UNIVERSITY OF LEEDS

Traveltime eikonal tomography with physics informed neural networks

Yunpeng Chen

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Earth and Environment

December 2024

Intellectual Property

The candidate conforms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given where reference has been made to the work of others.

The work in chapter 2 has been published as follows:

Chen, Y., de Ridder, S. A., Rost, S., Guo, Z., Wu, X., & Chen, Y. (2022). Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the northeastern margin of the Tibetan Plateau. *Geophysical Research Letters, 49*(21), e2022GL099053.

The work was initiated following suggestions from Sjoerd de Ridder and Sebastian Rost to utilize the PINN method, with Yongshun Chen and Zhen Guo recommending its application to the ChinArray-Himalaya II dataset. The candidate developed the methodology, conducted the formal analysis of results, and wrote the manuscript. Sjoerd de Ridder and Sebastian Rost provided comments and suggestions on the work. Xiaoyang Wu contributed to data acquisition and conventional eikonal tomography. Yongshun Chen and Zhen Guo secured funding for the project.

The work in chapter 3 has been published as follows:

Chen, Y., de Ridder, S. A., Rost, S., Guo, Z., Wu, X., Li, S., & Chen, Y. (2023). Physics-Informed Neural Networks for Elliptical-Anisotropy Eikonal Tomography: Application to Data From the Northeastern Tibetan Plateau. *Journal of Geophysical Research: Solid Earth, 128*(12), e2023JB027378.

The candidate conceptualized and designed the study in collaboration with Sjoerd de Ridder and Sebastian Rost. The candidate developed the methodology, conducted the formal analysis

of results, and wrote the manuscript. Sjoerd de Ridder and Sebastian Rost provided comments and suggestions on the work. Xiaoyang Wu contributed to data acquisition and double beamforming tomography, while Shilin Li assisted with the formal analysis. Zhen Guo and Yongshun Chen secured funding for the project.

The work in chapter 4 is a draft manuscript, ready for submission:
Chen, Y., de Ridder, S. A., Rost, S., Guo, Z., & Chen, Y. (2024). Teleseismic Rayleigh Wave Tomography in Northeastern Tibet Using Physics-Informed Neural Networks.
The candidate led the overall research process. Sjoerd de Ridder and Sebastian Rost provided comments and suggestions on the work. Yongshun Chen and Zhen Guo secured funding for the project.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Sjoerd de Ridder and Prof. Sebastian Rost. Choosing Sjoerd as my supervisor was the wisest decision I made during my PhD journey. He encouraged me with a Chinese proverb, comparing research to "crossing the river by feeling the stones," teaching me how to identify and solve problems. I am grateful for his guidance in selecting research directions, developing critical thinking, and advancing research skills. I also appreciate his help and support in daily life, which enabled me to quickly adapt to life in the UK. I am thankful to Sebastian for providing comprehensive support and guidance throughout my PhD, maintaining regular communication, offering insights to overcome research obstacles, and patiently providing advice for paper writing, which greatly contributed to my academic growth and knowledge. I would also like to thank Prof. Andy Nowacki for serving as a temporary second supervisor in my fourth year, bringing valuable suggestions and comments to my research.

I am deeply grateful to my supervisors at SUSTech, Prof. Yongshun Chen and Prof. Zhen Guo. Since joining the research group in 2018, they have helped me build a solid foundation in geophysics and taught me how to conduct scientific research, continuously influencing me with their rigorous academic attitude and pragmatic spirit. I thank Prof. Yongshun Chen for his generous financial support, which created a robust platform for my research, and for sharing valuable experiences and perspectives in planning my academic career. I appreciate Prof. Zhen Guo's careful guidance and patience in explaining the key points and limitations of my work during our progress discussions, which always inspired new solutions.

I would like to thank the members of the G$^{\text{i}}$G, particularly Prof. Longwei Chen, Afsaneh Mohammadzaheri, Bolin Li, and Ruhul Hajjaj, all talks and discussions with you have been enlightening. I am also grateful to the members of IGT and IAG, especially Prof. Chong Zhang,

Bo Li, Jin Fang, Jingru Wang, Lei Bao, Qi Ou, Sheng Chen, Xiaojiao Pang, Yuan Gao, Yan Li, Yikai Zhu, and Zhen Li, for their companionship and support during my time in Leeds. Special thanks to Felix Boschetty and Gamal Ahmed - working and communicating with you in the same office has created precious memories.

I extend my gratitude to all my colleagues in the Department of Ocean Science and Engineering at SUSTech. Thanks to Prof. Bin Luo, Prof. Chen Zhang, Prof. Lipeng He, Prof. Shilin Li, Prof. Wei Cai, Prof. Yong Yu, Anqi Zhang, Changqiao Zou, Guanzhi Wang, Haolan Liu, Junqing Tan, Junwei Ma, Keyuan Wu, Pengfei Ren, Qipeng Bai, Renxian Xie, Xiaoyang Wu, Xingli Fan, Xueqing Gao, Yang Zhao, Yang Tian, Zixuan Wu, and others. Both our academic interactions and the meaningful time we spent together have made my time at SUSTech incredibly wonderful.

I am deeply thankful to my family for their unconditional emotional encouragement and material support - home will always be my harbor. I thank my girlfriend Rongzi for accompanying me throughout my PhD journey, sharing countless beautiful moments together. I look forward to exploring more beautiful moments in the future with you.

As I reflect on my academic journey, I am grateful that both Leeds and Shenzhen blessed me not only with enriching research experiences but also with their magnificent sunsets - a daily reminder of the beauty in pursuing knowledge under different skies.

Abstract

This thesis develops and validates three innovative methodologies for surface wave tomography using physics-informed neural networks (PINNs): pinnET (PINN-based eikonal tomography), pinnEAET (PINN-based elliptical-anisotropic eikonal tomography), and pinnTET (PINN-based teleseismic eikonal tomography). These methods progressively address challenges in seismic tomography from isotropic to anisotropic tomography, and ambient noise to teleseismic earthquake applications. By integrating the eikonal equation as a physical constraint while leveraging neural networks' approximation capabilities, PINN-based surface wave tomography demonstrates several key advantages, including significant memory efficiency, physics-guided interpolation for sparse data regions, simultaneous multi-frequency processing, and flexible evaluation at arbitrary locations. However, challenges remain in computational efficiency and automated parameter optimization.

When applied to the seismic dense array in northeastern Tibetan Plateau using both ambient noise (10-40 s periods) and teleseismic data (20-80 s periods), these methods reveal significant lateral heterogeneity in velocity structure and azimuthal anisotropy. Notably, they achieve comparable resolution quality with only approximately 20% or even less of traditionally required data. The results indicate prominent low-velocity zones beneath the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane, contrasting with high-velocity zones in the Ordos Block and central Qinling Orogen. These findings provide new insights into the region's complex crustal and upper mantle structure while demonstrating the practical utility of PINN-based approaches in seismic tomography.

Comprehensive uncertainty analysis, checkerboard resolution tests and cross-validation with traditional approaches confirm the methods' reliability and resolution capabilities. This work demonstrates that PINN-based approaches provide valuable alternatives for seismic tomography,

particularly in regions with limited data coverage, while establishing a foundation for future developments in physics-constrained seismic tomography methods.

Contents

List of Figures

List of Abbreviations

AB	Alxa Block
ATF	Altyn-Tagh Fault
ANN	Artificial Neural Network
ASWMS	Automated Surface Wave Measuring System
BNN	Bayesian Neural Network
BS	Bivariate Statistics
CAOB	Central Asian Orogenic Belt
cPINN	Conservative Physics-Informed Neural Network
DL	Deep Learning
DNN	Deep Neural Network
PINN	Physics-Informed Neural Network
FBPINN	Finite Basis Physics-Informed Neural Network
FCNN	Fully Connected Neural Network
fPINN	Fractional Physics-Informed Neural Network
GAN	Generative Adversarial Network
GSDF	Generalized Seismological Data Functional
HB	Hexi Basin
HG	Hetao Graben

HLSF Helanshan Fault

HVZ High Velocity Zone

HYF Haiyuan Fault

JB Jilantai Basin

JTF Jintan Fault

KLF Kunlun Fault

LB Longzhong Basin

LPO Lattice Preferred Orientation

LPSF Liupanshan Fault

LR Logistic Regression

LSSF Longshoushan Fault

LSTM Long Short-term Memory

LVZ low Velocity Zone

ML Machine Learning

MSE Mean Squared Error

NE Northeastern

NN Neural Network

OB Ordos Block

PCA Principal Component Analysis

PDE Partial Differential Equation

PINN Physics-Informed Neural Network

PinnET PINN-based Eikonal Tomography

PinnEAET PINN-based Elliptical-anisotropic Eikonal Tomography

PinnTET PINN-based Teleseismic Eikonal Tomography

PREM Preliminary Reference Earth Model

QOB Qilian Orogenic Belt

RB Ruoergai Basin

RF Random Forest

RMSE Root Mean Square Error

RNN Recurrent Neural Network

SB Sichuan Basin

SciML Scientific Machine Learning

SGT Songpan-Ganzi Terrane

SNR Signal-to-noise Ratio

SOM Self-organizing Map

SPO Shaped Preferred Orientation

SVM Support Vector Machine

UTM Universal Transverse Mercator

VE-PINN Velocity-encoded Physics-Informed Neural Network

VMD Variational Mode Decomposition

vPINN Variational Physics-Informed Neural Network

WB Weihe Basin

WQF West-Qinling Fault

WQO Western Qinling Orogen

XPINN eXtended Physics-Informed Neural Network

YG Yinchuan Graben

Chapter 1

Introduction

1.1 Seismic tomography

1.1.1 Overview

Seismic tomography has significantly advanced our understanding of the Earth's interior by providing detailed images of subsurface structures. It employs seismic wave data from earthquakes, artificial sources or virtual sources using ambient noise to map variations in the Earth's seismic properties, such as velocity, anisotropy and attenuation. This process typically involves solving a complex inversion problem by using observational data including seismic wave traveltime, amplitude and/or other waveform characteristics. Methodologically, seismic tomography first requires establishing a relationship $\mathbf{d} = \mathbf{g}(\mathbf{m})$ between the seismic data $\mathbf{d}$ and the seismic model $\mathbf{m}$, where $\mathbf{g}$ is the forward operator that describes the physics of seismic wave propagation, then predicting the data $\mathbf{d}$ from a given model, and finally finding a model $\mathbf{m}$ that best fits the observations (Rawlinson et al. 2010). The concept of seismic tomography dates back to 1976, when Aki and Lee (1976) pioneered the construction of a 3-D crustal velocity structure beneath California based on the first-arrival P wave traveltime information from local earthquakes. In 1977, Aki et al. (1977) inverted for the 3-D velocity structure of the lithosphere beneath Norwegian Seismic Array based on the teleseismic P wave traveltime residuals. In the same year, Dziewonski et al. (1977) constructed a 3-D velocity model of the lower mantle on a global scale using the P wave traveltime residuals. The pioneering work of these researchers triggered a wave of interest in using seismic tomography to study the Earth's internal structure, laying a critical theoretical foundation for future research. Today, the technique remains pivotal

in seismology, contributing to insights into topics such as plate tectonics, Earth's dynamics and earthquake generation.

According to the type of seismic waves utilized, seismic tomography can be divided into body wave tomography and surface wave tomography. Body wave tomography focuses on the analysis of P and S waves in terms of their traveltimes and waveforms as they travel through the Earth's interior. On the other hand, surface wave tomography specifically targets the study of Rayleigh and Love waves, which propagate along the Earth's surface. These surface waves have unique characteristics that make them particularly useful for imaging the Earth's crust and upper mantle. One of the key methods employed in surface wave tomography is the use of dispersion curves. The dispersive character of surface wave refers to the phenomenon where different frequencies of the wavefield travel at different velocities, potentially allowing higher vertical resolution compared to body wave tomography (Romanowicz 2020). By measuring the phase or group velocities of surface waves across a wide range of periods, researchers can construct dispersion curves that provide direct information of elastic velocity about the subsurface. In the 1950s and early 1960s, surface waves were used to explore the properties of the Earth's structure, particularly the crust and upper mantle (Ewing and Press 1954; Press 1956; Dorman et al. 1960; Kuo et al. 1962), producing layered velocity models along the surface wave propagation path. Since the 1970s, with the advent of digital recording, the deployment of broadband networks and the subsequent methodological innovations, surface wave tomography has become a powerful tool to construct 3-D models of velocity (Woodhouse and Dziewonski 1984; Anderson 1987; Forsyth et al. 2005; Sabra et al. 2005; Shapiro et al. 2005; Nishida et al. 2009; Kaviani et al. 2020; Liu et al. 2021; Fan et al. 2024), anisotropy (Forsyth 1975; Tanimoto and Anderson 1985; Montagner and Tanimoto 1990; Montagner and Tanimoto 1991; Gung et al. 2003; Gaherty 2004; Hao et al. 2021; Magali et al. 2021), and attenuation (Dziewonski and Anderson 1981; Dalton and Ekström 2006; Dalton et al. 2008; Savage et al. 2010; Bao et al. 2016; Meng et al. 2021) in the crust and upper mantle on global, regional, and local scales. These studies have dramatically enhanced our understanding of crust-mantle structure, geodynamic processes, seismic source mechanics, mineral exploration and geohazard assessment. In this thesis, I employ surface wave tomography approaches to investigate the lithospheric structure and better constrain the geological implications in the study region.

1.1.2 Teleseismic surface wave tomography

Surface wave tomography using teleseismic data involves analyzing of surface waves generated by earthquakes to infer the Earth's structure. At teleseismic distances, surface waves carry the majority of long period energy (typically with periods exceeding 20 s) that provide constrains on upper mantle structure. Several tomography methods have been developed to explore the upper mantle structure based on teleseismic surface wave records. In the 1960s, the single-station method was introduced to extract and analyze average dispersion curves by directly utilizing surface wave signals generated by earthquakes (Brune et al. 1960). However, this method is often affected by uncertainties in the earthquake's origin time and hypocenter location. Teleseismic surface waves carry information about both the Earth's structure along the great circle path from the source to the station and the earthquake source itself, making it one of the long-standing challenges to separate the two (Romanowicz 2020). To eliminate the influence of source information on the results, Sato (1955) proposes the two-station method that uses two stations located approximately along the same great circle path to extract surface wave dispersion signals. Romanowicz (1982) uses two events for a single station instead of two stations to obtain seismic wave velocity from teleseismic records. In traditional surface wave tomography methods (such as the two-station method), it is commonly assumed that teleseismic surface waves arrive as plane waves. However, surface waves can deviate from the great circle path when they encounter heterogeneities during propagation, resulting in non-plane wave energy and systematic bias in velocity measurement (Friederich et al. 1994; Laske 1995). To handle the non-plane wave energy, Forsyth et al. (1998) and Forsyth et al. (2005) propose the two-plane-wave tomography method, which approximates the incident wavefield as the interference of two plane waves. By simulating the interaction of these two plane waves, this method can more accurately fit propagation characteristics of surface waves in complex geological structures.

Over the past two decades, significant improvements have been made in these methods for surface wave tomography. For the two-station method, Meier et al. (2004) apply cross-correlation functions to waveform of two stations and enhanced the accuracy of fundamental mode dispersion measurements through frequency-domain Gaussian filtering. Wu et al. (2009) present a wavelet transform approach to measure inter-station phase velocity. For two-plane-wave tomography method, Yang and Forsyth (2006) utilize the finite-frequency theory to account for the effects of heterogeneous structure on the wavefield, providing better resolution of local

structure. Li and Li (2015) extend the method to Love waves, demonstrating that the joint analysis of Rayleigh and Love waves enables simultaneous inversion for seismic velocity and radial anisotropy. These methods have been applied to different datasets and earthquakes at teleseimic distances to study the velocity structure in the crust and upper mantle. For example, Bourjot and Romanowicz (1992) reveal the crust and lithosphere structure beneath the Tibet using the two-event method based on Rayleigh wave data at 25 to 100 s. Curtis and Woodhouse (1997) present fundamental mode Rayleigh and Love wave phase velocities at periods across 32-200 s beneath the Tibetan Plateau and its surrounding regions through the two-event method. Li and Fu (2020) construct long-period phase velocity maps of Rayleigh waves and Love waves at periods between 20 and 143 s in the central and eastern parts of the Tibetan Plateau based on the two-plane-wave tomography method. These studies have demonstrated the capability of teleseismic surface wave tomography in resolving deep Earth structure and advancing our understanding of lithospheric deformation and mantle dynamics.

