PROPAGATION OF REGIONAL SEISMIC PHASES(S_n AND L_g) IN THE MIDDLE EAST AND EAST ASIA

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

HONGJUN HUI

Dr. Eric A. Sandvol, Dissertation Supervisor

JULY 2021

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled:

PROPAGATION OF REGIONAL SEISMIC PHASES (S_n and L_g)
IN THE MIDDLE EAST AND EAST ASIA

presented by Hongjun Hui,

a candidate for the degree of Doctor of Philosophy and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Eric Sandvol

Professor Mian Liu

Professor Francisco Gomez

Professor Martin Appold

Professor Scott H. Holan

ACKNOWLEDGMENTS

Firstly, I want to thank my dissertation advisor, Dr. Eric A. Sandvol. In the journey to pursue my doctoral degree, he has always been patient and passionate while advising me about my study and research. His grants have not only financially supported my research life through the years in Columbia, but also paid my travels to each academic meeting, where I experienced a 'hot spot' in earth science research and met people from all over the world. He has always been supportive and understanding when I have to request a leave of absence from work. I would also express my gratitude to Eric's family, Christine and Paul. They were friendly to invite me to dinner in their house. Christine also helped a lot with revising my comprehensive proposal and dissertation.

Secondly, I would like to acknowledge the members of my dissertation committee: Dr. Mian Liu, Dr. Francisco (Paco) Gomez, Dr. Martin Appold, Dr. Scott Holan and Dr. Eric Sandvol. I have learned both academic knowledge and research methods from classes by Mian and Eric. As Director of Graduate Studies, Paco has helped me through all the steps towards my doctoral degree. He has always been there helping to revise my study plan. I really appreciate Scott and Martin for agreeing to be my committee member although I asked them at the last moment.

Last but not least, my deepest gratitude goes to my family. I would not be who I am without them. My wife, Huiqian, and my mother and father always show their support and understanding no matter what. I would like to thank my lovely daughter Minlang for bringing a lot of fun to my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PROPAGATION OF REGIONAL SEISMIC PHASES (S_n AND L_g) IN THE MIDDLE EAST AND EAST ASIA

Hongjun Hui

Dr. Eric Sandvol, Dissertation Supervisor

ABSTRACT

I have studied the propagation of the regional seismic waves (S_n and L_g) in the Middle East (ME) and East Asia. The methods used in this study are : (1) the reverse two-station method (RTM) to determine the attenuation and site response, (2) the logistic regression/Bayesian lasso method to predict the probability of observing S_n and (3) the LSQR method to generate attenuation and predicted probability tomography.

I have constructed new frequency-dependent L_g and S_n attenuation models in China. The results show high L_g Qs in SE and NE China and the Tarim basin, while low L_g Qs are found in the Tibetan plateau. High S_n Qs are found in the Sichuan basin and the Ordos and Tibetan plateaus. Low S_n Qs are found in eastern China and northern, eastern Tibetan plateau. I have also estimated the site response with both L_g and S_n in China. With the differential site response results, I found that the effective S_n attenuation may include the effect of the crustal legs, which is supposed to be eliminated in RTM. I have also constructed a model that shows the spatial variability in S_n propagation in the ME and East Asia.

The model has successfully predicted the probability of observing S_n as well as being blocked. We observe high probabilities of observing S_n in stable and oceanic regions, such as the Arabian Plate, the Mediterranean Sea. In tectonically active areas, such as the Tibetan and Iranian plateaus, I have observed a low probability of observing S_n propagation. These models should help determine where existing attenuation models are significantly biased by systematic data censorship.

Chapter 1

Introduction

1.1 Regional Seismic Phases

Regional earthquakes are events that occur at an epicentral distance range from 3° to 20°. Regional seismograms are dominated by seismic energy propagating within the lithosphere (e.g., Bao et al., 2012). There are potentially four major high frequency phases observed on a typical regional seismogram. In the order of arrival time, they are P_n, P_g, S_n, and L_g (Figure 1.1). P_n and P_g are two compressional waves on regional seismograms, propagating within the uppermost mantle and crust, respectively. The two high frequency shear phases on regional seismograms are S_n and L_g, propagating within the uppermost mantle and crust, respectively. In this study, we use the S_n and L_g phases to study the blockage, attenuation and site response across the Middle East (ME) and East Asia.

L_g is usually the most prominent and stable phase on regional seismograms for continental paths (Sereno, 1990; Rapine et al., 1997; Baumgardt, 2001) and propagates within the crust as a guided shear wave with velocity and frequency band of 2.9 - 3.7 km/s and 0.5 - 5.0 Hz. L_g has been interpreted as the superposition of

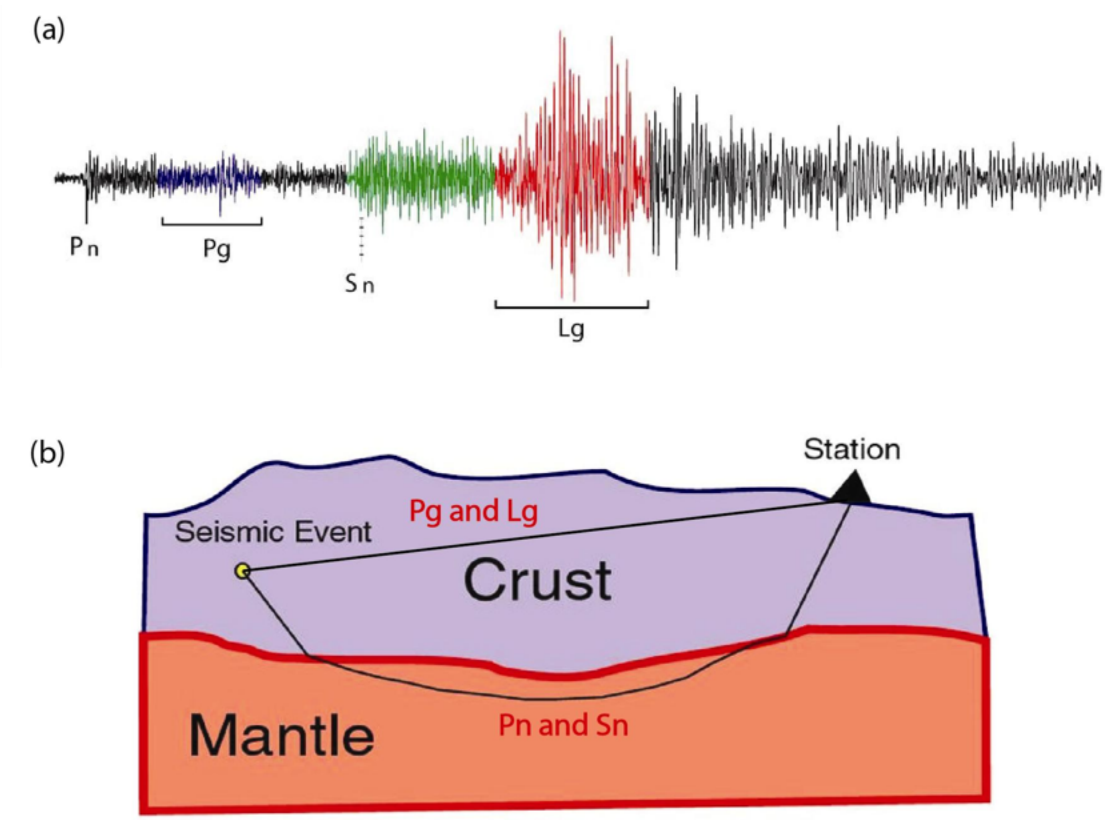


Figure 1.1: Regional seismogram example (a) and ray paths (b).

higher-mode Rayleigh wave that primarily propagates in the crust (Knopoff et al., 1973) and, alternatively, as the superposition of super-critically reflected shear waves in the continental crust (Campillo et al., 1985; Kennett, 1986). When observed on vertical component signal, L_g may be primarily associated with Rayleigh wave overtones, although scattering may mix Love and Rayleigh energy. L_g is likely fairly insensitive to earthquake radiation patterns; thus it has particular value for seismic magnitude estimation(Bao, 2011). Because L_g is dominated by shear wave energy, it tends to be more strongly excited by earthquakes than explosions (Fan and Lay, 2002). Because it fully samples the crust, L_g phase attenuation has widely been used to investigate the rheology and structure of the crust. L_g is strongly attenuated or blocked in some continental areas with significant changes in crustal thickness, such as mountain belts and basins (Zhao et al., 2003; Xie, 2002). Strong attenuation of L_g phase, either scattering or intrinsic, may lead to phase blockage. But it is still not possible to determine how three-dimensional L_g Q varies with depth(Bao, 2011).

The S_n phase is also a high-frequency guided wave. Unlike the L_g phase, S_n travels through the lithospheric mantle with a frequency that is usually between 0.5 and 4 Hz, with a velocity around 4.7 km/s (Sandvol et al., 2001), so that S_n is a function of the properties (i.e., velocity and attenuation) of the uppermost mantle. S_n arrives as a high-frequency wave-train lasting from tens of seconds to several minutes (Sandvol et al., 2001) with a faster velocity in stable continental and oceanic regions than in tectonically active regions. Although S_n propagates efficiently in stable continental and shield regions (Ni and Barazangi, 1983; Gök et al., 2000; Sandvol et al., 2001; Gök et al., 2003) and has been recorded with epicentral distances up to 35° (Molnar and Oliver, 1969; Huestis et al., 1973), it is usually blocked or highly attenuated in tectonically active regions with high heat flow (Molnar and Oliver, 1969; Kadinsky-Cade et al., 1981; Ni and Barazangi, 1983; McNamara and Owens, 1995; Gök et al., 2000; Calvert et al., 2000; Sandvol et al., 2001).

1.2 Regional Seismic Phase Attenuation

Seismic attenuation is a key parameter of seismic wave propagation which is a measure of amplitude reduction with distance. The quality factor, Q, is the reciprocal of attenuation. There are two attenuation mechanisms, scattering and intrinsic attenuation. Intrinsic attenuation is defined as the energy loss per cycle of oscillation resulting from internal friction across mineral grain boundaries. Therefore, intrinsic attenuation represents the anelasticity in the Earth and can be used to study the temperature throughout the mantle and crust (e.g. Karato, 1993; Knopoff, 1964).

While scattering attenuation is caused by an heterogeneous velocity structure along the seismic wave propagation path. Scattering generally reduces amplitude through the redirection of seismic energy away from the direction of propagation. Scattering attenuation is determined by the dimension of heterogeneous velocity structure in the Earth, which could be interpreted as slabs, faults, and other geological structures.

However, it is difficult to separate the two attenuation mechanisms in real data. Currently we can only calculate effective attenuation, which is the combined effect of scattering and intrinsic attenuation. Strong attenuation of regional seismic waves and low velocity area is generally thought to be related to partial melting and high temperatures in the lithosphere, while strong attenuation in high velocity areas may be interpreted as composition anomaly.

1.3 Tectonic Setting of the ME

This study focuses on two areas, the ME (Figures 1.2 and 1.3)and East Asia (Figure 1.4 and 1.5). Figures 1.2 and 1.3 show the topography and tectonic setting of the ME region, respectively. In Figure 1.3, NAF is North Anatolian Fault; EAF is East Anatolian Fault; AA is Aegean Arc; CA is Cyprian Arc; DSF is Dead Sea Fault; LS

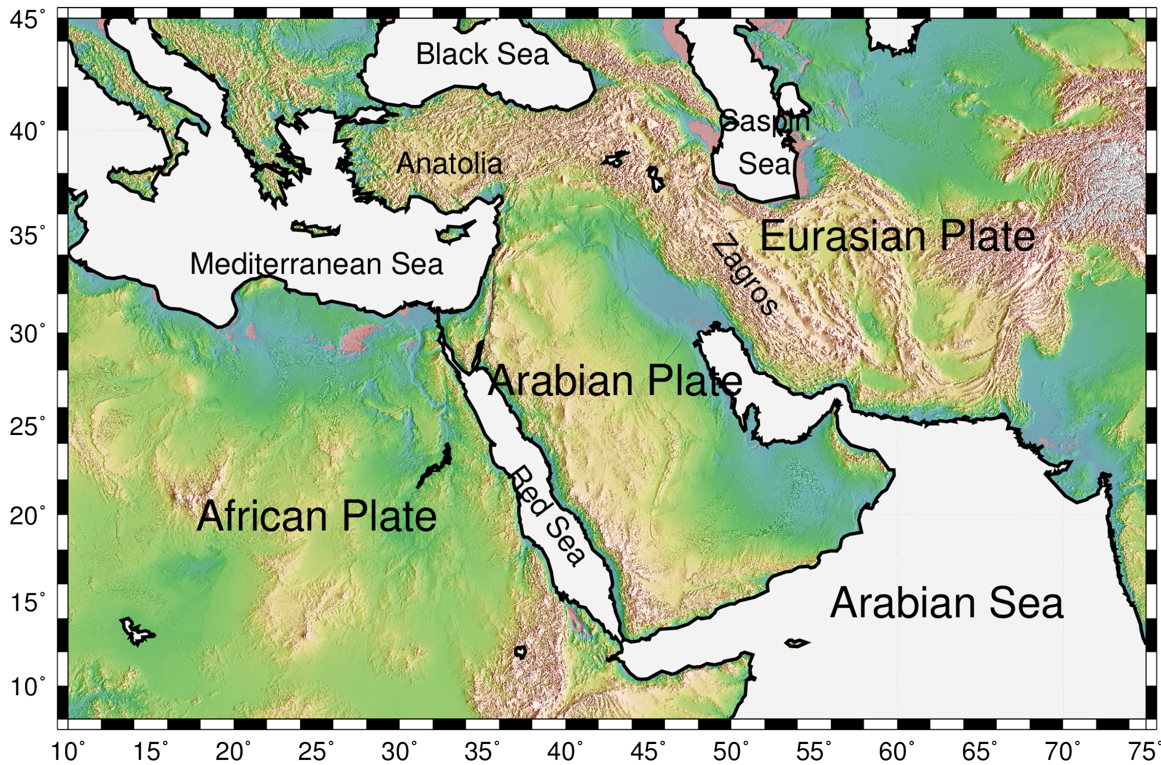


Figure 1.2: A topography plot of the ME.

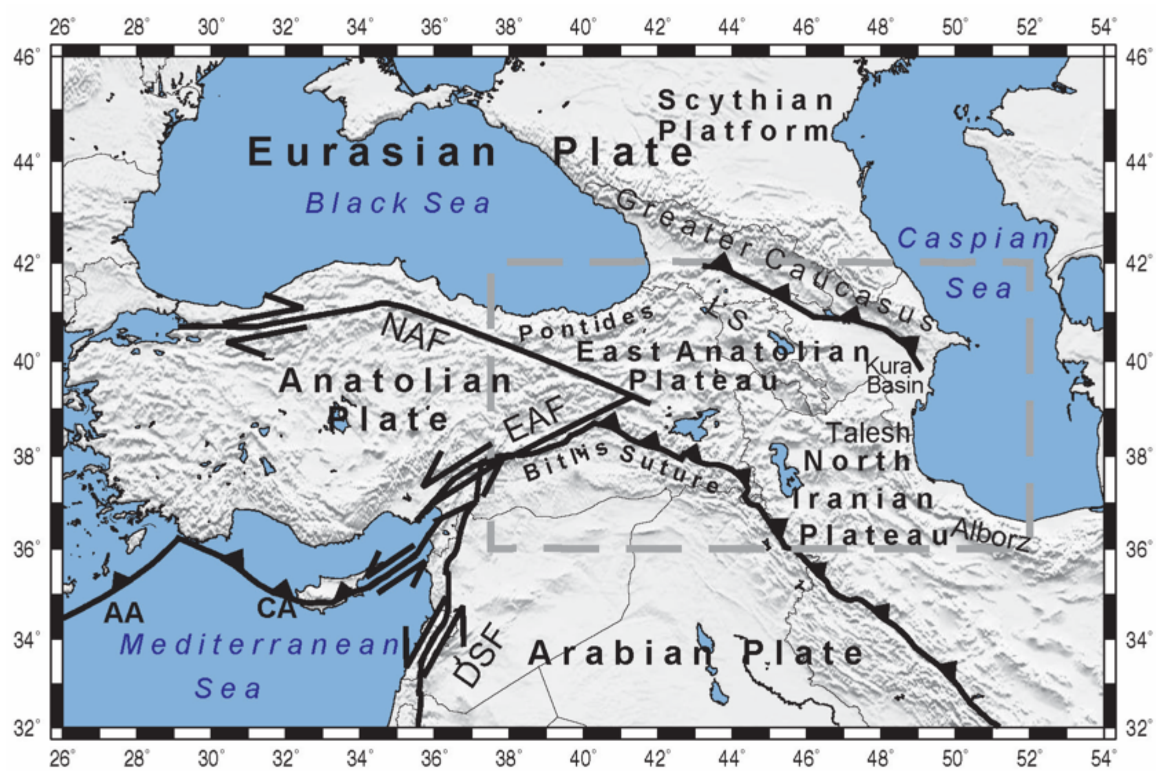


Figure 1.3: ME tectonic setting (Skobeltsyn, 2014).

is Lesser Caucasus.

The ME region is both tectonically and seismically active with continental break-up, collision, back-arc extension and westward escape tectonics. This region has been subject of extensive geological and geophysical studies during the past several decades. Blockage of the regional phases is widespread along the Bitlis-Zagros and the eastern and southern Anatolian plateau (Figure 1.3).

The subduction of Neo-Tethyan lithosphere beneath Eurasia, the geometry of the subducting slab, the timing of the eventual continental collisions, and the occurrence of possible slab break-off all vary along the strike of the collisional boundaries. Geological evidence and tectonic reconstructions suggest that the northward subduction of the Neo-Tethyan ocean beneath the southern margin of Eurasia initiated in the Early Jurassic to the Early Cretaceous (Agard et al., 2011; Richards, 2015). Hafkenscheid et al. (2006) suggest that the early slab break-off first occurred beneath the northern Zagros suture zone in the early Oligocene, followed by both eastward and westward propagation of a slab tear. Consisting of a Precambrian shield bounded by a sedimentary platform, the Arabian plate behaved as a rigid plate moving NE to NNE, producing spreading in the Red Sea and Gulf of Aden and collision with the Eurasian plate along the Bitlis-Zagros during the middle to late Pliocene (Phillip et al., 1989). Convergence in Iran is accommodated by distributed horizontal shortening across most of the plateau. The Anatolian plate moves at present coherently at average rate of $\sim$ 2 cm/yr, separated from the Eurasian plate by the NAF (Reilinger et al., 2006). As a result of the young and active continental collision, the Anatolian-Iranian plateau and Zagros mountains formed. Most of the Lesser and Greater Caucasus are believed to have formed within the same time frame as the Arabian-Eurasian collision (Phillip et al., 1989). The Anatolian block is escaping to the west, as evidenced by the right-lateral strike-slip movement along the NAF system and also by Global Positioning System (GPS) data (Ahadov and Jin, 2017). Faccenna et al. (2013) suggest that the progressive evolution of the Tethyan subduction resulted in back-arc extension in the Aegean.

1.4 Tectonic Setting of East Asia

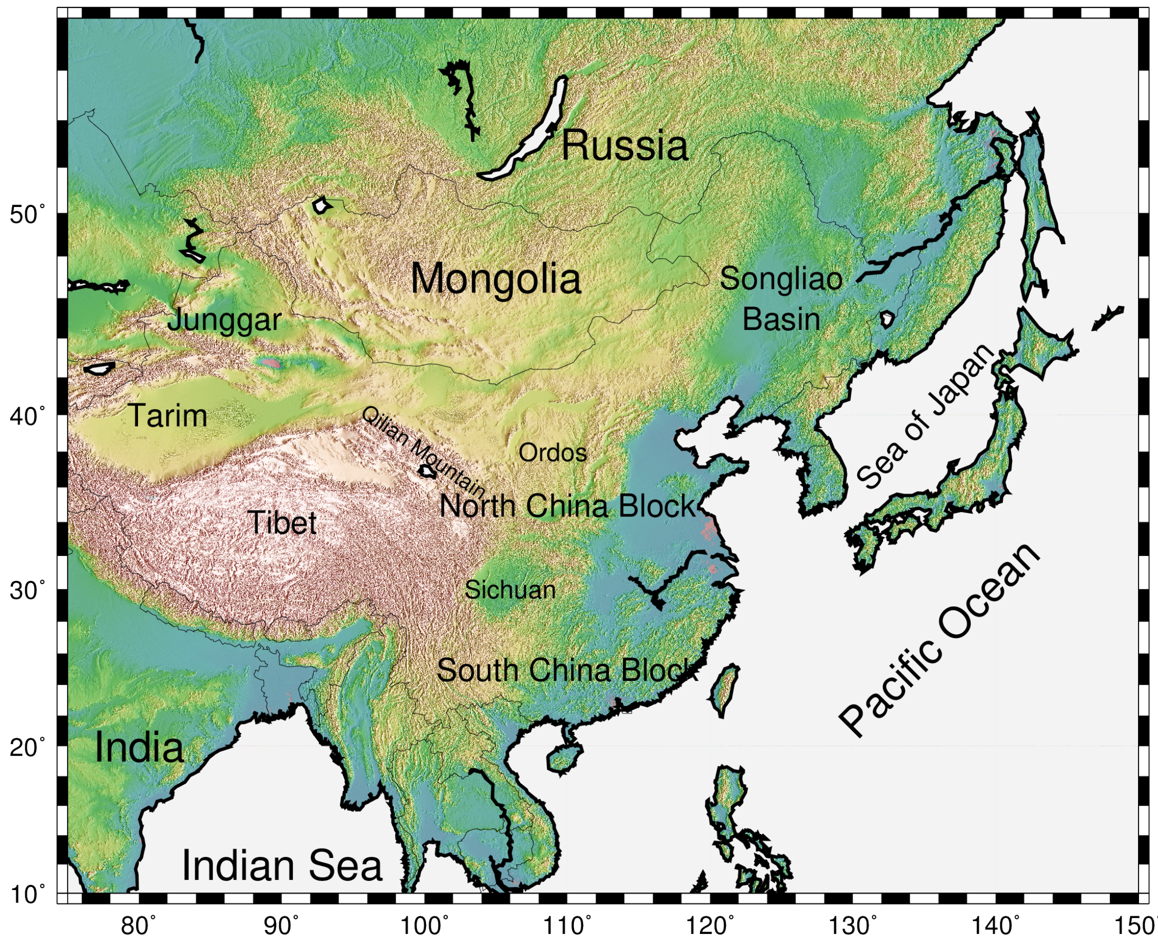


Figure 1.4: East Asia topography plot.

East Asia has a complex tectonic history with collision between the Indian and Eurasian plates, subduction of the Pacific plate beneath the Eurasian plate, and complex topographic change from Tibetan plateau to the Marianas Trench. To the west, the Tibetan plateau resulted from the collision between the Indian and Eurasian plates. This convergence began $\sim$ 50 Ma and the Indian plate is continuously subducting or underthrusting beneath Tibet (e.g., Yin and Harrison, 2000; Chen et al., 2017). The northward collision of the Indian plate with the Eurasian plate caused the Himalaya and Tien Shan orogenies and crustal shortening and uplifting of the Tibetan plateau, accompanied by eastward extrusion of portions of the plateau lithosphere (Yin and Harrison, 2000; Wang and Shen, 2020). At the northern end of our study area, there is also regional deformation related to the India-Eurasia convergence. Mongolia has been regarded as one of the most tectonically active intracontinental regions in the world (Choi et al., 2018), although compared to mainland China, GPS data show relatively small deformation rates (∼4 mm/yr, Wang and Shen, 2020). At the eastern margin of the Eurasia plate, the Pacific and Philippine Sea plates subducted underneath the Eurasian plate, causing destruction of parts of the North China craton with significant seismic and volcanic activity (Wang and Shen, 2020).

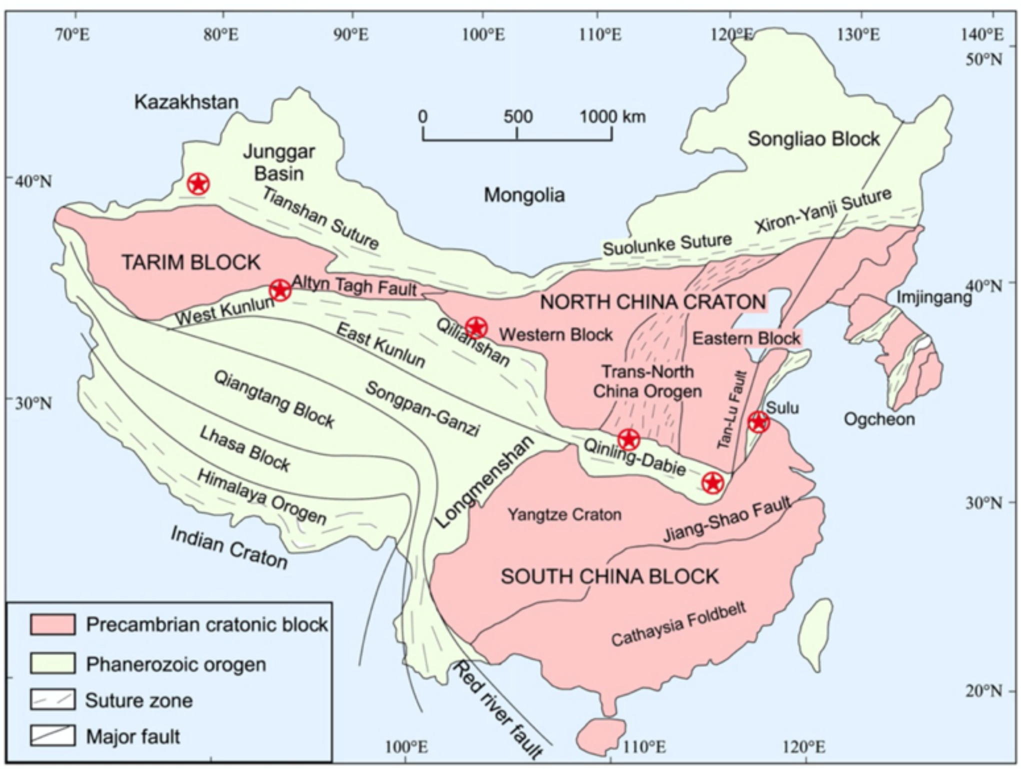


Figure 1.5: East Asia tectonic setting plot(Zheng et al., 2013).

China is composed of four major blocks: the Tarim Block, the Tibetan plateau, the North China Craton (NCC), and the South China Block (SCB) (Figure 1.5). These four blocks are formed by the Tethyan domain, the Western Pacific Domain and the Paleo Asian Ocean converging in a triangular setting, and divided by orogenic belts, faults and sutures (Zheng et al., 2013). China has tectonically undergone different stages of amalgamation, extension and compression.

The NCC is tectonically dominated by westward subduction of the Pacific plate and the India-Asia collision zone, bounded by faults and orogenic belts. The subduction of the Pacific plate has apparently generated an asthenospheric up-welling that has served to thermally and tectonically reactivate large portions of the NCC. These include the Ordos block and the eastern block, separated by the Trans-North China orogenic belts. The eastern block has undergone reactivation since the Mesozoic, characterized by thin lithosphere, high heat flow and extensive basins, while the western block (the Ordos plateau) remains a stable lithosphere, which is thought to be cold and thick. The eastern margin of the Ordos block is the Shanxi rift, an active rift zone formed during the Quaternary with intense seismicity and volcanoes, forming a distinct contrast to the stable lithosphere within the Ordos plateau (Molnar and Tapponnier, 1977). The Weihe Graben is located on the southern edge of the Ordos plateau, marked by a set of EW direction left lateral normal faults with steep dip (Wesnousky et al., 1984). Previous seismic studies reveal that the western block (Ordos) maintains a thick lithosphere and the eastern block progressively thins westward, from $\leq$ 200 km to $\sim$ 60 - 100 km (Tian et al., 2009; Zhao et al., 2008). The SCB is divided into the Yangtze Caton in the northwest and the Cathaysia Foldbelt in the southeast. These two blocks are considered to have collided along the Jiangnan orogenic belt. The Yangtze Craton consists of minor Archean–Paleoproterozoic crystalline basement (e.g. Kongling Complex) that is surrounded by late Mesoproterozoic to early Neoproterozoic orogens. The magmatic rocks are uncon- formably overlain by weakly metamorphosed Neoproterozoic strata (e.g., Banxi Group) and unmetamorphosed Sinian cover. Unlike the Yangtze Craton that contains the Archean basement, the Cathaysia Foldbelt is composed predominantly of Neoproterozoic basement rocks with a minor occurrence of Paleoproterozoic rocks in southwest Zhejiang and north Fujian, and Mesoproterozoic rocks in Hainan Island (Zheng et al., 2013). The SCB is separated from the Pacific subduction zone in the east by the Okinawa trough, a back-arc basin formed by the subduction process and on the west by the Longmenshan fault and the Xianshuihe fault, which are driven by the India-Eurasia collision. To the east along the Xianshuihe fault is the Chuan Dian fragment, which is bounded by the Sichuan basin on the northeast and the Songpan-Ganzi fold belt on the north. The collision of the SCB and NCC formed the Qinling-Dabie orogenic belt and the Sichuan Basin (Zheng et al., 2013). The Tarim block is the largest basin in China. It has Neoarchean basement overlain by sedimentary and volcanic strata (Zhao et al., 2012). The presence of Archean to Proterozoic rocks makes the Tarim basin a cratonic block that is bounded by the Kunlun orogen to the south and the Tianshan orogen to the north. At the southeastern boundary, the Altyn Tagh fault separates the Tarim block from the Qaidam basin and the Qilian orogenic belt. Ambient noise studies shows that velocity varies laterally in the lower crust and uppermost mantle beneath the basin, but overall the average shear wave velocity in the lower and uppermost mantle is higher than the surrounding areas, forming a steep velocity gradient (Li et al., 2012). The Tibetan plateau consist of several major blocks: the Lhasa, Qiangtang, and Songpan-Ganzi terranes. This orogenic system is formed by progressive closure of the Tethyan Ocean during the convergence between the Indian and Eurasian continents. It is bounded by Himalayan orogen at the southern edge. The topography changes sharply at both edges of the plateau. The Longmenshan fault zone separates the plateau from the SCB. The Tarim block is to the northwest and the Ordos block is to the northeast. Ambient noise and receiver function studies show thick crust beneath the Tibetan plateau with lateral velocity variations from the southern to the northern parts of the plateau (Li et al., 2012). Several models have been proposed to explain the evolution of the Tibetan plateau, including extrusion, shortening and ductile flow in the lower crust.

Chapter 2

Methodology

2.1 Regional Seismic Phase Attenuation

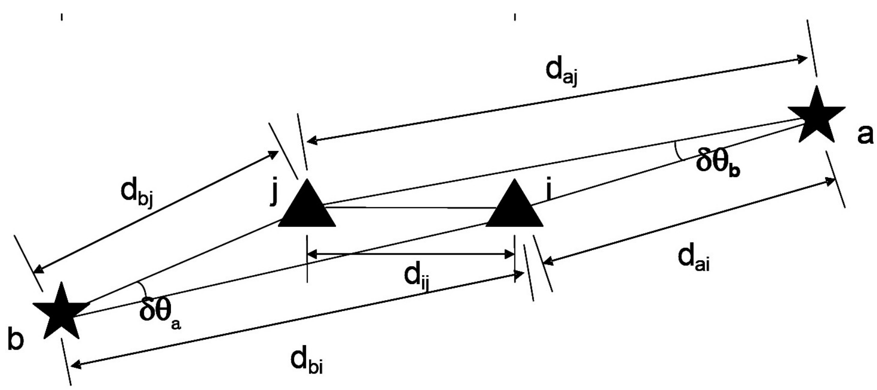


Figure 2.1: RTM scheme. a and b denote two sources. i and j are two stations. d denotes distance and $\delta\theta$ is azimuth difference.

Regional seismic attenuation (Q^{-1}) can provide important constraints on the properties of the lithosphere, including the rheology and temperature of the crust and uppermost mantle. In addition, understanding the propagation of regional phases is important for the development of seismic source discriminant and yield estimation of nuclear explosions (Bao, 2011). There are several deterministic methods used to calculate the regional seismic attenuation, and the discrepancies between methods are due to differing model parameterizations datasets and the nature of the methodologies (Ford et al., 2008).

The method used in this study is the reverse two-station method (RTM). RTM (Figure 2.1) is based upon the spectral amplitude ratio of the regional phases and the Q is more stable than that of other methods due to the separation of the source term, the station response, and the instrument response. In Figure 2.1, black triangles denote stations and black stars are sources. ds are distances and θs are azimuth difference.

The amplitude of regional phase in frequency domain can be determined by

$$A = S \cdot R \cdot G \cdot I_S \cdot S_S \cdot I_Q \cdot T_{FD} \cdot T_{SC} \cdot T_{AN}, \tag{2.1}$$

where S denotes source excitation function; R, the focal mechanism factor; G, the geometrical spreading function (e.g., $G = G_0 d^{-m}$ for L_g and S_n, d and m denote distance and a constant); I_S, the instrument response; S_S, the station site response; I_Q, intrinsic attenuation; T_{FD}, the coefficient of focusing and de-focusing; T_{SC}, the scattering coefficient; and T_{AN}, other effects.

T_{FD} is typically assumed to be 1 if we set a limit on standard errors of the measurements (Xie, 2002) and I_Q, T_{SC}, T_N are interpreted as effective attenuation. Thus, the deterministic representation of spectral amplitude can be simplified as

$$A(f,d) = S(f) \cdot R(f,\varphi) \cdot I(f) \cdot S_S(f) \cdot G(d) \cdot \exp(-\frac{\pi f d}{vQ}), \tag{2.2}$$

where f denotes frequency; d, the epicentral distance; φ, azimuth; and v, the seismic velocity.

