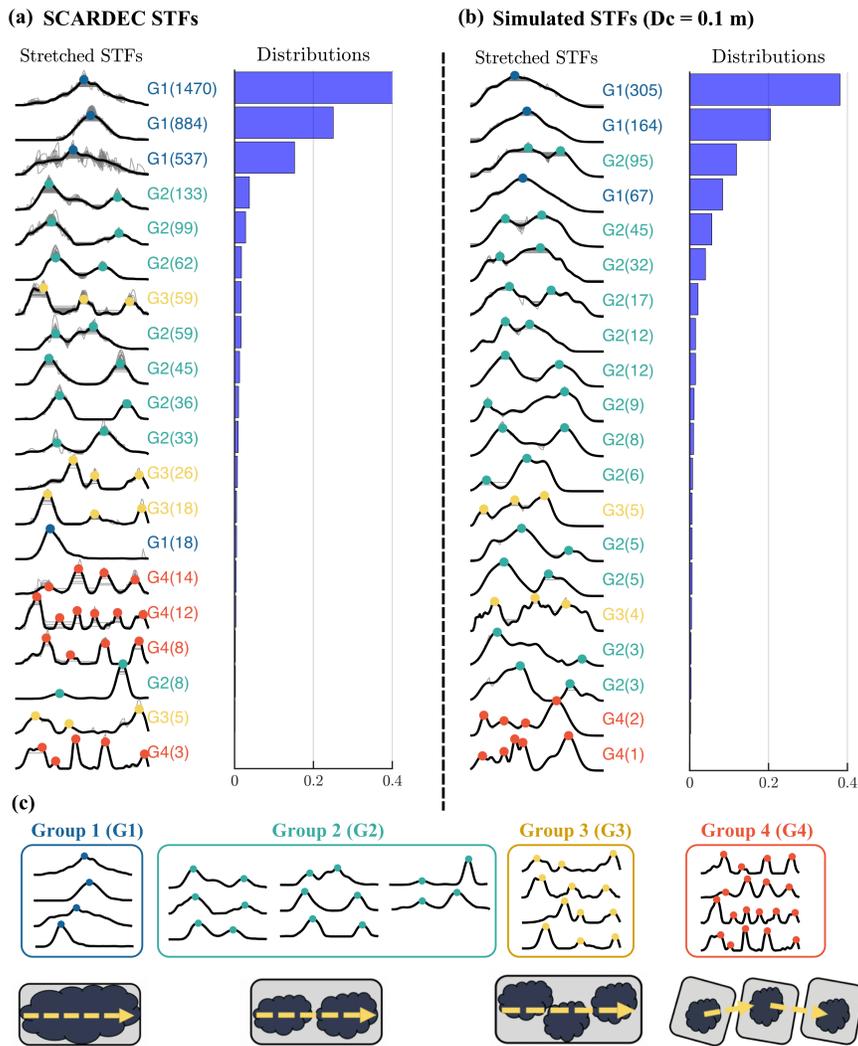
We first perform minimal pre-conditioning of the STF shapes. The STFs are built from the deconvolution of teleseismic P waves that are relatively well constrained at frequencies below  $1 \text{ Hz}^{311}$ . Given that the maximum duration of the STF in the catalog is about 100 s, we re-sample the data to 100 points giving a minimum sampling rate of 1 point per second. We then normalize the amplitude STFs to the event seismic moment. These two processing steps improve the stability of the warping. We have tested various strategies to resample and normalize the STFs, which did not affect the conclusions of this analysis.

Second, we apply the DTW to each pair of STFs. The DTW distance is the Euclidean distance between two STFs warped along the optimal warping path, and is chosen here as the measure of similarity between two STFs (see Fig.6.3 (a) - (b)). We apply a global search of the shortest warping distance, and thus allow for the maximum level of distortion. The warping follows causality such that the order of the peaks is kept. Knowing that the distortion induced by nonlinear stretching is unphysical, we will mainly focus on the STF general shape, which is well preserved when choosing an appropriate maximum distortion. Unlike for GSD results, here we do not seek to interpret individual stretched peaks as physical subevents.



**Figure 6.3:** Dynamic time warping (DTW) clustering of earthquake source time functions (STFs). (a) Point-to-point correspondence between two example STFs. (b) Optimal stretching path (white line) from the minimum differences for the two example STFs. (c) Hierarchical structure of all SCARDEC STFs from the DTW clustering.

Then, the STF shapes are clustered based on their DTW distance with a single-linkage hierarchical clustering analysis that provides the flexibility to form clusters at any desired level (Fig. 6.3 (c)). Here, we choose the threshold of clustering cutoff distance to be 0.45, which corresponds to 20 clusters. The DTW distance threshold, or choice in number of clusters, controls the degree of distortion. Thus, by choosing a proper number of clusters, we keep the diversity of clusters and limit the unphysical distortion. For each of these clusters, we choose the representative STF (defined as the centroid event) that has the minimum median distance to all of the other members of the

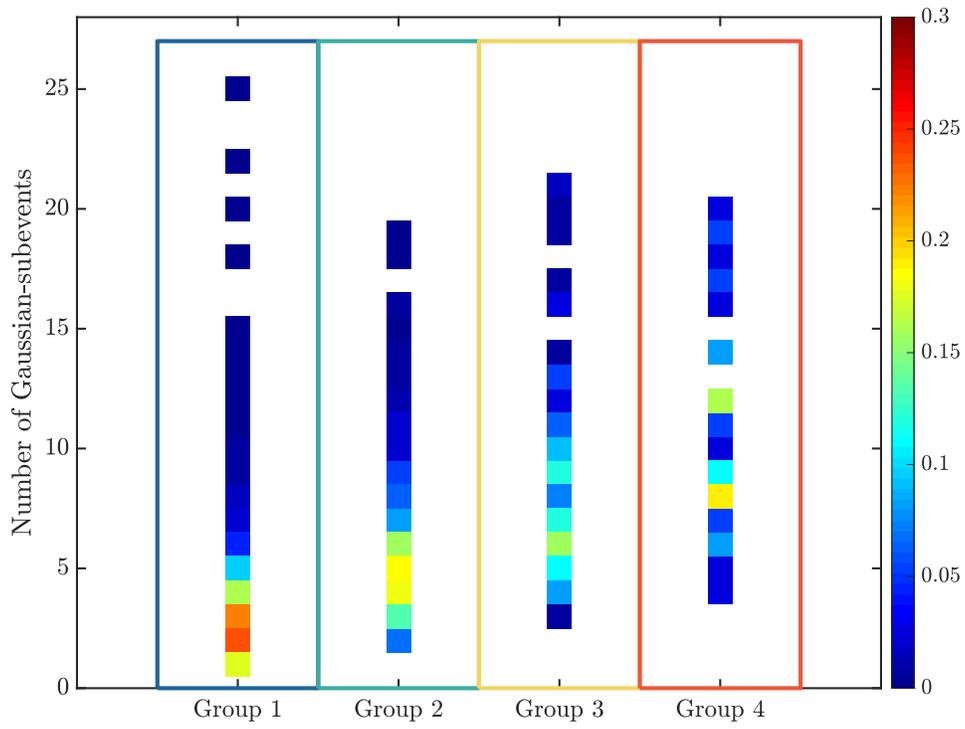


**Figure 6.4:** Source time function clustering, grouping, and conceptual interpretation. (a) Individual STFs after dynamic time warping and clustering are shown by gray thin lines. Black thick lines are the STFs of the centroid event of each cluster. Colored dots indicate the prominent peaks of the centroid STF as well as the associated group. Numbers in the parentheses are the number of STFs in each cluster. The corresponding population proportion of each cluster is shown in the right histograms. (b) Same as (a) but for the STFs from our dynamic simulations. (c) Cluster centroid STF shapes and conceptual models for G1-G4. In the model diagram, dark blocks represent major rupture asperities and the arrow indicates the rupture direction.

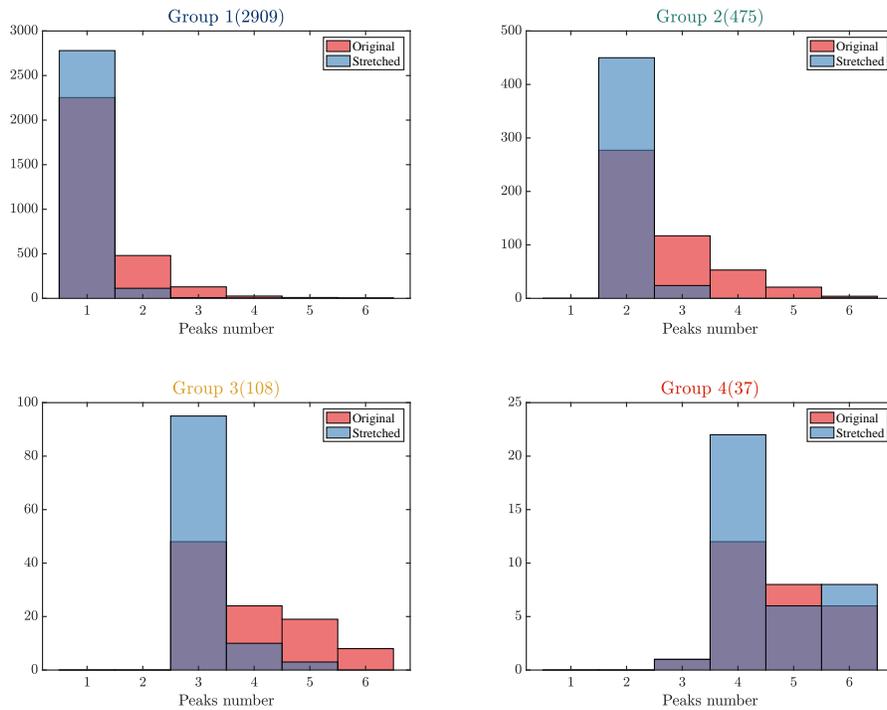
cluster. It is similar to the stack of all stretched STFs within each cluster (Fig.6.4), which, in turn, exhibits the common features of all cluster members.

Next, we parameterize the characteristic STF shape for each of these clusters by calculating the number of prominent peaks of each centroid event. The number of prominent peaks is commonly used for topographic relief analysis and is defined as the amplitude of the peak (hill summit) relative to the lowest amplitude point (valley) that does not contain a higher peak. In order to be counted as a peak, we choose a threshold of prominent peak amplitude to be 10% of the global maximum of the STF. The purpose of this step is to avoid counting the small-amplitude spurious peaks, which are usually caused by imperfect Green's function removal, as the false prominent peaks. Previous studies have used zero-crossing of the time derivative of the STF<sup>122,121,246,271</sup>, which is also sensitive to spurious peaks because it does not account for the signal amplitudes. We also test 1% and 5% threshold and the conclusions remain unchanged. We notice that the stretched STFs have a lot fewer prominent peaks than individual peaks from the GSD (Fig.6.5). Moreover, the stretched STFs have fewer prominent peaks than the raw STFs, but in general the same number of prominent peaks as the centroid event (Fig.6.6). For instance, a STF may have multiple separated amplitude peaks, but only one single prominent peak (Fig.6.4 (a) - (b)). These differences are because our new metric is defined for the complexity of general shape extracted from large number of STFs, instead of the detailed features of each individual STF.

Finally, we group the clusters based on the number of prominent peaks of the centroid event: G1 is the group where the centroid event has 1 prominent peak, G2 is the group where the centroid event has 2 prominent peaks, ... (Fig.6.4 (c)). G4 is the group where the centroid event has at least 4 prominent peaks. Examples of detected prominent peaks are found in Fig.6.4 (a) - (b). For DTW results, we define the STFs to be "complex" if their DTW stretched STFs have multiple prominent peaks. The first order result from the grouping is that most events have a single prominent peak whereas about 20% events are more complex.



**Figure 6.5:** Comparison between the GSD subevent number and DTW complexity groups. The color indicates the frequency of occurrence within each group.



**Figure 6.6:** Comparisons between prominent peak (0.1 of STF global maximum) number distributions of original raw STFs (red histograms) and DTW stretched STFs (blue histograms) in each group. Group numbers are also the prominent peak numbers of the centroid event within each group.

### 6.3.2 RESULTS OF DTW

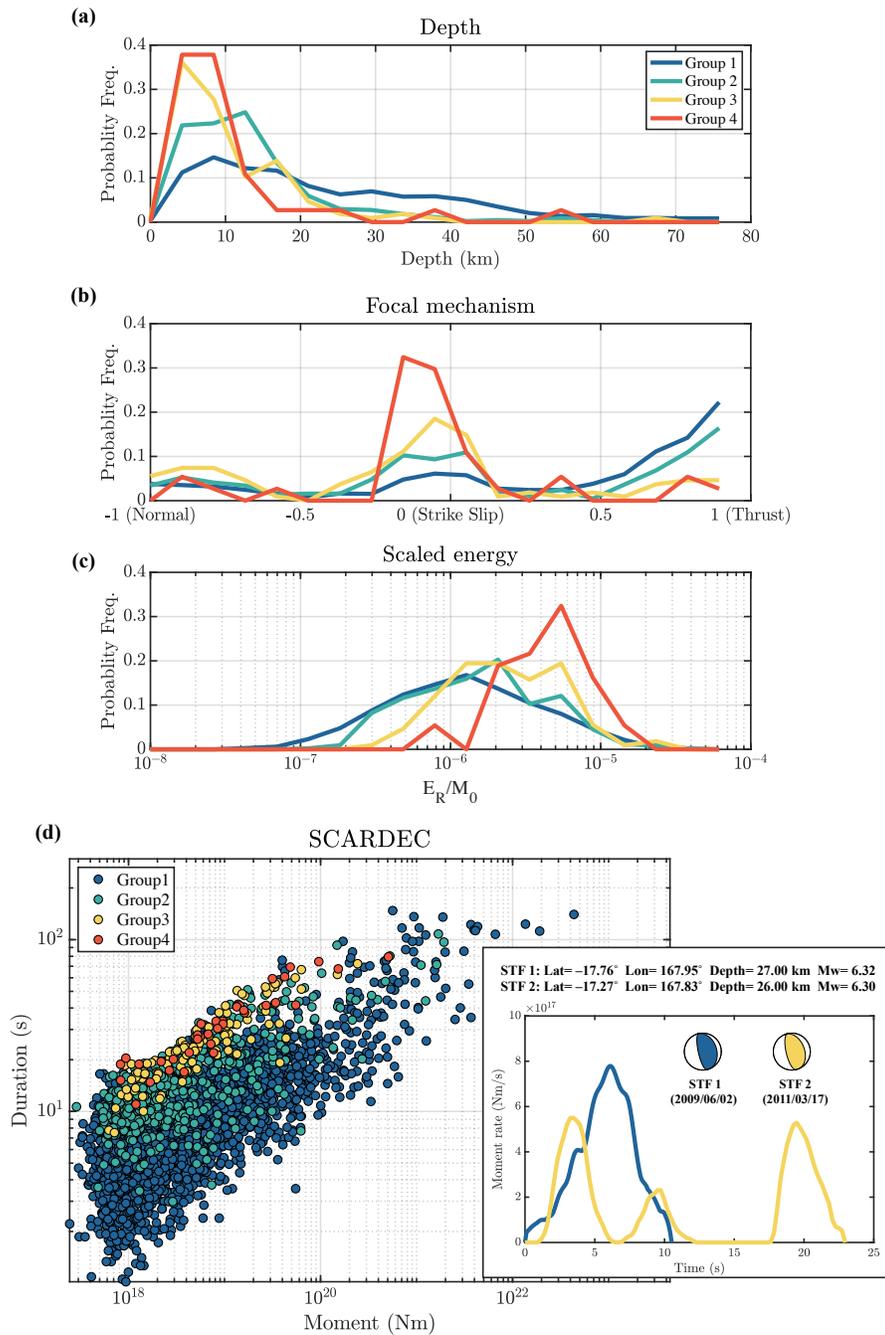
We now explore the correlation between grouping and several source parameters such as depth, focal mechanism, moment, duration, energy, and location.

The first property we investigate is the source depth. Complex STFs (groups G<sub>2</sub>-G<sub>4</sub>) are mostly shallow crustal events whereas the simple STFs (group G<sub>1</sub>) can be found at all depths (Fig.6.7 (a)): G<sub>1</sub> 35.37%; G<sub>2</sub> 73.68%; G<sub>3</sub> 86.11%; G<sub>4</sub> 89.19% with depth  $\leq 20$  km. Because co-located events have various degrees of complexity (Fig.6.7 (d), Fig.6.8), inaccuracy in the Green's function does not strongly bias our results.

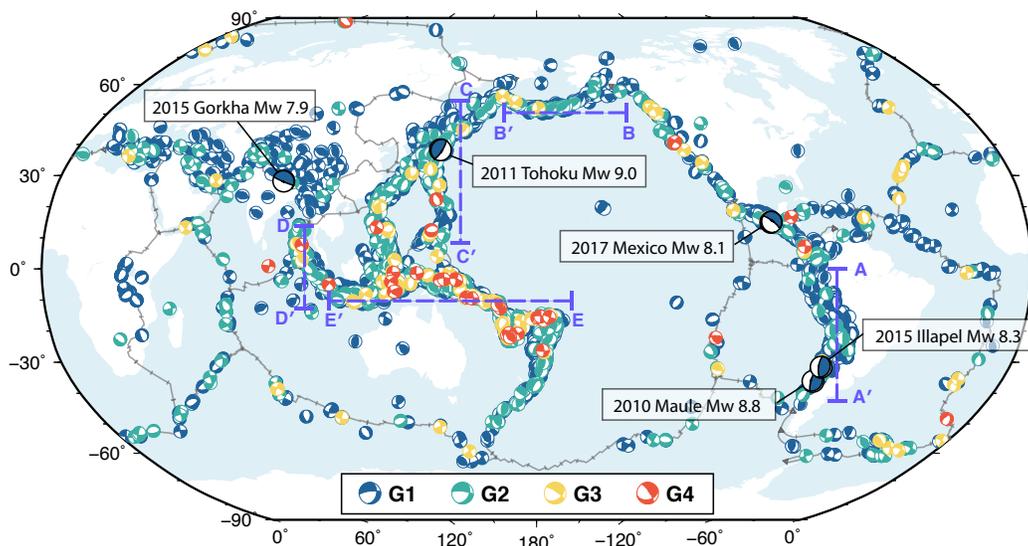
The second property we investigate is the focal mechanism (Fig.6.7 (b)). The focal mechanisms are solved simultaneously by the SCARDEC method<sup>310</sup>. Most of the thrust earthquakes have simple STFs (G<sub>1</sub> and G<sub>2</sub>), whereas the strike-slip earthquakes are dominated by complex STFs (G<sub>3</sub> and G<sub>4</sub>). There are too few normal events in the database (only 17.5 %) to give any significant conclusion regarding this mechanism.

There is no clear relation between earthquake size (seismic moment) and this metric of complexity (see Fig.6.7 (d)). For example in Fig.6.7 (d), we see that the largest events in SCARDEC database may only have one prominent peak in their stretched STF, while the events with smaller moments can be in any of those complexity groups.

We find a clear pattern that G<sub>3</sub>-G<sub>4</sub> events have an abnormally longer duration with respect to other events of similar magnitudes and relative to events of the other groups (Fig.6.7 (d)). It is illustrated in Fig.6.7 (d) by visualization of two STFs of co-located events and of similar magnitudes. For the same earthquake moment (or the STF integral), it is intuitive to understand that STFs in G<sub>4</sub> have multiple low amplitude prominent peaks and overall extended duration, compared to the G<sub>1</sub> STFs that have a single high amplitude and short duration peak. Simple models of crack ruptures yield a relation between moment, source duration, and stress drop<sup>46,92</sup> that could indicate low stress



**Figure 6.7:** Population distribution of four complexity groups and correlation with different source parameters: (a) centroid depth, (b) focal mechanism (scalar defined by<sup>283</sup> that varies from -1 (normal), 0 (strike-slip) to 1 (reverse)), (c) and scaled radiated energy  $e = E_R/M_0$ . Panel (d) shows the earthquake duration against earthquake moment, colored with the respective group labels. One pair of co-located events with different complexity are also shown in the inset.



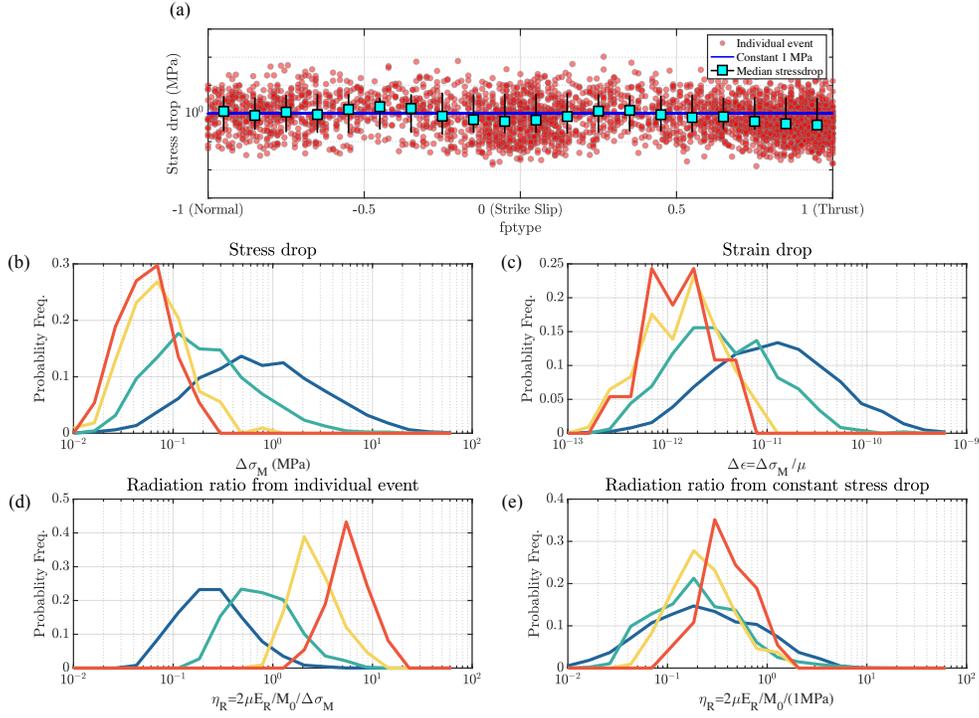
**Figure 6.8:** Map of focal mechanisms colored by their group label and overlay of the plate boundaries (gray thin lines). Several recent large megathrust earthquakes are highlighted. Blue dashed lines show the locations of profiles in Fig.6.10.

drops for the G<sub>4</sub> events (Fig.6.9 (a) - (c)).

We now explore the clustering results against the earthquake scaled energy. Here we calculate radiated energy from the squared time derivative of the STF (moment acceleration function  $\ddot{M}_0(t)$ ) using the relation  $E_R = (\frac{1}{15\pi\rho V_p^2} + \frac{1}{10\pi\rho V_s^2}) \int_0^\infty (\ddot{M}_0(t))^2 dt$ . We select depth-dependent bulk properties ( $V_p$  P-wave velocity,  $V_s$  shear-wave velocity,  $\rho$  density) from PREM<sup>89</sup>. Radiated energy scales almost linearly with seismic moment and we calculate the scaled energy as the ratio of both radiated energy and seismic moment, which is about invariant with earthquake size<sup>67,76</sup>. Fig.6.7 (c) shows the distribution of the scaled energy with respect to each group. G<sub>3</sub> and G<sub>4</sub> events have systematically larger scaled energy as G<sub>1</sub> and G<sub>2</sub> events. This is consistent with intuition that G<sub>3</sub> and G<sub>4</sub> events generally have rougher STFs.

The correlations between STF complexity and source depths and focal mechanism are consistent with the findings from previous studies<sup>70,121,309</sup>. In particular, shallow strike slip earthquakes are constrained geometrically by the Earth surface on the top and the seismogenic depth on the bottom. They also tend to be composed of segmented faults<sup>159</sup>. These geometrical settings control the evolution of rupture that tends to operate with moving energetic slip pulses<sup>147</sup> with repeated rupture acceleration and deceleration as they travel across segments<sup>54,144,248</sup>.

Since earthquake source parameters are closely related to the local tectonic regime, we also find that our observations from the clustering and grouping results (G<sub>1</sub> - G<sub>4</sub>) are consistent to the marked variation of tectonic environments (Fig.6.8). Many of the major subduction zones are dominated by the simpler types of events (G<sub>1</sub> and G<sub>2</sub>) and lack of more complex ones, likely because they are dominated by thrust events located along/within the subducting slabs at various depths. For example, since 1992, there have been only two events ( $M_w > 5.5$ ) belonging to the G<sub>3</sub> group along the Southern American and Aleutian subduction zones, respectively (Fig.6.10 (a) - (b)). Similarly, other subduction zone regions like in Japan and in Sumatra, the Indian-Eurasian collision zone are also dominated by simple-type earthquakes (Fig.6.10 (c) - (d)). In contrast, the complex group



**Figure 6.9:** Source parameter estimates of the SCARDEC database. Panel (a) shows the stress drop against with focal mechanisms parameters: the stress drop<sup>46,92</sup> is calculated as  $\Delta\tau = 7/16M_0/(0.32V_sT_D)^3$ , where  $T_D$  is the measured STF duration and  $V_s$  is taken at the source depth from PREM, and the focal mechanism parameter *fptype* is described in Shearer et al.<sup>283</sup>. In the simulations, the average stress drop of all models is approximately 1 MPa. Panel (b) shows the group distributions of these estimated stress drop according to grouping number. Note that the stress drop estimation based on duration may be underestimated for the very heterogeneous earthquake ruptures of the G3 and G4 groups<sup>235</sup>. Panels (c) and (d) show the group distributions of corresponding strain drop and radiation ratio calculated from stress drop, respectively, noting that radiation ratio might be overestimated<sup>235</sup>. Panel (e) also shows the group distributions of radiation ratio, but estimated based on the assumption that stress drop is a constant value of 1 MPa, for comparison.

(G<sub>3</sub> and G<sub>4</sub>) events are located mostly along the boundaries around the junction region of the Indo-Australian, western Pacific, Philippine plates and Eurasian plates (Fig.6.8 and Fig.6.10 (e)). Bird<sup>38</sup> explored and documented the kinematics at plate boundaries and found that this region is characterized by a particularly extensive number of micro plates, whose boundaries exhibit varied relative motions and kinematics (their Figure 6). Therefore, we propose that the complexity in the STF may reflect the complexity in the regional stress field.

#### 6.4 MODELING STF COMPLEXITY WITH DYNAMIC RUPTURE SIMULATION

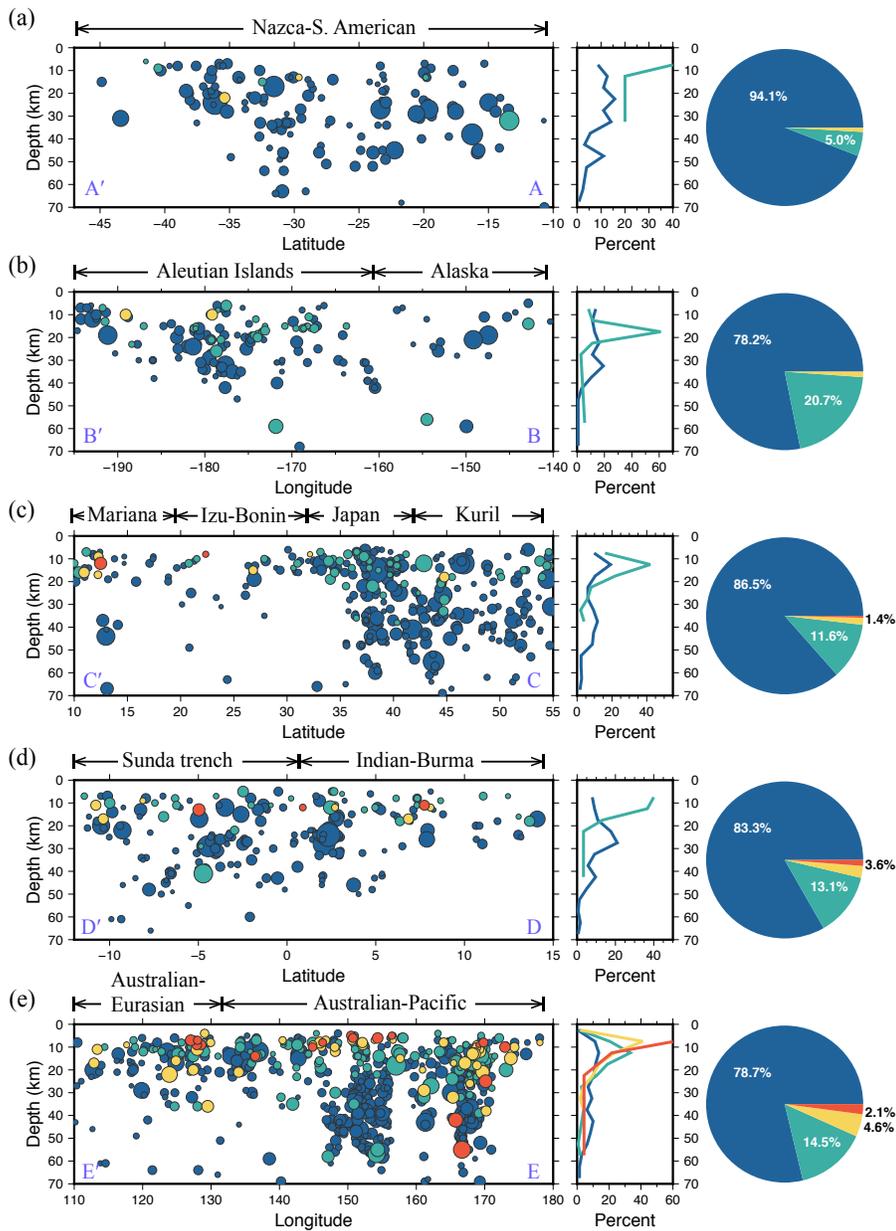
##### SIMULATIONS OF DYNAMIC RUPTURES USING STOCHASTIC DISTRIBUTIONS

Simulations of dynamic ruptures using stochastic distributions of fault-interface parameters are popular in the investigations of complex kinematic source models, realistic fault geometry and roughness models, and to simulate high-frequency ground motions<sup>106,196,197,260,305</sup>. In order to investigate possible factors that control the STF complexity patterns, we perform a large number of 2-dimensional dynamic rupture simulations with stochastic distributions of pre-stress, and apply the same clustering analysis to the resulting synthetic STFs as to the SCARDEC STFs.

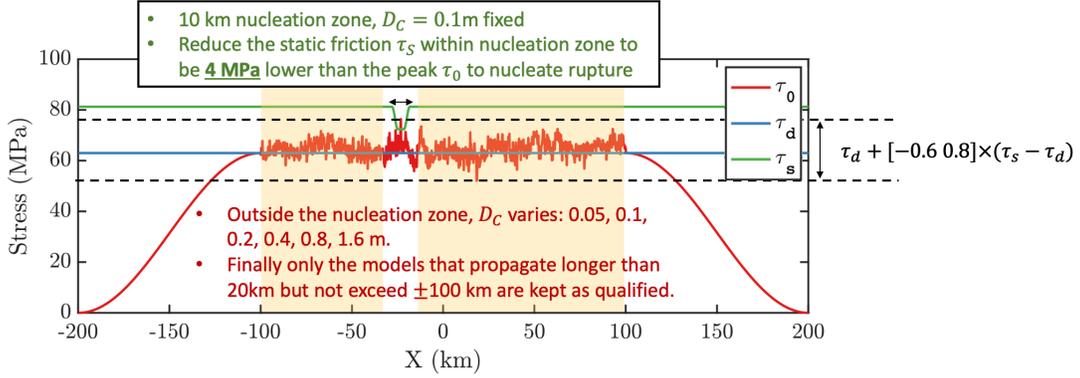
In this study, synthetic dynamic sources are generated in a 2-dimensional medium in an anti-plane setting. Pre-stress on the fault is constrained to follow a power-law amplitude distribution that approximates the scenario caused by natural fault roughness<sup>52</sup>. To generate diverse dynamic ruptures, we generate statistically similar shear pre-stress distributions  $\tau_0(x)$  on the fault plane. The distribution is the sum of a perturbation  $d\tau_0(x)$  and of a uniform level of shear stress  $\mu_d\sigma_0$ :

$$\tau_0(x) = \mu_d\sigma_0 + d\tau_0(x). \quad (6.1)$$

The power spectral density (PSD) of  $d\tau_0(x)$ ,  $dT_0(k)$ , follows power-law decay in the wavenumber



**Figure 6.10:** Earthquake distributions of different complexity groups on the vertical profiles (from 0–70 km, locations are indicated by blue dashed lines in Fig.6.8). The regional along-depth and total group distributions are also shown to the right.



**Figure 6.11:** Pre-stress (red curve) and frictional strengths (green curve: static friction; blue curve: dynamic friction) settings of the dynamic rupture simulations. Dashed lines indicate range of values of the randomly generated pre-stress. Finally, only the rupture models terminates within the yellow shadow regions are kept as the qualified models.

domain,

$$dT_0(k) = C|k|^{-\gamma}, \quad (6.2)$$

where  $\gamma = 0.8$  is based on observational constraints on the self-affine fault roughness<sup>52,87</sup>, and  $C$  is a normalization factor. Combining the PSD  $dT_0(k)$  with the random phases  $\varphi(k)$ , which are taken from a uniform distribution in  $[0, 2\pi]$ , we can generate various pre-stress distributions. For each realization of a pre-stress perturbation, we further scale the pre-stress perturbation amplitude to vary within the range from  $-0.6(\mu_s - \mu_d)\sigma_0$  to  $0.8(\mu_s - \mu_d)\sigma_0$ . Finally, we apply a Tukey-window to taper the 100 km on either end of the 400 km pre-stress distributions; this avoids an abrupt ending of rupture at the fault boundary, which is rarely observed in real STFs (Fig.6.11).

We assume a constant normal stress of 120 MPa and linear slip weakening friction law<sup>16,132,242</sup>. The linear slip weakening friction is used as a simple but general constitutive relation:

$$\mu = \begin{cases} \frac{(\mu_d - \mu_s)d}{D_c} + \mu_s, & d \leq D_c, \\ \mu_d, & d > D_c, \end{cases} \quad (6.3)$$

Linear slip weakening requires three parameters: the static friction coefficient (here chosen as  $\mu_s = 0.677$ ), the dynamic friction coefficient (here chosen as  $\mu_d = 0.525$ ), and the characteristic slip-weakening distance  $D_c$ . There is a trade-off between strength excess and  $D_c$  in controlling rupture velocity and the resulting ground motions<sup>107</sup>. Here we choose to focus on  $D_c$ . The use of dynamic simulation in this study is not intended to compare the shape of individual subevents, or to relate the time and shape of these subevents to physical properties on the field. Instead, we use these simulations to provide an ensemble of realistic STFs and analyze the statistical properties of their shapes. While we keep  $D_c$  constant within a single set of simulations, we carry several sets of experiments with values of  $D_c$  at various levels 0.05, 0.1, 0.2, 0.4, 0.8, and 1.6 m that are within bounds found in the literature.

For each  $D_c$ , we first generate a set of pre-stress distributions that we use in each simulations. The dynamic rupture is solved by 2D boundary integral method SBIEMLAB (<https://github.com/jpampuelo>, last accessed February 2021). We discard the rupture models that unsuccessfully nucleated with a source dimension less than 20 km, or rupture beyond the zone of heterogeneous pre-stress, and obtain 800 qualified simulations for each  $D_c$  value. Finally, the STFs are calculated from the integral of the moment-density-rate functions over the fault surface.

## COMPARING DYNAMIC RUPTURE SIMULATION RESULTS WITH STF OBSERVATIONS

With large number of simulated rupture models and their synthetic STFs, we can directly compare with the observations from SCARDEC STFs.

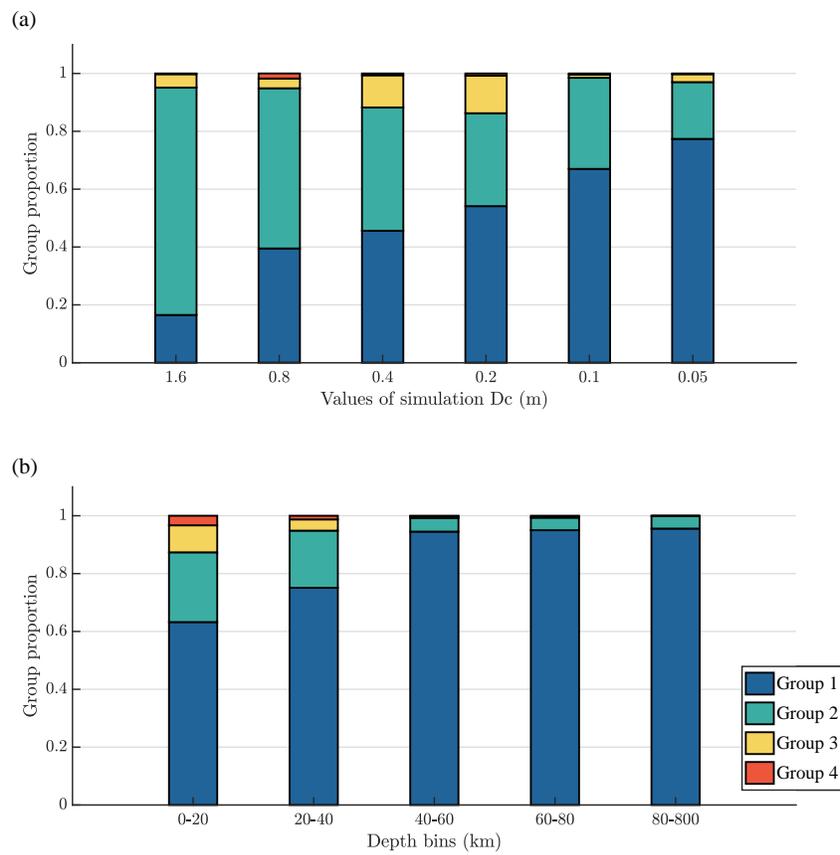
We first apply GSD to all the simulated STFs. We find a similar pattern of monotonic growth in the number of subevents in our numerical simulations (Fig.6.2 (b)). There is a close relation between pre-stress variations and rupture velocity. This relation is itself a basic aspect of elastodynamic crack propagation<sup>93,167</sup>. Subevents initiate when the rupture front accelerates due to a region of favorable pre-stress and terminate when the rupture front decelerates due to a region of unfavorable

pre-stress. More importantly, we find that heterogeneity in the stress field on the fault interface is necessary to reproduce the large STF variability and the monotonic growth of subevent number observed from GSD results. If we apply homogeneous pre-stress (or heterogeneous pre-stress with small variation) in our model settings, we are not able to get consistent results.

We also investigate the scaling relation from the simulated STF:  $\log_{10}(M_S) = 0.8\log_{10}(M_0) + 2.30$  (Fig.6.2 (d)). This scaling relation is different from what we see from realistic STFs. Our simulations are 2-D, whereas real earthquakes in nature occur on 3-D faults. For this reason, the two scaling relations (Fig.6.2 (c) - (d)) are not expected to be identical. We are, however, confident that a scaling slope lower than 1, which would support earthquake self-similarity, is a robust feature for both the observed and simulated STFs.

We further apply DTW to the simulated STFs. Our results indicate that the small values of  $D_c < 0.1$  m are probably necessary to produce the general level of complexity of the SCARDEC STFs (Fig.6.12 (a)). Furthermore, we notice that the distributions among group numbers vary systematically with depth in the SCARDEC database (Fig.6.12 (b)). The STFs of shallow crustal earthquakes present a diversity in complexity similar to that obtained in the simulations when using  $D_c \sim 0.1$  m. The STFs of deep mantle earthquakes present a diversity in complexity that can be obtained with much smaller values  $D_c (\leq 0.05\text{m})$ . Depth variations in  $D_c$  have been reported in earlier studies. Wibberley & Shimamoto<sup>342</sup> perform laboratory experiments on samples from the Median Tectonic Line in southwestern Japan, and estimate that  $D_c$  ought to vary with depth, with a deeper (6 km) values being systematically 30% smaller than the shallow (2 km) values. Kinematic source inversions also find a systematic depth variation of rise time, which they attribute to a systematic dependence in  $D_c$ <sup>135</sup>. Our results may provide a supporting evidence that the characteristic slip-weakening distance, or more generally the fracture energy that is proportional to the product of  $D_c$  and stress drop<sup>107</sup>, varies at depth over crustal scales.

Except for the on-fault frictional properties simulated in this study, different faulting mechanisms



**Figure 6.12:** Group proportion distributions: (a) simulated STFs clustering with different values of  $D_c$ . (b) Group proportions of real STFs (SCARDEC) within different depth bins.

for shallow (0 - 80 km) and deep (80 - 800 km) earthquakes can potentially be another factor for the variation of STF complexity along depth. For instance, the deep earthquakes are inferred to have different faulting mechanisms such as transformational faulting, dehydration embrittlement, and thermal runaway<sup>?</sup>. However, whether these different faulting mechanisms leads to various STF shapes is beyond the scope of our simulations in this study.

## 6.5 CONCLUSION AND DISCUSSION

In this study, we developed different metrics to extract physical information from STFs in SCARDEC database. We apply GSD to analyze earthquake STFs and their subevents. We make several observations based on GSD results: first, large earthquakes have more subevents than small earthquakes; second, the subevent moment scales with the main event moment with a power exponent of about 0.8. Moreover, we develop a dynamic time warping methodology to cluster a large number of earthquake source time functions into different complexity groups based on similarity of their general shapes. We find the patterns of STF shape complexity correlate with different source parameters such as depth, duration, focal mechanism, scaled energy.

Through dynamic rupture simulation, we find that simulations with spatially uniform fault parameters fail to produce subevents but that introducing a range of heterogeneity levels in the prestress yields similar results. We also find that the frictional parameter  $D_c$  affects the proportions of different complexity groups. Comparing the diversity in the STF complexity from SCARDEC database to that from the simulated STF, we suggest that small values of  $D_c$  are a viable explanation to the distribution between simple and complex events.

Moreover, the DTW results of SCARDEC STFs show that the shallow crust has relatively larger proportion of complex groups than the deep region, as does the STFs simulated with a large  $D_c$ . Based on this, the observed variations of complexity with depth could be explained with a depth

variation of  $D_c$ , or more generally, the fracture energy.

Furthermore, we compare our results from the two different metrics. The definition of complexity in GSD is the total number of individual peaks, which they referred to as subevents. GSD results show that strike slip events had more complex STF compared to thrust earthquakes. The DTW results is complementary to GSD results, and adds more detailed information in three ways. First, there is no obvious correlation between earthquake magnitude and the number of prominent peaks. This suggests that earthquakes have a limited number of large subevents (prominent peak). Second, DTW results help us to analyze the relation between degree of complexity and other source parameters, such as the scaling between duration and moment (sometimes used to estimate earthquake stress drop) and the ratio between radiated energy and moment. Taken together, it is reasonable to infer that the complex STFs exhibit large radiation ratio (proportion of radiated energy over available energy). Finally, the modeled STFs exhibit different degrees of complexity depending on the frictional properties.

There are several limitations to our approaches. First, the database we use is constructed from a Green's function in a radially symmetric Earth. Although this is unlikely to affect the overall results, Green's functions that account for laterally varying structure would improve the temporal resolution of the shallowest events. This requires better understanding of near surface scattering and attenuation. Second, our modeling approach is unable to characterize the correlation between focal mechanisms and STF complexity. Indeed, these parameters could be tested using a 3-dimensional dynamic rupture simulation framework, which is impractical to implement due to high computational expense and the employed statistical approaches. Nevertheless, because fault geometry and fault properties seem to play a dominant role in shaping the source and the resulting strong ground motions, further 3-dimensional modeling and observations are necessary.

## 6.6 DATA AND RESOURCES

All the source time functions are downloaded from SCARDEC source time function database (<http://scardec.projects.sismo.ipgp.fr/>). The dynamic rupture simulation code SBIEM-LAB is developed by Jean-Paul Ampuero (available on <https://github.com/jpampuero>). The GSD scripts and samples of data are available at [github.com/mdenolle/subevents2019/](https://github.com/mdenolle/subevents2019/). The DTW scripts can be obtained on the Github ([https://github.com/yinjiuxun/STF\\_DTW](https://github.com/yinjiuxun/STF_DTW)). Global maps are made by GMT<sup>341</sup> (available at <https://www.generic-mapping-tools.org/>).

7

# A multi-task encoder-decoder to separating earthquake and ambient noise signal in seismograms

## SUMMARY

Seismograms contain multiple sources of seismic waves, from distinct transient signals such as earthquakes to ambient seismic vibrations such as microseism. Ambient vibrations contaminate the earthquake signals, while the earthquake signals pollute the ambient noise's statistical properties necessary for ambient-noise seismology analysis. Separating ambient noise from earthquake signals would thus benefit multiple seismological analyses. This work develops a multi-task encoder-decoder network to separate transient signals from ambient signals directly in the time domain for 3-component seismograms. I choose the active-volcanic Big Island in Hawai'i as a natural laboratory given its richness in transients (tectonic and volcanic earthquakes) and diffuse ambient noise (strong microseism). The approach takes a noisy seismogram as input and independently predicts the earthquake and noise waveforms.

The model is trained on earthquake and noise waveforms from the STanford EArthquake Dataset (STEAD) and on the local noise of a seismic station. I estimate the network's performance using the Explained Variance (EV) metric on both earthquake and noise waveforms. I explore different network architectures and find that the long-short-term-memory bottleneck performs best

over other structures, which I refer to as the WaveDecompNet. Overall I find that WaveDecompNet provides satisfactory performance down to signal-to noise-ratio (SNR) of 0.1.

The potential of the method is 1) to improve broadband SNR of transient (earthquake) waveforms and 2) to improve local ambient noise to monitor the Earth structure using ambient noise signals. To test this, I apply a short-time-average to a long-time-average (STA/LTA) filter and improve the detection 27 times. I also measure single-station cross-correlation and autocorrelations of the recovered ambient noise and establish their improved coherence through time and over different frequency bands. I conclude that WaveDecompNet is a promising tool for a range of seismological research.

## 7.1 INTRODUCTION

A seismogram is a record of how the ground moves and usually contains a rich mix of different seismic signals. They may be transient such as the ground motions caused by earthquakes, surface processes (glacier sliding<sup>181,334</sup>, landslides<sup>149,337</sup>), human activities (cars, trains, ships, machinery from factories,<sup>274</sup>). They may be more diffuse such as the microseism<sup>55</sup>, the seismic hum<sup>255</sup>, river noise<sup>49</sup>, urban life<sup>176</sup>. The transient motions capture the seismic signature of their source (earthquakes, landslides, glacial sliding) and thus are essential information to understand these processes (event detection, location, discrimination, source properties). The diffuse ambient seismic field, on the other hand, found its use in correlation seismology to extract spatial and temporal variations in the Earth structure<sup>5,61,278,281</sup>. Therefore, separating the earthquake and ambient noise signals can significantly improve seismological studies from different perspectives: robust event source characterization and robust imaging and monitoring of the Earth's interior.