1.1.3 Ambient noise tomography

Ambient noise tomography involves extracting the Green's function between stations by performing cross-correlations on long-term seismic noise recordings at two stations. The surface wave dispersion curves obtained by Green's functions can then be further analyzed through tomography methods to resolve the velocity structure of the Earth's interior. While ambient noise was previously dismissed as lacking useful structural information, it has been found to reveal valuable insights into Earth's structure. In the early 2000s, a novel approach was introduced to extract the medium's Green's function between two points through cross-correlation of waveforms recorded at those points (Lobkis and Weaver 2001; Weaver and Lobkis 2001). In seismology, Campillo and Paul (2003) compute the cross-correlation functions of seismic coda data recorded at the Mexican National Seismological Network and successfully extracted the Green's function of fundamental mode Rayleigh and Love waves. Snieder (2004) also proves that the Green's function can be found by cross correlating scattered waveforms at two receivers. Shapiro and Campillo (2004) first perform cross-correlations on ambient noise data and measured the group velocity dispersion curves of fundamental mode Rayleigh waves. The proposal that Green's functions could be extracted from ambient noise through cross-correlation, and its subsequent confirmation through successful retrieval of surface wave signals, has laid a solid foundation for ambient noise tomography.

Compared to traditional tomography based on earthquakes or explosions, ambient noise tomography offers several advantages. Firstly, the Green's function obtained through cross-correlation reflects the information of the subsurface medium between the two stations, thereby avoiding errors associated with the earthquake's origin time and hypocenter location. Furthermore, the horizontal resolution of ambient noise tomography mainly depends on the spacing between stations and the wavelength of surface waves, rather than the ray distribution between stations and earthquakes. More importantly, as short-period seismic surface waves attenuate rapidly, traditional methods often fail to obtain high quality signals of short period surface wave. Due to these advantages, much research in recent years has focused on ambient noise tomography. Shapiro et al. (2005) apply this method to the USArray stations in southern California and successfully extracted high-frequency Rayleigh wave information from 7 to 18 s using only a month of continuous recordings and further analyzed group velocity distribution at 7.5 s and 15 s in this region. Yao et al. (2006) and Yao et al. (2008) demonstrate the variation of Rayleigh wave phase velocity and shear wave velocity beneath southeastern Tibet by combining empirical Green's function and the two-station method, suggesting a relationship between faults and crustal structure. Yang et al. (2007) extend the study area to cover the entire European continent using continuous seismic waveform recordings over 1 year and obtained Rayleigh wave group velocities from periods of 10 to 50 s, revealing high resolution maps of sedimentary basins and crust. These applications imply that ambient noise tomography has become a popular and reliable method for visualizing Earth's structure across diverse scales.

Ambient noise tomography provides high resolution of crustal structures by extracting shorter-period surface waves. This advantage in resolving shallow structures complements teleseismic surface wave tomography, which primarily constrains deeper structures. The integration of both methods enables comprehensive imaging across different depth ranges from the crust to upper mantle.

1.1.4 Seismic anisotropy

Seismic anisotropy refers to the phenomenon where seismic wave velocity varies with propagation direction, as caused by the Earth's elastic properties. This characteristic is often a result of crustal and upper mantle deformation, which can arise from either the Lattice Preferred Orientation (LPO) of anisotropic minerals or the Shaped Preferred Orientation (SPO) of isotropic materials with distinct shapes (Crampin and Booth 1985; Nicolas and Christensen 1987; Silver

1996; Montagner and Guillot 2002; Long and Becker 2010). Both LPO and SPO can cause substantial variations in wave velocities (up to approximately 10%) depending on the polarization or propagation direction of seismic waves (Becker 2020). These orientation-induced velocity differences can exceed those attributed to variations in composition or temperature within the Earth. A number of measurements have been proposed to reveal seismic anisotropy that can be categorized into body wave and surface wave methods. Body wave methods include shear wave splitting, receiver function, and P wave arrival methods (Hess 1964; Silver and Chan 1991; Savage 1999; Park and Levin 2002; Bokelmann 2002). Shear wave splitting, which utilizes both regional S arrivals to constrain crustal stress field variations and teleseismic arrivals like SKS or SKKS core phases to study anisotropy in the lithosphere and mantle, offers good lateral resolution ($< 50\ km$) but faces fundamental limitations in depth resolution as it cannot directly constrain the depth of anisotropic layers (Savage 1999). Receiver functions, on the other hand, offer evidence of crustal anisotropy through split pS conversions. P wave arrival methods usually contain Pn azimuthal anisotropy and P wave tomography. Hess (1964) first uses Pn wave to detect seismic anisotropy in the upper mantle. P wave tomography can also be used to infer mantle lithospheric anisotropy, though this requires knowledge of isotropic velocity variations (Becker 2020).

Surface wave methods offer complementary information about seismic anisotropy, particularly in terms of depth resolution. It is commonly used to study upper mantle anisotropy. Ambient noise tomography has extended the applications to shorter periods allowing for better resolution of crustal anisotropy. Furthermore, long-period surface waves and free oscillations of the Earth provide invaluable insights into deep Earth structure, including inner core anisotropy (Tromp 2001). Surface wave studies can provide constraints on both radial and azimuthal anisotropy, with the ability to construct 3-D models of anisotropy due to their dispersive character. The observation that Love waves generally travel faster than Rayleigh waves implies the existence of radial anisotropy in the upper mantle (Anderson 1965). This discovery led to the development of transversely isotropic models with vertical symmetry axes, which formed the basis for the Preliminary Reference Earth Model (PREM) (Dziewonski and Anderson 1981). In addition to radial anisotropy, surface waves also exhibit azimuthal anisotropy, where wave propagation velocities vary with azimuth relative to north (Forsyth 1975). The azimuthal variation of the Rayleigh and Love wave phase velocity is defined as (Smith and Dahlen 1973):

$$c(\omega,\theta) = A_0(\omega) + A_1(\omega)cos(2\theta) + A_2(\omega)sin(2\theta) + A_3(\omega)cos(4\theta) + A_4(\omega)sin(4\theta), \quad (1.1)$$

where θ is the azimuth of the wavenumber, ω is the angular frequency, and the A_i terms depend on the anisotropic elastic parameters within the medium. Note that both the phase velocity c and coefficients A_i implicitly depend on the spatial location $\mathbf{x} = (x, y)$, as the elastic properties of the medium vary with position. Romanowicz (2002) reviews the research on azimuthal anisotropy in the crust and upper mantle using this expression. Based on azimuthal anisotropy studies, Yuan and Romanowicz (2010) identify two distinct lithospheric layers throughout the stable North American continent, providing new insights into the formation and stability of cratons. Beghein et al. (2014) present the Gutenberg (G) discontinuity is primarily associated with vertical changes in azimuthal anisotropy. These applications have proved that surface wave azimuthal anisotropy is a promising way to detect layering. While their lateral resolution is generally lower than that of body wave methods, surface wave studies have been crucial in developing global and regional models of anisotropy, enabling more powerful petrological and geodynamic inferences about Earth's interior.

1.1.5 Eikonal tomography

Surface wave tomography is typically performed using either 'single-station' or 'array-based' methods. The single-station methods focus on measuring the traveltime of surface wave between seismic sources and receivers and utilize ray theory or finite frequency kernels to evaluate the wave speeds (Ekström et al. 1997; Yoshizawa and Kennett 2002; Yoshizawa and Kennett 2005). While the array-based methods focus on analyzing the phase differences from the waveform recorded across an array of stations to determine the dispersion characteristics, enabling high resolution regional tomography or imaging of small structures (Friederich 1998; Prindle and Tanimoto 2006). Compared to single-station methods, array-based methods have the advantage of enhancing the signal-to-noise ratio (SNR) of coherent signals over incoherent noise (Rost and Thomas 2009). With the development of dense arrays and seismic instruments, numerous imaging methods have been developed to improve the spatial resolution of crustal and upper mantle structure, such as beamforming tomography (Boué et al. 2014; Nakata 2016; Wu et al. 2023), wave gradiometry (Langston 2007; Liang and Langston 2009; De Ridder and Biondi 2015) and eikonal and Helmholtz tomography (Lin et al. 2009; Lin and Ritzwoller 2011; Jin

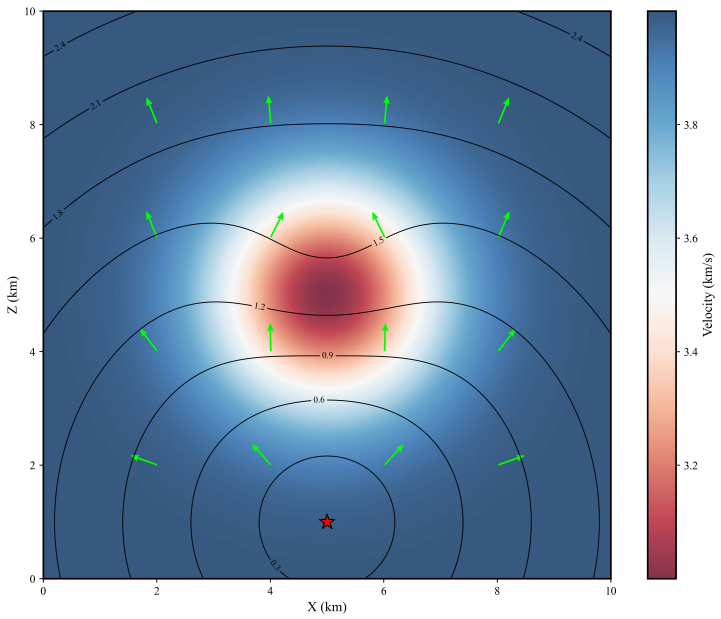


Figure 1.1: Traveltime surface and velocity relationship. A traveltime surface (or wavefront) is a surface in space that represents all points that a wave reaches at the same traveltime from a source. The surface is shaped by the velocity structure of the medium, where regions of lower velocity result in longer traveltimes (and regions of higher velocity result in shorter traveltimes). The red star indicates the source. The contours represent wavefronts, which are lines of constant traveltime. The green arrows indicate the direction of wave propagation, which is perpendicular to the wavefronts and follows the gradient of the traveltime surface.

and Gaherty 2015), which are widely used in analyzing surface wave signals recorded at large dense seismic networks. The detailed methods and applications of array seismology are reviewed in Rost and Thomas (2002) and Schweitzer et al. (2012).

Eikonal tomography is a seismic imaging technique that utilizes the eikonal equation to study the relationship between seismic wave traveltimes and subsurface velocity structures. In traditional methods, if the effects of high frequency components and the directionality of scattering are neglected, the traveltime of surface waves can be defined as (Lin et al. 2009):

$$\tau(\mathbf{x}_r, \mathbf{x}_s) = \int S(\mathbf{x}, \mathbf{x}_r, \mathbf{x}_s) \frac{dx^2}{c(\mathbf{x})}, \tag{1.2}$$

where $\tau(\mathbf{x}_r, \mathbf{x}_s) = \tau(\mathbf{x}_r, \mathbf{x}_s, \omega)$ is the frequency dependent phase traveltime between source $\mathbf{x}_s$ and receiver $\mathbf{x}_r$ at frequency ω, $S(\mathbf{x}, \mathbf{x}_r, \mathbf{x}_s)$ is the finite frequency sensitivity integral kernel, and $\mathbf{x}$ represent an arbitrary point, $c(\mathbf{x}) = c(\mathbf{x}, \omega)$ is the frequency dependent phase velocity at location $\mathbf{x}$ and frequency ω. Equation 1.2 links the surface wave traveltime to wave speed, but the traveltime acts as a global constraint.

In contrast, eikonal tomography utilize the eikonal equation to directly obtain the local phase velocities from interpolated phase traveltime surfaces (Lin et al. 2009). Figure 1.1 shows the concept of traveltime surface, which represents the time taken for a seismic wave to propagate from a source point to every location within a medium. By analyzing the traveltime surface and its gradient, this method captures variations in wave propagation caused by velocity anomalies. In regions with constant velocity, the traveltime surface forms concentric circular contours, with the gradient direction perpendicular to the contours. In low-velocity anomaly regions, the contours become denser, bend away from the source, and exhibit increased gradient magnitudes, reflecting slower wave propagation. These changes in the traveltime surface and its gradient provide direct insights into the bending of seismic ray paths in heterogeneous media, enabling the reconstruction of subsurface velocity models through the eikonal equation.

The eikonal equation governs wavefront propagation in heterogeneous media and is derived by applying the high-frequency approximation to the wave equation. For P-waves traveling through heterogeneous media, the scalar potential ϕ obeys the wave equation (Shearer 2019):

$$\nabla^2\phi - \frac{1}{\alpha^2(\mathbf{x})}\frac{\partial^2\phi}{\partial t^2} = 0, \tag{1.3}$$

where ϕ is the scalar for displacement $\mathbf{u}$ ($\mathbf{u} = \nabla\phi$), $\alpha(\mathbf{x})$ is P wave velocity at point $\mathbf{x}$. The solution to this equation in the high-frequency domain assumes a harmonic form:

$$\phi(\mathbf{x}, t) = A(\mathbf{x})e^{-i\omega[t-\tau(\mathbf{x})]}, \tag{1.4}$$

where $A(\mathbf{x})$ represents the amplitude function, $\tau(\mathbf{x})$ is the wavefront arrival time and ω is the angular frequency (with $\omega \gg 1$ for the high-frequency approximation). Substitution of this solution (Equation 1.4) into the wave equation (Equation 1.3) and subsequent differentiation yields:

$$\nabla^2 A - \omega^2 A|\nabla\tau|^2 - i\left[2\omega\nabla A \cdot \nabla T + \omega A\nabla^2 T\right] = -\frac{A\omega^2}{\alpha^2}, \tag{1.5}$$

Separation of the real and imaginary parts in Equation 1.5 produces two equations. The imaginary part corresponds to the amplitude transport equation, while the real part yields the

Helmholtz equation:

$$|\nabla\tau(\mathbf{x})|^2 = \frac{1}{\alpha^2(\mathbf{x})} + \frac{\nabla^2 A(\mathbf{x})}{A(\mathbf{x})\omega^2}, \tag{1.6}$$

At high frequencies ($\omega \gg 1$) or when the spatial variation of the amplitude is small compared to the traveltime gradients, the second term on the right side becomes negligible. This simplification leads to the eikonal equation (Aki and Richards 2002; Shearer 2019):

$$|\nabla\tau(\mathbf{x})|^2 = \frac{1}{\alpha^2(\mathbf{x})}, \tag{1.7}$$

While this equation is derived specifically for P-waves, the same mathematical form applies to other wave types. For surface wave propagating in a heterogeneous medium, the eikonal equation can be written in this form (Wielandt 1993; Lin et al. 2009):

$$\frac{\hat{k_i}}{c_i(\mathbf{x})} = \nabla\tau(\mathbf{x}_i, \mathbf{x}), \tag{1.8}$$

where $\hat{k}_i$ is the unit wave number vector for traveltime surface i at location $\mathbf{x}$, $c_i(\mathbf{x}) = c_i(\mathbf{x}, \omega)$ is the wave speeds for traveltime surface i at location $\mathbf{x}$ and frequency ω, $\nabla\tau(\mathbf{x}_i, \mathbf{x}) = \nabla\tau(\mathbf{x}_i, \mathbf{x}, \omega)$ denotes the gradient of traveltime surface i at location $\mathbf{x}$ and frequency ω. The process of eikonal tomography involves two main steps: wavefront or phase front tracking and gradient computation based on the eikonal equation at each spatial node. By compiling and averaging the local phase speeds and wave path directions from the traveltime surfaces centered on each station (or earthquake), a comprehensive phase velocity distribution map is constructed. Eikonal tomography avoids explicit regularization by replacing traditional ad hoc smoothing parameters with implicit regularization through phase-front tracking (Lin et al. 2009). Traveltime data are interpolated using minimum-curvature surface fitting (Smith and Wessel 1990), which naturally smooths gradients while preserving structural trends, and statistical averaging over multiple paths further stabilizes results.