We use i, j to denote station indices and a, b for event indices (Figure 2.1). The four spectral amplitude equations are

$$\begin{cases} A_{ai}(f, d_{ai}) = S_a(f)R_a(f, \varphi)I_i(f)S_{S_i}(f)G_{ai}(d_{ai}) \exp\left(-\frac{\pi f d_{ai}}{v_i Q_i}\right) \\ A_{aj}(f, d_{aj}) = S_a(f)R_a(f, \varphi)I_j(f)S_{S_j}(f)G_{aj}(d_{aj}) \exp\left(-\frac{\pi f d_{aj}}{v_j Q_j}\right) \\ A_{bi}(f, d_{bi}) = S_b(f)R_b(f, \varphi)I_i(f)S_{S_i}(f)G_{bi}(d_{bi}) \exp\left(-\frac{\pi f d_{bi}}{v_i Q_i}\right) \\ A_{bj}(f, d_{bj}) = S_b(f)R_b(f, \varphi)I_j(f)S_{S_j}(f)G_{bj}(d_{bj}) \exp\left(-\frac{\pi f d_{bj}}{v_j Q_j}\right) \end{cases} \tag{2.3}$$

Assume that the velocity structure is 1-D and effective Q is identical between stations i and j. From Equation 2.3, we get

$$\frac{A_{ai}A_{bj}}{A_{aj}A_{bi}} = \left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^{-m} \exp\left[\frac{\pi f}{vQ}(d_{aj} - d_{ai} - d_{bj} + d_{bi})\right]. \tag{2.4}$$

Thus, attenuation could be determined by,

$$\frac{1}{Q} = \frac{v}{\pi f(d_{aj} - d_{ai} - d_{bj} + d_{bi})} \times \ln\left[\frac{A_{ai}A_{bj}}{A_{aj}A_{bi}}\left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^{m}\right]. \tag{2.5}$$

To get Q tomography from ray-path Q (Equation 2.5), we discretize the study area into M discrete cells. Using this formulation, the effective attenuation of n-th path could be expressed as,

$$\exp\left(-f\pi\frac{d_n}{Q_n v_n}\right) = \exp\left[-f\pi\sum_{m=1}^{M}\left(\frac{d_{nm}}{Q_{nm}v_{nm}}\right)\right].$$

In RTM, because of the 1D velocity assumption ($v_{nm} \equiv v_n$), we get,

$$\frac{d_n}{Q_n} = \sum_{m=1}^{M}\left(\frac{d_{nm}}{Q_{nm}}\right). \tag{2.6}$$

To solve these linear equations, we use the LSQR algorithm (Paige and Saunders, 1982) to get 2D Q tomographic model. The geometrical spreading factor (m in

Equation 2.5) is assumed to be 1 for S_n and 0.5 for L_g. The azimuth tolerance $\delta\theta_a/\delta\theta_b$ is set to be $\pm 15°$ (Xie, 2002).

2.2 Regional Seismic Phase Site Response

The method we use to determine the site response is revised version of RTM. From Equation 2.3, we get

$$\begin{cases} \dfrac{A_{ai}}{A_{aj}} = \dfrac{I_i S_{S_i} G_{ai}}{I_j S_{S_j} G_{aj}} \exp\left(\dfrac{\pi f d_{aj}}{v_j Q_j} - \dfrac{\pi f d_{ai}}{v_i Q_i}\right) \\ \dfrac{A_{bi}}{A_{bj}} = \dfrac{I_i S_{S_i} G_{bi}}{I_j S_{S_j} G_{bj}} \exp\left(\dfrac{\pi f d_{bj}}{v_j Q_j} - \dfrac{\pi f d_{bi}}{v_i Q_i}\right) \end{cases} \tag{2.7}$$

By multiplying the two ratios in Equation 2.7, we get

$$\frac{A_{ai}A_{bi}}{A_{aj}A_{bj}} = \left(\frac{I_i S_{S_i}}{I_j S_{S_j}}\right)^2 \left(\frac{G_{ai}G_{bi}}{G_{aj}G_{bj}}\right) \exp\left(\frac{\pi f d_{aj}}{v_j Q_j} - \frac{\pi f d_{ai}}{v_i Q_i} + \frac{\pi f d_{bj}}{v_j Q_j} - \frac{\pi f d_{bi}}{v_i Q_i}\right). \tag{2.8}$$

Suppose that the effective Q is identical along the path and substitute 1D velocity v and $G = G_0 d^{-m}$,

$$\begin{aligned} \frac{A_{ai}A_{bi}}{A_{aj}A_{bj}} &= \left(\frac{I_i S_{S_i}}{I_j S_{S_j}}\right)^2 \left(\frac{d_{ai}d_{bi}}{d_{aj}d_{bj}}\right)^{-m} \exp\left[\frac{\pi f}{vQ}(d_{aj} - d_{ai} + d_{bj} - d_{bi})\right] \\ &= \left(\frac{I_i S_{Si}}{I_j S_{Sj}}\right)^2 \left(\frac{d_{ai}d_{bi}}{d_{aj}d_{bj}}\right)^{-m} \exp\left\{\frac{d_{aj} - d_{ai} + d_{bj} - d_{bi}}{d_{aj} - d_{ai} - d_{bj} + d_{bi}} \ln\left[\frac{A_{ai}A_{bj}}{A_{aj}A_{bi}}\left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^m\right]\right\}. \\ &= \left(\frac{I_i S_{Si}}{I_j S_{Sj}}\right)^2 \left(\frac{d_{ai}d_{bi}}{d_{aj}d_{bj}}\right)^{-m} \exp\left\{\ln\left[\frac{A_{ai}A_{bj}}{A_{aj}A_{bi}}\left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^m\right]^{\frac{d_{aj}-d_{ai}+d_{bj}-d_{bi}}{d_{aj}-d_{ai}-d_{bj}+d_{bi}}}\right\} \end{aligned} \tag{2.9}$$

Then, the ratio of the two site responses can be expressed in logarithmic form

$$\ln \frac{S_{Si}}{S_{Sj}} = \ln S_{Si} - \ln S_{Sj}$$
$$= \ln \frac{I_j}{I_i} + \frac{d_{aj} - d_{ai}}{d_{aj} + d_{bi} - d_{ai} - d_{bj}} \ln \frac{A_{ai} d_{ai}^m}{A_{aj} d_{aj}^m} + \frac{d_{bi} - d_{bj}}{d_{aj} + d_{bi} - d_{ai} - d_{bj}} \ln \frac{A_{bi} d_{bi}^m}{A_{aj} d_{bj}^m}. \tag{2.10}$$

The relative site response can be solved determinedly by an inversion problem

$$\begin{bmatrix} 1 & -1 & 0 & 0 & \cdots \\ 1 & 0 & -1 & 0 & \cdots \\ 0 & 1 & -1 & 0 & \cdots \\ 0 & 0 & 1 & -1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \ln S_{S1} \\ \ln S_{S2} \\ \ln S_{S3} \\ \ln S_{S4} \\ \vdots \end{bmatrix} = \begin{bmatrix} RS_{12} \\ RS_{13} \\ RS_{23} \\ RS_{34} \\ \vdots \end{bmatrix}, \tag{2.11}$$

where the RS_{ij} denotes the relative site response between station i and j, in other words, the right hand side of Equation 2.10. Then, we use the LSQR to solve the inversion problem to get the site response for each station.

2.3 S_n Efficiency Tomography

Sandvol et al. (2001) proposed methods to tomographically map L_g and S_n propagation efficiencies in the ME. Here we only use this the method to image geographical variations in S_n propagation. Following the derivation of Phillip et al. (1989) and the starting model of Cong et al. (1996), they assumed that S_n phase amplitude (a_{ij}) is

$$a_{ij}(f) = a_{0i} x_{ij}^{-m} s_j(f) c_i(f) \exp\left[-\alpha(f) x_{ij}\right], \tag{2.12}$$

where i and j are the event and station indices respectively, f is the phase frequency, a_{0i} is the amplitude of the i-th source, x_{ij} is the total ray path length, m is the geometrical spreading parameter for S_n (assumed to be 1), s_j is the station response for the j-th station, c_i is the source scaling term for the i-th seismic source, and α is the spatial average attenuation coefficient and is assumed to be constant as $\pi f x_{ij}/V$ over a given frequency band. V is S_n group velocity and assumed to be 4.5 km/s. Taking the natural logarithm of Equation 2.12, correcting the station response term, discretizing the spatial attenuation factor, and ignoring the source scaling term, we get:

$$\tilde{A}_{ij}^{disc.} = \log\left(\frac{a_{0i}}{a_{ij}}\right) = \log e \frac{\pi f}{V} \sum_{l} m x_{ijl}, \tag{2.13}$$

where x_{ijl} represents the ray path length corresponding to the i-th source, j-th station and l-th cell.

Following Sandvol et al. (2001), we reviewed and classified the S_n propagation efficiencies into three categories: blocked S_n, efficient S_n, and inefficient S_n. If the seismogram shows no evidence of a discernable S_n phase, we categorized that path as a blocked S_n path and $\tilde{A}_{ij}^{disc.}$ in Equation 2.13 is set to be zero. If the S_n wave train could be observed, regardless of its strength or amplitude, we designated it as efficient S_n path and $\tilde{A}_{ij}^{disc.}$ in Equation 2.13 is set to two. If there is an ambiguous signal in the seismogram that potentially could be an S_n signal, it was classified as inefficient S_n path and $\tilde{A}_{ij}^{disc.}$ in Equation 2.13 is set to be one. After trying different signal to noise ratios using pre-phase S_n noise, we found that the best way to set the efficiency level is to individually visually inspect the seismograms. Using these definitions, the model parameter in Equation 2.13, m , becomes the weighted average efficiency for all ray paths passing through that cell. The weights correspond to the cell path length. Solving the linear equation 2.13 using LSQR allows us to quantitatively map the qualitative S_n efficiencies.

2.4 Predicted Probability Tomography

We use a logistic regression/Bayesian lasso model to predict the likelihood of observing S_n based on efficiency datasets. Specifically, suppose that we have N seismic rays discretized into p sections. Define $\mathbf{X} = (X_{ij1}^T, \ldots, X_{ijp}^T)$ as a N x p design matrix corresponding to the discretized sub-distances X_{ijl} $(l = 1, \ldots, p)$ as defined in the previous section and let z_i be the response variable, assumed to have two possible outcomes, 0 if ray i is observed and 1 if it is blocked. It is important to note that some of the sub-distances can be equal to zero for a given ray path. Then if θ is the probability of being observed, we propose a simple Binomial logistic regression model for $\mathbf{z}$ and define it as

$$\begin{aligned} \mathbf{z}|\theta &\sim Bernoulli(\theta) \\ \text{logit}(\theta) &= \mathbf{X}\boldsymbol{\beta} \end{aligned} \tag{2.14}$$

where the logit or log-odds transformation is defined as $\text{logit}(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$ and $\boldsymbol{\beta} \in \mathbb{R}^p$ is the vector of unknown regression coefficients. Since N is typically large, $\mathbf{X}$ is a sparse matrix and thus if $rank(X) < p$ (e.g., this can happen when $p > N$), there are infinitely many solutions under an ordinary least squares (OLS) approach. Even if $rank(X) = p$, for a large p, the OLS estimates will have a lot of variability, resulting in overfitting and consequently poor predictions for future observations not used in model training. Furthermore, it is often the case that some or many of the variables used in a multiple regression setting like this are in fact not all strongly associated with the response. Including such variables leads to unnecessary computational complexity in the model. By removing these variables, say, by setting the corresponding coefficient estimates to zero we can obtain a model that is much easily interpreted. One way to deal with these issues is shrinkage or regularization which allows us to substantially reduce the variance at the cost of a negligible increase in bias. Here we adopt a Lasso regularization as introduced by Tibshirani (1996). Defined as the stricter l_1-penalty, the Lasso approach can set coefficients to exactly zero, making it a useful tool for feature selection with lower variability. For the model in (5.4), the Lasso estimates of the regression coefficients, $\hat{\boldsymbol{\beta}}$ are defined as

$$arg\ min \left\{ \sum_{i=1}^{N} \left[\mathbf{z}_i - \text{logit}^{-1} \left(\sum_{l=1}^{p} \beta_l X_{ijl} \right) \right]^2 \right\} \qquad \text{subject to } \sum_l |\beta_l| \leq t. \tag{2.15}$$

where $t \geq 0$ is a tuning parameter. The parameter t controls the amount of shrinkage that is applied to the estimates. The Lasso constraint $\sum\limits_l |\beta_l| \leq t$ is equivalent to the addition of a Lagrangian penalty $\lambda \sum\limits_l |\beta_l|$ to the residual sum of squares (Murray and Overton, 1981). Now $|\beta_l|$ is proportional to the negative log-density of the Laplace distribution and hence we can derive the Lasso estimate as the Bayesian posterior mode under independent double-exponential priors for the β_ls of the form

$$\pi(\beta) = \prod_{j=1}^{p} \frac{\lambda}{2} e^{-\lambda|\beta_j|}. \tag{2.16}$$

Park and Casella (2008) developed a Gibbs sampler implementation of a fully Bayesian adaptation of the Lasso regularization exploiting the representation of the Laplace distribution as a scale mixture of Normal distributions (with an exponential mixing density) as

$$\frac{a}{2} e^{-a|y|} = \int_0^\infty \frac{1}{\sqrt{2\pi s}} e^{-y^2/2s} \frac{a^2}{2} e^{-a^2 s/2} ds, \qquad a > 0.$$

Bae and Mallick (2004) assumed independent Laplace priors on β of the form in (5.6) to induce sparseness while Park and Casella (2008) assumed a conditional Laplace priors on β to ensure unimodal posterior distributions. We adopt both their approaches and define the following Binomial logistic regression model with Bayesian

Lasso regularization to predict the likelihood surface

$$
\begin{aligned}
\mathbf{z}|\boldsymbol{\theta} &\sim Bernoulli(\theta) \\
\text{logit}(\theta) &= \mathbf{X}\boldsymbol{\beta} \\
\boldsymbol{\beta}|\mathbf{D}_{\boldsymbol{\tau}} &\sim \mathbf{N}_p^+(\mathbf{0_p}, \mathbf{D}_\tau) \\
\mathbf{D}_\tau &= \text{diag}(\tau_1^2, \ldots, \tau_p^2) \\
\tau_l^2 &\sim Exponential(\lambda^2) \qquad l = 1, \ldots, p,
\end{aligned}
\tag{2.17}
$$

where $\mathbf{N}^+$ refers to the Truncated Gaussian Distribution, truncated below by 0. This assumption is based on physical properties of the elements of $\mathbf{X}$. λ is the Lasso tuning parameter. We discuss methods to determine this parameter in the Appendix. Given the prior distribution assumptions, we run into the issue of conjugacy due to the analytically inconvenient form of the model's likelihood function.

Chapter 3

S_n and L_g Attenuation in China

Abstract

With a large dataset recorded by densely deployed seismic stations in China from 2003 to 2011, using the reverse two-station method (RTM), we estimated lithospheric attenuation using the two regional seismic phases L_g and S_n. Using more than 200,000 RTM rays covering all over China, we have tomographically mapped both L_g and S_n Q to show the spatial variations of crustal and uppermost mantle attenuation with the LSQR method (Paige and Saunders, 1982). Both L_g and S_n Q results show strong spatial variation. High L_g Q values are observed across southeast (SE) and northeast (NE) China and the Tarim basin, while low L_g Q values are observed across most of the Tibetan plateau. Unlike L_g wave, high S_n Q values are found in the Sichuan basin, the Ordos basin, and the Tien Shan area, while low S_n Q values are found in eastern China and the northern and eastern Tibetan plateau. We also estimated L_g and S_n Q as a function of frequency. In regions with high Q, we see a positive linear relationship between Q and frequency for both L_g and S_n. We also examined the difference between the S_n and L_g Q models as a function of frequency to explore S_n

attenuation in the crust. The tomography of differential Q also shows strong spatial variation that correlates with the tectonic boundaries. Negative difference values are found in the Cathaysia Foldbelt and some parts of NE and NW China, while positive difference values are found in the Sichuan and Ordos basins and the Tibetan plateau. Our attenuation tomography results are consistent with most previous studies and the anomalies are strongly correlated with the major tectonic blocks in China.

Key words: L_g and S_n attenuation, China, Lithosphere, Tomography.

3.1 Introduction

Seismologists study seismic waves to understand earth's internal structure. Seismic attenuation is a key parameter used to quantitatively describe seismic wave propagation. Q is used to describe the amplitude reduction with travel distance. Q stands for "quality factor", which is proportional to the reciprocal of attenuation. There are two attenuation mechanisms, scattering and intrinsic attenuation. Intrinsic attenuation is defined as energy loss per cycle of oscillation because of internal friction along grain boundaries of a given mineral aggregate

$$\frac{1}{Q} = \frac{\Delta Q}{2\pi E}. \tag{3.1}$$

While scattering attenuation is caused by variations in seismic velocity structure along the seismic wave propagation path. The intrinsic attenuation is a function of the anelasticity in the media which could be used to study the rheology throughout the whole mantle and crust (e.g., Karato, 1993; Knopoff, 1964). The scattering attenuation is determined by the scale and size of the velocity inhomogeneities in the Earth. Significant velocity variations can be caused by slabs, faults, and other geological structures that lead to lateral variations in seismic wave velocity. Because of the complexity of the lithospheric seismic structure, it is difficult to separate the two attenuation mechanisms in real data. We can typically only calculate the effective attenuation, which is the combined effect of scattering and intrinsic attenuation on the amplitude reduction for a given seismic phase.

Regional earthquakes are events that occur at epicentral distances between 3° and 20°. Regional seismograms are dominated by seismic energy propagating within the lithosphere (e.g., Bao et al., 2012). The two high frequency shear phases on regional seismograms are S_n and L_g with different propagating paths. L_g is usually the most prominent phase on regional seismograms for continental paths (Sereno, 1990; Rapine et al., 1997; Baumgardt, 2001) and propagates within the crust as a guided shear wave with a velocity and frequency band of 2.8 - 4.0 km/s and 0.5 - 5.0 Hz respectively. L_g has been interpreted as the superposition of higher-mode Rayleigh wave that primarily propagates in the crust (Knopoff et al., 1973) and, alternatively, as the superposition of super-critically reflected shear wave in the continental crust (Campillo et al., 1985; Kennett, 1986). When observed on vertical seismograms, L_g can be primarily associated with Rayleigh wave overtone modes, although scattering may mix Love and Rayleigh energy. L_g is likely to be largely insensitive to earthquake radiation patterns. In addition, empirically it is particularly useful for seismic magnitude estimation for earthquakes and yield estimations for explosions. Because L_g is dominated by shear wave energy, it tends to be more strongly excited by earthquakes than explosions (Fan and Lay, 2002). Because it probably fully samples the crust, L_g phase attenuation has widely been used to investigate the rheology and structure of the crust. L_g is strongly attenuated or blocked in some continental areas with significant changes in crustal thickness or changes in sedimentary rock thickness, such as one would find across mountain belts or sedimentary basins (Zhao et al., 2003; Xie, 2002). Scattering or strong attenuation may also lead to blockage; but the three-dimensional L_g propagation features are very difficult to measure(Bao, 2011).

The S_n phase is also a high-frequency guided wave. Unlike the L_g phase, S_n travels along the lithospheric mantle with frequency that is usually between 0.5 and 4 Hz, with a velocity between 4.4 and 4.7 km/s (Sandvol et al., 2001), so that S_n velocity and attenuation can yield insight into the properties (e.g., velocity and attenuation) of the uppermost mantle. S_n arrives as a high-frequency wave-train lasting tens of seconds and up to several minutes (Sandvol et al., 2001), with velocity larger in stable continental and oceanic regions than in tectonically active regions. Although S_n propagates efficiently in stable continental and shield regions (Ni and Barazangi, 1983; Gök et al., 2000; Sandvol et al., 2001; Gök et al., 2003) and it has been recorded with epicentral distances up to 35° (Molnar and Oliver, 1969; Huestis et al., 1973); it is usually blocked or highly attenuated for paths crossing tectonically active regions with high heat flow (Molnar and Oliver, 1969; Kadinsky-Cade et al., 1981; Ni and Barazangi, 1983; McNamara and Owens, 1995; Gök et al., 2000; Calvert et al., 2000; Sandvol et al., 2001).

As a major part of East Asia, China is composed of highly diverse tectonic terrains with both ancient tectonic blocks and active orogenies. It has a complex geologic history involving the collision of the Indian and Eurasian plates, the subduction of the Pacific plate beneath the Eurasian plate, and complex topography change from Tibet to the Marianas Trench. All these tectonic processes lead to strong lateral variation in the seismic lithospheric structure beneath China.

In SW China, the Tibetan plateau is one of the most active continent-continent collisions on Earth. We find a number of features resulting from the collision between the Indian and Eurasian plate. The convergence of the Indian and Eurasian plates started approximately 50 Ma, and the Indian plate has been continuously subducting beneath Tibet (e.g., Yin and Harrison, 2000; Chen et al., 2017) since the initial collision. This collision has resulted in the formation of the Himalayas and the Tien Shan. It has also resulted in crustal shortening and uplifting of the Tibetan plateau, accom-

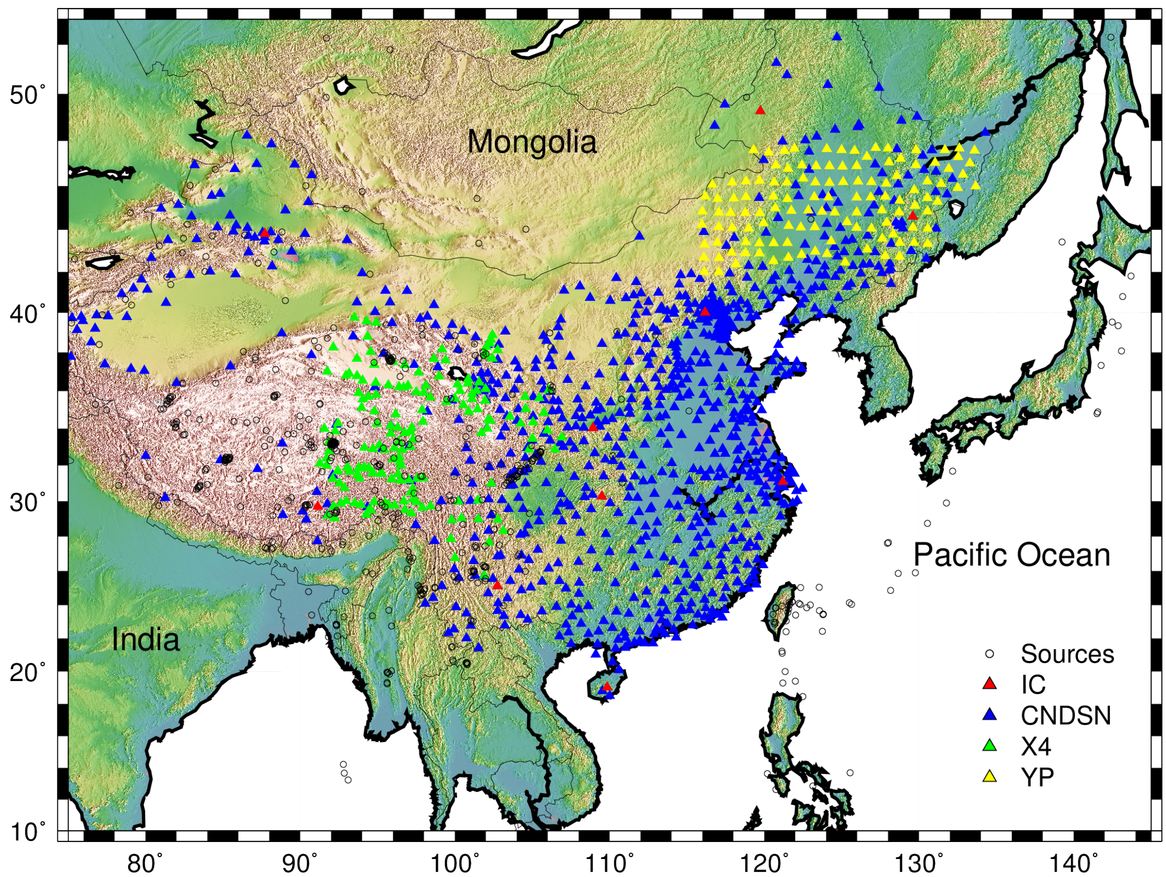


Figure 3.1: The study area and seismic stations used in this study. Blue triangles are stations from network CNDSN. Red triangles are from IC. Green triangles are from X4 and yellow triangles are stations from YP. Black circles are sources.

panied by eastern-ward extrusion of the plateau and the formation of large strike slip faults like the Altyn Tagh and Kunlun faults zones (Yin and Harrison, 2000; Wang and Shen, 2020). Mongolia is one of the more tectonically active intra-continental regions in the world (Choi et al., 2018), probably due to regional deformation related to the India-Eurasia collision. Compared to China, GPS data in Mongolia show relatively small crustal deformation (∼4 mm/yr, Wang and Shen, 2020). At the eastern margin of the Eurasia plate, the Pacific and Philippine Sea plates are subducting, causing destruction of the North China block at a significantly high rate (Wang and Shen, 2020). Wei et al. (2019) proposed that the deep subduction of the Pacific Plate has affected the formation of both Wudalianchi and Halaha volcanoes, which are located at the north and west edge of the Songliao Basin, respectively. The Changbaisha volcano is located in the Chingbaishan mountains, close to the boundary between NE China and North Korea. Magnetotelluric soundings show that low resistivity anomalies exist beneath the Changbaishan volcano in the crust. Seismic explosion experiments revealed low-velocity anomalies in the crust and upper mantle down to a depth of 40 km. These results suggest the existence of magma chambers under the Changbaishan volcano (Lei et al., 2013). The Quaternary Datong volcano is located in the northernmost portion of the Shanxi rift. There are 30 small volcanoes in the Datong volcanic field, and they are distributed about 3 km away from Datong county. Recent regional tomographic models showed low-V anomalies under the Datong volcano in the upper mantle (Lei et al., 2013). Our study focus on the lithospheric attenuation structure beneath the entire Chinese mainland.

3.2 Data

The data used in this study include 484 regional earthquakes recorded by 1188 stations (Figure 3.1). There are more than 200,000 RTM paths as shown in (Figure

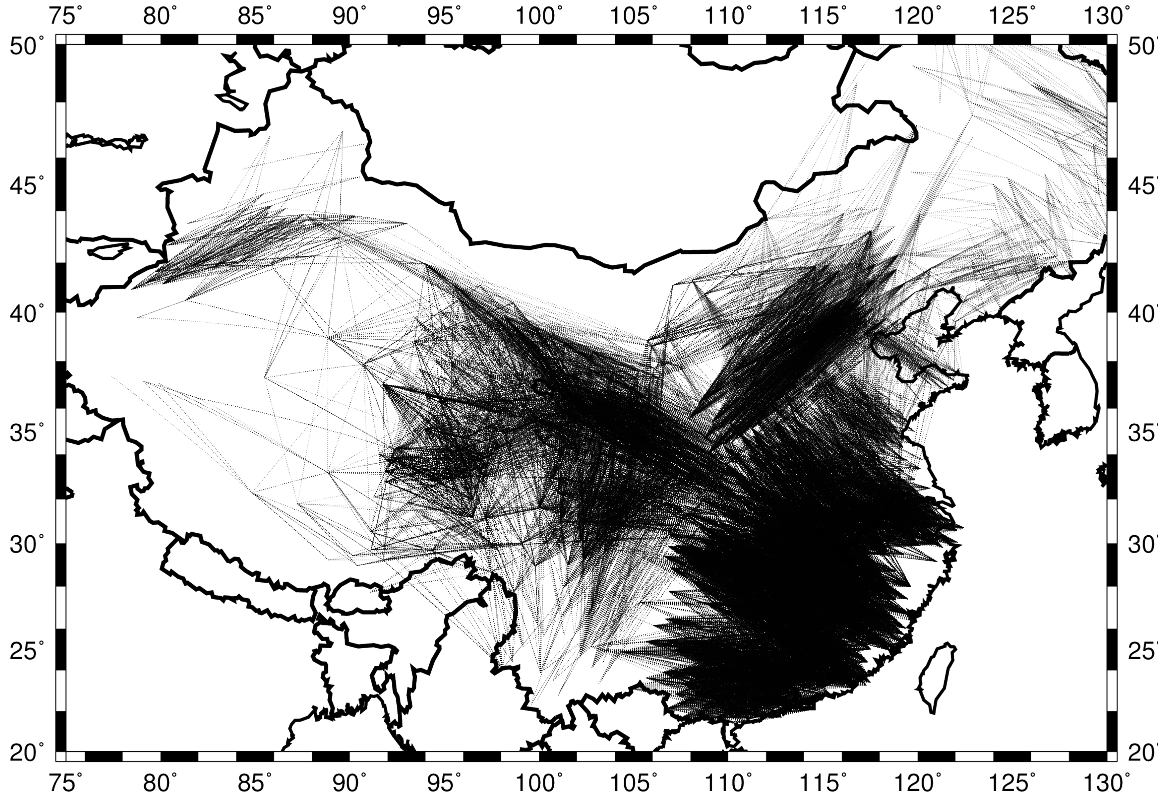


Figure 3.2: There are 212,828 RTM rays used in this study.

3.2). The stations are deployed throughout mainland China which includes several seismic networks. The network X4 includes 167 broadband stations covering the eastern Tibet from 5 major seismic networks: Indepth IV, Namche Barwa, MIT-China, NETS and ASCENT. The events we used were recorded between 2003 and 2009. The network CNDSN includes 891 stations covering almost the whole China mainland. The corresponding events are between 2009 and 2011. However, it was not possible to find overlapping events between these two data sets. To link these two data sets together, which is necessary in the RTM, we use 10 stations from network IC and 120 stations from the seismic network YP. YP covers the northeastern China, while IC stations cover nearly the entire Chinese mainland. All the events used in this study have magnitudes greater than 4.5 and occur within crust. In this study, we only use the vertical component data because the signal to noise ratios tend to be higher.

3.3 Methods

Regional seismic attenuation (Q^{-1}) provides an important constraint for a variety of geophysical and geologic studies, such as, rheology and temperature of the crust and uppermost mantle as well as used to create transportable earthquake-explosion discrimination algorithms (Bao, 2011). There are several methods used to calculate the regional seismic attenuation and the discrepancies between methods are due to differing parameterizations, employed data sets and the differences in assumptions made by different methodologies (Ford et al., 2008).

The method used in this study is the reverse two-station method (RTM). RTM is based upon the spectral amplitude ratio of the regional phases and the Q tends to be the most stable compared with all other methods due to the elimination of the source term, the station site response and instrument response.

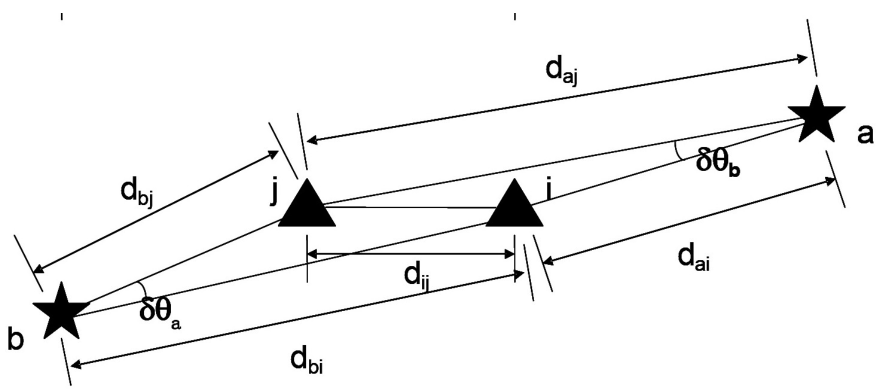


Figure 3.3: RTM scheme. a and b denote two sources. i and j are two stations. d denotes distance and $\delta\theta$ is azimuth difference.

The amplitude of regional phase in the frequency domain is given by

$$A = S \cdot R \cdot G \cdot I_S \cdot S_S \cdot I_Q \cdot T_{FD} \cdot T_{SC} \cdot T_{AN}, \tag{3.2}$$

where S denotes the source excitation function and R, the radiation pattern from the focal mechanism; G, the geometrical spreading function (e.g., $G = G_0 d^{-m}$ for L_g and

S_n, d and m denotes distance and a constant); I_S, the instrument response; S_S, the station site response; I_Q, intrinsic attenuation; T_{FD}, the coefficient of focusing and de-focusing; T_{SC}, scattering coefficient; T_{AN}, other effects.

T_{FD} is typically assumed to be 1 if we give a limit on standard errors of the measurements (Xie, 2002) and I_Q, T_{SC}, T_N are interpreted as apparent attenuation. Thus, the spectral amplitude is simplified as

$$A(f,d) = S(f) \cdot R(f,\varphi) \cdot I(f) \cdot S_S(f) \cdot G(d) \cdot \exp(-\frac{\pi f d}{vQ}), \tag{3.3}$$

where f denotes frequency and d, the epicentral distance; φ, azimuth; v, the seismic velocity.

We use i, j to denote station indices and a, b for event indices (Figure 3.3). The four spectral amplitude equations are

$$\begin{cases} A_{ai}(f,d_{ai}) = S_a(f)R_a(f,\varphi)I_i(f)S_{S_i}(f)G_{ai}(d_{ai})\exp\left(-\dfrac{\pi f d_{ai}}{v_i Q_i}\right) \\ A_{aj}(f,d_{aj}) = S_a(f)R_a(f,\varphi)I_j(f)S_{S_j}(f)G_{aj}(d_{aj})\exp\left(-\dfrac{\pi f d_{aj}}{v_j Q_j}\right) \\ A_{bi}(f,d_{bi}) = S_b(f)R_b(f,\varphi)I_i(f)S_{S_i}(f)G_{bi}(d_{bi})\exp\left(-\dfrac{\pi f d_{bi}}{v_i Q_i}\right) \\ A_{bj}(f,d_{bj}) = S_b(f)R_b(f,\varphi)I_j(f)S_{S_j}(f)G_{bj}(d_{bj})\exp\left(-\dfrac{\pi f d_{bj}}{v_j Q_j}\right) \end{cases} \tag{3.4}$$

Assume that the velocity structure is 1-D and apparent Q is identical between stations i and j. From Equation 4.3, we get

$$\frac{A_{ai}A_{bj}}{A_{aj}A_{bi}} = \left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^{-m} \exp\left[\frac{\pi f}{vQ}(d_{aj} - d_{ai} - d_{bj} + d_{bi})\right]. \tag{3.5}$$

Thus, attenuation could be deterministically determined by,

$$\frac{1}{Q} = \frac{v}{\pi f(d_{aj} - d_{ai} - d_{bj} + d_{bi})} \times \ln\left[\frac{A_{ai}A_{bj}}{A_{aj}A_{bi}}\left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^{m}\right]. \tag{3.6}$$

To get Q tomography, we discretize the study area into M meshes. Thus the apparent attenuation of the n-th path could be expressed as,

$$\exp\left(-f\pi\frac{d_n}{Q_n v_n}\right) = \exp\left[-f\pi\sum_{m=1}^{M}\left(\frac{d_{nm}}{Q_{nm}v_{nm}}\right)\right]. \tag{3.7}$$

In RTM, we assume 1D velocity structure, which is $v_{nm} \equiv v_n$. Thus,

$$\frac{d_n}{Q_n} = \sum_{m=1}^{M}\left(\frac{d_{nm}}{Q_{nm}}\right). \tag{3.8}$$

To solve these linear equations, we use the Sparse Equations and Least Squares (LSQR, Paige and Saunders, 1982) to get 2D Q tomography. The damping parameter used is 0.25 to make best visualization.