Many studies across the sciences have focused on removing the diffuse ambient data from the transient signals. In seismology, the diffuse field is often considered as “noise”. The task of denois-

ing earthquake signals aims to improve the signal quality, and the most commonly used method is Fourier-based spectral filtering. This approach assumes that the earthquake and ambient noise signals are well separated in the frequency domain. Although this technique has been proven to be effective in numerous cases, it can distort the signals and/or cause artifacts around the impulsive signals<sup>82,218</sup>.

Earthquake and ambient noise signals often overlap in the frequency domain, and direct filtering may be challenging to separate them<sup>202,247</sup>. Denoising using time-frequency representations of the signal is another widely applied and effective technique to separate the earthquake and ambient noise signals when they overlap. Many innovative algorithms and methods have been developed, for example, using time-frequency transforms such as the Stockwell S-transform<sup>292</sup>, the Radon transform, the wave-packet transform, the continuous wavelet transform, or others using f-x or f-k filtering, singular spectrum analysis, sparse transform-based denoising, which are extensively reviewed and discussed in Mousavi & Langston<sup>218</sup>. Most of these transform-based denoising methods achieve noise suppression through thresholding methods, that is, determining some hard-<sup>81</sup> or soft-<sup>56</sup> thresholds to separate seismic and noise signals. Although those transform-based methods are shown to be very practical and possible to be automated<sup>218,219</sup>; they still require manual intervention, and the parameter tuning is often performed using trial-and-error approaches.

With the recent leap of computational power, memory and data storage, machine learning (ML) has provided a diverse set of powerful tools in the geosciences<sup>31</sup>. Many ML algorithms are built to (1) automatically perform complex prediction task; (2) create a representation that approximates numerical simulations or captures relationships; (3) reveal new patterns, structures, or relationships from data<sup>31</sup>. ML algorithms are powerful in many different seismological tasks, including but not limited to waveform classification and earthquake detection<sup>34,142,163,180,217,221,245</sup>, phase picking and association<sup>180,182,204,217,322,370</sup>, source location and characterization<sup>170,216,223,245,254,312,367</sup>, earthquake early warning<sup>180,224</sup>, and many others.

Machine-learning methods have also been developed to denoise and decompose the seismic data<sup>57,266,288,299,371</sup>. Saad et al.<sup>266</sup> develop a stacked denoising auto-encoder to smooth the ambient noise and output a time series for better arrival time detection, in a similar fashion to an impulsive filter<sup>9</sup>. Zhu et al.<sup>371</sup> develop a DeepDenoiser network, which applies deep learning to the short-time Fourier transform and focuses on classification on pixels of the spectrogram of seismograms<sup>299,371</sup>.

This study aims to separate the waveforms of both earthquake and ambient noise signals through the ML network, thereby benefiting both earthquake and ambient-noise seismology. I develop a new multi-task encoder-decoder ML network, which I name WaveDecompNet, to separate both types of signal simultaneously. I treat the problem as a time series extrinsic regression<sup>296</sup> and directly work on the seismic data in the time domain. This setting can save human intervention by not tuning parameters for the time-frequency representation and provides the excellent potential to be applied in a near real-time framework given its lower computation cost. The performance of the WaveDecompNet is noticeable even at a low signal-to-noise ratio.

I take the Big Island of Hawai'i as a natural laboratory of the seismically complex environment. Oceanic islands record strong microseismic signals, which are the basis for ambient noise seismology<sup>184</sup>. The volcanic region also exhibits dynamic tectonics and volcanic activities. I choose the broadband seismic station IU.POHA at Pohakuloa, Hawai'i. The richness of this dynamical system presents particular challenges in monitoring the volcano-tectonic activities and the temporal evolution of structural changes from ambient-noise seismology.

I further test the applicability of my technique by performing two standard single-station measurements. I first test the improvements on detecting transients by applying the standard short-time-average through long-time-average (STA/LTA) method<sup>9</sup> and find increased trigger rates and potentially improved picking accuracy. I also measure the coherence of single-station ambient-noise cross-correlation functions and find increased stability in the coda of these functions. My results

highlight the potential for WaveDecompNet to improve seismograms used in both event-based seismology and noise-based seismology.

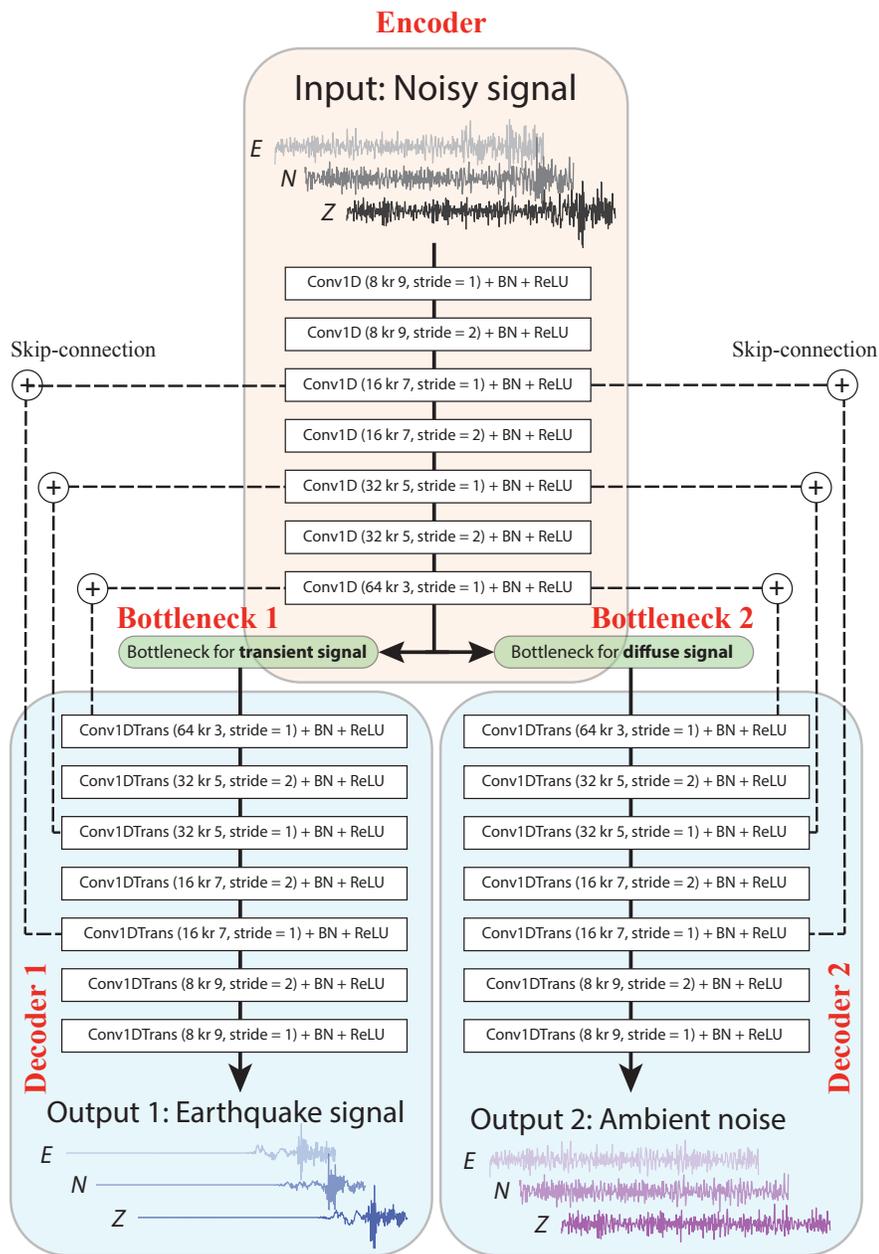
## 7.2 NETWORK DESIGN OF WAVEDECOMPNET

The multi-task encoder-decoder network handles the time series seismic data directly. The input to the network is 1-minute long, 3-component (East-West, North-South, and up-down) raw seismograms. The output of the network is two 1-minute long, 3-component seismograms (earthquake and noise, Fig.7.1). In order to seek computational efficiency, no pre-processing of the data is applied except the amplitude normalization of the waveforms.

### 7.2.1 NETWORK ARCHITECTURE

The encoder-decoder network is a popular network design in ML problems, such as generating dialogues<sup>279</sup>, semantic image labelling<sup>23,24</sup>, detection of image forgeries<sup>26</sup>, and prediction of vehicle trajectory<sup>243</sup>. In this study, my encoder-decoder network consists of 3 major parts: the encoder branch, the two decoder branches, and two bottleneck blocks in between:

1. The role of the encoder is to extract useful, high-level features from the seismic time series. Through training with sufficient data and updating its parameters, the encoder aims to learn features of the input data that can help characterize the earthquake and ambient noise signals. I use one-dimensional (1D) convolutional layers with an increasing number of kernels to extract high-level features with a minimal number of parameters (Fig.7.1). The stride of the convolution is adjusted to down-sample the time series along the time axis. I have tested the use of MaxPooling instead of convolutional strides but found poorer network performance. After each 1D convolutional layer, batch normalization is applied to normalize the output to zero-mean and unit variance. Finally, a rectified linear unit function (ReLU) is



**Figure 7.1:** The multi-task encoder-decoder separates earthquake and ambient noise signals. The network consists of 5 main blocks: the encoder branch, two bottlenecks, and two decoder branches. The encoder and decoder branches contain 7 one-dimensional convolutional and transpose convolutional layers. The layer parameters “x kr y, stride=z” refer to x kernels with y features and stride of z. Each convolutional or transpose convolutional layer is followed by a batch normalization (BN) layer and a ReLU activation layer. The structure of the bottleneck block is tested and discussed in the main text for details. 6 residual connections (skip-connection layer by summation) directly connect the encoder to the decoder to improve the convergence of training and prediction performance<sup>116,261,370</sup>.

used as the activation function to add non-linearity to the network for better regression from time series to time series.

2. The role of the decoder is to translate the learned features from the encoder branch and reconstruct the separated earthquake and noise time series. A two-branch decoder block handles both the separated earthquake and noise waveforms individually and performs better than a single branch network that only outputs the earthquake signal. The branches are composed of 1D transpose convolutional layers. In symmetry with the encoder block, the number of kernels gradually decreases, and I manually select the stride to incorporate the high-level features back into the time domain. Like the encoder block, I also apply batch-normalization and ReLU activation following each 1D transpose convolutional layer. The parameters of the two branches are learned independently.
3. The bottleneck blocks link the encoder and decoder blocks. Their purpose is to learn the mapping relation between the encoder-extracted features of the composite waveform (earthquakes and noise) and the features of the separated earthquake and noise time series, respectively. The design of the bottleneck block greatly impacts the performance of the algorithm and is subject to investigation in this study.

### 7.2.2 DATA

I use the earthquake waveform data from the STEAD<sup>220</sup> (STandford EArthquake Dataset, available at <https://github.com/smousavi05/STEAD>) because of its broad coverage of global earthquakes.

This data set is curated to provide many high signal-to-noise ratio waveforms of local (source-receiver distance less than 350 km) earthquakes and a set of “noise” (non-earthquake) signals recorded globally. There are 234,526 samples of 3-component seismograms of ambient noise and 1,030,231 samples of 1-minute 3-component seismograms associated with 450,000 earthquakes located at various

regions globally<sup>220</sup>. I randomly pick 144,000 STEAD earthquake waveforms based on their SNRs, which is defined as:

$$SNR = \frac{\|S\|^2}{\|N\|^2}, \quad (7.1)$$

where  $\|S\|^2$  and  $\|N\|^2$  are the power of signal and noise, respectively. I only keep the highest quality earthquake waveforms with  $SNR > 10^4$  in the STEAD dataset to approximate a noise-free signal, then lowpass-filter the waveform below 5 Hz and down-sample them from 100 Hz to 10 Hz.

The data set for ambient noise windows combines a “global noise” from STEAD noise waveforms with a “local noise” from IU.POHA station. I randomly pick 100,000 3-component STEAD samples, low-pass filtering first and then down-sample them to 10 Hz. For the “local noise”, I select noise waveforms from 1-month-long continuous seismic data recorded by IU.POHA (from July 31, 2021, to September 1, 2021), and down-sample the data to 10 Hz. The continuous data may contain known and unknown earthquakes. To keep the spectral features of the noise and reduce the effects from these transient signals, I shuffle the phases as follows. First, I transform the seismic data into the Fourier domain, which gives the amplitude and phase spectra. I keep the amplitude spectrum but assign a random phase using a uniform distribution  $-\pi$  to  $\pi$  to each frequency value, then I apply the inverse Fourier transform. I obtain 44,000 samples of 1-minute ambient noise time series.

To represent more time series for better generalization, I apply two strategies of data augmentation<sup>372</sup>. First, I shift the arrival time of earthquake signals randomly in a uniform distribution of -30 to 60 s to handle the uncertain arrival time of earthquake signals in the real application, and allow for a redistribution of the weights in the encoder 1D convolutional layers. Second, I randomly scale the amplitude of earthquake waveforms related to the noise signals to increase the range of SNR from  $10^{-2}$  to  $10^4$ . I linearly stack the shifted and scaled earthquake and ambient noise waveforms to

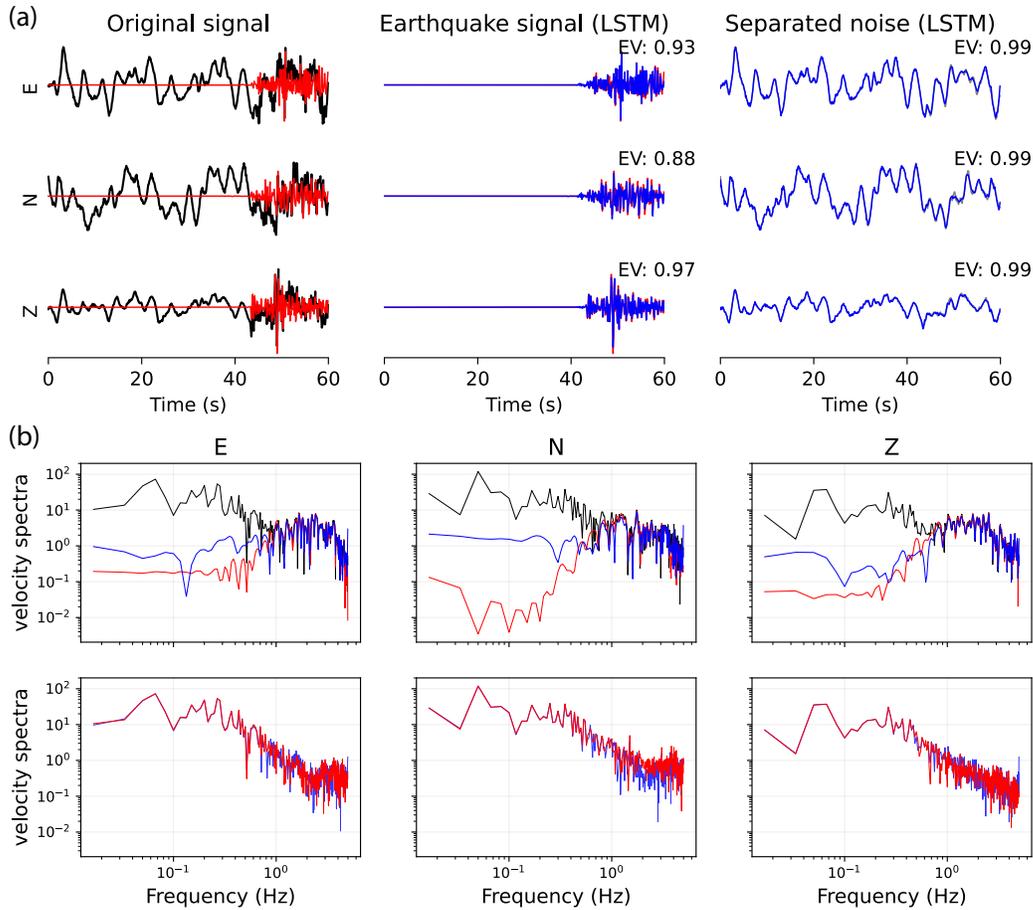
generate the composite waveforms. I report that the resulting SNR of my training data is uniformly distributed in the logspace.

Normalization of the input data is necessary to stabilize the optimization. I use a standard scaler to normalize the time series to zero-mean and unit-variance. I rescale the earthquake and noise signals using that factor after the two decoder branches of WaveDecompNet. The loss error is the mean square error (MSE) function, and the training is performed using the Adam optimizer<sup>154</sup>.

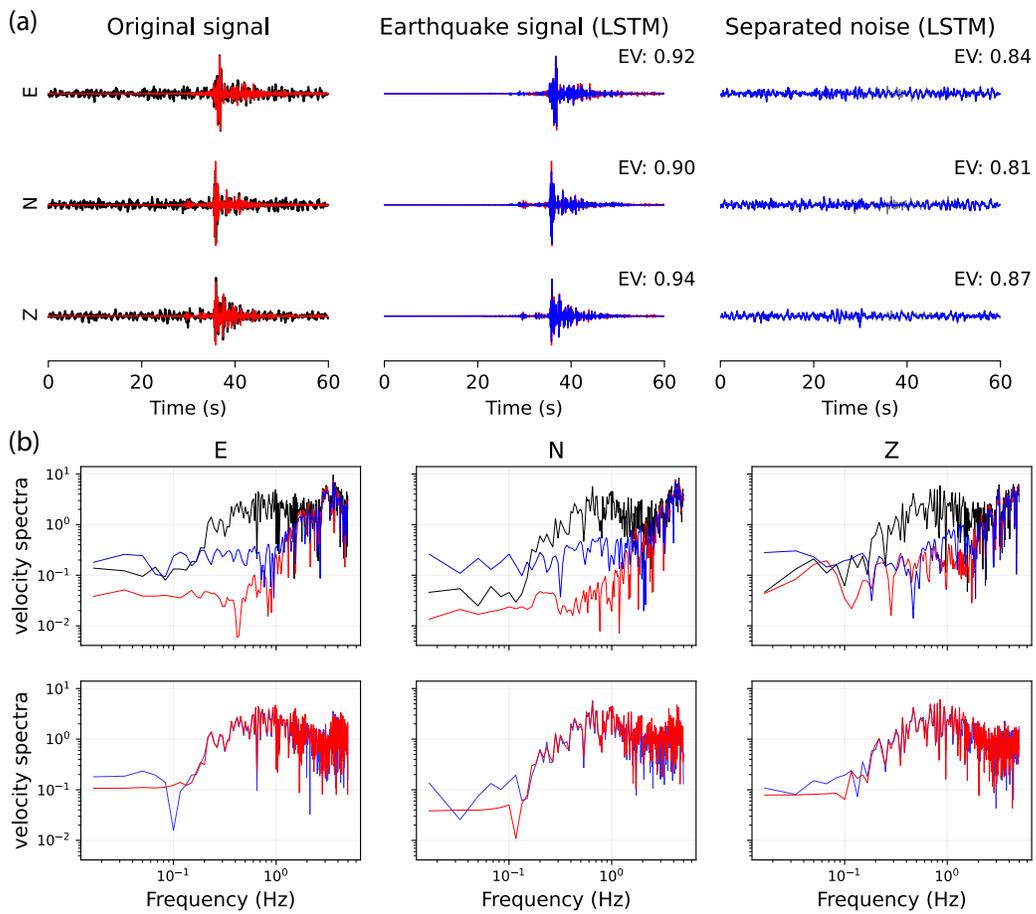
The training, validation, and testing data sets are split using 60%-20%-20%. Only the training data are used to train the network, update the model parameters and minimize the loss function. I use a batch size of 128 during training. The validation data is used to track over-fitting during the training. Over-fitting is also mitigated with an early stopping strategy (<https://github.com/Bjarten/early-stopping-pytorch>) with the patience of 10; that is, the training automatically terminates if the value of validation loss remains unimproved for 10 epochs.

After training, I use the test data set to evaluate the model performance, especially for choosing the best one from models using different bottlenecks (more details follow in the next section).

After training, I evaluate the model performance on the test data set. Figures 7.2 - 7.3 show two examples with different types of noise signals: Figure 7.2 contains the regional noise from the IU.POHA station, which is energetic mostly low-frequency below 1 Hz and characteristic of micro-seism at a dominant period of 7 s, Fig.7.3 contain the STEAD noise signal, which is broadband and rich in high frequencies in a band that overlaps with the earthquake signal. Visually, the earthquake waveforms are relatively well recovered over a broad range of frequencies. In particular, it is able to decompose the signals with overlapping frequency content (Figures 7.2 - 7.3 (b)), which is often a challenge for filtering-based denoising methods<sup>82,218</sup>.



**Figure 7.2:** Example 1 of waveform separation using one of the bottleneck architecture (LSTM, see section 7.2.3). (a) 3-component (E-East, N-North, Z-vertical from top to bottom) velocity seismograms normalized with the same scaling factor of maximum amplitude in each component. (left panels) One STEAD earthquake waveform (in red) and IU.POHA local noise is stacked to get the noise input waveform (in black). (Middle panels) Comparison between the separated earthquake waveforms (blue) with the ground truth earthquake waveform (red). (right panels) Comparison of the separated noise waveform (blue) and ground truth noise waveform (red). (b) 3-component waveform Fourier amplitude spectra. (Top panels) The spectrum of the input waveform is shown in black, with the ground truth earthquake spectrum (in red), the separated earthquake spectrum (in blue). (Bottom panels) The ground truth noise spectrum is shown in red, and the separated noise spectrum is shown in blue.



**Figure 7.3:** Example 2 of waveform separation. Same as in Figure 7.2 except that the noise waveform is from the STEAD dataset.

### 7.2.3 CHOICE OF THE NETWORK BOTTLENECK

The bottleneck block aims at learning the mapping relation between features in the encoder and decoder, and those features are necessary to reconstruct the separated signals in the decoder blocks. There are multiple choices for the bottleneck structure in time series analysis. I explore five of them and evaluate their impacts on model performance:

1. None: no specified bottleneck. The encoder and decoder are directly connected. The total number of trainable parameters in the network is 78,090.
2. Linear: a linear regression layer between the encoder and the decoder. The total number of trainable parameters in the network is 86,410.
3. LSTM: a bidirectional long-short-term-memory (LSTM) layer between the encoder and the decoder<sup>119</sup>. The total number of trainable parameters in the network is 178,442.
4. Attention mechanism: a multi-head attention layer between the encoder and the decoder<sup>314</sup>. I use a 4-head dot-product self-attention layer with a dropout probability of 0.2. Other numbers of heads were tested but did not significantly affect the results. The total number of trainable parameters in the network is 110,858.
5. Transformer: the standard transformer encoder layer made up of self-attention and feed-forward network<sup>314</sup>. The transformer model has been shown to be a powerful tool in different seismological applications such as earthquake detection and phase picking<sup>217</sup>, earthquake source characterization<sup>223</sup> and early warning<sup>224</sup>. I only use one layer but find adding more layers can greatly downgrade the model performance, which I attribute to insufficient training. The total number of trainable parameters in the network is 640,394.

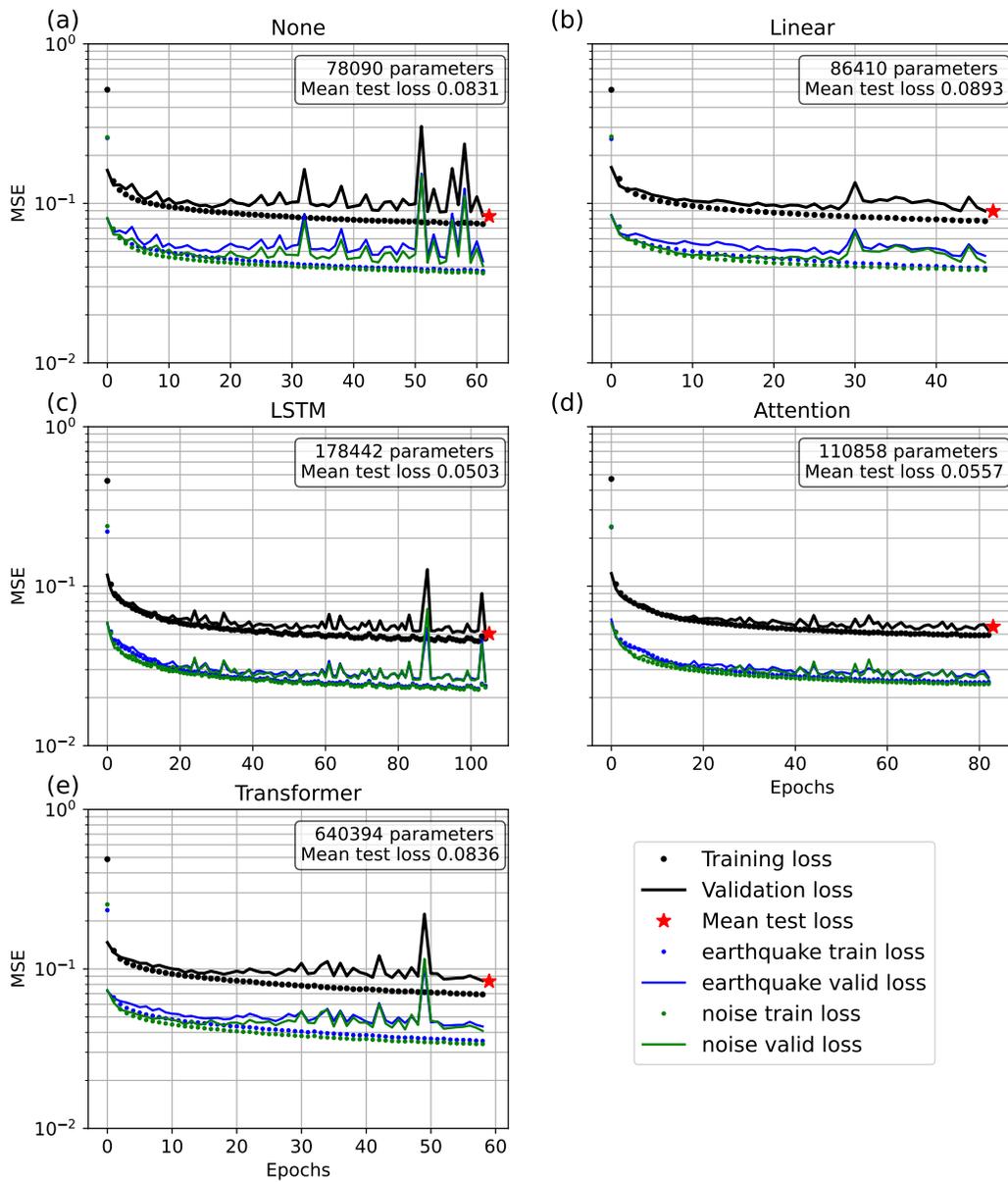
#### 7.2.4 MODEL TRAINING

I use the identical training and validation data set to train all these networks. I train and validate over the batch size of 128. I require each network to be trained for at least 30 epochs. After 30 epochs, I apply the same early stopping patience of 10 epochs. These steps can help guarantee the convergence of each model's optimization while avoiding over-fitting (Fig.7.4).

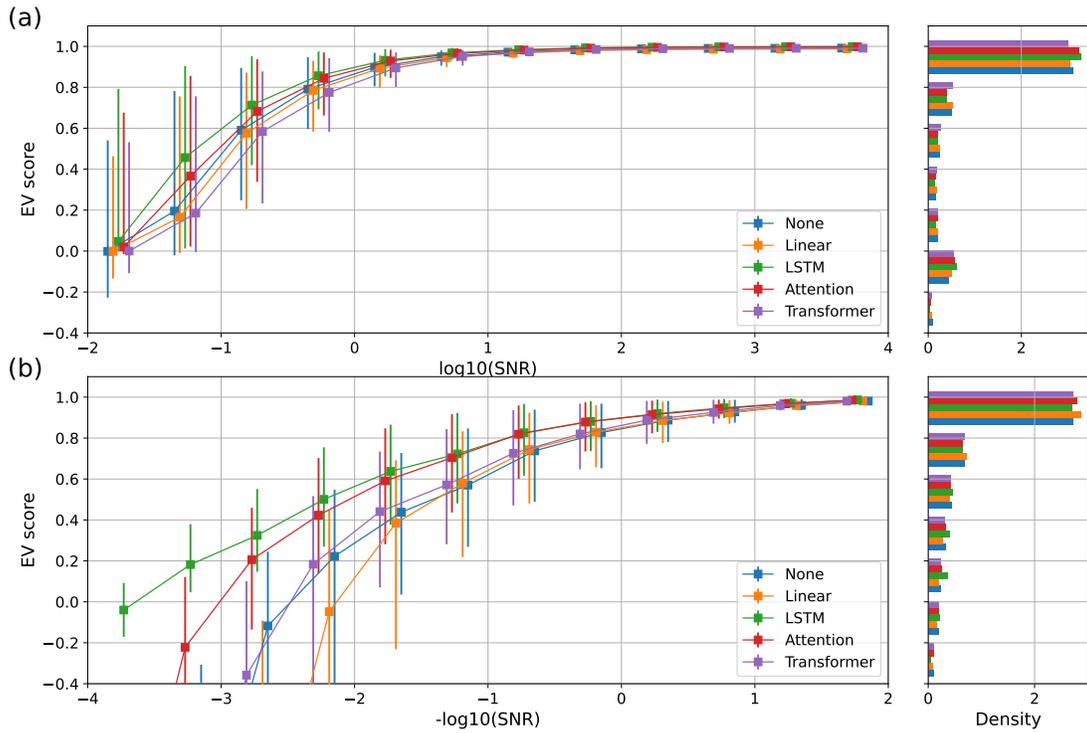
The overall training behaves properly for all models. During model training, the training loss curve keeps decreasing. The validation loss curve decreases and approaches the training loss curve as the training goes, and finally stays almost constant at some epoch, which indicates the convergence of optimization and the model is no longer improved. LSTM and attention models achieve the minimum final loss value for both training and validation data sets (Figs. 7.4 (c)-(d)). The transformer model, although with more model parameters than any other models, shows a higher validation loss value than that of the LSTM and attention bottlenecks (Fig.7.4 (e)). The minimal None and Linear bottlenecks exhibit the highest final loss values for both training and validation loss (Figs. 7.4 (a)-(b)). I also show the partial loss curves from individual branches of the earthquake and the noise waveforms, approximately half of the total loss. The validation loss for the earthquake waveforms is slightly higher than that of the earthquake loss probably due to the complexity of transient earthquake signals.

Next, I test the trained models with the 28,800 samples of the test data set. The test data set is not included during the training process, so it can be used to evaluate the model performance. The LSTM and Attention models have achieved the minimum mean test loss value of 0.0503 and 0.0557, respectively. The transformer model has a mean test loss value of 0.0836. The None and Linear models have mean test loss values of 0.0831 and 0.0893, respectively.

Furthermore, I inspect the waveform fitting for different models. For the same input  $\mathbf{X}$  (composite waveform), I obtain the predicted output/waveform  $\tilde{\mathbf{y}}$  and compare it with the ground truth



**Figure 7.4:** Training curves of networks with different bottlenecks: (a) None, (b) Linear; (c) LSTM; (d) Attention; (e) Transformer. Dots and solid lines indicate the loss from training data set and validation data set, respectively. Colors indicate different part of loss: total loss in black, earthquake waveform (decoder 1) in blue and noise waveform (decoder 2) in green. The red star indicates the loss from test data set.



**Figure 7.5:** Comparison between networks with different bottlenecks. (a) Variation of earthquake EV score with SNR of the noisy input signal. (b) Variation of noise EV score with SNR of the noisy input signal. Colored lines show the median EV score from networks with different bottlenecks. The error bars are calculated from the median values of EV score for test samples above and below the median EV score. Right panels show the histograms of the EV score of each network with the same color scheme.

waveform  $y$ . Other seismic denoising studies have reported improved SNR values as performance metrics<sup>299,370</sup>. Tibi et al.<sup>299</sup> also use the signal-to-distortion ratio (SDR) metric<sup>227</sup>, but the SDR metric is unbounded. Here, I calculate the EV score for both separated earthquake waveforms and ambient noise waveforms:  $EV = 1 - \frac{Var(y-\tilde{y})}{Var(y)}$ ,  $Var$  means variance of the time series. The best possible EV score is 1.0, corresponding to perfect waveform reconstruction. An EV score of 0.0 means that no waveform has been reconstructed ( $\tilde{y} = 0$ ). A negative EV score means a false waveform reconstruction, for example, in the time window where there is no earthquake waveform ( $y \approx 0$ ) but the network reconstructs a spurious waveform ( $\tilde{y} \neq 0$ ).

The comparative results are shown in Fig.7.5. All models can reconstruct both earthquake and noise waveforms with over half of the tested samples that achieve a high EV score around 1 (Fig.7.5). The network with the LSTM bottleneck recovers most test samples with a high EV score of around 1 for the earthquake waveforms (Fig.7.5 (a)). There is a bimodal distribution in the EV scores for earthquake waveforms. All models show two peaks around EV scores of 1 and 0, especially when the earthquake-to-noise ratio is less than 1. The bimodal pattern is suggestive of the behaviors of this encoder-decoder network. In cases that the network can recognize, the network can reconstruct the waveforms accurately. In cases that the network can hardly recognize, the network tends to output a time series of zeroes, which leads to almost 0 EV score. All networks show similar performance in EV score for the noise waveforms (Fig.7.5 (b)). I find no obvious bimodal pattern in the ambient noise reconstruction, which indicates a lower likelihood of outputting strictly zero noise. There are, however, spurious reconstructions of the noise waveforms for lower noise-to-earthquake amplitude ratios, or when transient signals dominate the time series (Fig.7.5 (b)).

Moreover, I explore how the EV score varies with SNR for both earthquake (Fig.7.5 (c)) and ambient noise waveforms (Fig.7.5 (d)). First, all models present the same pattern that the EV score monotonically increases with the corresponding amplitude of signals, quantified by SNR for earthquake waveforms or  $1/\text{SNR}$  for ambient noise waveforms, respectively. This is well expected because it is easier for the ML network to learn the intrinsic features of the signal waveforms and reconstruct the earthquake and ambient noise signals for higher amplitude. All models perform similarly. Take the case of earthquake waveform as an example (Fig.7.5 (a)). For  $\text{SNR} > 10^1$ , all models can correctly separate the earthquake waveform almost perfectly (EV score  $\sim 1$ ). The performance of any model drops as the SNR decreases. For example, at  $\text{SNR} = 1$ , the median EV score of models is about 0.8 to 0.9, and LSTM has the best performance. The discrepancy between model performance is exacerbated at low SNR. For instance, when the SNR is small,  $= 10^{-1} = 0.1$ , it is visually difficult to extract the earthquake signal. However, the LSTM model can still achieve a median EV

score of about 0.6, the Attention model has a median EV score of 0.5, the None model has a median EV score of 0.48, the Linear model has a median EV score of about 0.4, and the Transformer model has a median EV score of 0.3. For even smaller SNR values,  $= 10^{-1.8} = 0.02$ , all models tend to fail with most EV score being 0. The variance of the EV scores also increases with decreasing SNRs, indicating that there are more uncertainties in the reconstructed waveforms. Similar behaviors can be observed for the EV score of ambient noise part, with larger amplitudes of noise yields to better model performance (Fig.7.5 (b)).

One possible explanation for the different performances between bottlenecks architectures is the difference in model complexity. None and Linear models have fewer parameters than other models, so they may not be enough to understand the internal features of the seismograms properly. The None and Linear models have higher training loss (about 0.08) than other models, suggesting a not good model. The Linear model presents a larger loss value (Fig.7.4 (b)) and poorer waveform fitting (Fig.7.5) than the None model, implying the inability of linear regression as the bottleneck layer for this waveform decomposition problem. On the other hand, the Transformer model is more complex than the other models. Its mean test loss (0.0836) is slightly lower than that of the Linear model (0.0893), and the earthquake waveform fitting is almost the same as the Linear model (Fig.7.5 (a)) but the noise waveform fitting is similar to that of the None model (Fig.7.5 (b)). The LSTM and attention models share a similar overall complexity and achieve the lowest test loss and the most stable training. The similarity and systematically low values of the training and validation losses for the LSTM and the attention bottleneck may also indicate that those two models have already well “learned” the features in the training data sets (Fig.7.4 (c) and (d)).

To summarize, I evaluate the performance of models with different types of bottleneck models by testing the same test data set. I find that the model performance can differ due to model complexity. In general, the order of performance of my model is LSTM > attention > None > Linear  $\approx$  Transformer based on the variation of EV score with SNR. The LSTM bottleneck outperforms

other bottlenecks in reconstructing both earthquake and noise waveform, especially for a situation with low SNR (Fig.7.5 (c)). It is interesting to note that LSTM performs better than the attention and Transformer models, which implies that the sequential information is essential for reconstructing the waveforms. I speculate that the feature extraction of the encoder branch suffices at representing the temporal sequencing in the bottleneck layer. The conventional limitations of LSTM that long memory is not long enough are no longer important.

### 7.3 APPLICATION TO CONTINUOUS SEISMIC DATA

I now apply WaveDecompNet to continuous time series. It is straightforward to apply the model to any continuous data, provided that it has the same sampling rate. I select continuous recordings at IU.POHA from July 31, 2021, to September 1, 2021. I first down-sample the three-component 1-month-long waveforms to 10 Hz. Next, the typical pre-processing steps to apply machine-learning models is a) windowing to 1-minute long time series (600 samples) without overlap and b) applying the data normalization using the standard scaler. The most intuitive order to apply these processing steps are a), then b). I found that ordering a) then b) leads to spurious effects when concatenating back the 1-minute waveforms into a 1-month long waveform due to offset (means) and trends that rendered the application to continuous time series unpractical. Instead, I experimented with the order of b) then a) and found much better performance without artifacts when stitching back the waveforms.

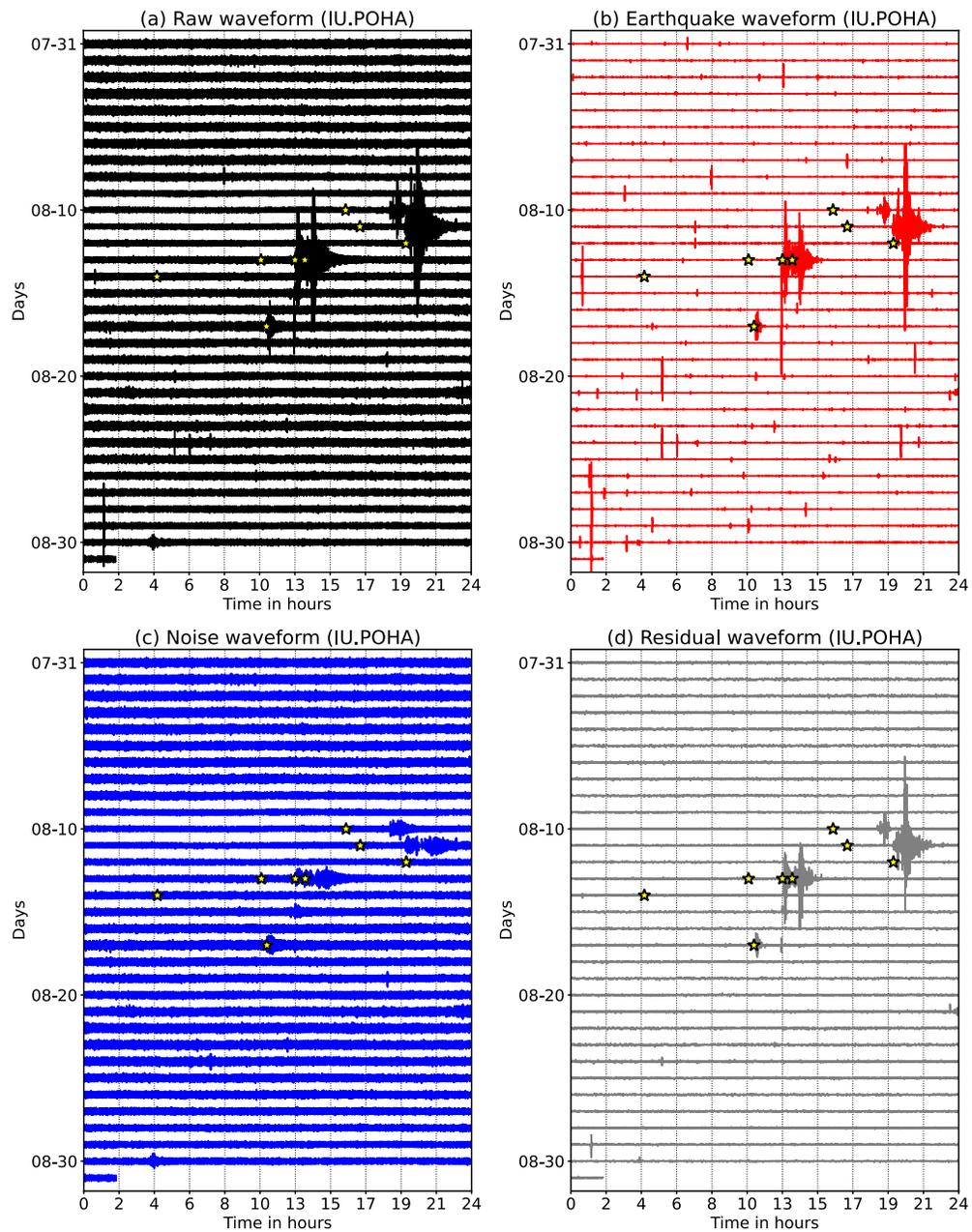
I normalize the entire month-long time series by removing its mean and scaling with its standard deviation (STD) to have the zero-mean, unit-variance time series. I slide through the data with 1-minute long windows (600 samples) without overlap. I apply the WaveDecompNet to all 1-minute long windows, concatenate all ML-filtered windows, and scale back the 1-month long time series with the standard deviation and mean for both the earthquake/transient time series and the noise

time series.

I show the results of separating transients and noise waveforms in Figure 7.6. Most of the transient signals have been well separated, with a significantly suppressed noise level (Fig.7.6 (b)). The residuals between the original waveform and the reconstructed waveforms are obtained from subtracting the earthquake and noise recovered waveform from the input waveform (Fig.7.6 (d)). Overall, the residuals are low. However, they are large between August 10, 2021, and August 20, 2021, and these are due to the teleseismic ( $30^\circ$  to  $100^\circ$  angular distance) earthquakes. I mark the P arrival, calculated from TauP with IASP91 Earth model, of these large  $M_{5.5+}$  teleseismic earthquakes to illustrate that in Figure 7.6. No teleseismic waveform was used during the training, in part because the input data length is restricted to one minute. Therefore, my model does not handle longer seismic periods at this stage, and the coda reconstruction of these long waveforms is imperfect. Nevertheless, the general envelop pattern of those earthquake waveforms can still be recovered. In the following section, I test the validity and usefulness of these transient waveforms by applying a standard impulsivity filter most commonly used detection method in seismology.

The separated noise waveforms exhibit more leveled, constant amplitudes throughout the month (Fig.7.6 (c)). Some transient signals remain, especially in the coda of teleseismic earthquakes. For additional evaluation of the usefulness of this network, I apply the single-station correlation functions used in ambient-seismic noise monitoring in a later section.

Most of the previous denoising networks, such as the DeepDenoiser<sup>371</sup>, construct the noise time series from direct subtraction of the “denoised” earthquake waveforms from the raw data. Unlike the DeepDenoiser, the WaveDecompNet has two branches that learn features of the earthquake and noise waveforms, somewhat independently since their only connection is through a residual connection to the encoder branch. Because the noise window is not the linear difference between the original and the transient/earthquake signal, I also investigate the waveform residuals and show them in Figure 7.6 (d). In general, the amplitudes of the residual waveforms are small (about 10



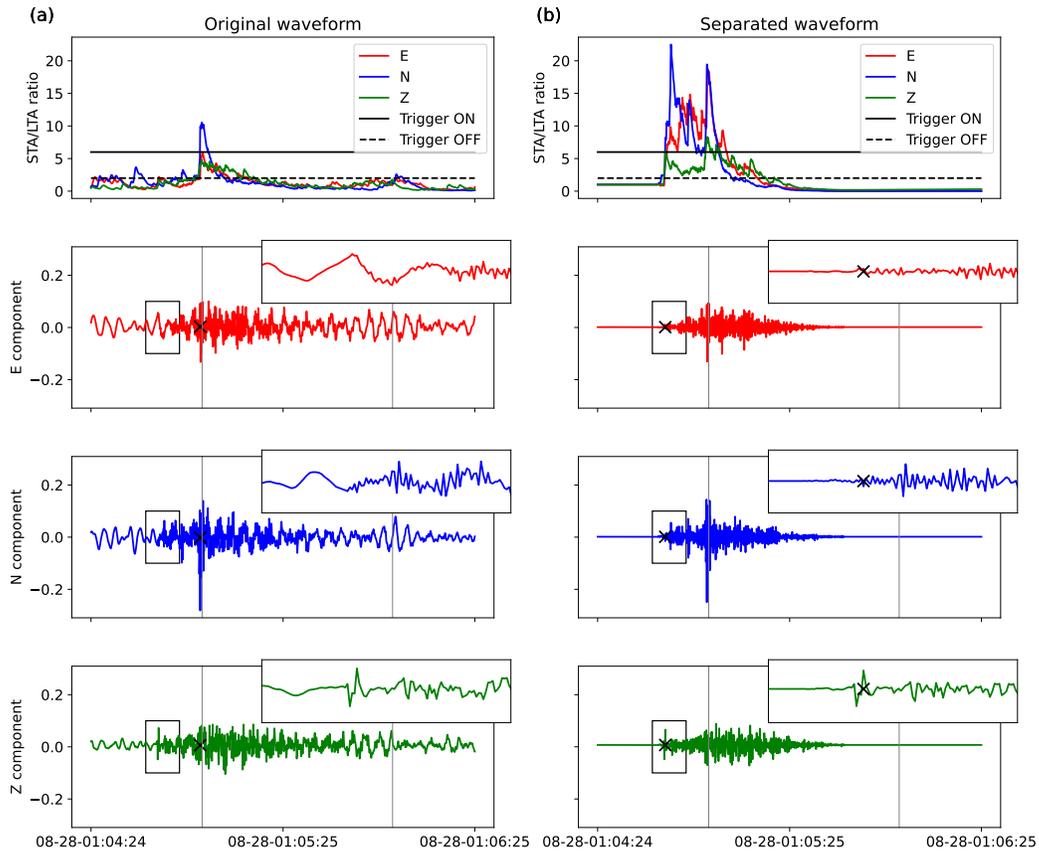
**Figure 7.6:** Application to continuous seismic data from an island station IU.POHA. (a) One-month raw waveform from IU.POHA; (b) Separated earthquake waveform; (c) Separated noise waveform; (d) Waveform residuals from subtracting the separated earthquake and noise waveforms from the raw waveform. The yellow stars label the P wave arrivals, TauP calculates with IASP91 velocity model<sup>150</sup>, of large earthquakes (M5.5+) between August 10, 2021, to August 20, 2021, from the International Seismological Centre catalog provided by default by Obspy and IRIS FDSN event server.

times lower than the ambient noise in the standard deviation of waveform amplitude, Fig.7.6 (c)) except for some large residuals during large teleseismic events. Including more teleseismic earthquake waveforms in the model training can potentially help to mitigate those large residuals, and I leave as a future direction to explore.