Based on a high frequency analysis when entering an elliptical definition of c into the Helmholtz equation, De Ridder et al. (2015) proposed an elliptical-anisotropic eikonal equation:

$$1 = \begin{bmatrix} \partial_x \tau & \partial_y \tau \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \partial_x \tau \\ \partial_y \tau \end{bmatrix}. \tag{1.9}$$

where τ represents the traveltime field, and its spatial derivatives $\partial_x \tau$ and $\partial_y \tau$ describe the components of the slowness (the reciprocal of velocity) vector. The matrix elements are expressed as:

$$M_{11} = M_{11}(\mathbf{x}, \omega) = (c_f^2 - c_s^2) sin^2(\alpha) + c_s^2, \tag{1.10}$$

$$M_{12} = M_{21} = M_{12}(\mathbf{x}, \omega) = (c_f^2 - c_s^2) sin(\alpha) cos(\alpha), \tag{1.11}$$

$$M_{22} = M_{22}(\mathbf{x}, \omega) = (c_f^2 - c_s^2) cos^2(\alpha) + c_s^2. \tag{1.12}$$

where $c_f = c_f(\mathbf{x}, \omega)$ and $c_s = c_s(\mathbf{x}, \omega)$ are the fast and slow velocities, respectively, and $\alpha = \alpha(\mathbf{x}, \omega)$ is the azimuth of the fast direction. In this way, the eigenvalues and eigenvectors of the matrix $\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$ indicate the fast and slow velocities and directions, which enables explicit spatial regularization during inversion (De Ridder et al. 2015).

Eikonal tomography complements traditional surface wave tomography by avoiding the forward and inversion processes, accounting for bent rays and directly measuring azimuth-dependent phase velocities at each point without making ad hoc assumptions. Eikonal tomography has been applied to both regional and continental scale arrays and has become a promising method to investigate the crust and mantle structure (De Ridder 2011; Qiu et al. 2019; Tong 2021; Hao et al. 2021; Liu and Tong 2021; Kästle et al. 2022).

1.2 Physics informed machine learning

1.2.1 Machine learning for seismology

Seismology relies on observations to study the internal structures and properties of Earth and other planets. Through analysis of these observations, seismologists can investigate natural

phenomena such as faults and volcanoes, as well as the impact of human activities. With the dramatic increase in observational data (e.g. volume) and advancements in computational power, machine learning (ML) has emerged as a vital tool in seismological research, primarily leveraging supervised and unsupervised learning techniques. The detailed ML applications in seismology have been discussed in multiple review papers (Bergen et al. 2019; Kong et al. 2019; Yu and Ma 2021; Mousavi and Beroza 2022; Li et al. 2023; Zhao et al. 2024). In the following, I summarize main points relevant for my work.

Supervised learning is a type of machine learning where a model is trained on a labeled dataset, meaning that each training example includes both input features and the correct output (label). The main task of supervised learning is to learn a mapping function from input features to output labels, so that the model can predict the correct new output. Supervised learning tasks encompass classification and regression, with commonly used algorithms in seismology including logistic regression (LR), support vector machines (SVM), random forests (RF), and neural network (NN). LR is a statistical model used for binary classification that provides good interpretability and computational efficiency but may struggle with non-linear relationships commonly encountered in seismic problem. SVM uses kernels functions to separate data classes by maximizing margins that excels in handling high-dimensional data, though kernel selection can be challenging. RF is an ensemble of decision trees for classification or regression that is robust to noise and outliers and handles both numerical and categorical features well, but requires significant memory for large datasets. NN is a computational model inspired by biological neurons that offers flexible architecture for complex patterns and can automatically learn features, but it has a complex training process and needs careful parameter tuning. These distinct characteristics are suitable for different seismic tasks. For example, Xu et al. (2012) compare bivariate statistics (BS), LR, artificial neural networks (ANNs), and three kinds of SVM models for susceptibility mapping of earthquake-triggered landslides in a tributary watershed of the Fu River affected by the 2008 Wenchuan earthquake, finding that LR models provided the highest success and prediction rates. Hibert et al. (2017) validate the performance of the RF algorithm in classifying seismic signals recorded at Piton de la Fournaise volcano, distinguishing rockfalls and volcano-tectonic earthquakes with sensitivity exceeding 99% when sufficient training samples are available and remaining above 90% even with limited samples. Paitz et al. (2018) use ANN to identify seismic time series suitable for noise tomography, achieving better performance compared to manually classification with errors below 20%. The high accuracy achieved across

these studies demonstrates the potential of supervised learning in automating, modeling and inversion, which significantly improves traditional seismological analysis tasks.

Unsupervised learning, on the other hand, involves training a model on a dataset that contains only input features and no labels. The main task of unsupervised learning is to discover hidden patterns or data groupings within the dataset, without any external guidance on what those patterns should be. Unsupervised learning tasks contain clustering and dimensionality reduction, with commonly used algorithms in seismology including k-means, self-organizing map (SOM), principal component analysis (PCA) and dictionary learning. k-means is a simple and efficient clustering algorithm that partitions data into k clusters but requires pre-specified cluster numbers and is sensitive to initial values. SOM is a NN-based algorithm with topological preservation and adaptive learning capabilities though it needs careful parameter tuning and can be computationally intensive. PCA projects data onto lower-dimensional orthogonal axes that offers linear dimensionality reduction and noise filtering while maintaining computational efficiency, but may miss important non-linear relationships and requires an understanding of the data for result interpretation. Dictionary learning represents data sparsely via learned basis functions that enables adaptive sparse representation of signals and effective noise suppression, though it demands significant computational resources and careful dictionary design. In seismic applications, Galvis et al. (2017) employ k-means to detect and classify surface waves in seismic data acquired in Colombia, presenting well-differentiated zones that improves seismic interpretation. Roden et al. (2015) utilize PCA to determine the most significant seismic attributes, which are then used as inputs for SOM training, revealing previous unidentified geological features. Nazari Siahsar et al. (2017) propose a data-driven 3-D dictionary learning algorithm with multitasking strategy for seismic data processing, achieving efficient and effective random noise attenuation of synthetic and real 3-D data when compared to several established filtering methods. These studies highlight how unsupervised learning can reveal hidden patterns and structures in seismic data that might be missed by traditional analysis methods, particularly valuable when dealing with the growing volume of observational data.

NNs are computational models inspired by the human brain's neural structure. A standard NN consists of interconnected nodes (neurons) organized in layers: an input layer, one or more hidden layers, and an output layer (Figure 1.2). Each connection between neurons carries a weight that is adjusted during training to optimize the network's performance. These weights

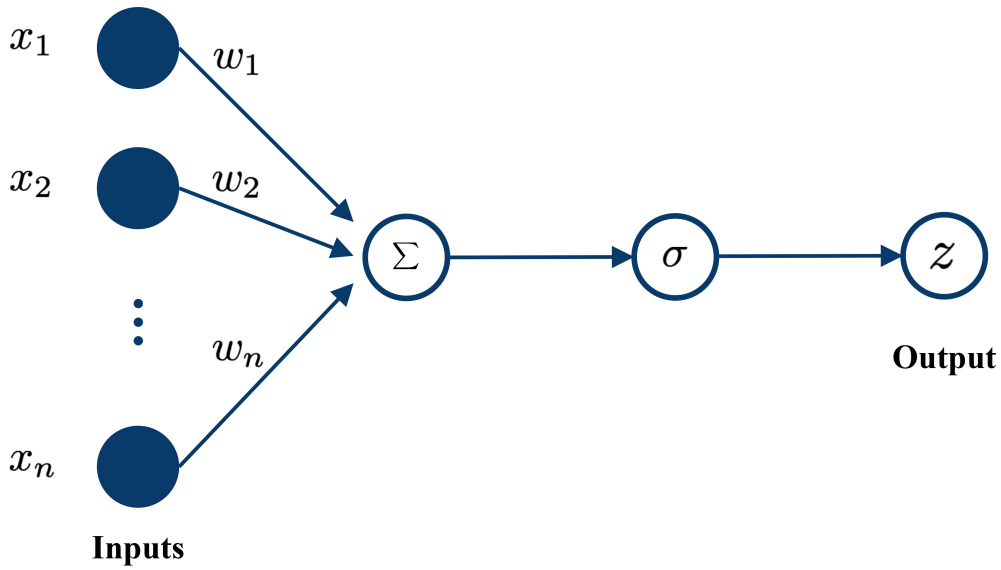


Figure 1.2: Basic structure diagram of neural networks (modified from Bishop (1994)). $x_1, x_2, ..., x_n$ are inputs, $w_1, w_2, ..., w_n$ are weights, $\sum$ represents the weighted sum of inputs, σ represents non-linear activation function and z is output.

are typically updated using automatic differentiation, a computational technique that efficiently calculates derivatives by tracking operations during forward computation, enabling the network to learn from its errors. In a feedforward NN, information flows unidirectionally from input to output. Each neuron performs two operations: (1) computing a weighted sum of its inputs and a bias term, and (2) applying a non-linear activation function to this sum. Mathematically,

$$z = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right), \tag{1.13}$$

where w_i represents the weight and b represents the bias. A weighted sum of the inputs plus a bias is then transformed using a non-linear activation function σ to give a final output z (Bishop 1994). NNs are trained using training data, which is carefully selected examples that represent the problem space, allowing the model to learn patterns and relationships. During training, the network's predictions are compared to the actual expected outputs, and the difference is quantified as a loss, which is a numerical measure of prediction error that guides the weight adjustment process. This architecture forms the basis of deep learning, where multiple hidden layers enable hierarchical feature learning. Hinton and Salakhutdinov (2006) extend the classical ANN by integrating numerous hidden layers to construct deep neural networks (DNNs) that are capable

of extracting complex features from massive amounts of data. This approach, which involves training DNNs to learn high-level abstractions from data, is known as deep learning (DL). By leveraging the power of multiple layers of non-linear processing units, DL models can automatically discover and represent complex patterns and hierarchical structures in the data (LeCun et al. 2015). Deep learning algorithms also encompass both supervised and unsupervised tasks with commonly used architectures such as fully connected neural network (FCNN), convolutional neural network (CNN), recurrent neural network (RNN) and generative adversarial network (GAN). A FCNN consists of multiple layers of neurons where each neuron in one layer is connected to every neuron in the next layer, providing a simple and intuitive NN architecture. A CNN typically consists of convolutional layers, pooling layers, and fully-connected layers. It significantly reduces the number of parameters by featuring parameter sharing and sparse connections. RNNs have a loop in their architecture that allows them to maintain a hidden state over time, offering a strong ability in handling sequence data. A GAN is composed of a generator and a discriminator that are trained simultaneously in an adversarial manner, enabling the remarkable ability to generate new data that resembles the real data distribution. Though these DNN architectures share the advantages of strong non-linear fitting capabilities and wide application adaptability, they also face issues such as high computational resource requirements, over-fitting risks, and poor model interpretability.

The DL applications in seismology can be discussed from several aspects, such as seismic denoising, arrival time picking and earthquake location. Zhu et al. (2019) propose a denoising/decomposition method called DeepDenoiser based on a DNN, which simultaneously learns a sparse representation of data in the time-frequency domain and maps it into masks for seismic signal and noise separation. By automatically learning noise statistics from data, DeepDenoiser significantly improves SNR and generalizes well to real noisy data sets. A challenge in DL based seismic denoising is to obtain clean seismic data or noise and an effective solution is create semisynthetic white noise. Wu et al. (2019) use a modified denoising CNN with variational mode decomposition (VMD) that simulates the user-generated white noise, presenting effectively rejection on white noise and migration artifacts. Unsupervised learning such as GAN can also be used for noise attenuation by learning the domain mapping from noisy data domain to effective signal data domain (Li et al. 2020). With the ability of DNNs to identify predefined targets in images, and given the vast amount of seismic data along with manually labeled examples, seismic arrival picking has become a popular application of DL in seismic data

processing. Zhu and Beroza (2019) apply U-Net, a kind of CNN, to identify the first arrivals of P-waves and S-waves based on training data from the Northern California Seismic Network catalog, enabling automatic seismic phase picking, particularly effective in reliably extracting S-wave arrivals. Yuan et al. (2020) combine CNN and RNN for first-arrival picking that uses an RNN to optimize picking after U-net segmentation, improving the picking performance and reducing human intervention. For earthquake location determination, Zhang et al. (2020b) utilize DNN to determine the locations of small earthquakes in Oklahoma and are able to locate small event ($ML \geq 2.0$) with a mean epicenter error of 4-6 km. Mousavi and Beroza (2020) present a single-station earthquake location method based on Bayesian neural network that is capable of estimating the epicenter, origin time, and depth along with their uncertainties. Other DL applications in seismology include solving seismic forward and inverse problems (DeVries et al. 2018; Araya-Polo et al. 2018; Das et al. 2019; Gatti and Clouteau 2020; Münchmeyer et al. 2021; Jianguo and Ntibahanana 2024). DL has already shown great potential in different seismic tasks and will likely soon become a routine approach for various seismic data analyses.

The evolution of ML applications in seismology demonstrates a clear progression from traditional methods to more sophisticated approaches. Traditional supervised learning methods like LR, SVM, and RF offer reliable performance with good interpretability for specific tasks, while unsupervised learning techniques enable the discovery of hidden patterns in unlabeled seismic data. The emergence of DL has further transformed seismic data analysis by providing powerful tools for handling complex, large-scale datasets. As each ML method has its own advantages and disadvantages, careful consideration of task requirements, data characteristics, and desired level of model interpretability is essential when selecting a specific approach. These methods have revolutionized seismological research by enabling automatic analysis of large datasets, solving forward and inverse problems, and discovering new patterns or relationships, thereby enhancing our understanding of Earth's structure and dynamics.

1.2.2 Scientific machine learning

As a purely data-driven method, traditional ML faces several fundamental challenges. First, it heavily relies on large-scale datasets for model training, which presents significant barriers for complex physical systems where data collection is costly and limited. Second, the design of NN architectures lacks solid theoretical foundations, relying primarily on empirical experimentation rather than systematic principles. Third, the internal decision-making process of DNN

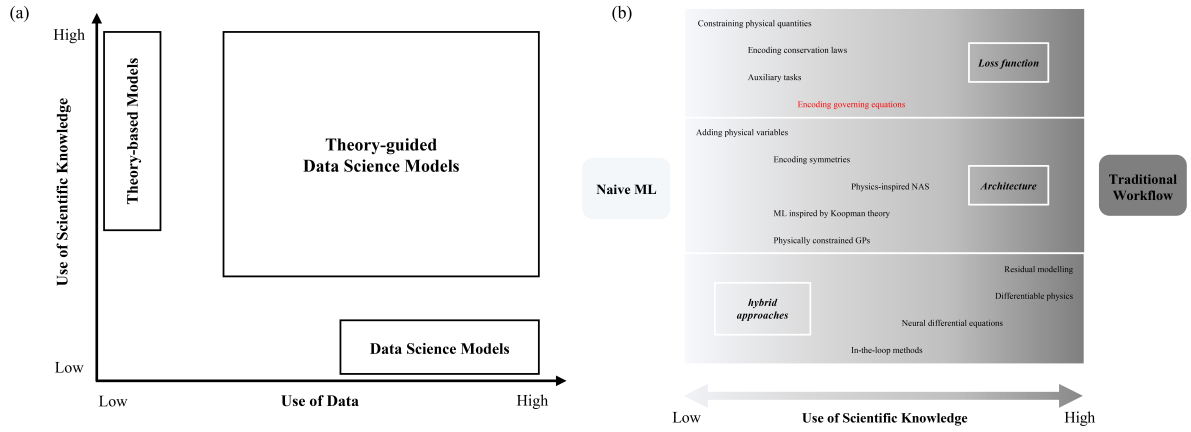


Figure 1.3: (a) A representation of balance between data and scientific knowledge (modified from Karpatne et al. (2017)). The x-axis indicates the use of data while the y-axis indicates the use of scientific knowledge. Theory-based models rely heavily on scientific principles with minimal data, while data science models use extensive data with less emphasis on theory. Theory-guided data science models integrate both, enhancing scientific understanding by combining data-driven approaches with theoretical knowledge. (b) Illustration of SciML methods (modified from Moseley (2022)). The three colored rectangular blocks represent different ways of integrating scientific knowledge into ML models. Each block contains specific implementation methods. The color intensity (along the x-axis) indicates the degree of scientific knowledge constraint imposed by each approach. Note that the scientific constraint's strength is a rather vague concept. Here the methods closer to the traditional workflow enforce stronger physical principles.

functions as a "black-box", making it difficult to understand and validate their predictions. Fourth, despite their strong representational power, these data-based models typically exhibit poor generalization capability beyond their training distribution (Lazer et al. 2014). Before the development of data-driven approaches, model-driven approaches (or physics-based models) served as the primary framework for understanding physical systems. These approaches construct mathematical representations based on fundamental physical laws, typically expressed through partial differential equations (PDEs). They transform conceptual understanding into quantifiable frameworks, ranging from simple process models to complex multi-system integrations. The key advantages of model-driven approaches include their mechanistic interpretability, explicit incorporation of physical principles, and reduced reliance on large datasets. However, purely model-driven method also faces limitations. Simpler models may oversimplify complex phenomena and introduce significant errors, while more sophisticated models often become computationally intractable. The inherent complexity of real-world systems, combined with our limited understanding of relationships and trade-offs, makes it difficult to fully capture the relationships between various factors and their dependencies. Additionally, these models

struggle with inverse problems, particularly in complex geometries or high-dimensional spaces, which can significantly limit their practical applications.