In this study, we only use the vertical component of the seismograms. We manually picked the arrival times of P_n and S_n phases to get the spectral amplitude; however, we used a fixed velocity window from 2.9 km/s to 3.5 km/s for the L_g phase. We used pre-P_n signal as background noise and only used data with signal-to-noise ratio (SNR) larger than 2.0 to estimate Q. The geometrical spreading factor (m in Equation 3.6) is 1 for S_n and 0.5 for L_g. The azimuth tolerance $\delta\theta_a/\delta\theta_b$ is set to be $\pm 15°$ (Xie, 2002).

3.4 Results

To determine the resolution of our data, we created several checkerboard models with $\pm 15\%$ random noise. Then, using the same RTM ray paths and an LSQR tomographic model, we developed a spatially and frequency dependent Q model. The anomalies in the checkerboard test that had a size $\geq 3°$ are well recovered (Figure 3.4). The model resolves most of our study area with some smearing at the edges. We find relatively poor resolution in the Tarim Basin because of lack of data coverage. Because the

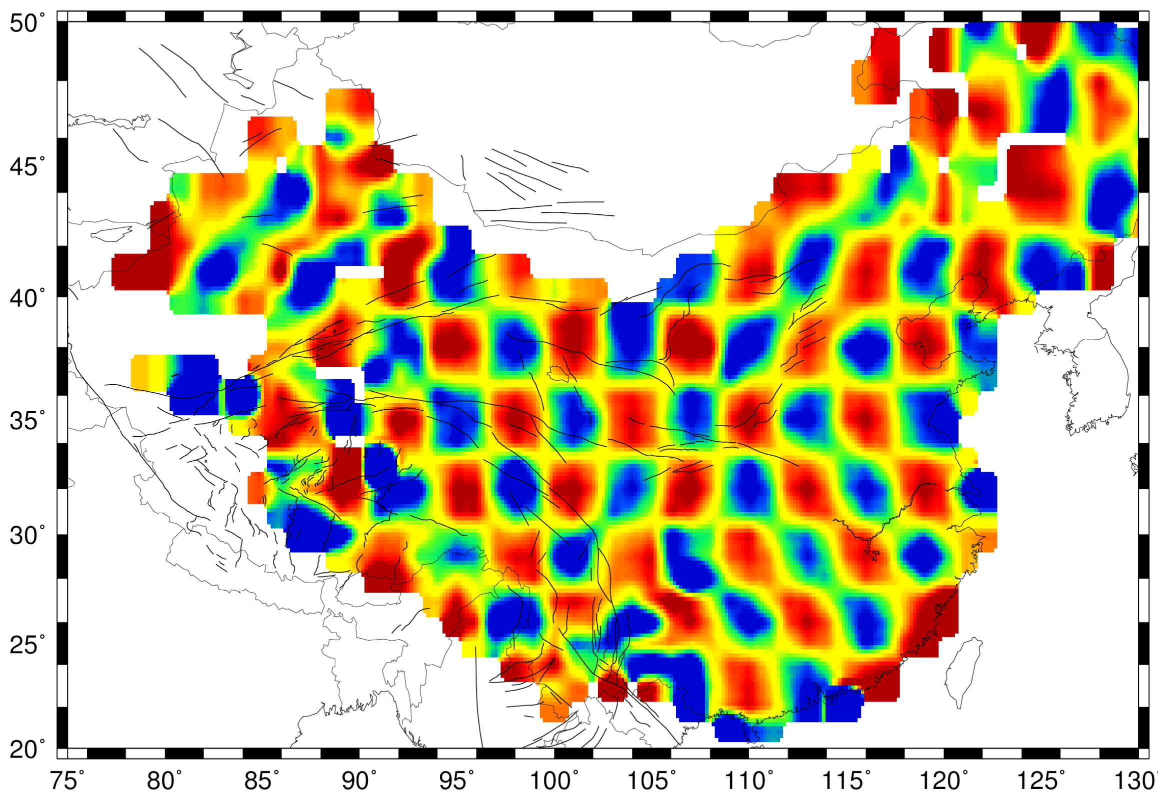
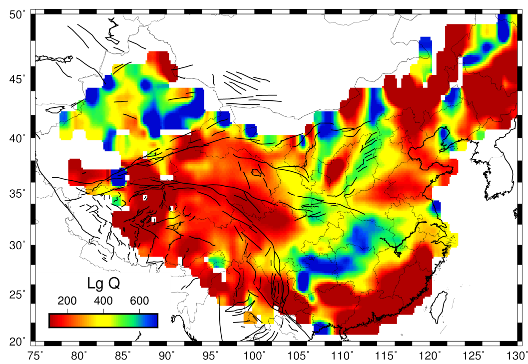


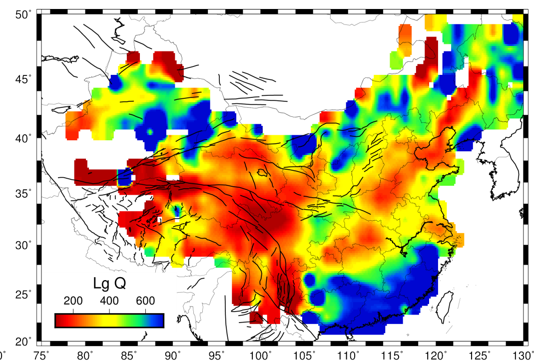
Figure 3.4: Resolution test result with 3° anomaly of RTM ray data. Blue denotes high Q area and red denotes low Q area. The noise level of the checkerboard test is ±15%.

raypaths that cross the Junggar Basin and the South China Block mostly have east-west oriented azimuths, the resolved model has some smearing. Even though there are fewer rays across the Songliao Block compared with other regions, the model is resolved reasonably well because the rays there have a large azimuthal range, which results in many crossing ray paths.

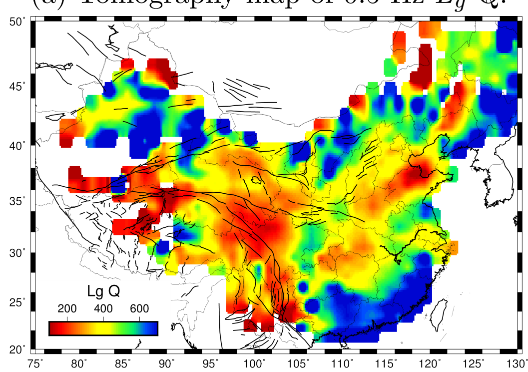
3.4.1 L_g Attenuation Tomography



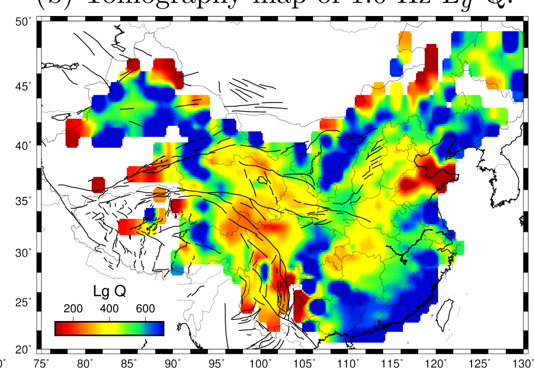
(a) Tomography map of 0.5 Hz L_g Q.



(b) Tomography map of 1.0 Hz L_g Q.



(c) Tomography map of 1.5 Hz L_g Q.



(d) Tomography map of 2.0 Hz L_g Q.

Figure 3.5: Tomography maps of L_g Q at four different frequencies.

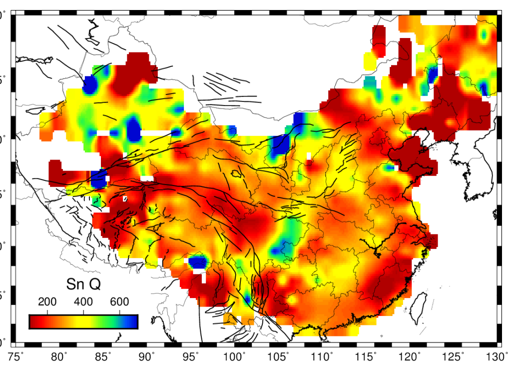
We observe strong lateral variations in L_g Q across the China mainland, from Q $<$ 100 within the Tibetan Plateau to Q $>$ 700 across the Cathaysia Foldbelt. The L_g Q patterns are relatively similar at different frequencies; however, unlike the high frequency L_g results, we observe very high L_g attenuation (Q $<$ 100) across the South China Block with 0.5 Hz signal. This strong difference in L_g Q indicates that scattering may be the primary attenuation mechanism.

Using a low frequency (0.5 Hz) L_g signal (Figure 3.5a), we observe high attenuation across some parts of the Songliao Basin, at the northeastern edge of the North China Craton (NCC), in the Cathaysia Foldbelt, in SW China, and in the Tibetan Plateau. We also observe high crustal attenuation in the middle of the Ordos plateau. In areas like the Tien Shan mountain range, the edges of the Ordos plateau, the Shanxi rift, and part of the Qinling and Yangtze Cratons, we observe large areas of low L_g attenuation. There is also low L_g attenuation at the western edge of the Changbaishan mountain belt.

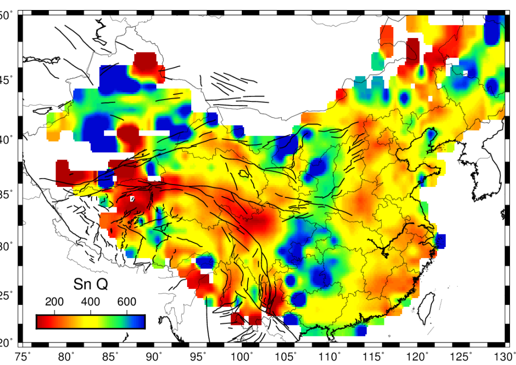
Unlike the 0.5 Hz L_g result, with frequency $\geq 1Hz$ we observe low attenuation in the Cathaysia Foldbelt and high attenuation in the Yangtze Craton ($\geq 1Hz$, Figure 3.5b, 3.5c, and 3.5d). We observe low crustal attenuation across most of the Ordos plateau, the Changbaishan basin, and the northeastern North China Craton with high frequency L_g. We also observe low crustal attenuation (high L_g Q), at the eastern edge of the Tibetan Plateau. It should be noted that we use the same color bar for all the tomography to compare the absolute values of Q at different frequencies. In general, we observe an increase in L_g Q with increasing frequency. There are some low Q regions that do not change very much with changes in frequency.

3.4.2 S_n Attenuation Tomography

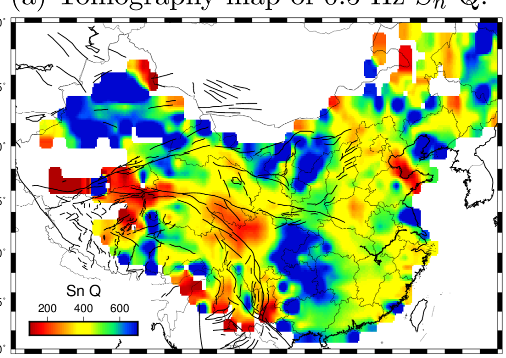
Unlike our L_g Q results, which range from low (0.5 Hz, Figure 3.5a) to high (2.0 Hz, Figure 3.5d) depending on frequency, our S_n attenuation is consistent across frequencies. We observe low S_n Q throughout the northeastern Tibetan Plateau and the Shanxi rift. High upper mantle attenuation is also found throughout almost all of the North China Craton as well as the South China Block. We also observe high Q in Ordos, Sichuan and Junggar Basins, which is consistent with the existence of a tectonically stable (no internal deformation) and thick continental lithosphere. S_n Q



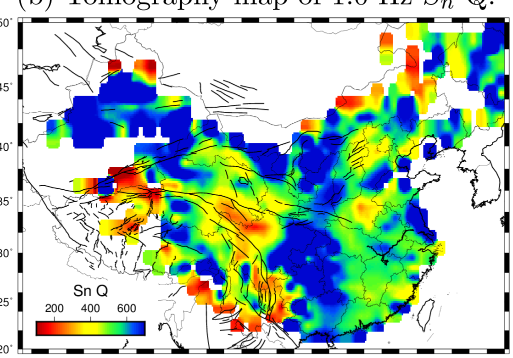
(a) Tomography map of 0.5 Hz S_n Q.



(b) Tomography map of 1.0 Hz S_n Q.



(c) Tomography map of 1.5 Hz S_n Q.

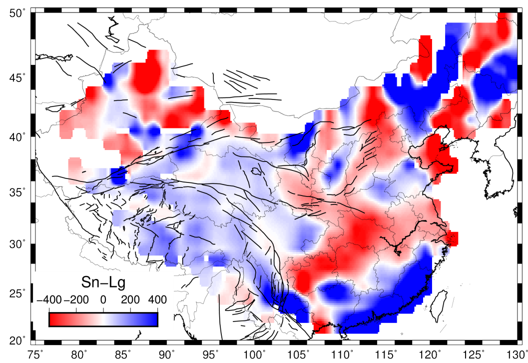


(d) Tomography map of 2.0 Hz S_n Q.

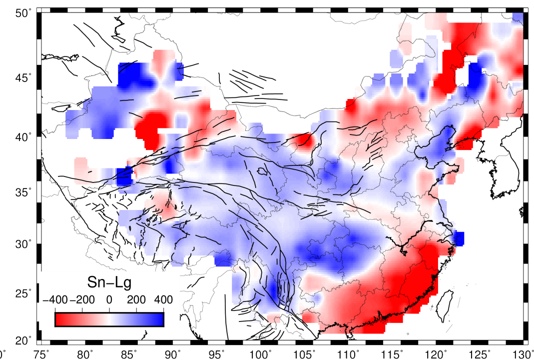
Figure 3.6: Tomography map of S_n Q at four different frequencies.

also consistently increases with increasing frequency.

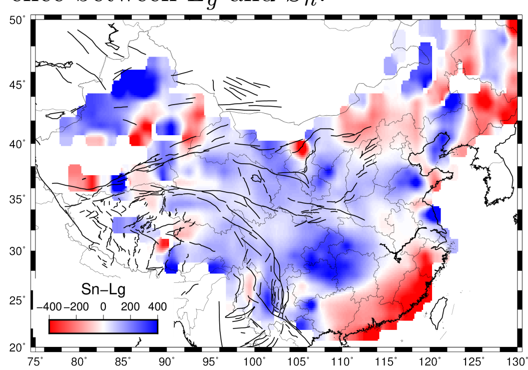
3.4.3 Attenuation Difference Tomography



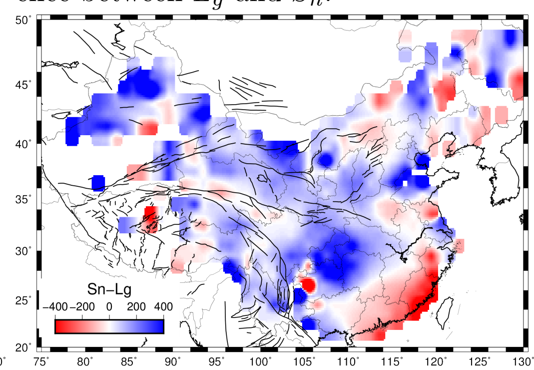
(a) Tomography map of 0.5 Hz Q difference between L_g and S_n.



(b) Tomography map of 1.0 Hz Q difference between L_g and S_n.



(c) Tomography map of 1.5 Hz Q difference between L_g and S_n.



(d) Tomography map of 2.0 Hz Q difference between L_g and S_n.

Figure 3.7: Tomography map of Q difference between L_g and $S_n$at four different frequencies.

Due to their different paths, S_n and L_g Qs reveal the attenuation of different part of the earth. Both of the Q_{S_n} and Q_{L_g} models are consistent at the higher frequency bands, although across eastern China, the 0.5 Hz results are opposite. In other words, in eastern China we observe a positive L_g Q anomaly and negative S_n Q anomaly. Examining 0.5 Hz amplitude data, we observe high and positive $Q_{S_n} - Q_{L_g}$ within the Cathayasia Foldbelt and Songliao Basin. We also observe moderate and positive $Q_{S_n} - Q_{L_g}$ in the Tibetan Plateau, the Tianshan suture, the Ordos and the Trans-

North China Orogen. Low $Q_{S_n} - Q_{L_g}$ regions are found in the Yangtze Craton, Shanxi rift, Qinling-Dabie mountains, Junggar Basin and part of the Songliao Basin (Figure 3.7a). With high frequency data, we observe high negative Q anomalies $Q_{S_n} - Q_{L_g}$ in the Cathayasia Foldbelt, the Songliao Basin, and the eastern edge of the Tarim Basin. We also see negative $Q_{S_n} - Q_{L_g}$ within the Qinling-Dabie zone. The high frequency positive $Q_{S_n} - Q_{L_g}$ are found in the Junggar Basin, Tibetan Plateau, North China Craton and Sichuan Basin (Figure 3.7b, 3.7c and 3.7d).

3.5 Discussion and Conclusion

Using a dense seismic network that covers all of mainland China, we estimate high resolution lithosphere attenuation. We observe dramatic spatial variation in both L_g and S_n Q.

3.5.1 Results error analysis

There are often more than one rays traveling in the same RTM paths, named by path-repeating. The different rays traveling in the same RTM paths are called repeated paths. The Q from repeated paths is constant without any error or uncertainties. The standard deviation of all the Q values of repeating RTM paths depends on the error/uncertainty level of the data.

Figures 3.8 and 3.9 show the standard deviation of Q values over RTM paths with a repeating number larger than 5 by colored rays and tomography, respectively. The standard variation of repeated-path Q is less than 200 almost all over China. The standard variation of repeated-path Q larger than 200 are densely located at Cathayasia Foldbeld zone. Figure 3.10 is the histogram plot of Q standard deviation of all the repeated paths. The majority values are less than 200, meaning that most of our Q results are stable.

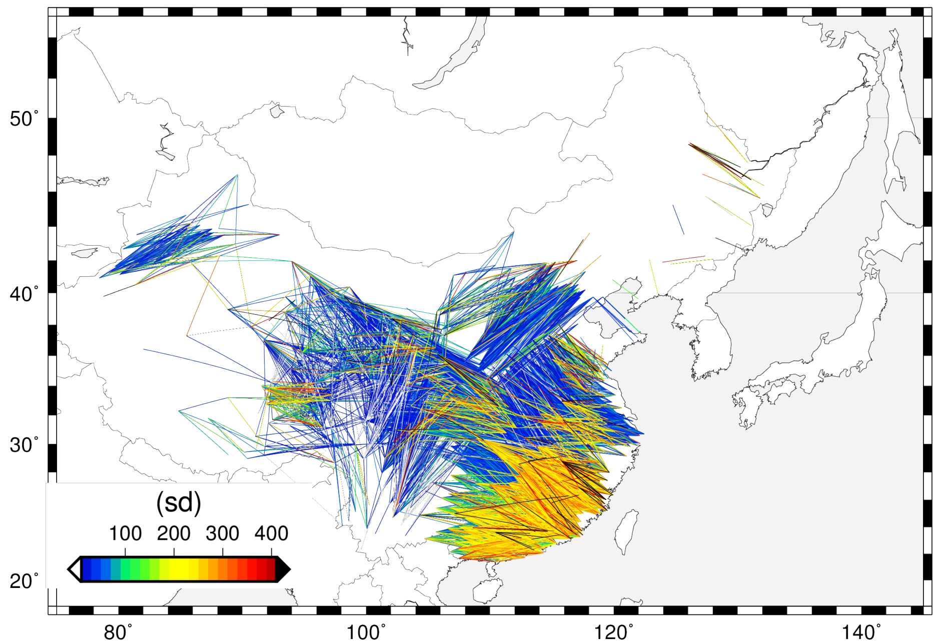


Figure 3.8: A map showing standard variation calculated from Q values over RTM paths with a repeating number larger than 5.

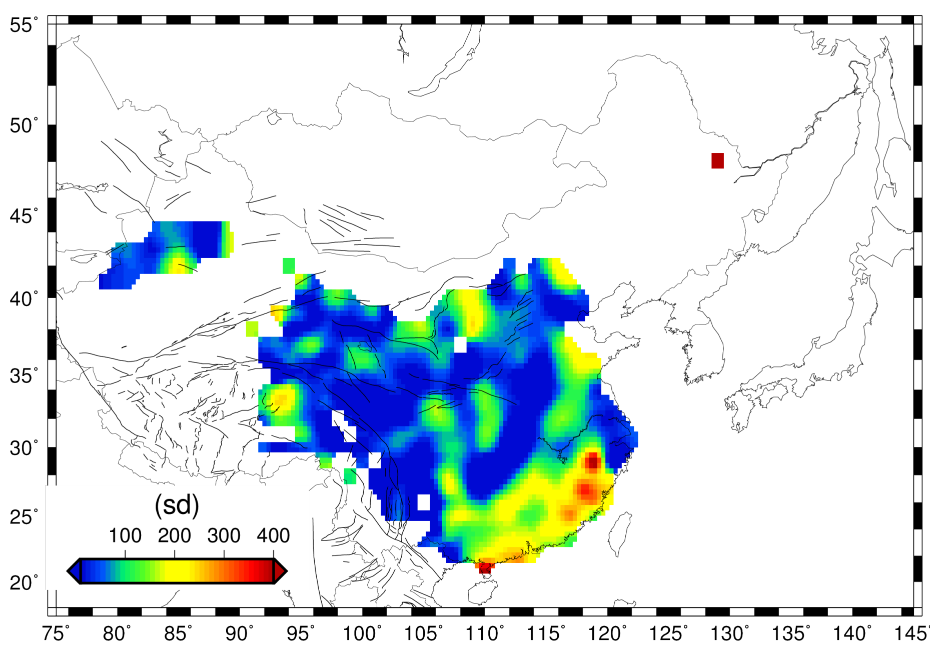


Figure 3.9: A tomography map showing standard variation calculated from Q values over RTM paths with a repeating number larger than 5.

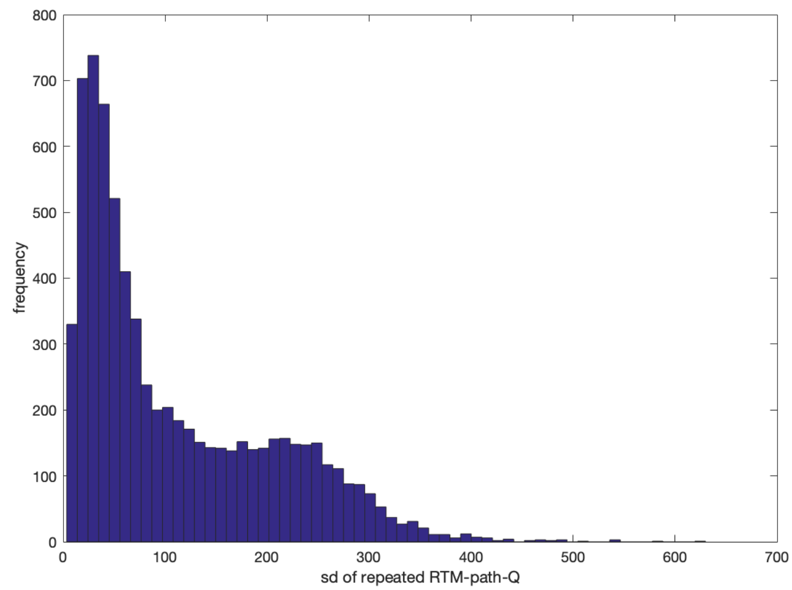


Figure 3.10: A histogram plot showing standard variation calculated from Q values over RTM paths with a repeating number larger than 5.

3.5.2 Crustal attenuation tomography

The L_g Q tomographic model is different from the S_n Q model. This may be an indication that there is no widespread energy leakage from S_n to L_g or vice versa. The high frequency L_g Q pattern is similar to others (Phillips et al., 2005) with higher resolution. We observe high L_g Q across SE China, the Ordos plateau, the Sichuan basin, the Tarim basin, and NE China. We find low L_g Q throughout most of the Tibetan Plateau as well as the North China Craton. In the South China Block, we observe high L_g Q in the Yangtze Craton and low L_g Q in the Cathayasia fold and thrust belt. Also we find low L_g Q in the Yangtze Craton and high L_g Q in the Cathayasia fold belt for higher frequency Lg signals. The reason for this strong frequency dependence may be that the crust is too thin ($<$ 30 km, Li et al., 2018) to support or generate long wavelength L_g. In other words, the 0.5 Hz L_g has sampled uppermost mantle.

In the Qinling mountain area, the Moho is deeper than in the adjacent regions ($\sim$

40 km, Li et al., 2018). In the Tibetan Plateau, the effective crustal attenuation shows significant spatial variation, but the anomalies are not strongly frequency dependent (Fan and Lay, 2002, 2003; Bao et al., 2011, 2012). These faults do not seem to affect Sn attenuation either, which may indicate that they do not extend into the uppermost mantle or do not affect S_n attenuation as much as they do L_g. Compared to adjacent areas, the low L_g Q in the Tibetan Plateau is the result of high temperature anomalies that are linked to tectonic activity (Bao et al., 2012).

In the North China Craton, we observe widespread high crustal attenuation. This result is consistent with widespread low shear wave velocities in the crust (Li et al., 2018).

3.5.3 Uppermost mantle attenuation tomography

S_n Q is often used to interpret the uppermost mantle rheology and structure. Low S_n Q is consistent with hot, partially molten, or even absent uppermost mantle. High S_n Q indicates cold and thick lithosphere, which is typically found in cratonic regions. In addition, Chen and Niu (2016) found high shear wave velocities in the uppermost mantle beneath the Sichuan basin, which is consistent with our high S_n Q results. Both of these observations are consistent with the stable thick mantle lithosphere underlying the Sichuan basin. Rapine and Ni (2003) found regions with both inefficient and efficient S_n propagation at the western edge of NE China. Our S_n Q tomography results are consistent with prior studies in this regionadd references. We observe high Q in efficient areas and low Q in inefficient areas. Furthermore, we observe low S_n Q across the North China Plain, where Rapine and Ni (2003) found extensive S_n blockage.

3.5.4 Attenuation difference between S_n and L_g

$Q_{S_n} - Q_{L_g}$ is high for the Ordos and most of the North China Craton. All of these high $Q_{S_n} - Q_{L_g}$ are associated with tectonically stable lithospheric mantle and crust with significant sedimentary layers. $Q_{S_n} - Q_{L_g}$ is positive and large within the Tibetan Plateau. The Tibetan Plateau is composed of several blocks with many faults, so that crustal scattering attenuation is high, leading to S_n Q higher than L_g Q. Within the Sichuan basin, L_g Q is low, which explains why $Q_{S_n} - Q_{L_g}$ is positive and high. The lower Lg Q is likely linked with the young sedimentary rocks within the upper portion of the crust.

Chapter 4

S_n and L_g Site Response in China

Abstract

Using a large data set recorded across the China from 2003 to 2011, we estimated the frequency dependent site amplification of both S_n and L_g phases with a revised reverse two-station method (RTM). Our data set, with over 20,000 RTM rays, covers all of China with high density. We used the LSQR algorithm (Paige and Saunders, 1982) to estimate the absolute site response of all the stations with with respect to one station. It is worth noting that the absolute site amplification inverse problem is inherently under-determined. Our site response models for both S_n and L_g show strong positive amplification for all of the basins across China. For example we observe strong amplification in the Songliao Basin and Ordos Basin; however, we generally observe negative amplification (deamplification) across most of the mountain ranges, such as Tibet and Changbaishan. It is important to note that the site response results for S_n and L_g are not identical to each other. While the L_g site response primarily correlates with crustal structure, the S_n site response likely includes both mantle lithosphere and crustal effects. The absolute values of the site response become

larger with higher frequency, which means that the correlation between site response and topography is stronger at higher frequencies. We also compared the site response difference between S_n and L_g ($SR_{S_n} - SR_{L_g}$) with S_n Q tomography. We found that almost all positive site response difference stations are located within low S_n Q regions, such as the South China Block and the Tibetan Plateau. We think this is the effect of the crustal legs of S_n, which are not considered separately in the RTM. There are two crustal legs, the event side and the station side, and this chapter focuses primarily on the station side.

Key words: Site amplification, L_g and S_n, China, Lithosphere.

4.1 Introduction

Regional seismograms are dominated by seismic energy propagating within the lithosphere (e.g. Bao et al., 2012). The four high frequency regional phases are P_n, P_g, S_n and L_g, all with different propagating paths. L_g is usually the most prominent phase on regional seismograms (Baumgardt, 2001), propagating within the crust primarily as a guided shear wave, but appearing on all three components of a three component seismogram. Because L_g likely samples most of the crust as a guided wave, it has been widely used to investigate the rheology and structure of the crust (e.g. Knopoff et al., 1973; Kennett and Mykkeltveit, 1984; Kennett, 1986; Gibson Jr and Campillo, 1994; Benz et al., 1997; Ranasinghe et al., 2015). L_g is usually the most prominent and stable phase on regional seismograms for continental paths (Sereno, 1990; Rapine et al., 1997; Baumgardt, 2001) and propagates within the crust as a guided shear wave with velocity and frequency bands of 2.9 - 3.7 km/s and 0.5 - 5.0 Hz, respectively. L_g has been interpreted as the superposition of higher-mode Rayleigh waves that primarily propagate in the crust (Knopoff et al., 1973) or, alternatively, as the superposition of super-critical reflected shear waves in the continental crust

(Campillo et al., 1985; Kennett, 1986). When observed on vertical seismograms, L_g can be primarily associated with Rayleigh wave overtone modes, although scattering may mix Love and Rayleigh energy. It is worth noting that we typically see little to no dispersion of L_g waves across China, suggesting that L_g is more complex than just higher order model Rayleigh waves. L_g is largely insensitive to earthquake radiation patterns, thus it has particular value for seismic magnitude estimation. Because L_g is dominated by shear wave energy, it tends to be more strongly excited by earthquakes than by explosions (Fan and Lay, 2002).

L_g is strongly attenuated or blocked in some continental areas with significant changes in crustal thickness, such as mountain belts or basins (Zhao et al., 2003; Xie, 2002). These observations would suggest that scattering or strong attenuation may cause the blockage. But the factors affecting three-dimensional L_g propagation are not fully understood(Bao, 2011). In addition to scattering attenuation, there is some evidence that regions of the crust with partial melt may also lead to L_g blockage. S_n is another high frequency (0.5 - 5 Hz) regional phase traveling primarily within the lithospheric mantle as a shear wave. Thus, S_n is commonly used to study the properties (i.e., velocity and attenuation) of the uppermost mantle. S_n arrives as a high-frequency wave-train lasting from tens of seconds up to one to two minutes (Sandvol et al., 2001). S_n velocities are larger in stable continental and oceanic regions than in tectonically active regions. Although S_n propagates efficiently in stable continental and shield regions (Ni and Barazangi, 1983; Gök et al., 2000; Sandvol et al., 2001; Gök et al., 2003), and has been recorded with epicentral distances of up to 35^o (Molnar and Oliver, 1969), it is usually blocked for paths that cross tectonically active regions with high heat flow (Molnar and Oliver, 1969; Kadinsky-Cade et al., 1981; Ni and Barazangi, 1983; McNamara and Owens, 1995; Gök et al., 2000; Calvert et al., 2000; Sandvol et al., 2001). Site response is often used in seismic hazard analysis to understand smaller scale variation in ground motion and usually is interpreted as focusing and defocusing very near a recording site (Gao et al., 1996) on a scale of of only hundreds of meters (Imtiaz et al., 2015; Gao et al., 1996). Additionally, contrasts in elastic impedance can have a very important effect on site response (Murphy et al., 1971). We use the RTM to estimate both S_n and L_g site response to study to better understand the factor that can influence ground motion amplification at regional distances in China.

East Asia has a complex tectonic history, much of which is related to the collision between the Indian and Eurasian plates, the subduction of the Pacific plate beneath the Eurasian plate, and the complex topography change from Tibet to the South China sea. As a major part of eastern Asia, China is highly diverse in geologic structure, with both ancient tectonic blocks and active orogenies. All of these tectonic processes lead to seismic heterogeneities in the lithosphere beneath China that can effect the amplification of ground motion at regional distances.