### 7.3.1 APPLICATION TO DETECTING EARTHQUAKES USING STA/LTA TRIGGER

I apply a recursive short-term-average (STA) to long-term-average (LTA) trigger method (STA/LTA) to the continuous data<sup>9,313,343</sup>. This particular STA/LTA algorithm produces a decaying exponential impulse response, and that is sharper impulse than the original STA/LTA algorithm<sup>343</sup>. The settings of the STA/LTA parameters are chosen from Trnkoczy<sup>303</sup> and also from trial-and-error tests. The short time window length is set at 2.0 s, the long time window length is set to 60.0 s, the on-threshold is 6.0, and the off-threshold is 2.0. Because I simultaneously run STA/LTA detection on 3-component waveforms, I perform a coincidence trigger with a threshold of 2, which means that a detection trigger occurs when the STA/LTA ratios of any of the two components exceed the on-threshold.

The STA/LTA time series are a lot cleaner in the separated earthquake waveforms than in the original seismograms (Fig.7.7), which is manifested in two aspects. First, the increased signal-to-noise ratio of the separated earthquake waveforms improves the accuracy of the detection time automated by STA/LTA triggers. For the example shown in Fig.7.7, the arrival time cannot be correctly picked in the raw data using automated STA/LTA thresholding detector (Fig.7.7 (a)). Nevertheless, with the noise separated by WaveDecompNet, the event arrival can be easily detected, and the accuracy of picking the first arrivals can be improved by about 15 seconds (Fig.7.7 (b)). Second, I can detect many smaller signals (either smaller magnitude or more distant events) buried in the noise, which is suggested by the increased number of coincidence triggers in the separated waveform, from 38 in the original time series to 1031 in the separated waveforms. The wavefield separation increases



**Figure 7.7:** Example of STA/LTA detection algorithm: (a) original waveform, (b) separated earthquake waveform. (Top panels) Recursive STA/LTA ratio from the waveform in the chosen window. Black solid and dashed lines indicate the trigger thresholds on and off, respectively. (bottom panels) Red, blue and green lines indicate E, N, Z components. Black crosses show the picks from STA/LTA. Gray vertical bars indicate the edges of the 1-minute time windows when applying WaveDecompNet. The inset figures show the zoom-in waveforms within the boxes.

the detection by a factor of 68 in the current STA/LTA settings. This ratio varies with the choice of threshold, from 4 (1899 events vs. 478 events) for a coincidence threshold of 1 and 7 (147 events vs. 19 events) for a coincidence threshold of 3. Tuning the parameters of this detector is not the scope of this study but would be necessary in the deployment of this algorithm in specific cases.

### 7.3.2 APPLICATION TO AMBIENT NOISE MONITORING USING SINGLE-STATION CROSS-CORRELATIONS

Single-station correlations are related to the zero-offset Green's function<sup>61,62,83,272</sup>. Monitoring phase changes in the single-station measurements have enabled the monitoring of changes in the near-surface environment that occur during earthquakes<sup>316,336</sup>, volcanic unrest<sup>73</sup>, and to monitor shallow hydrology<sup>136</sup>. Here, I do not attempt to verify that the single-station correlation is proportional to the Green's function. Instead, I evaluate the temporal stability of the single-station cross-correlations.

I calculate all 9 components of the correlation tensor. I select 1-minute long windows, pad them with zeroes from 600 samples to 2048 samples (204.8 s). I then follow the spectral method from Viens et al.<sup>317</sup> to calculate the ambient noise single-station correlation function (ACF):

$$ACF_{ij}(t) = F^{-1}\left(\frac{\hat{a}_i \hat{a}_j^*}{|\hat{a}_i| |\hat{a}_j|}\right), \quad (7.2)$$

where  $i, j$  corresponds to components (E, N, and Z),  $\hat{a}_i$  is the Fourier transform of the  $i$ -component waveform, \* represents the complex conjugate,  $F^{-1}(\cdot)$  is the inverse Fourier transform. I whiten the amplitude spectrum using a running mean as in conventional processing<sup>30</sup>  $|\cdot|$  of 32 samples in the frequency domain.

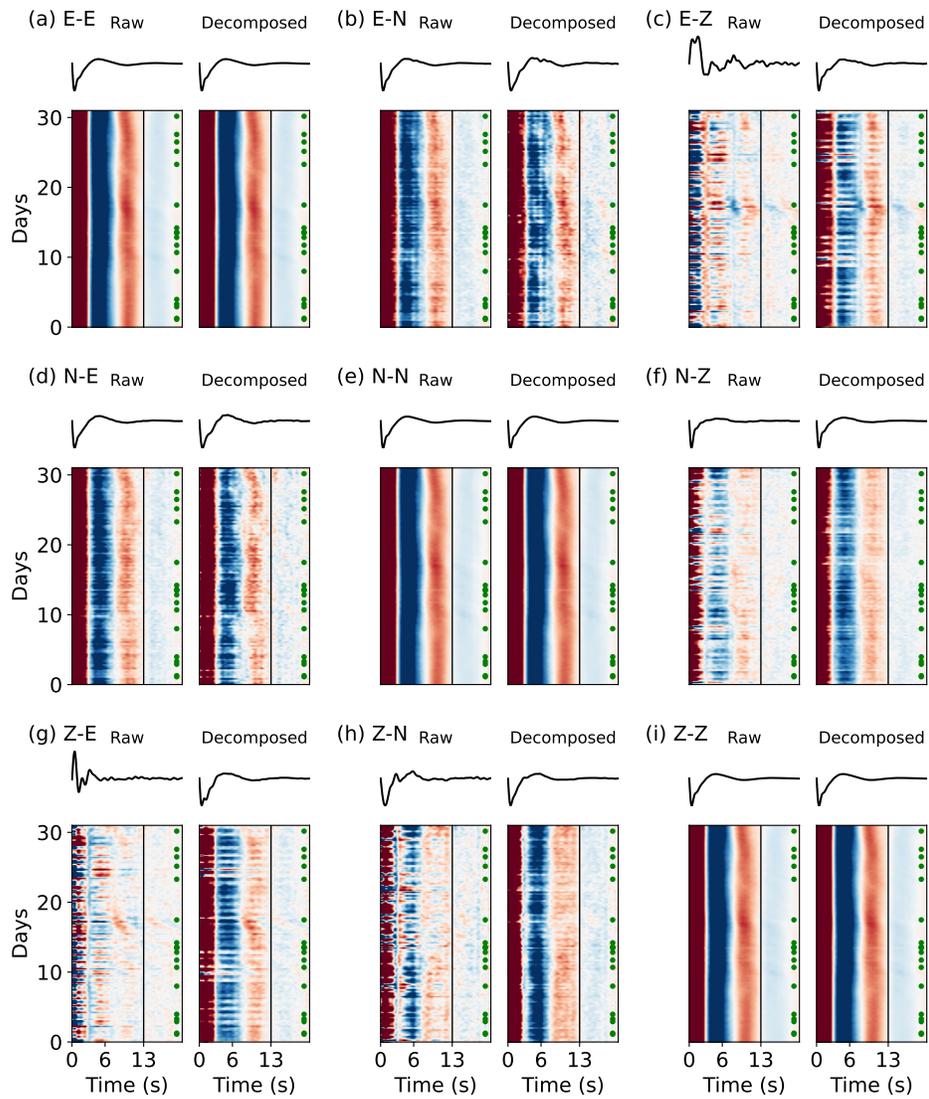
I sub-stack the correlations functions every 6 hours to evaluate their stability through time. I show the causal part (positive lag) of correlation functions in different frequency bands: Low Fre-

quency (LF) 0.1 - 1.0 Hz (Fig.7.8); Medium Frequency (MF) 1.0 - 2.0 Hz (Fig.7.9); High Frequency (HF) 2.0 - 4.0 Hz (Fig.7.10). Each figure shows the single-station correlations from the original raw waveforms and ones obtained from the separated waveform.

I also stack all of the correlations to form a reference stack, from July 31, 2021, to September 1, 2021. I calculate the correlation coefficient between each 6-hour stack and the month-stack reference waveform and show them in Figure 7.11. I use the stability of cross-correlation as a success metric of ambient seismic noise recovery.

Because of transient earthquake signals in the raw waveforms, I can see large fluctuations (mostly reduced amplitudes) in the correlation functions at all frequency bands, especially for the E-Z, N-Z, Z-E, Z-N functions (Figs.7.8 - 7.10). These fluctuations in the correlation functions and drops in their coherence arise from transient signals in the original time series. However, many of these fluctuations disappear when using the separated noise signals to calculate the correlation functions. Furthermore, some of the coda phases that are weak in the original correlation functions appear clearly in the correlation functions built from the separated noise. These coda phases potentially correspond to seismic wavespeed interfaces or discontinuity beneath the seismic station. With the transient earthquake signals removed, WaveDecompNet can help constrain the velocity structure underneath the seismic station. Additional work remains to be done to verify the nature of these coda phases and whether they can be related to Earth structure.

As expected, the improvement on the correlation functions coherence is substantial (Fig.7.11). As shown in Figures 7.8 - 7.10, the transient earthquake signals can break the coherence among correlation functions, and lead to low correlation coefficients between each function and the reference (see Fig.7.11). On the other hand, the correlation coefficients from separated noise mostly have stable values closer to 1, confirming the enhanced coherence of the cross-correlation functions from continuous ambient noise data. I find that the coherence from separated noise drops in some time windows (for example, day 1 - day 7 in Fig.7.11 (a) and day 24 - day 30 in Fig.7.11 (b)). This can



**Figure 7.8:** Single-station cross-correlation and auto-correlation functions filtered in the LF low frequency band (0.1 - 1.0 Hz) for the original raw waveforms (a) and the separated noise waveforms (b). Green dots show the P wave arrivals of M5.5+ teleseismic earthquakes.

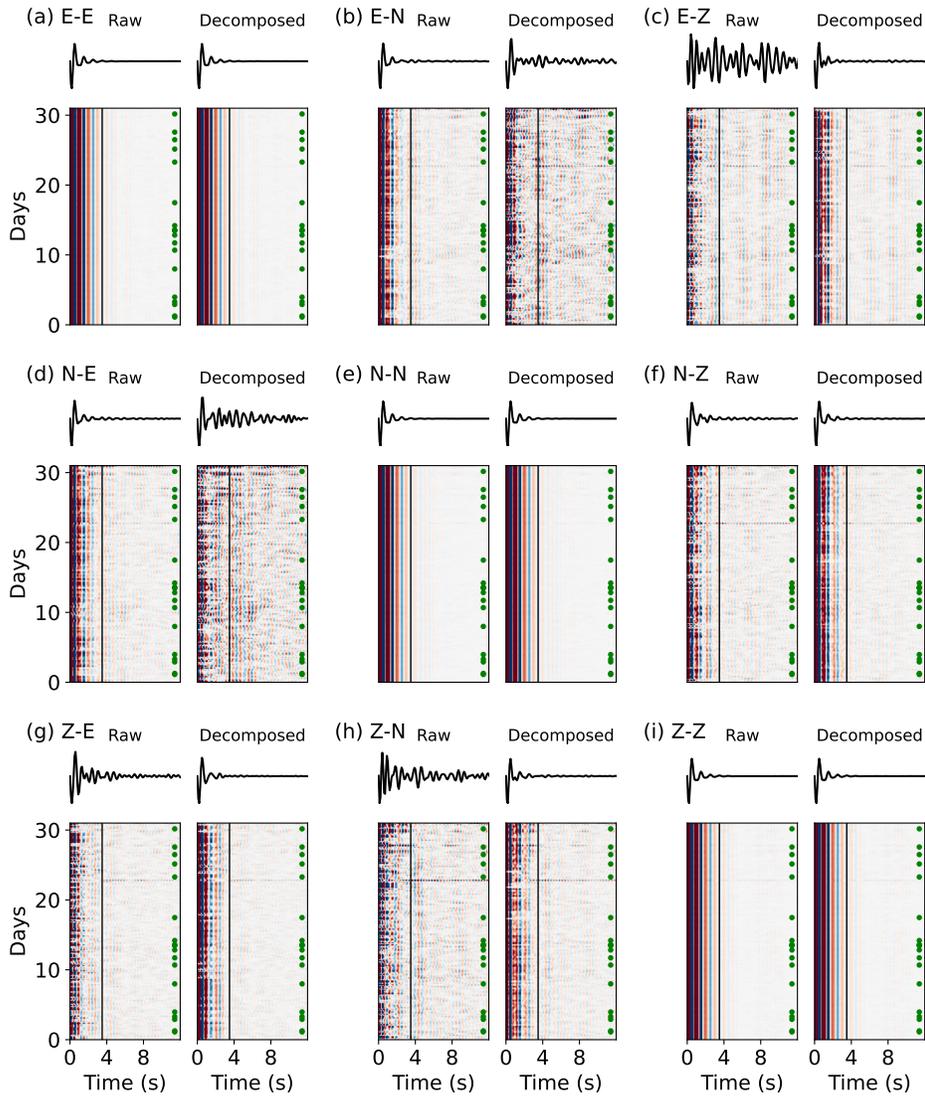


Figure 7.9: Same as Figure 7.8 for the MF medium frequency band 1.0-2.0 Hz.

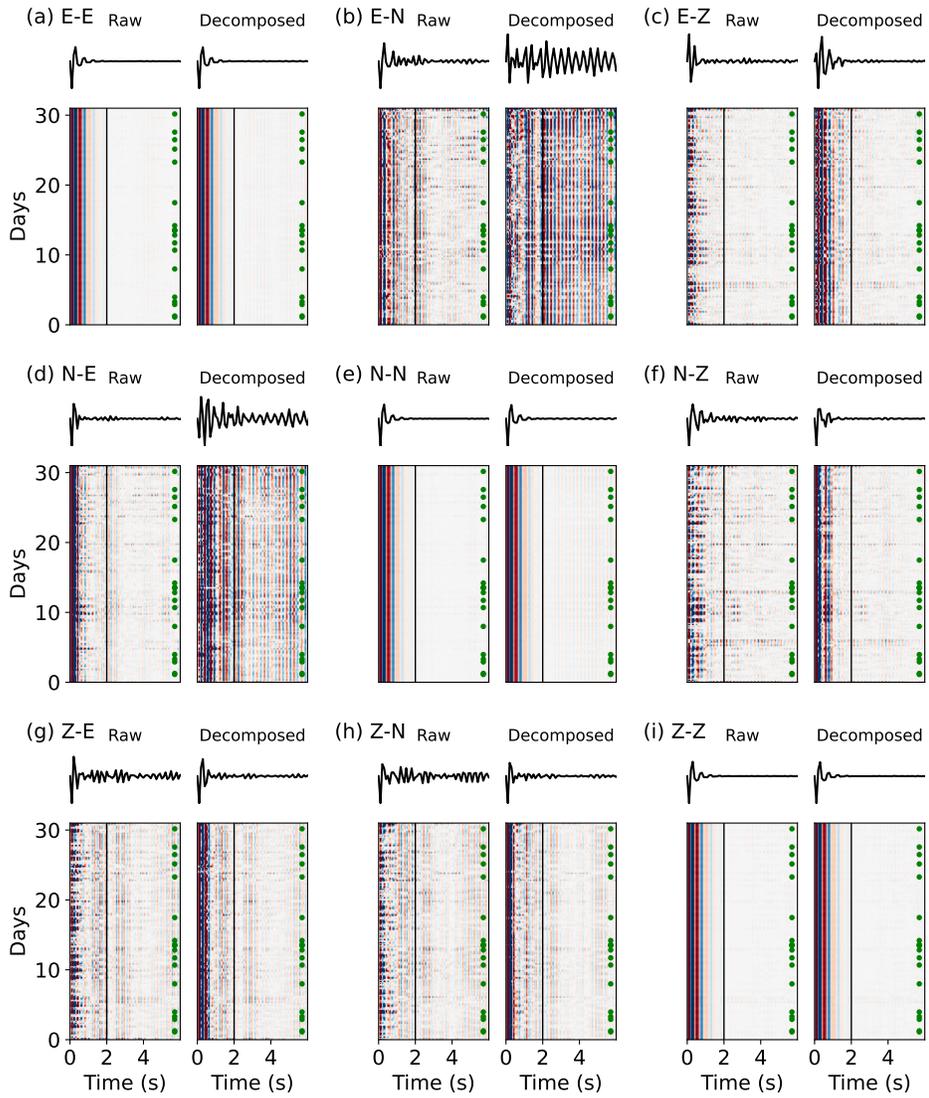
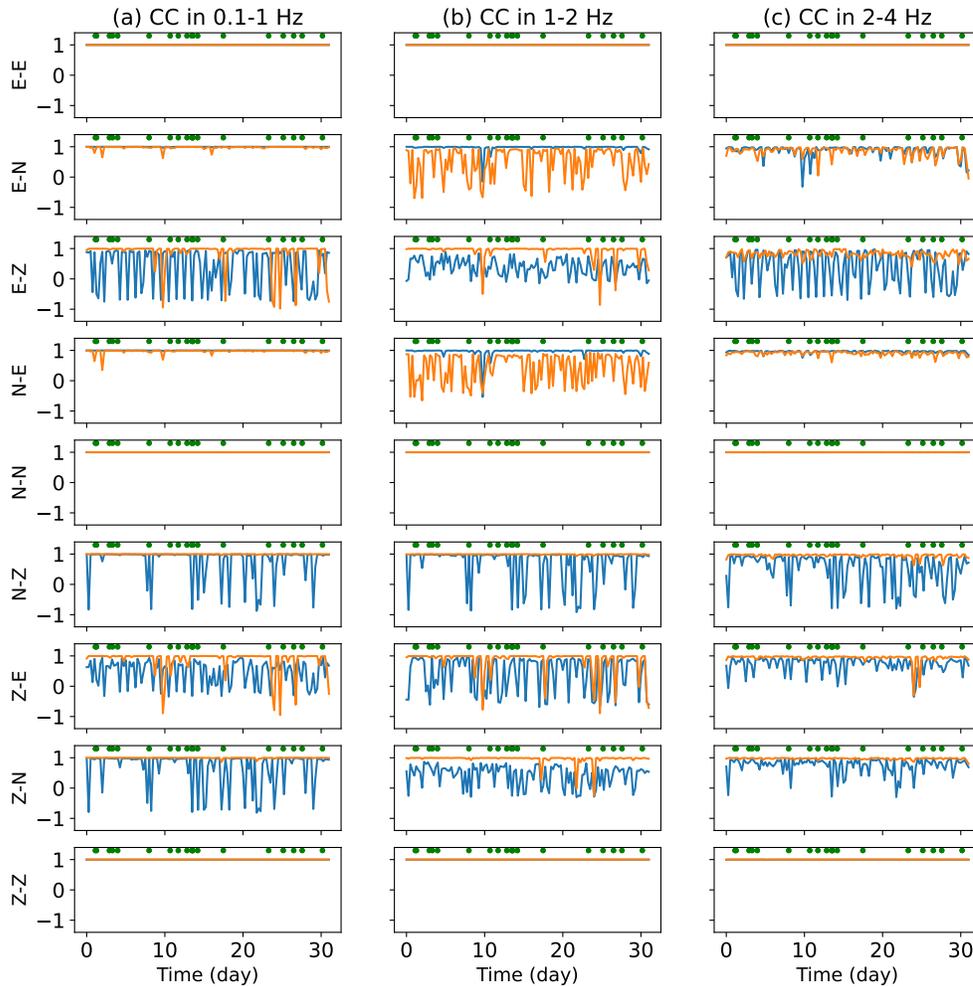


Figure 7.10: Same as Figure 7.8 for the HF high frequency band 2.0-4.0 Hz.



**Figure 7.11:** Coherence of the single-station cross-correlation and auto-correlation functions at different frequency bands (a) 0.1 - 1.0 Hz; (b) 1.0 - 2.0 Hz; (c) 2.0 - 4.0 Hz. Blue lines indicate the coherence from the original waveforms, orange lines indicate the coherence from separated ambient noise waveform. The coherence is quantified by the correlation coefficients between each 6-hour averaged correlation function and the 1-month averaged reference correlation function. Green dots show the P wave arrivals of M5.5+ earthquakes in the month calculated using TauP in the IASP91 model.

be possibly attributed to the poor reconstruction of the ambient noise signals. I also notice that for some components, the separated noise does not improve the coherence at all, for example, the N-Z component in MF (Fig.7.11 (b)) and Z-E component in HF (Fig.7.11 (c)). Further study is needed to understand these less-dominant issues in ambient noise applications.

#### 7.4 CONCLUSION AND DISCUSSION

I develop a machine-learning-based model, WaveDecompNet, to separate earthquake and ambient noise signals from raw seismic data. I combine the STEAD and local ambient noise to form a sufficient overall data set to train and test WaveDecompNet. My network consists of three parts: one encoder branch, two decoder branches, and two bottlenecks. I systematically explore the performance of models using different types of bottlenecks, and I find the network using LSTM bottleneck has the best performance. Next, I test how well my network can be applied to observed continuous data. I apply the trained model directly to a 1-month continuous seismic data at IU.POHA and successfully separate the corresponding earthquake and noise signals, except for the long-duration teleseismic signals. Next, I apply an automated transient detector (STA/LTA) and an established ambient-noise seismology monitoring method to the separated earthquake and noise signals, respectively. My results show that the quality of both separated earthquake and noise signals has been improved significantly. With the same ML filter, I can obtain more STA/LTA triggers and a highly coherent ambient-noise correlation function.

However, there are some limitations to my current method. First, it only includes waveforms from local earthquakes ( $< 350$  km). The lack of teleseismic waveforms, especially those from large earthquakes, leads to the poor performance of WaveDecompNet when handling the time windows with teleseismic earthquake waveforms. While I extract the general patterns of the teleseismic earthquake waveforms correctly, there remain large residuals in the ambient noise waveforms and residual

waveforms. Second, I only include the local noise from a single island station, IU.POHA. I also test with other stations from the Hawaiian Volcano Observatory and find the network trained from IU.POHA can still successfully separate the earthquake, but the coherence of ambient noise worsens, and there are more waveform residuals. Therefore, a good direction to improve the network performance is to include additional and different types of data. For example, I can include teleseismic data for better separation of earthquake waveforms, and I should also include the ambient noise waveforms from other stations and regions for a specific regional or global ambient noise study.

Future developments may involve the integration of multiple stations. The combination of multiple stations to combine the automated triggered events help reduce the false (non-tectonic) detections. It also helps locate the event and build a more complete earthquake catalog. Furthermore, a modification of the network to add more stations may help improve the stability of the inter-station cross-correlations, which in turn can be used for better Earth imaging.

## 7.5 DATA AND RESOURCES

The continuous seismic data from IU.POHA (IU: [doi:10.7914/SN/IU](https://doi.org/10.7914/SN/IU)) are downloaded using Obspy (available at <https://github.com/obspy/obspy/wiki>). PyTorch machine learning framework (<https://pytorch.org>) is to build and train the network. The module of self-attention bottleneck is based on Chapter 10.5 of the online book “Dive into Deep Learning” (available at <https://d2l.ai/index.html>). All the codes to reproduce this work are hosted on Github at <https://github.com/yinjiuxun/WaveDecompNet-paper>, WaveDecompNet is hosted on <https://github.com/yinjiuxun/WaveDecompNet>.



## Conclusion

During my Ph.D., I worked on several different projects on earthquake seismology with my advisor Marine Denolle and other collaborators. I contributed by developing various tools and solving a few long-existing scientific questions in earthquake seismology. I hope my Ph.D. research can have a significant long-term impact on the earthquake science community.

In Chapter 2 and Chapter 3, I developed innovative observational methods to constrain the earthquake kinematics and dynamics better. My methodologies are an attempt to bridge kinematic

observations with earthquake dynamics. I applied my observational method to the 2015 Illapel, Chile  $M_w$ 8.3 earthquake as a specific case study. I found a distinct evolution of the low-frequency radiation compared to the high-frequency radiation. Low-frequency radiation is not surprisingly more prominent in large earthquakes and appears consistently shallower than high-frequency radiation. This is called depth-frequency relation and has been observed during many large megathrust earthquakes. I also reviewed a few possible mechanisms, either from pre-stress, fault geometry, and/or friction, that can explain the depth-frequency relation.

In Chapter 4, I showed the physical interpretation of back-projection (BP) images through theoretical study. The Back-projection method is widely used to recover the rupture image of a large earthquake. I built a theoretical formulation of the linear BP algorithm to show that the BP image is indeed related to the slip motion on the fault, granted a spatial smoothing. Furthermore, I proposed a resolvability parameter, which can help quantify the resolution of BP methods better. Finally, I construct a relation between resolvable area and seismic frequencies. Given the scaling of earthquake size with source length, my analysis provides simple guidelines to the lower bounds of seismic frequencies required to image details of the source provided earthquake magnitude.

In Chapter 5, I managed to explain the ubiquitous depth-frequency relation observed during most megathrust earthquakes. Global databases of BP images show a systematic depth variation of the frequency content in source radiation. It is a recurring pattern among most moderate-to-large subduction zone earthquakes. I find that the inclusion of Earth's free surface is sufficient to explain this ubiquitous observation through dynamic rupture simulation. I propose that the dynamics of shallow rupture are dominated by free-surface effects that are, in turn, the first-order factor in explaining the depth-frequency relation. The second-order effect is the evolution of earthquake rupture in a realistic velocity structure. The presence of anomalously low  $V_S$ , relative to  $V_P$ , also impacts the rupture behavior that further enhances seismic radiation's depth-dependence.

Except for the detailed rupture process of the largest megathrust earthquakes, we further studied

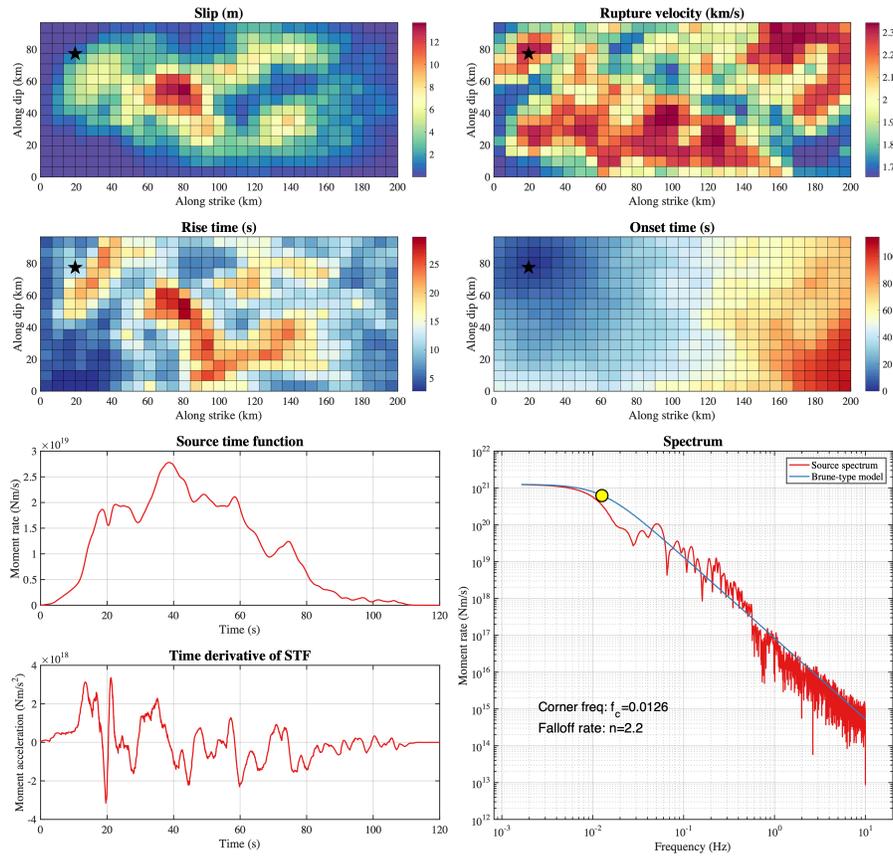
the earthquake source time functions (STFs) in Chapter 6. We developed different metrics to extract physical information from STFs in the SCARDEC database, including Gaussian-subevent decomposition (GSD) and dynamic time warping (DTW). We apply GSD to analyze earthquake STFs and their subevents. We find that large earthquakes have more subevents than small earthquakes, and the subevent moment scales with the main event moment with a power exponent of about 0.8. We also apply DTW to cluster many earthquake STFs into different complexity groups based on the similarity of their general shapes. We find the patterns of STF shape complexity correlate with different source parameters such as depth, duration, focal mechanism, scaled energy. Through dynamic rupture simulation, we show that simulations with spatially uniform fault parameters fail to produce subevents. Introducing a range of heterogeneity levels in the pre-stress yields similar results. We also find that the frictional parameter  $D_c$  affects the proportions of different complexity groups. Comparing the diversity in the STF complexity from the SCARDEC database to that from the simulated STF, we suggest that small values of  $D_c$  are viable explanations for the distribution between simple and complex events.

In the last Chapter 7, I develop a new auto-encoder network to separate transient earthquake signals from ambient noise signals directly in the time domain for 3-component seismograms. I explore different network architectures and find that the long-short-term-memory (LSTM) bottleneck outperforms other bottleneck structures. Finally, I evaluate the performance of my network in the actual application by applying the trained network to separate earthquake and noise signals for the continuous seismic raw data at an island station IU.POHA on Hawaii. I use STA/LTA and autocorrelation to the separated earthquake and noise signals, respectively. My results show that the quality of both separated earthquake and noise signals has been improved significantly. The promising results show that my developed encoder-decoder network for the separation of earthquake and noise signals can substantially help to improve the quality of seismic data, especially for those stations installed in noisy environments such as ocean islands or ocean bottom.

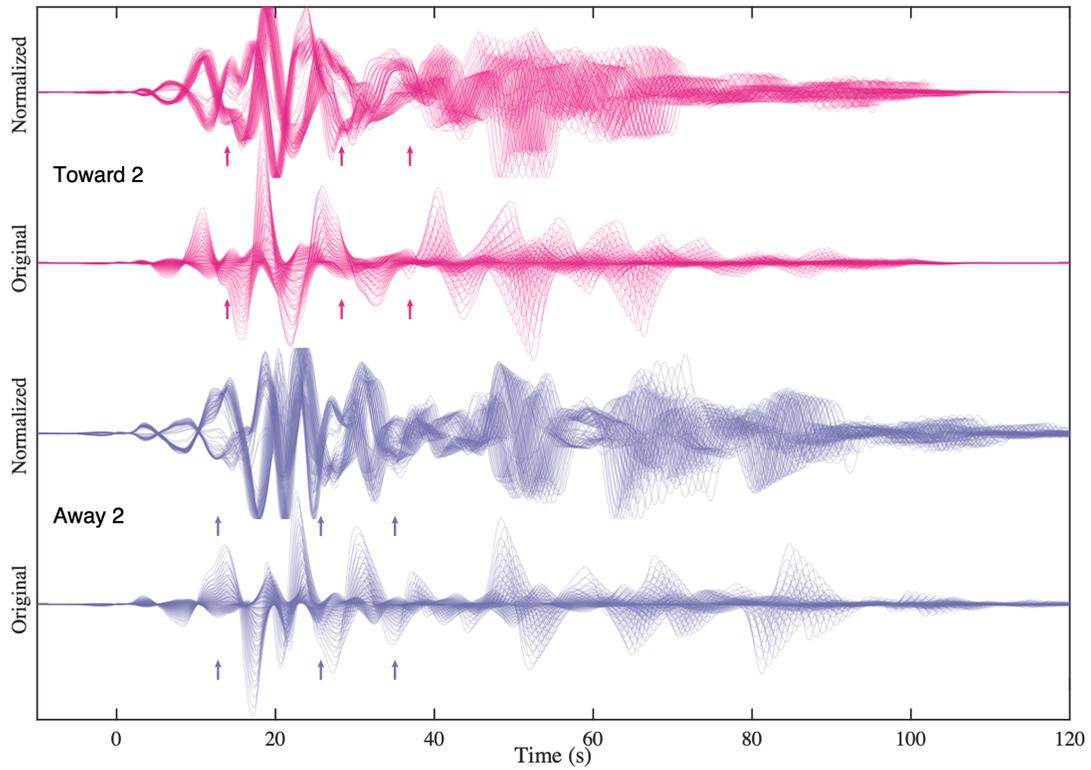
A

Supplementary materials for Chapter 4,  
“Relating teleseismic backprojection images  
to earthquake kinematics”

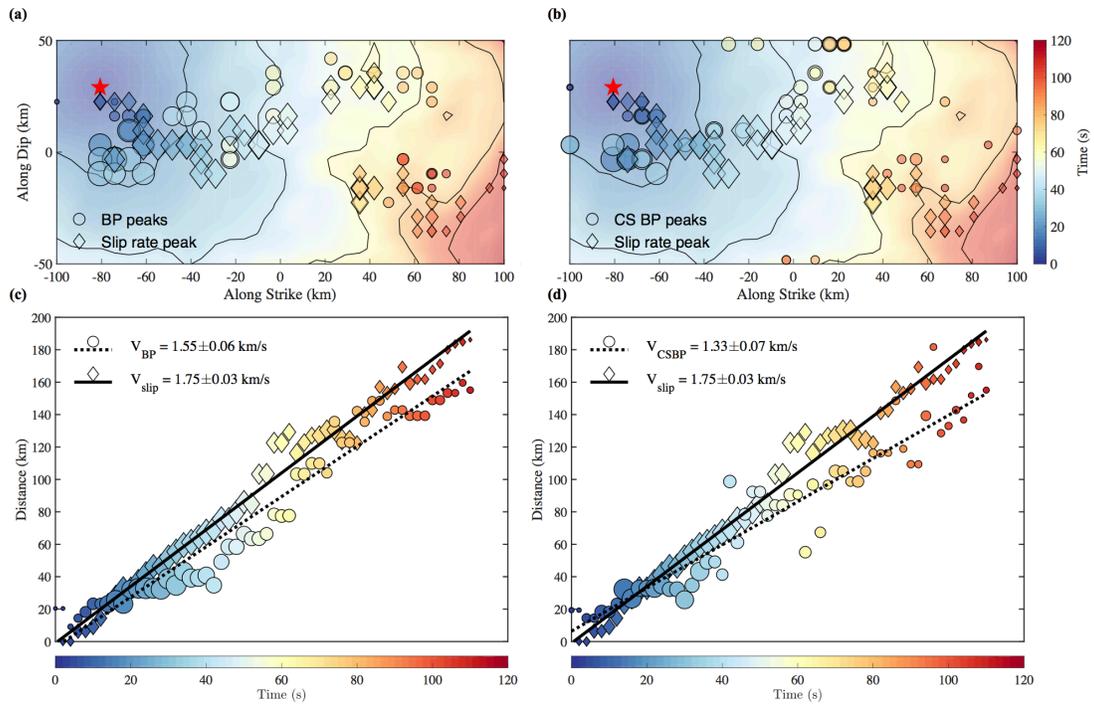
In this Appendix, the supplementary figures for Chapter 4, “Relating teleseismic backprojection images to earthquake kinematics”, are provided.



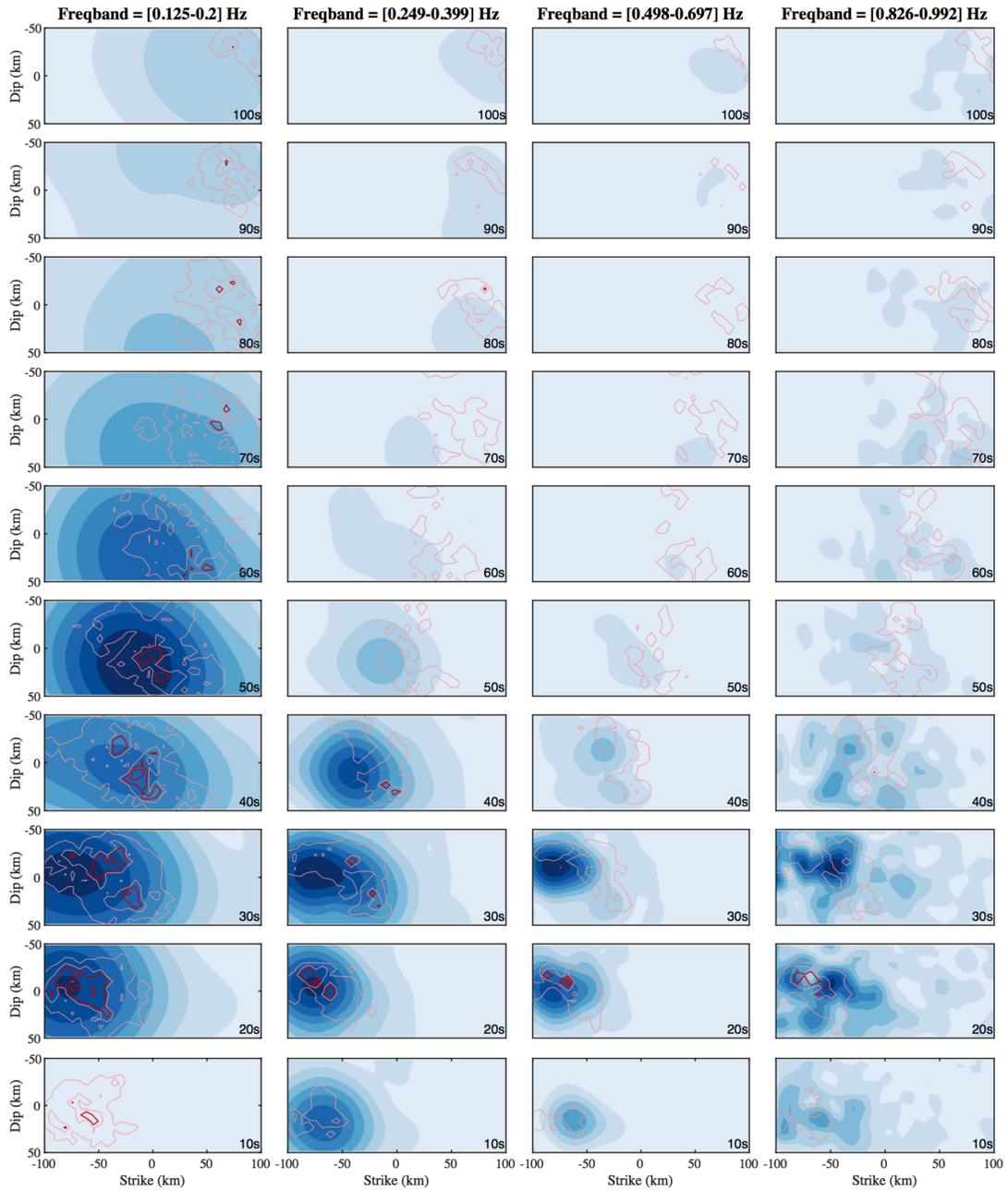
**Figure A.1:** One example of the kinematic source used in this study and its relevant source parameters: (a) coseismic slip distribution; (b) rupture velocity distribution; (c) rise time distribution; (d) Onset time distribution; (e) moment rate function (source time function); (f) moment acceleration (time derivative of the source time function) and (g) source spectrum and the best-fit Brune type model.



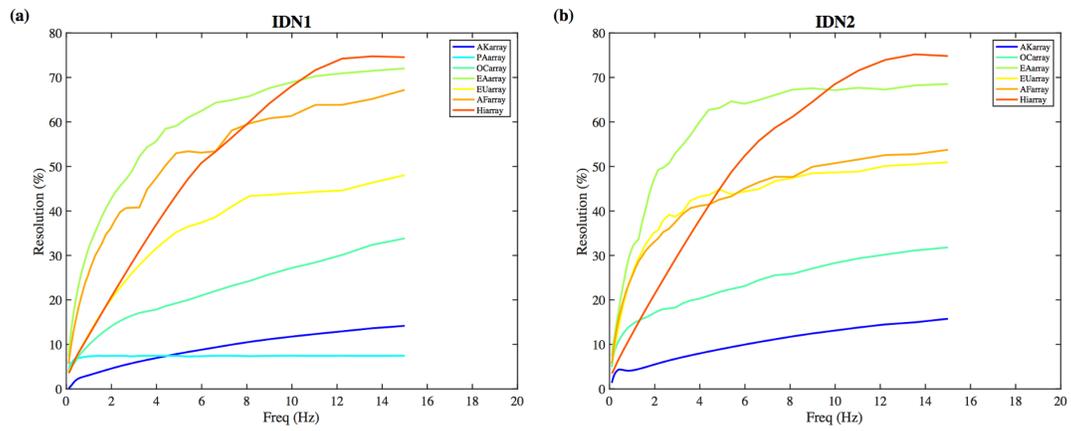
**Figure A.2:** Waveforms recorded by the arrays near the nodal plane of focal mechanism (Toward2 and Away2 arrays). For each array, both the original waveforms (same as Fig.2 (b) in the main text) and the waveforms after amplitude normalization by their maxima are shown here. The normalized waveforms here more clearly indicate the secondary polarity flipping caused by rupture propagation, which is also highlighted by the arrows.



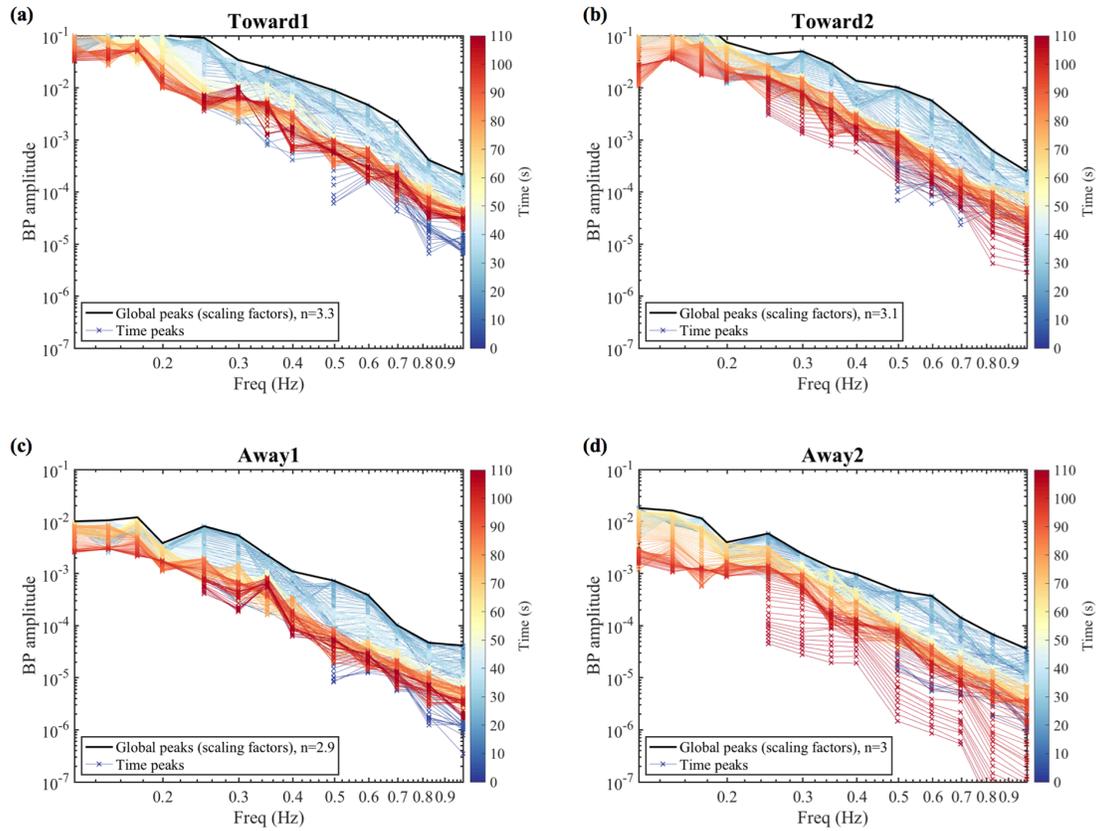
**Figure A.3:** (a) Spatiotemporal evolution of Linear BP peaks (circles) from Away1 array compared with slip rate peaks (diamonds) and onset time distribution (background image) from the kinematic model. Color corresponds to both the peak time and onset time. Red star indicates the epicenter of the kinematic model. (b) same as (a) but for the CSBP peaks. (c) Distance from the BP peaks to the epicenter varying with time. The black bold line shows the linear fitting for the slip peaks while the black dashed line corresponds to the linear fitting for the BP peaks. Color of symbols and x-axis both indicate the peak time. (d) Same as (c) but for the CSBP peaks.



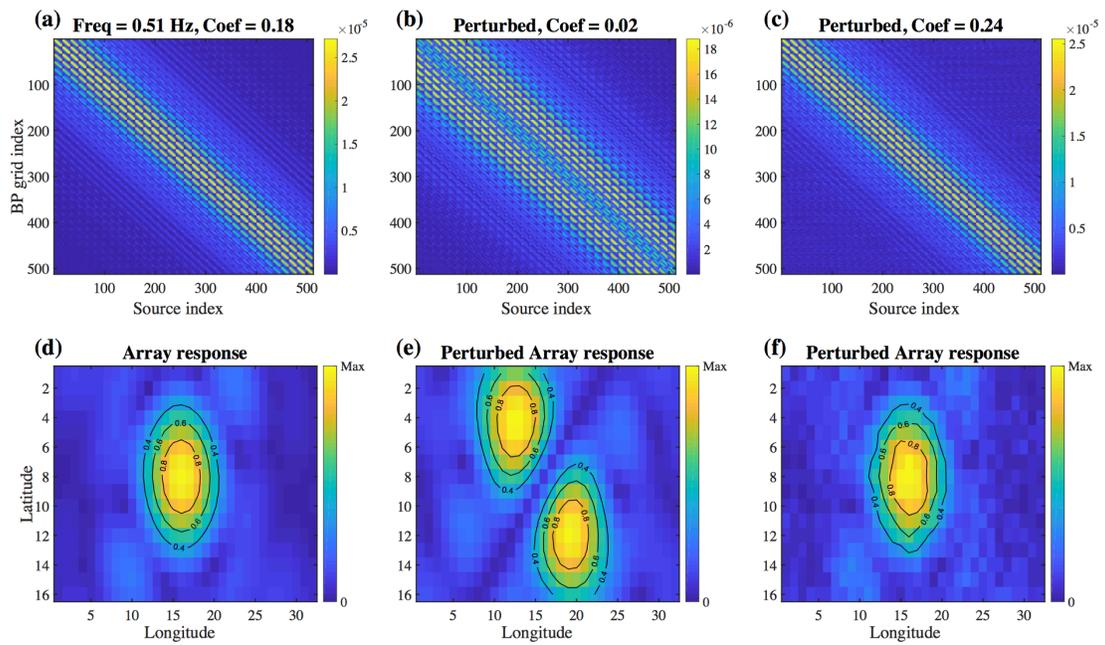
**Figure A.4:** BP image (blue images) in each sub-frequency-band from Away1 array compared with the corresponding filtered slip rate distribution (red thick contours correspond to 10% of maximum amplitude while pink thin contours correspond to 1% of maximum amplitude) from the same kinematic source shown in Fig.3.