Given these limitations, new methods are needed that can learn from limited or noisy data while incorporating physical knowledge to handle the complexity of dynamic Earth systems. As shown in Figure 1.3 (a), theory-based and data science models represent the two extremes of knowledge discovery and neither is sufficient for tackling complex scientific problems independently. A better way is to combine the strengths of both through theory-guided data science models. To address this, a new and rapidly advancing field known as scientific machine learning (SciML) has established (Baker et al. 2019; Karniadakis et al. 2021; Willard et al. 2022). SciML aims to integrate existing scientific knowledge into ML algorithms, generating more sophisticated algorithms that can leverage our prior scientific understanding while maintaining the powerful learning capabilities of ML approaches. This integration is achieved through loss function modification, architecture design or hybrid approaches (Moseley 2022). Figure 1.3 (b) presents the specific SciML methods and illustrates the different strengths to which these methods impose scientific principles into ML algorithm. For loss function modification methods, additional physical constraint terms are added to the loss function, serving as regularizers or physical priors. These constraints guide models toward physically consistent solutions while maintaining learning flexibility (Beucler et al. 2021; Daw et al. 2022; Raissi et al. 2019). For architecture design methods, traditional "black-box" architectures are transformed into physics-informed systems by embedding scientific principles directly into their design. This introduces strong inductive biases, narrowing the model's hypothesis space (Daw et al. 2020; Panju and Ghodsi 2020; Wang et al. 2020b). For hybrid methods, ML algorithms are integrated with traditional scientific methods to create hybrid systems that simultaneously leverage the adaptability of data-driven learning and established physical knowledge (Chen et al. 2018; Pawar et al. 2021; Um et al. 2020). These methods encompass varying degrees of scientific constraints, and there is no single best approach for incorporating underlying theories. By embedding scientific principles, SciML creates more interpretable, generalizable, and efficient models that address the challenges posed by data limitations and complex physical equations.

1.2.3 Physics informed neural network

Physics-informed neural network (PINNs), introduced by Raissi et al. (2019), are designed to solve various PDEs using DL techniques. PINNs represent a novel type of SciML framework,

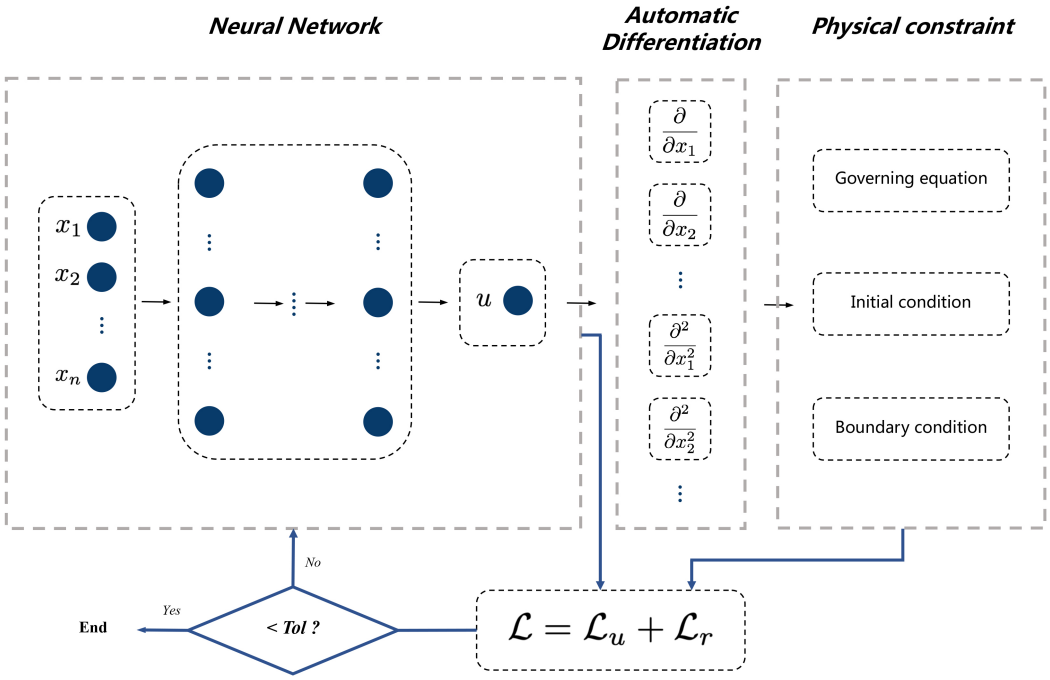


Figure 1.4: Schematic of PINN for solving a general PDE. Inputs $x_1, x_2, ..., x_n$ are processed by neural network $N(\theta; x)$, producing the predicted solution u. Automatic differentiation is used to compute different order of derivatives $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, ..., \frac{\partial^2}{\partial x_1^2}, \frac{\partial^2}{\partial x_2^2}, ...$, which is essential for enforcing physical constraints. In the PINN algorithm, the loss function $\mathcal{L}$ is composed of two parts: the loss of data $\mathcal{L}_u$ and the loss of physical constraint $\mathcal{L}_r$. This physical constraint contains the governing equation along with appropriate initial and boundary conditions.

which embedding the governing equations of a system into the training process (see 'Encoding governing equation' in Figure 1.3 (b)). Rather than simply incorporating constraints on a system's physical quantities into the loss function, integrating knowledge of underlying governing equation is more effective (Moseley 2022). PINNs leverage the automatic differentiation to incorporate the differential form constraints from PDEs directly into the loss function, thereby merging data- and physics-based models (Baydin et al. 2018). This approach functions as a mesh-free technique, which uses NNs to directly approximate functions at arbitrary points without the need for spatial discretization. A generic schematic illustrating the network structure of PINN is shown in Figure 1.4. Consider $u(x)$ is the unknown solution in a PDE, a PINN is a NN, $N(\theta; x)$, with trainable parameters θ to approximate the solution $u(x)$. The loss function consists of two components: the data loss $\mathcal{L}_u$, which measures the discrepancy between the NN's predictions and available observations, and the physics loss $\mathcal{L}_r$, which ensures that the NN's predictions respect the governing PDEs. This is achieved by penalizing the residuals of the PDE at collocation points distributed across the domain. The loss function is given by:

$$\mathcal{L} = \mathcal{L}_u + \mathcal{L}_r, \tag{1.14}$$

where

$$\mathcal{L}_u = \sum_{i}^{N_d} ||N(\theta; x_i) - u(x_i)||^2, \tag{1.15}$$

$$\mathcal{L}_r = \sum_{j}^{N_p} ||r(x_j)||^2, \tag{1.16}$$

where N_d and N_p are the numbers of data and physics points. $r(x_j)$ represents the residual of PDE at a set of points x_j over the entire domain.

By incorporating physical laws into the NN architecture, PINNs offer several advantages over both purely data-driven and model-driven approaches. At the foundational level, PINNs achieve better generalization and provide a more interpretable framework compared to conventional NNs, as their solutions naturally respect underlying physical principles. From a computational perspective, their mesh-free nature yields continuous, differentiable solutions without requiring discretization schemes, significantly simplifying the modeling process. Furthermore, PINNs demonstrate remarkable data efficiency, requiring substantially less training data than traditional NNs while enabling effective extrapolation beyond the given points. Finally, PINNs offer a unified framework for solving both forward and inverse problems, and the same architecture can be adapted with minimal modifications to handle both types of problems, even in cases involving complex geometries or high-dimensional spaces where traditional numerical methods struggle. These characteristics make PINNs particularly valuable for practical applications in scientific computing.

The success of the basic PINN framework has generated numerous methodological extensions, each addressing specific challenges in physical modeling. These extensions can be broadly categorized into two groups: physics-oriented variants and NN architectural adaptations. The physics-oriented variants encompass conservative PINNs (cPINNs) (Jagtap et al. 2020), fractional PINNs (fPINNs) (Pang et al. 2019), nonlocal PINNs (Pang et al. 2020), variational PINNs (vPINNs) (Rojas et al. 2024) and extended PINNs (XPINNs) (Jagtap and Karniadakis

2020). The NN architectural adaptations include Bayesian neural networks (BNNs) (Yang et al. 2021), CNNs (Fang 2021; Gao et al. 2021), RNNs (Viana et al. 2021; Lu et al. 2024), long short-term memory (LSTM) networks (Zhang et al. 2020a), GAN (Yang et al. 2020; Yang and Ma 2023). The diversity of these extensions reflects both the potential and challenges of PINNs, suggesting a trend toward more specialized and sophisticated implementations while maintaining the core philosophy of combining ML and physics-based models. Researchers have applied PINNs to diverse fields, including fluid mechanics (Raissi et al. 2020; Mao et al. 2020; Cai et al. 2021), solid mechanics (Goswami et al. 2020; Tao et al. 2020), material science (Fang and Zhan 2019; Chen et al. 2020b; Islam et al. 2021), biomedicine (Kissas et al. 2020; Sun et al. 2020) and power systems (Misyris et al. 2020). In addition, several software packages have been developed to implement efficient PINNs training based on different framework, for example, DeepXDE (Lu et al. 2021), NeuroDiffEq (Chen et al. 2020a), NVIDIA SimNET (Hennigh et al. 2021) and SciANN (Haghighat and Juanes 2021). These frameworks facilitate the practical implementation and broader adoption of PINNs. A comprehensive literature review on PINNs can be found in Karniadakis et al. (2021) and Cuomo et al. (2022).

In seismology, PINNs have demonstrated remarkable potential in various applications, addressing both forward and inverse problems. For forward problems, PINN has been applied to solve isotropic and anisotropic eikonal equation (Smith et al. 2020; Waheed et al. 2021; Waheed et al. 2020; Taufik et al. 2022) and wave equation in the time and frequency domains for isotropic and anisotropic media (Alkhalifah et al. 2020; Karimpouli and Tahmasebi 2020; Moseley et al. 2020a; Moseley et al. 2020b; Song et al. 2021; Huang and Alkhalifah 2022a; Huang and Alkhalifah 2022b; Song et al. 2022; Song and Wang 2023; Chai et al. 2024; Alkhalifah and Huang 2024; Zou et al. 2024). For inverse problems, PINNs have been applied to traveltime tomography (Waheed et al. 2021; Agata et al. 2023; Gou et al. 2023; Song et al. 2024), wavefield reconstruction inversion (Song and Alkhalifah 2021) and full waveform inversion (Rasht-Behesht et al. 2022; Yang and Ma 2023; Lu et al. 2024). During these applications, several key methodological advances have enhanced PINN's performance. Alkhalifah et al. (2020) introduced frequency-domain scattered wavefield simulation to avoid point-source singularity challenges. Song et al. (2022) developed adaptive sinusoidal activation functions to achieve faster convergence during training. Huang and Alkhalifah (2022a) employed frequency upscaling and neuron splitting to model high-frequency wavefields. For multi-frequency wavefield solutions, Huang and Alkhalifah (2022b) incorporated a single reference frequency into the loss function, while Song and

Alkhalifah (2021) introduced Fourier feature PINNs. Moseley et al. (2020b) and Moseley et al. (2020a) proposed "curriculum learning" for training efficiency and developed finite basis PINN (FBPINN) for large computational domains. Agata et al. (2023) and Gou et al. (2023) utilized Bayesian PINNs to estimate uncertainty in seismic tomography. For full waveform inversion, Yang and Ma (2023) proposed physics-informed GAN to handle inversions without requiring training datasets. Alkhalifah and Huang (2024) developed physics-informed neural wavefields that models the wavefield as linear combinations of Gabor basis functions governed by the wave equation, improving the efficiency and accuracy of NN solutions. Zou et al. (2024) proposed a velocity-encoded PINN (VE-PINN) that introduces feature parameters to represent different layered velocity models, enabling generalization across various initial conditions.

While PINNs have demonstrated remarkable achievements in seismological applications, they also face significant challenges in transitioning from theory to practice. First, the scalability issue, where computational resource requirements grow exponentially when handling large-scale, complex seismic data, limiting their practical applicability. Second, the optimization challenge of network architecture and training workflows, where different seismological problems require different network structures and parameter settings, yet systematic configuration standards are lacking. Third, the balance between model complexity and interpretability, where despite enhanced interpretability through physics incorporation, a better balance between the "black-box" nature of NNs and the transparency of traditional methods is still needed. Addressing these challenges, future research should focus on three directions: developing robust methods for complex geological environments, establishing standardized practical guidelines, and improving uncertainty quantification for field applications. Through these efforts, PINNs are expected to successfully transition from theoretical achievements to practical applications, ultimately playing a more significant role in seismological research and applications.

1.3 Geological background

The northeastern Tibetan Plateau, characterized by the interaction between the stable Ordos Block, North China Craton, and the actively deforming Tibetan Plateau, preserves critical evidence of the ongoing India-Asia continental collision. The dense seismic networks deployed across the region provide excellent data coverage, while its well-documented geological complexity allows us to rigorously test our method's resolution capabilities. Furthermore, the wealth

of previous seismic studies using conventional approaches provides a robust framework for evaluating our results, making northeastern Tibet particularly well-suited for both demonstrating our methodological innovations and advancing our understanding of continental dynamics.

1.3.1 Crust-mantle structure of northeastern Tibetan Plateau

The Tibetan Plateau (Figure 1.5), with its complex and intense tectonic activities, provides one of the most ideal environments on Earth to study continental crust-mantle deformation. However, the tectonic evolution of the Tibetan Plateau, especially the mechanisms of its uplift and deformation, remains debated. There are several dynamical models proposed to explain these mechanisms (Figure 1.6), including underthrusting, oblique continental subduction, convective removal of mantle lithosphere, lower crustal flow, slab tearing/break-off and lithosphere delamination. Among these, the two most representative models are the rigid block extrusion model (Molnar and Tapponnier 1975) and the lower crustal flow model (Royden et al. 1997). The rigid block extrusion model suggests that internal continental deformation relates to plate tectonic theory, primarily through the southward oblique subduction of blocks along ancient sutures, leading to the formation of large strike-slip faults and crustal shortening (Molnar and Tapponnier 1975; Tapponnier et al. 2001). The lower crustal flow model suggests that the uplift and crustal thickening of the Tibetan Plateau results in a low-strength, low-viscosity layer in the middle-lower crust, which flows outward under pressure (Royden et al. 1997; Clark and Royden 2000; Royden et al. 2008). This model indicates that the middle-lower crustal flow divides into two branches: one flowing toward the southeastern margin of the Tibetan Plateau, and the other toward its northeastern margin (Clark et al. 2005).