In the southwestern portion of our study region, the Tibetan plateau is the highest and largest continental plateau on planet Earth. It is also one of the most active continent-continent collisions and is a result of the convergence between the Indian and Eurasian plates. This convergence started $\sim$ 50 Ma and the Indian plate is continuously subducting beneath Tibet (e.g. Yin and Harrison, 2000; Chen et al., 2017). The northward collision of the Indian plate with the Eurasian plate caused the Himalayan and Tien Shan orogenies and crustal shortening and uplifting of the Tibetan plateau, accompanied by eastward extrusion of portions of the plateau (Yin and Harrison, 2000; Wang and Shen, 2020). To the north, the regional deformation related to the India-Eurasia convergence continues into Mongolia. This region is regarded as one of the most tectonically active intra-continental regions in the world (Choi et al., 2018), although compared to mainland China, GPS data show only small a little bit of upper crustal deformation ($\sim$4 mm/yr, Wang and Shen, 2020). At the eastern margin of the Eurasian plate, the Pacific and Philippine Sea plates subduct

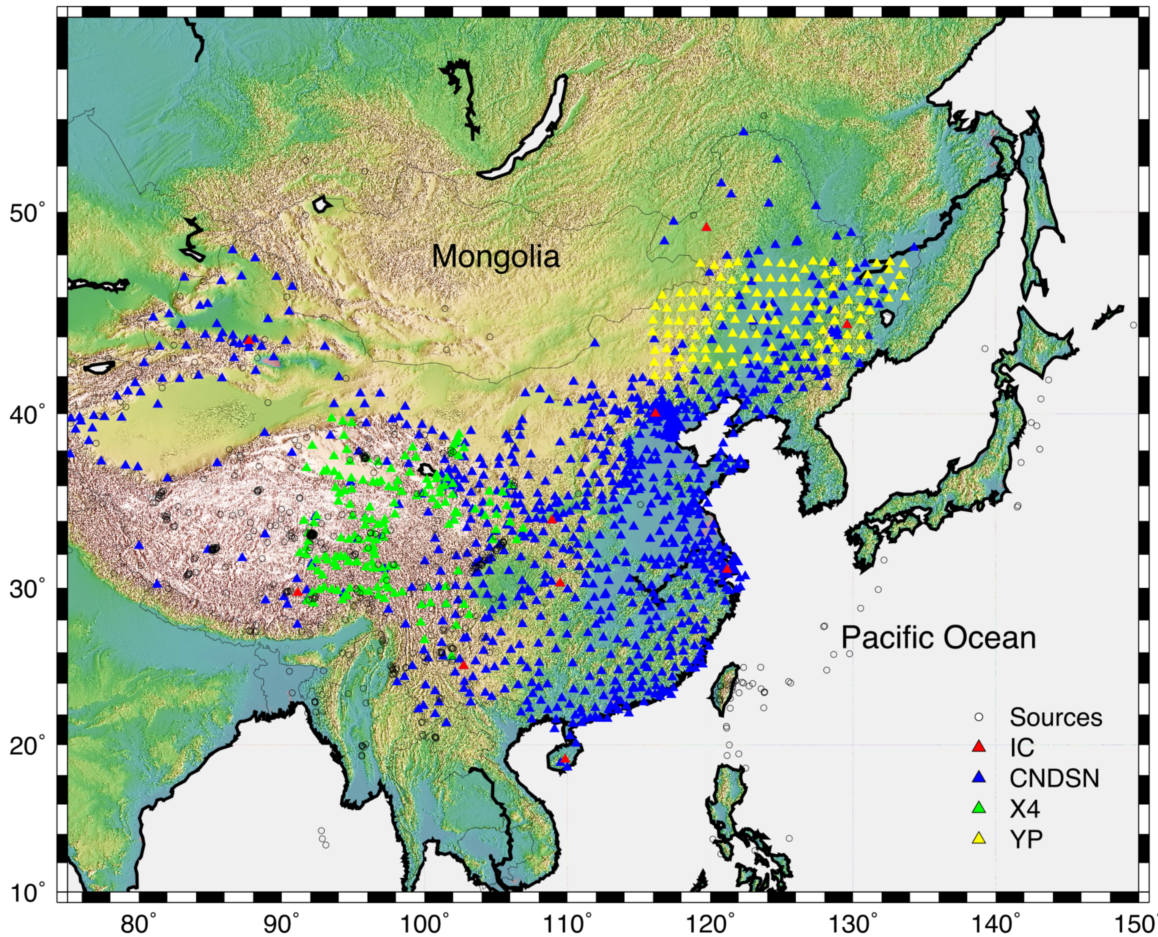


Figure 4.1: The distribution of events and stations.

underneath the Eurasian plate, likely leading to the reactivation of the North China block at a significantly high rate (Wang and Shen, 2020). All of this tectonic activity have created a series of mountain ranges and basins that we have used to investigate the factors that influence site amplification.

4.2 Data

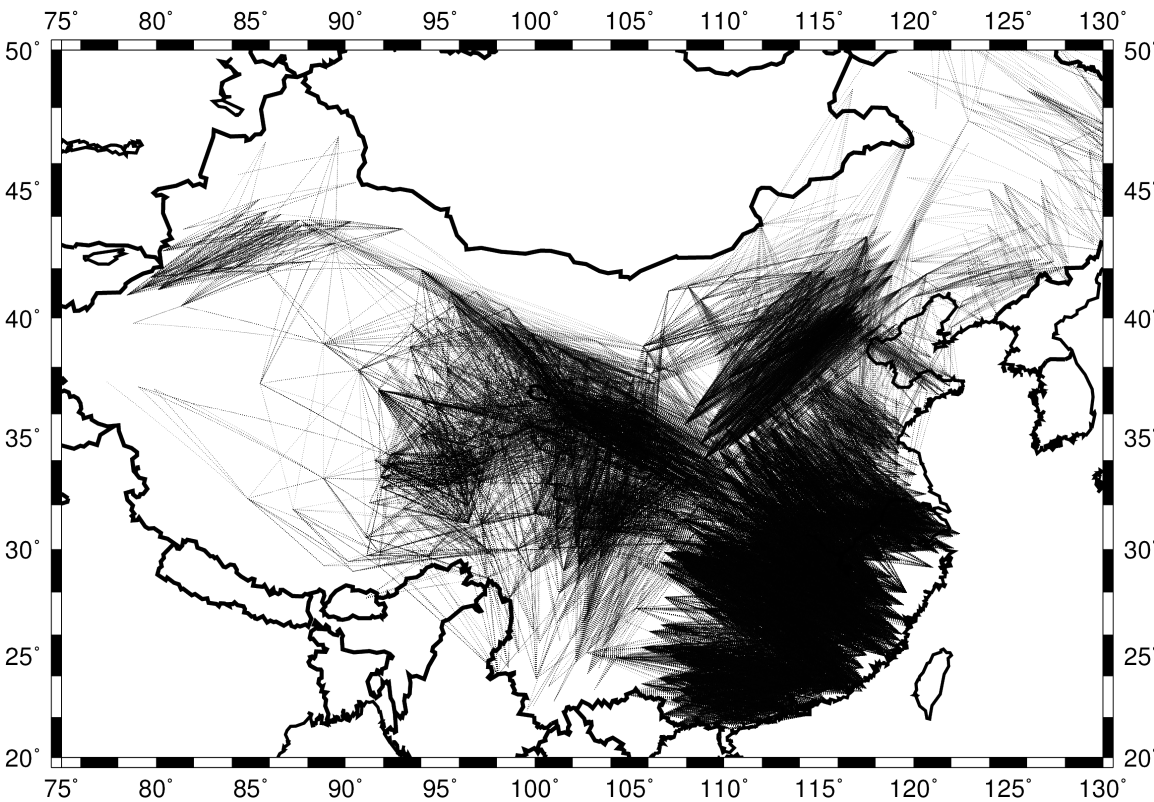


Figure 4.2: RTM Ray coverage of data used in this study.

The data used in this study include 484 regional earthquakes recorded by 1188 stations (Figure 4.1). There are more than 200,000 RTM paths as shown in (Figure 4.2). The stations are deployed throughout China and are from several seismic networks. The network X4 includes 167 broadband stations covering eastern Tibet from 5 major seismic networks: Indepth IV, Namche Barwa, MIT-China, NETS and ASCENT. The events we used were recorded between 2003 and 2009. The network

CNDSN includes 891 stations covering almost the whole China mainland. The events are between 2009 and 2011. However, it was not possible to find overlapping events between these two data sets. To link these two data sets together, which is necessary in the RTM, we used 10 stations from network IC and 120 stations from the seismic network YP. YP covers northeastern China, while IC stations cover nearly all of China. All the events used in this study have magnitudes greater than 4.5 and occur within the crust. In this study, we only use the vertical component data because the signal to noise ratios tend to be higher.

4.3 Methods

The method we use is the reverse two station (RTM) method. Suppose the amplitude of regional phase in frequency domain can be denoted by

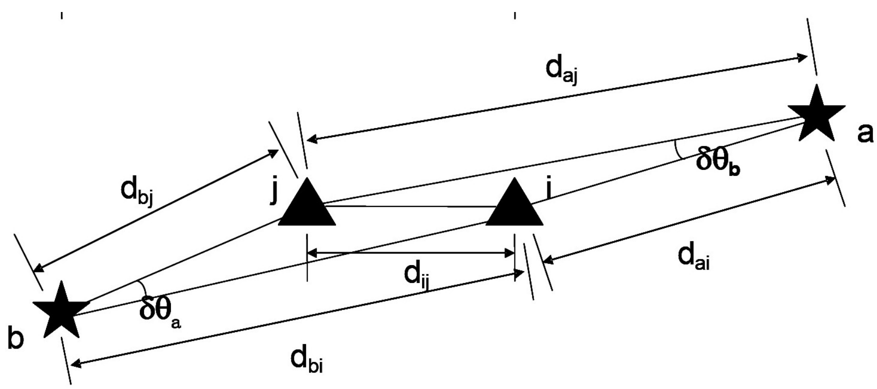


Figure 4.3: Scheme of the reverse two-station method (RTM).

$$A = S \cdot R \cdot G \cdot I_S \cdot S_S \cdot I_Q \cdot T_{FD} \cdot T_{SC} \cdot T_{AN}, \tag{4.1}$$

where S denotes the source excitation function; R, the focal mechanism factor; G, the geometrical spreading (e.g., $G = G_0 d^{-m}$ for L_g and S_n); I_S, the instrument response; S_S, the station site response; I_Q, intrinsic attenuation; T_{FD}, the coefficient of focusing and defocusing; T_{SC}, the scattering coefficient; and T_{AN}, other effects. T_{FD} is typically assumed to be one if we put a limit on standard errors of the measurements (Xie, 2002) and I_Q, T_{SC}, T_{AN} are interpreted as apparent Q. Thus, the spectral amplitude is simplified as

$$A(f,d) = S(f) \cdot R(f,\varphi) \cdot I_S(f) \cdot S_S(f) \cdot G(d) \exp\left(-\frac{\pi f d}{vQ}\right), \tag{4.2}$$

where f denotes frequency and d, the epicentral distance; φ, azimuth; v, the regional seismic phase velocity.

We use i, j to denote stations and a, b for events (Figure 4.3). The four spectral amplitude equations are

$$\begin{cases} A_{ai}(f,d_{ai}) = S_a(f)R_a(f,\varphi)I_i(f)S_{S_i}(f)G_{ai}(d_{ai}) \exp\left(-\dfrac{\pi f d_{ai}}{v_i Q_i}\right) \\ A_{aj}(f,d_{aj}) = S_a(f)R_a(f,\varphi)I_j(f)S_{S_j}(f)G_{aj}(d_{aj}) \exp\left(-\dfrac{\pi f d_{aj}}{v_j Q_j}\right) \\ A_{bi}(f,d_{bi}) = S_b(f)R_b(f,\varphi)I_i(f)S_{S_i}(f)G_{bi}(d_{bi}) \exp\left(-\dfrac{\pi f d_{bi}}{v_i Q_i}\right) \\ A_{bj}(f,d_{bj}) = S_b(f)R_b(f,\varphi)I_j(f)S_{S_j}(f)G_{bj}(d_{bj}) \exp\left(-\dfrac{\pi f d_{bj}}{v_j Q_j}\right) \end{cases} \tag{4.3}$$

If the tolerance azimuth difference ($\delta\theta$ in Figure 4.3) is small enough, we can treat R of two stations from the same events as non-function of φ. From Equation 4.3, we get

$$\begin{cases} \dfrac{A_{ai}}{A_{aj}} = \dfrac{I_i S_{S_i} G_{ai}}{I_j S_{S_j} G_{aj}} \exp\left(\dfrac{\pi f d_{aj}}{v_j Q_j} - \dfrac{\pi f d_{ai}}{v_i Q_i}\right) \\ \dfrac{A_{bi}}{A_{bj}} = \dfrac{I_i S_{S_i} G_{bi}}{I_j S_{S_j} G_{bj}} \exp\left(\dfrac{\pi f d_{bj}}{v_j Q_j} - \dfrac{\pi f d_{bi}}{v_i Q_i}\right) \end{cases} \tag{4.4}$$

By multiplying the two ratios in Equation 4.4, we get

$$\frac{A_{ai}A_{bi}}{A_{aj}A_{bj}} = \left(\frac{I_i S_{S_i}}{I_j S_{S_j}}\right)^2 \left(\frac{G_{ai}G_{bi}}{G_{aj}G_{bj}}\right) \exp\left(\frac{\pi f d_{aj}}{v_j Q_j} - \frac{\pi f d_{ai}}{v_i Q_i} + \frac{\pi f d_{bj}}{v_j Q_j} - \frac{\pi f d_{bi}}{v_i Q_i}\right). \tag{4.5}$$

Now, suppose that the apparent Q is identical along the path and substitute 1D velocity v and $G = G_0 d^{-m}$,

$$\begin{aligned}
\frac{A_{ai}A_{bi}}{A_{aj}A_{bj}} &= \left(\frac{I_i S_{S_i}}{I_j S_{S_j}}\right)^2 \left(\frac{d_{ai}d_{bi}}{d_{aj}d_{bj}}\right)^{-m} \exp\left[\frac{\pi f}{vQ}(d_{aj} - d_{ai} + d_{bj} - d_{bi})\right] \\
&= \left(\frac{I_i S_{Si}}{I_j S_{Sj}}\right)^2 \left(\frac{d_{ai}d_{bi}}{d_{aj}d_{bj}}\right)^{-m} \exp\left\{\frac{d_{aj} - d_{ai} + d_{bj} - d_{bi}}{d_{aj} - d_{ai} - d_{bj} + d_{bi}} \ln\left[\frac{A_{ai}A_{bj}}{A_{aj}A_{bi}}\left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^m\right]\right\}. \\
&= \left(\frac{I_i S_{Si}}{I_j S_{Sj}}\right)^2 \left(\frac{d_{ai}d_{bi}}{d_{aj}d_{bj}}\right)^{-m} \exp\left\{\ln\left[\frac{A_{ai}A_{bi}}{A_{aj}A_{bi}}\left(\frac{d_{ai}d_{bj}}{d_{aj}d_{bi}}\right)^m\right]^{\frac{d_{aj}-d_{ai}+d_{bj}-d_{bi}}{d_{aj}-d_{ai}-d_{bj}+d_{bi}}}\right\}
\end{aligned} \tag{4.6}$$

Then, the ratio of the two site responses can be expressed in logarithmic form:

$$\begin{aligned}
\ln\frac{S_{Si}}{S_{Sj}} &= \ln S_{Si} - \ln S_{Sj} \\
&= \ln\frac{I_j}{I_i} + \frac{d_{aj} - d_{ai}}{d_{aj} + d_{bi} - d_{ai} - d_{bj}} \ln\frac{A_{ai}d_{ai}^m}{A_{aj}d_{aj}^m} + \frac{d_{bi} - d_{bj}}{d_{aj} + d_{bi} - d_{ai} - d_{bj}} \ln\frac{A_{bi}d_{bi}^m}{A_{aj}d_{bj}^m}.
\end{aligned} \tag{4.7}$$

The relative site response can be deterministically solved by an inversion problem:

$$\begin{bmatrix} 1 & -1 & 0 & 0 & \cdots \\ 1 & 0 & -1 & 0 & \cdots \\ 0 & 1 & -1 & 0 & \cdots \\ 0 & 0 & 1 & -1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \ln S_{S1} \\ \ln S_{S2} \\ \ln S_{S3} \\ \ln S_{S4} \\ \vdots \end{bmatrix} = \begin{bmatrix} RS_{12} \\ RS_{13} \\ RS_{23} \\ RS_{34} \\ \vdots \end{bmatrix}, \tag{4.8}$$

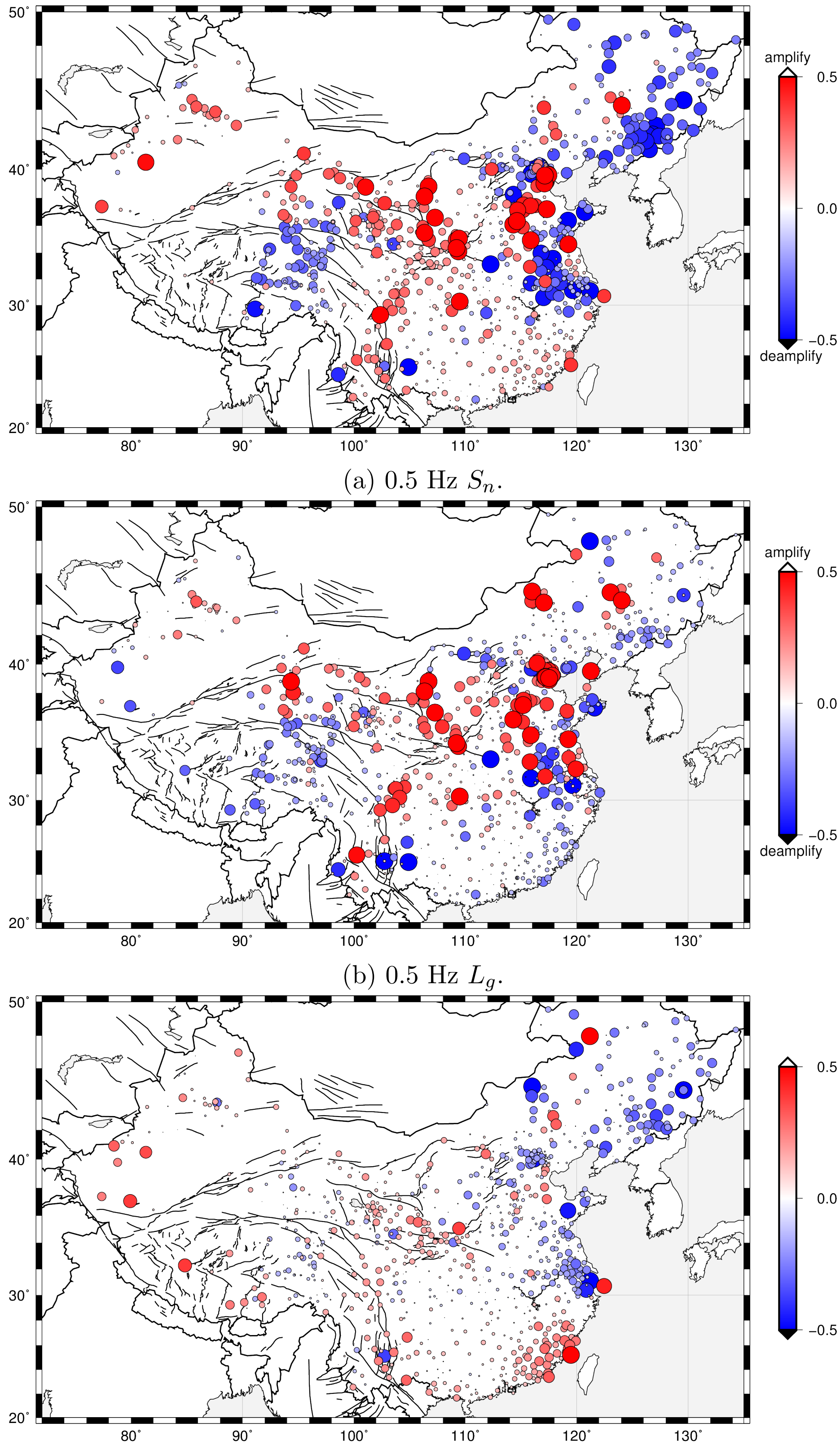
where the RS_{ij} denotes the relative site response between station i and j, in other words, the right hand side of Equation 4.7. Then, we use the LSQR Paige and Saunders (1982) algorithm to solve the inversion problem to get the site response for each station. The damping value used is 0.25. We should note that this inverse problem is largely underdetermined in terms of the absolute site amplification. We can, however, determine with a high degree of confidence the variations in site amplification.

4.4 Results

We have estimated the frequency-dependent site amplification for all the stations with both S_n and L_g phases. Then we subtract L_g site amplification from S_n site amplification to get the differential site amplification ($SR_{S_n} - SR_{L_g}$). We see high site amplification at the edge of basins, such as Sichuan basin, the Songliao basin, the Junggar basin, and the west edge of the Ordos basin and the northeast edge of Tibet. However, we see site de-amplification in mountain ranges, such as Tibet and Changbaishan, and in the North China Plain. There are also several stations with de-amplification found at the northeastern edge of Ordos. The site response results of L_g and S_n show a similar pattern.

4.4.1 Site response of S_n

The S_n site amplification patterns are similar at all of the frequency bands we have examined. We observe strong S_n site amplification at stations in the Tarim and Sichuan basins. The S_n site amplification stations are also found along the Tienshan Suture, Qilianshan and Longmenshan. All these stations are located at the edge of the Tarim, Qaidam, and Sichuan basins, respectively. We observe strong site amplification at the western edge of the Ordos plateau and the eastern block of the North China Craton. It is worth noting there are basins all along the edges of the Ordos plateau. The medial site amplification stations are found in the Junggar Basin, the Sichuan Basin, and the South China Block. There are several stations with strong site amplifications found at the edges of the Erlian and Songliao Basins. We find evidence of site de-amplification for S_n phases for stations located within the Changbaishan mountain belt, the North China Plain, and the Tibetan Plateau. In general we see de-amplification in high topography. There are also several S_n stations with strong de-amplification found in the middle of the Tarim Basin and the

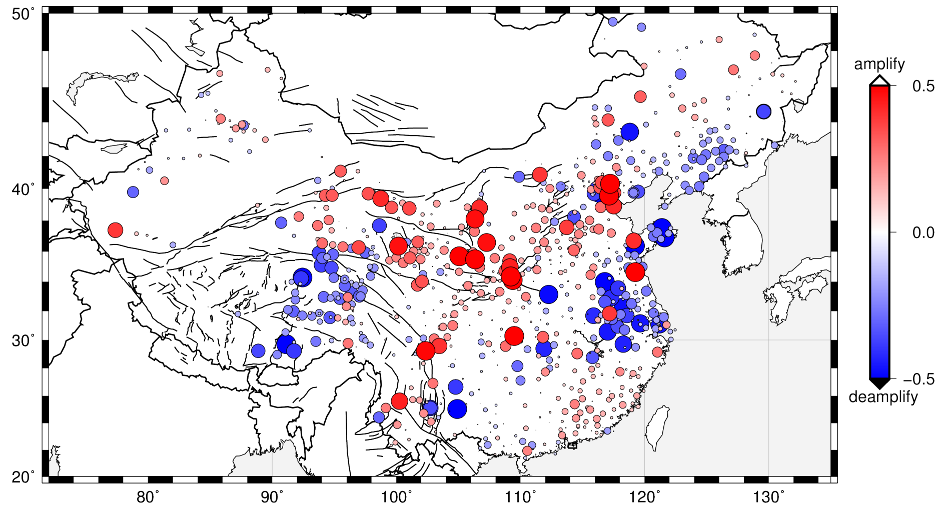


(a) 0.5 Hz S_n.

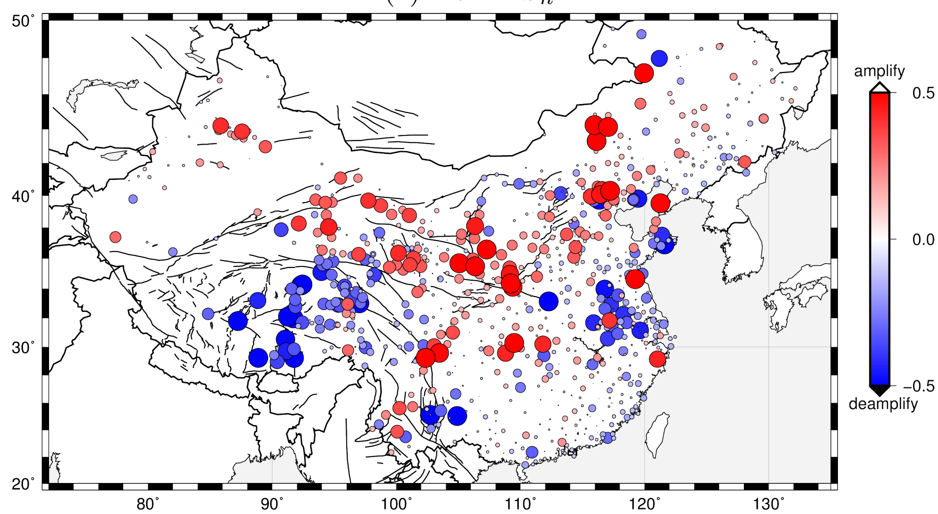
(b) 0.5 Hz L_g.

(c) 0.5 Hz $SR_{S_n} - SR_{L_g}$.

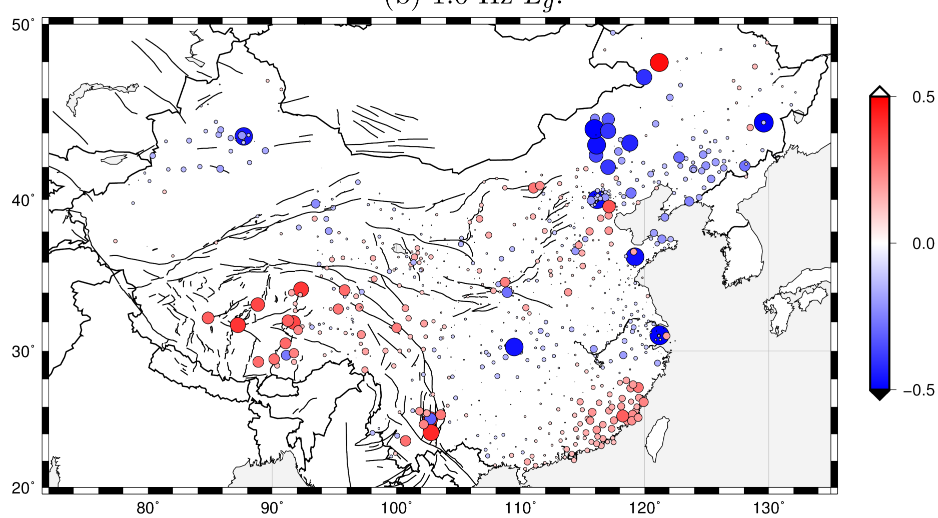
Figure 4.4: Site response results (0.5 Hz).



(a) 1.0 Hz S_n.

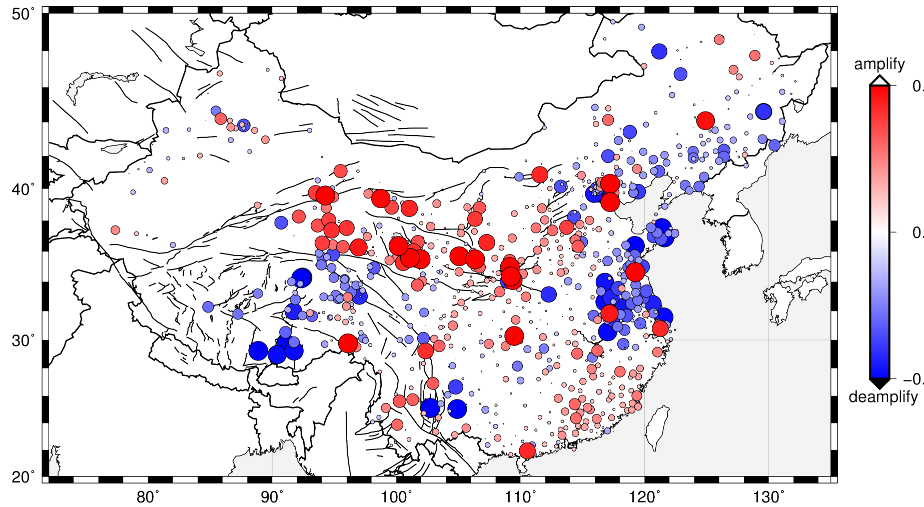


(b) 1.0 Hz L_g.

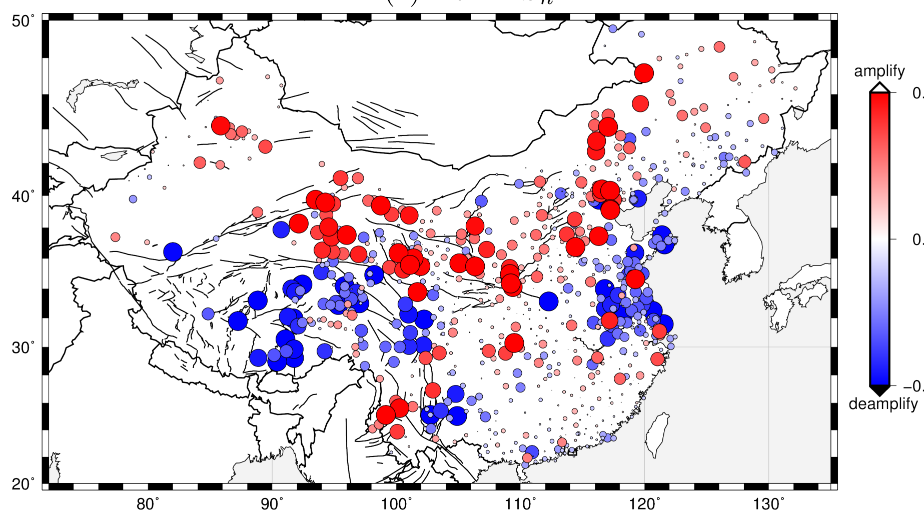


(c) 1.0 Hz $SR_{S_n} - SR_{L_g}$.

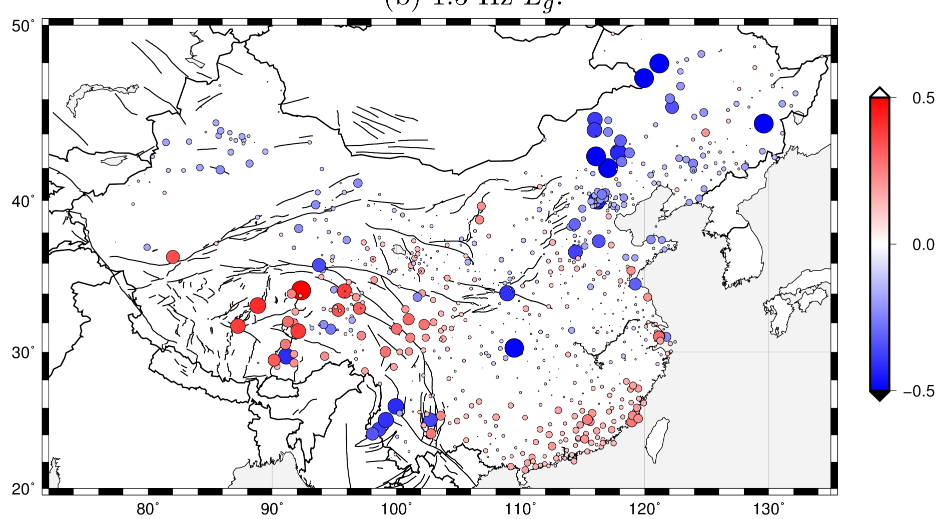
Figure 4.5: Site response results (1.0 Hz).



(a) 1.5 Hz S_n.

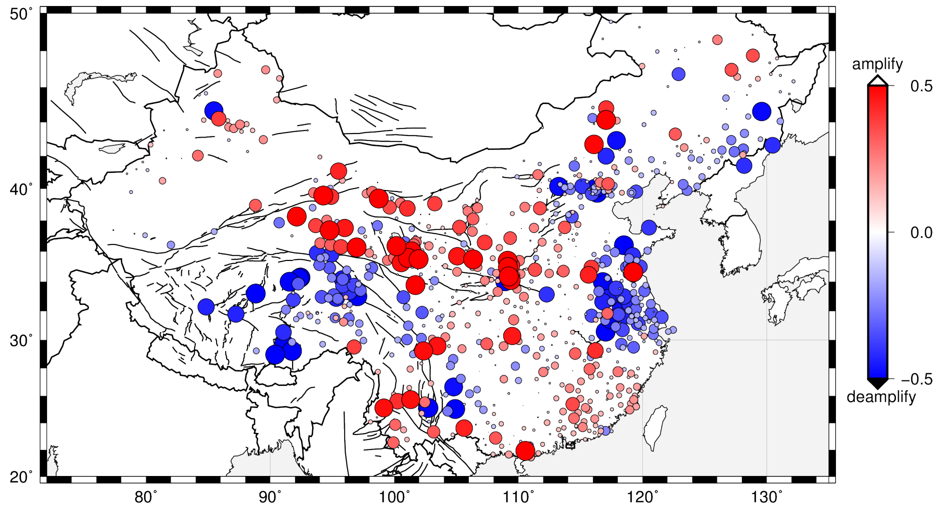


(b) 1.5 Hz L_g.

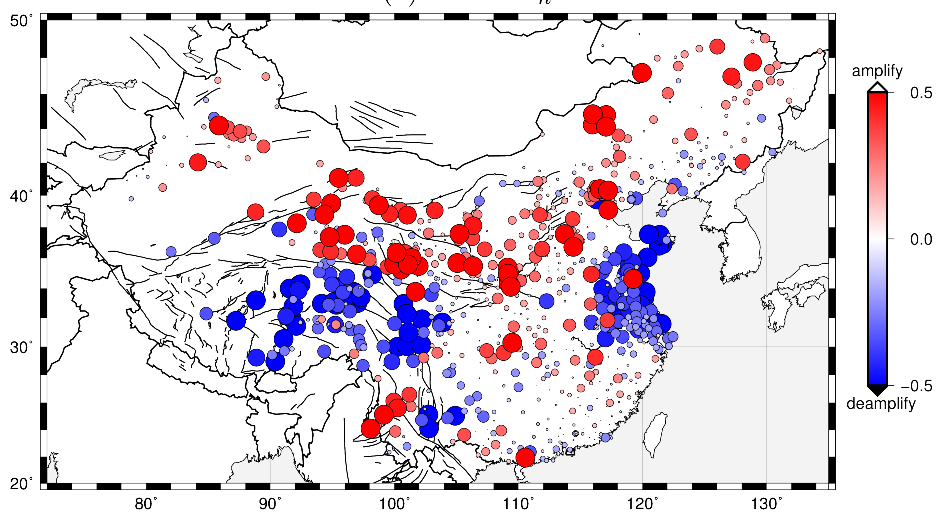


(c) 1.5 Hz $SR_{S_n} - SR_{L_g}$.

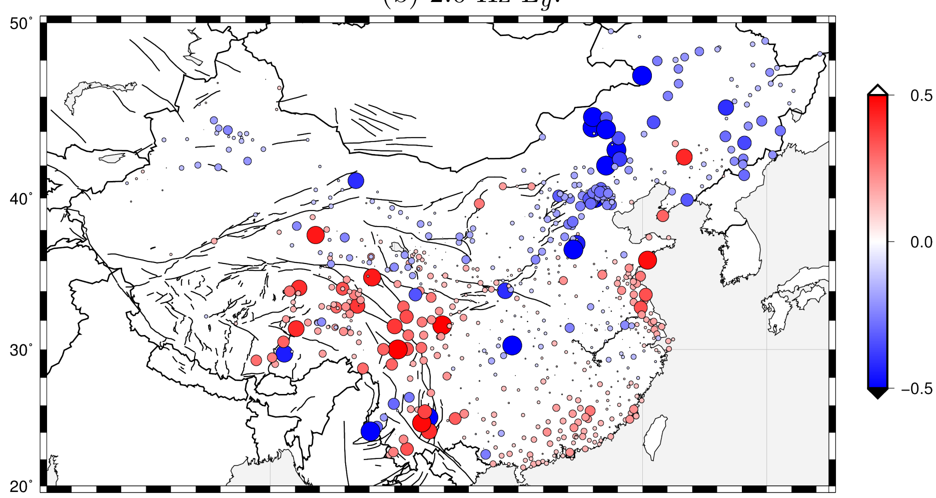
Figure 4.6: Site response results (1.5 Hz).