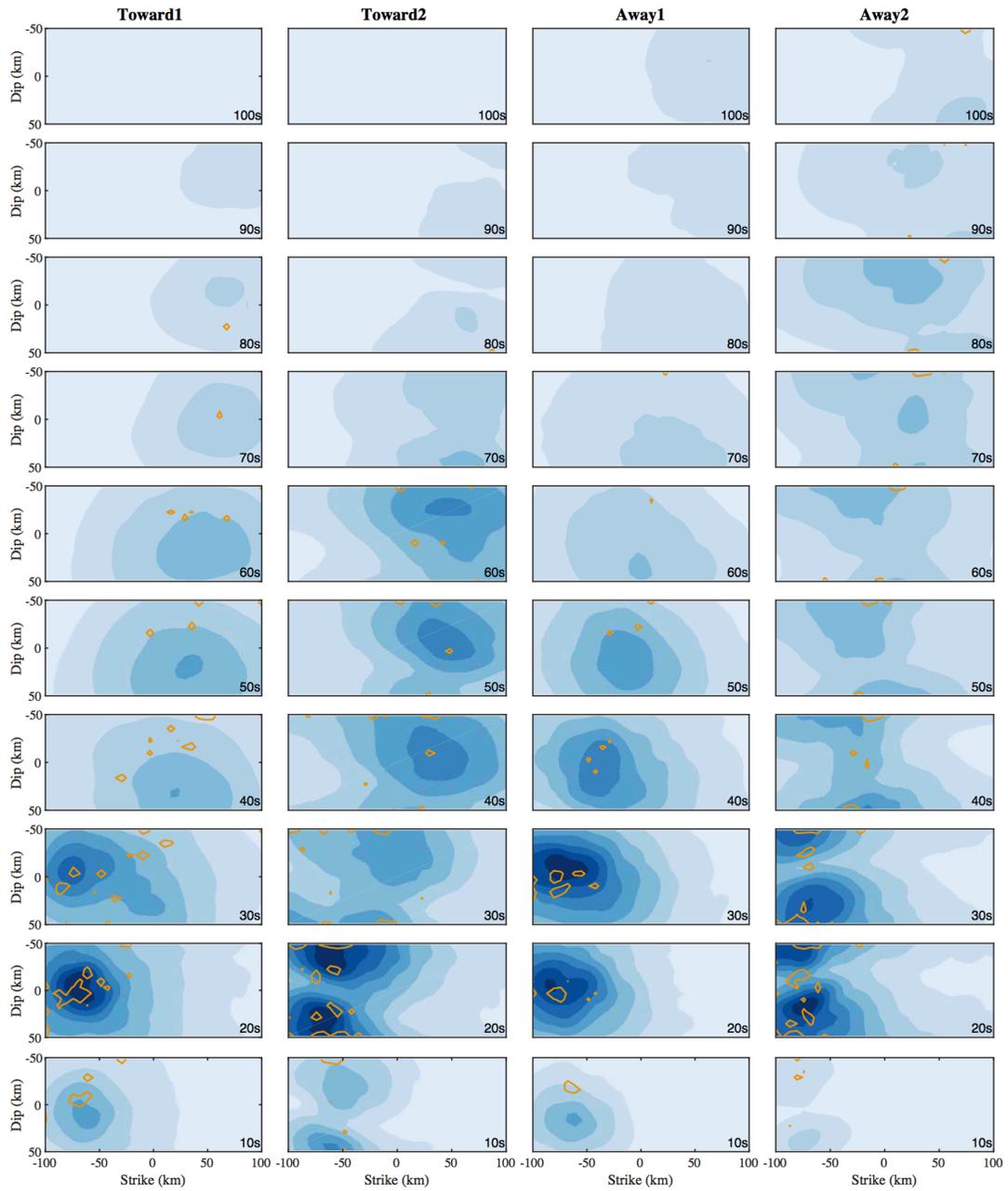
**Figure A.5:** The variation of resolvability  $\epsilon_I$  for all the available arrays within the teleseismic distance range to (a) IDN1 (Java) and (b) IDN2 (Sumatra) regions in a broad frequency band from 0.1 to 15 Hz.



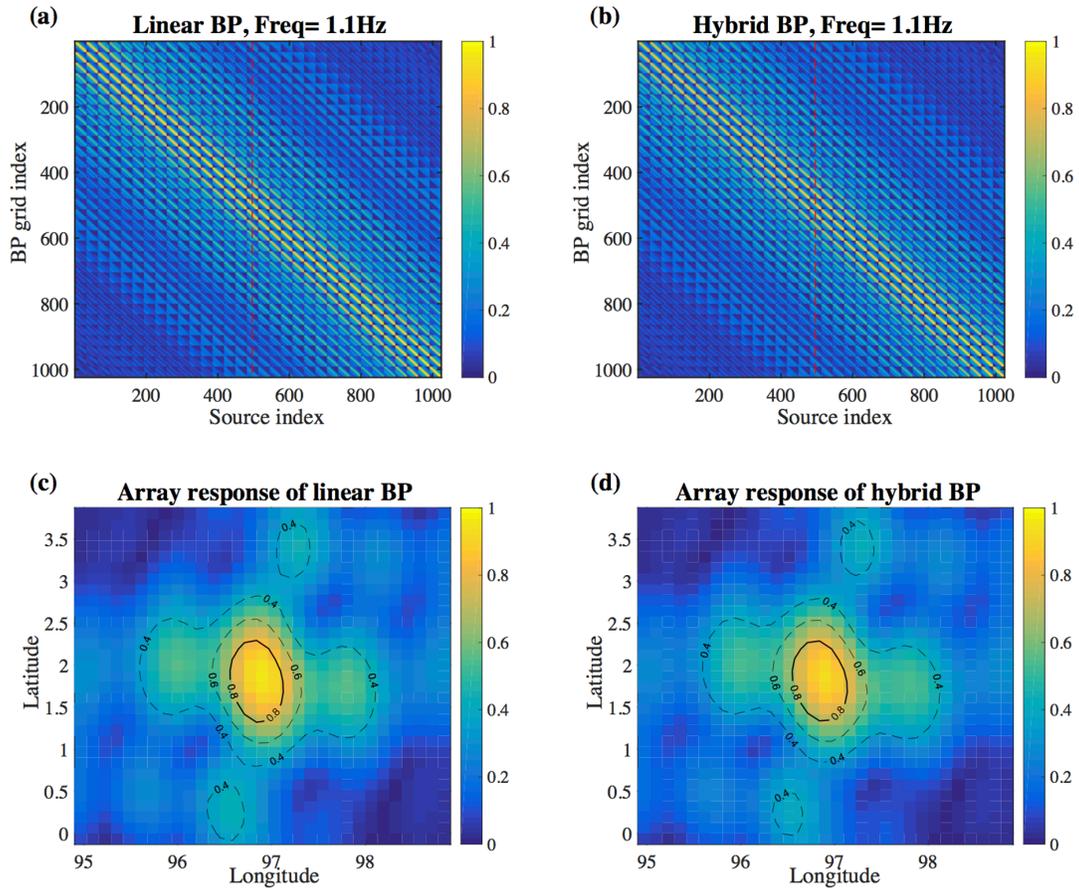
**Figure A.6:** The spectral falloff of the BP peak amplitude for the kinematic source model from (a) Towards1 array; (b) Towards2 array; (c) Away1 array and (d) Away2 results. Crosse-lines show the BP peak amplitudes at the corresponding frequency in each time step that indicated by line colors. Black curves indicate the total BP peaks, which are used to normalized the BP images, during entire rupture at each frequency.



**Figure A.7:** Calculated resolution matrices and their corresponding array response at the center (248th column of the resolution matrices): (a) and (d) from travel time without perturbation; (b) and (e) from systematically perturbed travel time; (c) and (f) from randomly perturbed travel time.



**Figure A.8:** Comparison between the Linear BP image (blue images) with the CSBP results (yellow contours correspond to 20% of maximum CSBP amplitude) in each frequency band.



**Figure A.9:** (a) Resolution matrix of the Hi-net array towards the IDN2 (Sumatra) region. It is same as that in Fig.1 (b) in the main text. (b) The corresponding resolution matrix calculated from the HyBP for the same configuration (Hi-net to IND2). Both resolution matrices are normalized to their maximum element. (c) and (d) show the corresponding array response at the same source location indicated by the red dashed lines in (a) and (b).

B

# Supplementary materials I for Chapter 5

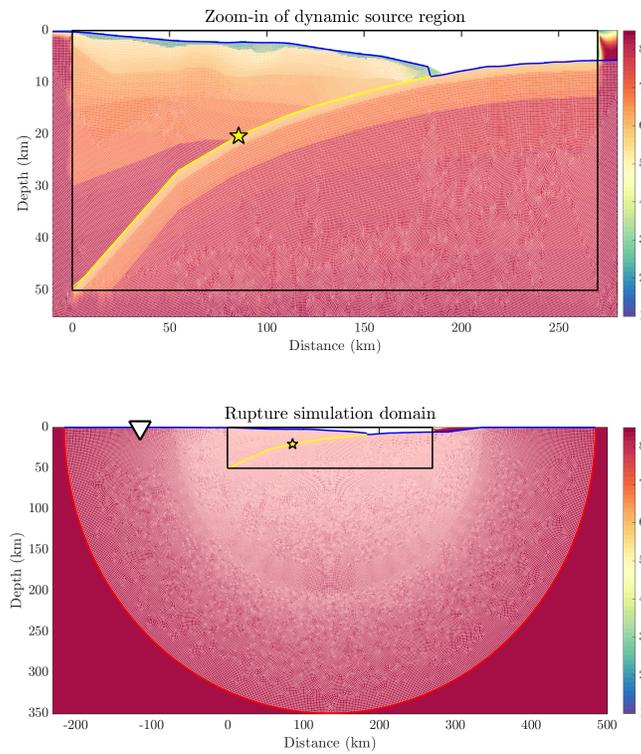
## “The Earth’s surface controls the depth-dependent seismic radiation of megathrust earthquakes”

In this Appendix, the supplementary figures for Chapter 5 “The Earth’s surface controls the depth-dependent seismic radiation of megathrust earthquakes”, are provided. Text S1 presents detailed information about dynamic rupture modeling. The model setup details include model setting, friction, initial stress, and software information. I also show other supplementary figures mentioned in Chapter 5 here.

### B.1 TEXT S1. DETAILS ON THE DYNAMIC RUPTURE SIMULATIONS

#### B.1.1 MODEL SETTING

The simulation domain is a semicircle domain with a radius of 350 km and centered at  $X = 150$  km,  $Y = 0$  km, and a traction-free surface. The simulation domain consists of 1) a near-source, small grid-size, rectangular structure of dimension 270 km  $\times$  50 km (black box area in Fig.B.1 (a)), and 2) a far-source homogeneous half-space (Fig.B.1).



**Figure B.1:** Domain of dynamic simulations. (top) The near-source region with various model settings: Blue and yellow lines indicate the free surface and dynamic fault, respectively. The colormap shows the P wave velocity from Miura et al. <sup>213</sup>. The star indicates the hypocenter of simulated megathrust earthquakes. (bottom) Entire simulation domain: The red semicircle indicates the domain boundary with absorbing conditions. The unstructured mesh is shown in white on top of the simulation domain.

In the near-source region, I test different structural settings: a planar fault embedded in a homogeneous velocity structure and flat topography (Model 1 and Model 15); a curved fault embedded in a homogeneous velocity structure and flat topography (Model 2 and Model 16); and a curved fault embedded in a homogeneous velocity structure and realistic topography (Model 3 and Model 17). The rest of the models use a curved fault embedded in heterogeneous velocity structure and realistic topography. I use the P-wave velocity model directly from tomography<sup>213</sup>. I use the empirical relation of Brocher<sup>43</sup> to calculate density from the  $V_p$  values,  $\rho = 1.74(V_p)^{0.25}$ . The S-wave velocity  $V_s$  is calculated from a  $V_p/V_s$  ratio structure. For most of the simulation domain, I fix the  $V_p/V_s$  ratio constant of  $\sqrt{3} \approx 1.73$ , assuming a Poisson medium. For specific regions detailed as the blue outlined region in Fig.5.3 (b) (also see Fig.B.2), I raise the  $V_p/V_s$  ratio to the following values: 1.83 (Models 8 and 22), 1.94 (Models 9 and 23), 2.04 (Models 10 and 24), 2.14 (Models 11 and 25), 2.24 (Models 12 and 26), 2.34 (Models 13 and 27), and 2.45 (Models 14 and 28). For other heterogeneous models, the  $V_p/V_s$  ratio is fixed constant  $\sqrt{3} \approx 1.73$  (Models 4-7 and Models 18-21). Finally, I can get the shear modulus  $\mu = \rho V_s^2$ .

For the homogeneous models in the far-source region, I have  $V_p = 6.93$  km/s and  $V_s = 4$  km/s, which are the same as those in the near-source region of Models 1-3 and 15-17. For the heterogeneous models,  $V_p = 8.30$  km/s is chosen as the maximum P wave velocity in the model of Miura et al.<sup>213</sup> and  $V_s = 4.79$  km/s, corresponding to  $V_p/V_s$  ratio =  $\sqrt{3}$ . To avoid strong wave reflections from sharp velocity contrasts between the two simulation domains, I set a 5-km wide transition zone with a smooth gradient in the velocity values from the near-source to the far-source regions. At the boundaries of the simulation domain, I set the traction-free boundary condition on the top surface (blue line in Fig.B.1), and absorbing boundary conditions along the borders of the semicircle domain (red line in Fig.B.1).

As a benchmark case for the free-surface effects, I also run one model in a homogeneous full-space (no free surface, Model 29, also referred to as Full in the main text). The simulation domain

of the full-space model is a sufficiently large circular domain with the same radius of 350 km, and an absorbing boundary condition encloses the entire domain. The same curved fault is embedded in the center of the simulation domain, and all other model parameters are kept identical to the homogeneous half-space model.

### B.1.2 FRICTION

I use a linear slip weakening friction for most of my simulations (except Models 4-6 and 18-20). The parameters of linear slip weakening are constant from the surface down to 40 km depth (static friction coefficient  $\mu_s = 0.677$ ; dynamic friction coefficient  $\mu_d = 0.2$ ; the critical slip of slip weakening  $D_c = 0.4$  m. Below 40 km, I increase the dynamic friction coefficient to 0.99 to force the termination of the rupture. While the focus of this study is not to explore all frictional relations, I test several different friction relations above 10.8 km depth (at the base of the frontal prism) to be slip neutral/stable ( $\mu_s = \mu_d = 0.677$  above 10.8 km depth, Models 4 and 18) or slip hardening/strengthening ( $\mu_s = 0.677, \mu_s < \mu_d = 0.85$  and  $D_c = 2$  m above 10.8 km depth, Models 5 and 19). Finally, I include a model with the same lab-based exponential slip weakening proposed by Murphy et al.<sup>225</sup> in Models 6 and 20. I use the same relations (See their equations (1) and (2)) to set up the stress and frictional parameters.

### B.1.3 INITIAL STRESS

In my simulations, the effective normal stress  $\bar{\sigma}_n$  is reduced from the fault normal stress  $\sigma_L$  due to pore pressure  $p$ ,  $\bar{\sigma}_n = \sigma_L - p$ . Because of the relatively low dip angle of the fault, I approximate the normal stress  $\sigma_L$  as the lithostatic stress that is calculated based on the density structure  $\rho(x, h)$  of each model:  $\sigma_L(x) = \int_{h_{slab}}^{h_0} \rho(x, h) g dh$ , where  $h_{slab}$  and  $h_0$  are the depths of slab surface and top free surface,  $g$  is the gravitation constant. I use the fluid pressure ratio  $\lambda$  to quantify the pore pressure:

$p = \lambda\sigma_L$ . This quantification is introduced by Hubbert & Rubey<sup>130</sup> and has been used in many previous studies<sup>225,186</sup>. Finally, I assume the effective normal stress  $\bar{\sigma}_n$  is bounded at 40 MPa, at which the over-pressurized pore pressure becomes lithostatic<sup>256</sup>, and this is similar to the settings in Lotto et al.<sup>186</sup>. In this study, I mainly vary  $\lambda$  for the stress setting variations of models and include cases of  $\lambda = 0.9$  and  $\lambda = 0.7$ . This parameter controls how pore pressure varies along the depth and where the pore fluid becomes lithostatic (see Fig.5.3 (c) in the Chapter 5).

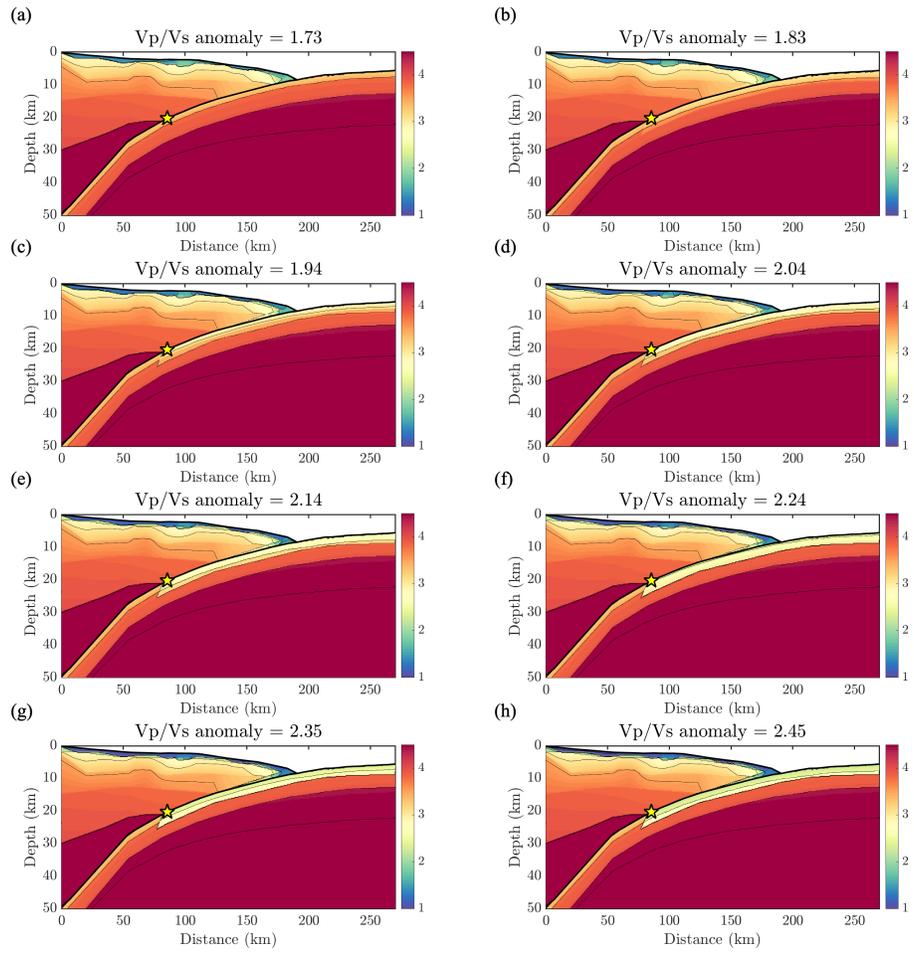
I assume a relatively low initial shear stress  $\tau_0$  on the fault and calculate it using the seismic S ratio (Fig. 2c), which is used to measure how close the initial stress is to the level of failure<sup>71</sup>:

$$S = \frac{\tau_s - \tau_0}{\tau_0 - \tau_d} = 2.77, \quad (\text{B.1})$$

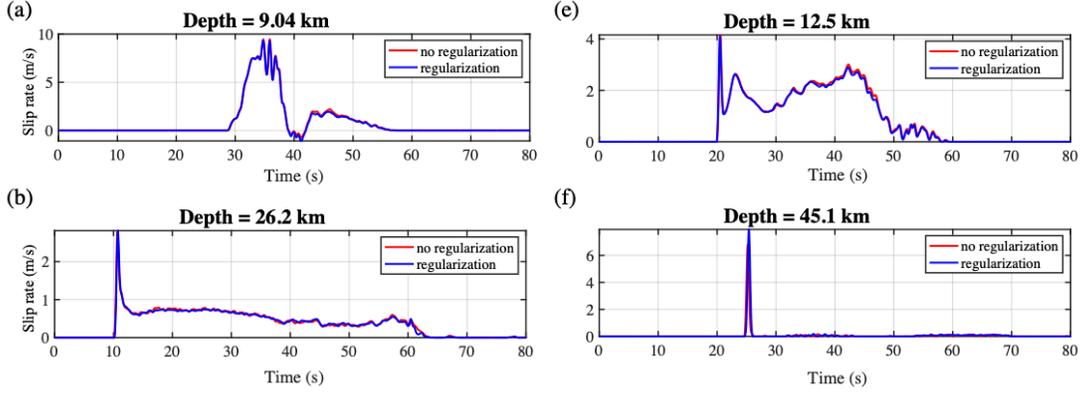
where  $\tau_s = \bar{\sigma}_n\mu_s$  and  $\tau_d = \bar{\sigma}_n\mu_d$  are the static friction (yielding stress) and dynamic friction, respectively. This high seismic S ratio is set to avoid the unwanted supershear rupture that arises from high initial stress and resulting high dynamic stress drop<sup>17,86</sup>. Finally, I use over-stress nucleation to start the spontaneous dynamic rupture for all models. I increase the initial shear stress to  $1.016\tau_s$  within a 2-km patch on fault centered at a depth of 20 km (Fig.5.3 (c)). The only exceptions are the models with exponential slip weakening friction (Models 6 and 20). I have to set a larger nucleation zone of about 14 km to nucleate megathrust rupture successfully. I have checked the results of those models (Models 6 and 20) and can assure that this large nucleation patch has negligible effects on the later dynamic rupture process.

#### B.1.4 NUMERICAL SOLVER

The entire domain is discretized with unstructured mesh using software CUBIT (<https://cubit.sandia.gov/>), the mesh script is written based on Huang et al.<sup>129</sup>). To determine the element grid size, I estimate the corresponding cohesive zone size  $\Lambda_0$  based on Palmer & Rice<sup>242</sup>:  $\Lambda_0 = \frac{2\pi}{32} \frac{\mu}{(1-\nu)} \frac{D_c}{(\tau_s - \tau_d)}$ ,



**Figure B.2:** Corresponding S wave velocity from different settings of  $V_p/V_S$  ratios: (a)  $V_p/V_S = 1.73$ ; (b)  $V_p/V_S = 1.84$ ; (c)  $V_p/V_S = 1.94$ ; (d)  $V_p/V_S = 2.04$ ; (e)  $V_p/V_S = 2.14$ ; (f)  $V_p/V_S = 2.24$ ; (g)  $V_p/V_S = 2.34$ ; (h)  $V_p/V_S = 2.45$ .



**Figure B.3:** Comparisons between slip-rate functions with (blue) or without (red) normal stress regularization after the Gaussian time window smoothing, extracted at different depths: (a) 9.04 km; (b) 12.5 km; (c) 26.2 km and (d) 45.1 km.

where  $\nu = \frac{1}{2} \frac{(V_p/V_S)^2 - 2}{(V_p/V_S)^2 - 1}$  is the Poisson's ratio. For the homogeneous model  $V_p = 6.93$  km/s,  $V_p/V_S = \sqrt{3}$  and  $\tau_s - \tau_d = 40$  MPa, the corresponding cohesive zone size  $\Lambda_0 = 1114.4$  m. For the heterogeneous model, I take the case of  $V_p = 4$  km/s,  $V_p/V_S = 2.45$  and  $\tau_s - \tau_d = 8$  MPa as a representative lower bond estimation, which gives the cohesive zone size  $\Lambda_0 = 1012.6$  m. Based on the estimation of the cohesive zone size, I set the element grid size  $dl = 500$  m  $< \Lambda_0/2$  in the source domain (Fig.B.1) to ensure sufficient numerical resolution<sup>72</sup>. Accordingly, the frequency resolution is determined by  $dl$  and the minimum S wave wavelength. I require at least  $n = 4$  grids within the minimum wavelength, so I can estimate the maximum resolvable frequency of my simulations. This varies for different models. For the models with homogeneous velocity structure (Models 1-3, 15-17),  $V_S = 4$  km/s and the maximum frequency I can resolve is  $f = V_S/4dl = 2$  Hz. For the models with heterogeneous velocity structures, the maximum resolvable frequency varies with minimum  $V_S$ . The minimum shear wave speed in all the velocity models is 0.6 km/s, corresponding to  $f = \min(V_S)/4dl = 0.3$  Hz. In my results, I will interpret radiation below this maximum frequency.

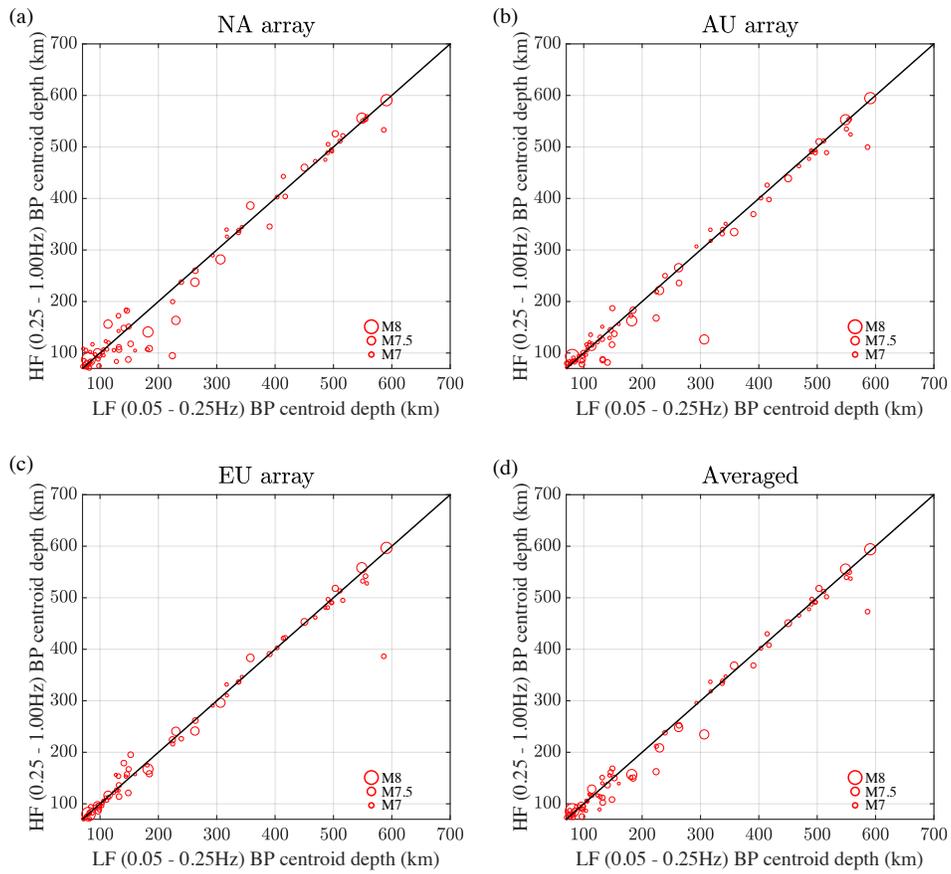
I use the 2D spectral element-based code SEM2DPACK<sup>14</sup> (available at <https://github.com/jpampuelo/sem2dpack>, last accessed on 06/08/2021) to solve for the dynamic rupture. This code

has been well validated and applied in some previous studies<sup>125,129</sup> to simulate the megathrust earthquakes as well as the wave fields.

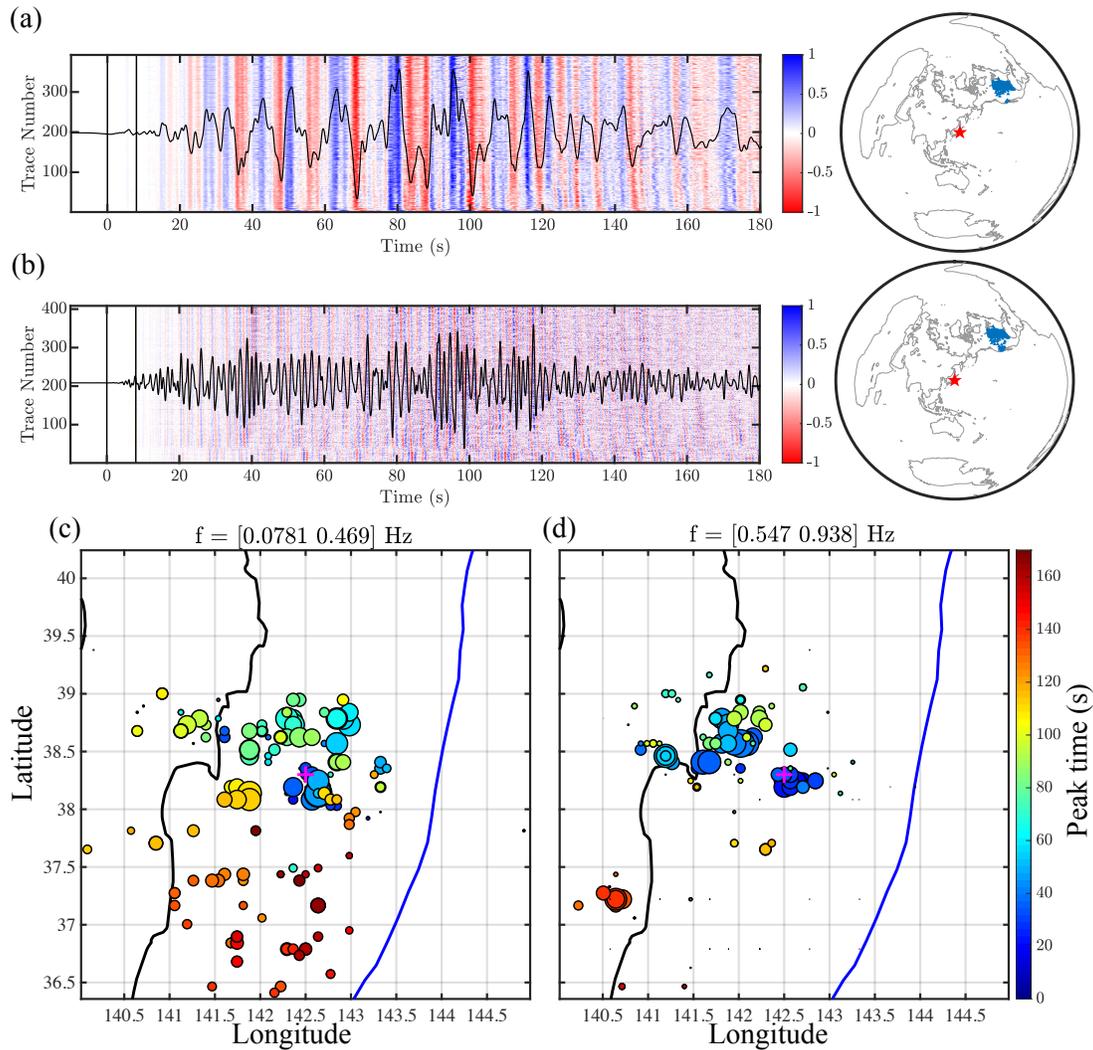
In most of my simulations, I include the realistic velocity models, which have significant material contrasts in the downdip regions (Fig. 5.3 (b) or Fig.B.2). The material contrasts can lead to ill-posedness in the numerical solution and regularization is needed<sup>63</sup>. As proposed by Rubin & Ampuero<sup>263</sup>, Ampuero & Ben-Zion<sup>15</sup>, Huang<sup>124</sup>, the material contrasts can cause normal stress perturbation during dynamic rupture. They suggest using a regularization  $\dot{\sigma}^* = \frac{V^*}{D_\sigma}(\sigma - \sigma^*)$  to force the normal stress to evolve continuously.  $\sigma$  and  $\sigma^*$  are the actual normal stress and the regularized normal stress (referred to as an “effective” normal stress but here I use “regularized” to differentiate from the one related to pore pressure). The reference velocity  $V^*$  and slip distance  $D_\sigma$  are the two constitutive parameters. In my simulations, since I am focusing on the fault slip within the frequency band below 0.3 Hz, I apply a 1-s-long Gaussian window to smooth out the numerical noise in the slip rate functions. I compare models processed by different schemes and find that the slip-rate functions are almost indistinguishable (Fig.B.3).

**Table B.1:** Range in  $V_p$  values in the downgoing slab low velocity zone LVZ ( $V_{LVZ}$ ) and in the overhanging continental crust ( $V_{cont}$ ) for various subduction zones.

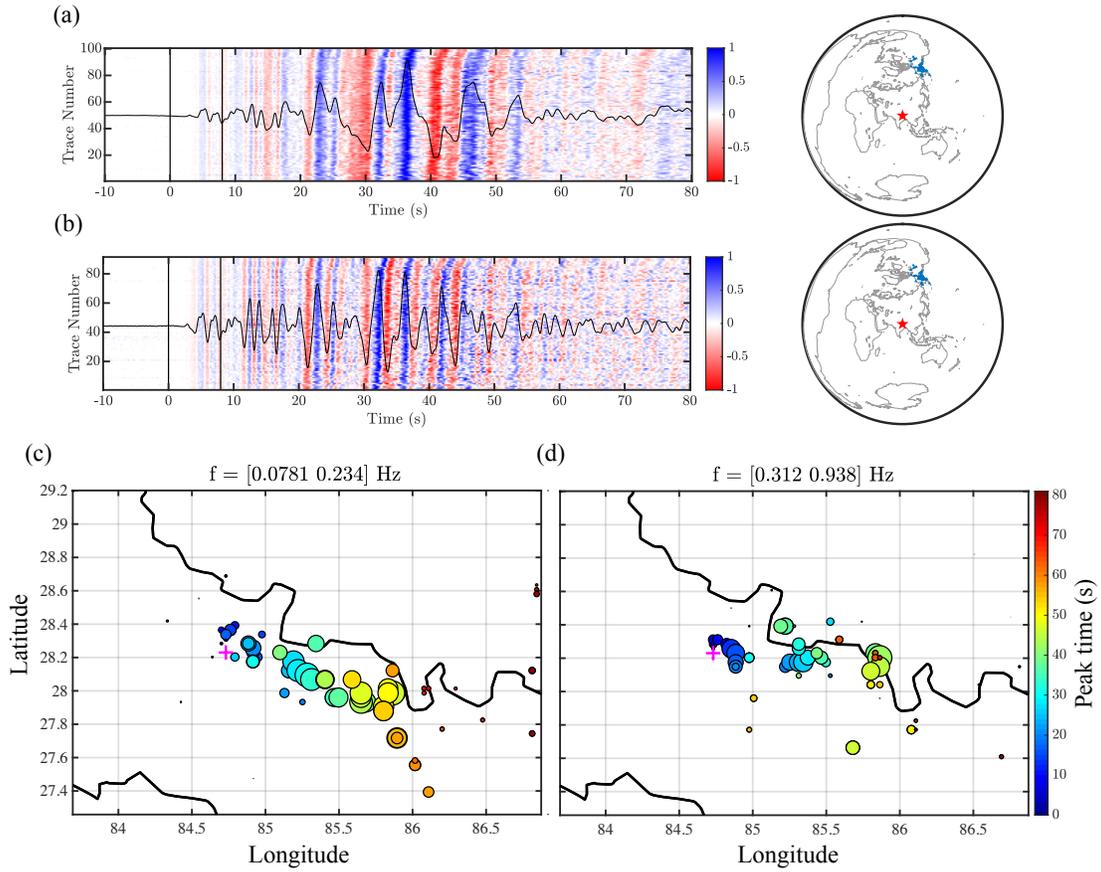
Subduction zone	Reference	$V_{LVZ}$ (km/s)	$V_{cont}$ (km/s)
Alaska	Ye et al. <sup>356</sup>	4.9 - 5.1	4.6 - 5.1
Antilles	Kopp et al. <sup>165</sup>	5.5 - 6.0	6.5 - 8.0
Cascadia	Horning et al. <sup>120</sup>	4.0 - 4.5	4.5 - 6.5
Chile 1	Contreras-Reyes et al. <sup>65</sup>	3.5 - 4.8	5.5 - 6.0
Chile 2	Scherwath et al. <sup>273</sup>	4.5 - 5.0	5.0 - 7.0
Chile 3	Moscoso et al. <sup>215</sup>	4.5 - 6.0	6.0 - 6.9
Chile 4	Contreras-Reyes et al. <sup>64</sup>	4.0 - 5.0	5.5 - 7.0
Costa Rica 1	Walther et al. <sup>323</sup>	5.5 - 6.0	5.7 - 8.3
Costa Rica 2	Sallarès et al. <sup>269</sup>	5.0 - 6.3	5.9 - 7.2
Costa Rica 3	Zhu et al. <sup>369</sup>	3.0 - 4.0	4.5 - 6.0
Costa Rica 4	Martínez-Loriente et al. <sup>200</sup>	4.0 - 5.0	4.0 - 6.5
Ecuador 1	Graindorge et al. <sup>104</sup>	5.0 - 6.0	6.0 - 6.7
Ecuador 2	Gailler et al. <sup>101</sup>	4.5 - 6.0	4.5 - 6.5
Ecuador 3	Agudelo et al. <sup>4</sup>	4.5 - 6.0	6.0 - 7.0
Izu Bonin	Takahashi et al. <sup>294</sup>	4.7 - 6.4	5.7 - 7.4
Java 1	Planert et al. <sup>250</sup>	3.0 - 4.5	5.0 - 7.6
Java 2	Shulgin et al. <sup>287</sup>	5.0 - 6.0	5.0 - 7.5
Kuril	Nakanishi et al. <sup>229</sup>	4.5 - 6.0	6.0 - 8.0
Nankai Trough 1	Kodaira et al. <sup>162</sup>	5.2 - 5.8	5.2 - 6.7
Nankai Trough 2	Nakanishi et al. <sup>230</sup>	4.2 - 5.4	5.0 - 6.8
New Zealand	Bassett et al. <sup>28</sup>	4.9 - 6.3	6.8 - 8.5
Nicaragua 1	Walther et al. <sup>323</sup>	5.5 - 6.9	5.9 - 8.3
Peru 1	Hampel et al. <sup>110</sup>	4.5 - 5.0	4.2 - 5.5
Peru 2	Krabbenhöft et al. <sup>169</sup>	4.0 - 6.1	5.7 - 6.5
Ryukyu	Nishizawa et al. <sup>233</sup>	5.0 - 6.0	5.0 - 7.0
Sumatra	Klingelhoefer et al. <sup>158</sup>	5.0 - 6.0	5.0 - 8.0
Solomon	Miura et al. <sup>212</sup>	5.0 - 6.3	5.3 - 6.9
Taiwan	Klingelhoefer et al. <sup>157</sup>	5.5 - 6.0	4.5 - 7.0
Tohoku	Miura et al. <sup>213</sup>	5.5 - 6.6	5.5 - 8.0
Tonga 1	Contreras-Reyes et al. <sup>66</sup>	5.5 - 6.5	6.0 - 7.5
Tonga 2	Bassett et al. <sup>27</sup>	3.8 - 4.5	4.5 - 7.9



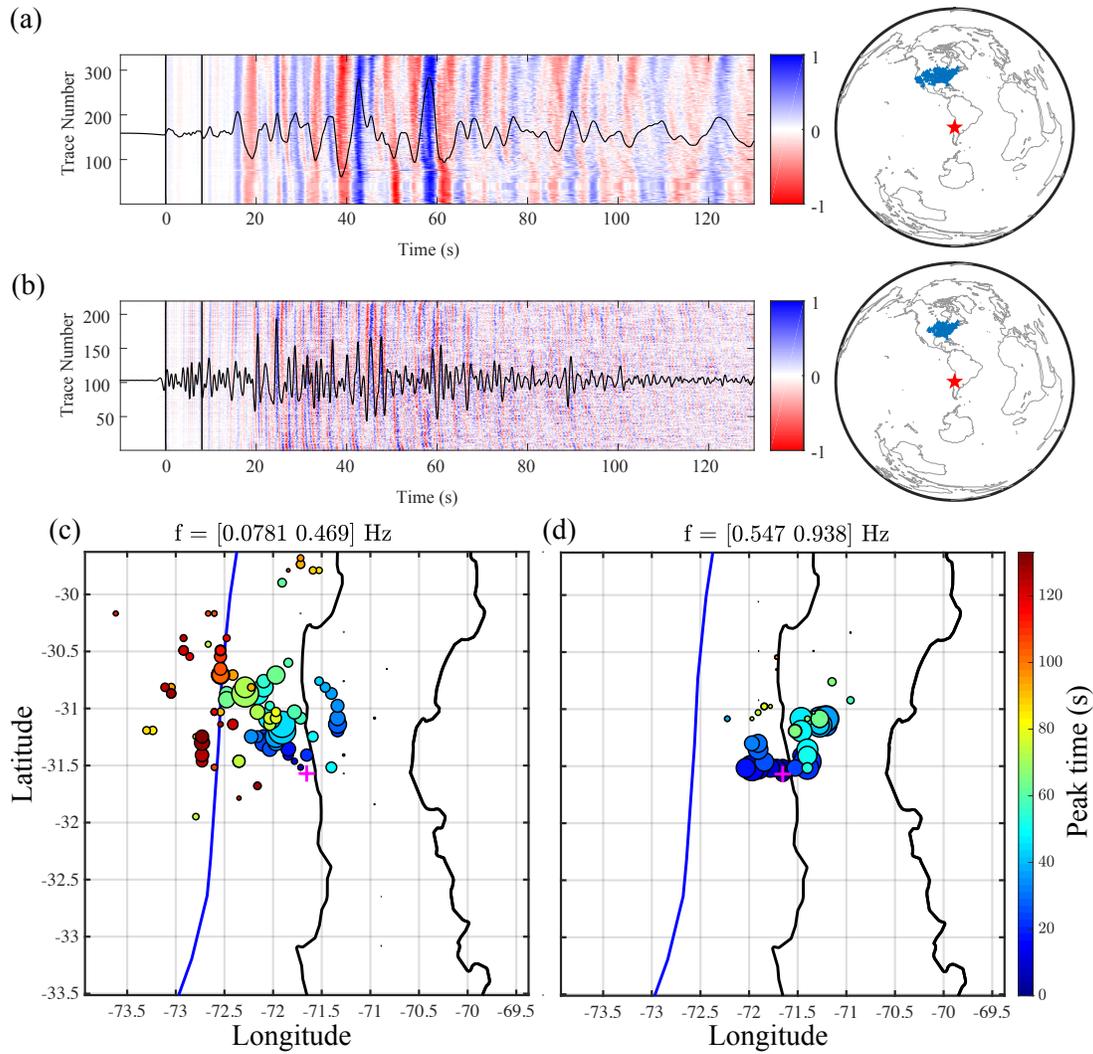
**Figure B.4:** Comparison between the LF BP centroid depth from GSN and HF BP centroid depth from (a) NA array; (b) AU array; (c) EU array and (d) three-array-average for the deep earthquakes (70 - 700 km) in the IRIS back-projection database.



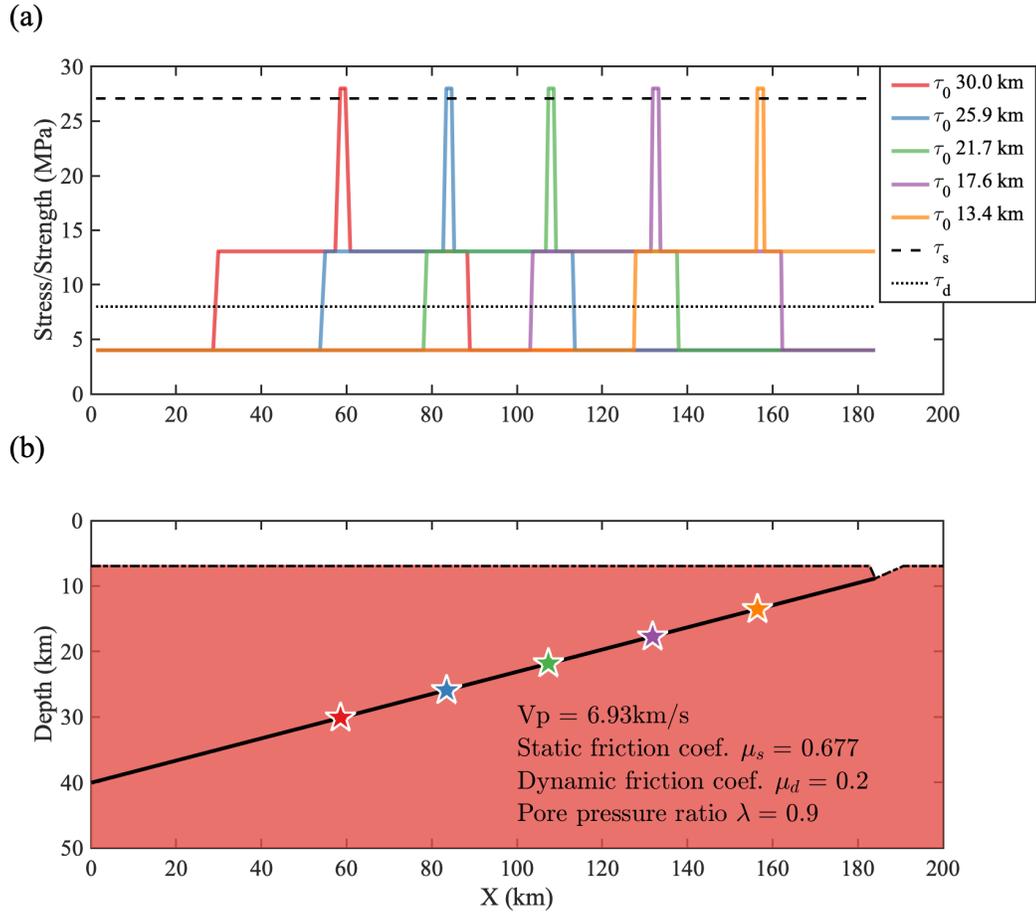
**Figure B.5:** Data and back-projection results of the Mw 9.0 2011 Tohoku-oki earthquake. (a) Teleseismic P wave velocity seismograms filtered in the LF band (0.05 - 0.5 Hz) and the corresponding TA array distribution (blue triangles to the right and the red star indicates the location of the epicenter). The aligned waveforms recorded by the array are shown by the red-to-blue image and the stacked waveform is also shown on top of the image. (b) Same as (a) but for the teleseismic P wave velocity seismograms filtered in the high-frequency band (0.5 - 1 Hz). (c) imCS-BP results in the low-frequency band (0.05 - 0.5 Hz): the circles indicate the energy bursts, their colors correspond to the time of the burst since the onset of the earthquake, and their sizes are proportional to the amplitude power of energy bursts. The purple cross indicates the location of the epicenter. (d) The imCS-BP results in the high-frequency band (0.5 - 1 Hz) and the symbols have the same meanings as (c).