The northeastern margin of the Tibetan Plateau is located at the junction of the Alxa Block, Ordos Block, western Qinling Orogen, Sichuan Basin, Songpan-Ganzi Terrane and Qilian Orogenic Belt (Figure 1.5 (b)). It is internally developed with complex large faults, such as the Altyn-Tagh Fault, Haiyuan Fault and Kunlun Fault. This area is the front part of the plateau's internal expansion towards the northeast, where intense orogenesis lead to the rapid uplift of orogenic belts such as the Qilian Orogen and western Qinling Orogen (Yin 2010). It is also a significant area for the tectonic extrusion of plateau materials towards the east (Tapponnier et al. 2001). Seismic data and deep structure contribute significantly to understanding these processes and to validating dynamic models. Ye et al. (2015) identify a low-velocity layer in the middle-lower crust beneath the northeastern Tibetan Plateau through P-wave receiver

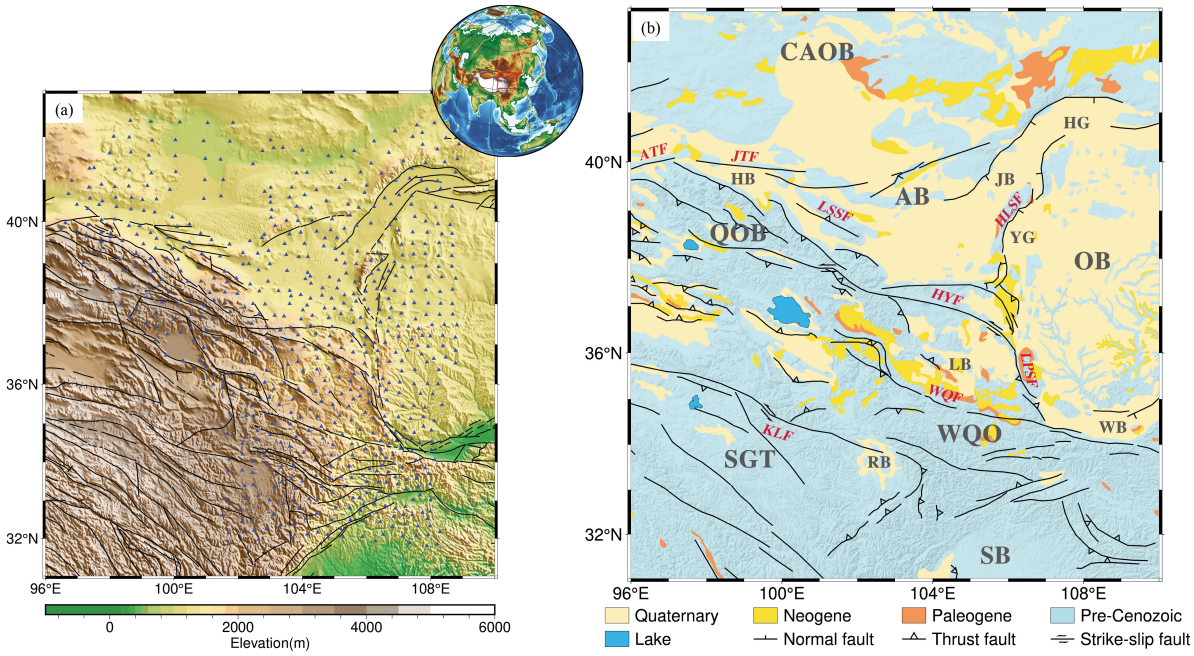


Figure 1.5: Overview of the tectonics, topography and distribution of seismic stations in northeastern Tibetan Plateau. (a) Dense array geometry used in the following chapters. Black lines indicate the main faults, blue triangles indicate stations. (b) Geological settings of the study area (modified from Sethian (1999) and Ren et al. (2024)). The main faults are labeled as: ATF = Altyn-Tagh Fault; HLSF = Helanshan Fault; HYF = Haiyuan Fault; JTF = Jintan Fault; KLF = Kunlun Fault; LPSF = Liupanshan Fault; LSSF = Longshoushan Fault; WQF = West-Qinling Fault. The main tectonic blocks are labeled as AB = Alxa Block; CAOB = Central Asian Orogenic Belt; HB = Hexi Basin; HG = Hetao Graben; JB = Jilantai Basin; LB = Longzhong Basin; OB = Ordos Block; QOB = Qilian Orogenic Belt; RB = Ruoergai Basin; SB = Sichuan Basin; SGT = Songpan-Ganzi Terrane; WB = Weihe Basin; WQO = Western Qinling Orogen; YG = Yinchuan Graben.

functions, suggesting that the lithospheric mantle of the North China Craton has underthrust beneath the Qilian Orogen. Yu and Chen (2016) perform SKS wave splitting to estimate the seismic anisotropy in the upper mantle beneath the southern Ordos and Qinling-Dabie Orogen, finding that a narrow eastward channel flow along the Weihe Basin and Qinling-Dabie Orogen could provide a pathway for asthenospheric flow from the northeastern Tibetan Plateau. Guo and Chen (2017) illustrate two separated low velocity zones in the upper mantle beneath the North China craton and northeastern Tibetan Plateau using joint inversion of ambient noise tomography and receiver functions, showing that the eastward asthenospheric flow from the Tibetan Plateau does not extend beneath the Qinling orogen. Sun and Zhao (2020) present P wave velocity and azimuthal anisotropy in the crust and uppermost mantle beneath the northeastern Tibet using body wave traveltime tomography. Their findings indicate that the low-velocity zone in the middle-lower crust extends northeastward but is resisted by surrounding rigid blocks, preventing further eastward extrusion of crustal flow between the Ordos Orogen

and Sichuan Basin. These observations reveal a complex crust-mantle structure in the north-eastern Tibetan Plateau, highlighting the need for new constraints and advanced methodologies to further clarify the region's tectonic evolution and dynamics.

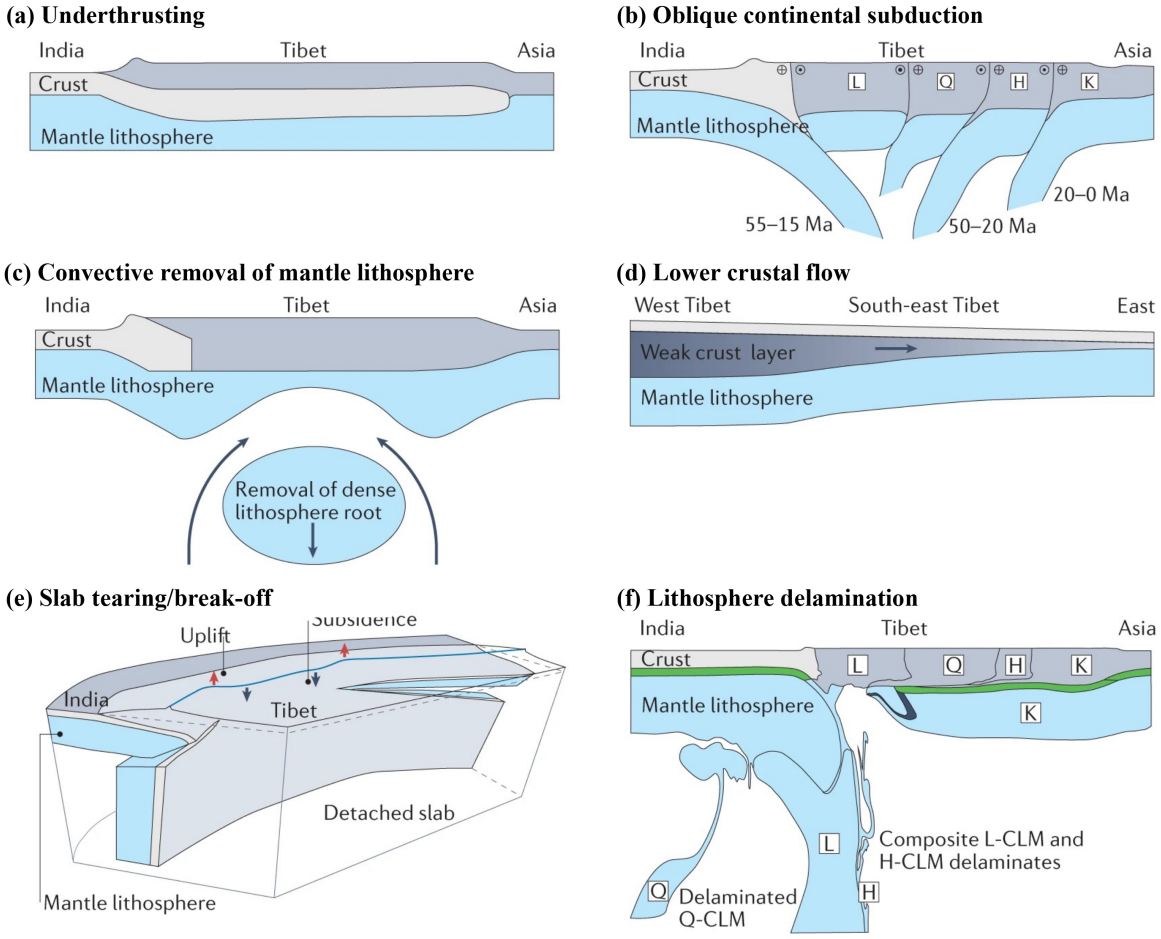


Figure 1.6: Model of the major geodynamic mechanisms proposed to explain the formation and evolution of the Tibetan Plateau (modified from Ding et al. (2022)). (a) Underthrusting (Powell and Conaghan 1973). The Indian lithosphere subducted horizontally beneath the crust of the Tibetan Plateau, led to crustal thickening. (b) Oblique continental subduction (Tapponnier et al. 2001). The northward growth of the Tibetan Plateau was driven by the subduction of continental lithosphere along sutures. (c) Convective removal of mantle lithosphere (Molnar et al. 1993). The uniform thickening of the lithosphere, followed by the removal of its dense lower part, led to a surface uplift by 1.0–2.5 km within a few million years. (d) Lower crustal flow (Royden et al. 2008). The flow of weak lower crust driven by topography could have uplifted eastern and southeastern Tibet. (e) Slab tearing/break-off (Webb et al. 2017). Dynamic surface subsidence occurs during the break-off of a dense slab, followed by surface uplift after the break-off. (f) Lithosphere delamination (Kelly et al. 2020). The northward advancing subduction of the Indian lithosphere induced thickening and delamination of the weaker Asian lithosphere.

1.3.2 Surface wave studies in northeastern Tibetan Plateau

Surface wave tomography methods have been widely used in the northeastern Tibetan Plateau, providing reliable characteristics of crust-mantle velocity structure and anisotropic structure in this region. Early studies, such as Romanowicz (1982), obtain phase velocities of Rayleigh waves and Love waves at 30-90 s in the Tibet, finding a low velocity zone (LVZ) at a depth of 100 to 150 km. With increasing seismic station density, more detailed studies emerged. For example, Bao et al. (2013) utilize ambient noise and earthquake data from Chinese regional digital seismic networks and temporary seismic arrays to reveal S wave velocity structure down to 120 km in the North China Craton and northeastern Tibetan Plateau. The results illustrate an isolated mid-crustal LVZ in the Qilian Orogen, with relatively higher velocity compared to other LVZs at similar depth, suggesting a young LVZ that may reconcile vertically coherent deformation model. Jiang et al. (2014) construct Rayleigh wave maps at 10–60 s and a 3-D Vsv model using ambient noise data from 280 stations. The LVZs in the middle crust exhibit significant west–east variations along the Kunlun Fault, which is consistent with the north-eastward flow. Guo and Chen (2016) construct a crustal model of the eastern Qinling Orogenic Belt through joint analysis of surface waves, receiver functions and gravity data from 35 portable seismometers and 22 permanent stations. The low velocities in the middle-lower crust indicates no mafic lower crust.

The deployment of 676 seismic stations of ChinArray-Himalaya II (Figure 1.5 (a)) enables more comprehensive surface wave tomography studies in this region. Li et al. (2017) reveal Rayleigh wave phase velocity maps at 10-80 s of the northeastern Tibetan Plateau and western Sino-Korea Craton using two-station method, showing a relatively thick lithosphere beneath the Ordos and northeastern Alxa block and thinning lithosphere beneath the southwestern Alxa block and Qilian Orogen. Zhong et al. (2017) obtain Rayleigh wave phase velocity maps at 12-60 s using eikonal tomography. The low velocities at 16-60 s beneath the western Qilian Orogen, western Qinlin Orogen and Songpan-Ganzi Terrane indicate low mechanical strength of the crust and uppermost mantle in these areas, which is prone to deformation under intense tectonic stress. Wang et al. (2020a) construct 2-D Rayleigh wave phase velocity maps at 7-35 s using beamforming method. The velocity model presents two LVZs in the mid-to lower crust beneath the Songpan-Ganzi Terrane and northwestern Qilian Orogen, providing an evidence for lower crustal flow. Li et al. (2022) invert for radial anisotropy by joint analysis of Rayleigh

and Love wave dispersion curves. The results show positive crustal radial anisotropy beneath the Songpan-Ganzi Terrane and negative radial anisotropy beneath the Qilian Orogen, which suggests crustal shortening deformation may took place in the early stages of plateau evolution, followed by crustal channel flow in later stages. Wu et al. (2023) obtain a 3-D azimuthal anisotropic velocity model using double beamforming tomography method. The results indicate different deformation mechanisms of the Songpan-Ganzi Terrane and Qilian Orogen and decoupled crust and mantle deformation beneath the eastern Alxa and Ordos block. These findings from different surface wave tomography methods have provided essential constraints on the heterogeneous lithospheric structure of the northeastern Tibetan Plateau, demonstrating the effectiveness of surface wave tomography in investigating regional deep dynamics and deformation mechanisms. The methodological diversity and robust results from these studies provide valuable references for developing and validating new approaches to surface wave tomography, particularly for PINN-based surface wave tomography.

1.4 Scientific gap

Current studies on seismic tomography and PINNs have made significant progress, yet critical limitations remain:

1. Conventional eikonal tomography methods rely on generic interpolation algorithms (e.g., linear or bicubic spline) to reconstruct traveltime surfaces, introducing velocity biases and resolution loss, especially in regions with sparse or irregular station coverage. These approaches lack physical consistency, as interpolation does not inherently enforce wave propagation physics, leading to potential artifacts in velocity models.

2. PINN applications in seismology lack uncertainty quantification and adaptive optimization, limiting their reliability for real-world tomography. Most studies focus on synthetic or idealized cases, neglecting the impact of field data noise, irregular sampling, and complex geology on model stability and generalizability. Additionally, computational inefficiencies prevent applications on large-scale datasets from dense arrays like ChinArray-Himalaya II.

1.5 Aims and objectives

The aims of this dissertation are:

1. Develop and validate a novel surface wave tomography method that leverages physics-informed neural networks (PINNs) to advance the technology behind surface wave tomography.
2. Investigate practical advantages and limitations of PINN-based surface wave tomography using field data recorded by a dense network (ChinArray-Himalaya II) deployed in northeastern Tibetan Plateau.

To this aim the following objectives are identified:

1. Reshape eikonal tomography as a PINNs training problem where the neural networks (NNs) predict the medium properties and the traveltime observations.
2. Quantify and analyze the uncertainty in the estimated phase velocity models.
3. Develop an approach for choosing the hyper-parameters of the NNs to ensure their ability of accurately representing the effect of complex geology on phase velocity and traveltimes.
4. Adapt the setup of the NNs and validate PINN-based surface wave tomography for phase velocity tomography across range of periods using both ambient noise and teleseismic data.
5. Compare the surface wave phase velocity maps obtained using PINN-based surface wave tomography to traditional eikonal and beamforming tomography techniques.

1.6 Thesis Outline

In chapter2, I propose a novel eikonal tomography approach that leverages physics-informed neural networks (PINNs) to derive Rayleigh wave phase velocities, grounded in the eikonal equation. Initially, I introduce the framework of the PINN-based eikonal tomography (pinnET). This approach is then compared with classical eikonal tomography methods. To demonstrate its application, I utilize 25 s Rayleigh wave data from ChinArray-Himalaya II in the northeastern Tibetan plateau. Furthermore, the results are validated against those obtained through conventional eikonal tomography. This work has been published as Chen et al. (2022).

In chapter 3, I introduce a multi-frequency, elliptical-anisotropic eikonal tomography approach using PINNs, termed pinnEAET. The development of pinnEAET focuses on simultaneously estimating medium properties and reconstructing traveltimes across several frequencies. This approach demonstrates comparable results even with limited inputs. To showcase its application, I utilize ambient noise data from ChinArray-Himalaya II in the northeastern Tibetan Plateau. The pinnEAET approach produces anisotropic phase velocity maps for Rayleigh waves

within a 10-40 s period range. Additionally, these results are compared with those obtained through conventional tomography methods. This work has been published as Chen et al. (2023).

In chapter 4, I extend the PINN-based eikonal tomography method to teleseismic Rayleigh wave tomography (pinnTET). I demonstrate the adaptation of the PINNs' structure and parameters for the analysis of teleseismic data, its application to teleseismic Rayleigh wave data from the northeastern Tibetan Plateau, and a comparison of the results with those obtained through ambient noise tomography presented in previous chapters. Additionally, this work discusses the advantages and challenges associated with using PINNs for teleseismic tomography. This work is ready for submission.

In chapter 5, I discuss the comprehensive development and evaluation of PINN-based surface wave tomography. The discussion covers five key aspects: 1) a summary of the research contributions and achievements, 2) a comprehensive evaluation of the methodological advantages and implementation considerations, 3) an analysis of current limitations and challenges, 4) implications for understanding the structure of the northeastern Tibetan Plateau, and 5) future research directions. Through this systematic evaluation, I demonstrate both the theoretical advances and practical applications of PINNs in seismic tomography while providing a foundation for future developments in physics-constrained approaches.

References

Agata, Ryoichiro, Kazuya Shiraishi, and Gou Fujie (2023). "Bayesian seismic tomography based on velocity-space Stein variational gradient descent for physics-informed neural network". In: *IEEE Transactions on Geoscience and Remote Sensing*.

Aki, Keiiti, Anders Christoffersson, and Eystein S Husebye (1977). "Determination of the three-dimensional seismic structure of the lithosphere". In: *Journal of Geophysical Research* 82.2, pp. 277–296.