(a) 2.0 Hz S_n.

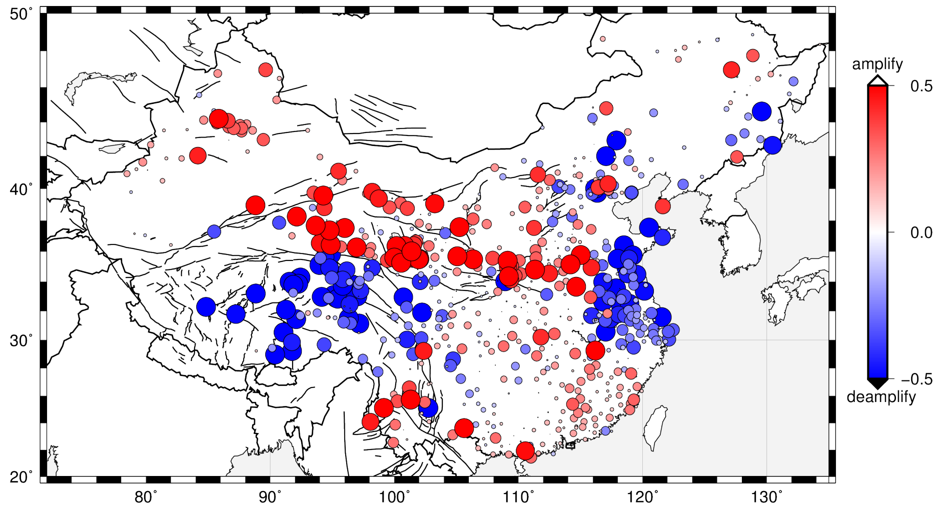


(b) 2.0 Hz L_g.

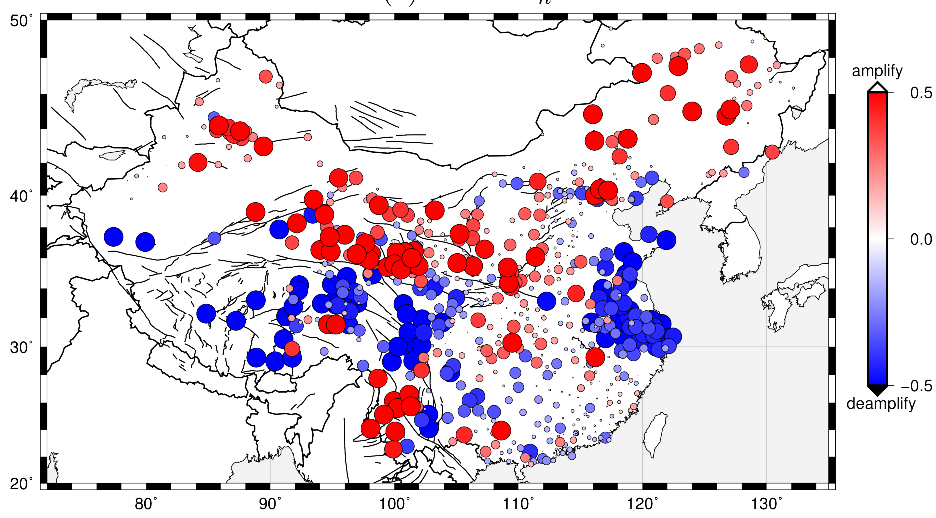


(c) 2.0 Hz $SR_{S_n} - SR_{L_g}$.

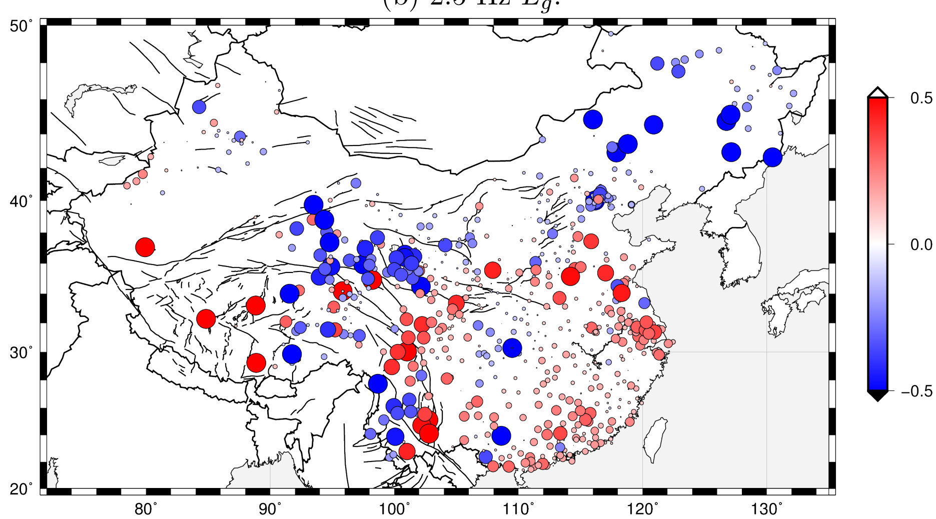
Figure 4.7: Site response results (2.0 Hz).



(a) 2.5 Hz S_n.

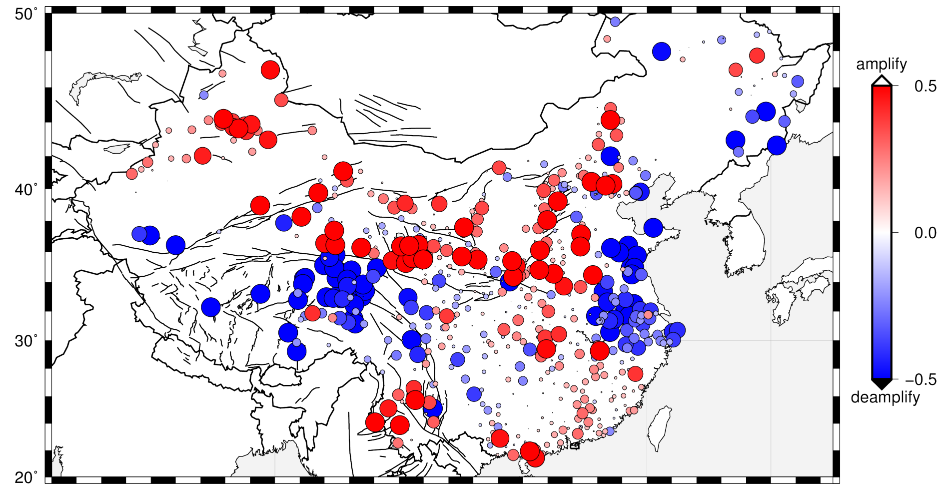


(b) 2.5 Hz L_g.

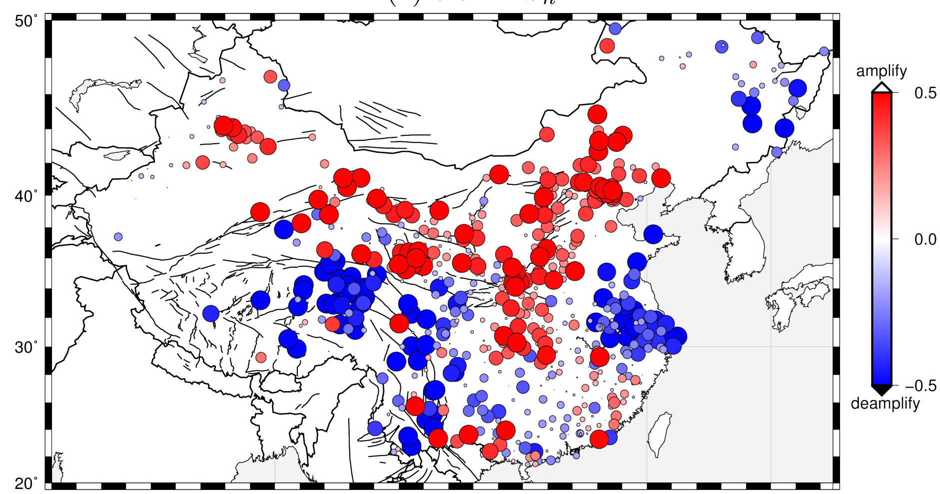


(c) 2.5 Hz $SR_{S_n} - SR_{L_g}$.

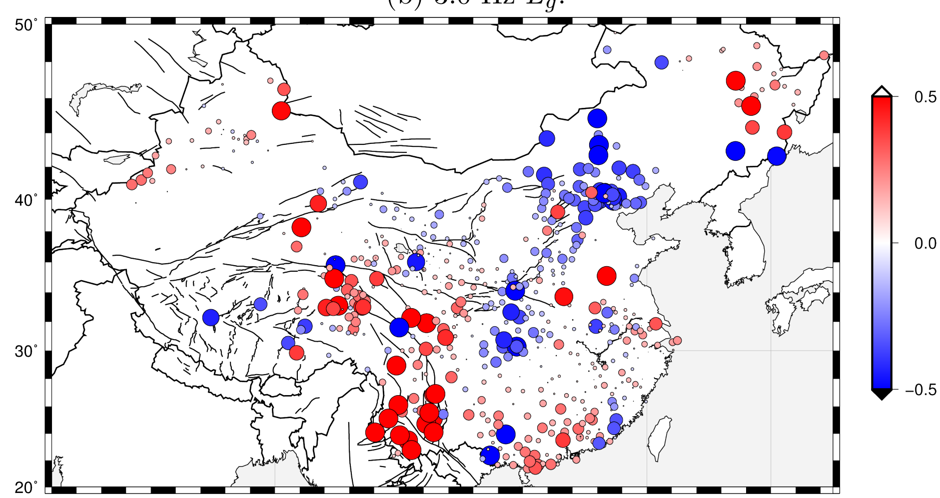
Figure 4.8: Site response results (2.5 Hz).



(a) 3.0 Hz S_n.

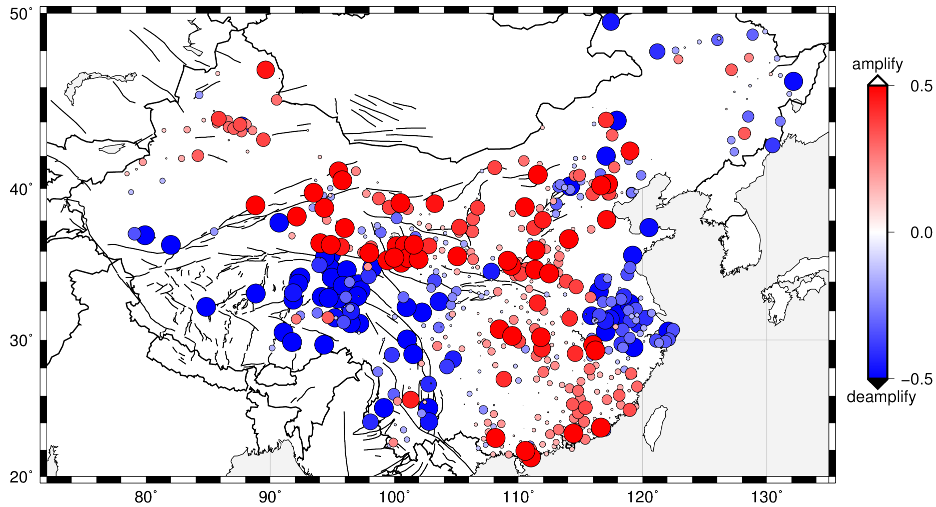


(b) 3.0 Hz L_g.

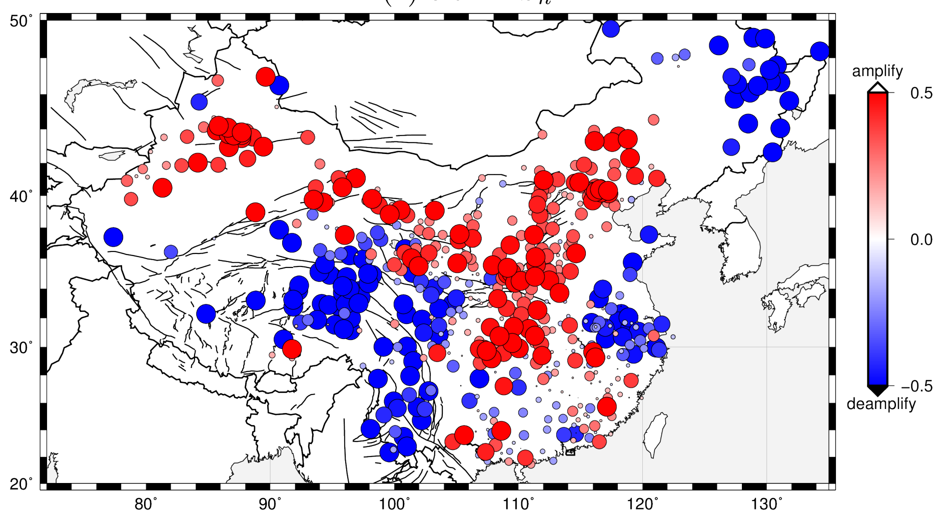


(c) 3.0 Hz $SR_{S_n} - SR_{L_g}$.

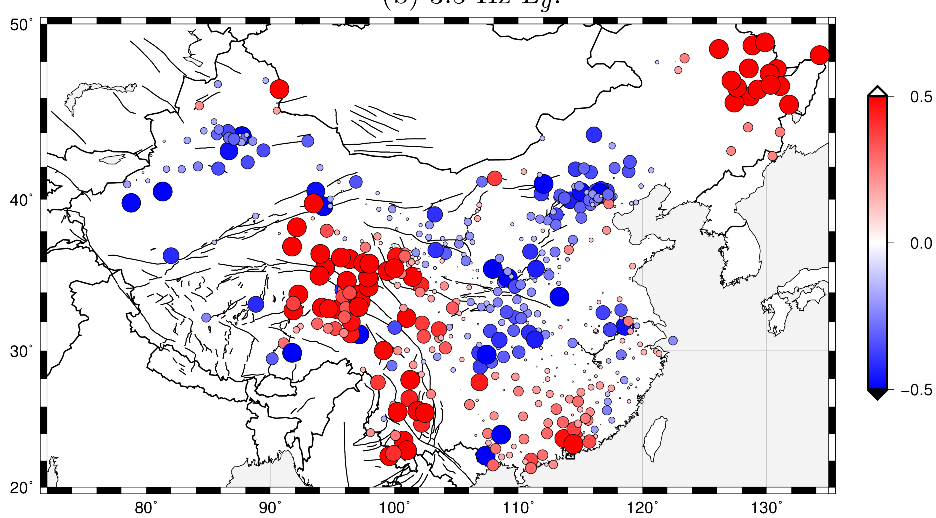
Figure 4.9: Site response results (3.0 Hz).



(a) 3.5 Hz S_n.

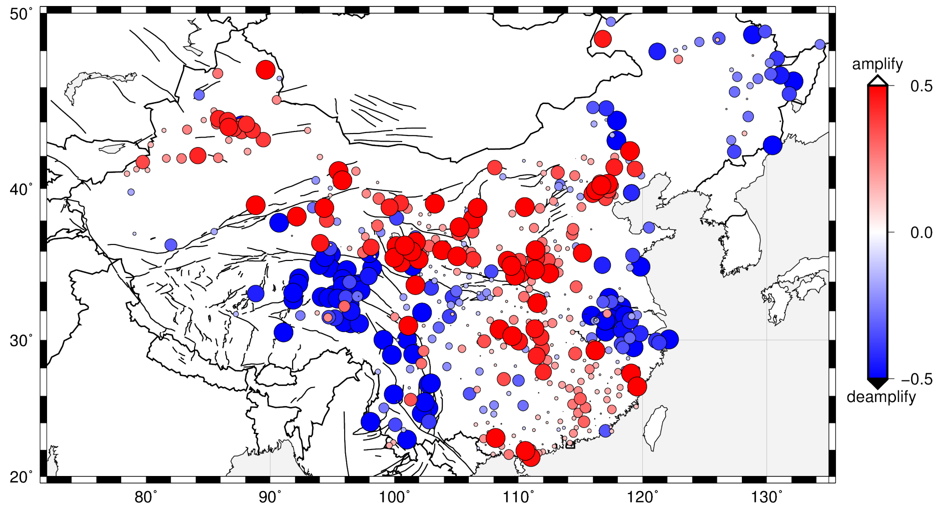


(b) 3.5 Hz L_g.

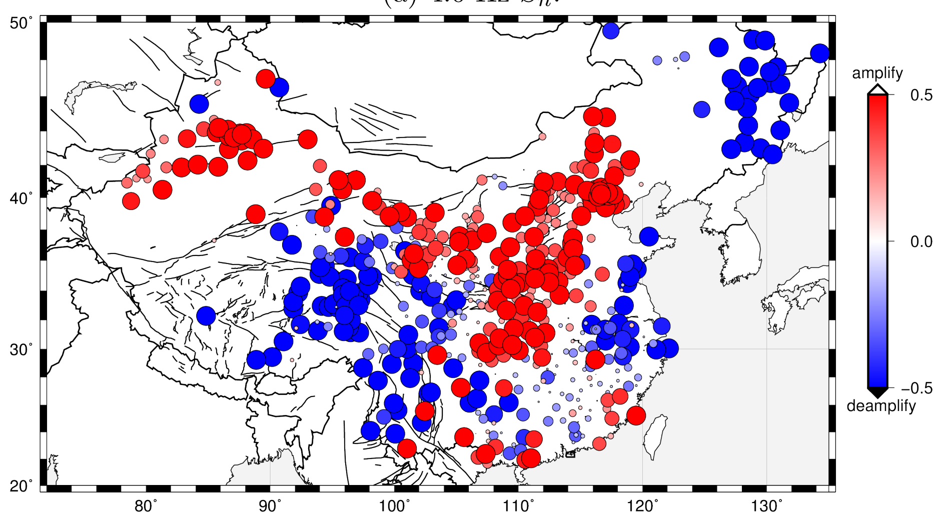


(c) 3.5 Hz $SR_{S_n} - SR_{L_g}$.

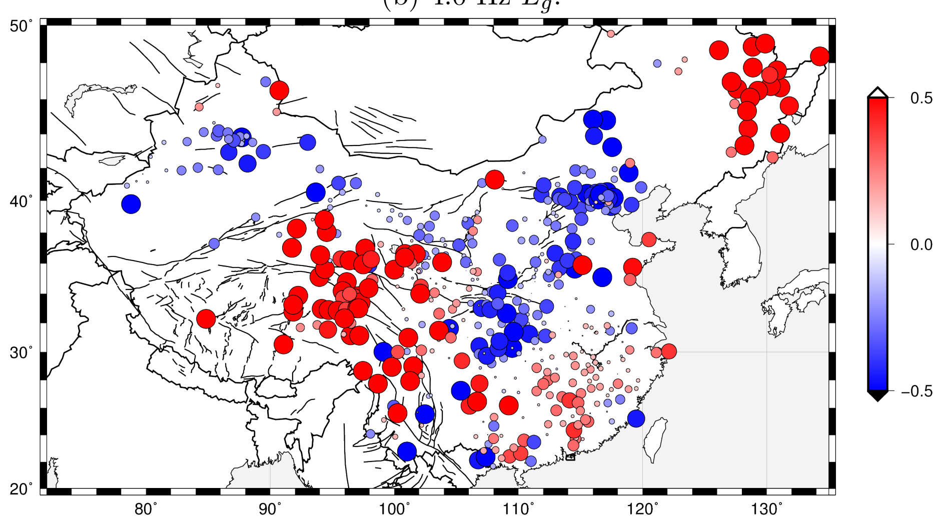
Figure 4.10: Site response results (3.5 Hz).



(a) 4.0 Hz S_n.



(b) 4.0 Hz L_g.



(c) 4.0 Hz $SR_{S_n} - SR_{L_g}$.

Figure 4.11: Site response results (4.0 Hz).

Sichuan Basin and at the eastern edge of the Ordos Basin. However, these station results are not constant with frequency. We do not see them on 1 Hz results. With increasing frequency, the absolute value of both amplification and de-amplification becomes higher, which means that the site response results are more sensitive for higher frequency S_n phases. We estimate the site response of fewer stations with higher frequency because signal to noise tends to get worse with increases in frequency. We do not have high frequency S_n signal for a number of stations.

4.4.2 Site response of L_g

For the L_g phase, we found that the site response results have a similar pattern with that of S_n phase. For lower frequency (0.5 Hz) L_g phase, we found positive site amplification at stations at the edge of Junggar Basin; however, the stations in the middle of the Tarim Block are found to have de-amplification (Figure 4.4b). We also observe positive site amplification at stations at the edge of Tibetan Plateau and the de-amplification stations in the middle of Tibetan Plateau (Figure 4.5b), which is consistent with the previous study of Bao et al. (2011). We observe site de-amplification for stations located within the Lesser Xing'an Range and the Songen-Zhangguangcai Range and site amplification stations in the Songliao Basin (Figure 4.4b), which is consistent with the results of Ranasinghe et al. (2018). We found consistently positive site amplification stations within the Ordos plateau and the Sichuan Basin and site de-amplification stations in the North China Plain. However, we found that the stations in the South China Block do not have a strong positive site response for lower frequency L_g phases (Figure 4.4b, 4.5b and 4.7b). Using higher frequencies, such as 4.0 Hz L_g, we found strong de-amplification and amplification stations mixed across the South China Block (Figure 4.11b). This complexity may reflect a complex geologic structure throughout the block and thus no coherent amplification.

4.4.3 Site response difference between S_n and L_g

The L_g site response patterns are similar to the S_n site response with RTM because site response is likely a function of the local seismic structure, properties of the near surface rocks, and the local topography. However, it is worth noting that when we compare the site response between S_n and L_g phases (Figure 4.4), we found that they are not identical with each other. For lower frequency (0.5 Hz) data, we observe $SR_{S_n} < SR_{L_g}$ at nearly every station across nearly all of Northeastern China (the Greater Xing'an Range, the Lesser Xing'an Range and the Songen-Zhangguangcai Range), the North China Plain, and the Ordos (Figure 4.4c). We also found $SR_{S_n} < SR_{L_g}$ for stations located within the Tibetan Plateau; however, we found $SR_{S_n} > SR_{L_g}$ for stations across the South China Block, the Junggar Basin, the Tarim Block, and at the eastern edge of the Tibetan Plateau with low frequency data. We also found $SR_{S_n} > SR_{L_g}$ for stations located along the western margin of the Ordos plateau. There are several stations with $SR_{S_n} > SR_{L_g}$ in the eastern Block of the North China Craton (Figure 4.4c). With high frequency data ($\geq$ 1.0 Hz), we found $SR_{S_n} < SR_{L_g}$ for stations in the Junggar Basin and the Tarim Block (e.g., Figure 4.7c). We observe $SR_{S_n} > SR_{L_g}$ in the Tibetan Plateau and the South China Block. We also found $SR_{S_n} > SR_{L_g}$ in the North China Plain when the data frequency is $>$ 2.0 Hz (Figure 4.7c and 4.11c). Almost all the stations in the Lesser Xing'an Range and Songen-Zhangguangcai Range have $SR_{S_n} > SR_{L_g}$ for the 4.0 Hz frequency band (Figure 4.11c). The absolute values of differential site response are much smaller for low frequency ($<$ 2.0 Hz) data than for high frequency data. Almost all the positive differential site response ($SR_{S_n} > SR_{L_g}$) stations are found in areas with high topography with high frequency data, such as Changbaishan and Tibet, except those in the Cathaysia Foldbelt. However, with low frequency data ($<$ 2.0 Hz), we don't see positive differential site response stations in the Changbaishan area, although we still see them in Tibet and the Cathaysia Foldbelt.

4.5 Discussion and Conclusion

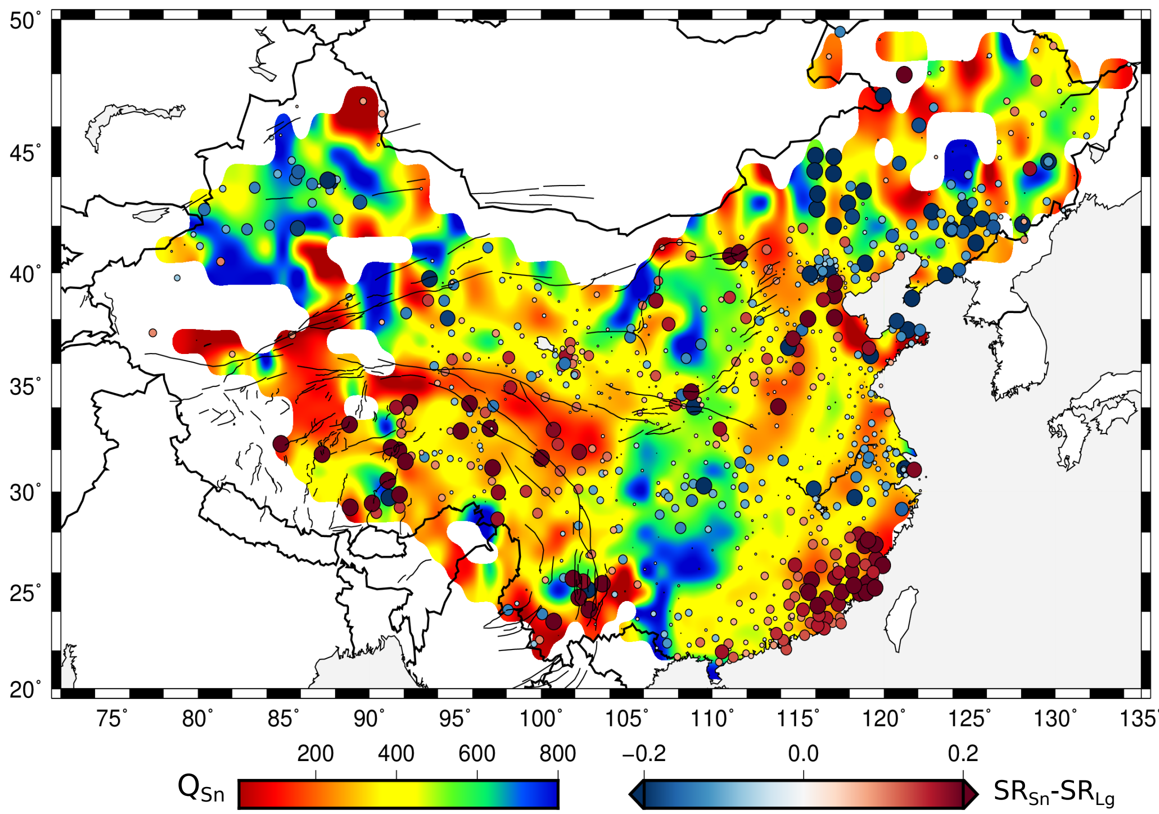


Figure 4.12: S_n Q and difference between S_n and L_g site amplification at 1 Hz.

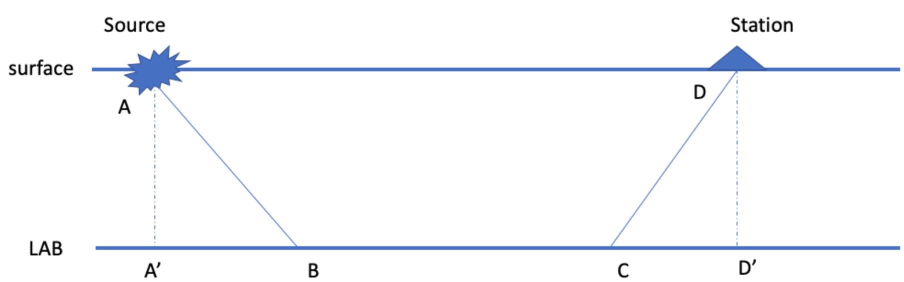


Figure 4.13: Sn propagation path plot.

The L_g site response patterns are similar to the S_n site response, perhaps because site amplification is to some degree a function of the local effect of topography, but they are not identical, possibly because the RTM may not entirely separate path attenuation from site response. These observations suggest that the site response and the path attenuation are both factors. While L_g site response is very likely

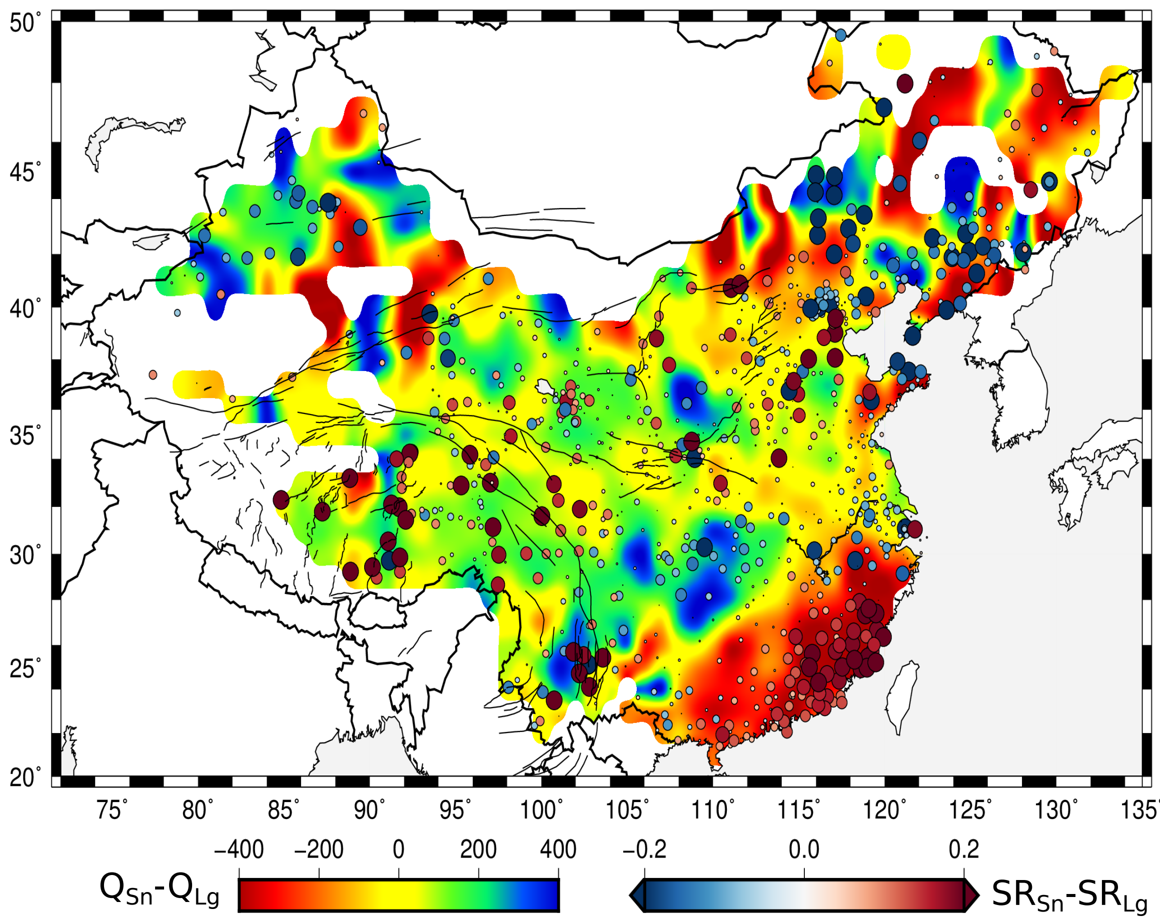


Figure 4.14: Q and site amplification difference at 1 Hz.

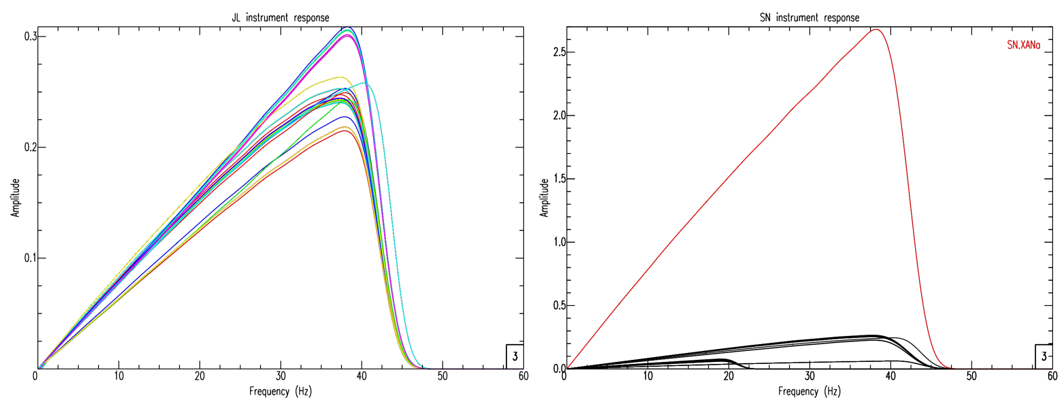
only due to the near station earth structure, the S_n site response may also be a function of the crustal portion of the S_n ray path. We have evaluated this idea by subtracting the L_g site amplification from S_n site amplification. In effect this difference in amplification should be a function of the S_n crustal leg effect. We have also compared the site response difference between S_n and L_g with S_n Q. Figure 4.12 indicates the difference between S_n site amplification and L_g site amplification plotted over our S_n Q tomography model. Almost all the stations with positive site response difference are located in the regions with low S_n Q, such as Tibet. The stations with negative site response difference are located in the regions with high S_n Q, such as the Changbaishan and Sichuan Basins.

Figure 4.13 is the plot of S_n propagation. Using our RTS assumptions, we assume that the attenuation for path ABCD is the same as the attenuation for path A′BCD′. However, in high S_n attenuation regions, attenuation A′B may be larger than that of AB. If we still keep the assumption, that leads to determined Q being higher than the true Q. The rest of the S_n energy may leak to the surface and amplify the high frequency S_n amplitude. For a region with low S_n attenuation, the attenuation A′B may be less than that of AB. That leads to the effective Q being lower than the true Q, so that the site amplification energy would leak into path attenuation. Thus, in a low S_n Q zone, the S_n and L_g site response difference tends to be positive. Otherwise, in a high S_n Q zone, the site response difference tends to be negative (i.e. $Q_{S_n} < Q_{L_g}$) .