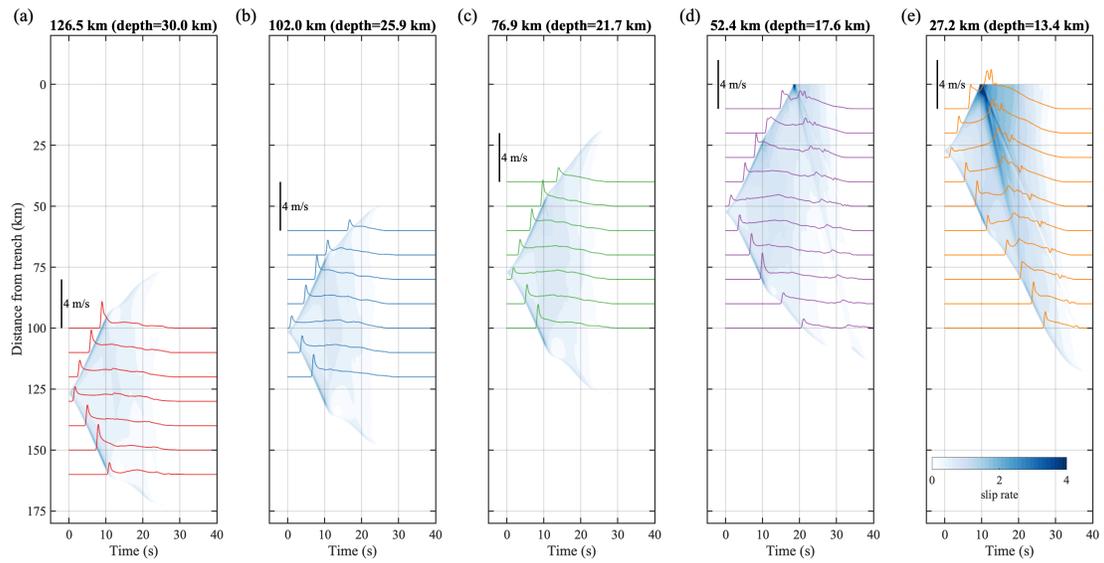
**Figure B.6:** Data and back-projection results of the Mw 7.9 2015 Gorkha earthquake. (a) Teleseismic P-wave velocity seismograms filtered in the low-frequency band (0.05 - 0.25 Hz). (b) Same as (a) but for the teleseismic P-wave velocity seismograms filtered in the high-frequency band (0.25 - 1 Hz). (c) imCS-BP results in the low-frequency band (0.05 - 0.25 Hz). (d) imCS-BP results in the high-frequency band (0.25 - 1 Hz) and all other symbols have the same meanings as Fig. B.5.



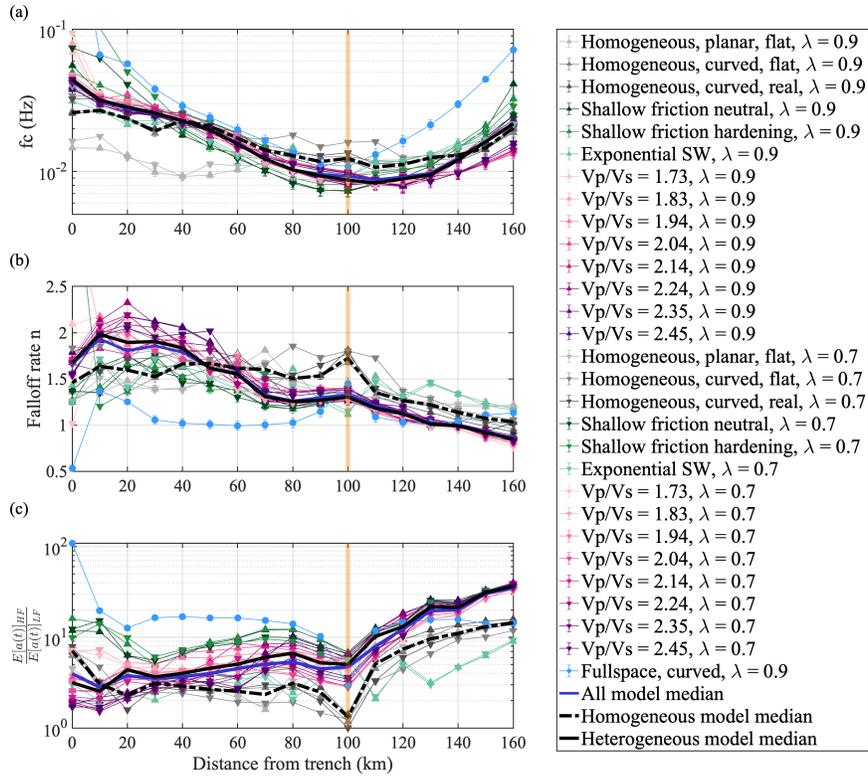
**Figure B.7:** Data and back-projection results of the Mw 8.3 2015 Illapel earthquake. (a) Teleseismic P-wave velocity seismograms filtered in the LF band (0.05 - 0.5 Hz). (b) Same as (a) but for the teleseismic P-wave velocity seismograms filtered in the high-frequency band (0.5 - 1 Hz). (c) imCS-BP results in the low-frequency band (0.05 - 0.5 Hz). (d) imCS-BP results in the high-frequency band (0.5 - 1 Hz) and all other symbols have the same meanings as Fig. B.5.



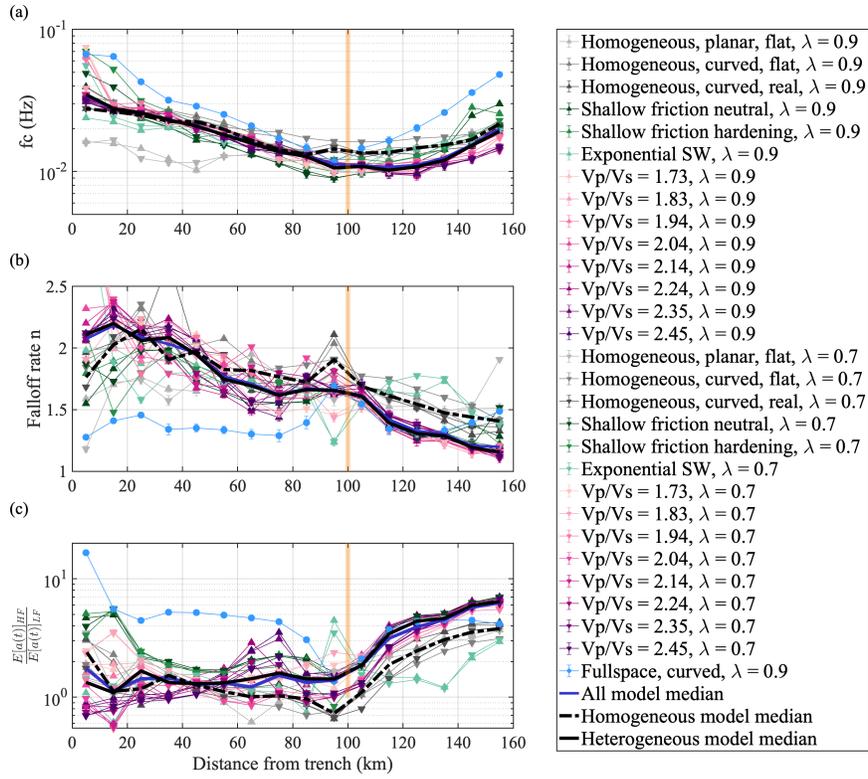
**Figure B.8:** Model settings of the five small megathrust earthquake models. (a) Stress/strength distribution along the slab (in X coordinate): the black dotted line and dashed line show the dynamic friction  $\tau_d$  and static friction  $\tau_s$ , respectively. Colored lines indicate the initial shear stress  $\tau_0$  for earthquakes nucleated at different depths: red - 30.0 km; blue - 25.9 km; green - 21.7 km; purple - 17.6 km; orange - 13.4 km. (b) Simulation domain for a homogeneous medium with planar slab geometry and flat topography for the small rupture models. The colored stars indicate the location of nucleation/hypocenters of the small earthquakes.



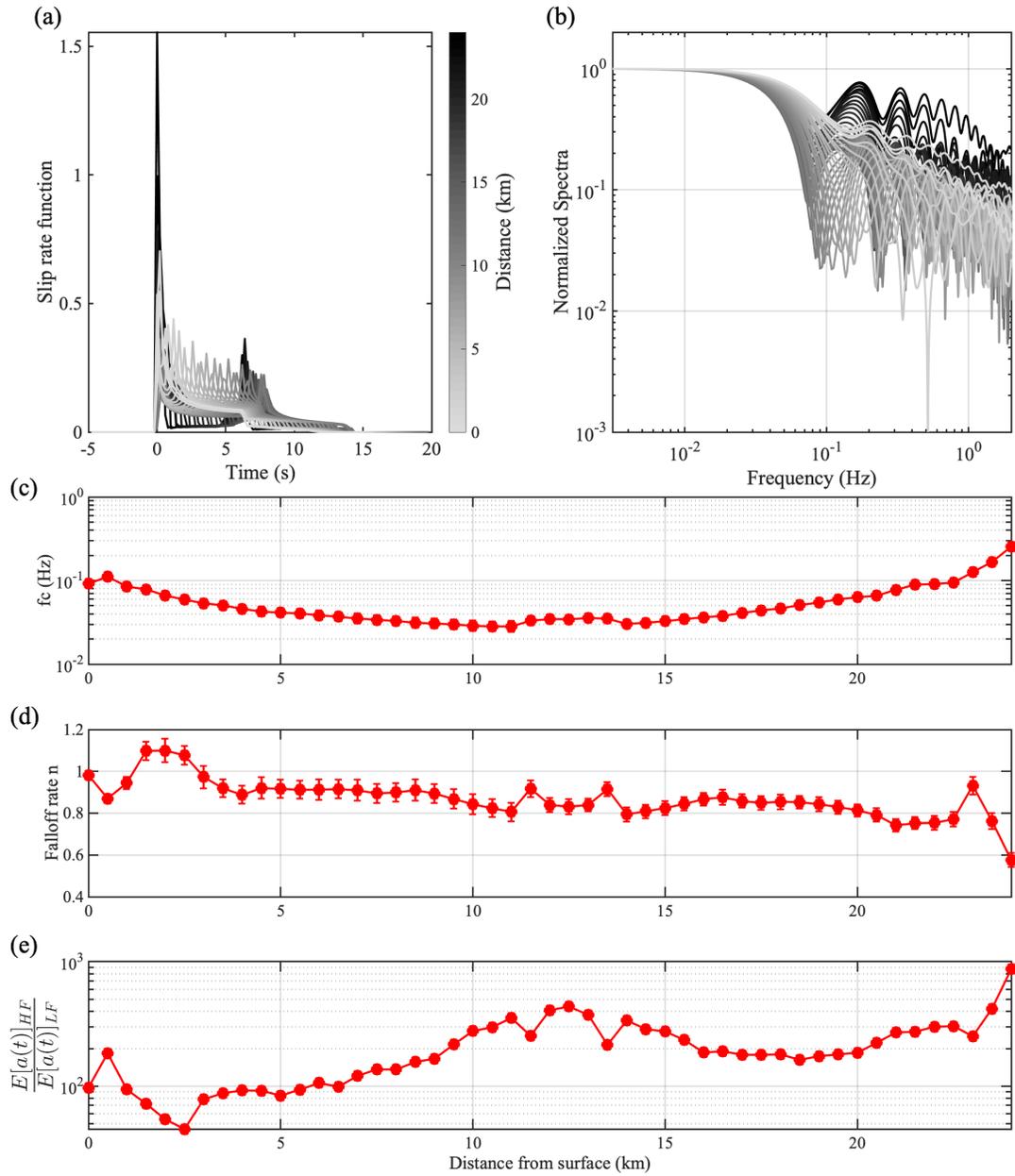
**Figure B.9:** Space-time plot of the slip histories for all small earthquake models nucleated at different depths measured by distance from trench (depth): (a) distance = 126.5 km / depth = 30.0 km; (b) distance = 102.0 km / depth = 25.9 km; (c) distance = 76.9 km / depth = 21.7 km; (d) distance = 52.4 km / depth = 17.6 km and (e) distance = 27.2 km / depth = 13.4 km. The slip-rate functions at different points are also shown in colored lines.



**Figure B.10:** Results of the spectral content of the slip-rate function extracted at individual point every 10 km along dip: (a) corner frequency  $f_c$ ; (b) spectral falloff rate  $n$ ; (c) HF/LF power ratio of slip acceleration. Yellow bars indicate the location where rupture is nucleated.



**Figure B.11:** Results of the spectral content of the slip-rate functions averaged over 10-km subfault along dip: (a) corner frequency  $f_c$ ; (b) spectral falloff rate  $n$ ; (c) HF/LF power ratio of slip acceleration. Yellow bars indicate the location where rupture is nucleated.



**Figure B.12:** Results of a mode-III (anti-plane) rupture model on a vertical fault intersecting the free surface in a homogeneous medium. Fault length is 25 km, and the rupture is nucleated at 12.5 km depth with over-stress nucleation. Other model parameters are:  $V_p=6.9$  km/s,  $V_s=4.0$  km/s,  $D_c=0.4$  m,  $\mu_s=0.677$ ,  $\mu_D=0.2$ ,  $\bar{\sigma}_n=40$  MPa. (a) - (b) Slip-rate function and slip-rate spectrum at different depths. (c) - (e) Along-depth variation of corner frequency  $f_c$ , spectral falloff rate  $n$  and HF/LF power ratio of slip acceleration.

C

## Supplementary materials II for Chapter 5

# “The Earth’s surface controls the depth-dependent seismic radiation of megathrust earthquakes”

In this Appendix C, I provide details about all simulation settings and results of each model of Chapter 5, including the model parameters ((a) structures, (b) stress, (c) friction), simulations results ((d) - (g)) and the fitting of spectral parameters ((h) corner frequency  $f_c$  and spectral falloff rate  $n$ ).

Caption of the figures. Each figure contains 8 subfigures to show model settings and results. (a) The structure of model: topography, fault geometry, P wave velocity. The blue outlined region (if any) indicates the region where I set the  $V_P/V_S$  ratio to the given value. (b) Initial stress distributions along depth (black line: initial shear stress  $\tau_0$ ; gray line: initial effective normal stress  $\bar{\sigma}_0$ ). (c) Parameters of used friction law along depth: upper X-axis shows the friction coefficients (red dashed line: dynamic friction coefficient  $\mu_d$ ; red solid line: static friction coefficient  $\mu_s$ ), bottom X-axis shows the critical slip  $D_c$  in black line. (d) Space-time evolution of the rupture (in blue colormap) and of selected points on the fault (black lines), including the one at the trench/surface (thick black line). Gray and red lines show the updip- and downdip-propagating rupture front, respectively. I

estimate the rupture velocity by linear fitting the location and time of rupture front. Light purple lines (if any) indicate the super-shear rupture front triggered by free surface and shallow compliant structures. (e) Slip-rate functions at each fault segment, aligned to their onset time (when rupture front arrives). The location of the fault segment center taken as the along-dip distance from the trench is indicated by the gray colormap. (f) Normalized Fourier amplitude spectra corresponding to the slip-rate functions shown in (e). The same color scheme is used to indicate the fault segment location. (g) moment-rate density function averaged along the entire fault. (h) The along-dip best-fit spectral parameters of the spectra in (f) as well as its 95% confidence interval. The right Y-axis shows the corner frequency  $f_c$  in red. The left Y-axis shows the spectral falloff rate  $n$  in blue.

Model 1: Homogeneous media, planar fault, flat topography,  $\lambda = 0.9$

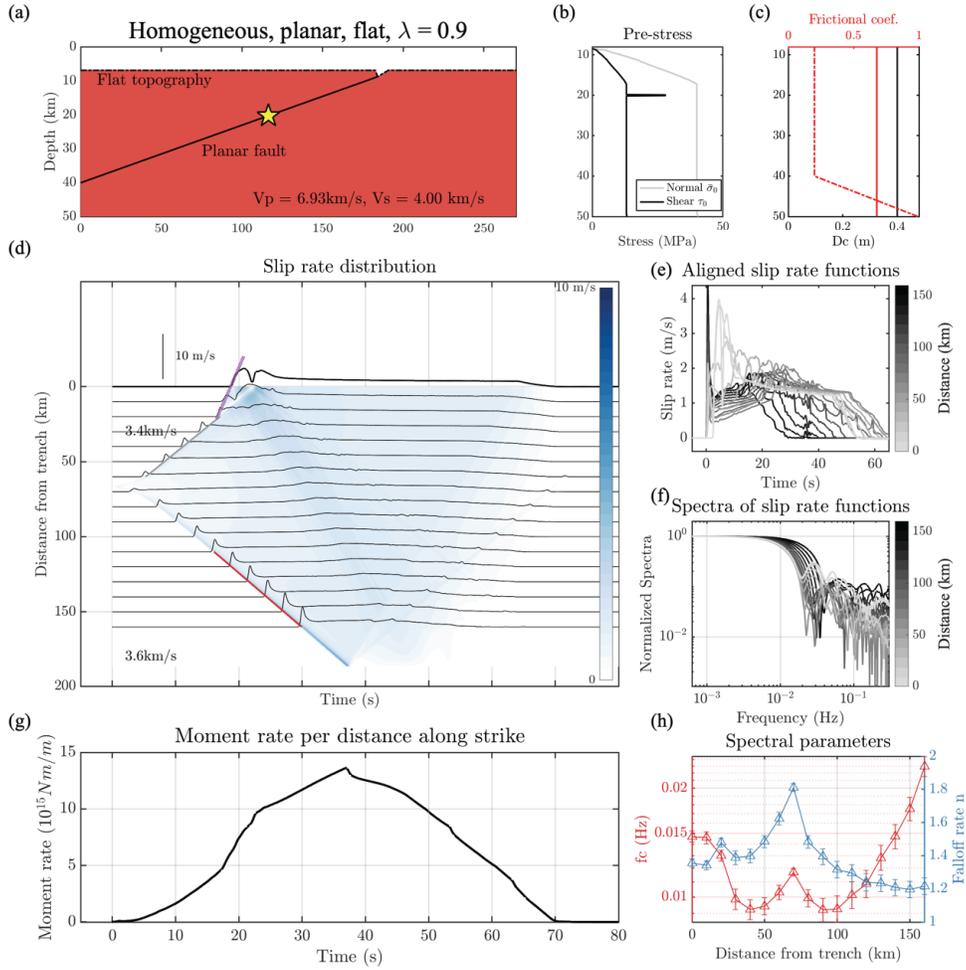


Figure C.1: Model settings and simulation results of model 1.

Model 2: Homogeneous media, curved fault, flat topography,  $\lambda = 0.9$

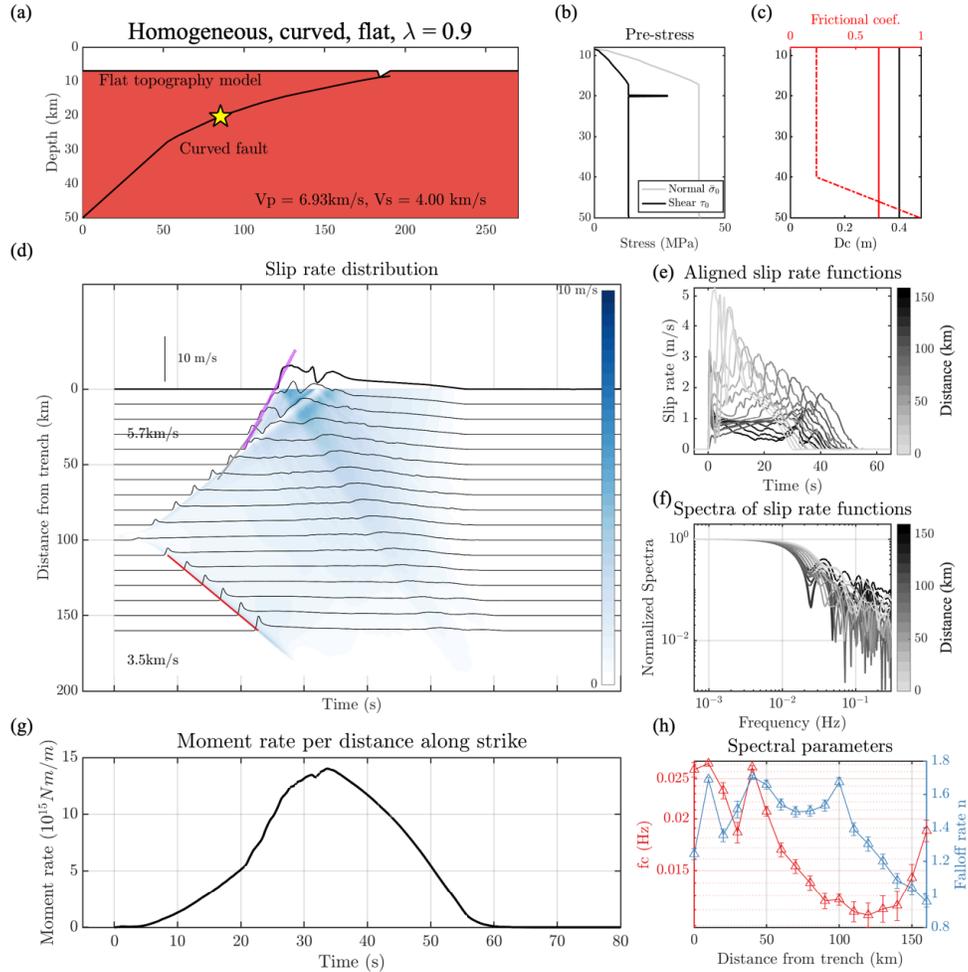


Figure C.2: Model settings and simulation results of model 2.

Model 3: Homogeneous media, curved fault, realistic topography,  $\lambda = 0.9$

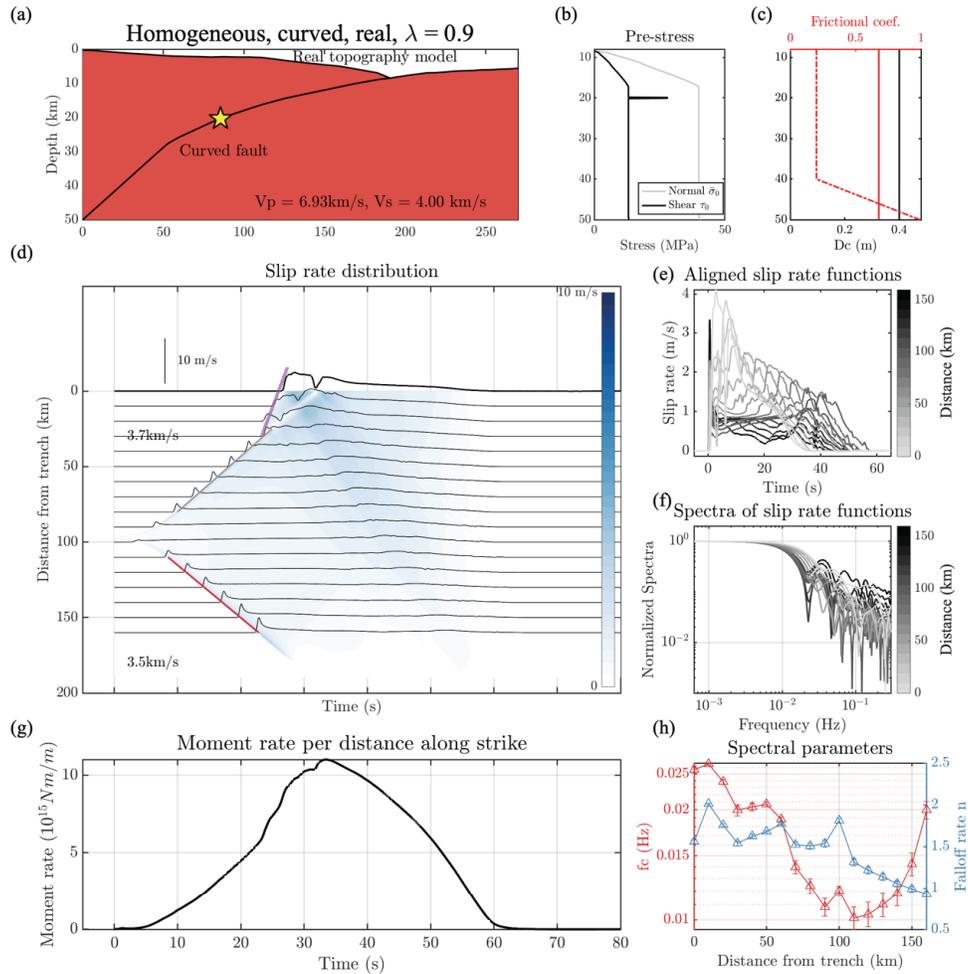


Figure C.3: Model settings and simulation results of model 3.

Model 4: Heterogeneous media, curved fault, realistic topography, shallow neutral friction,  
 $V_p/V_s = \sqrt{3}, \lambda = 0.9$

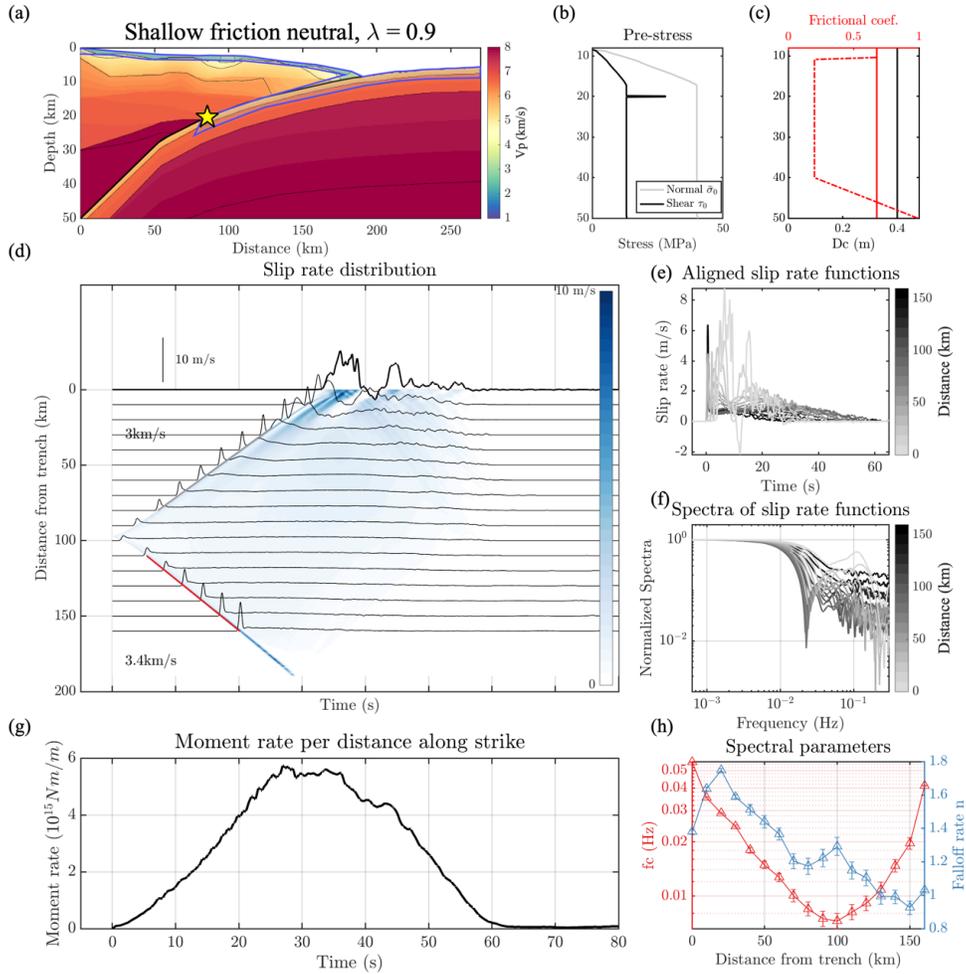


Figure C.4: Model settings and simulation results of model 4.

Model 5: Heterogeneous media, curved fault, realistic topography, shallow hardening/strengthening friction,  $V_P/V_S = \sqrt{3}, \lambda = 0.9$

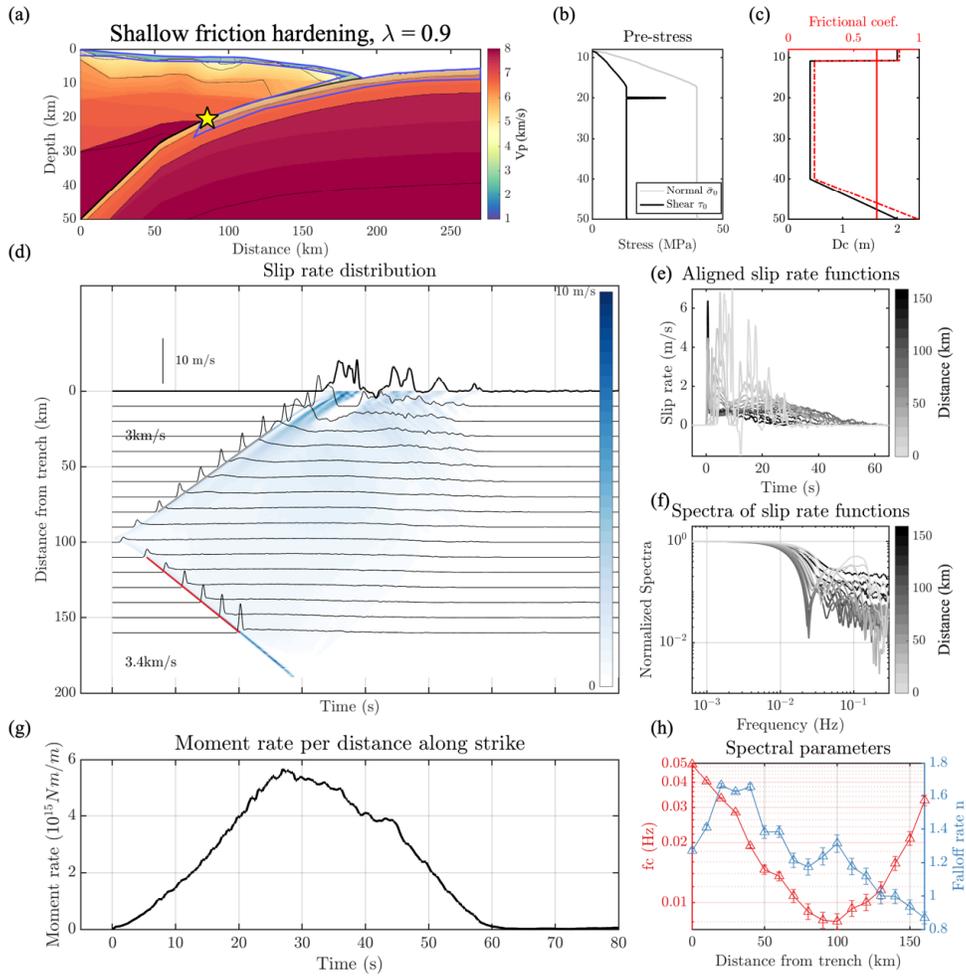


Figure C.5: Model settings and simulation results of model 5.

Model 6: Heterogeneous media, curved fault, realistic topography, exponential slip-weakening friction,  $V_p/V_S = \sqrt{3}, \lambda = 0.9$

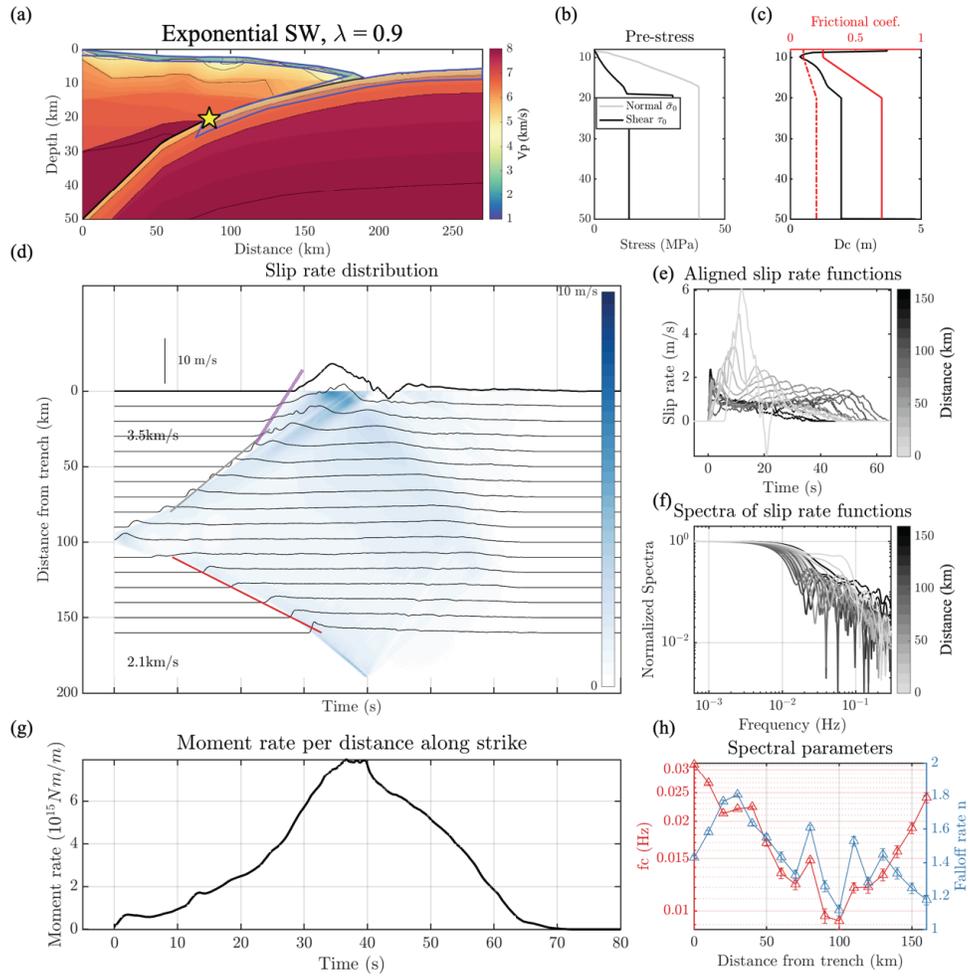


Figure C.6: Model settings and simulation results of model 6.

Model 7: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = \sqrt{3}, \lambda = 0.9$

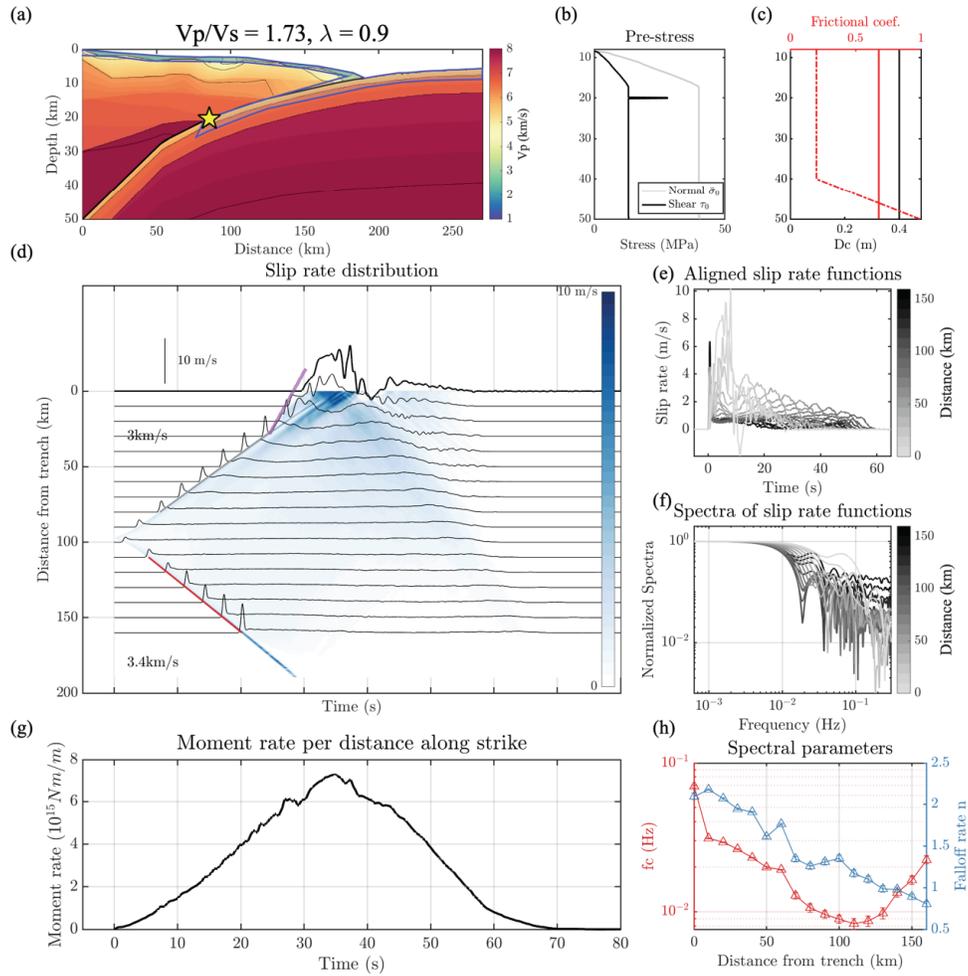


Figure C.7: Model settings and simulation results of model 7.

Model 8: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 1.83, \lambda = 0.9$

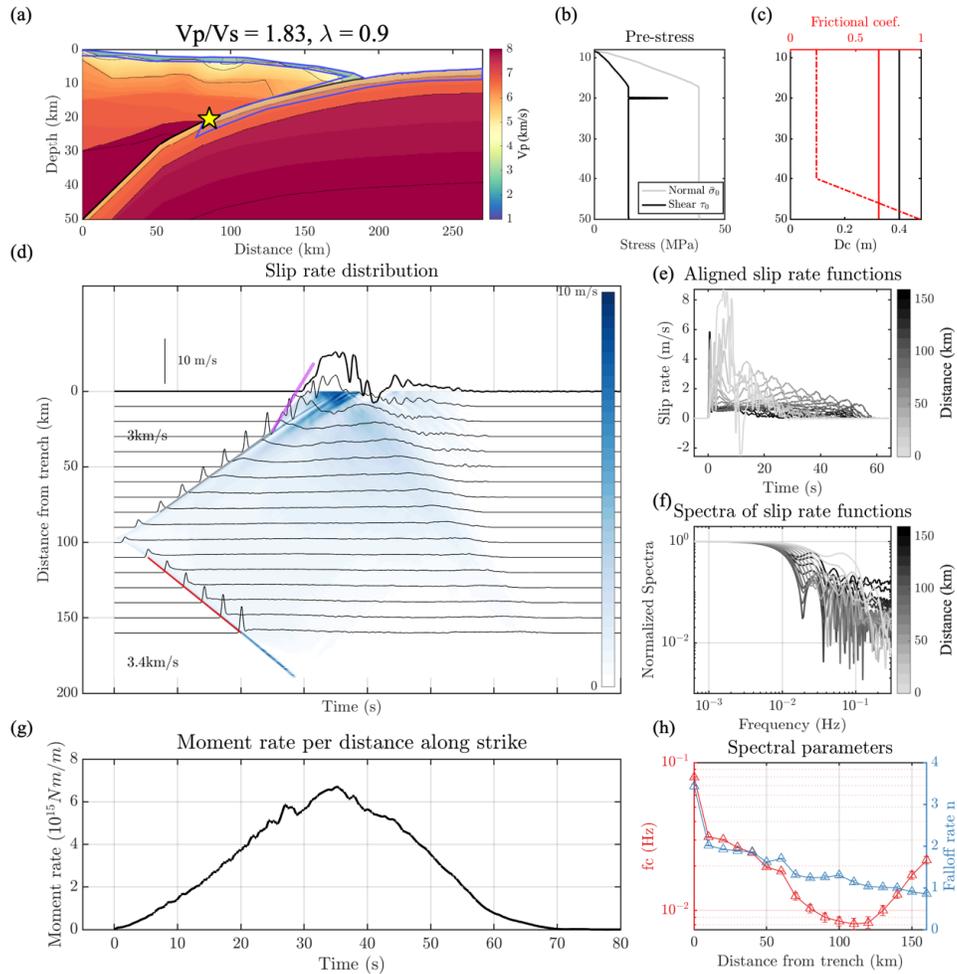


Figure C.8: Model settings and simulation results of model 8.

Model 9: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 1.94, \lambda = 0.9$

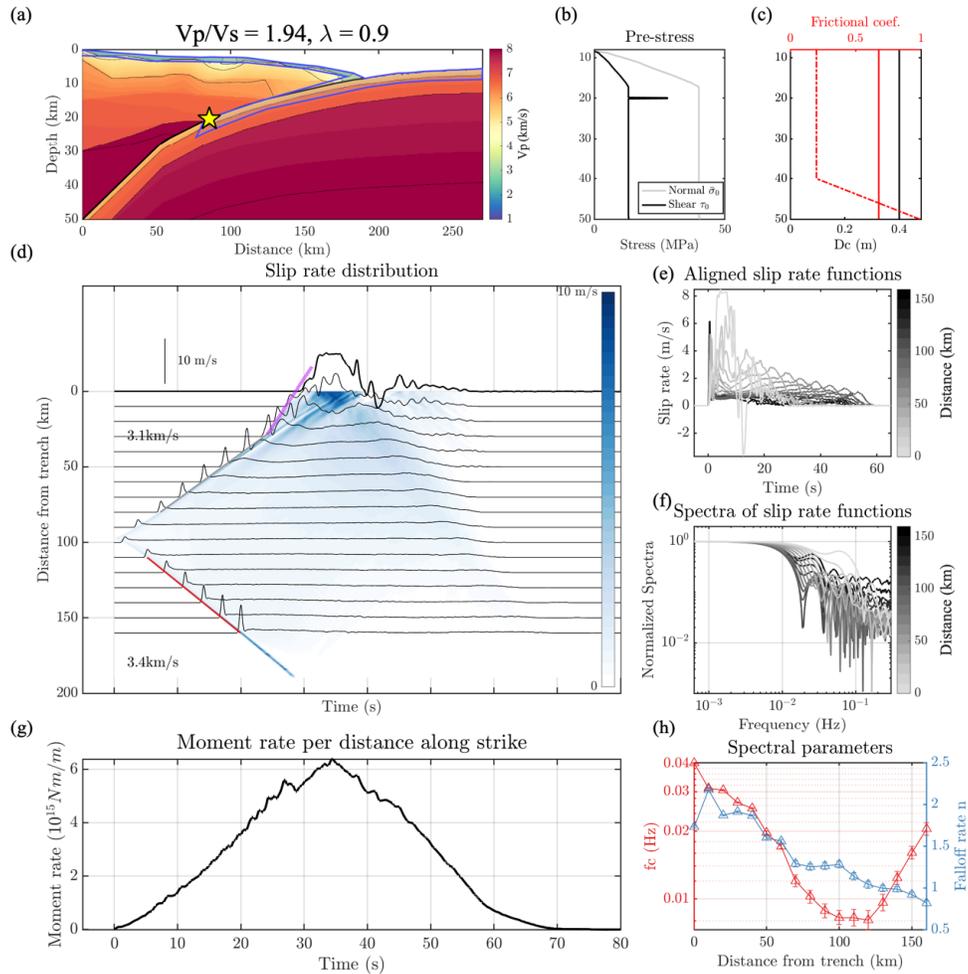


Figure C.9: Model settings and simulation results of model 9.

Model 10: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.04, \lambda = 0.9$

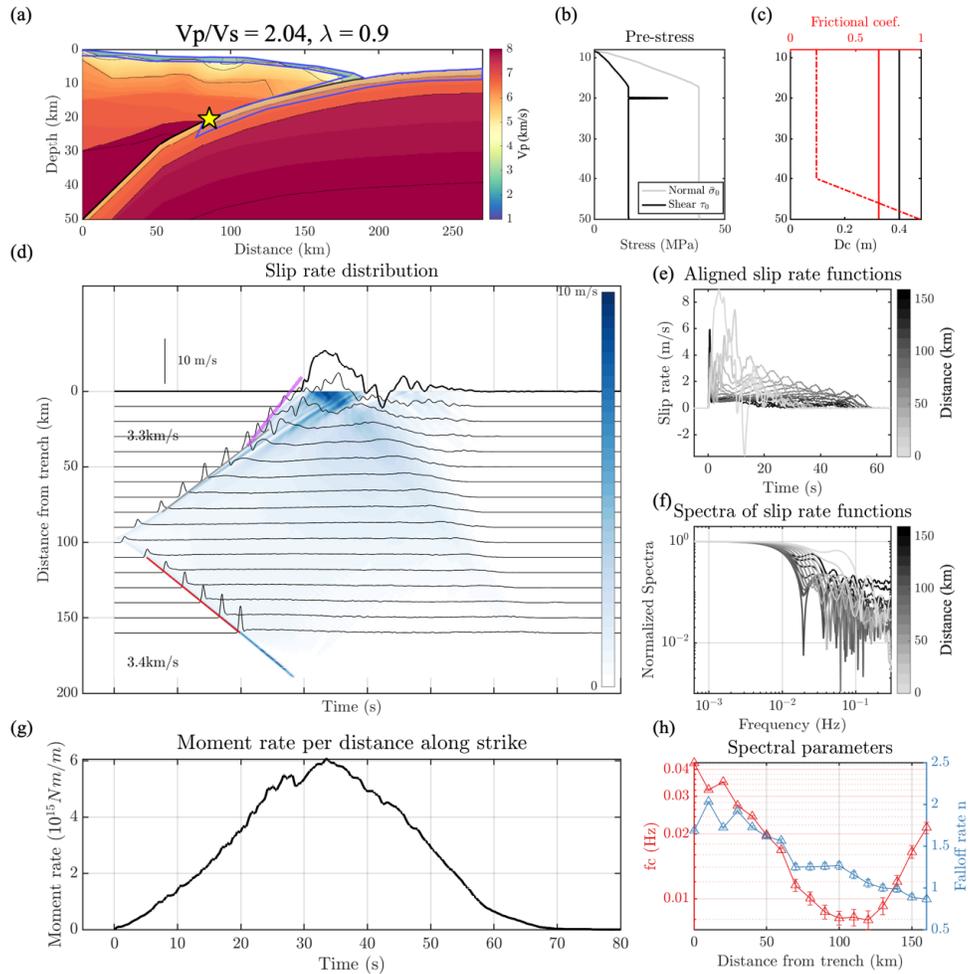


Figure C.10: Model settings and simulation results of model 10.