Aki, Keiiti and WHK Lee (1976). "Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model". In: *Journal of Geophysical research* 81.23, pp. 4381–4399.

Aki, Keiiti and Paul G Richards (2002). *Quantitative seismology*.

Alkhalifah, Tariq and Xinquan Huang (2024). "Physics-informed neural wavefields with Gabor basis functions". In: *Neural Networks* 175, p. 106286.

Alkhalifah, Tariq, Chao Song, U Bin Waheed, and Qi Hao (2020). "Wavefield solutions from machine learned functions that approximately satisfy the wave equation". In: *EAGE 2020 annual conference & exhibition online*. Vol. 2020. 1. European Association of Geoscientists & Engineers, pp. 1–5.

Anderson, Don L (1965). "Recent evidence concerning the structure and composition of the Earth's mantle". In: *Physics and Chemistry of the Earth* 6, pp. 1–131.

– (1987). "Global mapping of the uppermantle by surface wave tomography". In: *Composition, Structure and Dynamics of the Lithosphere-Asthenosphere System* 16, pp. 89–97.

Araya-Polo, Mauricio, Joseph Jennings, Amir Adler, and Taylor Dahlke (2018). "Deep-learning tomography". In: *The Leading Edge* 37.1, pp. 58–66.

Baker, Nathan, Frank Alexander, Timo Bremer, Aric Hagberg, Yannis Kevrekidis, Habib Najm, Manish Parashar, Abani Patra, James Sethian, Stefan Wild, et al. (2019). *Workshop report on basic research needs for scientific machine learning: Core technologies for artificial intelligence*. Tech. rep. USDOE Office of Science (SC), Washington, DC (United States).

Bao, Xuewei, Xiaodong Song, Mingjie Xu, Liangshu Wang, Xiaoxiao Sun, Ning Mi, Dayong Yu, and Hua Li (2013). "Crust and upper mantle structure of the North China Craton and the NE Tibetan Plateau and its tectonic implications". In: *Earth and Planetary Science Letters* 369, pp. 129–137.

Bao, Xueyang, Colleen A Dalton, and Jeroen Ritsema (2016). "Effects of elastic focusing on global models of Rayleigh wave attenuation". In: *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society* 207.2, pp. 1062–1079.

Baydin, Atilim Gunes, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind (2018). "Automatic differentiation in machine learning: a survey". In: *Journal of Machine Learning Research* 18, pp. 1–43.

Becker, Thorsten W (2020). "Seismic anisotropy". In: *Encyclopedia of Solid Earth Geophysics*. Springer, pp. 1–11.

Beghein, Caroline, Kaiqing Yuan, Nicholas Schmerr, and Zheng Xing (2014). "Changes in seismic anisotropy shed light on the nature of the Gutenberg discontinuity". In: *Science* 343.6176, pp. 1237–1240.

Bergen, Karianne J, Paul A Johnson, Maarten V de Hoop, and Gregory C Beroza (2019). "Machine learning for data-driven discovery in solid Earth geoscience". In: *Science* 363.6433, eaau0323.

Beucler, Tom, Michael Pritchard, Stephan Rasp, Jordan Ott, Pierre Baldi, and Pierre Gentine (2021). "Enforcing analytic constraints in neural networks emulating physical systems". In: *Physical review letters* 126.9, p. 098302.

Bishop, Chris M (1994). "Neural networks and their applications". In: *Review of scientific instruments* 65.6, pp. 1803–1832.

Bokelmann, Götz HR (2002). "Convection-driven motion of the North American craton: evidence from P-wave anisotropy". In: *Geophysical Journal International* 148.2, pp. 278–287.

Boué, Pierre, Philippe Roux, Michel Campillo, and Xavier Briand (2014). "Phase velocity tomography of surface waves using ambient noise cross correlation and array processing". In: *Journal of Geophysical Research: Solid Earth* 119.1, pp. 519–529.

Bourjot, Laurence and Barbara Romanowicz (1992). "Crust and upper mantle tomography in Tibet using surface waves". In: *Geophysical Research Letters* 19.9, pp. 881–884.

Brune, James N, John E Nafe, and Jack E Oliver (1960). "A simplified method for the analysis and synthesis of dispersed wave trains". In: *Journal of Geophysical Research* 65.1, pp. 287–303.

Cai, Shengze, Zhiping Mao, Zhicheng Wang, Minglang Yin, and George Em Karniadakis (2021). "Physics-informed neural networks (PINNs) for fluid mechanics: A review". In: *Acta Mechanica Sinica* 37.12, pp. 1727–1738.

Campillo, Michel and Anne Paul (2003). "Long-range correlations in the diffuse seismic coda". In: *Science* 299.5606, pp. 547–549.

Chai, Xintao, Zhiyuan Gu, Hang Long, Shaoyong Liu, Taihui Yang, Lei Wang, Fenglin Zhan, Xiaodong Sun, and Wenjun Cao (2024). "Modeling multisource multifrequency acoustic wavefields by a multiscale Fourier feature physics-informed neural network with adaptive activation functions". In: *Geophysics* 89.3, T79–T94.

Chen, Feiyu, David Sondak, Pavlos Protopapas, Marios Mattheakis, Shuheng Liu, Devansh Agarwal, and Marco Di Giovanni (2020a). "Neurodiffeq: A python package for solving differential equations with neural networks". In: *Journal of Open Source Software* 5.46, p. 1931.

Chen, Ricky TQ, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud (2018). "Neural ordinary differential equations". In: *Advances in neural information processing systems* 31.

Chen, Yunpeng, Sjoerd AL de Ridder, Sebastian Rost, Zhen Guo, Xiaoyang Wu, and Yongshun Chen (2022). "Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the northeastern margin of the Tibetan Plateau". In: *Geophysical Research Letters* 49.21, e2022GL099053.

Chen, Yunpeng, Sjoerd AL de Ridder, Sebastian Rost, Zhen Guo, Xiaoyang Wu, Shilin Li, and Yongshun Chen (2023). "Physics-Informed Neural Networks for Elliptical-Anisotropy Eikonal Tomography: Application to Data From the Northeastern Tibetan Plateau". In: *Journal of Geophysical Research: Solid Earth* 128.12, e2023JB027378.

Chen, Yuyao, Lu Lu, George Em Karniadakis, and Luca Dal Negro (2020b). "Physics-informed neural networks for inverse problems in nano-optics and metamaterials". In: *Optics express* 28.8, pp. 11618–11633.

Clark, Marin Kristen, MA House, LH Royden, KX Whipple, BC Burchfiel, X Zhang, and W Tang (2005). "Late Cenozoic uplift of southeastern Tibet". In: *Geology* 33.6, pp. 525–528.

Clark, Marin Kristen and Leigh Handy Royden (2000). "Topographic ooze: Building the eastern margin of Tibet by lower crustal flow". In: *Geology* 28.8, pp. 703–706.

Crampin, Stuart and David C Booth (1985). "Shear-wave polarizations near the North Anatolian Fault–II. Interpretation in terms of crack-induced anisotropy". In: *Geophysical Journal International* 83.1, pp. 75–92.

Cuomo, Salvatore, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli (2022). "Scientific machine learning through physics–informed neural networks: Where we are and what's next". In: *Journal of Scientific Computing* 92.3, p. 88.

Curtis, Andrew and John H Woodhouse (1997). "Crust and upper mantle shear velocity structure beneath the Tibetan plateau and surrounding regions from interevent surface wave phase velocity inversion". In: *Journal of Geophysical Research: Solid Earth* 102.B6, pp. 11789–11813.

Dalton, Colleen A and Göran Ekström (2006). "Global models of surface wave attenuation". In: *Journal of Geophysical Research: Solid Earth* 111.B5.

Dalton, Colleen A, Göran Ekström, and Adam M Dziewoński (2008). "The global attenuation structure of the upper mantle". In: *Journal of Geophysical Research: Solid Earth* 113.B9.

Das, Vishal, Ahinoam Pollack, Uri Wollner, and Tapan Mukerji (2019). "Convolutional neural network for seismic impedance inversion". In: *Geophysics* 84.6, R869–R880.

Daw, Arka, Anuj Karpatne, William D Watkins, Jordan S Read, and Vipin Kumar (2022). "Physics-guided neural networks (pgnn): An application in lake temperature modeling". In: *Knowledge guided machine learning*. Chapman and Hall/CRC, pp. 353–372.

Daw, Arka, R Quinn Thomas, Cayelan C Carey, Jordan S Read, Alison P Appling, and Anuj Karpatne (2020). "Physics-guided architecture (pga) of neural networks for quantifying uncertainty in lake temperature modeling". In: *Proceedings of the 2020 siam international conference on data mining*. SIAM, pp. 532–540.

De Ridder, SA (2011). "Ambient seismic noise tomography for exploration seismology at Valhall". In: *AGU Fall Meeting Abstracts*. Vol. 2011, S31F–04.

De Ridder, SAL and BL Biondi (2015). "Near-surface Scholte wave velocities at Ekofisk from short noise recordings by seismic noise gradiometry". In: *Geophysical Research Letters* 42.17, pp. 7031–7038.

De Ridder, SAL, BL Biondi, and D Nichols (2015). "Elliptical-anisotropic eikonal phase velocity tomography". In: *Geophysical Research Letters* 42.3, pp. 758–764.

DeVries, Phoebe MR, Fernanda Viégas, Martin Wattenberg, and Brendan J Meade (2018). "Deep learning of aftershock patterns following large earthquakes". In: *Nature* 560.7720, pp. 632–634.

Ding, Lin, Paul Kapp, Fulong Cai, Carmala N Garzione, Zhongyu Xiong, Houqi Wang, and Chao Wang (2022). "Timing and mechanisms of Tibetan Plateau uplift". In: *Nature Reviews Earth & Environment* 3.10, pp. 652–667.

Dorman, James, Maurice Ewing, and Jack Oliver (1960). "Study of shear-velocity distribution in the upper mantle by mantle Rayleigh waves". In: *Bulletin of the Seismological Society of America* 50.1, pp. 87–115.

Dziewonski, Adam M and Don L Anderson (1981). "Preliminary reference Earth model". In: *Physics of the earth and planetary interiors* 25.4, pp. 297–356.

Dziewonski, Adam M, Bradford H Hager, and Richard J O'Connell (1977). "Large-scale heterogeneities in the lower mantle". In: *Journal of Geophysical Research* 82.2, pp. 239–255.

Ekström, Göran, Jeroen Tromp, and Erik WF Larson (1997). "Measurements and global models of surface wave propagation". In: *Journal of Geophysical Research: Solid Earth* 102.B4, pp. 8137–8157.

Ewing, Maurice and Frank Press (1954). "An investigation of mantle Rayleigh waves". In: *Bulletin of the Seismological Society of America* 44.2A, pp. 127–147.

Fan, Xingli, Zhen Guo, Ling Chen, Qi-Fu Chen, Ting Yang, and Chuntao Liang (2024). "Focused mantle upwelling beneath the southeastern Asian Basalt Province revealed by seismic surface wave tomography". In: *Geophysical Research Letters* 51.2, e2023GL104336.

Fang, Zhiwei (2021). "A high-efficient hybrid physics-informed neural networks based on convolutional neural network". In: *IEEE Transactions on Neural Networks and Learning Systems* 33.10, pp. 5514–5526.

Fang, Zhiwei and Justin Zhan (2019). "Deep physical informed neural networks for metamaterial design". In: *IEEE Access* 8, pp. 24506–24513.

Forsyth, Donald W (1975). "The early structural evolution and anisotropy of the oceanic upper mantle". In: *Geophysical Journal International* 43.1, pp. 103–162.

Forsyth, Donald W, Aibing Li, A Levander, and G Nolet (2005). "Array analysis of two-dimensional variations in surface wave phase velocity and azimuthal anisotropy in the presence of multipathing interference". In: *Geophysical Monograph-American Geophysical Union* 157, p. 81.

Forsyth, Donald W, Spahr C Webb, LeRoy M Dorman, and Yang Shen (1998). "Phase velocities of Rayleigh waves in the MELT experiment on the East Pacific Rise". In: *Science* 280.5367, pp. 1235–1238.

Friederich, Wolfgang (1998). "Wave-theoretical inversion of teleseismic surface waves in a regional network: phase-velocity maps and a three-dimensional upper-mantle shear-wave-velocity model for southern Germany". In: *Geophysical Journal International* 132.1, pp. 203–225.

Friederich, Wolfgang, Erhard Wielandt, and Stefan Stange (1994). "Non-plane geometries of seismic surface wavefields and their implications for regional surface-wave tomography". In: *Geophysical Journal International* 119.3, pp. 931–948.

Gaherty, James B (2004). "A surface wave analysis of seismic anisotropy beneath eastern North America". In: *Geophysical Journal International* 158.3, pp. 1053–1066.

Galvis, Ivan Sánchez, Yenni Villa, César Duarte, Daniel Sierra, and William Agudelo (2017). "Seismic attribute selection and clustering to detect and classify surface waves in multicomponent seismic data by using k-means algorithm". In: *The Leading Edge* 36.3, pp. 239–248.

Gao, Han, Luning Sun, and Jian-Xun Wang (2021). "PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain". In: *Journal of Computational Physics* 428, p. 110079.

Gatti, Filippo and Didier Clouteau (2020). "Towards blending physics-based numerical simulations and seismic databases using generative adversarial network". In: *Computer Methods in Applied Mechanics and Engineering* 372, p. 113421.

Goswami, Somdatta, Cosmin Anitescu, Souvik Chakraborty, and Timon Rabczuk (2020). "Transfer learning enhanced physics informed neural network for phase-field modeling of fracture". In: *Theoretical and Applied Fracture Mechanics* 106, p. 102447.

Gou, Rongxi, Yijie Zhang, Xueyu Zhu, and Jinghuai Gao (2023). "Bayesian physics-informed neural networks for the subsurface tomography based on the eikonal equation". In: *IEEE Transactions on Geoscience and Remote Sensing* 61, pp. 1–12.

Gung, Yuancheng, Mark Panning, and Barbara Romanowicz (2003). "Global anisotropy and the thickness of continents". In: *Nature* 422.6933, pp. 707–711.

Guo, Zhen and Y John Chen (2016). "Crustal structure of the eastern Qinling orogenic belt and implication for reactivation since the Cretaceous". In: *Tectonophysics* 683, pp. 1–11.

Guo, Zhen and Yongshun John Chen (2017). "Mountain building at northeastern boundary of Tibetan Plateau and craton reworking at Ordos block from joint inversion of ambient noise

tomography and receiver functions". In: *Earth and Planetary Science Letters* 463, pp. 232–242.

Haghighat, Ehsan and Ruben Juanes (2021). "Sciann: A keras/tensorflow wrapper for scientific computations and physics-informed deep learning using artificial neural networks". In: *Computer Methods in Applied Mechanics and Engineering* 373, p. 113552.

Hao, Shijie, Zhouchuan Huang, Cunrui Han, Liangshu Wang, Mingjie Xu, Ning Mi, and Dayong Yu (2021). "Layered crustal azimuthal anisotropy beneath the northeastern Tibetan Plateau revealed by Rayleigh-wave Eikonal tomography". In: *Earth and Planetary Science Letters* 563, p. 116891.

Hennigh, Oliver, Susheela Narasimhan, Mohammad Amin Nabian, Akshay Subramaniam, Kaustubh Tangsali, Zhiwei Fang, Max Rietmann, Wonmin Byeon, and Sanjay Choudhry (2021). "NVIDIA SimNet™: An AI-accelerated multi-physics simulation framework". In: *International conference on computational science*. Springer, pp. 447–461.

Hess, HH (1964). "Seismic anisotropy of the uppermost mantle under oceans". In: *Nature* 203.4945, pp. 629–631.

Hibert, Clément, Floriane Provost, Jean-Philippe Malet, Alessia Maggi, André Stumpf, and Valérie Ferrazzini (2017). "Automatic identification of rockfalls and volcano-tectonic earthquakes at the Piton de la Fournaise volcano using a Random Forest algorithm". In: *Journal of Volcanology and Geothermal Research* 340, pp. 130–142.

Hinton, Geoffrey E and Ruslan R Salakhutdinov (2006). "Reducing the dimensionality of data with neural networks". In: *science* 313.5786, pp. 504–507.

Huang, Xinquan and Tariq Alkhalifah (2022a). "PINNup: Robust neural network wavefield solutions using frequency upscaling and neuron splitting". In: *Journal of Geophysical Research: Solid Earth* 127.6, e2021JB023703.