To evaluate our theory, we compared the site amplification difference with the Q difference between L_g and S_n (Figure 4.14). This figure shows S_n Q subtracted from L_g Q as the Q difference and S_n site response minus the L_g site response as the site response difference. Except for some stations in Tibet, almost all the positive site response difference stations are located in a region with a negative Q difference (i.e. $Q_{S_n} < Q_{L_g}$) regions. In conclusion, in regions where the lithospheric mantle Q is less than crustal Q, the effective S_n Q may be larger than the true Q for the crust, which leads to the S_n energy leaking into the site amplification. In regions where the lithospheric mantle Q is larger than the crustal Q, the effective S_n Q may be less than the true Q, which suggests that the site amplification energy is leaking into the path.

Appendix

Data Processing



(a) Instrument response of stations in network JL.

(b) Instrument response of stations in network SN.

Figure 4.15: The instrument response examples.

After we checked the station information by location, we found that there is one station with a clearly wrong location (in the ocean, black triangle in Figure 4.17). We also checked the instrument response by plotting both the instrument response in the frequency domain and the amplitudes recorded by stations from the same network. We found that there are seven stations with no instrument response file converted by our code (green triangles in Figure 4.17). There are 16 stations with infinite response at all frequencies (red triangles in Figure 4.17). If the instrument

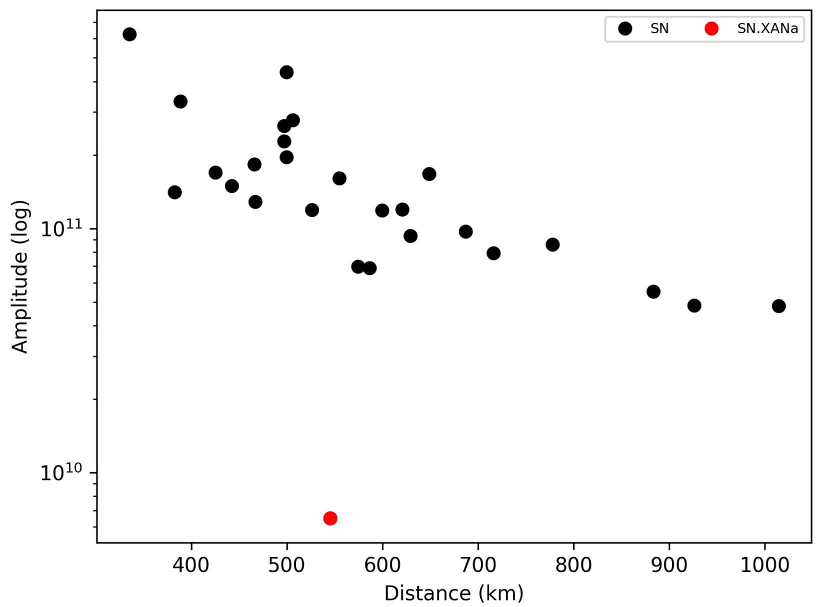


Figure 4.16: Amplitude of stations from network SN.

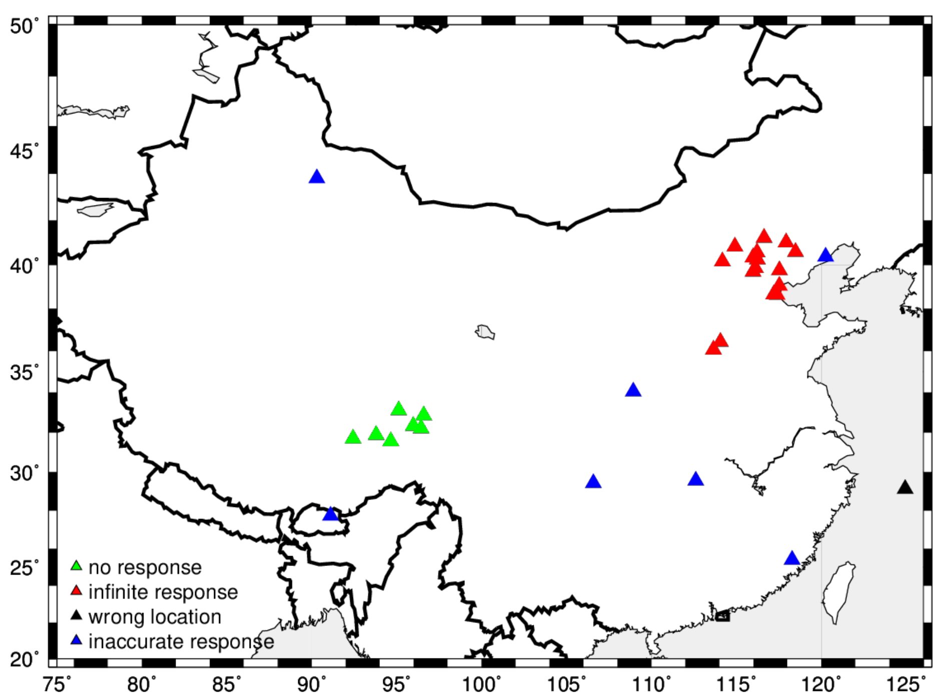


Figure 4.17: Stations not used in this study.

response curve is similar to others in the same network (Figure 4.15a), we treat the data as usable. There also are some stations with a response that is clearly different from the others in the same network (Figure 4.15b). We plot the amplitudes of the same event data (Figure 4.16) with hypocentral distance. Because we are using the same event, we know that the amplitude is a function of site response, instrument response and apparent attenuation (Eq. 4.2). The stations in the same network are located close to each other, so that we could ignore the apparent attenuation effect. So, for stations with clearly different instrument response curves, if the amplitude with distance looks correct, we treat the station data as usable. However, if the amplitude looks clearly different from others, we could not distinguish the site response effect from the instrument response effect. We do not use those station data (blue triangles in Figure 4.17). After all the preprocessing, we found 31 stations not usable (Figure 4.17).

After checking all the instrument responses by plotting the response in the frequency domain (e.g. Figure 4.15), we found that there are 16 stations with a wrong instrument response file. Their response in the frequency domain is infinite. The stations we did not use due to the wrong instrument response files are: CHDc, CXTb, FENb, KUCb, SXTa, ZJKb, DHCb, LBPb, NKYb, SJZa, SSLb, T23a, XAZa, XBZb, SNZa and YAYa. If there are several stations' instrument response files that are clearly different from others in the same network (Figure 4.15b), especially in the low-pass (under 10 Hz) frequency range, we double-checked the signal amplitude with distance. There is one station with a clearly wrong location: BJTc. There are seven stations from X4 without available instrument files generated by our code: F003, F004, F005, F006, F014, F015 and F016.

Chapter 5

S_n Blockage in the ME and East Asia

Abstract

High frequency seismic wave phase blockage is often the result of strong attenuation. The regional phase S_n is more prone to blockage than any of the other regional phases including L_g. In addition, regions with extensive S_n phase blockage are typically indicative of a thin to absent lithospheric mantle. Furthermore, widespread blockage can lead to difficulty in trying to estimate source parameters or path attenuation. In this paper, we have applied two approaches to map phase blockage: (1) the relatively standardized efficiency tomography and (2) we have developed a Bayesian logistic regression model to predict the likelihood (probability) of phase blockage using a Bayesian Lasso algorithm. We applied our methods on both simulated efficiency data and real efficiency data obtained from earthquakes and stations from the Middle East (ME) and the Eastern Asia. Our models successfully predict the probability of blockage zones with relatively high accuracy ($> 75\%$). Additionally, we observe both low probability of S_n blockage and efficient S_n propagation in tectonically stable conti-

nental lithosphere, such as the Arabian Plate, the Mediterranean Sea, northeastern Iran, the Ordos plateau, and the Sichuan basin. All of these regions have stable lithospheric mantle. The regions with a high probability of S_n blockage or inefficient S_n propagation zones are in the tectonically active areas, such as the Tibetan and Iranian plateaus. Our probability of blockage model can also be used in phase identification to image the regions where S_n Q models are likely to be biased due to left censored data. Lastly, as a byproduct of our Bayesian approach, we can also provide uncertainties of the predictions which can further show regions of propagation complexity. The propagation complexity may imply tectonic complexity.

Keywords: Bayesian Lasso, S_n phase, Predicted propagation, Efficiency tomography.

5.1 Introduction

Regional Seismic Phases

Seismic attenuation is an important parameter describing seismic wave propagation amplitude reductions with distance; however, high attenuation can cause seismic phases to not be observed when the signal amplitude is lower than the noise level. This is how we define phase blockage. It is important to note that the phase blockage does not occur randomly and that the blockage is spatially systematic. Thus, by excluding those blocked paths we are systematically biasing attenuation models, leading to predicted amplitudes being larger than the true values. S_n is a high-frequency (0.5 – 5 Hz) shear wave propagating within the lithospheric mantle lid with a velocity between 4.3 – 4.7 km/s. It is commonly used to investigate the parameters (e.g. shear velocity and attenuation) of lithosphere-asthenosphere boundary (LAB). Although S_n propagates efficiently in stable continental and shield regions (Ni and

Barazangi, 1983; Gök et al., 2000; Sandvol et al., 2001; Gök et al., 2003) and has been observed at distances up to 35° (Molnar and Oliver, 1969; Huestis et al., 1973), it is usually partially or totally blocked when traveling through tectonically active regions with high heat flow (Molnar and Oliver, 1969; Kadinsky-Cade et al., 1981; Ni and Barazangi, 1983; McNamara and Owens, 1995; Gök et al., 2000; Calvert et al., 2000; Sandvol et al., 2001; Al-Damegh et al., 2004). Sandvol et al. (2001) proposed a method to tomographically image the regions of efficient and blocked wave propagation (L_g and S_n) using collected data from the ME using a traditional Sparse Equations and Least Squares (LSQR, Paige and Saunders, 1982) method to solve the linear system of equations. Their results show consistent features between efficiency tomography and geological structures. This method has become a relatively standard approach (Al-Damegh et al., 2004) to objectively map S_n blockage zones. Tibshirani (1996) proposed a technique for regularizing linear regression estimates called the Lasso or "least absolute shrinkage and selection operator", wherein it shrinks some of the coefficient estimates and sets others to zero. This method aims to improve the ordinary least squares (OLS) estimates by combining the best features of both subset selection and ridge regression.

In this study, we first apply the efficiency tomography method on the visually picked data to study the geological structures in the ME and East Asia. Then, motivated by Tibshirani (1996); Park and Levin (2002), and Sandvol et al. (2001), we develop a logistic regression model where the parameter coefficients are regularised by a Bayesian lasso approach to predict the probability of regional phase blockage. The effectiveness of our proposed method is evaluate using both simulated and real data. Our results demonstrate that our new approach allows us to not only predict S_n propagation with high probability but also to accurately predict the phase propagation efficiency.

Tectonic Setting

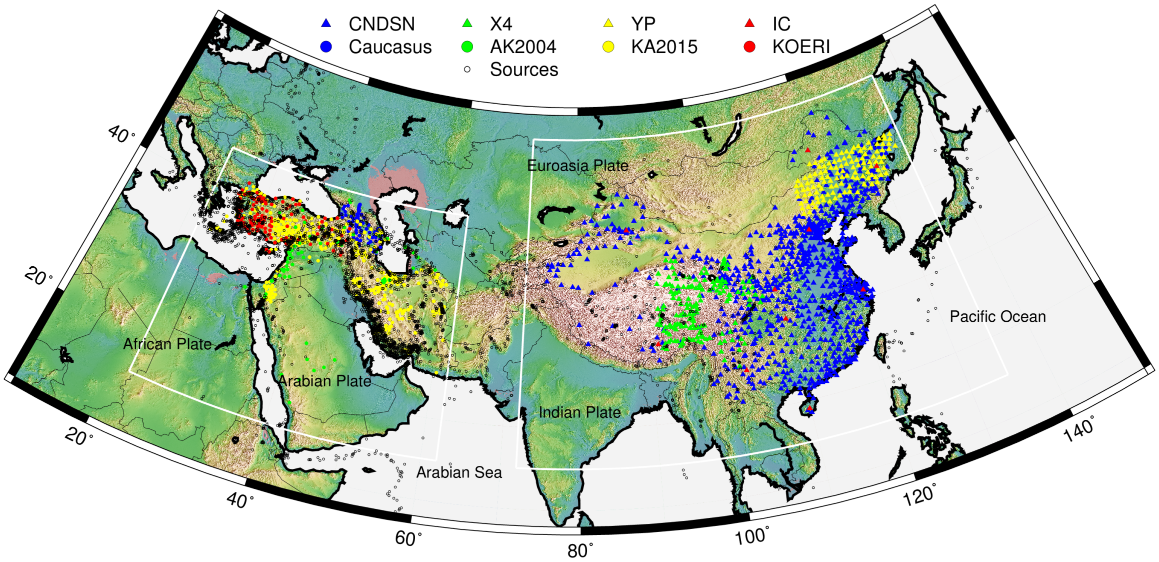


Figure 5.1: A topography map of study areas and the distribution of stations and events.

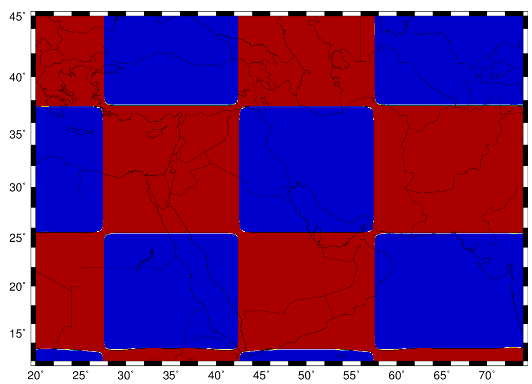
This study focuses on two areas, the ME and Eastern Asia (Figure 5.1). Since there is no overlap of our two study regions, we have processed our data separately.

The ME region is tectonically and seismically active and includes tectonic regions with continental break-up, collision, back-arc extension, and westward escape tectonics. This region has been the subject of extensive geological and geophysical studies during the past several decades. Blockage of the regional phases is widespread along the Bitlis-Zagros and the eastern and southern Anatolian plateau. The subduction of Neo-Tethyan lithosphere beneath Eurasia, the geometry of the subducting slab, the timing of the eventual continental collisions, and the occurrence of possible slab break-off all varied along the strike of the southern edge of the Eurasian plate. Geological evidence and tectonic reconstructions suggest that the northward subduction of the Neo-Tethyan ocean beneath the southern margin of Eurasia has initialed in the Early Jurassic to the Early Cretaceous (Agard et al., 2011; Richards, 2015). Hafkenscheid et al. (2006) suggest that the early slab break-off first occurred beneath the northern Zagros suture zone in the early Oligocene, followed by both eastward and westward propagation of the slab tear. This has lead to a very thin to absent mantle lithosphere beneath much of the Iranian and Anatolian plateau. Much of the Arabian plate consists of a Precambrian shield bounded by a sedimentary platform, Arabian plate behaves as a rigid plate moving NE to NNE, producing spreading in the Red Sea and Gulf of Aden and collision against Eurasian plate along the Bitlis-Zagros during the middle to late Pliocene (Phillip et al., 1989). While convergence in Iran is accommodated by distributed horizontal shortening, the Anatolian plate moves westward at present with an average velocity of $\sim$ 2cm/yr (Reilinger et al., 2006). As a result of the young continental collision, the Anatolian-Iranian plateau and Zagros mountains formed. Most of the Lesser and Greater Caucasus are believed to have also formed within the same time frame as the Arabian-Eurasian collision (Phillip et al., 1989). The Anatolian block is escaping to the west, as evidenced by the right-lateral strike-slip movement along the North Anatolian fault system and measured by Global Positioning System (GPS) data (Ahadov and Jin, 2017). Faccenna et al. (2013) suggest that the progressive evolution of the Tethyan convection result in back-arc extension in the Aegean.

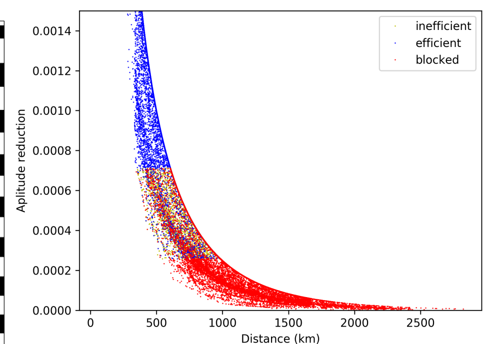
The East Asia region has a complex tectonic history with the collision between the Indian and Eurasian plate, subduction of Pacific plate beneath Eurasian plate and complex topography change from Tibet to Mariana Trench. To the west, the Tibetan plateau is one of the largest active continental-continental collisions features on Earth and is the result of the collision between the Indian and Eurasian plate. The convergence of the Indian plate and Eurasia plate started since $\sim$ 50 Ma and Indian plate is continuously subducting beneath Tibet (e.g. Yin and Harrison, 2000; Chen et al., 2017). Other than the Tibetan plateau, the northward collision of the Indian plate with the Eurasian plate led to the formation of the Himalaya and Tien Shan mountain belt and crustal shortening and uplifting of the Tibetan plateau, accompanied with eastward extrusion of northeastern portions of the plateau (Yin and

Harrison, 2000; Wang and Shen, 2020). At the norther edge of our study area, we find regional deformation related that is likely related to the India-Eurasia convergence. Mongolia has been regarded as one of the most tectonically active intracontinental regions in the world (Choi et al., 2018), although compared to mainland China, GPS data show very small surface deformation ($\sim$ 4 mm/yr, Wang and Shen, 2020). At the eastern margin of the Eurasia plate, the Pacific and Philippine Sea plates are subducing beneath the Eurasian plate, causing reactivation of the Sino-Korean craton including active volcanism and the Shanxi rift (Wang and Shen, 2020).

5.2 Data



(a) Checkerboard model for simulation.



(b) Amplitude reduction plot with distance.

Figure 5.2: The model of simulation data.

In order to evaluate the Propagation tomography approach we have simulated datasets for both the ME and the Eastern Asia. We conducted the simulations using the same geometry (ME and Eastern Asia ray coverage) with different simulated attenuation structure.

The technique we use to generate simulated data is to set the efficiency level based on the amplitude reduction value. The amplitude reduction is determined by

$$A^* = d^{-m} \times \exp\left(\frac{-\pi f d}{vQ}\right). \tag{5.1}$$

where A^* denotes the amplitude reduction, d is distance the ray propagates and m is geometrical spreading parameter, which is assumed to be 1 for S_n. f, v and Q are frequency, S_n group velocity, and quality factor, respectively. We model blockage at 1 Hz signal and assume the S_n group velocity is 4.5 km/s. We obtain the amplitude reduction values for all ray paths for various attenuation models (Figure 5.2a), then classify the efficient and blocked paths based on them (Figure 5.2b). In this study, we set the rays with top 30% A^* as efficient, the middle 30% as inefficient and the bottom 40% as blocked. We then added noise to simulated efficiency data by converting 25% of the inefficient rays to efficient and 25% of the inefficient rays to blocked randomly (Figures 5.2b and 5.3).

All the events used in this study are with magnitude greater than 4.5 and hypocenters located within the crust (black circles in Figure 5.1). The epicentral distances are between 3° and 20°. The whole dataset used in this study includes 30612 ray paths for the ME and 33366 ray path for East Asia. In Figure 5.1, the stations labeled as blue dots are from Caucasus Array, including network AB (35 stations), CW (15 stations), GO (7 stations), and XA (50 stations). The earthquake data are collected from 211 events recorded between June 2017 to December 2018. The stations shown as green points in Figure 5.1 includes 195 stations and data from 1985 events recorded 1978-1979 (102 events), 1992-1999 (1460 events). For more information about this dataset see Al-Damegh et al. (2004). The stations labeled as red points are stations from the Kandilli Observatory And Earthquake Research Institute (KOERI) between 2006 and 2008 (452 events). The stations labeled as yellow dots include 556 stations along with the corresponding data recorded 1995-2014 (1254 events). For more information

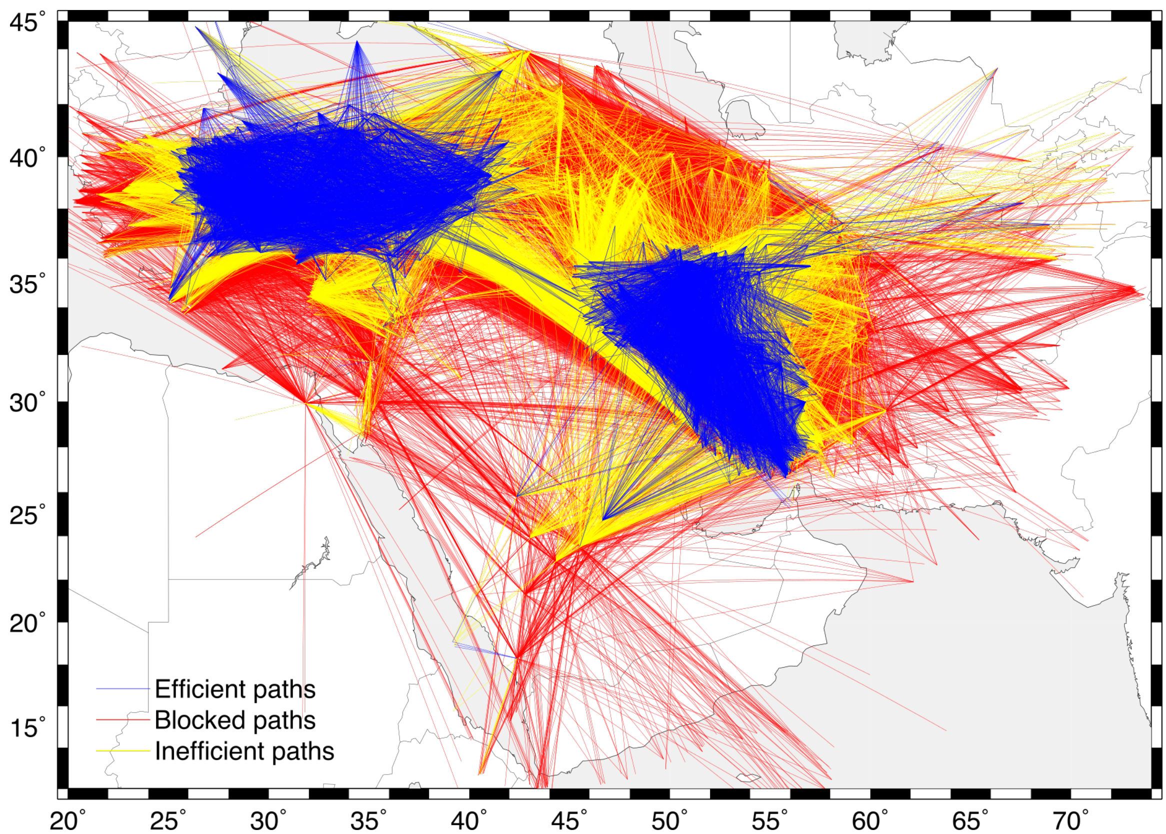


Figure 5.3: Simulation data plotted by colored rays.

about this dataset see Sandvol et al. (2001) and Kaviani et al. (2015).

The network code X4 includes 167 broadband stations covering nearly all of eastern Tibet from 5 seismic networks: Indepth IV, Namche Barwa, MIT-China, NETS, and ASCENT. The corresponding events were recorded between 2003 and 2009. The network CNDSN includes 891 stations covering the whole China mainland. The corresponding events occurred between 2009 and 2011. For more information about this dataset see Chen and Niu (2016). We also use 10 stations from network code IC and 120 stations from network code YP. YP covers northeastern China, while IC station is spread out over the China mainland. We also included data from 484 regional events (epicentral distance between 3° and 20°) recorded by 1188 stations. All the event magnitudes are greater than 4.5 and all the hypocenters are with crust. In total there are 33,366 source-receiver pair used in this study.

5.3 Methods

S_n Efficiency Tomography

Sandvol et al. (2001) proposed methods to tomographically image L_g and S_n propagation on the collected data in the ME. Here we only use the method to image S_n propagation. Following the derivation of Phillip et al. (1989) and the starting model of Cong et al. (1996), they assumed (without error) that S_n phase amplitude (a_{ij}) is

$$a_{ij}(f) = a_{0i} x_{ij}^{-m} s_j(f) c_i(f) \exp\left[-\alpha(f) x_{ij}\right], \tag{5.2}$$

where i and j are the event and station indices respectively, f is the phase frequency, a_{0i} is the amplitude of the i-th source, x_{ij} is the total ray-path length, m is the geometrical spreading parameter for S_n (assumed to be 1), s_j is the station response for the j-th station, c_i is the source scaling term for the i-th seismic source, and α

which the spatial average attenuation coefficient and is assumed to be constant as $\pi f x_{ij}/V$ over a given frequency band. V is S_n group velocity and assumed to be 4.5 km/s. Taking the logarithm of (5.2), correcting the station response term, discretizing the spatial attenuation factor, and ignoring the source scaling term, they obtained

$$\tilde{A}_{ij}^{disc.} = \log\left(\frac{a_{0i}}{a_{ij}}\right) = \log e \frac{\pi f}{V} \sum_{l} m x_{ijl}, \tag{5.3}$$

where x_{ijl} represents the ray-path length corresponding to the i-th source, j-th station and l-th mesh block. Following Sandvol et al. (2001), we partitioned the S_n propagation efficiencies into three categories: no S_n, efficient S_n, and inefficient S_n. If the seismogram showed no evidence of a discernable S_n phase, we categorized that path as a blocked S_n path and $\tilde{A}_{ij}^{disc.}$ in (5.3) is set to be zero. If some S_n waves could be observed, regardless of its strength or amplitude, we designated it as an efficient S_n path and $\tilde{A}_{ij}^{disc.}$ in (5.3) is set to be two. If there were some ambiguous signal in the seismogram that potentially could be an S_n phase, it was classified as an inefficient S_n path and $\tilde{A}_{ij}^{disc.}$ in (5.3) is set to be one. After trying different automated signal to noise ratios using pre-S_n noise, we found that the optimal method to set the efficiency level is to visually pick the seismogram manually. Using these definitions, the model parameter m, in (5.3), becomes the average phase efficiency for paths crossing through that tomographic cell. Since we assume that there is no attenuation, the extinction path length is maximum for areas with efficient S_n and minimum for areas with blocked or no S_n. Solving the linear equation (5.3) using LSQR allows us to objectively and quantitatively map the S_n efficiencies.

The Bayesian Lasso Propagation Probability

We use a logistic regression/Bayesian lasso model to predict the likelihood of observing S_n based on efficiency datasets. Specifically, suppose that we have N seismic rays discretized into p sections. Define $\mathbf{X} = (X_{ij1}^T, \ldots, X_{ijp}^T)$ as a N x p design matrix corresponding to the discretized sub-distances X_{ijl} $(l = 1, \ldots, p)$ as defined in the previous section and let z_i be the response variable, assumed to have two possible outcomes, 0 if ray i is observed and 1 if it is blocked. It is important to note that some of the sub-distances can be equal to zero for a given ray path. Then if θ is the probability of being observed, we propose a simple Binomial logistic regression model for $\mathbf{z}$ and define it as

$$\begin{aligned} \mathbf{z}|\theta &\sim Bernoulli(\theta) \\ \text{logit}(\theta) &= \mathbf{X}\boldsymbol{\beta} \end{aligned} \tag{5.4}$$

where the logit or log-odds transformation is defined as $\text{logit}(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$ and $\boldsymbol{\beta} \in \mathbb{R}^p$ is the vector of unknown regression coefficients. Since N is typically large, $\mathbf{X}$ is a sparse matrix and thus if $rank(X) < p$ (e.g., this can happen when $p > N$), there are infinitely many solutions under an ordinary least squares (OLS) approach. Even if $rank(X) = p$, for a large p, the OLS estimates will have a lot of variability, resulting in overfitting and consequently poor predictions for future observations not used in model training. Furthermore, it is often the case that some or many of the variables used in a multiple regression setting like this are in fact not all strongly associated with the response. Including such variables leads to unnecessary computational complexity in the model. By removing these variables, say, by setting the corresponding coefficient estimates to zero we can obtain a model that is much easily interpreted. One way to deal with these issues is shrinkage or regularization which allows us to substantially reduce the variance at the cost of a negligible increase in bias. Here we adopt a Lasso regularization as introduced by Tibshirani (1996). Defined as the stricter l_1-penalty, the Lasso approach can set coefficients to exactly zero, making it a useful tool for feature selection with lower variability. For the model in (5.4), the

Lasso estimates of the regression coefficients, $\hat{\boldsymbol{\beta}}$ are defined as

$$arg\ min \left\{ \sum_{i=1}^{N} \left[\mathbf{z}_i - \text{logit}^{-1} \left(\sum_{l=1}^{p} \beta_l X_{ijl} \right) \right]^2 \right\} \qquad \text{subject to } \sum_{l} |\beta_l| \leq t. \tag{5.5}$$

where $t \geq 0$ is a tuning parameter. The parameter t controls the amount of shrinkage that is applied to the estimates. The Lasso constraint $\sum_l |\beta_l| \leq t$ is equivalent to the addition of a Lagrangian penalty $\lambda \sum_l |\beta_l|$ to the residual sum of squares (Murray and Overton, 1981). Now $|\beta_l|$ is proportional to the negative log-density of the Laplace distribution and hence we can derive the Lasso estimate as the Bayesian posterior mode under independent double-exponential priors for the β_ls of the form

$$\pi(\beta) = \prod_{j=1}^{p} \frac{\lambda}{2} e^{-\lambda|\beta_j|}. \tag{5.6}$$

Park and Casella (2008) developed a Gibbs sampler implementation of a fully Bayesian adaptation of the Lasso regularization exploiting the representation of the Laplace distribution as a scale mixture of Normal distributions (with an exponential mixing density) as

$$\frac{a}{2} e^{-a|y|} = \int_0^\infty \frac{1}{\sqrt{2\pi s}} e^{-y^2/2s} \frac{a^2}{2} e^{-a^2 s/2} ds, \qquad a > 0.$$

Bae and Mallick (2004) assumed independent Laplace priors on β of the form in (5.6) to induce sparseness while Park and Casella (2008) assumed a conditional Laplace priors on β to ensure unimodal posterior distributions. We adopt both their approaches and define the following Binomial logistic regression model with Bayesian

Lasso regularization to predict the likelihood surface

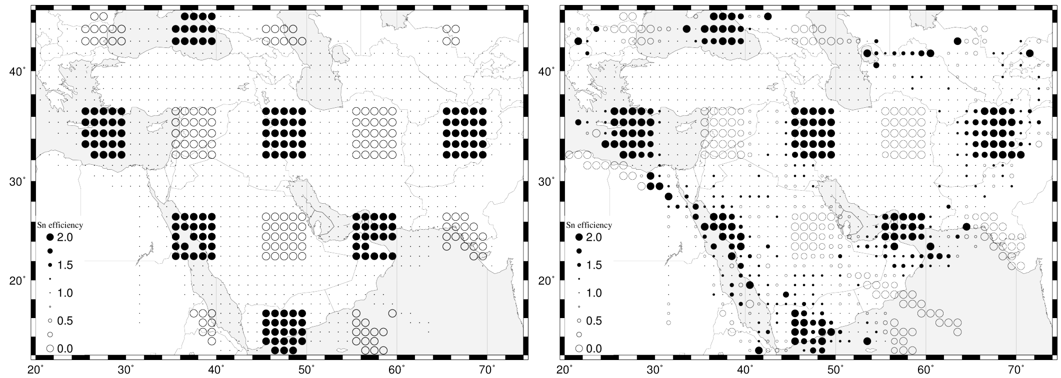
$$\begin{aligned}
\mathbf{z}|\boldsymbol{\theta} &\sim Bernoulli(\theta) \\
\text{logit}(\theta) &= \mathbf{X}\boldsymbol{\beta} \\
\boldsymbol{\beta}|\mathbf{D}_{\boldsymbol{\tau}} &\sim \mathbf{N}_p^+(\mathbf{0_p}, \mathbf{D}_\tau) \\
\mathbf{D}_\tau &= \text{diag}(\tau_1^2, \ldots, \tau_p^2) \\
\tau_l^2 &\sim Exponential(\lambda^2) \qquad l = 1, \ldots, p,
\end{aligned} \tag{5.7}$$

where $\mathbf{N}^+$ refers to the Truncated Gaussian Distribution, truncated below by 0. This assumption is based on physical properties of the elements of $\mathbf{X}$. λ is the Lasso tuning parameter. We discuss methods to determine this parameter in the Appendix. Given the prior distribution assumptions, we run into the issue of conjugacy due to the analytically inconvenient form of the model's likelihood function. Hence we adopt a data augmentation strategy and further details about the Gibbs sampler implementation of our model has been discussed in Appendix.

5.4 Results

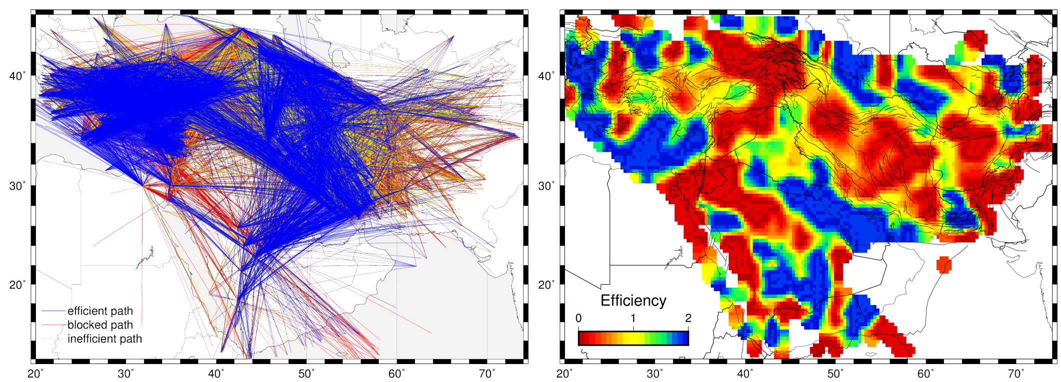
5.4.1 Efficiency tomography

Figures 5.5 and 5.7 show the efficiency data and tomography in the ME and Eastern Asia, respectively. To check our data resolution, we have created a checkerboard model with 1° mesh size and 5° anomaly. As is commonly done in seismic tomography we have created a synthetic dataset using a hypothetical Q model with alternating low and high attenuation zones which can be used to generate synthetic Sn efficiencies. We have used these synthetic efficiencies to test how well we can resolve high and low attenuation zones. Since the ray coverage is very good for most of our model, we can resolve the study area reasonably well with some smearing effects at the edges of the



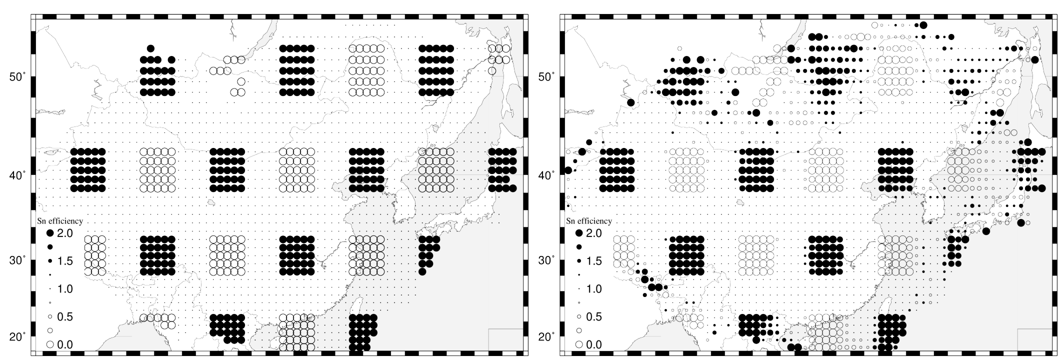
(a) The checkerboard model. (b) The checkerboard test resolution.