Model 11: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_S = 2.14, \lambda = 0.9$

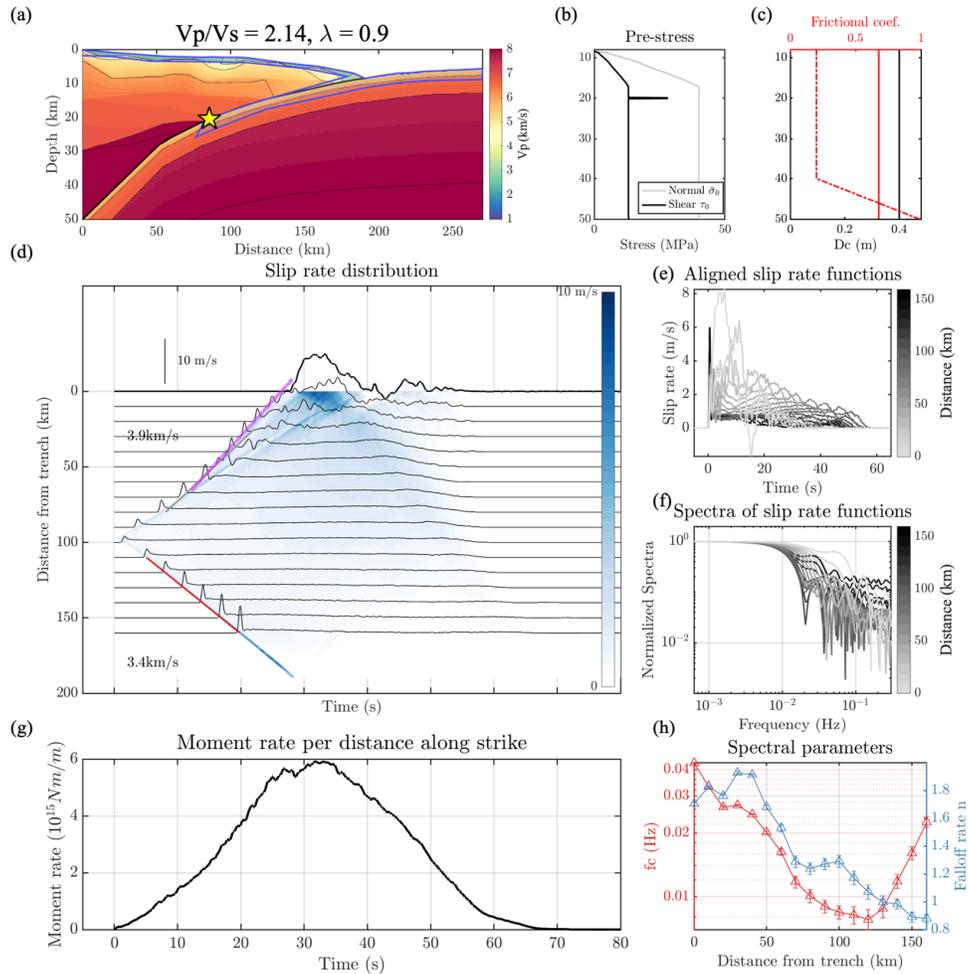


Figure C.11: Model settings and simulation results of model 11.

Model 12: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.24, \lambda = 0.9$

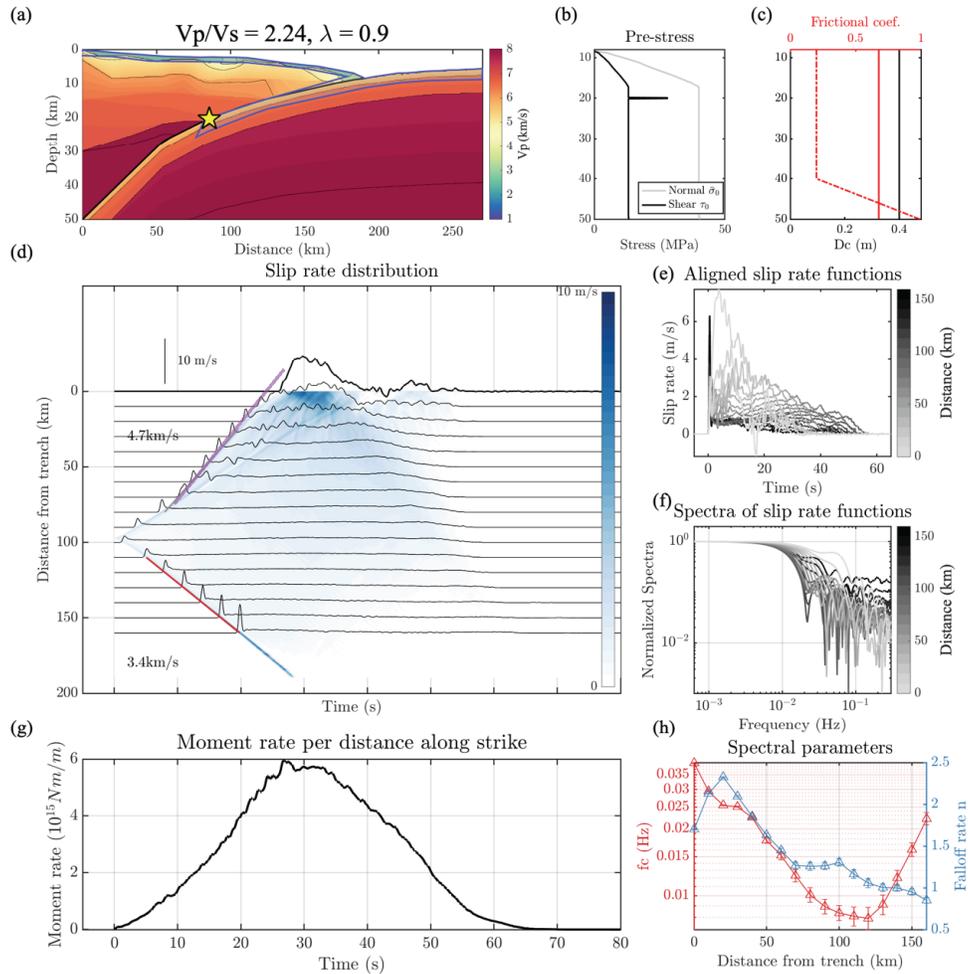


Figure C.12: Model settings and simulation results of model 12.

Model 13: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.35, \lambda = 0.9$

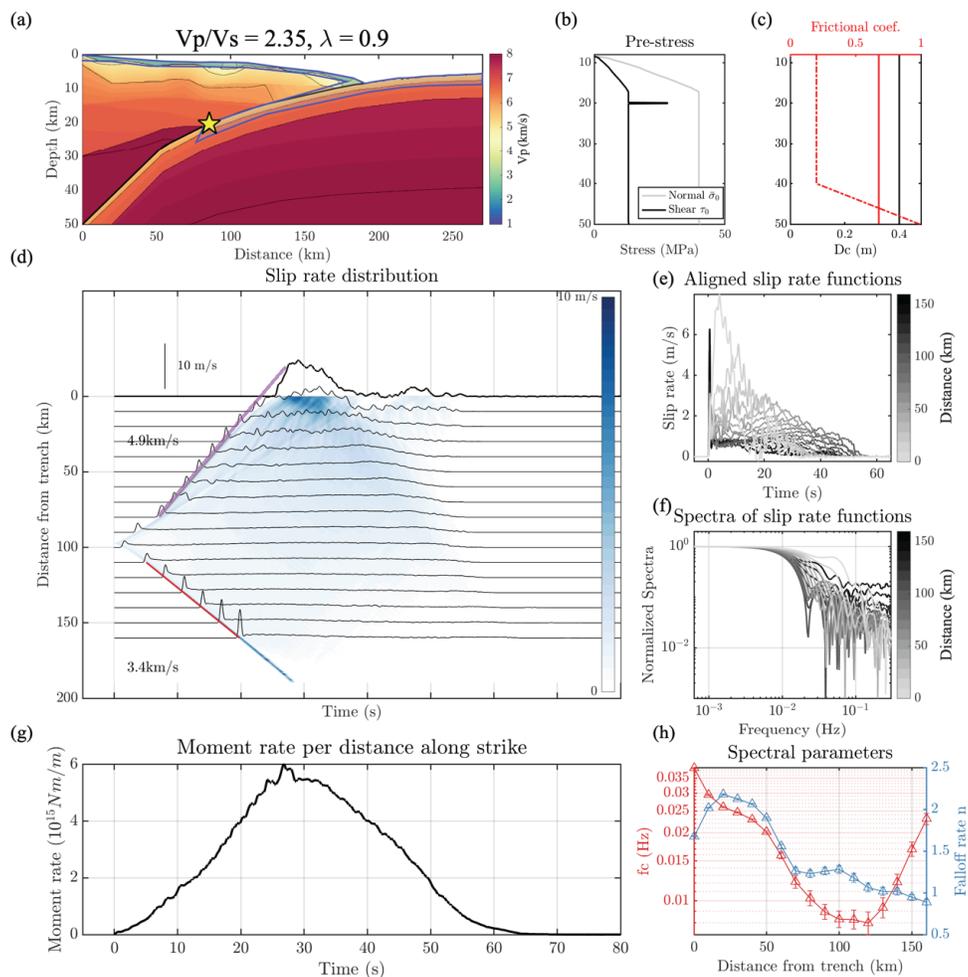


Figure C.13: Model settings and simulation results of model 13.

Model 14: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.45, \lambda = 0.9$

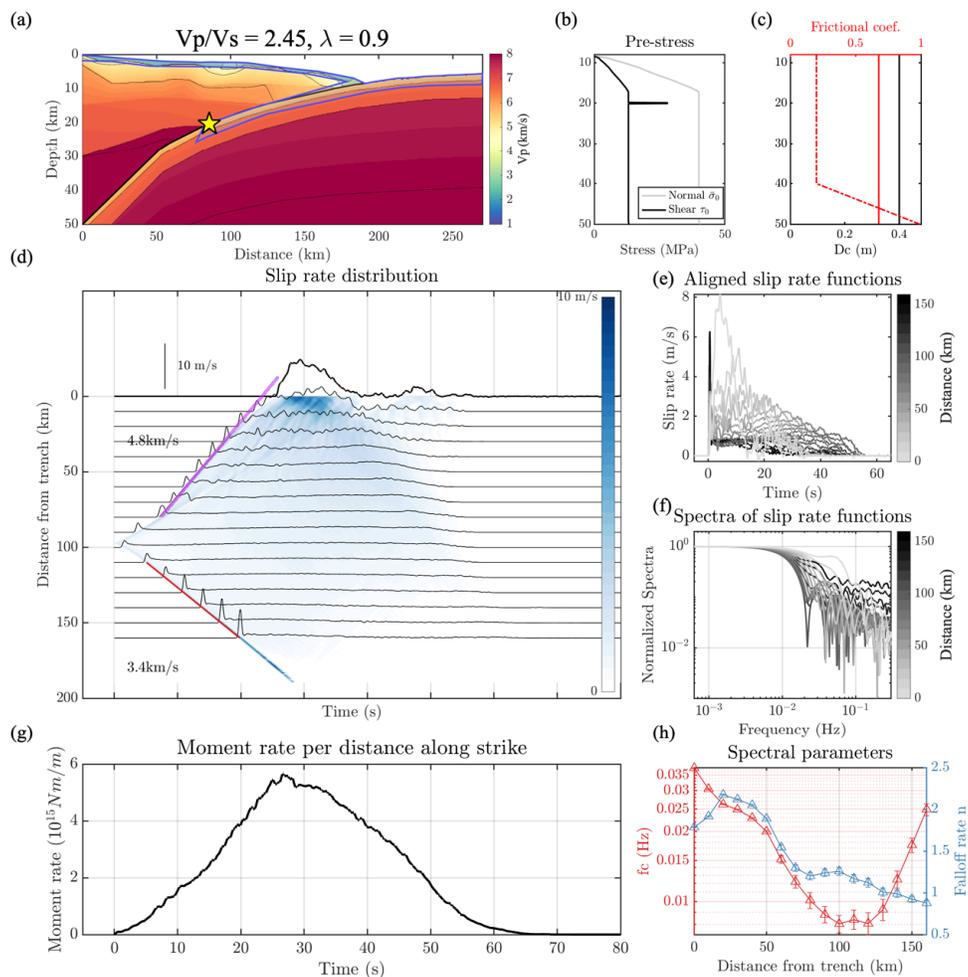


Figure C.14: Model settings and simulation results of model 14.

Model 15: Homogeneous media, planar fault, flat topography,  $\lambda = 0.7$

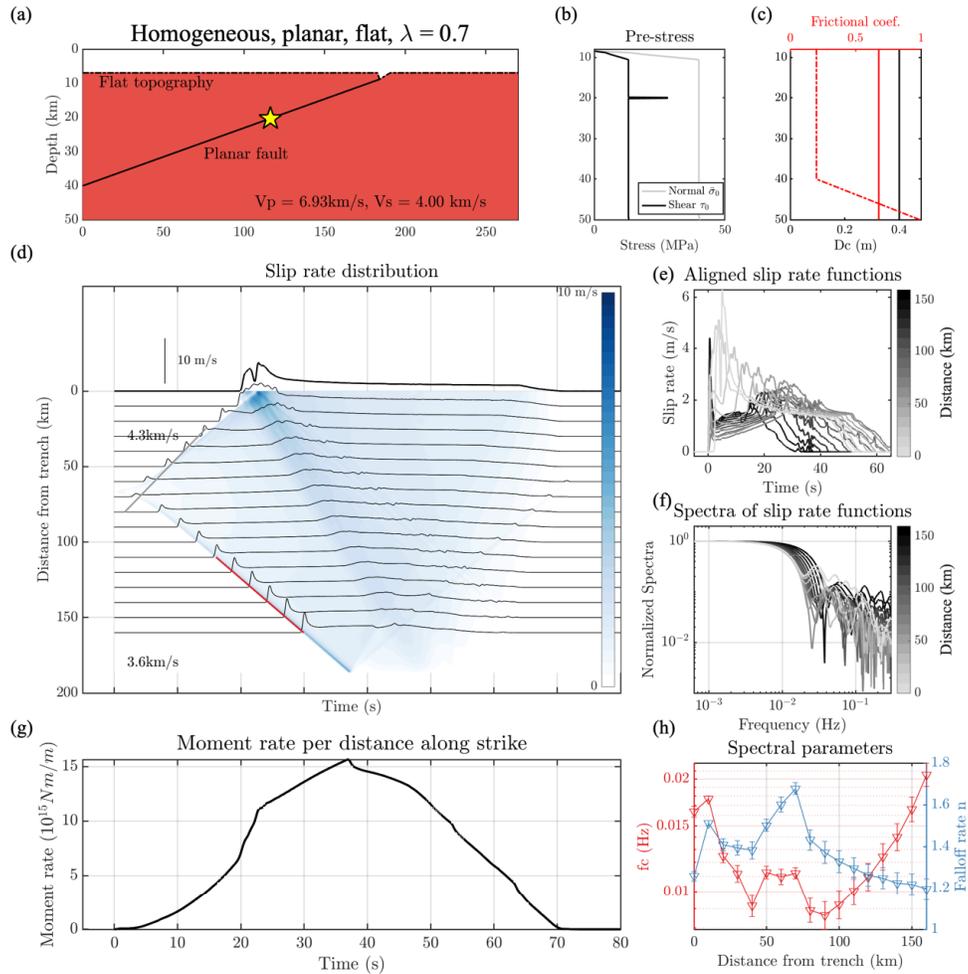


Figure C.15: Model settings and simulation results of model 15.

Model 16: Homogeneous media, curved fault, flat topography,  $\lambda = 0.7$

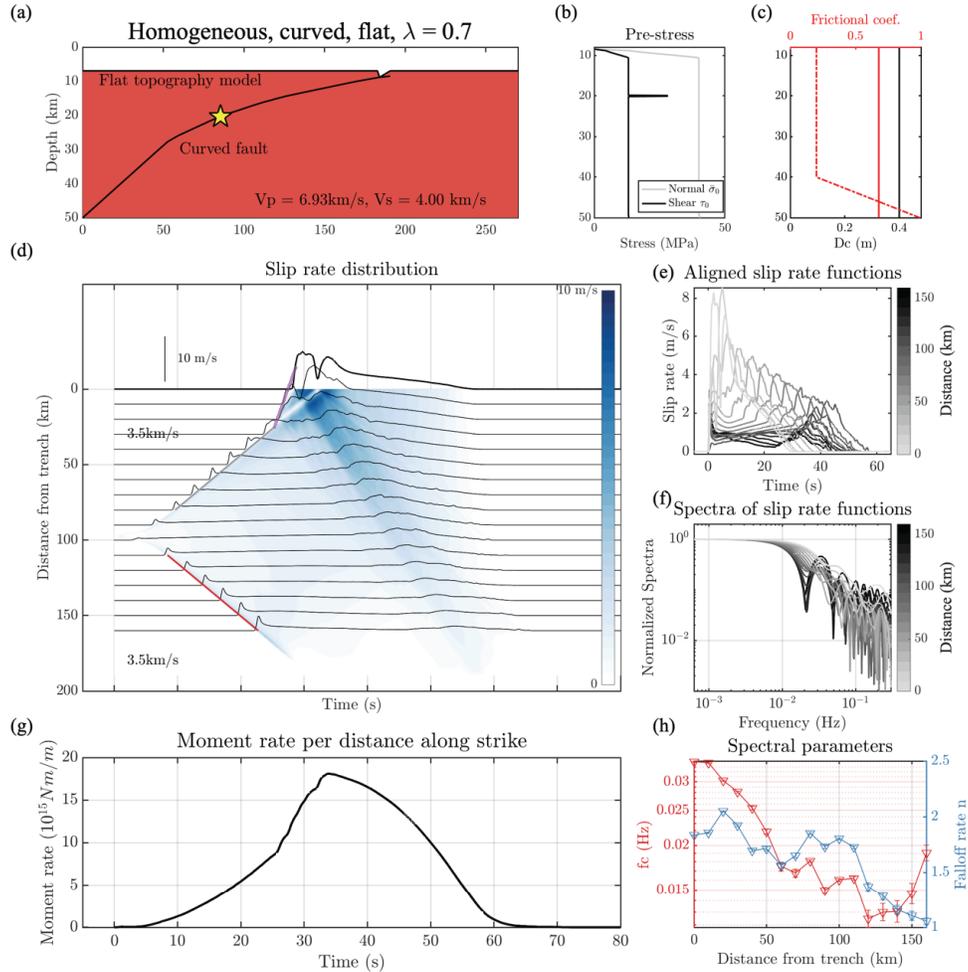


Figure C.16: Model settings and simulation results of model 16.

Model 17: Homogeneous media, curved fault, realistic topography,  $\lambda = 0.7$

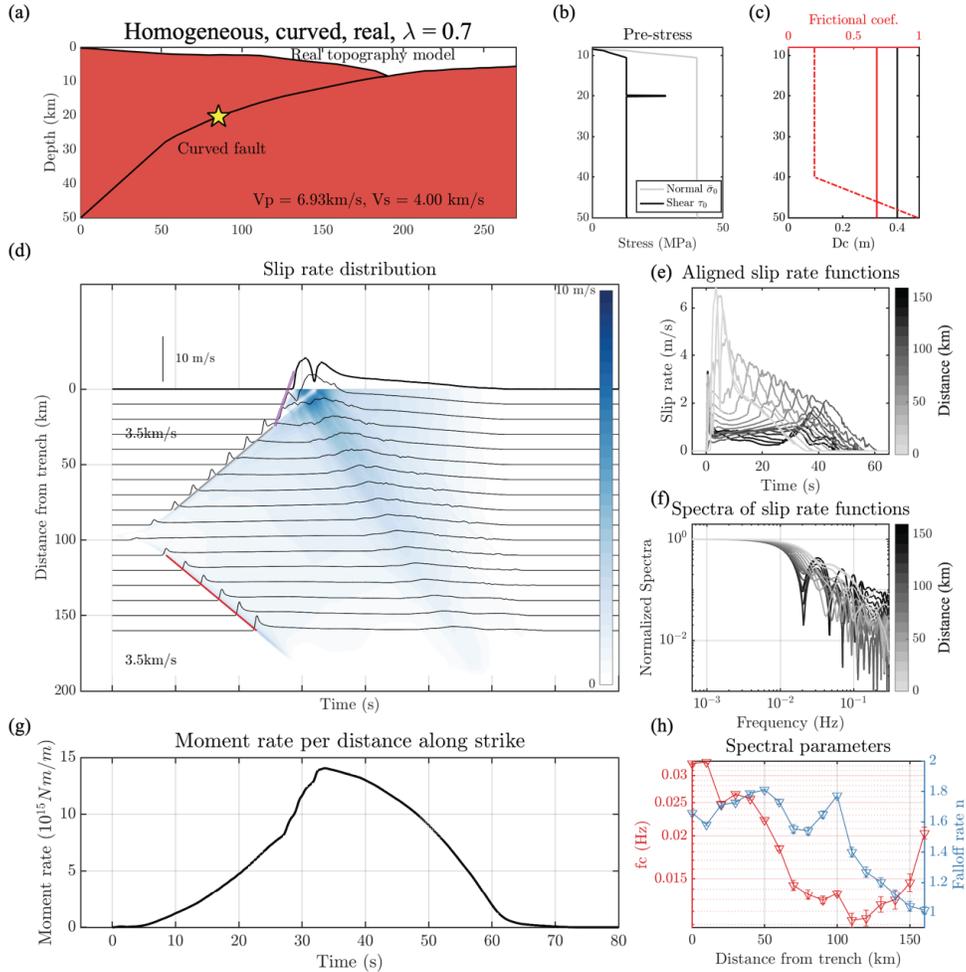


Figure C.17: Model settings and simulation results of model 17.

Model 18: Heterogeneous media, curved fault, realistic topography, shallow neutral friction,  $V_p/V_s = \sqrt{3}, \lambda = 0.7$

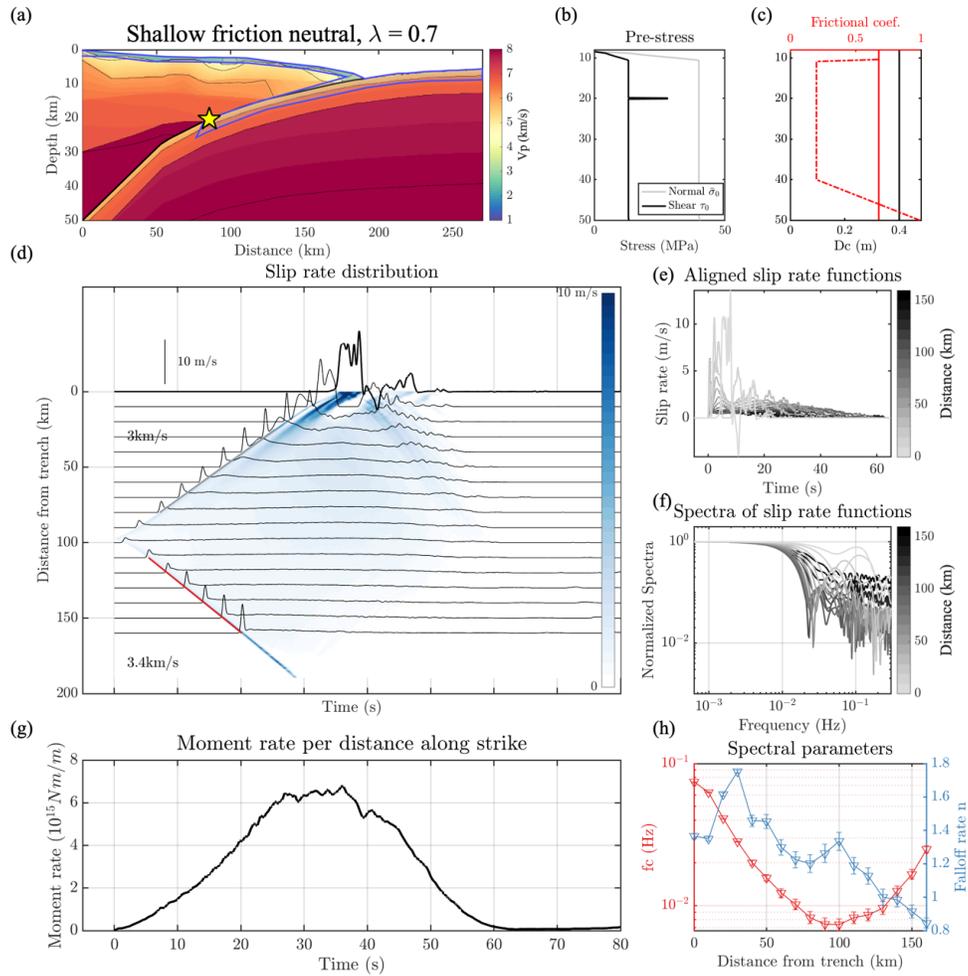


Figure C.18: Model settings and simulation results of model 18.

Model 19: Heterogeneous media, curved fault, realistic topography, shallow hardening/strengthening friction,  $V_p/V_s = \sqrt{3}$ ,  $\lambda = 0.7$

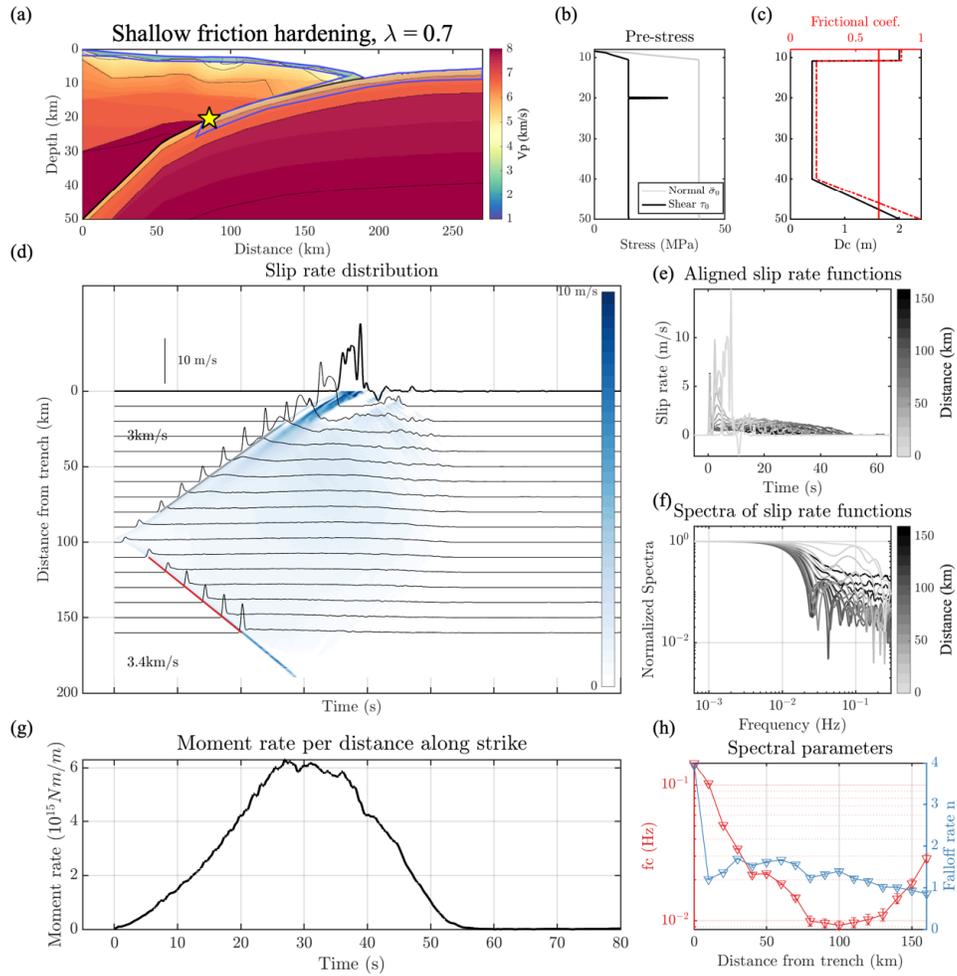


Figure C.19: Model settings and simulation results of model 19.

Model 20: Heterogeneous media, curved fault, realistic topography, exponential slip-weakening friction,  $V_p/V_s = \sqrt{3}, \lambda = 0.7$

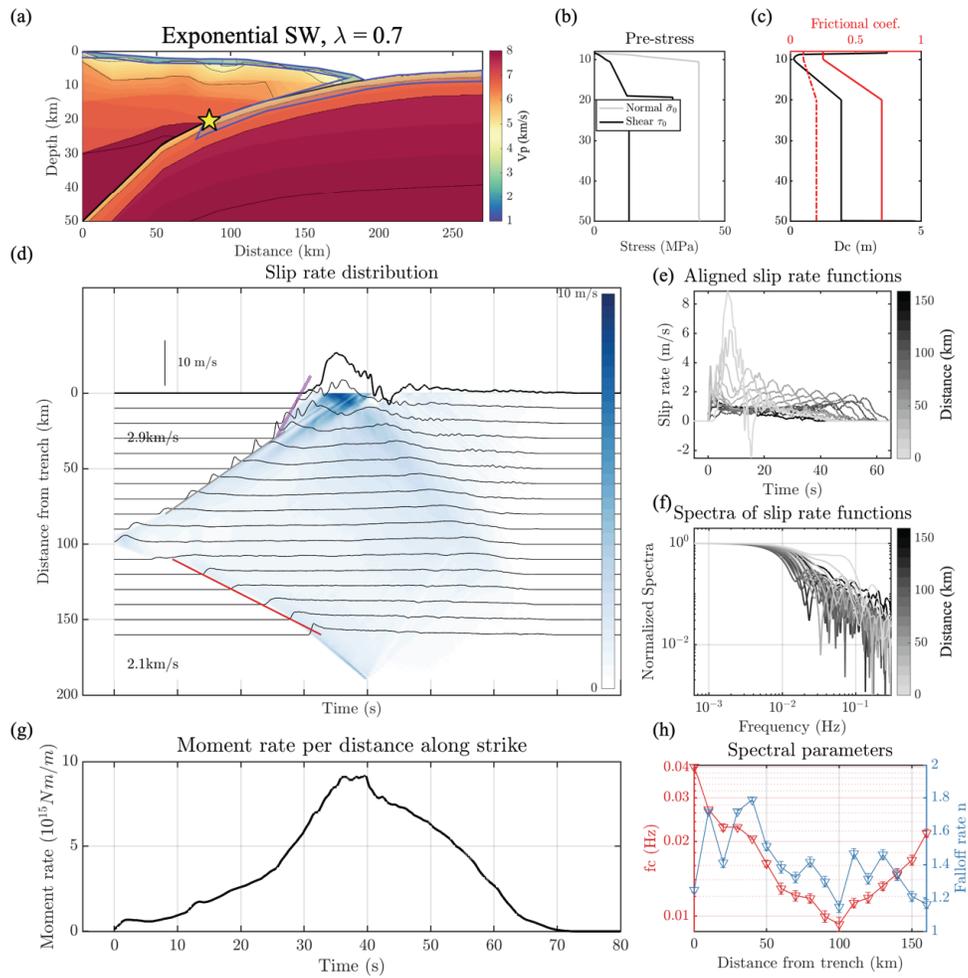


Figure C.20: Model settings and simulation results of model 20.

Model 21: Heterogeneous media, curved fault, realistic topography, slip-weakening friction,  
 $V_p/V_s = \sqrt{3}, \lambda = 0.7$

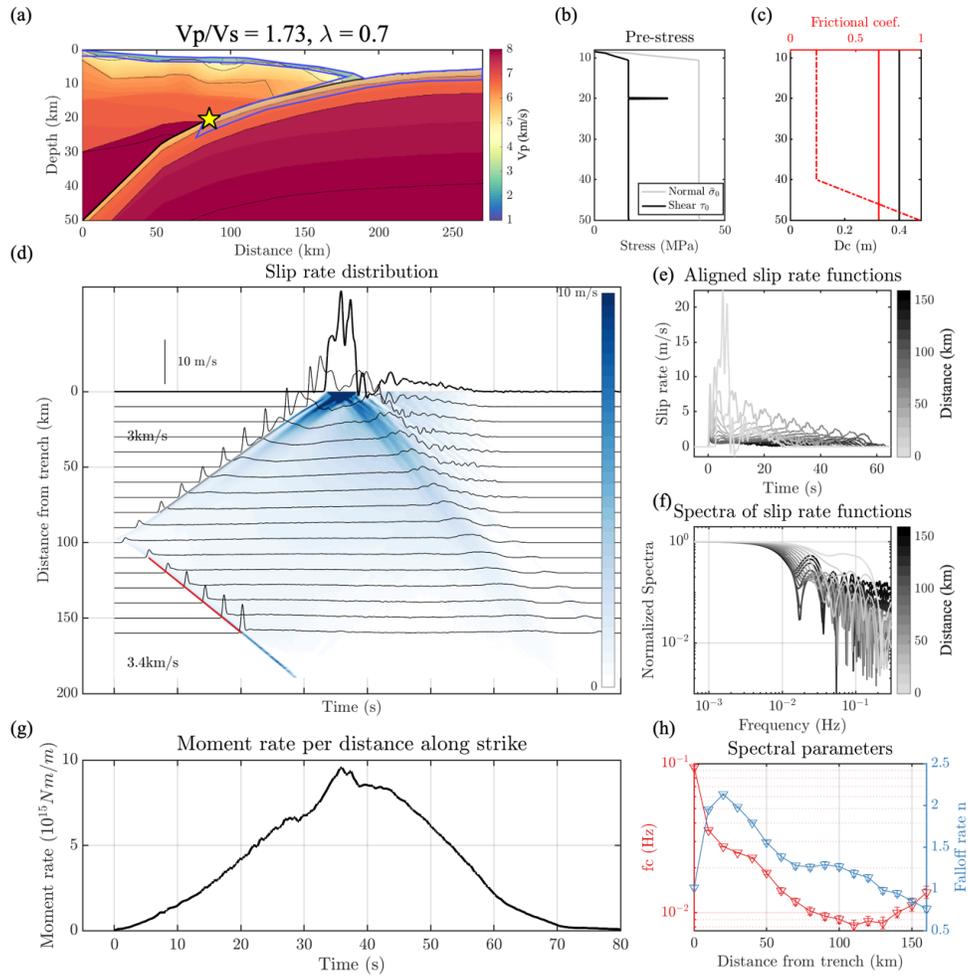


Figure C.21: Model settings and simulation results of model 21.

Model 22: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 1.83, \lambda = 0.7$

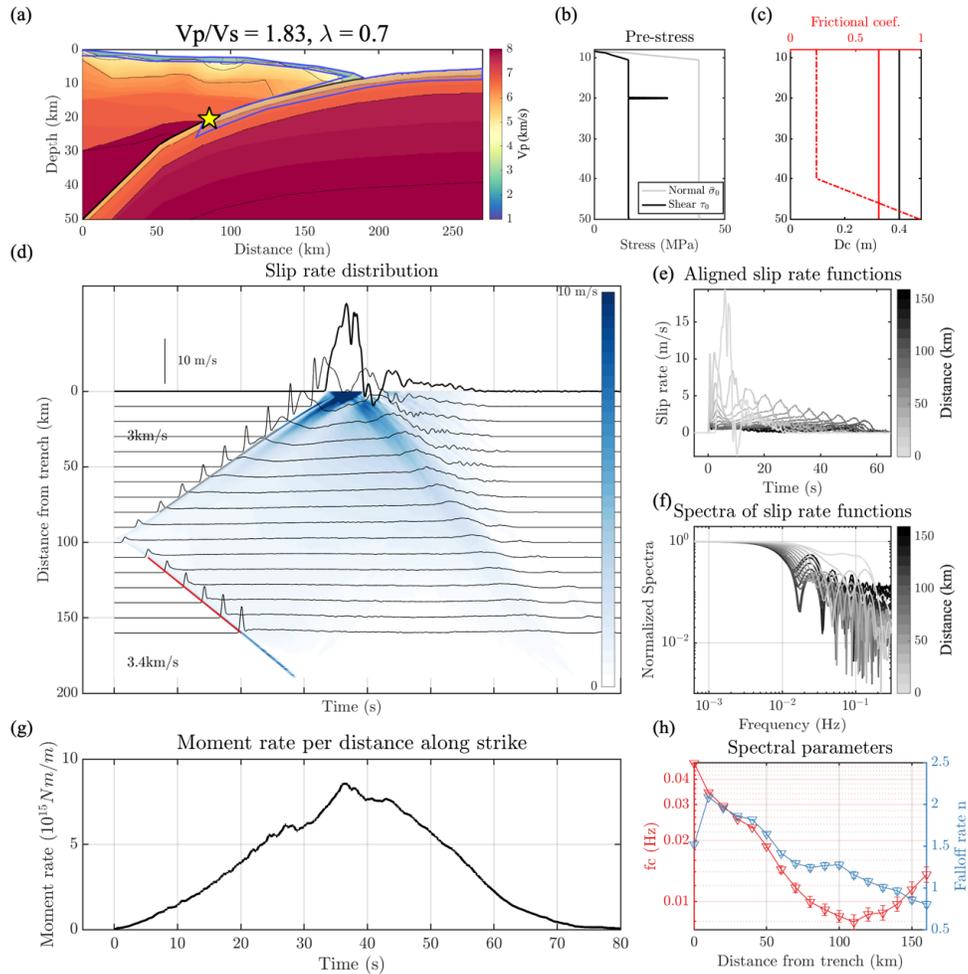


Figure C.22: Model settings and simulation results of model 22.

Model 23: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 1.94, \lambda = 0.7$

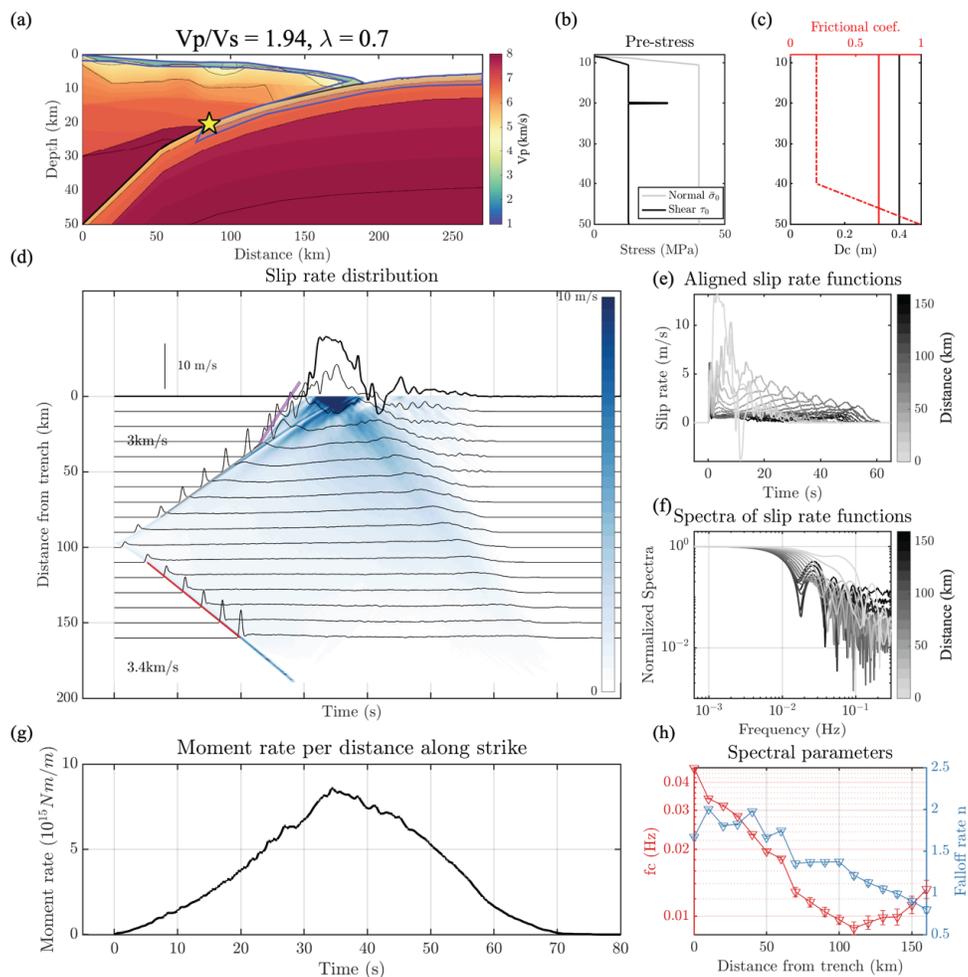


Figure C.23: Model settings and simulation results of model 23.

Model 24: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.04, \lambda = 0.7$

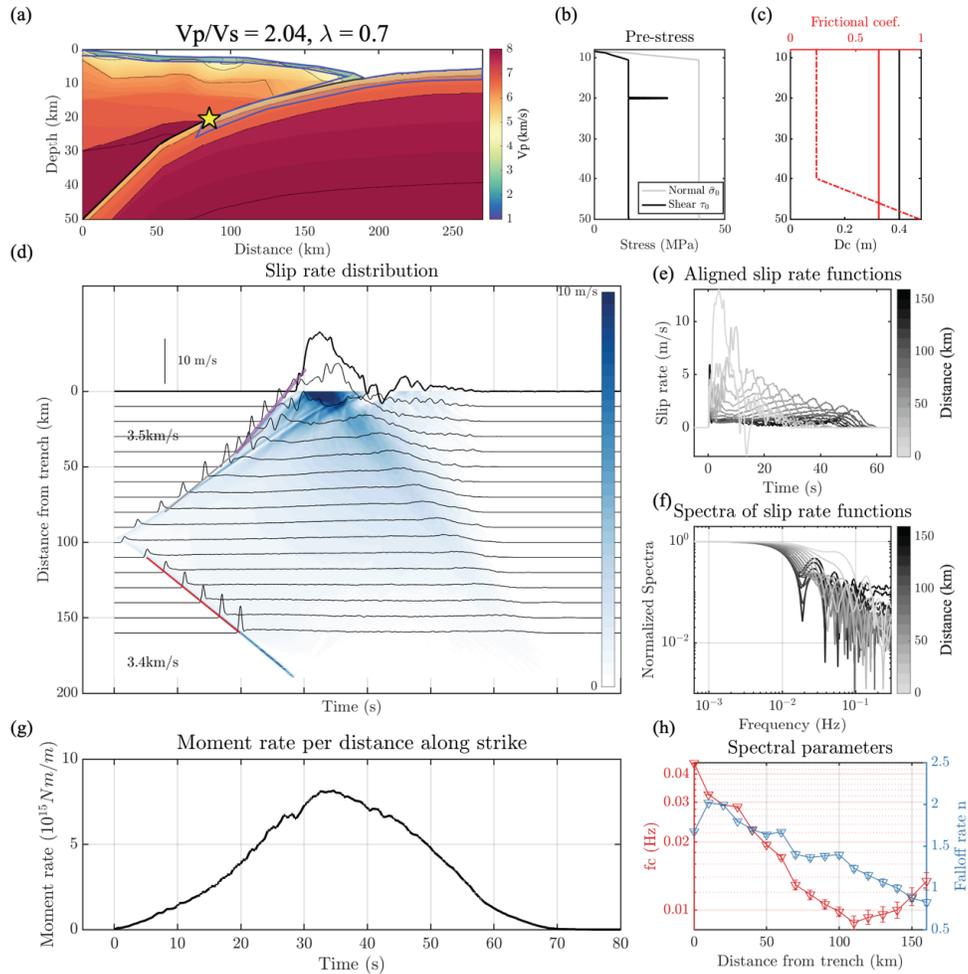


Figure C.24: Model settings and simulation results of model 24.

Model 25: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.14, \lambda = 0.7$

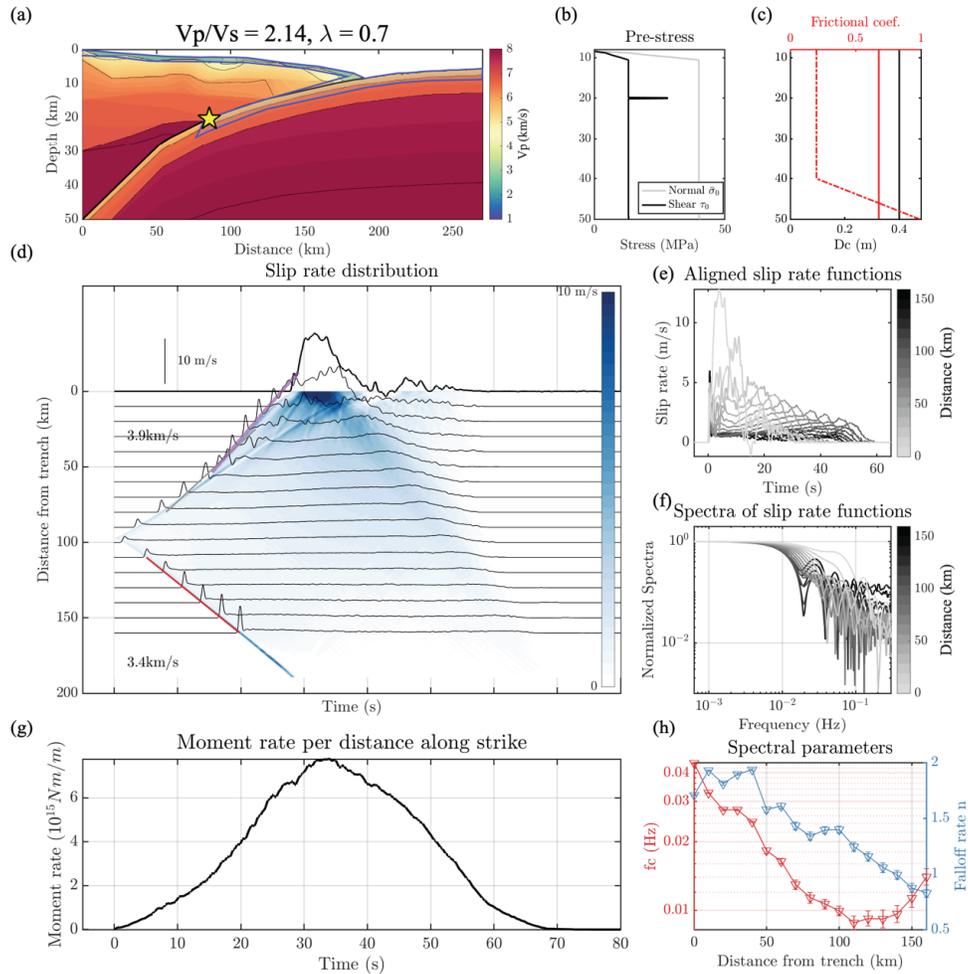


Figure C.25: Model settings and simulation results of model 25.

Model 26: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.24, \lambda = 0.7$

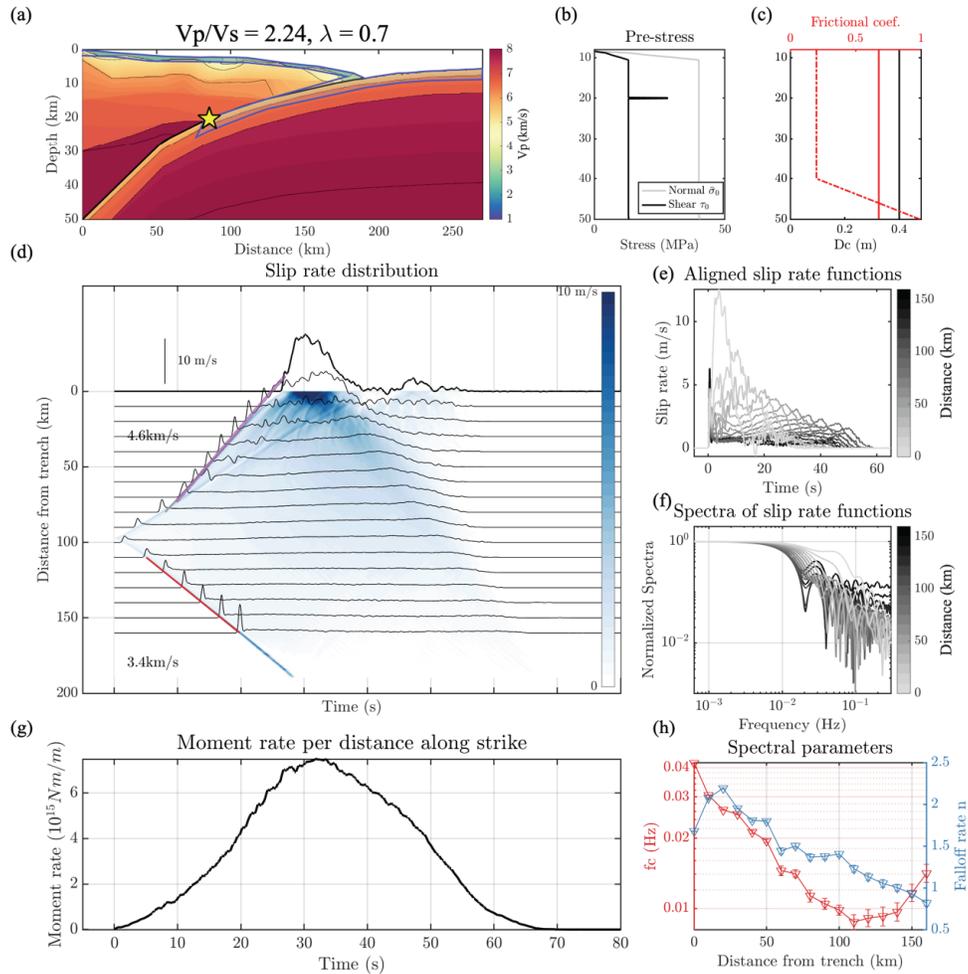


Figure C.26: Model settings and simulation results of model 26.