– (2022b). "Single Reference Frequency Loss for Multi-frequency Wavefield Representation using Physics-Informed Neural Networks". In: *IEEE Geoscience and Remote Sensing Letters*.

Islam, Mahmudul, Md Shajedul Hoque Thakur, Satyajit Mojumder, and Mohammad Nasim Hasan (2021). "Extraction of material properties through multi-fidelity deep learning from molecular dynamics simulation". In: *Computational Materials Science* 188, p. 110187.

Jagtap, Ameya D and George Em Karniadakis (2020). "Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations". In: *Communications in Computational Physics* 28.5.

Jagtap, Ameya D, Ehsan Kharazmi, and George Em Karniadakis (2020). "Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems". In: *Computer Methods in Applied Mechanics and Engineering* 365, p. 113028.

Jiang, Chengxin, Yingjie Yang, and Yong Zheng (2014). "Penetration of mid-crustal low velocity zone across the Kunlun Fault in the NE Tibetan Plateau revealed by ambient noise tomography". In: *Earth and Planetary Science Letters* 406, pp. 81–92.

Jianguo, Song and Munezero Ntibahanana (2024). "Developing deep learning methods for pre-stack seismic data inversion". In: *Journal of Applied Geophysics* 222, p. 105336.

Jin, Ge and James B Gaherty (2015). "Surface wave phase-velocity tomography based on multichannel cross-correlation". In: *Geophysical Journal International* 201.3, pp. 1383–1398.

Karimpouli, Sadegh and Pejman Tahmasebi (2020). "Physics informed machine learning: Seismic wave equation". In: *Geoscience Frontiers* 11.6, pp. 1993–2001.

Karniadakis, George Em, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang (2021). "Physics-informed machine learning". In: *Nature Reviews Physics* 3.6, pp. 422–440.

Karpatne, Anuj, Gowtham Atluri, James H Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar (2017). "Theory-guided data science: A new paradigm for scientific discovery from data". In: *IEEE Transactions on knowledge and data engineering* 29.10, pp. 2318–2331.

Kästle, Emanuel David, Irene Molinari, Lapo Boschi, Edi Kissling, and AlpArray Working Group (2022). "Azimuthal anisotropy from eikonal tomography: example from ambient-noise measurements in the AlpArray network". In: *Geophysical Journal International* 229.1, pp. 151–170.

Kaviani, Ayoub, Anne Paul, Ali Moradi, Paul Martin Mai, Simone Pilia, Lapo Boschi, Georg Rümpker, Yang Lu, Zheng Tang, and Eric Sandvol (2020). "Crustal and uppermost mantle shear wave velocity structure beneath the Middle East from surface wave tomography". In: *Geophysical Journal International* 221.2, pp. 1349–1365.

Kelly, Sean, Christopher Beaumont, and Jared P Butler (2020). "Inherited terrane properties explain enigmatic post-collisional Himalayan-Tibetan evolution". In: *Geology* 48.1, pp. 8–14.

Kissas, Georgios, Yibo Yang, Eileen Hwuang, Walter R Witschey, John A Detre, and Paris Perdikaris (2020). "Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks". In: *Computer Methods in Applied Mechanics and Engineering* 358, p. 112623.

Kong, Qingkai, Daniel T Trugman, Zachary E Ross, Michael J Bianco, Brendan J Meade, and Peter Gerstoft (2019). "Machine learning in seismology: Turning data into insights". In: *Seismological Research Letters* 90.1, pp. 3–14.

Kuo, John, James Brune, and Maurice Major (1962). "Rayleigh wave dispersion in the Pacific Ocean for the period range 20 to 140 seconds". In: *Bulletin of the Seismological Society of America* 52.2, pp. 333–357.

Langston, Charles A (2007). "Wave gradiometry in two dimensions". In: *Bulletin of the Seismological Society of America* 97.2, pp. 401–416.

Laske, G (1995). "Global observation of off-great-circle propagation of long-period surface waves". In: *Geophysical Journal International* 123.1, pp. 245–259.

Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani (2014). "The parable of Google Flu: traps in big data analysis". In: *science* 343.6176, pp. 1203–1205.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). "Deep learning". In: *nature* 521.7553, pp. 436–444.

Li, Aibing and Lun Li (2015). "Love wave tomography in southern Africa from a two-plane-wave inversion method". In: *Geophysical Journal International* 202.2, pp. 1005–1020.

Li, Lun and Yuanyuan V Fu (2020). "Surface-wave tomography of eastern and central Tibet from two-plane-wave inversion: Rayleigh-wave and Love-wave phase velocity maps". In: *Bulletin of the Seismological Society of America* 110.3, pp. 1359–1371.

Li, Shilin, Zhen Guo, Yong Yu, Xiaoyang Wu, and Yongshun John Chen (2022). "Imaging the Northeastern crustal boundary of the Tibetan Plateau with radial anisotropy". In: *Geophysical Research Letters* 49.23, e2022GL100672.

Li, Yonghua, Jiatie Pan, Qingju Wu, and Zhifeng Ding (2017). "Lithospheric structure beneath the northeastern Tibetan Plateau and the western Sino-Korea Craton revealed by Rayleigh wave tomography". In: *Geophysical Journal International* 210.2, pp. 570–584.

Li, Yue, Hongzhou Wang, and Xintong Dong (2020). "The denoising of desert seismic data based on cycle-GAN with unpaired data training". In: *IEEE Geoscience and Remote Sensing Letters* 18.11, pp. 2016–2020.

Li, Yunyue Elita, Daniel O'malley, Greg Beroza, Andrew Curtis, and Paul Johnson (2023). *Machine Learning Developments and Applications in Solid-Earth Geosciences: Fad or Future?*

Liang, Chuntao and Charles A Langston (2009). "Wave gradiometry for USArray: Rayleigh waves". In: *Journal of Geophysical Research: Solid Earth* 114.B2.

Lin, Fan-Chi and Michael H Ritzwoller (2011). "Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure". In: *Geophysical Journal International* 186.3, pp. 1104–1120.

Lin, Fan-Chi, Michael H Ritzwoller, and Roel Snieder (2009). "Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array". In: *Geophysical Journal International* 177.3, pp. 1091–1110.

Liu, Yiduo, Lun Li, Jolante van Wijk, Aibing Li, and Yuanyuan V Fu (2021). "Surface-wave tomography of the Emeishan large igneous province (China): Magma storage system, hidden hotspot track, and its impact on the Capitanian mass extinction". In: *Geology* 49.9, pp. 1032–1037.

Liu, Yongsheng and Ping Tong (2021). "Eikonal equation-based P-wave seismic azimuthal anisotropy tomography of the crustal structure beneath northern California". In: *Geophysical Journal International* 226.1, pp. 287–301.

Lobkis, Oleg I and Richard L Weaver (2001). "On the emergence of the Green's function in the correlations of a diffuse field". In: *Journal of the Acoustical Society of America* 110.6, 3011–3017".

Long, Maureen D and Thorsten W Becker (2010). "Mantle dynamics and seismic anisotropy". In: *Earth and Planetary Science Letters* 297.3-4, pp. 341–354.

Lu, Cai, Yunchen Wang, Xuyang Zou, Jingjing Zong, and Qin Su (2024). "Elastic Full-Waveform Inversion via Physics-Informed Recurrent Neural Network". In: *IEEE Transactions on Geoscience and Remote Sensing*.

Lu, Lu, Xuhui Meng, Zhiping Mao, and George Em Karniadakis (2021). "DeepXDE: A deep learning library for solving differential equations". In: *SIAM review* 63.1, pp. 208–228.

Magali, John Keith, Thomas Bodin, Navid Hedjazian, Henri Samuel, and Suzanne Atkins (2021). "Geodynamic tomography: constraining upper-mantle deformation patterns from Bayesian inversion of surface waves". In: *Geophysical Journal International* 224.3, pp. 2077–2099.

Mao, Zhiping, Ameya D Jagtap, and George Em Karniadakis (2020). "Physics-informed neural networks for high-speed flows". In: *Computer Methods in Applied Mechanics and Engineering* 360, p. 112789.

Meier, T, K Dietrich, B Stöckhert, and H-P Harjes (2004). "One-dimensional models of shear wave velocity for the eastern Mediterranean obtained from the inversion of Rayleigh wave phase velocities and tectonic implications". In: *Geophysical Journal International* 156.1, pp. 45–58.

Meng, Haoran, Yehuda Ben-Zion, and Christopher W Johnson (2021). "Analysis of seismic signals generated by vehicle traffic with application to derivation of subsurface Q-values". In: *Seismological Research Letters* 92.4, pp. 2354–2363.

Misyris, George S, Andreas Venzke, and Spyros Chatzivasileiadis (2020). "Physics-informed neural networks for power systems". In: *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, pp. 1–5.

Molnar, Peter, Philip England, and Joseph Martinod (1993). "Mantle dynamics, uplift of the Tibetan Plateau, and the Indian monsoon". In: *Reviews of Geophysics* 31.4, pp. 357–396.

Molnar, Peter and Paul Tapponnier (1975). "Cenozoic Tectonics of Asia: Effects of a Continental Collision: Features of recent continental tectonics in Asia can be interpreted as results of the India-Eurasia collision". In: *science* 189.4201, pp. 419–426.

Montagner, Jean-Paul and Laurent Guillot (2002). "Seismic anisotropy and global geodynamics". In: *Reviews in mineralogy and geochemistry* 51.1, pp. 353–385.

Montagner, Jean-Paul and Toshiro Tanimoto (1990). "Global anisotropy in the upper mantle inferred from the regionalization of phase velocities". In: *Journal of Geophysical Research: Solid Earth* 95.B4, pp. 4797–4819.

– (1991). "Global upper mantle tomography of seismic velocities and anisotropies". In: *Journal of Geophysical Research: Solid Earth* 96.B12, pp. 20337–20351.

Moseley, Ben, Andrew Markham, and Tarje Nissen-Meyer (2020a). "Solving the wave equation with physics-informed deep learning". In: *arXiv preprint arXiv:2006.11894*.

Moseley, Ben, Tarje Nissen-Meyer, and Andrew Markham (2020b). "Deep learning for fast simulation of seismic waves in complex media". In: *Solid Earth* 11.4, pp. 1527–1549.

Moseley, Benjamin (2022). "Physics-informed machine learning: from concepts to real-world applications". PhD thesis. University of Oxford.

Mousavi, S Mostafa and Gregory C Beroza (2020). "Bayesian-Deep-Learning Estimation of Earthquake Location From Single-Station Observations". In: *IEEE Transactions on Geoscience and Remote Sensing* 58.11, pp. 8211–8224.

– (2022). “Deep-learning seismology”. In: *Science* 377.6607, eabm4470.

Münchmeyer, Jannes, Dino Bindi, Ulf Leser, and Frederik Tilmann (2021). “The transformer earthquake alerting model: A new versatile approach to earthquake early warning”. In: *Geophysical Journal International* 225.1, pp. 646–656.

Nakata, Nori (2016). “Near-surface S-wave velocities estimated from traffic-induced Love waves using seismic interferometry with double beamforming”. In: *Interpretation* 4.4, SQ23–SQ31.

Nazari Siahsar, Mohammad Amir, Saman Gholtashi, Amin Roshandel Kahoo, Wei Chen, and Yangkang Chen (2017). “Data-driven multitask sparse dictionary learning for noise attenuation of 3D seismic data”. In: *Geophysics* 82.6, pp. V385–V396.

Nicolas, Adolphe and Nikolas I Christensen (1987). “Formation of anisotropy in upper mantle peridotites-A review”. In: *Composition, structure and dynamics of the lithosphere-asthenosphere system* 16, pp. 111–123.

Nishida, Kiwamu, Jean-Paul Montagner, and Hitoshi Kawakatsu (2009). “Global surface wave tomography using seismic hum”. In: *Science* 326.5949, pp. 112–112.

Paitz, Patrick, Alexey Gokhberg, and Andreas Fichtner (2018). “A neural network for noise correlation classification”. In: *Geophysical Journal International* 212.2, pp. 1468–1474.

Pang, Guofei, Marta D’Elia, Michael Parks, and George E Karniadakis (2020). “nPINNs: nonlocal Physics-Informed Neural Networks for a parametrized nonlocal universal Laplacian operator. Algorithms and Applications”. In: *Journal of Computational Physics* 422, p. 109760.

Pang, Guofei, Lu Lu, and George Em Karniadakis (2019). “fPINNs: Fractional physics-informed neural networks”. In: *SIAM Journal on Scientific Computing* 41.4, A2603–A2626.

Panju, Maysum and Ali Ghodsi (2020). “A neuro-symbolic method for solving differential and functional equations”. In: *arXiv preprint arXiv:2011.02415*.

Park, Jeffrey and Vadim Levin (2002). “Seismic anisotropy: tracing plate dynamics in the mantle”. In: *Science* 296.5567, pp. 485–489.

Pawar, Suraj, Omer San, Burak Aksoylu, Adil Rasheed, and Trond Kvamsdal (2021). "Physics guided machine learning using simplified theories". In: *Physics of Fluids* 33.1.

Powell, C McA and PJ Conaghan (1973). "Plate tectonics and the Himalayas". In: *Earth and Planetary Science Letters* 20.1, pp. 1–12.

Press, Frank (1956). "Determination of crustal structure from phase velocity of Rayleigh waves part I: southern California". In: *Geological Society of America Bulletin* 67.12, pp. 1647–1658.

Prindle, K and T Tanimoto (2006). "Teleseismic surface wave study for S-wave velocity structure under an array: Southern California". In: *Geophysical Journal International* 166.2, pp. 601–621.

Qiu, Hongrui, Fan-Chi Lin, and Yehuda Ben-Zion (2019). "Eikonal tomography of the Southern California plate boundary region". In: *Journal of Geophysical Research: Solid Earth* 124.9, pp. 9755–9779.

Raissi, Maziar, Paris Perdikaris, and George E Karniadakis (2019). "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In: *Journal of Computational Physics* 378, pp. 686–707.

Raissi, Maziar, Alireza Yazdani, and George Em Karniadakis (2020). "Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations". In: *Science* 367.6481, pp. 1026–1030.

Rasht-Behesht, Majid, Christian Huber, Khemraj Shukla, and George Em Karniadakis (2022). "Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions". In: *Journal of Geophysical Research: Solid Earth* 127.5, e2021JB023120.

Rawlinson, Nicholas, S Pozgay, and S Fishwick (2010). "Seismic tomography: a window into deep Earth". In: *Physics of the Earth and Planetary Interiors* 178.3-4, pp. 101–135.

Ren, Pengfei, Zhen Guo, Yiduo Liu, Bin Luo, and Xiaoyang Wu (2024). "Seismic evidence of basin development in NE Tibetan Plateau in response to deep crustal dynamics from joint

inversion of Rayleigh wave ellipticity and phase velocity". In: *Geophysical Research Letters* 51.18, e2024GL108589.

Roden, Rocky, Thomas Smith, and Deborah Sacrey (2015). "Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps". In: *Interpretation* 3.4, SAE59–SAE83.

Rojas, Sergio, Paweł Maczuga, Judit Muñoz-Matute, David Pardo, and Maciej Paszyński (2024). "Robust Variational Physics-Informed Neural Networks". In: *Computer Methods in Applied Mechanics and Engineering* 425, p. 116904.

Romanowicz, Barbara (2002). "Inversion of surface waves: a review". In: *International Geophysics Series* 81.A, pp. 149–174.

– (2020). "Surface waves". In: *Encyclopedia of solid earth geophysics*, pp. 1–13.

Romanowicz, Barbara A (1982). "Constraints on the structure of the Tibet Plateau from pure path phase velocities of Love and Rayleigh waves". In: *Journal of Geophysical Research: Solid Earth* 87.B8, pp. 6865–6883.

Rost, Sebastian and Christine Thomas (2002). "Array seismology: Methods and applications". In: *Reviews of geophysics* 40.3, pp. 2–1.

– (2009). "Improving seismic resolution through array processing techniques". In: *Surveys in geophysics* 30, pp. 271–299.

Royden, Leigh H, B Clark Burchfiel, and Robert D van der Hilst (2008). "The geological evolution of the Tibetan Plateau". In: *science* 321.5892, pp. 1054–1058.

Royden, Leigh H, B Clark Burchfiel, Robert W King, Erchie Wang, Zhiliang Chen, Feng Shen, and Yuping Liu (1997). "Surface deformation and lower crustal flow in eastern Tibet". In: *science* 276.5313, pp. 788–790.

Sabra, Karim G, Peter Gerstoft, Philippe Roux, WA Kuperman, and Michael C Fehler (2005). "Surface wave tomography from microseisms in Southern California". In: *Geophysical Research Letters* 32.14.