Figure 5.4: The ME data resolution.



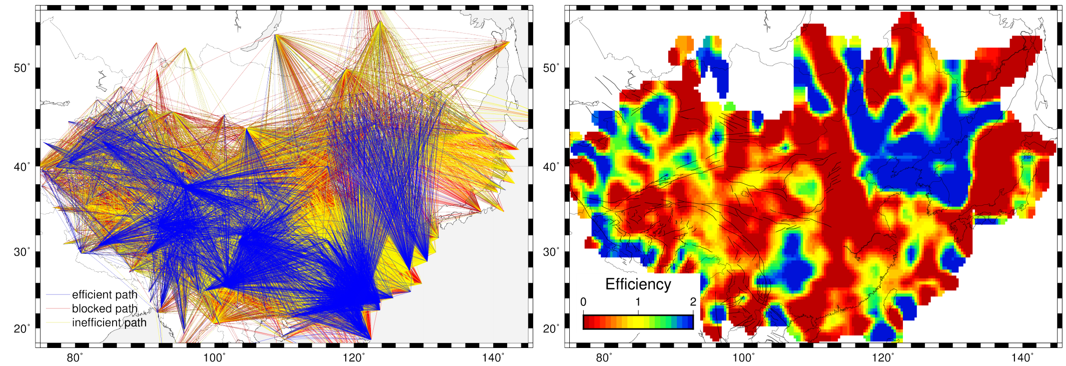
(a) The ME efficiency data. (b) The ME efficiency tomography.

Figure 5.5: The ME efficiency tomography.



(a) The checkerboard model. (b) The checkerboard test resolution.

Figure 5.6: The East Asia data resolution.



(a) The East Asia efficiency data. (b) The East Asia efficiency tomography.

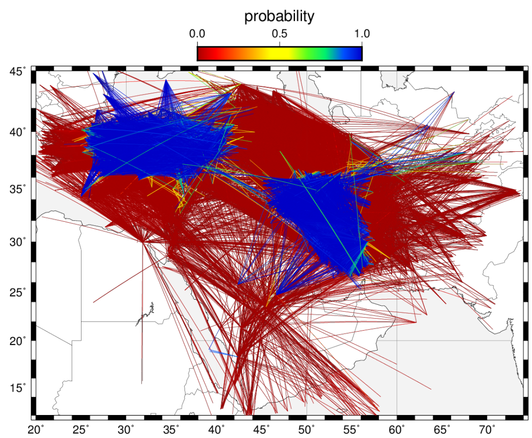
Figure 5.7: The East Asia efficiency tomography.

model. We find evidence of poorly resolved synthetic anomalies along the Red Sea, at the southern edge of the Arabian Plate and at the southern edge of the Iranian Plateau. The model is not well resolved for portions of the Caucasus. Using the East Asia data, we resolve all of the China mainland as well. The northwestern Pacific and Mongolian portions of the model are not resolved well enough for the 5° anomaly.

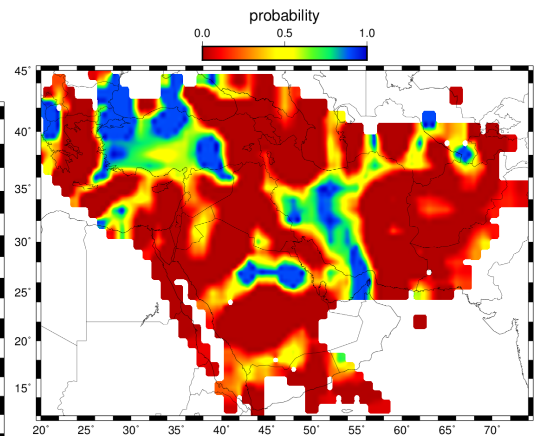
Figures 5.5a and 5.7a show the efficiency data used in this study. We manually picked all the S_n data and set the efficiency level to 2 for efficient Sn, 1 for inefficient Sn and 0 for blocked Sn phases. The efficient ray-paths are plotted as blue, inefficient ray-paths as yellow and blocked ray-paths as red. We have used an efficiency tomography method to map variations in propagation efficiency (Figures 5.5b and 5.7b). We observed blocked S_n across most of the Iranian and northeastern Anatolian Plateau. We also found an S_n blockage zone at the eastern edge of the Mediterranean sea. These observations are consistent with previous studies (Sandvol et al., 2001) and (Al-Damegh et al., 2004). The S_n blockage zone is also found along the Red Sea (Al-Damegh et al., 2004). Further, we identified a blocked S_n zone at western edge of the Arabian Plate and inefficient S_n zone in the middle of Anatolian Plateau, which is not observed by Sandvol et al. (2001); Al-Damegh et al. (2004). We observe efficient S_n across most of the Arabian Plate, the Mediterranean Sea, and Caspian Sea, which is consistent with Al-Damegh et al. (2004).

We found large efficient S_n zone across the Northwestern Pacific basin, which is consistent with the studies of Molnar and Oliver (1969) and Rapine et al. (1997). We also observe efficient S_n for paths within the Ordos, Sichuan and Songliao basins. The inefficient S_n is also observed across the China East Sea (Molnar and Oliver, 1969).

5.4.2 Predicted probability tomography

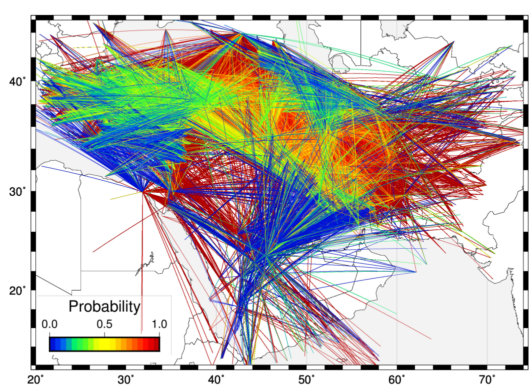


(a) The predicted probability of rays.

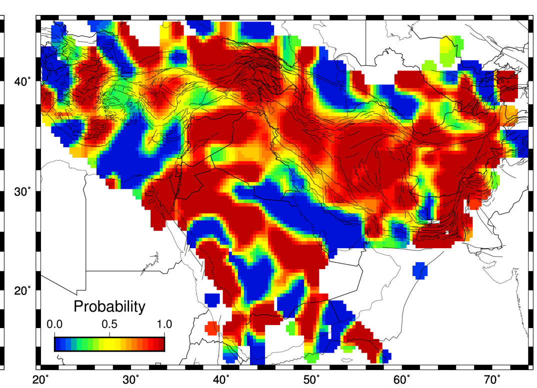


(b) The predicted probability tomography.

Figure 5.8: Predicted probability results of simulated data.



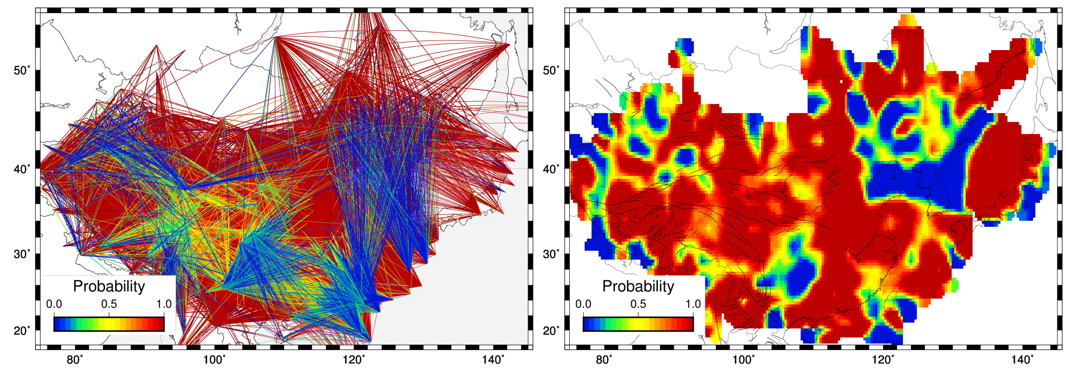
(a) Predicted probability rays.



(b) Predicted probability tomography.

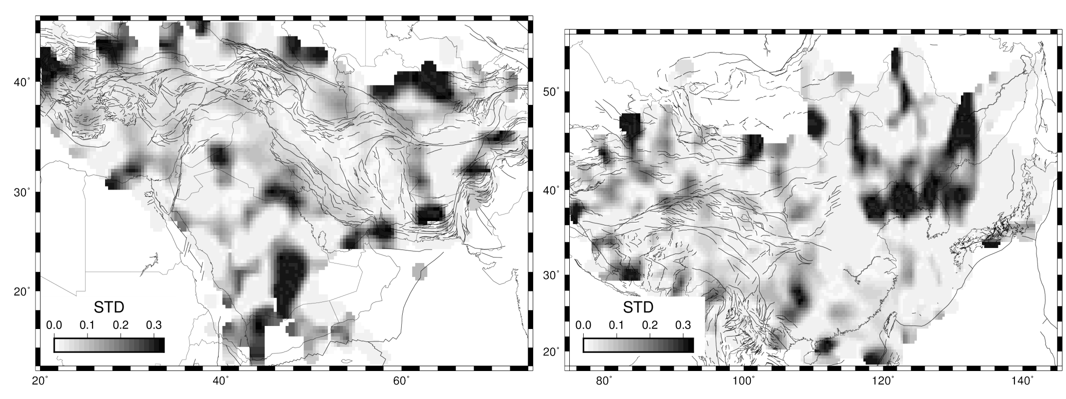
Figure 5.9: Predicted probability of ME data.

Note that we discard all the inefficient rays when using predicted probability tomography method since our method is based on logit Bayesian method. We evaluate



(a) Predicted probability rays. (b) Predicted probability tomography.

Figure 5.10: Predicted probability of the East Asia data.



(a) ME data. (b) The East Asia data.

Figure 5.11: STD error of Bayesian Lasso method.

our method on simulated data derived from a hypothetical attenuation structure (Figure 5.2). The results of these simulations show that our method is able to correctly predict the probability of being blocked for each ray (Figure 5.8a) with high accuracy (83.48%, Table 5.1). The area under the curve (AUC) metric is also used to asses the accuracy of our predictions. The AUC results are largely consistent with our prediction accuracies. The tomographic results (Figure 5.8b) for the simulated data is not similar to the checkerboard model (Figure 5.2a), since the ray-path coverage is poor in this part of the study area..

Figures 5.9 and 5.10 are the blockage probability results of the ME and East Asia, respectively. We predict a high probability of blockage for the low inefficient S_n zones with high accuracy (Table 5.1). We also can predict the blockage probability of each ray with high accuracy (79.77%, Table 5.1) for the ME data set (Figure 5.9a). We observe the high probability of S_n blockage for most of the Iranian Plateau, the eastern Anatolian Plateau, Dead Sea Fault Zone, and along the Red Sea (Figure 5.9b). The low probability of blockage is found across most of the Arabian Plate, the Mediterranean Sea, and the Caspian Sea. All of these regions have primarily efficient S_n propagation (Sandvol et al., 2001; Al-Damegh et al., 2004). We also estimate the standard deviation (STD) of the predicted probability tomography results (Figure 5.18). The STD is lower than 0.33 for all of the study area. The regions with large STD are at the edge of the model where we do not have good ray coverage (Figure 5.11a).

The predicted blockage probability of each ray with high accuracy (86.67%, Table 5.1) for the East Asia data (Figure 5.10a). The lower blockage probability is found within Sichuan and Ordos basins as well as the Northwest Pacific Basin. The large STD is also found for the Northwestern Pacific Basin (Figure 5.11b).

5.5 Discussion

5.5.1 Efficiency tomography

Our Sn efficiency tomography model for the Middle East (Figure 5.5b) clearly imaged the major tectonic across the ME, although some of the anomalies are likely distorted especially within the Arabian plate. From the S_n efficiency data, we observed consistent efficient S_n propagation in the Arabian plate, the Mediterranean, and the Caspian Seas. We have also observed efficient S_n propagation for paths across western Iraq and northwestern Arabia; these anomalies reflect the portion of the Arabian plate that is tectonically stable with no volcanism and is known to have a stable lithospheric mantle. We observed inefficient S_n propagation in nearly all of the Anatolian Iranian plateaus. This is a well-known blockage zone that corresponds with low uppermost mantle seismic velocities and Miocene volcanism (e.g. Keskin, 2003). We observed blocked S_n propagation along the Dead Sea Fault Zone, northwestern Anatolian plateau, and most parts of the Iranian plateau. In the Mesopotamium Foredeep, we observe an efficient S_n region, which is consistent with the results of Molnar and Oliver (1969). A number of studies have found efficient S_n propagation in the Zagros mountains (e.g., see Kadinsky-Cade et al., 1981) and S_n blockage in the Anatolian and Tibetan Plateaus (Molnar and Oliver, 1969). We have mapped the efficient S_n region for paths that cross the Caspian Sea, Black Sea and Mediterranean Sea. This is consistent with efficient S_n propagation within the oceanic lithosphere. The Arabian plate and Gulf of Oman also have efficient S_n transmission due to the presence of a stable continental lithospheric mantle. These observations are consistent with all of the prior results of Kadinsky-Cade et al. (1981); Rodgers et al. (1997); Sandvol et al. (2003); Al-Damegh et al. (2004). Rodgers et al. (1997) observe inefficient S_n zone restricted to the northern Turkish plateau, the Lesser Caucasus, the western Greater Caucasus, the northern Iranian plateau, and the Dead Sea Fault system. Our image of S_n propagation in the Caucasus region is mostly new and it appears there is evidence of S_n blockage across the western two thirds of the Greater Caucasus. This would suggest that the western Greater Caucasus lithospheric mantle is fundamentally different than the easternmost Greater Caucasus. Furthermore, our observation shows that the inefficient S_n zone extends to the southern part of Zagros mountain and Anatolian plateau. We observe efficient S_n transmission across most of eastern portion of the Arabian plate. S_n phases are blocked for paths along the Dead Sea Fault system, which is consistent with the low seismic velocities seen in a wide range of studies of (e.g. Mellors et al., 1999; Al-Lazki et al., 2003; Simmons et al., 2011; Kaviani et al., 2020). This broad zone of blocked S_n along nearly the entire western edge of the Arabian plate is consistent with widespread Cenozoic volcanism that may be related to an asthenospheric up-welling.

Our East Asia S_n efficiency tomography model (Figure 5.7b) has imaged most of the major tectonic structures in the East Asia. We observe high S_n efficiency in Ordos plateau, Sichuan basin, Songliao basin, Qiadam basin, and Tarim basin. We observe large S_n efficient area at Northwest Pacific basin, from NE China to the Korean peninsula. All of these are tectonically stable continental lithospheric blocks with very little internal deformation and almost completely aseismic (Wang and Shen, 2020). We would expect efficient S_n propagation. Besides these regions, one of the most striking aspects of our model is how widespread S_n blockage is throughout much of East Asia when compared with the ME. It is worth noting that our efficiency database is focused on high frequency S_n not lower frequency S_n that can be observed consistently in some regions like the northern Tibetan plateau (e.g. Barron and Priestley, 2009). The regions with the most consistent high frequency S_n blockage areas are observed in Tibetan plateau, Yunnan providence, Shanxi rift, and eastern China. One of the most surprising results is the spatially extensive S_n blockage zone associated with the Shanxi rift zone. The blockage zone appears to extend well beyond the margins of the rift zone suggesting that perhaps strong scattering attenuation also contributes to S_n blockage in this part of northern China. The Sea of Japan is another S_n blockage zone. This observation is consistent with the prior studies (Molnar and Oliver, 1969). Like Molnar and Oliver (1969), we also observe inefficient S_n in the China East Sea. To evaluate the quality of our efficiency data, we developed a method to check the instability of the efficiency data (see Appendix).

5.5.2 Predicted Probability Tomography

Figure 5.18 shows the 95% credible interval of the estimates from our Bayesian Logistic model for the ME and East Asia models. In most of the study areas, the standard errors are quite low, with some pockets exhibiting higher standard errors. It is interesting to note that the higher error regions do not necessarily correlate well with the our estimates of the noisiest data. This suggests that the larger errors are likely due to complexity in the seismic attenuation structure. Figure 5.9b shows the predicted probability of S_n propagation across the ME. We observe high S_n propagation probabilities in almost all S_n efficient regions (Figure 5.5b). We see some probabilities of approximately of observing S_n in areas within the Arabian platform, the middle of Turkmenistan, southern Caspian Sea and the Mediterranean Sea. Almost all of these regions are regions with oceanic lithosphere or a continental stable platform. We see lower probability of observing S_n in western Iraq and southeastern Iranian plateau. The predicted probability results of all these regions are more than that of the efficiency tomography. We see a probability of .4 in observing S_n propagation in Turkey and the Anatolian plateau, that is consistent with the inefficient propagation we observed in these regions. Similar to our efficiency tomography model, we see very low probabilities ($< .1$) of observing S_n in northeastern and southeastern Turkey. The Dead Sea Fault system also has a very low probability of S_n being observed (i.e. this is a S_n blockage zone). This is a well known S_n blockage zone (Sandvol et al., 2001) and is consistent with presence of Quaternary volcanism. We see more continuous blocked zones beneath the Iranian plateau and in Baluchistan. The SW-NE oriented blocked zone in northwestern Arabia is not shown in the S_n probability model. This blockage zone is not shown in Al-Damegh et al. (2004). We do not see the attenuation block in the Aegean Sea in our S_n probability model. Our predicted probability models of S_n propagation agree with the prior work of Sandvol et al. (2001) and Al-Damegh et al. (2004). From the S_n propagation observations, we see high probability of observing S_n for the Sichuan basin, Tarim basin, Turpan basin, Dzungaria basin, Ordos basin, and Sino-Korean craton. All these regions indicate stable continental blocks. We also see high probability of blockage zones (less than .10 probability of being observed) in the Tibetan plateau and eastern China.

Another important potential application of our model for the probability of blockages is the application to earthquake bulletins and the reduction of mis-identification of secondary regional phases. Regional phase identification can be difficult due to blockage. Our model can be used to predict the probability of observing a particular phase for a given path. This can be used to filter out potentially misidentified secondary phase arrivals for any given network's earthquake bulletin. This model can also be used to guide both manual and automated phase identification. In future work we hope to use these types of models to be able to automatically identify and process regional waveforms.

Table 5.1: Prediction Accuracy and confustion matrices.

Model	**Accuracy**	**AUC**	**Confusion Matrices**			
			T0P0	T1P0	T0P1	T1P1
Simulation	83.48 %	0.83	0.83	0.16	0.17	0.84
ME	79.77%	0.80	0.70	0.13	0.30	0.87
East Asia	86.67%	0.82	0.58	0.05	0.42	0.95

5.6 Conclusion

In this study, we developed a Bayesian Lasso model to adapt the logit tomography to the likelihood of S_n observation, which is incorporated into a spatial probability model that will estimate the probability of S_n being observed. Since our model is only for binomial case, we have to delete the inefficient paths or treat them as inefficient or blocked to make the data set as binary. However, like we showed in this study, the way we treat the inefficient paths would matter for the accuracy of the prediction. In future, we will extend this to work on multinomial logistic models to deal with data having more than two categories.

Appendix

Stability of the Data

In the real data, we often observe two rays close to each other have different efficiency. We often see rays traveling along similar path, however, with different efficiency. To evaluate the consistency of our data, we developed a method to determine the instability of the efficiency data. We discretize the study area into cells with uniform attenuation ($1/Q$). If the end points for each ray are in the same cell, we assume the rays travel through a very similar portion of the Earth. For all ray-paths with the same begin-and-end points, suppose the count of blocked, inefficient and efficient path are l, m and n, respectively. Now, suppose that the modal value of the three paths count is l ($l > m, n$), then the instability of the efficiency data in the cell-pairs is determined by $(m + n)/(l + m + n) \times 1.5$. The range of the efficiency instability is from 0.0 to 1.0.

We have estimated the mean instability of the whole dataset (Table 5.2). The mean instability increases with larger mesh size. With a larger mesh, since the number of ray paths passing through a cell pair increases, the probability that a cell-pair will have different efficiencies gets higher. We also plot the instability for each cell-pair using colder colors to indicate high spatial variability in the efficiency data (Figure 5.12 and 5.13). Essentially we have equated instability with the noise in the efficiency data. The instability may be the result of the noise in seismograms or the complexity of the attenuation structure. Overall we find the dataset is stable. The higher efficiency instability is found within the portion of the Iranian and Anatolian Plateaus for the ME data. The instability is quite low within the Mediterranean Sea and the Arabian Plate. The higher efficiency instability is found in paths crossing the eastern Tibetan Plateau and the area of China with several tectonic blocks, such as Ordos, Sichuan Basin, the North China Craton, and the South China Block.

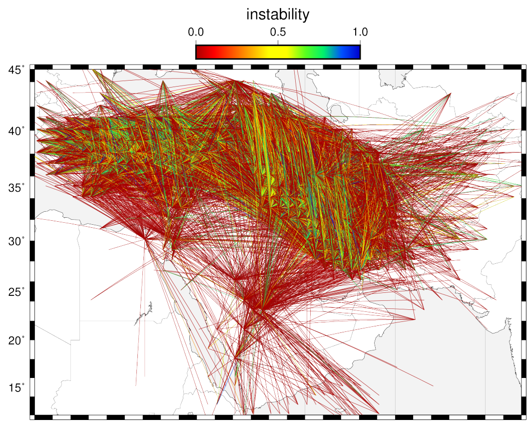
Table 5.2: Mean instability of the data used in this study.

Mesh size (°)	1.0	2.0	3.0	4.0	5.0	6.0
East Asia	0.179046	0.30610	0.360699	0.401719	0.428902	0.451979
ME	0.213896	0.310876	0.365953	0.378041	0.383538	0.413873

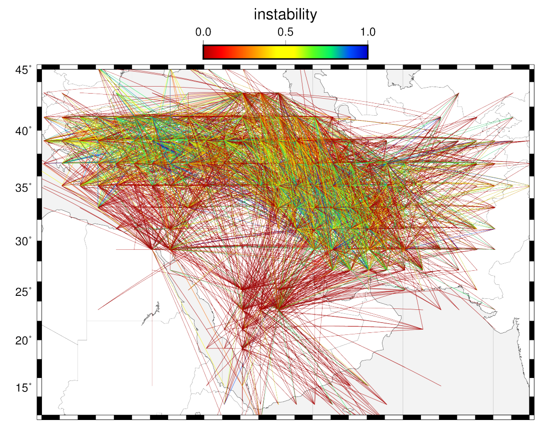
Determining the lasso tuning parameter

Tibshirani (1996) described three methods for the estimation of the lasso parameter λ: cross-validation (see Stone (1974)), generalized cross-validation (see Golub et al. (1979)) and a third method based on Stein's unbiased estimate of risk. James et al. (2013) discusses and illustrates how to use these methods to determine the tuning parameter, using the *glmnet* package in the R software. The Bayesian lasso also offers uniquely Bayesian alternatives: empirical Bayes via marginal maximum likelihood (see Casella (2001), Park and Casella (2008)) or choosing an appropriate hyperprior. Park and Casella (2008) suggests a gamma prior on λ^2 of the form

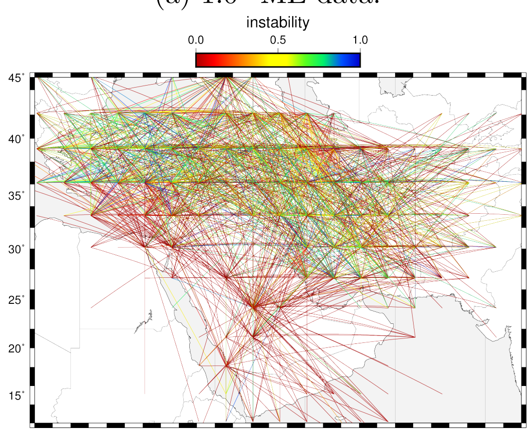
$$\pi(\lambda^2) = \frac{\delta^r}{\Gamma(r)}(\lambda^2)^{r-1}e^{-\delta\lambda^2} \qquad \lambda^2 > 0 \quad (r > 0, \delta > 0). \tag{5.8}$$



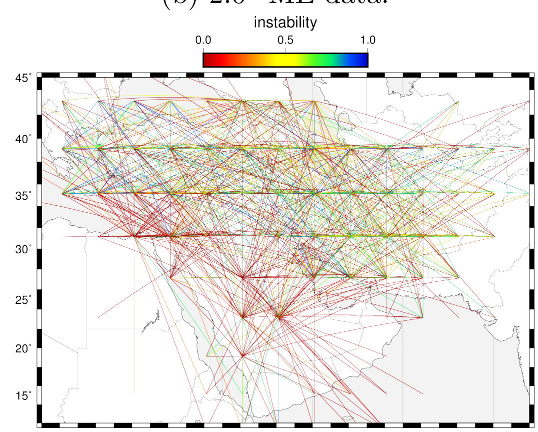
(a) 1.0° ME data.



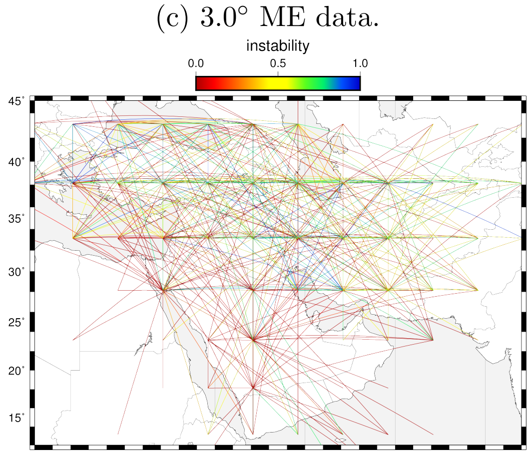
(b) 2.0° ME data.



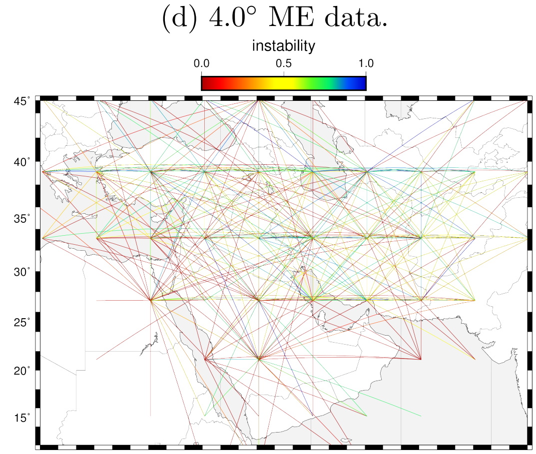
(c) 3.0° ME data.



(d) 4.0° ME data.

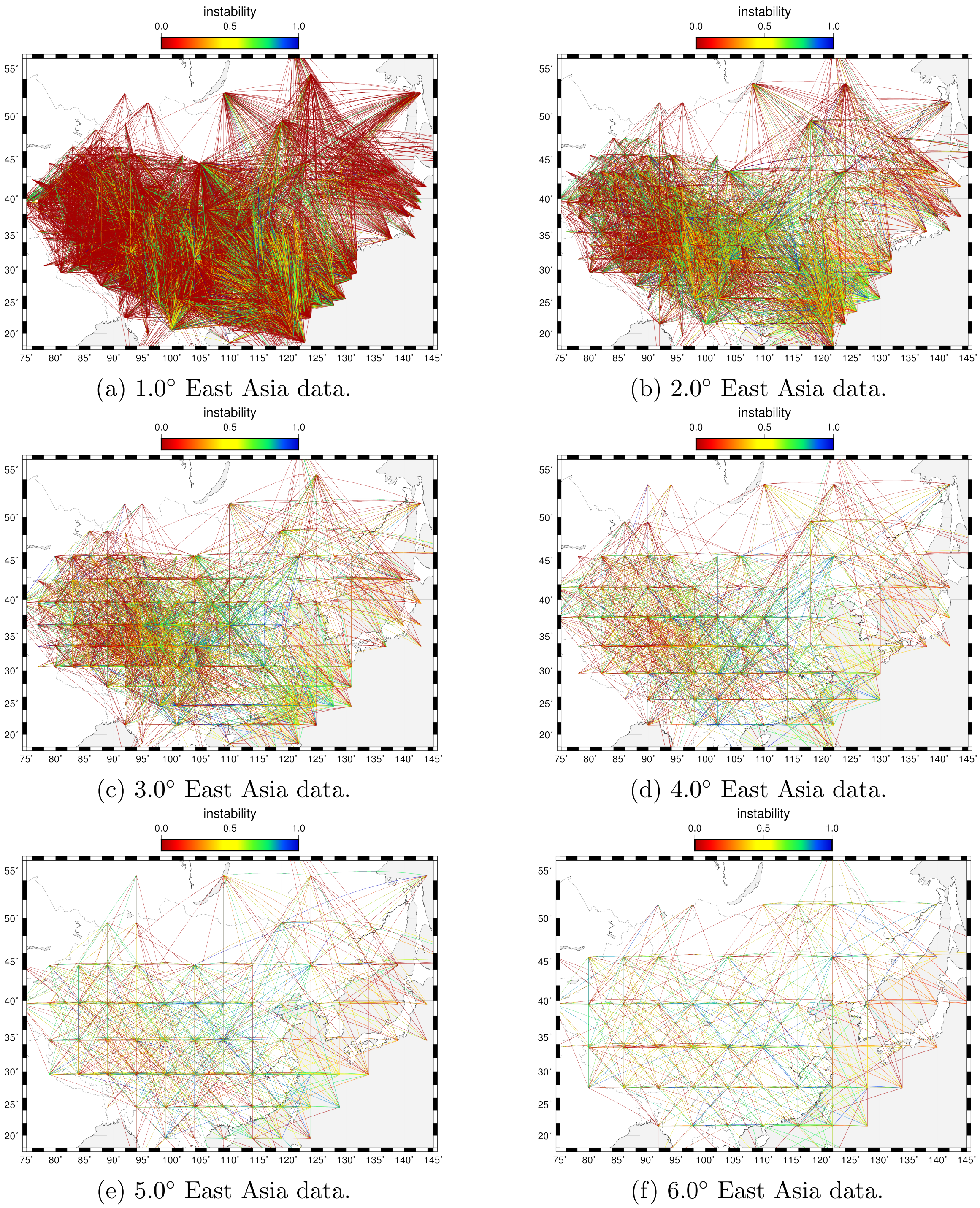


(e) 5.0° ME data.



(f) 6.0° ME data.

Figure 5.12: The instability of the ME data.



(a) 1.0° East Asia data.

(b) 2.0° East Asia data.

(c) 3.0° East Asia data.

(d) 4.0° East Asia data.

(e) 5.0° East Asia data.

(f) 6.0° East Asia data.

Figure 5.13: The instability of Eastern Asia data.

as it helps maintain conjugacy properties in a Gibbs sampler. When this gamma prior is used in the hierarchical structure in (5.7), the product of the factors in the joint density involving the the term λ^2 takes the form,

$$(\lambda^2)^{p+r-1} \exp\left\{-\lambda^2\left(\frac{1}{2}\sum_{l=1}^{p}\tau_l^2+\delta\right)\right\} \tag{5.9}$$

Thus the full conditional distribution of λ^2 is a Gamma distribution with shape parameter $p + r$ and rate parameter . $\frac{1}{2}\sum_{l=1}^{p}\tau_l^2+\delta$. The parameter δ needs to be reasonably larger than zero to avoid computational and conceptual problems, (see Park and Casella (2008)).

The Pólya-Gamma data augmentation

In contrast to the probit model, where data augmentation strategies have been available for some time,(Albert and Chib (1993)), Bayesian data analysis in the logistic model has until recently been computationally difficult due to lack of an analogous algorithm. Polson et al. (2013) developed a real analogue to the Albert and Chib (1993) algorithm for Bayesian logistic regression. The main difference between the two algorithms being the truncated normal distributions in Albert and Chib (1993) algorithm are now replaced by Polya-Gamma random variables. Let $PG(.,.)$ denote a Pólya-Gamma distribution, Polson et al. (2013) showed that binomial likelihoods parametrized by the logodds of success can be represented as mixtures of Gaussian densities with respect to a Pólya-Gamma distribution. The fundamental integral identity at the heart of their approach is that, for $b > 0$,

$$\frac{(\exp\{\psi\})^a}{(1+\exp\{\psi\})^b} = 2^{-b}e^{\kappa\psi}\int_0^\infty e^{-\omega\psi^2/2}p(\omega)d\omega, \tag{5.10}$$

where $\kappa = a - b/2$, $\omega \sim PG(b, 0)$ and $\psi = \mathbf{X}\boldsymbol{\beta}$ is a linear function of predictors. This makes the integrand in the hierarchy in (5.7) the kernel of a Gaussian likelihood in β. This allows a simple strategy for Gibbs sampling : Gaussian draws for the main parameters, and Pólya-Gamma draws for a single step of latent variables. Incorporating this technique in our hierarchical model, we get the following form of the Binomial Logistic regression model with a Bayesian lasso regularization using a Pólya-Gamma data augmentation

$$\begin{aligned} \mathbf{z}|\mathbf{X}, \boldsymbol{\beta}, \sigma^2, \lambda &\sim Bernoulli(\theta) \\ logit(\theta) &= \mathbf{X}\boldsymbol{\beta} \\ \boldsymbol{\omega_i}|\boldsymbol{\beta} &\sim P.G.(n_i, \mathbf{X}_i'\boldsymbol{\beta}) \qquad i = 1, \ldots, N \\ \boldsymbol{\beta}|\boldsymbol{\omega}, \mathbf{D}_{\boldsymbol{\tau}} &\sim N_p^+(m_\omega, V_\omega) \end{aligned} \tag{5.11}$$

where $\mathbf{V}_\omega = (\mathbf{X}'\mathbf{D}_\omega\mathbf{X} + \mathbf{D}_\tau{}^{-1})^{-1}$, $m_\omega = \mathbf{V}_\omega(\mathbf{X}'\boldsymbol{\kappa})$, $\kappa = y_i - n_i/2$, $\mathbf{D}_\tau = diag(\tau_1^2, \ldots, \tau_p^2)$, and $\tau_l^2 \sim Exponential(\lambda^2), l = 1, \ldots, p$.