Model 27: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.35, \lambda = 0.7$

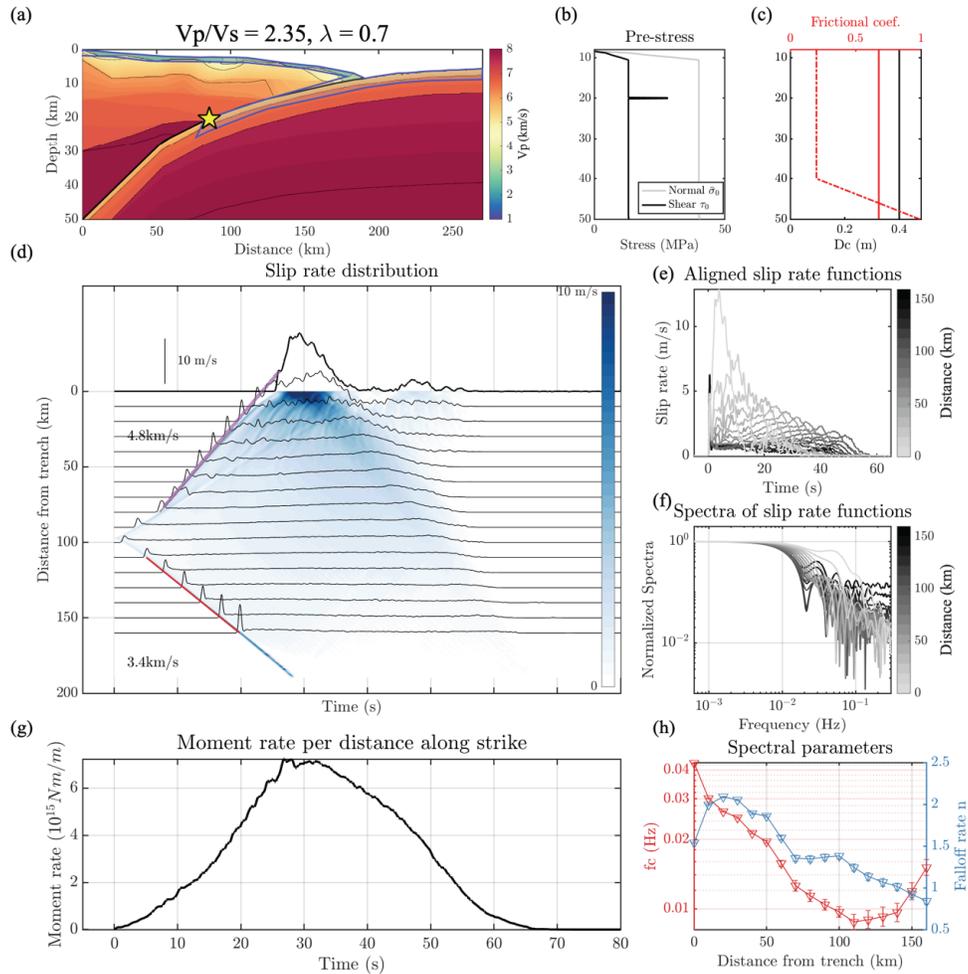


Figure C.27: Model settings and simulation results of model 27.

Model 28: Heterogeneous media, curved fault, realistic topography, linear slip-weakening friction,  $V_p/V_s = 2.45, \lambda = 0.7$

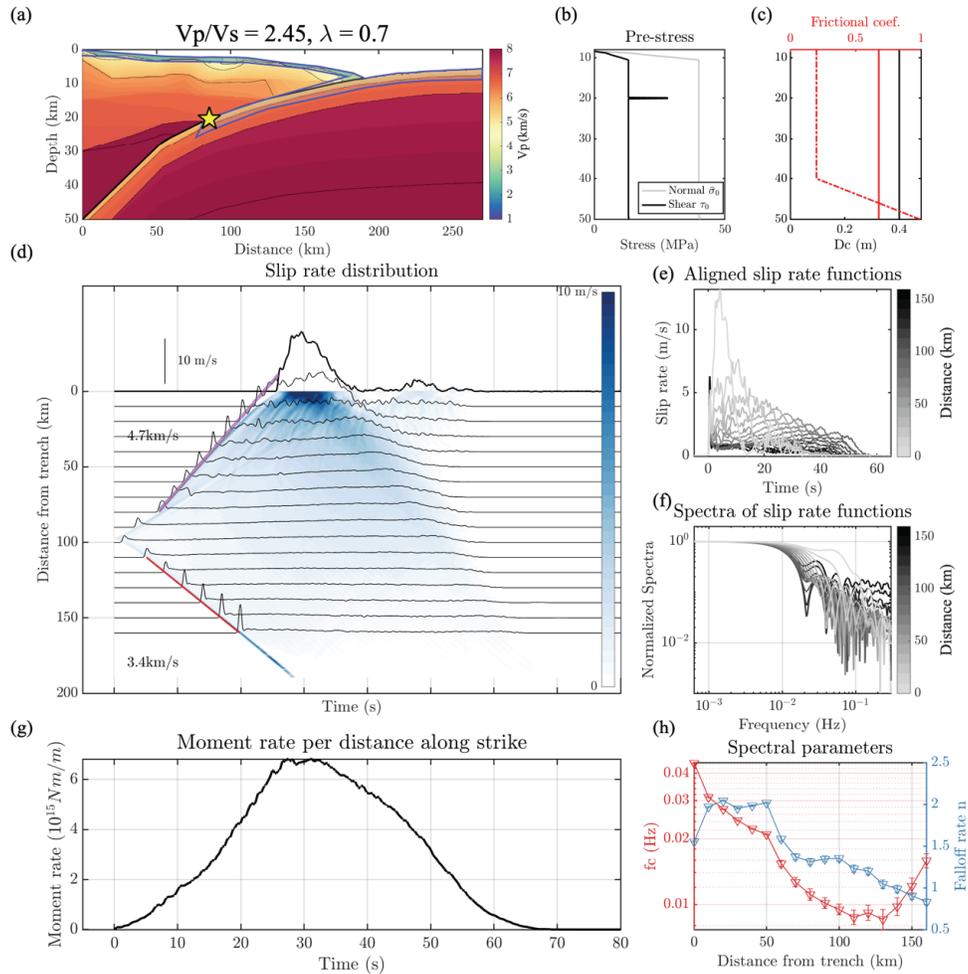


Figure C.28: Model settings and simulation results of model 28.

Model 29: Homogeneous full-space media, curved fault, linear slip-weakening friction,  $\lambda = 0.9$

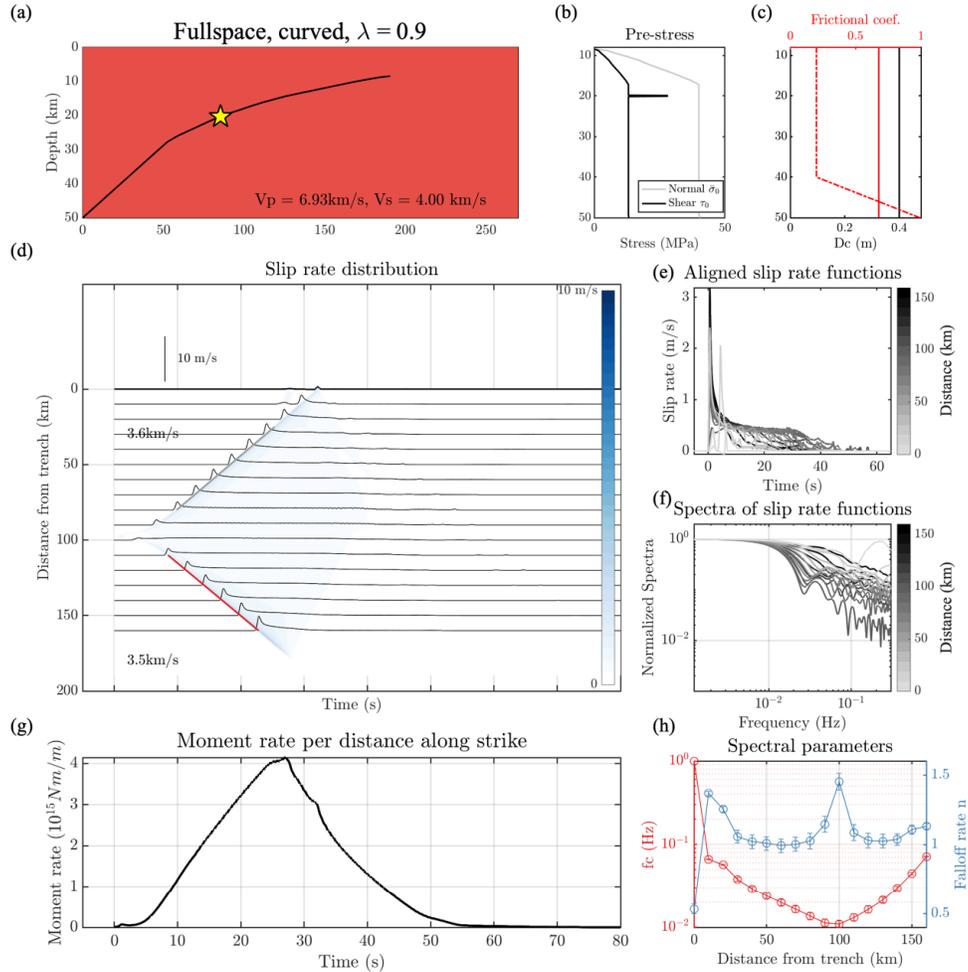


Figure C.29: Model settings and simulation results of model 29.

## References

- [1] (2007). Dynamic Time Warping. In M. Müller (Ed.), *Information Retrieval for Music and Motion* (pp. 69–84). Berlin, Heidelberg: Springer.
- [2] Abercrombie, R. E., Antolik, M., Felzer, K., & Ekström, G. (2001). The 1994 Java tsunami earthquake: Slip over a subducting seamount. *Journal of Geophysical Research: Solid Earth*, 106(B4), 6595–6607.
- [3] Abercrombie, R. E. & Rice, J. R. (2005). Can observations of earthquake scaling constrain slip weakening? *Geophysical Journal International*, 162(2), 406–424.
- [4] Agudelo, W., Ribodetti, A., Collot, J.-Y., & Operto, S. (2009). Joint inversion of multichannel seismic reflection and wide-angle seismic data: Improved imaging and refined velocity model of the crustal structure of the north Ecuador–south Colombia convergent margin. *Journal of Geophysical Research: Solid Earth*, 114(B2).
- [5] Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute*, 35, 415–456.
- [6] Aki, K. (1967). Scaling law of seismic spectrum. *Journal of Geophysical Research*, 72(4), 1217–1231.
- [7] Aki, K. & Richards, P. G. (2002). *Quantitative Seismology, 2nd Ed.*
- [8] Akuhara, T. & Mochizuki, K. (2015). Hydrous state of the subducting Philippine Sea plate inferred from receiver function image using onshore and offshore data. *Journal of Geophysical Research: Solid Earth*, 120(12), 2015JB012336.
- [9] Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bulletin of the seismological society of America*, 68(5), 1521–1532.
- [10] Allmann, B. P. & Shearer, P. M. (2009). Global variations of stress drop for moderate to large earthquakes. *Journal of Geophysical Research: Solid Earth*, 114(B1), B01310.
- [11] Ammon, C. J., Ji, C., Thio, H.-K., Robinson, D., Ni, S., Hjorleifsdottir, V., Kanamori, H., Lay, T., Das, S., Helmberger, D., Ichinose, G., Polet, J., & Wald, D. (2005). Rupture Process of the 2004 Sumatra-Andaman Earthquake. *Science*, 308(5725), 1133–1139.

- [12] Ammon, C. J., Kanamori, H., & Lay, T. (2008). A great earthquake doublet and seismic stress transfer cycle in the central Kuril islands. *Nature*, 451(7178), 561–565.
- [13] Ammon, C. J., Kanamori, H., Lay, T., & Velasco, A. A. (2006). The 17 July 2006 Java tsunami earthquake. *Geophysical Research Letters*, 33(24).
- [14] Ampuero, J. P. (2012). SEM2DPACK, a spectral element software for 2D seismic wave propagation and earthquake source dynamics, v2.3.8. More information at <https://github.com/jpampuero/sem2dpack>.
- [15] Ampuero, J.-P. & Ben-Zion, Y. (2008). Cracks, pulses and macroscopic asymmetry of dynamic rupture on a bimaterial interface with velocity-weakening friction. *Geophysical Journal International*, 173(2), 674–692.
- [16] Andrews, D. J. (1976). Rupture propagation with finite stress in antiplane strain. *Journal of Geophysical Research*, 81(20), 3575–3582.
- [17] Andrews, D. J. (1985). Dynamic plane-strain shear rupture with a slip-weakening friction law calculated by a boundary integral method. *Bulletin of the Seismological Society of America*, 75(1), 1–21.
- [18] Andrews, D. J. & Ben-Zion, Y. (1997). Wrinkle-like slip pulse on a fault between different materials. *Journal of Geophysical Research: Solid Earth*, 102(B1), 553–571.
- [19] Asano, K. & Iwata, T. (2012). Source model for strong ground motion generation in the frequency range 0.1–10 Hz during the 2011 Tohoku earthquake. *Earth, Planets and Space*, 64(12), 6.
- [20] Atik, L. A., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., & Kuehn, N. (2010). The Variability of Ground-Motion Prediction Models and Its Components. *Seismological Research Letters*, 81(5), 794–801.
- [21] Avouac, J.-P. (2015). From Geodetic Imaging of Seismic and Aseismic Fault Slip to Dynamic Modeling of the Seismic Cycle. *Annual Review of Earth and Planetary Sciences*, 43(1), 233–271.
- [22] Avouac, J.-P., Meng, L., Wei, S., Wang, T., & Ampuero, J.-P. (2015). Lower edge of locked Main Himalayan Thrust unzipped by the 2015 Gorkha earthquake. *Nature Geoscience*, 8(9), 708–711.
- [23] Badrinarayanan, V., Handa, A., & Cipolla, R. (2015). Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *arXiv preprint arXiv:1505.07293*.

- [24] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12), 2481–2495.
- [25] Baltay, A. S., Beroza, G. C., & Ide, S. (2014). Radiated Energy of Great Earthquakes from Teleseismic Empirical Green’s Function Deconvolution. *Pure and Applied Geophysics*, 171(10), 2841–2862.
- [26] Bappy, J. H., Simons, C., Nataraj, L., Manjunath, B., & Roy-Chowdhury, A. K. (2019). Hybrid lstm and encoder–decoder architecture for detection of image forgeries. *IEEE Transactions on Image Processing*, 28(7), 3286–3300.
- [27] Bassett, D., Kopp, H., Sutherland, R., Henrys, S., Watts, A. B., Timm, C., Scherwath, M., Grevemeyer, I., & Ronde, C. E. J. d. (2016). Crustal structure of the Kermadec arc from MANGO seismic refraction profiles. *Journal of Geophysical Research: Solid Earth*, 121(10), 7514–7546.
- [28] Bassett, D., Sutherland, R., Henrys, S., Stern, T., Scherwath, M., Benson, A., Toulmin, S., & Henderson, M. (2010). Three-dimensional velocity structure of the northern Hikurangi margin, Raukumara, New Zealand: Implications for the growth of continental crust by subduction erosion and tectonic underplating. *Geochemistry, Geophysics, Geosystems*, 11(10).
- [29] Ben-Zion, Y. & Sammis, C. G. (2003). Characterization of fault zones. *Pure and applied geophysics*, 160(3), 677–715.
- [30] Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., Shapiro, N., & Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3), 1239–1260.
- [31] Bergen, K. J., Johnson, P. A., Hoop, M. V. d., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363(6433), eaau0323.
- [32] Bernard, P. & Madariaga, R. (1984). High-frequency seismic radiation from a buried circular fault. *Geophysical Journal of the Royal Astronomical Society*, 78(1), 1–17.
- [33] Berndt, D. J. & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *KDD workshop*, volume 10 (pp. 359–370): Seattle, WA.
- [34] Beroza, G. C., Segou, M., & Mostafa Mousavi, S. (2021). Machine learning and earthquake forecasting—next steps. *Nature Communications*, 12(1), 1–3.
- [35] Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3), 530–533.

- [36] Bilek, S. L. & Lay, T. (1999a). Rigidity variations with depth along interplate megathrust faults in subduction zones. *Nature*, 400(6743), 443–446.
- [37] Bilek, S. L. & Lay, T. (1999b). Rigidity variations with depth along interplate megathrust faults in subduction zones. *Nature*, 400(6743), 443–446.
- [38] Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, 4(3).
- [39] Boatwright, J. & Choy, G. L. (1986). Teleseismic estimates of the energy radiated by shallow earthquakes. *Journal of Geophysical Research: Solid Earth*, 91(B2), 2095–2112.
- [40] Bostock, M. G. (2013). The Moho in subduction zones. *Tectonophysics*, 609, 547–557.
- [41] Boyd, S., Boyd, S. P., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.
- [42] Brocher, T. M. (2005a). Empirical Relations between Elastic Wavespeeds and Density in the Earth's Crust. *Bulletin of the Seismological Society of America*, 95(6), 2081–2092.
- [43] Brocher, T. M. (2005b). Empirical Relations between Elastic Wavespeeds and Density in the Earth's Crust. *Bulletin of the Seismological Society of America*, 95(6), 2081–2092.
- [44] Brown, L., Wang, K., & Sun, T. (2015). Static stress drop in the Mw 9 Tohoku-oki earthquake: Heterogeneous distribution and low average value. *Geophysical Research Letters*, 42(24), 2015GL066361.
- [45] Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research*, 75(26), 4997–5009.
- [46] Brune, J. N. (1971). Correction (to Brune, 1970). *J. geophys. Res*, 76, 5002.
- [47] Brune, J. N. (1996). Particle motions in a physical model of shallow angle thrust faulting. *Proceedings of the Indian Academy of Sciences - Earth and Planetary Sciences*, 105(2), 197.
- [48] Burridge, R. & Knopoff, L. (1964). Body force equivalents for seismic dislocations. *Bulletin of the Seismological Society of America*, 54(6A), 1875–1888.
- [49] Burtin, A., Bollinger, L., Vergne, J., Cattin, R., & Nábělek, J. (2008). Spectral analysis of seismic noise induced by rivers: A new tool to monitor spatiotemporal changes in stream hydrodynamics. *Journal of Geophysical Research: Solid Earth*, 113(B5).
- [50] Campillo, M., Gariel, J., Aki, K., & Sanchez-Sesma, F. (1989). Destructive strong ground motion in Mexico city: Source, path, and site effects during great 1985 Michoacán earthquake. *Bulletin of the Seismological Society of America*, 79(6), 1718–1735.

- [51] Campillo, M. & Paul, A. (2003). Long-Range Correlations in the Diffuse Seismic Coda. *Science*, 299(5606), 547–549.
- [52] Candela, T., Renard, F., Klinger, Y., Mair, K., Schmittbuhl, J., & Brodsky, E. E. (2012). Roughness of fault surfaces over nine decades of length scales. *Journal of Geophysical Research: Solid Earth*, 117(B8).
- [53] Candes, E. J., Romberg, J. K., & Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8), 1207–1223.
- [54] Cesca, S., Zhang, Y., Mouslopoulou, V., Wang, R., Saul, J., Savage, M., Heimann, S., Kufner, S. K., Oncken, O., & Dahm, T. (2017). Complex rupture process of the Mw 7.8, 2016, Kaikoura earthquake, New Zealand, and its aftershock sequence. *Earth and Planetary Science Letters*, 478, 110–120.
- [55] Cessaro, R. K. (1994). Sources of primary and secondary microseisms. *Bulletin of the Seismological Society of America*, 84(1), 142–148.
- [56] Chang, S. G., Yu, B., & Vetterli, M. (2000). Adaptive wavelet thresholding for image denoising and compression. *IEEE transactions on image processing*, 9(9), 1532–1546.
- [57] Chen, Y., Ma, J., & Fomel, S. (2016). Double-sparsity dictionary for seismic noise attenuation. *Geophysics*, 81(2), V103–V116.
- [58] Chounet, A. & Vallée, M. (2018). Global and interregion characterization of subduction interface earthquakes derived from source time functions properties. *Journal of Geophysical Research: Solid Earth*, 123(7), 5831–5852.
- [59] Chounet, A., Vallée, M., Causse, M., & Courboux, F. (2018). Global catalog of earthquake rupture velocities shows anticorrelation between stress drop and rupture velocity. *Tectonophysics*, 733, 148–158.
- [60] Chu, R., Wei, S., Helmberger, D. V., Zhan, Z., Zhu, L., & Kanamori, H. (2011). Initiation of the great Mw 9.0 Tohoku–Oki earthquake. *Earth and Planetary Science Letters*, 308(3), 277–283.
- [61] Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *GEOPHYSICS*, 33(2), 264–269. Publisher: Society of Exploration Geophysicists.
- [62] Clayton, R. W. (2020). Imaging the Subsurface with Ambient Noise Autocorrelations. *Seismological Research Letters*, 91(2A), 930–935.
- [63] Cochard, A. & Rice, J. R. (2000). Fault rupture between dissimilar materials: Ill-posedness, regularization, and slip-pulse response. *Journal of Geophysical Research: Solid Earth*, 105(B11), 25891–25907.

- [64] Contreras-Reyes, E., Becerra, J., Kopp, H., Reichert, C., & Díaz-Naveas, J. (2014). Seismic structure of the north-central Chilean convergent margin: Subduction erosion of a paleo-magmatic arc. *Geophysical Research Letters*, 41(5), 1523–1529.
- [65] Contreras-Reyes, E., Grevemeyer, I., Flueh, E. R., & Reichert, C. (2008). Upper lithospheric structure of the subduction zone offshore of southern Arauco peninsula, Chile, at ~38°S. *Journal of Geophysical Research: Solid Earth*, 113(B7).
- [66] Contreras-Reyes, E., Grevemeyer, I., Watts, A. B., Flueh, E. R., Peirce, C., Moeller, S., & Papenberg, C. (2011). Deep seismic structure of the Tonga subduction zone: Implications for mantle hydration, tectonic erosion, and arc magmatism. *Journal of Geophysical Research: Solid Earth*, 116(B10).
- [67] Convers, J. A. & Newman, A. V. (2011). Global evaluation of large earthquake energy from 1997 through mid-2010. *J. Geophys. Res. Solid Earth*.
- [68] Crempien, J. G. F. & Archuleta, R. J. (2014). UCSB Method for Simulation of Broadband Ground Motion from Kinematic Earthquake Sources. *Seismological Research Letters*.
- [69] Dalguer, L. A. & Day, S. M. (2009). Asymmetric rupture of large aspect-ratio faults at bimaterial interface in 3D. *Geophysical Research Letters*, 36(23).
- [70] Danré, P., Yin, J., Lipovsky, B. P., & Denolle, M. A. (2019). Earthquakes Within Earthquakes: Patterns in Rupture Complexity. *Geophysical Research Letters*, 46(13), 7352–7360.
- [71] Day, S. M. (1982). Three-dimensional simulation of spontaneous rupture: The effect of nonuniform prestress. *Bulletin of the Seismological Society of America*, 72(6A), 1881–1902.
- [72] Day, S. M., Dalguer, L. A., Lapusta, N., & Liu, Y. (2005). Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture. *Journal of Geophysical Research: Solid Earth*, 110(B12).
- [73] De Plaen, R. S., Lecocq, T., Caudron, C., Ferrazzini, V., & Francis, O. (2016). Single-station monitoring of volcanoes using seismic ambient noise. *Geophysical Research Letters*, 43(16), 8511–8518.
- [74] Denolle, M. A. (2019). Energetic Onset of Earthquakes. *Geophysical Research Letters*, 46(5), 2458–2466.
- [75] Denolle, M. A., Fan, W., & Shearer, P. M. (2015). Dynamics of the 2015 M7.8 Nepal earthquake. *Geophysical Research Letters*, 42(18), 2015GL065336.
- [76] Denolle, M. A. & Shearer, P. M. (2016). New perspectives on self-similarity for shallow thrust earthquakes. *Journal of Geophysical Research: Solid Earth*, 121(9), 2016JB013105.

- [77] Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research: Solid Earth*, 99(B2), 2601–2618.
- [78] Dieterich, J. H. (1992). Earthquake nucleation on faults with rate-and state-dependent strength. *Tectonophysics*, 211(1), 115–134.
- [79] Ding, M. & Lin, J. (2014). Post-seismic viscoelastic deformation and stress transfer after the 1960 M9.5 Valdivia, Chile earthquake: effects on the 2010 M8.8 Maule, Chile earthquake. *Geophysical Journal International*, 197(2), 697–704.
- [80] Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4), 1289–1306.
- [81] Donoho, D. L. & Johnstone, J. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *biometrika*, 81(3), 425–455.
- [82] Douglas, A. (1997). Bandpass filtering to reduce noise on seismograms: is there a better way? *Bulletin of the Seismological Society of America*, 87(3), 770–777.
- [83] Draganov, D., Campman, X., Thorbecke, J., Verdel, A., & Wapenaar, K. (2009). Reflection images from ambient seismic noise. *GEOPHYSICS*, 74(5), A63–A67. Publisher: Society of Exploration Geophysicists.
- [84] Duan, B. & Oglesby, D. D. (2005). Multicycle dynamics of nonplanar strike-slip faults. *Journal of Geophysical Research: Solid Earth*, 110(B3), B03304.
- [85] Duan, B. & Oglesby, D. D. (2007). Nonuniform prestress from prior earthquakes and the effect on dynamics of branched fault systems. *Journal of Geophysical Research: Solid Earth*, 112(B5), B05308.
- [86] Dunham, E. M. (2007). Conditions governing the occurrence of supershear ruptures under slip-weakening friction. *Journal of Geophysical Research: Solid Earth*, 112(B7).
- [87] Dunham, E. M., Belanger, D., Cong, L., & Kozdon, J. E. (2011). Earthquake Ruptures with Strongly Rate-Weakening Friction and Off-Fault Plasticity, Part 2: Nonplanar Faults. *Bulletin of the Seismological Society of America*, 101(5), 2308–2322.
- [88] Duputel, Z. & Rivera, L. (2017). Long-period analysis of the 2016 Kaikoura earthquake. *Physics of the Earth and Planetary Interiors*, 265, 62–66.
- [89] Dziewonski, A. M. & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25(4), 297–356.

- [90] Dziewonski, A. M., Chou, T.-A., & Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research: Solid Earth*, 86(B4), 2825–2852.
- [91] Eberhart-Phillips, D., Haessler, P. J., Freymueller, J. T., Frankel, A. D., Rubin, C. M., Craw, P., Ratchkovski, N. A., Anderson, G., Carver, G. A., Crone, A. J., Dawson, T. E., Fletcher, H., Hansen, R., Harp, E. L., Harris, R. A., Hill, D. P., Hreinsdóttir, S., Jibson, R. W., Jones, L. M., Kayen, R., Keefer, D. K., Larsen, C. F., Moran, S. C., Personius, S. F., Plafker, G., Sherrod, B., Sieh, K., Sitar, N., & Wallace, W. K. (2003). The 2002 Denali Fault Earthquake, Alaska: A Large Magnitude, Slip-Partitioned Event. *Science*, 300(5622), 1113–1118.
- [92] Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. 241, 376–396.
- [93] Eshelby, J. D. (1969). The elastic field of a crack extending non-uniformly under general anti-plane loading. *Journal of the Mechanics and Physics of Solids*, 17(3), 177–199.
- [94] Fan, W. & Shearer, P. M. (2015). Detailed rupture imaging of the 25 April 2015 Nepal earthquake using teleseismic P waves. *Geophysical Research Letters*, 42(14), 2015GL064587.
- [95] Fan, W. & Shearer, P. M. (2016). Local near instantaneously dynamically triggered aftershocks of large earthquakes. *Science*, 353(6304), 1133–1136.
- [96] Frankel, A. (2009). A Constant Stress-Drop Model for Producing Broadband Synthetic Seismograms: Comparison with the Next Generation Attenuation Relations Constant Stress-Drop Model for Producing Broadband Synthetic Seismograms. *Bulletin of the Seismological Society of America*, 99(2A), 664–680.
- [97] Frankel, A. (2013). Rupture History of the 2011 M 9 Tohoku Japan Earthquake Determined from Strong-Motion and High-Rate GPS Recordings: Subevents Radiating Energy in Different Frequency Bands Rupture History of the 2011 M 9 Tohoku Earthquake from Strong-Motion and High-Rate GPS Recordings. *Bulletin of the Seismological Society of America*, 103(2B), 1290–1306.
- [98] Fujie, G., Kodaira, S., Yamashita, M., Sato, T., Takahashi, T., & Takahashi, N. (2013). Systematic changes in the incoming plate structure at the Kuril trench. *Geophysical Research Letters*, 40(1), 88–93.
- [99] Fukahata, Y., Yagi, Y., & Rivera, L. (2014). Theoretical relationship between back-projection imaging and classical linear inverse solutions. *Geophysical Journal International*, 196(1), 552–559.
- [100] Gabuchian, V., Rosakis, A. J., Bhat, H. S., Madariaga, R., & Kanamori, H. (2017). Experimental evidence that thrust earthquake ruptures might open faults. *Nature*, 545(7654), 336–339.

- [101] Gailler, A., Charvis, P., & Flueh, E. R. (2007). Segmentation of the Nazca and South American plates along the Ecuador subduction zone from wide angle seismic profiles. *Earth and Planetary Science Letters*, 260(3), 444–464.
- [102] Galetzka, J., Melgar, D., Genrich, J. F., Geng, J., Owen, S., Lindsey, E. O., Xu, X., Bock, Y., Avouac, J.-P., Adhikari, L. B., Upreti, B. N., Pratt-Sitaula, B., Bhattarai, T. N., Sitaula, B. P., Moore, A., Hudnut, K. W., Szeliga, W., Normandeau, J., Fend, M., Flouzat, M., Bollinger, L., Shrestha, P., Koirala, B., Gautam, U., Bhattarai, M., Gupta, R., Kandel, T., Timsina, C., Sapkota, S. N., Rajaura, S., & Maharjan, N. (2015). Slip pulse and resonance of the Kathmandu basin during the 2015 Gorkha earthquake, Nepal. *Science*, 349(6252), 1091–1095.
- [103] Galvez, P., Ampuero, J.-P., Dalguer, L. A., Somala, S. N., & Nissen-Meyer, T. (2014). Dynamic earthquake rupture modelled with an unstructured 3-D spectral element method applied to the 2011 M9 Tohoku earthquake. *Geophysical Journal International*, 198(2), 1222–1240.
- [104] Graindorge, D., Calahorrano, A., Charvis, P., Collot, J.-Y., & Bethoux, N. (2004). Deep structures of the Ecuador convergent margin and the Carnegie Ridge, possible consequence on great earthquakes recurrence interval. *Geophysical Research Letters*, 31(4).
- [105] Grandin, R., Vallée, M., Satriano, C., Lacassin, R., Klinger, Y., Simoes, M., & Bollinger, L. (2015). Rupture process of the Mw=7.9 2015 Gorkha earthquake (Nepal): Insights into Himalayan megathrust segmentation. *Geophysical Research Letters*, 42(20), 2015GL066044.
- [106] Graves, R. & Pitarka, A. (2016). Kinematic Ground-Motion Simulations on Rough Faults Including Effects of 3D Stochastic Velocity Perturbations Kinematic Ground-Motion Simulations on Rough Faults. *Bulletin of the Seismological Society of America*, 106(5), 2136–2153.
- [107] Guatteri, M. & Spudich, P. (2000). What Can Strong-Motion Data Tell Us about Slip-Weakening Fault-Friction Laws? *Bulletin of the Seismological Society of America*, 90(1), 98–116.
- [108] Guo, Y., Koketsu, K., & Miyake, H. (2016). Propagation mechanism of long-period ground motions for offshore earthquakes along the Nankai Trough: Effects of the accretionary wedge. *Bulletin of the Seismological Society of America*, 106(3), 1176–1197.
- [109] Hamling, I. J., Hreinsdóttir, S., Clark, K., Elliott, J., Liang, C., Fielding, E., Litchfield, N., Villamor, P., Wallace, L., Wright, T. J., et al. (2017). Complex multifault rupture during the 2016 Mw 7.8 Kaikōura earthquake, New Zealand. *Science*, 356(6334).
- [110] Hampel, A., Kukowski, N., Bialas, J., Huebscher, C., & Heinbockel, R. (2004). Ridge subduction at an erosive margin: The collision zone of the Nazca Ridge in southern Peru. *Journal of Geophysical Research: Solid Earth*, 109(B2).

- [111] Harris, R. A., Barall, M., Aagaard, B., Ma, S., Roten, D., Olsen, K., Duan, B., Liu, D., Luo, B., Bai, K., et al. (2018). A suite of exercises for verifying dynamic earthquake rupture codes. *Seismological Research Letters*, 89(3), 1146–1162.
- [112] Harris, R. A. & Day, S. M. (1997). Effects of a low-velocity zone on a dynamic rupture. *Bulletin of the Seismological Society of America*, 87(5), 1267–1280.
- [113] Haskell, N. A. (1964). Total energy and energy spectral density of elastic wave radiation from propagating faults. *Bulletin of the Seismological Society of America*, 54(6A), 1811–1841.
- [114] Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. *Science*, 362(6410), 58–61.
- [115] Hayes, G. P., Wald, D. J., & Johnson, R. L. (2012). Slab1.0: A three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research: Solid Earth*, 117(B1), B01302.
- [116] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).
- [117] Heidarzadeh, M., Murotani, S., Satake, K., Ishibe, T., & Gusman, A. R. (2015). Source model of the 16 September 2015 Illapel, Chile Mw 8.4 earthquake based on teleseismic and tsunami data. *Geophysical Research Letters*, (pp. 2015GL067297).
- [118] Hicks, S. P., Rietbrock, A., Ryder, I. M., Lee, C.-S., & Miller, M. (2014). Anatomy of a megathrust: The 2010 M8.8 Maule, Chile earthquake rupture zone imaged using seismic tomography. *Earth and Planetary Science Letters*, 405, 142–155.
- [119] Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- [120] Horning, G., Canales, J. P., Carbotte, S. M., Han, S., Carton, H., Nedimović, M. R., & Keken, P. E. v. (2016). A 2-D tomographic model of the Juan de Fuca plate from accretion at axial seamount to subduction at the Cascadia margin from an active source ocean bottom seismometer survey. *Journal of Geophysical Research: Solid Earth*, 121(8), 5859–5879.
- [121] Houston, H. (2001). Influence of depth, focal mechanism, and tectonic setting on the shape and duration of earthquake source time functions. *Journal of Geophysical Research: Solid Earth*, 106(B6), 11137–11150.
- [122] Houston, H., Benz Harley M., & Vidale John E. (1998). Time functions of deep earthquakes from broadband and short-period stacks. *Journal of Geophysical Research: Solid Earth*, 103(B12), 29895–29913.

- [123] Houston, H. & Kanamori, H. (1986). Source spectra of great earthquakes: Teleseismic constraints on rupture process and strong motion. *Bulletin of the Seismological Society of America*, 76(1), 19–42.
- [124] Huang, Y. (2018). Earthquake Rupture in Fault Zones With Along-Strike Material Heterogeneity. *Journal of Geophysical Research: Solid Earth*, 123(11), 9884–9898.
- [125] Huang, Y. & Ampuero, J.-P. (2011). Pulse-like ruptures induced by low-velocity fault zones. *Journal of Geophysical Research: Solid Earth*, 116(B12).
- [126] Huang, Y., Ampuero, J.-P., & Helmberger, D. V. (2014a). Earthquake ruptures modulated by waves in damaged fault zones. *Journal of Geophysical Research: Solid Earth*, 119(4), 3133–3154.
- [127] Huang, Y., Ampuero, J.-P., & Helmberger, D. V. (2016). The potential for supershear earthquakes in damaged fault zones—theory and observations. *Earth and Planetary Science Letters*, 433, 109–115.
- [128] Huang, Y., Ampuero, J.-P., & Kanamori, H. (2014b). Slip-Weakening Models of the 2011 Tohoku-Oki Earthquake and Constraints on Stress Drop and Fracture Energy. *Pure and Applied Geophysics*, 171(10), 2555–2568.
- [129] Huang, Y., Meng, L., & Ampuero, J.-P. (2012). A dynamic model of the frequency-dependent rupture process of the 2011 Tohoku-Oki earthquake. *Earth, Planets and Space*, 64(12), 1.
- [130] Hubbert, M. K. & Rubey, W. W. (1959). Role of fluid pressure in mechanics of overthrust faulting: I. mechanics of fluid-filled porous solids and its application to overthrust faulting. *GSA Bulletin*, 70(2), 115–166.
- [131] Husseini, M. I. (1977). Energy balance for motion along a fault. *Geophysical Journal International*, 49(3), 699–714.
- [132] Ida, Y. (1972). Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy. *Journal of Geophysical Research*, 77(20), 3796–3805.
- [133] Ide, S., Baltay, A., & Beroza, G. C. (2011). Shallow dynamic overshoot and energetic deep rupture in the 2011 mw 9.0 tohoku-oki earthquake. *Science*, 332(6036), 1426–1429.
- [134] Ide, S. & Beroza, G. C. (2001). Does apparent stress vary with earthquake size. *Geophys. Res. Lett*, 28(17), 3349–3352.
- [135] Ide, S. & Takeo, M. (1997). Determination of constitutive relations of fault slip based on seismic wave analysis. *Journal of Geophysical Research: Solid Earth*, 102(B12), 27379–27391.

- [136] Illien, L., Andermann, C., Sens-Schönfelder, C., Cook, K., Baidya, K., Adhikari, L., & Hovius, N. (2021). Subsurface moisture regulates himalayan groundwater storage and discharge. *AGU Advances*, 2(2), e2021AV000398.
- [137] Incorporated Research Institutions for Seismology Data Management Center (2011). Data Services Products: BackProjection P-wave back-projection rupture imaging.
- [138] Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2005). Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature*, 435(7044), 933.
- [139] Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2007). Teleseismic P wave imaging of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array. *Journal of Geophysical Research: Solid Earth*, 112(B11), B11307.
- [140] Ji, C., Wald, D. J., & Helmberger, D. V. (2002a). Source Description of the 1999 Hector Mine, California, Earthquake, Part I: Wavelet Domain Inversion Theory and Resolution Analysis. *Bulletin of the Seismological Society of America*, 92(4), 1192–1207.
- [141] Ji, C., Wald, D. J., & Helmberger, D. V. (2002b). Source Description of the 1999 Hector Mine, California, Earthquake, Part II: Complexity of Slip History. *Bulletin of the Seismological Society of America*, 92(4), 1208–1226.
- [142] Johnson, P. A., Rouet-Leduc, B., Pyrak-Nolte, L. J., Beroza, G. C., Marone, C. J., Hulbert, C., Howard, A., Singer, P., Gordeev, D., Karaflos, D., et al. (2021). Laboratory earthquake forecasting: A machine learning competition. *Proceedings of the National Academy of Sciences*, 118(5).
- [143] Kanamori, H. (1972). Mechanism of tsunami earthquakes. *Physics of the Earth and Planetary Interiors*, 6(5), 346–359.
- [144] Kanamori, H., Hong-Kie, T., Doug, D., Egill, H., & Heaton, T. (1992). Initial investigation of the Landers, California, Earthquake of 28 June 1992 using TERRAScope. *Geophysical Research Letters*, 19(22), 2267–2270.
- [145] Kanamori, H. & Kikuchi, M. (1993). The 1992 Nicaragua earthquake: a slow tsunami earthquake associated with subducted sediments. *Nature*, 361(6414), 714–716.
- [146] Kaneko, Y., Fukuyama, E., & Hamling, I. J. (2017). Slip-weakening distance and energy budget inferred from near-fault ground deformation during the 2016 Mw7.8 Kaikōura earthquake. *Geophysical Research Letters*, 44(10), 4765–4773.
- [147] Kaneko, Y. & Lapusta, N. (2010). Supershear transition due to a free surface in 3-D simulations of spontaneous dynamic rupture on vertical strike-slip faults. *Tectonophysics*, 493(3), 272–284.

- [148] Kaneko, Y. & Shearer, P. M. (2014). Seismic source spectra and estimated stress drop derived from cohesive-zone models of circular subshear rupture. *Geophysical Journal International*, 197(2), 1002–1015.
- [149] Keefer, D. K. (1984). Landslides caused by earthquakes. *Geological Society of America Bulletin*, 95(4), 406–421.
- [150] Kennett, B. L. N. & Engdahl, E. R. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2), 429–465.
- [151] Kikuchi, M. & Kanamori, H. (1982). Inversion of complex body waves. *Bulletin of the Seismological Society of America*, 72(2), 491–506.
- [152] Kikuchi, M. & Kanamori, H. (1986). Inversion of complex body waves-II. *Physics of the Earth and Planetary Interiors*, 43(3), 205–222.
- [153] Kikuchi, M. & Kanamori, H. (1991). Inversion of complex body waves—III. *Bulletin of the Seismological Society of America*, 81(6), 2335–2350.
- [154] Kingma, D. P. & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [155] Kiser, E. & Ishii, M. (2017). Back-Projection Imaging of Earthquakes. *Annual Review of Earth and Planetary Sciences*, 45(1), 271–299.
- [156] Kiser, E., Ishii, M., Langmuir, C. H., Shearer, P. M., & Hirose, H. (2011). Insights into the mechanism of intermediate-depth earthquakes from source properties as imaged by back projection of multiple seismic phases. *Journal of Geophysical Research: Solid Earth*, 116(B6), B06310.
- [157] Klingelhoefer, F., Berthet, T., Lallemand, S., Schnurle, P., Lee, C. S., Liu, C. S., McIntosh, K., & Theunissen, T. (2012). P-wave velocity structure of the southern Ryukyu margin east of Taiwan: Results from the ACTS wide-angle seismic experiment. *Tectonophysics*, 578, 50–62.
- [158] Klingelhoefer, F., Gutscher, M.-A., Ladage, S., Dessa, J.-X., Graindorge, D., Franke, D., André, C., Permana, H., Yudistira, T., & Chauhan, A. (2010). Limits of the seismogenic zone in the epicentral region of the 26 December 2004 great Sumatra-Andaman earthquake: Results from seismic refraction and wide-angle reflection surveys and thermal modeling. *Journal of Geophysical Research: Solid Earth*, 115(B1).
- [159] Klinger, Y. (2010). Relation between continental strike-slip earthquake segmentation and thickness of the crust. *Journal of Geophysical Research: Solid Earth*, 115(B7).

- [160] Klinger, Y., Xu, X., Tapponnier, P., Van der Woerd, J., Lasserre, C., & King, G. (2005). High-resolution satellite imagery mapping of the surface rupture and slip distribution of the m w 7.8, 14 november 2001 kokoxili earthquake, kunlun fault, northern tibet, china. *Bulletin of the Seismological Society of America*, 95(5), 1970–1987.
- [161] Ko, J. Y.-T. & Kuo, B.-Y. (2016). Low radiation efficiency of the intermediate-depth earthquakes in the Japan subduction zone. *Geophysical Research Letters*, 43(22), 2016GL070993.
- [162] Kodaira, S., Takahashi, N., Park, J.-O., Mochizuki, K., Shinohara, M., & Kimura, S. (2000). Western Nankai Trough seismogenic zone: Results from a wide-angle ocean bottom seismic survey. *Journal of Geophysical Research: Solid Earth*, 105(B3), 5887–5905.
- [163] Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, 90(1), 3–14.
- [164] Koper, K. D., Hutko, A. R., Lay, T., Ammon, C. J., & Kanamori, H. (2011). Frequency-dependent rupture process of the 2011 M w 9.0 Tohoku Earthquake: Comparison of short-period P wave backprojection images and broadband seismic rupture models. *Earth, planets and space*, 63(7), 599–602.
- [165] Kopp, H., Weinzierl, W., Becel, A., Charvis, P., Evain, M., Flueh, E. R., Gailler, A., Galve, A., Hirn, A., Kandilarov, A., Klaeschen, D., Laigle, M., Papenberg, C., Planert, L., & Roux, E. (2011). Deep structure of the central Lesser Antilles Island Arc: Relevance for the formation of continental crust. *Earth and Planetary Science Letters*, 304(1), 121–134.
- [166] Kostrov, B. V. (1964). Selfsimilar problems of propagation of shear cracks. *Journal of Applied Mathematics and Mechanics*, 28(5), 1077–1087.
- [167] Kozdon, J. E. & Dunham, E. M. (2013). Rupture to the Trench: Dynamic Rupture Simulations of the 11 March 2011 Tohoku Earthquake. *Bulletin of the Seismological Society of America*, 103(2B), 1275–1289.
- [168] Kozdon, J. E., Dunham, E. M., & Nordström, J. (2013). Simulation of Dynamic Earthquake Ruptures in Complex Geometries Using High-Order Finite Difference Methods. *Journal of Scientific Computing*, 55(1), 92–124.
- [169] Krabbenhöft, A., Bialas, J., Kopp, H., Kukowski, N., & Hübcher, C. (2004). Crustal structure of the Peruvian continental margin from wide-angle seismic studies. *Geophysical Journal International*, 159(2), 749–764.
- [170] Kuang, W., Yuan, C., & Zhang, J. (2021). Real-time determination of earthquake focal mechanism via deep learning. *Nature communications*, 12(1), 1–8.