Gibbs Sampler Implementation

To derive the full conditional distribution of $\boldsymbol{\beta}$ we utilize Theorem 1 in Polson et al. (2013) and write the likelihood contribution of observation i as

$$L_i(\beta) \propto \exp\left(\kappa_i \mathbf{X}\beta\right) \int_0^\infty \exp\left\{-\omega_i(\mathbf{X}\boldsymbol{\beta})^2/2\right\} p(\omega_i|n_i, 0)$$

where $\kappa_i = y_i - n_i/2$, and where $p(\omega_i|n_i, 0)$ is the density of a Pólya-Gamma random variable with parameters $(n_i; 0)$. Assuming a $\mathbf{N}_p^+(\mathbf{0_p}, \mathbf{D}_\tau)$ prior on $\boldsymbol{\beta}$ and the above likelihood, the full conditional distribution of $\boldsymbol{\beta}$ takes the form,

$$\begin{aligned} p(\beta|\omega, y) &\propto p(\beta) \prod_{i=1}^{N} L_i(\beta|\omega_i) \\ &\propto \exp\left\{-\frac{1}{2}\boldsymbol{\beta}^T \mathbf{D}_\tau^{-1}\boldsymbol{\beta}\right\} \exp\left\{-\frac{1}{2}(\mathbf{w} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{D}_\omega (\mathbf{w} - \mathbf{X}\boldsymbol{\beta})\right\} \end{aligned} \tag{5.12}$$

where $\mathbf{w} = (\kappa_1/\omega_1, \ldots, \kappa_n/\omega_N)$, $\mathbf{D}_\omega = diag(\omega_1, \ldots, \omega_N)$, and $\mathbf{D}_\tau = diag(\tau_1^2, \ldots, \tau_p^2)$. Thus the full conditional distribution of $\boldsymbol{\beta}$ is Gaussian with mean $\mathbf{V}_\omega(\mathbf{X}'\boldsymbol{\kappa} + \mathbf{D}_\tau^{-1}\mathbf{0}_p)$ and variance $\mathbf{V}_\omega = (\mathbf{X}'\mathbf{D}_\omega\mathbf{X} + \mathbf{D}_\tau{}^{-1})^{-1}$.

Subsequently using Theorem 1, we note that the full conditional distribution of $\boldsymbol{\omega}$ is in the Pólya-Gamma class as $\boldsymbol{\omega_i}|\boldsymbol{\psi} \sim P.G.(n_i, \boldsymbol{\psi})$ where $\boldsymbol{\psi} = \mathbf{X}\boldsymbol{\beta}$. To sample from this distribution we use the *BayesLogit* package in R (see Polson et al. (2013)) available at https://cran.r-project.org/web/packages/BayesLogit/BayesLogit.pdf.

For each $l = 1, \ldots, p$, the portion of the joint density involving τ_l^2 is

$$(\tau_l^2)^{-1/2} \exp\left\{-\frac{1}{2}\left(\frac{\beta_l^2}{\tau_l^2} + \lambda^2\tau_l^2\right)\right\}$$

which is proportional to a popular parameterization of the density of the reciprocal of an inverse Gaussian random variable (Folks and Chhikara (1978)). Thus, the conditional distribution of $1/\tau_l^2$ is Inverse Gaussian with parameters

$$\text{mean: } \mu = \sqrt{\frac{\lambda^2}{\beta_l^2}}, \quad \text{scale: } \lambda' = \lambda^2.$$

To simulate from an Inverse Gaussian distribution, we used the *statmod* package in R available at https://cran.r-project.org/web/packages/statmod/statmod.pdf

For the lasso tuning parameter, if we assume a prior distribution on λ^2, the Gibbs sampler samples from the conditional distribution in (5.9) which happens to be a Gamma distribution. However, for this paper, given the computational cost associated with datasets of dimensions that we have used here, we chose to use the cross validation method. We used the *glmnet* package in R (see James et al. (2013)) available at `https://cran.r-project.org/web/packages/glmnet/glmnet.pdf`

The Gibbs sampler simply samples cyclically from the distributions of $\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\tau^2}$, and λ^2 conditional on the current values of the each parameter. 25% of the dataset was reserved as a validation set, and the model was trained on the remaining 75%. Once the posterior estimates were obtained, $\hat{\boldsymbol{\theta}}$ was computed, which is the probability of the each ray being blocked or efficient and based on a threshold value, classifications were made. Confusion matrices and Area under the Curve (AUC) was computed to measure the performance of our proposed model. Determining the threshold value is left at the researcher's discretion as it is specific to the application. A threshold value of 0.5 would represent the equally likely case, which means that a ray path has equal chances of being either blocked or efficient. However, researcher prior knowledge may be contrary to this and hence, is to be determined according to the problem at hand.

Information from Inefficient Paths

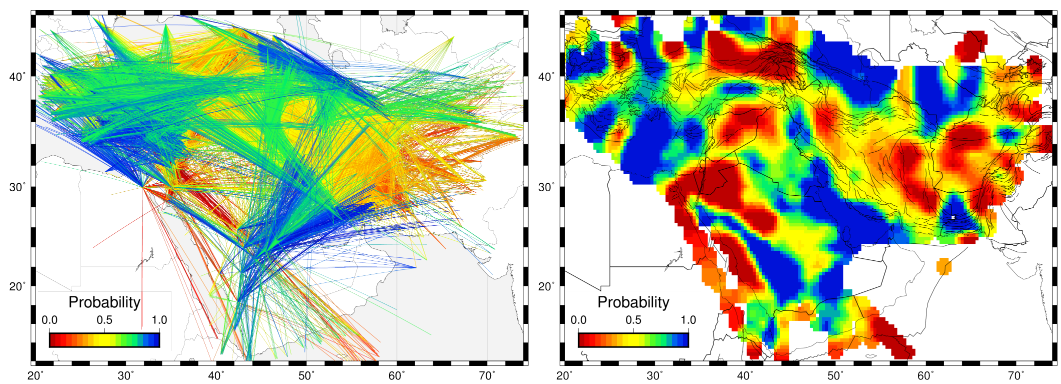
We evaluated different way to use the inefficient paths when we predict the probability of S_n propagation: (case 1) discard the inefficient paths; (case 2) convert all the inefficient paths to efficient; (case 3) convert all the inefficient paths to blocked.

Table 5.3: Prediction Accuracy and the area under the curve (AUC) for case 2 and case 3.

Model	**Accuracy**		**AUC**	
	In-Sample	Out-Sample	In-Sample	Out-Sample
ME case 2	64.64%	64.03%	0.6351	0.6303
ME case 3	77.06%	75.94%	0.7481	0.7311
East Asia case 2	66.87%	66.67%	0.6683	0.6668
East Asia case 3	86%	85.38%	0.7315	0.7034

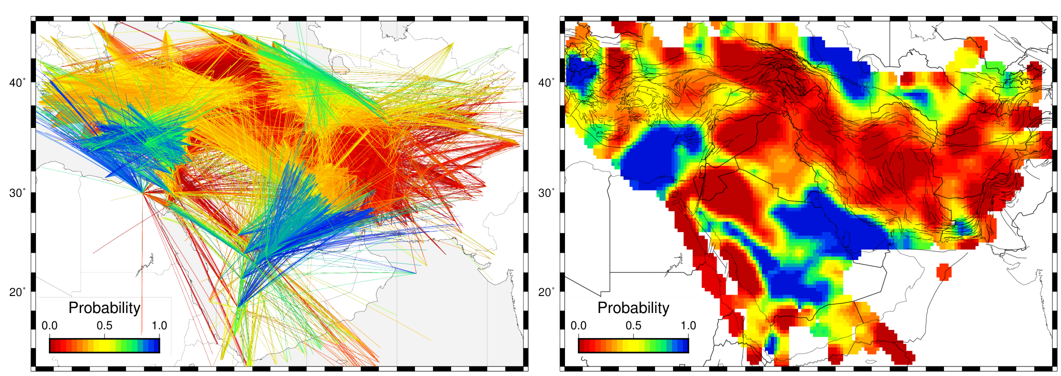
Table 5.4: Confusion Matrices for case 2 and case 3.

Model	**T0P0**	**T2P0**	**T0P2**	**T2P2**
ME case 2	0.5846	0.3138	0.4154	0.6862
ME case 3	0.9388	0.6413	0.0612	0.3587
East Asia case 2	0.7093	0.3763	0.2907	0.6237
East Asia case 3	0.9832	0.8649	0.0168	0.1351



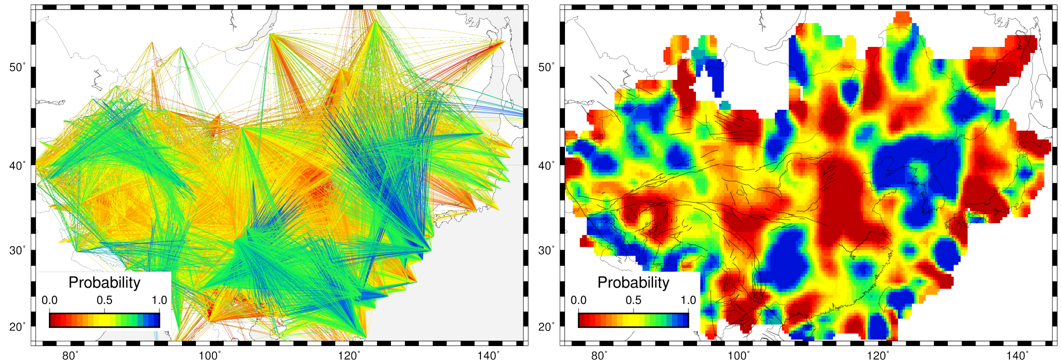
(a) Predicted probability rays. (b) Predicted probability tomography.

Figure 5.14: Predicted probability of ME data for case 2.



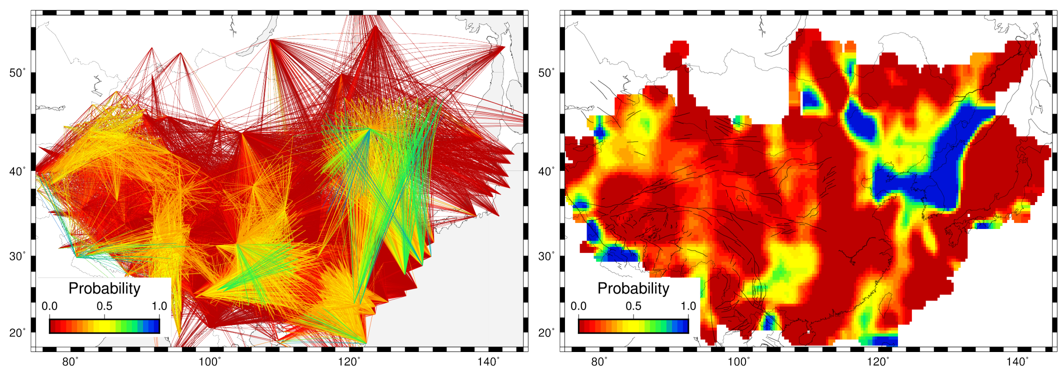
(a) Predicted probability rays. (b) Predicted probability tomography.

Figure 5.15: Predicted probability of ME data for case 3.



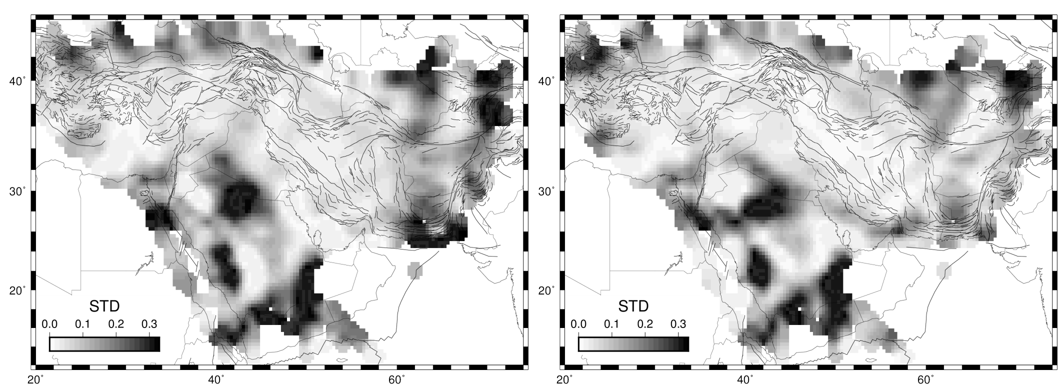
(a) Predicted probability rays. (b) Predicted probability tomography.

Figure 5.16: Predicted probability of the East Asia data for case 2.



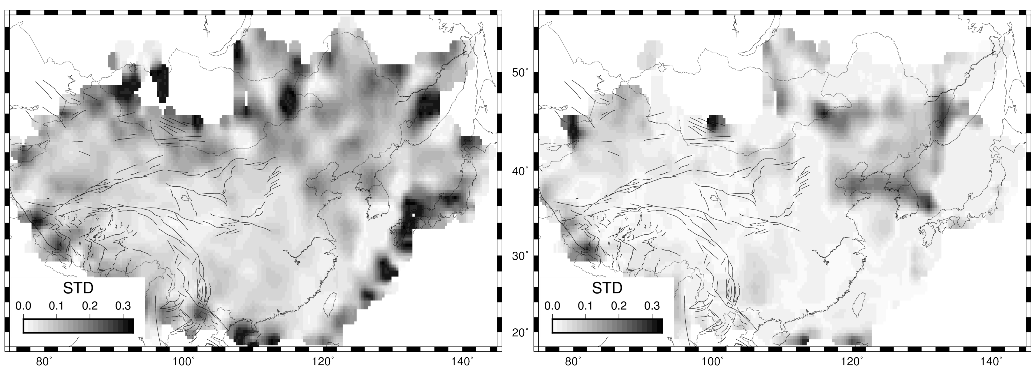
(a) Predicted probability rays. (b) Predicted probability tomography.

Figure 5.17: Predicted probability of the East Asia data for case 3.



(a) ME data for case 2. (b) ME data for case 3.

Figure 5.18: STD error of Bayesian lasso method for ME data.



(a) The East Asia data for case 2. (b) The East Asia data for case 3.

Figure 5.19: STD error of Bayesian lasso method for the East Asia data.

Chapter 6

Summary and Concluding Remarks

In this dissertation my focus has been on regional seismic phase blockage and propagation properties across two portions of Asia: the Middle East and East Asia. Using a variety of seismic waveforms recorded in both regions, we conducted three main research projects: (1) tomographically map variations in S_n and L_g attenuation for China, (2) estimate site amplification of both L_g and S_n in China, and (3) finally developing methodology to build a model that is able to predict the probability of blockage for S_n phases.

In the third chapter of this dissertation, the Reverse Two Stations method was applied to estimate the attenuation of regional seismic shear waves (S_n and L_g). S_n and L_g attenuation are proxies for lithoshpere asthenosphere boundary and crustal attenuation, respectively. This approach and our results show the following:

- We see fairly good correspondence between geologic structures and our Q map.
- We see strong evidence of scattering attenuation at passive margins; for example along the southeastern Chinese coastline.
- The L_g phase attenuation is generally low across the Tibetan plateau.

- We see a good correlation between uppermost mantle shear wave velocity models and our S_n Q map.
- In particular, we see very high Q values below the Sichuan basin and Ordos plateau.
- We see strong S_n attenuation (low S_n Q) in all regions with active volcanism, such as Datong volcano and Changbaishan.
- There are substantial differences in the S_n and L_g tomographic Q models, suggesting that there is no widespread energy leakage from S_n to L_g or vice versa.

We also compared S_n and L_g site responses across China in order to better understand the relationship between the two. Site response is often used in seismic hazards analysis. Recorded seismic data from a large number of different seismic arrays were analysed and the site response was calculated. We simulated local and regional amplification and fundamental frequency of a 2D model with point-source radiation. We found:

- Models using simple geologic structures show consistent results between regional and local site responses while complex geologic structures show differences.
- The effect of topography causes frequency dependent site amplification.
- The L_g phase is strongly amplified by the basins.
- In particular, the edges of the basins strongly amplify regional high-frequency seismic waves.
- The S_n amplitude given by RTM may include the effect of crustal legs.

Finally, we developed a method to be able to objectively predict the likelihood of phase blockage. This method used a logistic regression model that in addition to making predictions of blockage also provide confidence limits. We have applied this new method to efficiency data from the Middle East and East Asia:

- We created simulated blockage data sets in order to test our method. We find that for simulated data our method does an excellent job nearing an accuracy rate greater than 95%. This suggests that most of the inaccuracy seen in observed data is from noise in the propagation efficiencies.

- We have also developed a method to characterize noise in blockage data sets.

- In general, we see good correlation between our attenuation models and the regions with the highest probability of blockage.

- We find extensive regions with high probabilities of blockage of the S_n phase across all active plate boundaries in East Asia and the Middle East.

Bibliography

Agard, P., Omrani, J., Jolivet, L., Whitechurch, H., Vrielynck, B., Spakman, W., Monié, P., Meyer, B., and Wortel, R. (2011). Zagros orogeny: a subduction-dominated process. *Geological Magazine*, 148(5-6):692–725.

Ahadov, B. and Jin, S. (2017). Present-day kinematics in the eastern mediterranean and caucasus from dense gps observations. *Physics of the Earth and Planetary Interiors*, 268:54–64.

Al-Damegh, K., Sandvol, E., Al-Lazki, A., and Barazangi, M. (2004). Regional seismic wave propagation (lg and sn) and pn attenuation in the arabian plate and surrounding regions. *Geophysical Journal International*, 157(2):775–795.

Al-Lazki, A. I., Seber, D., Sandvol, E., Turkelli, N., Mohamad, R., and Barazangi, M. (2003). Tomographic pn velocity and anisotropy structure beneath the anatolian plateau (eastern turkey) and the surrounding regions. *Geophysical Research Letters*, 30(24).

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association*, 88(422):669–679.

Bae, K. and Mallick, B. K. (2004). Gene selection using a two-level hierarchical bayesian model. *Bioinformatics*, 20(18):3423–3430.

Bao, X. (2011). *Seismic attenuation of regional phases in the northern middle east and the Tibetan Plateau*. phdthesis, University of Missouri - Columbia.

Bao, X., Sandvol, E., Chen, Y. J., Ni, J., Hearn, T., and Shen, Y. (2012). Azimuthal anisotropy of lg attenuation in eastern tibetan plateau. *Journal of Geophysical Research*, 117:B10309.

Bao, X., Sandvol, E., Ni, J., Hearn, T., Chen, Y. J., and Shen, Y. (2011). High resolution regional seismic attenuation tomography in eastern tibetan plateau and adjacent regions. *Geophysical Research Letters*, 38.

Barron, J. and Priestley, K. (2009). Observations of frequency-dependentSnpropagation in northern tibet. *Geophysical Journal International*, 179(1):475–488.

Baumgardt, D. R. (2001). Sedimentary basins and the blockage of lg wave propagation in the continents. *Pure and applied geophysics*, 158(7):1207–1250.

Benz, H. M., Frankel, A., and Boore, D. M. (1997). Regional lg attenuation for the continental united states. *Bulletin of the Seismological Society of America*, 87(3):606–619.

Calvert, A., Sandvol, E., Seber, D., Barazangi, M., Vidal, F., Alguacil, G., and Jabour, N. (2000). Propagation of regional seismic phases (lg and sn) and pn velocity structure along the africa-iberia plate boundary zone: tectonic implications. *Geophysical Journal International*, 142(2):384–408.

Campillo, M., Plantet, J.-L., and Bouchon, M. (1985). Frequency-dependent attenuation in the crust beneath central france from lg waves: data analysis and numerical modeling. *Bulletin of the Seismological Society of America*, 75(5):1395–1411.

Casella, G. (2001). Empirical Bayes Gibbs sampling. *Biostatistics*, 2(4):485–500.

Chen, L., Capitanio, F. A., Liu, L., and Gerya, T. V. (2017). Crustal rheology controls on the tibetan plateau formation during india-asia convergence. *Nature communications*, 8:15992.

Chen, Y. and Niu, F. (2016). Joint inversion of receiver functions and surface waves with enhanced preconditioning on densely distributed cndsn stations: Crustal and upper mantle structure beneath china. *Journal of Geophysical Research: Solid Earth*.

Choi, J.-H., Klinger, Y., Ferry, M., Ritz, J.-F., Kurtz, R., Rizza, M., Bollinger, L., Davaasambuu, B., Tsend-Ayush, N., and Demberel, S. (2018). Geologic inheritance and earthquake rupture processes: The 1905 m $\geq$ 8 tsetserleg-bulnay strike-slip earthquake sequence, mongolia. *Journal of Geophysical Research: Solid Earth*, 123(2):1925–1953.

Cong, L., Xie, J., and Mithcell, B. J. (1996). Excitation and propagation of lg from earthquakes in central asia with implications for explosion/earthquake discrimination. *Geophysical Reserach: Solid Earth*, 101(B12):27779–27789.

Faccenna, C., Becker, T. W., Jolivet, L., and Keskin, M. (2013). Mantle convection in the middle east: Reconciling afar upwelling, arabia indentation and aegean trench rollback. *Earth and Planetary Science Letters*, 375:254–269.

Fan, G. and Lay, T. (2002). Characteristics of lg attenuation in the tibetan plateau. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 107(B10):ESE–14.

Fan, G.-W. and Lay, T. (2003). Strong lg attenuation in the tibetan plateau. *Bulletin of the Seismological Society of America*, 93(5):2264–2272.

Folks, J. L. and Chhikara, R. S. (1978). The inverse gaussian distribution and its statistical application—a review. *Journal of the Royal Statistical Society: Series B (Methodological)*, 40(3):263–275.

Ford, S. R., Dreger, D. S., Mayeda, K., Walter, W. R., Malagnini, L., and Phillips, W. S. (2008). Regional attenuation in northern california: A comparison of five 1d q methods. *Bulletin of the Seismological Society of America*, 98(4):2033–2046.

Gao, S., Liu, H., Davis, P. M., and Knopoff, L. (1996). Localized amplification of seismic waves and correlation with damage due to the Northridge earthquake: Evidence for focusing in Santa Monica. *Bulletin of the Seismological Society of America*, 86(1B):S209–S230.

Gibson Jr, R. L. and Campillo, M. (1994). Numerical simulation of high-and low-frequency lg-wave propagation. *Geophysical Journal International*, 118(1):47–56.

Gök, R., Sandvol, E., Türkelli, N., Seber, D., and Barazangi, M. (2003). Sn attenuation in the anatolian and iranian plateau and surrounding regions. *Geophysical Research Letters*, 30(24):8042.

Gök, R., Türkelli, N., Sandvol, E., Seber, D., and Barazangi, M. (2000). Regional wave propagation in turkey and surrounding regions. *Geophysical Research Letters*, 27(3):429–432.

Golub, G. H., Heath, M., and Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223.

Hafkenscheid, E., Wortel, M. J. R., and Spakman, W. (2006). Subduction history of the tethyan region derived from seismic tomography and tectonic reconstructions. *Journal of Geophysical Research: Solid Earth*, 111(B8).

Huestis, S., Molnar, P., and Oliver, J. (1973). Regional sn velocities and shear velocity in the upper mantle. *Bulletin of the Seismological Society of America*, 63(2):469–475.

Imtiaz, A., Cornou, C., Bard, P.-Y., and Zerva, A. (2015). Spatial coherence of seismic ground motion and geometric structure of the sub-surface: an example in argostoli, greece. In *9ème Colloque National AFPS 2015*.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Kadinsky-Cade, K., Barazangi, M., Oliver, J., and Isacks, B. (1981). Lateral variations of high-frequency seismic wave propagation at regional distances across the turkish and iranian plateaus. *Journal of Geophysical Research: Solid Earth*, 86(B10):9377–9396.

Karato, S.-i. (1993). Importance of anelastic1ty in the interpretation of seismic tomography. *Geophysical research letters*, 20(15):1623–1626. low velocity anomalies in the mantle will be accounted for by solid state mechanisms. However, low velocity should not be interpreted to low density anomalies.

Kaviani, A., Paul, A., Moradi, A., Mai, P. M., Pilia, S., Boschi, L., Rümpker, G., Lu, Y., Tang, Z., and Sandvol, E. (2020). Crustal and uppermost mantle shear wave velocity structure beneath the Middle East from surface wave tomography. *Geophysical Journal International*, 221(2):1349–1365.

Kaviani, A., Sandvol, E., Bao, X., Rümpker, G., and Gök, R. (2015). The structure of the crust in the turkish–iranian plateau and zagros using lg q and velocity. *Geophysical Journal International*, 200(2):1254–1268.

Kennett, B. (1986). Lg waves and structural boundaries. *Bulletin of the Seismological Society of America*, 76(4):1133–1141.

Kennett, B. and Mykkeltveit, S. (1984). Guided wave propagation in laterally varying media—ii. lg-waves in north-western europe. *Geophysical Journal International*, 79(1):257–267.

Keskin, M. (2003). Magma generation by slab steepening and breakoff beneath a subduction-accretion complex: An alternative model for collision-related volcanism in eastern anatolia, turkey. *Geophysical Research Letters*, 30(24).

Knopoff (1964). Q. *Review of Geophysics*, 2(4):625–660.

Knopoff, L., Schwab, F., and Kauselt, E. (1973). Interpretation of lg. *Geophysical Journal International*, 33(4):389–404.

Lei, J., Xie, F., Fan, Q., and Santosh, M. (2013). Seismic imaging of the deep structure under the chinese volcanoes: An overview. *Physics of the Earth and Planetary Interiors*, 224:104–123.

Li, H., Li, S., Song, X., Gong, M., Li, X., and Jia, J. (2012). Crustal and uppermost mantle velocity structure beneath northwestern china from seismic ambient noise tomography. *Geophysical Journal International*, 188(1):131–143.

Li, H., Song, X., Lü, Q., Yang, X., Deng, Y., Ouyang, L., Li, J., Li, X., and Jiang, G. (2018). Seismic imaging of lithosphere structure and upper mantle deformation beneath east-central china and their tectonic implications. *Journal of Geophysical Research: Solid Earth*, 123(4):2856–2870.

McNamara, D. E. and Owens, T. J. (1995). Observations of regional phase propagation across the tibetan plateau. *Journal of Geophysical Research*, 100(B11):22215–22229.

Mellors, R. J., Camp, V. E., Vernon, F. L., Al-Amri, A., and Ghalib, A. (1999). Regional waveform propagation in the arabian peninsula. *Journal of Geophysical Research: Solid Earth*, 104(B9):20221–20235.

Molnar, P. and Oliver, J. (1969). Lateral variations of attenuation in the upper mantle and discontinuities in the lithosphere. *Journal of Geophysical Research*, 74(10):2648–2682.

Molnar, P. and Tapponnier, P. (1977). Relation of the tectonics of eastern china to the india-eurasia collision: Application of slip-line field theory to large-scale continental tectonics. *Geology*, 5(4):212–216.

Murphy, J., Davis, A., and Weaver, N. (1971). Amplification of seismic body waves by low-velocity surface layers. *Bulletin of the Seismological Society of America*, 61(1):109–145.

Murray, W. and Overton, M. L. (1981). A projected lagrangian algorithm for non-linear l_1 optimization. *SIAM Journal on Scientific and Statistical Computing*, 2(2):207–224.

Ni, J. and Barazangi, M. (1983). High-frequency seismic wave propagation beneath the indian shield, himalyan arc, tibetan plateau and surrounding regions: high uppermost mantle velocities and efficient sn propagation beneath tibet. *Geophys. J. R. astr. Soc.*, 72:665–689.

Paige, C. C. and Saunders, M. A. (1982). Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8(1):43–71.

Park, J. and Levin, V. (2002). Seismic anisotropy: tracing plate dynamics in the mantle. *Science*, 296(5567):485–489.

Park, T. and Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686.

Phillip, H., Cisternas, A., Gvishiani, A., and Gorshkov, A. (1989). The caucasus:

an actual example of the initial stages of continental collision. *Tectonophysics*, 161(1-2):1–21.

Phillips, W. S., Hartse, H. E., and Rutledge, J. T. (2005). Amplitude ratio tomography for regional phase q. *Geophysical Research Letters*, 32(21).

Polson, N. G., Scott, J. G., and Windle, J. (2013). Bayesian inference for logistic models using pólya–gamma latent variables. *Journal of the American statistical Association*, 108(504):1339–1349.

Ranasinghe, N. R., Gallegos, A., Hearn, T., Ni, J., and Sandvol, E. (2018). Frequency-dependent lg attenuation in northeast china and its implications. *Geophysical Journal International*, 212(3):2131–2142.

Ranasinghe, N. R., Gallegos, A. C., Trujillo, A. R., Blanchette, A. R., Sandvol, E. A., Ni, J., Hearn, T. M., Tang, Y., Grand, S. P., Niu, F., et al. (2015). Lg attenuation in northeast china using necessarray data. *Geophysical Journal International*, 200(1):67–76.

Rapine, R. R. and Ni, J. F. (2003). Propagation characteristics of sn and lg in northeastern china and mongolia. *Bulletin of the Seismological Society of America*, 93(2):939–945.

Rapine, R. R., Ni, J. F., and Hearn, T. M. (1997). Regional wave propagation in china and its surrounding regions. *Bulletin of the Seismological Society of America*, 87(6):1622–1636.

Reilinger, R., McClusky, S., Vernant, P., Lawrence, S., Ergintav, S., Cakmak, R., Ozener, H., Kadirov, F., Guliev, I., Stepanyan, R., et al. (2006). Gps constraints on continental deformation in the africa-arabia-eurasia continental collision zone and implications for the dynamics of plate interactions. *Journal of Geophysical Research: Solid Earth*, 111(B5).

Richards, J. P. (2015). Tectonic, magmatic, and metallogenic evolution of the tethyan orogen: From subduction to collision. *Ore Geology Reviews*, 70:323 – 345.

Rodgers, A. J., Ni, J. F., and Hearn, T. M. (1997). Propagation characteristics of short-period sn and lg in the middle east. *Bulletin of the Seismological Society of America*, 87(2):396–413.

Sandvol, E., Al-Damegh, K., Calvert, A., Seber, D., Barazangi, M., Mohamad, R., Gök, R., Türkelli, N., and Gürbüz, C. (2001). Tomographic imaging of lg and sn propagation in the middle east. *Pure and Applied Geophysics*, 158(7):1121–1163.

Sandvol, E., Turkelli, N., Zor, E., Gok, R., Bekler, T., Gurbuz, C., Seber, D., and Barazangi, M. (2003). Shear wave splitting in a young continent-continent collision: An example from eastern turkey. *Geophysical Research Letters*, 30(24).

Sereno, T. J. (1990). Frequency-dependent attenuation in eastern kazakhstan and implications for seismic detection thresholds in the soviet union. *Bulletin of the Seismological Society of America*, 80(6):2089–2105.

Simmons, N., Myers, S., and Johannesson, G. (2011). Global-scale p wave tomography optimized for prediction of teleseismic and regional travel times for middle east events: 2. tomographic inversion. *Journal of Geophysical Research: Solid Earth*, 116(B4).

Skobeltsyn, G. A. (2014). *Upper mantle shear wave velocity structure of the east Anatolian-Caucasus region*. phdthesis, University of Missouri - Columbia.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2):111–133.

Tian, Y., Zhao, D., Sun, R., and Teng, J. (2009). Seismic imaging of the crust and upper mantle beneath the north china craton. *Physics of the Earth and Planetary Interiors*, 172(3-4):169–182.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Wang, M. and Shen, Z.-K. (2020). Present-day crustal deformation of continental china derived from gps and its tectonic implications. *Journal of Geophysical Research: Solid Earth*, 125(2):e2019JB018774. e2019JB018774 2019JB018774.

Wei, W., Hammond, J. O. S., Zhao, D., Xu, J., Liu, Q., and Gu, Y. (2019). Seismic evidence for a mantle transition zone origin of the wudalianchi and halaha volcanoes in northeast china. *Geochemistry, Geophysics, Geosystems*, 20(1):398–416.

Wesnousky, S., Jones, L., Scholz, C., and Deng, Q. (1984). Historical seismicity and rates of crustal deformation along the margins of the ordos block, north china. *Bulletin of the Seismological Society of America*, 74(5):1767–1783.

Xie, J. (2002). Lg q in the eastern tibetan plateau. *Bulletin of the Seismological Society of America*, 92(2):871–876.

Yin, A. and Harrison, T. M. (2000). Geologic evolution of the himalayan-tibetan orogen. *Annual Review of Earth and Planetary Sciences*, 28(1):211–280.

Zhao, C., Kennett, B., and Furumura, T. (2003). Contrasts in regional seismic wave propagation to station wmq in central asia. *Geophysical Journal International*, 155(1):44–56.

Zhao, G., Cawood, P. A., Li, S., Wilde, S. A., Sun, M., Zhang, J., He, Y., and Yin, C. (2012). Amalgamation of the north china craton: key issues and discussion. *Precambrian Research*, 222:55–76.

Zhao, L., Zheng, T., and Lü, G. (2008). Insight into craton evolution: Constraints from shear wave splitting in the north china craton. *Physics of the Earth and Planetary Interiors*, 168(3-4):153–162.

Zheng, Y.-F., Xiao, W.-J., and Zhao, G. (2013). Introduction to tectonics of china.

VITA

Hongjun Hui was born in Hebei, China. He received the Bachelor of Science in applied physics from Harbin Institute of Technology in 2011. He received the Master of Science in geophysics from Peking University in 2014. He started work with Eric Sandvol in the fall of 2014. At the University of Missouri, he become proficient in numerical simulation and seismic data processing.

Hongjun Hui married Huiqian Zhang in November 2019. Their daughter Claire was born in 2020. Hongjun will start working in Beijing after receiving doctoral degree.