- [171] Kurahashi, S. & Irikura, K. (2011). Source model for generating strong ground motions during the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space*, 63(7), 11.
- [172] Langston, C. A. (1978). Moments, corner frequencies, and the free surface. *Journal of Geophysical Research: Solid Earth*, 83(B7), 3422–3426.
- [173] Lay, T. (2015). The surge of great earthquakes from 2004 to 2014. *Earth and Planetary Science Letters*, 409, 133–146.
- [174] Lay, T., Ammon, C. J., Kanamori, H., Yamazaki, Y., Cheung, K. F., & Hutko, A. R. (2011). The 25 October 2010 Mentawai tsunami earthquake (Mw 7.8) and the tsunami hazard presented by shallow megathrust ruptures. *Geophysical Research Letters*, 38(6), L06302.
- [175] Lay, T., Kanamori, H., Ammon, C. J., Koper, K. D., Hutko, A. R., Ye, L., Yue, H., & Rushing, T. M. (2012). Depth-varying rupture properties of subduction zone megathrust faults. *Journal of Geophysical Research: Solid Earth*, 117(B4), B04311.
- [176] Lecocq, T., Hicks, S. P., Van Noten, K., Van Wijk, K., Koelemeijer, P., De Plaen, R. S., Massin, F., Hillers, G., Anthony, R. E., Apoloner, M.-T., et al. (2020). Global quieting of high-frequency seismic noise due to covid-19 pandemic lockdown measures. *Science*, 369(6509), 1338–1343.
- [177] Lee, S.-J., Yeh, T.-Y., Lin, T.-C., Lin, Y.-Y., Song, T.-R. A., & Huang, B.-S. (2016). Two-stage composite megathrust rupture of the 2015 Mw8.4 Illapel, Chile, earthquake identified by spectral-element inversion of teleseismic waves. *Geophysical Research Letters*, 43(10), 2016GL068843.
- [178] Li, L., Lay, T., Cheung, K. F., & Ye, L. (2016). Joint modeling of teleseismic and tsunami wave observations to constrain the 16 September 2015 Illapel, Chile, Mw 8.3 earthquake rupture process. *Geophysical Research Letters*, 43(9), 2016GL068674.
- [179] Li, Y.-G., Aki, K., Vidale, J. E., Lee, W. H. K., & Marone, C. J. (1994). Fine Structure of the Landers Fault Zone: Segmentation and the Rupture Process. *Science*, 265(5170), 367–370.
- [180] Li, Z., Meier, M.-A., Hauksson, E., Zhan, Z., & Andrews, J. (2018). Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophysical Research Letters*, 45(10), 4773–4779.
- [181] Lipovsky, B. P., Meyer, C. R., Zoet, L. K., McCarthy, C., Hansen, D. D., Rempel, A. W., & Gimbert, F. (2019). Glacier sliding, seismicity and sediment entrainment. *Annals of Glaciology*, 60(79), 182–192.
- [182] Liu, M., Zhang, M., Zhu, W., Ellsworth, W. L., & Li, H. (2020). Rapid characterization of the July 2019 ridgecrest, California, earthquake sequence from raw seismic data using machine-learning phase picker. *Geophysical Research Letters*, 47(4), e2019GL086189.

- [183] Liu, P., Archuleta, R. J., & Hartzell, S. H. (2006). Prediction of Broadband Ground-Motion Time Histories: Hybrid Low/High-Frequency Method with Correlated Random Source Parameters. *Bulletin of the Seismological Society of America*, 96(6), 2118–2130.
- [184] Longuet-Higgins, M. S. (1950). A theory of the origin of microseisms. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 243(857), 1–35.
- [185] Lotto, G. C., Dunham, E. M., Jeppson, T. N., & Tobin, H. J. (2017). The effect of compliant prisms on subduction zone earthquakes and tsunamis. *Earth and Planetary Science Letters*, 458, 213–222.
- [186] Lotto, G. C., Jeppson, T. N., & Dunham, E. M. (2018). Fully Coupled Simulations of Megathrust Earthquakes and Tsunamis in the Japan Trench, Nankai Trough, and Cascadia Subduction Zone. *Pure and Applied Geophysics*.
- [187] Luo, Y., Ampuero, J.-P., Miyakoshi, K., & Irikura, K. (2018). Surface Rupture Effects on Earthquake Moment-Area Scaling Relations. In L. A. Dalguer, Y. Fukushima, K. Irikura, & C. Wu (Eds.), *Best Practices in Physics-based Fault Rupture Models for Seismic Hazard Assessment of Nuclear Installations*, Pageoph Topical Volumes (pp. 7–18). Cham: Springer International Publishing.
- [188] Ma, S. (2012). A self-consistent mechanism for slow dynamic deformation and tsunami generation for earthquakes in the shallow subduction zone. *Geophysical Research Letters*, 39(11), L11310.
- [189] Ma, S. & Beroza, G. C. (2008). Rupture Dynamics on a Bimaterial Interface for Dipping Faults. *Bulletin of the Seismological Society of America*, 98(4), 1642–1658.
- [190] Ma, S. & Hirakawa, E. T. (2013). Dynamic wedge failure reveals anomalous energy radiation of shallow subduction earthquakes. *Earth and Planetary Science Letters*, 375, 113–122.
- [191] Ma, S. & Nie, S. (2019). Dynamic wedge failure and along-arc variations of tsunamigenesis in the Japan trench margin. *Geophysical Research Letters*, 46(15), 8782–8790.
- [192] Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, 66(3), 639–666.
- [193] Madariaga, R. (1977). High-frequency radiation from crack (stress drop) models of earthquake faulting. *Geophysical Journal International*, 51(3), 625–651.
- [194] Madariaga, R. (1983). High frequency radiation from dynamic earthquake fault models. *Annales de Geophysique*, 1(1), 17–23.

- [195] Madariaga, R., Ampuero, J. P., & Adda-Bedia, M. (2006). Seismic Radiation from Simple Models of Earthquakes. In R. Abercrombie, A. McGarr, G. D. Toro, & H. Kanamori (Eds.), *Earthquakes: Radiated Energy and the Physics of Faulting* (pp. 223–236). American Geophysical Union.
- [196] Mai, P. M. & Beroza, G. C. (2002). A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research: Solid Earth*, 107(B11), ESE 10–1.
- [197] Mai, P. M., Galis, M., Thingbaijam, K. K. S., Vyas, J. C., & Dunham, E. M. (2017). Accounting for Fault Roughness in Pseudo-Dynamic Ground-Motion Simulations. *Pure and Applied Geophysics*, 174(9), 3419–3450.
- [198] Mai, P. M., Schorlemmer, D., Page, M., Ampuero, J.-P., Asano, K., Causse, M., Custodio, S., Fan, W., Festa, G., Galis, M., Galovic, F., Imperatori, W., Käser, M., Malytsky, D., Okuwaki, R., Pollitz, F., Passone, L., Razafindrakoto, H. N. T., Sekiguchi, H., Song, S. G., Somala, S. N., Thingbaijam, K. K. S., Twardzik, C., Driel, M. v., Vyas, J. C., Wang, R., Yagi, Y., & Zielke, O. (2016). The Earthquake-Source Inversion Validation (SIV) Project. *Seismological Research Letters*.
- [199] Marone, C. & Saffer, D. M. (2007). Fault friction and the upper transition from seismic to aseismic faulting. *The Seismogenic Zone of Subduction Thrust Faults*, (pp. 346–369).
- [200] Martínez-Loriente, S., Sallarès, V., Ranero, C. R., Ruh, J. B., Barckhausen, U., Grevemeyer, I., & Bangs, N. (2019). Influence of Incoming Plate Relief on Overriding Plate Deformation and Earthquake Nucleation: Cocos Ridge Subduction (Costa Rica). *Tectonics*, 38(12), 4360–4377.
- [201] Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K., & Rabaute, T. (1993). The displacement field of the Landers earthquake mapped by radar interferometry. *Nature*, 364(6433), 138–142.
- [202] McNamara, D. E., Boaz, R. I., Nakata, N., Gualtieri, L., & Fichtner, A. (2019). Visualization of the seismic ambient noise spectrum. *Seismic Ambient Noise*, (pp. 1–29).
- [203] Meier, M.-A., Ampuero, J. P., & Heaton, T. H. (2017). The hidden simplicity of subduction megathrust earthquakes. *Science*, 357(6357), 1277–1281.
- [204] Meier, M.-A., Ross, Z. E., Ramachandran, A., Balakrishna, A., Nair, S., Kundzicz, P., Li, Z., Andrews, J., Hauksson, E., & Yue, Y. (2019). Reliable real-time seismic signal/noise discrimination with machine learning. *Journal of Geophysical Research: Solid Earth*, 124(1), 788–800.
- [205] Melgar, D., Fan, W., Riquelme, S., Geng, J., Liang, C., Fuentes, M., Vargas, G., Allen, R. M., Shearer, P. M., & Fielding, E. J. (2016). Slip segmentation and slow rupture to the trench

- during the 2015, Mw8.3 Illapel, Chile earthquake. *Geophysical Research Letters*, 43(3), 2015GL067369.
- [206] Meng, L., Ampuero, J.-P., Luo, Y., Wu, W., & Ni, S. (2012a). Mitigating artifacts in back-projection source imaging with implications for frequency-dependent properties of the Tohoku-Oki earthquake. *Earth, Planets and Space*, 64(12), 5.
- [207] Meng, L., Ampuero, J.-P., Stock, J., Duputel, Z., Luo, Y., & Tsai, V. C. (2012b). Earthquake in a maze: Compressional rupture branching during the 2012 Mw 8.6 Sumatra earthquake. *Science*, 337(6095), 724–726.
- [208] Meng, L., Bao, H., Huang, H., Zhang, A., Bloore, A., & Liu, Z. (2018). Double pincer movement: Encircling rupture splitting during the 2015 Mw 8.3 Illapel earthquake. *Earth and Planetary Science Letters*, 495, 164–173.
- [209] Meng, L., Inbal, A., & Ampuero, J.-P. (2011). A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake. *Geophysical Research Letters*, 38(7), L00G07.
- [210] Meng, L., Zhang, A., & Yagi, Y. (2016). Improving Back-Projection Imaging with a Novel Physics-Based Aftershock Calibration Approach: A Case Study of the 2015 Gorkha Earthquake. *Geophysical Research Letters*.
- [211] Michel, S., Avouac, J.-P., Lapusta, N., & Jiang, J. (2017). Pulse-like partial ruptures and high-frequency radiation at creeping-locked transition during megathrust earthquakes. *Geophysical Research Letters*, 44(16), 8345–8351.
- [212] Miura, S., Suyehiro, K., Shinohara, M., Takahashi, N., Araki, E., & Taira, A. (2004). Seismological structure and implications of collision between the Ontong Java Plateau and Solomon Island Arc from ocean bottom seismometer–airgun data. *Tectonophysics*, 389(3), 191–220.
- [213] Miura, S., Takahashi, N., Nakanishi, A., Tsuru, T., Kodaira, S., & Kaneda, Y. (2005). Structural characteristics off Miyagi forearc region, the Japan Trench seismogenic zone, deduced from a wide-angle reflection and refraction study. *Tectonophysics*, 407(3), 165–188.
- [214] Miyake, H., Iwata, T., & Irikura, K. (2003). Source Characterization for Broadband Ground-Motion Simulation: Kinematic Heterogeneous Source Model and Strong Motion Generation Area. *Bulletin of the Seismological Society of America*, 93(6), 2531–2545.
- [215] Moscoso, E., Grevemeyer, I., Contreras-Reyes, E., Flueh, E. R., Dzierma, Y., Rabbal, W., & Thorwart, M. (2011). Revealing the deep structure and rupture plane of the 2010 Maule, Chile earthquake (Mw=8.8) using wide angle seismic data. *Earth and Planetary Science Letters*, 307(1), 147–155.

- [216] Mousavi, S. M. & Beroza, G. C. (2019). Bayesian-deep-learning estimation of earthquake location from single-station observations. *arXiv preprint arXiv:1912.01144*.
- [217] Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, 11(1), 1–12.
- [218] Mousavi, S. M. & Langston, C. A. (2017). Automatic noise-removal/signal-removal based on general cross-validation thresholding in synchrosqueezed domain and its application on earthquake data. *Geophysics*, 82(4), V211–V227.
- [219] Mousavi, S. M., Langston, C. A., & Horton, S. P. (2016). Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform. *Geophysics*, 81(4), V341–V355.
- [220] Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019a). Stanford earthquake dataset (stead): A global data set of seismic signals for ai. *IEEE Access*, 7, 179464–179476.
- [221] Mousavi, S. M., Zhu, W., Sheng, Y., & Beroza, G. C. (2019b). Cred: A deep residual network of convolutional and recurrent units for earthquake signal detection. *Scientific reports*, 9(1), 1–14.
- [222] Mueller, C. S. & Cranswick, E. (1985). Source parameters from locally recorded aftershocks of the 9 january 1982 miramichi, new brunswick, earthquake. *Bulletin of the Seismological Society of America*, 75(2), 337–360.
- [223] Münchmeyer, J., Bindi, D., Leser, U., & Tilmann, F. (2021a). Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network. *Geophysical Journal International*, 226(2), 1086–1104.
- [224] Münchmeyer, J., Bindi, D., Leser, U., & Tilmann, F. (2021b). The transformer earthquake alerting model: a new versatile approach to earthquake early warning. *Geophysical Journal International*, 225(1), 646–656.
- [225] Murphy, S., Di Toro, G., Romano, F., Scala, A., Lorito, S., Spagnuolo, E., Aretusini, S., Festa, G., Piatanesi, A., & Nielsen, S. (2018). Tsunamigenic earthquake simulations using experimentally derived friction laws. *Earth and Planetary Science Letters*, 486, 155–165.
- [226] Naif, S., Key, K., Constable, S., & Evans, R. L. (2015). Water-rich bending faults at the Middle America Trench. *Geochemistry, Geophysics, Geosystems*, 16(8), 2582–2597.
- [227] Nakajima, H., Takahashi, Y., Kondo, K., & Hisaminato, Y. (2018). Monaural source enhancement maximizing source-to-distortion ratio via automatic differentiation. *arXiv preprint arXiv:1806.05791*.

- [228] Nakamura, Y., Kodaira, S., Cook, B. J., Jeppson, T., Kasaya, T., Yamamoto, Y., Hashimoto, Y., Yamaguchi, M., Obana, K., & Fujie, G. (2014). Seismic imaging and velocity structure around the JFAST drill site in the Japan Trench: low  $V_p$ , high  $V_p/V_s$  in the transparent frontal prism. *Earth, Planets and Space*, 66(1), 121.
- [229] Nakanishi, A., Kurashimo, E., Tatsumi, Y., Yamaguchi, H., Miura, S., Kodaira, S., Obana, K., Takahashi, N., Tsuru, T., Kaneda, Y., Iwasaki, T., & Hirata, N. (2009). Crustal evolution of the southwestern Kuril Arc, Hokkaido Japan, deduced from seismic velocity and geochemical structure. *Tectonophysics*, 472(1), 105–123.
- [230] Nakanishi, A., Takahashi, N., Park, J.-O., Miura, S., Kodaira, S., Kaneda, Y., Hirata, N., Iwasaki, T., & Nakamura, M. (2002). Crustal structure across the coseismic rupture zone of the 1944 Tonankai earthquake, the central Nankai Trough seismogenic zone. *Journal of Geophysical Research: Solid Earth*, 107(B1), EPM 2–1–EPM 2–21.
- [231] Nielsen, S. B. (1998). Free surface effects on the propagation of dynamic rupture. *Geophysical Research Letters*, 25(1), 125–128.
- [232] Nishimura, K., Uehara, S., & Mizoguchi, K. (2019). An alternative origin of high  $v_p/v_s$  anomalies in slow slip regions: Experimental constraints from the elastic wave velocity evolution of highly fractured rock. *Journal of Geophysical Research: Solid Earth*, 124(5), 5045–5059.
- [233] Nishizawa, A., Kaneda, K., Oikawa, M., Horiuchi, D., Fujioka, Y., & Okada, C. (2017). Variations in seismic velocity distribution along the Ryukyu (Nansei-Shoto) Trench subduction zone at the northwestern end of the Philippine Sea plate. *Earth, Planets and Space*, 69(1), 86.
- [234] Noda, H. & Lapusta, N. (2013). Stable creeping fault segments can become destructive as a result of dynamic weakening. *Nature*, 493(7433), 518–521.
- [235] Noda, H., Lapusta, N., & Kanamori, H. (2013). Comparison of average stress drop measures for ruptures with heterogeneous stress change and implications for earthquake physics. *Geophysical Journal International*, 193(3), 1691–1712.
- [236] Obara, K., Kasahara, K., Hori, S., & Okada, Y. (2005). A densely distributed high-sensitivity seismograph network in Japan: Hi-net by National Research Institute for Earth Science and Disaster Prevention. *Review of Scientific Instruments*, 76(2), 021301.
- [237] Oglesby, D. D., Archuleta, R. J., & Nielsen, S. B. (1998). Earthquakes on Dipping Faults: The Effects of Broken Symmetry. *Science*, 280(5366), 1055–1059.
- [238] Oglesby, D. D., Archuleta, R. J., & Nielsen, S. B. (2000). Dynamics of dip-slip faulting: Explorations in two dimensions. *Journal of Geophysical Research: Solid Earth*, 105(B6), 13643–13653.

- [239] Okada, Y., Kasahara, K., Hori, S., Obara, K., Sekiguchi, S., Fujiwara, H., & Yamamoto, A. (2004). Recent progress of seismic observation networks in Japan —Hi-net, F-net, K-NET and KiK-net—. *Earth, Planets and Space*, 56(8), xv–xxviii.
- [240] Olsen, K. B. & Mayhew, J. E. (2010). Goodness-of-fit Criteria for Broadband Synthetic Seismograms, with Application to the 2008 Mw 5.4 Chino Hills, California, Earthquake. *Seismological Research Letters*, 81(5), 715–723.
- [241] Olsen-Kettle, L. M., Weatherley, D., Saez, E., Gross, L., Mühlhaus, H.-B., & Xing, H. L. (2008). Analysis of slip-weakening frictional laws with static restrengthening and their implications on the scaling, asymmetry, and mode of dynamic rupture on homogeneous and bimaterial interfaces. *Journal of Geophysical Research: Solid Earth*, 113(B8).
- [242] Palmer, A. C. & Rice, J. R. (1973). The growth of slip surfaces in the progressive failure of over-consolidated clay. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 332(1591), 527–548.
- [243] Park, S. H., Kim, B., Kang, C. M., Chung, C. C., & Choi, J. W. (2018). Sequence-to-sequence prediction of vehicle trajectory via lstm encoder-decoder architecture. In 2018 *IEEE Intelligent Vehicles Symposium (IV)* (pp. 1672–1678).: IEEE.
- [244] Peacock, S., Westbrook, G. K., & Bais, G. (2010). S-wave velocities and anisotropy in sediments entering the Nankai subduction zone, offshore Japan. *Geophysical Journal International*, 180(2), 743–758.
- [245] Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2), e1700578.
- [246] Persh, S. E. & Houston, H. (2004). Strongly depth-dependent aftershock production in deep earthquakes. *Bulletin of the Seismological Society of America*, 94(5), 1808–1816.
- [247] Peterson, J. et al. (1993). Observations and modeling of seismic background noise.
- [248] Peyrat, S., Olsen, K., & Madariaga, R. (2001). Dynamic modeling of the 1992 Landers earthquake. *Journal of Geophysical Research: Solid Earth*, 106(B11), 26467–26482.
- [249] Pimienta, L., Schubnel, A., Violay, M., Fortin, J., Guéguen, Y., & Lyon-Caen, H. (2018). Anomalous vp/vs ratios at seismic frequencies might evidence highly damaged rocks in subduction zones. *Geophysical Research Letters*, 45(22), 12,210–12,217.
- [250] Planert, L., Kopp, H., Lueschen, E., Mueller, C., Flueh, E. R., Shulgin, A., Djajadihardja, Y., & Krabbenhoft, A. (2010). Lower plate structure and upper plate deformational segmentation at the Sunda-Banda arc transition, Indonesia. *Journal of Geophysical Research: Solid Earth*, 115(B8).

- [251] Poli, P. & Prieto, G. A. (2016). Global rupture parameters for deep and intermediate-depth earthquakes. *Journal of Geophysical Research: Solid Earth*, 121(12), 2016JB013521.
- [252] Qin, W. & Yao, H. (2017). Characteristics of subevents and three-stage rupture processes of the 2015 Mw 7.8 Gorkha Nepal earthquake from multiple-array back projection. *Journal of Asian Earth Sciences*, 133, 72–79.
- [253] Ramos, M. D. & Huang, Y. (2019). How the Transition Region Along the Cascadia Megathrust Influences Coseismic Behavior: Insights From 2-D Dynamic Rupture Simulations. *Geophysical Research Letters*, 46(4), 1973–1983.
- [254] Ren, C. X., Hulbert, C., Johnson, P. A., & Rouet-Leduc, B. (2020). Machine learning and fault rupture: a review. *Advances in Geophysics*, 61, 57–107.
- [255] Rhie, J. & Romanowicz, B. (2004). Excitation of earth's continuous free oscillations by atmosphere–ocean–seafloor coupling. *Nature*, 431(7008), 552–556.
- [256] Rice, J. R. (1992). Chapter 20 Fault Stress States, Pore Pressure Distributions, and the Weakness of the San Andreas Fault. In B. Evans & T.-f. Wong (Eds.), *International Geophysics*, volume 51 of *Fault Mechanics and Transport Properties of Rocks* (pp. 475–503). Academic Press.
- [257] Rice, J. R. (1993). Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research: Solid Earth*, 98(B6), 9885–9907.
- [258] Rice, J. R. (2006). Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research: Solid Earth*, 111(B5), B05311.
- [259] Rice, J. R., Lapusta, N., & Ranjith, K. (2001). Rate and state dependent friction and the stability of sliding between elastically deformable solids. *Journal of the Mechanics and Physics of Solids*, 49(9), 1865–1898.
- [260] Ripperger, J., Ampuero, J.-P., Mai, P. M., & Giardini, D. (2007). Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress. *Journal of Geophysical Research: Solid Earth*, 112(B4).
- [261] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234–241): Springer.
- [262] Rost, S. & Thomas, C. (2002). Array Seismology: Methods and Applications. *Reviews of Geophysics*, 40(3), 1008.
- [263] Rubin, A. M. & Ampuero, J.-P. (2007). Aftershock asymmetry on a bimaterial interface. *Journal of Geophysical Research: Solid Earth*, 112(B5).

- [264] Ruina, A. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research: Solid Earth*, 88(B12), 10359–10370.
- [265] Ruiz, S., Madariaga, R., Astroza, M., Saragoni, G. R., Lancieri, M., Vigny, C., & Campos, J. (2012). Short-Period Rupture Process of the 2010 Mw 8.8 Maule Earthquake in Chile. *Earthquake Spectra*, 28(S1), S1–S18.
- [266] Saad, O. M., Inoue, K., Shalaby, A., Samy, L., & Sayed, M. S. (2018). Automatic arrival time detection for earthquakes based on stacked denoising autoencoder. *IEEE Geoscience and Remote Sensing Letters*, 15(11), 1687–1691.
- [267] Saffer, D. M. & Marone, C. (2003). Comparison of smectite- and illite-rich gouge frictional properties: application to the updip limit of the seismogenic zone along subduction megathrusts. *Earth and Planetary Science Letters*, 215(1–2), 219–235.
- [268] Saffer, D. M. & Tobin, H. J. (2011). Hydrogeology and mechanics of subduction zone forearcs: Fluid flow and pore pressure. *Annual Review of Earth and Planetary Sciences*, 39, 157–186.
- [269] Sallarès, V., Dañobeitia, J. J., & Flueh, E. R. (2001). Lithospheric structure of the Costa Rican Isthmus: Effects of subduction zone magmatism on an oceanic plateau. *Journal of Geophysical Research: Solid Earth*, 106(B1), 621–643.
- [270] Sallarès, V. & Ranero, C. R. (2019). Upper-plate rigidity determines depth-varying rupture behaviour of megathrust earthquakes. *Nature*, 576(7785), 96–101.
- [271] Sato, K. & Mori, J. (2006). Relationship between rupture process complexity and earthquake size. *Journal of Geophysical Research: Solid Earth*, 111(B5).
- [272] Saygin, E., Cummins, P. R., & Lumley, D. (2017). Retrieval of the P wave reflectivity response from autocorrelation of seismic noise: Jakarta Basin, Indonesia. *Geophysical Research Letters*, 44(2), 792–799. [\\_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2016GL071363](https://onlinelibrary.wiley.com/doi/pdf/10.1002/2016GL071363).
- [273] Scherwath, M., Contreras-Reyes, E., Flueh, E. R., Grevemeyer, I., Krabbenhoft, A., Papenberg, C., Petersen, C. J., & Weinrebe, R. W. (2009). Deep lithospheric structures along the southern central Chile margin from wide-angle P-wave modelling. *Geophysical Journal International*, 179(1), 579–600.
- [274] Schippkus, S., Garden, M., & Bokelmann, G. (2020). Characteristics of the ambient seismic field on a large-n seismic array in the vienna basin. *Seismological Society of America*, 91(5), 2803–2816.
- [275] Schmedes, J., Archuleta, R. J., & Lavallée, D. (2010). Correlation of earthquake source parameters inferred from dynamic rupture simulations. *Journal of Geophysical Research: Solid Earth*, 115(B3), B03304.

- [276] Schmedes, J., Archuleta, R. J., & Lavallée, D. (2013). A kinematic rupture model generator incorporating spatial interdependency of earthquake source parameters. *Geophysical Journal International*, 192(3), 1116–1131.
- [277] Scholz, C. H. (1998). Earthquakes and friction laws. *Nature*, 391(6662), 37–42.
- [278] Sens-Schönfelder, C. & Wegler, U. (2006). Passive image interferometry and seasonal variations of seismic velocities at merapi volcano, indonesia. *Geophysical research letters*, 33(21).
- [279] Serban, I., Sordoni, A., Lowe, R., Charlin, L., Pineau, J., Courville, A., & Bengio, Y. (2017). A hierarchical latent variable encoder-decoder model for generating dialogues. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.
- [280] Shao, G., Ji, C., & Zhao, D. (2011). Rupture process of the 9 march, 2011 mw 7.4 sanriku-oki, japan earthquake constrained by jointly inverting teleseismic waveforms, strong motion data and gps observations. *Geophysical Research Letters*, 38(7).
- [281] Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307(5715), 1615–1618.
- [282] Shearer, P. M., Abercrombie, R. E., Trugman, D. T., & Wang, W. (2019). Comparing egf methods for estimating corner frequency and stress drop from p wave spectra. *Journal of Geophysical Research: Solid Earth*, 124(4), 3966–3986.
- [283] Shearer, P. M., Prieto, G. A., & Hauksson, E. (2006). Comprehensive analysis of earthquake source spectra in southern California. *Journal of Geophysical Research: Solid Earth*, 111(B6).
- [284] Shelly, D. R., Beroza, G. C., Zhang, H., Thurber, C. H., & Ide, S. (2006). High-resolution subduction zone seismicity and velocity structure beneath Ibaraki Prefecture, Japan. *Journal of Geophysical Research: Solid Earth*, 111(B6).
- [285] Shi, Z. & Ben-Zion, Y. (2006). Dynamic rupture on a bimaterial interface governed by slip-weakening friction. *Geophysical Journal International*, 165(2), 469–484.
- [286] Shlomag, H. & Fineberg, J. (2016). The structure of slip-pulses and supershear ruptures driving slip in bimaterial friction. *Nature Communications*, 7(1), 11787.
- [287] Shulgin, A., Kopp, H., Mueller, C., Planert, L., Lueschen, E., Flueh, E. R., & Djajadihardja, Y. (2011). Structural architecture of oceanic plateau subduction offshore Eastern Java and the potential implications for geohazards. *Geophysical Journal International*, 184(1), 12–28.
- [288] Siahshar, M. A. N., Gholtashi, S., Abolghasemi, V., & Chen, Y. (2017). Simultaneous denoising and interpolation of 2d seismic data using data-driven non-negative dictionary learning. *Signal Processing*, 141, 309–321.

- [289] Simons, M., Minson, S. E., Sladen, A., Ortega, F., Jiang, J., Owen, S. E., Meng, L., Ampuero, J.-P., Wei, S., Chu, R., Helmberger, D. V., Kanamori, H., Hetland, E., Moore, A. W., & Webb, F. H. (2011). The 2011 Magnitude 9.0 Tohoku-Oki Earthquake: Mosaicking the Megathrust from Seconds to Centuries. *Science*, 332(6036), 1421–1425.
- [290] Somerville, P. G., Smith, N. F., Graves, R. W., & Abrahamson, N. A. (1997). Modification of Empirical Strong Ground Motion Attenuation Relations to Include the Amplitude and Duration Effects of Rupture Directivity. *Seismological Research Letters*, 68(1), 199–222.
- [291] Spudich, P. & Frazer, L. N. (1984). Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop. *Bulletin of the Seismological Society of America*, 74(6), 2061–2082.
- [292] Stockwell, R. G., Mansinha, L., & Lowe, R. (1996). Localization of the complex spectrum: the s transform. *IEEE transactions on signal processing*, 44(4), 998–1001.
- [293] Sufri, O., Koper, K. D., & Lay, T. (2012). Along-dip seismic radiation segmentation during the 2007 Mw 8.0 Pisco, Peru earthquake. *Geophysical Research Letters*, 39(8), L08311.
- [294] Takahashi, N., Suyehiro, K., & Shinohara, M. (1998). Implications from the seismic crustal structure of the northern Izu–Bonin arc. *Island Arc*, 7(3), 383–394.
- [295] Tal, Y., Rubino, V., Rosakis, A. J., & Lapusta, N. (2020). Illuminating the physics of dynamic friction through laboratory earthquakes on thrust faults. *Proceedings of the National Academy of Sciences*, 117(35), 21095–21100.
- [296] Tan, C. W., Bergmeir, C., Petitjean, F., & Webb, G. I. (2021). Time series extrinsic regression. *Data Mining and Knowledge Discovery*, 35(3), 1032–1060.
- [297] Tanioka, Y. & Ruff, L. J. (1997). Source Time Functions. *Seismological Research Letters*, 68(3), 386–400.
- [298] Tanioka, Y. & Satake, K. (1996). Tsunami generation by horizontal displacement of ocean bottom. *Geophysical Research Letters*, 23(8), 861–864.
- [299] Tibi, R., Hammond, P., Brogan, R., Young, C. J., & Koper, K. (2021). Deep learning denoising applied to regional distance seismic data in Utah. *Bulletin of the Seismological Society of America*, 111(2), 775–790.
- [300] Tinti, E., Fukuyama, E., Piatanesi, A., & Cocco, M. (2005a). A Kinematic Source-Time Function Compatible with Earthquake Dynamics. *Bulletin of the Seismological Society of America*, 95(4), 1211–1223.
- [301] Tinti, E., Spudich, P., & Cocco, M. (2005b). Earthquake fracture energy inferred from kinematic rupture models on extended faults. *Journal of Geophysical Research: Solid Earth*, 110(B12).

- [302] Tocheport, A., Rivera, L., & Van der Woerd, J. (2006). A study of the 14 November 2001 Kokoxili earthquake: History and geometry of the rupture from teleseismic data and field observations. *Bulletin of the Seismological Society of America*, 96(5), 1729–1741.
- [303] Trnkoczy, A. (2009). Understanding and parameter setting of sta/lta trigger algorithm. In *New Manual of Seismological Observatory Practice (NMSOP)* (pp. 1–20). Deutsches Geo-ForschungsZentrum GFZ.
- [304] Tromp, J., Komatitsch, D., & Liu, Q. (2008). Spectral-element and adjoint methods in seismology. *Communications in Computational Physics*, 3(1), 1–32.
- [305] Trugman, D. T. & Dunham, E. M. (2014). A 2D Pseudodynamic Rupture Model Generator for Earthquakes on Geometrically Complex Faults A 2D Pseudodynamic Rupture Model Generator for Earthquakes on Geometrically Complex Faults. *Bulletin of the Seismological Society of America*, 104(1), 95–112.
- [306] Trugman, D. T. & Shearer, P. M. (2017). Application of an improved spectral decomposition method to examine earthquake source scaling in Southern California. *Journal of Geophysical Research: Solid Earth*, 122(4), 2890–2910.
- [307] Tsuji, T., Dvorkin, J., Mavko, G., Nakata, N., Matsuoka, T., Nakanishi, A., Kodaira, S., & Nishizawa, O. (2011). VP/VS ratio and shear-wave splitting in the Nankai Trough seismogenic zone: Insights into effective stress, pore pressure, and sediment consolidation Tsuji et al. VP/VS ratio & S-wave splitting from OBS data. *Geophysics*, 76(3), WA71–WA82.
- [308] Uchide, T., Yao, H., & Shearer, P. M. (2013). Spatio-temporal distribution of fault slip and high-frequency radiation of the 2010 El Mayor-Cucapah, Mexico earthquake. *Journal of Geophysical Research: Solid Earth*, 118(4), 1546–1555.
- [309] Vallée, M. (2013). Source time function properties indicate a strain drop independent of earthquake depth and magnitude. *Nature Communications*, 4, 2606.
- [310] Vallée, M., Charléty, J., Ferreira, A. M. G., Delouis, B., & Vergoz, J. (2011). SCARDEC: a new technique for the rapid determination of seismic moment magnitude, focal mechanism and source time functions for large earthquakes using body-wave deconvolution. *Geophysical Journal International*, 184(1), 338–358.
- [311] Vallée, M. & Douet, V. (2016). A new database of source time functions (STFs) extracted from the SCARDEC method. *Physics of the Earth and Planetary Interiors*, 257, 149–157.
- [312] van den Ende, M. P. & Ampuero, J.-P. (2020). Automated seismic source characterization using deep graph neural networks. *Geophysical Research Letters*, 47(17), e2020GL088690.
- [313] Vanderkulk, W., Rosen, F., & Lorenz, S. (1965). Large aperture seismic array signal processing study. *IBM Final Report, ARPA Contract Number SD-296*.

- [314] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998–6008).
- [315] Venkataraman, A. & Kanamori, H. (2004). Observational constraints on the fracture energy of subduction zone earthquakes. *Journal of Geophysical Research: Solid Earth*, 109(B5), B05302.
- [316] Viens, L., Denolle, M. A., Hirata, N., & Nakagawa, S. (2018). Complex near-surface rheology inferred from the response of greater tokyo to strong ground motions. *Journal of Geophysical Research: Solid Earth*, 123(7), 5710–5729.
- [317] Viens, L., Jiang, C., & Denolle, M. (2020). Imaging the kanto basin bedrock with noise and earthquake autocorrelations.
- [318] Viesca, R. C. & Garagash, D. I. (2015). Ubiquitous weakening of faults due to thermal pressurization. *Nature Geoscience*, 8(11), 875–879.
- [319] Von Huene, R., Ranero, C. R., & Scholl, D. W. (2009). Convergent margin structure in high-quality geophysical images and current kinematic and dynamic models. In *Subduction zone geodynamics* (pp. 137–157). Springer.
- [320] Walker, K. T., Ishii, M., & Shearer, P. M. (2005). Rupture details of the 28 March 2005 Sumatra Mw 8.6 earthquake imaged with teleseismic P waves. *Geophysical Research Letters*, 32(24), L24303.
- [321] Walker, K. T. & Shearer, P. M. (2009). Illuminating the near-sonic rupture velocities of the intracontinental Kokoxili Mw 7.8 and Denali fault Mw 7.9 strike-slip earthquakes with global P wave back projection imaging. *Journal of Geophysical Research: Solid Earth*, 114(B2).
- [322] Walter, J. I., Ogwari, P., Thiel, A., Ferrer, F., & Woelfel, I. (2021). Easyquake: Putting machine learning to work for your regional seismic network or local earthquake study. *Seismological Society of America*, 92(1), 555–563.
- [323] Walther, C. H. E., Flueh, E. R., Ranero, C. R., Von Huene, R., & Strauch, W. (2000). Crustal structure across the Pacific margin of Nicaragua: evidence for ophiolitic basement and a shallow mantle sliver. *Geophysical Journal International*, 141(3), 759–777.
- [324] Wang, D. & Mori, J. (2011). Frequency-dependent energy radiation and fault coupling for the 2010 Mw8.8 Maule, Chile, and 2011 Mw9.0 Tohoku, Japan, earthquakes. *Geophysical Research Letters*, 38(22).
- [325] Wang, D., Mori, J., & Uchide, T. (2012). Supershear rupture on multiple faults for the Mw 8.6 Off Northern Sumatra, Indonesia earthquake of April 11, 2012. *Geophysical Research Letters*, 39(21).

- [326] Wang, D., Takeuchi, N., Kawakatsu, H., & Mori, J. (2016). Estimating high frequency energy radiation of large earthquakes by image deconvolution back-projection. *Earth and Planetary Science Letters*, 449, 155–163.
- [327] Wang, K. & Bilek, S. L. (2011). Do subducting seamounts generate or stop large earthquakes? *Geology*, 39(9), 819–822.
- [328] Wang, K., Sun, T., Brown, L., Hino, R., Tomita, F., Kido, M., Inuma, T., Kodaira, S., & Fujiwara, T. (2018). Learning from crustal deformation associated with the M9 2011 Tohoku-oki earthquake. *Geosphere*, 14(2), 552–571.
- [329] Wang, W. & Shearer, P. M. (2019). An improved method to determine coda-q, earthquake magnitude, and site amplification: Theory and application to southern california. *Journal of Geophysical Research: Solid Earth*, 124(1), 578–598.
- [330] Wang, Y., Day, S. M., & Denolle, M. A. (2019). Geometric Controls on Pulse-Like Rupture in a Dynamic Model of the 2015 Gorkha Earthquake. *Journal of Geophysical Research: Solid Earth*, 124(2), 1544–1568.
- [331] Warren, L. M. & Shearer, P. M. (2000). Investigating the frequency dependence of mantle Q by stacking P and PP spectra. *Journal of Geophysical Research: Solid Earth*, 105(B11), 25391–25402.
- [332] Warren, L. M. & Shearer, P. M. (2002). Mapping lateral variations in upper mantle attenuation by stacking P and PP spectra. *Journal of Geophysical Research: Solid Earth*, 107(B12), 2342.
- [333] Warren, L. M. & Shearer, P. M. (2005). Using the Effects of Depth Phases on P-wave Spectra to Determine Earthquake Depths. *Bulletin of the Seismological Society of America*, 95(1), 173–184.
- [334] Weaver, C. S. & Malone, S. D. (1979). Seismic evidence for discrete glacier motion at the rock–ice interface. *Journal of Glaciology*, 23(89), 171–184.
- [335] Weertman, J. (1980). Unstable slippage across a fault that separates elastic media of different elastic constants. *Journal of Geophysical Research: Solid Earth*, 85(B3), 1455–1461.
- [336] Wegler, U. & Sens-Schönfelder, C. (2007). Fault zone monitoring with passive image interferometry. *Geophysical Journal International*, 168(3), 1029–1033.
- [337] Weichert, D., Horner, R. B., & Evans, S. G. (1994). Seismic signatures of landslides: The 1990 brenda mine collapse and the 1965 hope rockslides. *Bulletin of the Seismological Society of America*, 84(5), 1523–1532.

- [338] Wells, D. L. & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*, 84(4), 974–1002.
- [339] Wen, Y.-Y., Ma, K.-F., & Fry, B. (2018). Multiple-fault, slow rupture of the 2016 m w 7.8 kaikōura, new zealand, earthquake: Complementary insights from teleseismic and geodetic data. *Bulletin of the Seismological Society of America*, 108(3B), 1774–1783.
- [340] Weng, H., Yang, H., Zhang, Z., & Chen, X. (2016). Earthquake rupture extents and coseismic slips promoted by damaged fault zones. *Journal of Geophysical Research: Solid Earth*, 121(6), 4446–4457.
- [341] Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic Mapping Tools: Improved Version Released. *Eos, Transactions American Geophysical Union*, 94(45), 409–410.
- [342] Wibberley, C. A. J. & Shimamoto, T. (2005). Earthquake slip weakening and asperities explained by thermal pressurization. *Nature*, 436(7051), 689–692.
- [343] Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., & Trujillo, J. (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bulletin of the Seismological Society of America*, 88(1), 95–106.
- [344] Xu, Y., Koper, K. D., Sufri, O., Zhu, L., & Hutko, A. R. (2009). Rupture imaging of the Mw 7.9 12 May 2008 Wenchuan earthquake from back projection of teleseismic P waves. *Geochemistry, Geophysics, Geosystems*, 10(4), Q04006.
- [345] Yagi, Y. & Fukahata, Y. (2011). Rupture process of the 2011 Tohoku-oki earthquake and absolute elastic strain release. *Geophysical Research Letters*, 38(19).
- [346] Yagi, Y., Nakao, A., & Kasahara, A. (2012). Smooth and rapid slip near the Japan Trench during the 2011 Tohoku-oki earthquake revealed by a hybrid back-projection method. *Earth and Planetary Science Letters*, 355–356, 94–101.
- [347] Yagi, Y. & Okuwaki, R. (2015). Integrated seismic source model of the 2015 Gorkha, Nepal, earthquake. *Geophysical Research Letters*, 42(15), 2015GL064995.
- [348] Yang, H., Lin, J., Yin, J., & Yao, H. (2015). Tectonic settings of the 2015 m w 8.3 coquimbo, chile earthquake and its implications on megathrust earthquakes. *Chinese Science Bulletin*, 60(36), 3549–3556.
- [349] Yao, H., Gerstoft, P., Shearer, P. M., & Mecklenbräuker, C. (2011). Compressive sensing of the Tohoku-Oki Mw 9.0 earthquake: Frequency-dependent rupture modes. *Geophysical Research Letters*, 38(20), L20310.

- [350] Yao, H., Shearer, P. M., & Gerstoft, P. (2012). Subevent location and rupture imaging using iterative backprojection for the 2011 Tohoku Mw 9.0 earthquake. *Geophysical Journal International*, 190(2), 1152–1168.
- [351] Yao, H., Shearer, P. M., & Gerstoft, P. (2013). Compressive sensing of frequency-dependent seismic radiation from subduction zone megathrust ruptures. *Proceedings of the National Academy of Sciences*, 110(12), 4512–4517.
- [352] Ye, L., Kanamori, H., & Lay, T. (2018). Global variations of large megathrust earthquake rupture characteristics. *Science Advances*, 4(3), ea04915.
- [353] Ye, L., Lay, T., Kanamori, H., & Koper, K. D. (2015). Rapidly Estimated Seismic Source Parameters for the 16 September 2015 Illapel, Chile Mw 8.3 Earthquake. *Pure and Applied Geophysics*, (pp. 1–12).
- [354] Ye, L., Lay, T., Kanamori, H., & Rivera, L. (2016a). Rupture characteristics of major and great ( $m_w \geq 7.0$ ) megathrust earthquakes from 1990 to 2015: 1. source parameter scaling relationships. *Journal of Geophysical Research: Solid Earth*, 121(2), 826–844.
- [355] Ye, L., Lay, T., Kanamori, H., & Rivera, L. (2016b). Rupture characteristics of major and great ( $m_w \geq 7.0$ ) megathrust earthquakes from 1990 to 2015: 2. depth dependence. *Journal of Geophysical Research: Solid Earth*, 121(2), 845–863.
- [356] Ye, S., Flueh, E. R., Klaeschen, D., & von Huene, R. (1997). Crustal structure along the EDGE transect beneath the Kodiak shelf off Alaska derived from OBH seismic refraction data. *Geophysical Journal International*, 130(2), 283–302.
- [357] Yin, J. & Denolle, M. A. (2019). Relating teleseismic backprojection images to earthquake kinematics. *Geophysical Journal International*, 217(2), 729–747.
- [358] Yin, J., Denolle, M. A., & Yao, H. (2018). Spatial and Temporal Evolution of Earthquake Dynamics: Case Study of the Mw 8.3 Illapel Earthquake, Chile. *Journal of Geophysical Research: Solid Earth*, 123(1), 344–367.
- [359] Yin, J., Yang, H., Yao, H., & Weng, H. (2016). Coseismic radiation and stress drop during the 2015 Mw 8.3 Illapel, Chile megathrust earthquake. *Geophysical Research Letters*, 43(4), 1520–1528.
- [360] Yin, J. & Yao, H. (2016). Rupture and frequency-dependent seismic radiation of the 2012 Mw 8.6 Sumatra strike-slip earthquake. *Geophysical Journal International*, 205(3), 1682–1693.
- [361] Yin, J., Yao, H., Yang, H., Qin, W., Jing, L.-Z., & Zhang, H. (2017). Frequency-dependent rupture process, stress change, and seismogenic mechanism of the 25 April 2015 Nepal Gorkha Mw 7.8 earthquake. *SCIENCE CHINA Earth Sciences*, 60(4), 796–808.

- [362] Yue, H., Castellanos, J. C., Yu, C., Meng, L., & Zhan, Z. (2017). Localized water reverberation phases and its impact on backprojection images. *Geophysical Research Letters*, 44(19), 9573–9580.
- [363] Yue, H., Lay, T., & Koper, K. D. (2012). En échelon and orthogonal fault ruptures of the 11 April 2012 great intraplate earthquakes. *Nature*, 490(7419), 245–249.
- [364] Yue, H., Simons, M., Duputel, Z., Jiang, J., Fielding, E., Liang, C., Owen, S., Moore, A., Riel, B., Ampuero, J. P., & Samsonov, S. V. (2016). Depth varying rupture properties during the 2015 Mw 7.8 Gorkha (Nepal) earthquake. *Tectonophysics*.
- [365] Zhan, Z., Kanamori, H., Tsai, V. C., Helmberger, D. V., & Wei, S. (2014). Rupture complexity of the 1994 Bolivia and 2013 Sea of Okhotsk deep earthquakes. *Earth and Planetary Science Letters*, 385, 89–96.
- [366] Zhang, H., van der Lee, S., & Ge, Z. (2016). Multi-array rupture imaging of the devastating 2015 gorkha, nepal earthquake sequence. *Geophysical Research Letters*.
- [367] Zhou, L., Zhao, C., Zhang, M., Xu, L., Cui, R., Zhao, C., Duan, M., & Luo, J. (2021). Machine-learning-based earthquake locations reveal the seismogenesis of the 2020 mw 5.0 qiaojia, yunnan earthquake. *Geophysical Journal International*.
- [368] Zhu, J., Canales, J. P., Han, S., Carbotte, S. M., Arnulf, A., & Nedimović, M. R. (2020a). Vp/vs ratio of incoming sediments off cascadia subduction zone from analysis of controlled-source multicomponent obs records. *Journal of Geophysical Research: Solid Earth*, 125(6), e2019JB019239.
- [369] Zhu, J., Kopp, H., Flueh, E. R., Klaeschen, D., Papenberg, C., & Planert, L. (2009). Crustal structure of the central Costa Rica subduction zone: Implications for basal erosion from seismic wide-angle data. *Geophysical Journal International*, 178(2), 1112–1131.
- [370] Zhu, L., Peng, Z., McClellan, J., Li, C., Yao, D., Li, Z., & Fang, L. (2019a). Deep learning for seismic phase detection and picking in the aftershock zone of 2008 mw7.9 wenchuan earthquake. *Physics of the Earth and Planetary Interiors*, 293, 106261.
- [371] Zhu, W., Mousavi, S. M., & Beroza, G. C. (2019b). Seismic signal denoising and decomposition using deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(11), 9476–9488.
- [372] Zhu, W., Mousavi, S. M., & Beroza, G. C. (2020b). Seismic signal augmentation to improve generalization of deep neural networks. *Advances in Geophysics*, 61, 151–177.