Sato, Yasuo (1955). "Analysis of dispersed surface waves by means of Fourier transform I". In: *Bull. Earthq. Res. Inst* 33, pp. 33–48.

Savage, Brian, Dimitri Komatitsch, and Jeroen Tromp (2010). "Effects of 3D attenuation on seismic wave amplitude and phase measurements". In: *Bulletin of the Seismological Society of America* 100.3, pp. 1241–1251.

Savage, MK (1999). "Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting?" In: *Reviews of Geophysics* 37.1, pp. 65–106.

Schweitzer, Johannes, Jan Fyen, Svein Mykkeltveit, Steven J Gibbons, Myrto Pirli, Daniela Kühn, and Tormod Kværna (2012). "Seismic arrays". In: *New manual of seismological observatory practice 2 (NMSOP-2)*. Deutsches GeoForschungsZentrum GFZ, pp. 1–80.

Sethian, James A (1999). "Fast marching methods". In: *SIAM review* 41.2, pp. 199–235.

Shapiro, Nikolai M and Michel Campillo (2004). "Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise". In: *Geophysical Research Letters* 31.7.

Shapiro, Nikolai M, Michel Campillo, Laurent Stehly, and Michael H Ritzwoller (2005). "High-resolution surface-wave tomography from ambient seismic noise". In: *Science* 307.5715, pp. 1615–1618.

Shearer, Peter M (2019). *Introduction to seismology*. Cambridge university press.

Silver, Paul G (1996). "Seismic anisotropy beneath the continents: Probing the depths of geology". In: *Annual review of earth and planetary sciences* 24, pp. 385–432.

Silver, Paul G and W Winston Chan (1991). "Shear wave splitting and subcontinental mantle deformation". In: *Journal of Geophysical Research: Solid Earth* 96.B10, pp. 16429–16454.

Smith, Jonathan D, Kamyar Azizzadenesheli, and Zachary E Ross (2020). "Eikonet: Solving the eikonal equation with deep neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 59.12, pp. 10685–10696.

Smith, Martin L and FA Dahlen (1973). "The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium". In: *Journal of Geophysical Research* 78.17, pp. 3321–3333.

Smith, Walter HF and Paul Wessel (1990). "Gridding with continuous curvature splines in tension". In: *Geophysics* 55.3, pp. 293–305.

Snieder, Roel (2004). "Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase". In: *Physical Review Letters* 69.8, p. 046610.

Song, Chao, Tariq Alkhalifah, and UB Waheed (2022). "A versatile framework to solve the Helmholtz equation using physics-informed neural networks". In: *Geophysical Journal International* 228.3, pp. 1750–1762.

Song, Chao, Tariq Alkhalifah, and Umair Bin Waheed (2021). "Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks". In: *Geophysical Journal International* 225.2, pp. 846–859.

Song, Chao and Tariq A Alkhalifah (2021). "Wavefield reconstruction inversion via physics-informed neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–12.

Song, Chao, Hang Geng, Umair bin Waheed, and Cai Liu (2024). "PINNPStomo: Simultaneous P-and S-wave seismic traveltime tomography using physics-informed neural networks with a new factored eikonal equation". In: *arXiv preprint arXiv:2407.16439*.

Song, Chao and Yanghua Wang (2023). "Simulating seismic multifrequency wavefields with the Fourier feature physics-informed neural network". In: *Geophysical Journal International* 232.3, pp. 1503–1514.

Sun, Anhui and Dapeng Zhao (2020). "Anisotropic tomography beneath Northeast Tibet: evidence for regional crustal flow". In: *Tectonics* 39.7, e2020TC006161.

Sun, Luning, Han Gao, Shaowu Pan, and Jian-Xun Wang (2020). "Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data". In: *Computer Methods in Applied Mechanics and Engineering* 361, p. 112732.

Tanimoto, Toshiro and Don L Anderson (1985). "Lateral heterogeneity and azimuthal anisotropy of the upper mantle: Love and Rayleigh waves 100–250 s". In: *Journal of Geophysical Research: Solid Earth* 90.B2, pp. 1842–1858.

Tao, Fei, Xin Liu, Haodong Du, and Wenbin Yu (2020). "Physics-informed artificial neural network approach for axial compression buckling analysis of thin-walled cylinder". In: *AIAA Journal* 58.6, pp. 2737–2747.

Tapponnier, Paul, Xu Zhiqin, Francoise Roger, Bertrand Meyer, Nicolas Arnaud, Gérard Wittlinger, and Yang Jingsui (2001). "Oblique stepwise rise and growth of the Tibet Plateau". In: *science* 294.5547, pp. 1671–1677.

Taufik, Mohammad Hasyim, Umair bin Waheed, and Tariq A Alkhalifah (2022). "Upwind, no more: Flexible traveltime solutions using physics-informed neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–12.

Tong, Ping (2021). "Adjoint-state traveltime tomography: Eikonal equation-based methods and application to the Anza area in southern California". In: *Journal of Geophysical Research: Solid Earth* 126.5, e2021JB021818.

Tromp, Jeroen (2001). "Inner-core anisotropy and rotation". In: *Annual Review of Earth and Planetary Sciences* 29.1, pp. 47–69.

Um, Kiwon, Robert Brand, Yun Raymond Fei, Philipp Holl, and Nils Thuerey (2020). "Solver-in-the-loop: Learning from differentiable physics to interact with iterative pde-solvers". In: *Advances in neural information processing systems* 33, pp. 6111–6122.

Viana, Felipe AC, Renato G Nascimento, Arinan Dourado, and Yigit A Yucesan (2021). "Estimating model inadequacy in ordinary differential equations with physics-informed neural networks". In: *Computers & Structures* 245, p. 106458.

Waheed, Umair bin, Ehsan Haghighat, Tariq Alkhalifah, Chao Song, and Qi Hao (2021). "PINNeik: Eikonal solution using physics-informed neural networks". In: *Computers & Geosciences* 155, p. 104833.

Waheed, Umair bin, Tariq Alkhalifah, Ehsan Haghighat, Chao Song, and Jean Virieux (2021). "PINNtomo: Seismic tomography using physics-informed neural networks". In: *arXiv preprint arXiv:2104.01588*.

Waheed, Umair bin, Ehsan Haghighat, and Tariq Alkhalifah (2020). "Anisotropic eikonal solution using physics-informed neural networks". In: *SEG international exposition and annual meeting*. SEG, D041S100R003.

Wang, Kaiming, Laiyu Lu, Valérie Maupin, Zhifeng Ding, Chen Zheng, and Shijun Zhong (2020a). "Surface wave tomography of northeastern Tibetan Plateau using beamforming of seismic noise at a dense array". In: *Journal of Geophysical Research: Solid Earth* 125.4, e2019JB018416.

Wang, Rui, Robin Walters, and Rose Yu (2020b). "Incorporating symmetry into deep dynamics models for improved generalization". In: *arXiv preprint arXiv:2002.03061*.

Weaver, Richard L and Oleg I Lobkis (2001). "Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies". In: *Physical Review Letters* 87.13, p. 134301.

Webb, A Alexander G, Hongcheng Guo, Peter D Clift, Laurent Husson, Thomas Müller, Diego Costantino, An Yin, Zhiqin Xu, Hui Cao, and Qin Wang (2017). "The Himalaya in 3D: Slab dynamics controlled mountain building and monsoon intensification". In: *Lithosphere* 9.4, pp. 637–651.

Wielandt, E (1993). "Propagation and structural interpretation of non-plane waves". In: *Geophysical Journal International* 113.1, pp. 45–53.

Willard, Jared, Xiaowei Jia, Shaoming Xu, Michael Steinbach, and Vipin Kumar (2022). "Integrating scientific knowledge with machine learning for engineering and environmental systems". In: *ACM Computing Surveys* 55.4, pp. 1–37.

Woodhouse, John H and Adam M Dziewonski (1984). "Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms". In: *Journal of Geophysical Research: Solid Earth* 89.B7, pp. 5953–5986.

Wu, Hao, Bo Zhang, Tengfei Lin, Fangyu Li, and Naihao Liu (2019). "White noise attenuation of seismic trace by integrating variational mode decomposition with convolutional neural network". In: *Geophysics* 84.5, pp. V307–V317.

Wu, Qingju, Xiufen Zheng, Jiatie Pan, Fengxue Zhang, and Guangcheng Zhang (2009). "Measurement of interstation phase velocity by wavelet transformation". In: *Earthquake Science* 22, pp. 425–429.

Wu, Xiaoyang, Zhen Guo, Shilin Li, Yong Yu, Qipeng Bai, and Yongshun John Chen (2023). "Seismic azimuthal anisotropy of northeastern Tibetan Plateau from ambient noise double beamforming tomography: Implications for crustal deformation". In: *Journal of Geophysical Research: Solid Earth* 128.6, e2022JB026109.

Xu, Chong, Xiwei Xu, Fuchu Dai, and Arun K Saraf (2012). "Comparison of different models for susceptibility mapping of earthquake triggered landslides related with the 2008 Wenchuan earthquake in China". In: *Computers & Geosciences* 46, pp. 317–329.

Yang, Fangshu and Jianwei Ma (2023). "FWIGAN: Full-Waveform Inversion via a Physics-Informed Generative Adversarial Network". In: *Journal of Geophysical Research: Solid Earth* 128.4, e2022JB025493.

Yang, Liu, Xuhui Meng, and George Em Karniadakis (2021). "B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data". In: *Journal of Computational Physics* 425, p. 109913.

Yang, Liu, Dongkun Zhang, and George Em Karniadakis (2020). "Physics-informed generative adversarial networks for stochastic differential equations". In: *SIAM Journal on Scientific Computing* 42.1, A292–A317.

Yang, Yingjie and Donald W Forsyth (2006). "Regional tomographic inversion of the amplitude and phase of Rayleigh waves with 2-D sensitivity kernels". In: *Geophysical Journal International* 166.3, pp. 1148–1160.

Yang, Yingjie, Michael H Ritzwoller, Anatoli L Levshin, and Nikolai M Shapiro (2007). "Ambient noise Rayleigh wave tomography across Europe". In: *Geophysical Journal International* 168.1, pp. 259–274.

Yao, Huajian, Caroline Beghein, and Robert D Van Der Hilst (2008). "Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis-II. Crustal and upper-mantle structure". In: *Geophysical Journal International* 173.1, pp. 205–219.

Yao, Huajian, Robert D van Der Hilst, and Maarten V De Hoop (2006). "Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—I. Phase velocity maps". In: *Geophysical Journal International* 166.2, pp. 732–744.

Ye, Zhuo, Rui Gao, Qiusheng Li, Hongshuang Zhang, Xuzhang Shen, Xuzhou Liu, and Chen Gong (2015). "Seismic evidence for the North China plate underthrusting beneath northeastern Tibet and its implications for plateau growth". In: *Earth and Planetary Science Letters* 426, pp. 109–117.

Yin, An (2010). "Cenozoic tectonic evolution of Asia: A preliminary synthesis". In: *Tectonophysics* 488.1-4, pp. 293–325.

Yoshizawa, K and BLN Kennett (2002). "Determination of the influence zone for surface wave paths". In: *Geophysical Journal International* 149.2, pp. 440–453.

– (2005). "Sensitivity kernels for finite-frequency surface waves". In: *Geophysical Journal International* 162.3, pp. 910–926.

Yu, Siwei and Jianwei Ma (2021). "Deep learning for geophysics: Current and future trends". In: *Reviews of Geophysics* 59.3, e2021RG000742.

Yu, Yong and Yongshun John Chen (2016). "Seismic anisotropy beneath the southern Ordos block and the Qinling-Dabie orogen, China: Eastward Tibetan asthenospheric flow around the southern Ordos". In: *Earth and Planetary Science Letters* 455, pp. 1–6.

Yuan, Huaiyu and Barbara Romanowicz (2010). "Lithospheric layering in the North American craton". In: *Nature* 466.7310, pp. 1063–1068.

Yuan, Pengyu, Shirui Wang, Wenyi Hu, Xuqing Wu, Jiefu Chen, and Hien Van Nguyen (2020). "A robust first-arrival picking workflow using convolutional and recurrent neural networks". In: *Geophysics* 85.5, U109–U119.

Zhang, Ruiyang, Yang Liu, and Hao Sun (2020a). "Physics-informed multi-LSTM networks for metamodeling of nonlinear structures". In: *Computer Methods in Applied Mechanics and Engineering* 369, p. 113226.

Zhang, Xiong, Jie Zhang, Congcong Yuan, Sen Liu, Zhibo Chen, and Weiping Li (2020b). "Locating induced earthquakes with a network of seismic stations in Oklahoma via a deep learning method". In: *Scientific reports* 10.1, p. 1941.

Zhao, Tianjie, Sheng Wang, Chaojun Ouyang, Min Chen, Chenying Liu, Jin Zhang, Long Yu, Fei Wang, Yong Xie, Jun Li, et al. (2024). "Artificial intelligence for geoscience: Progress, challenges and perspectives". In: *The Innovation*.

Zhong, Shi-Jun, Jian-Ping Wu, Li-Hua Fand, Wei-Lai Wang, Li-Ping Fan, and Huai-Fu Wang (2017). "Surface wave Eikonal tomography in and around the northeastern margin of the Tibetan plateau". In: *Chinese Journal of Geophysics* 60.6, pp. 2304–2314.

Zhu, Weiqiang and Gregory C Beroza (2019). "PhaseNet: a deep-neural-network-based seismic arrival-time picking method". In: *Geophysical Journal International* 216.1, pp. 261–273.

Zhu, Weiqiang, S Mostafa Mousavi, and Gregory C Beroza (2019). "Seismic signal denoising and decomposition using deep neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 57.11, pp. 9476–9488.

Zou, Jingbo, Cai Liu, Pengfei Zhao, and Chao Song (2024). "Seismic wavefields modeling with variable horizontally-layered velocity models via velocity-encoded PINN". In: *IEEE Transactions on Geoscience and Remote Sensing*.

Chapter 2

Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the northeastern margin of the Tibetan Plateau

Yunpeng Chen[1,2], Sjoerd A.L. de Ridder[1], Sebastian Rost[1], Zhen Guo[2], Xiaoyang Wu[2], and Yongshun Chen[2]

[1] *School of Earth and Environment, University of Leeds, LS29JT, Leeds, UK*

[2] *Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

Citation: Chen, Y., de Ridder, S. A., Rost, S., Guo, Z., Wu, X., & Chen, Y. (2022). Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the northeastern margin of the Tibetan Plateau. *Geophysical Research Letters, 49*(21), e2022GL099053.

Abstract

We present a novel eikonal tomography approach using physics-informed neural networks (PINNs) for Rayleigh wave phase velocities based on the eikonal equation. The PINN eikonal tomography (pinnET) neural network utilizes deep neural networks as universal function approximators and extracts traveltimes and velocities of the medium during the optimization process. Whereas classical eikonal tomography uses a generic non-physics based interpolation and regularization step to reconstruct traveltime surfaces, optimizing the network parameters in pinnET means solving a physics constrained traveltime surface reconstruction inversion tackling measurement noise and satisfying physics. We demonstrate this approach by applying it to 25 s surface wave data from ChinArray II sampling the northeastern (NE) Tibetan plateau. We validate our results by comparing them to results from conventional eikonal tomography in the same area and find good agreement.

Plain Language Summary

Eikonal tomography is an efficient approach to resolve velocity structure from surface wave data. Classical ambient noise eikonal tomography constrains the local phase velocity with all station pairs and uses a generic interpolation to reconstruct traveltime surfaces. Here we implement physics-informed neural networks for eikonal tomography. Unlike traditional neural networks that only rely on a large volume of data that is agnostic to physical laws, physics-informed neural network eikonal tomography (pinnET) can combine the data-driven model and theory-based model, which includes eikonal equation and boundary conditions as physical constrains on the system. We apply this pinnET to NE Tibetan plateau and our Rayleigh wave phase velocity results are quite similar with other established methods but use much less data.

2.1 Introduction

Eikonal tomography was proposed to directly obtain the phase velocity of surface waves from interpolated phase traveltime surfaces (Lin et al. 2009) and is often applied to ambient noise surface wave data. In ambient noise eikonal tomography, each station is considered as a 'source' and the traveltime surface can be estimated between this source and other stations whose recordings have been cross-correlated with those of the main source. Eikonal tomography constrains the phase velocity by locally evaluating the eikonal equation. The eikonal equation is a first-order nonlinear partial differential equation (PDE) that represents a high frequency approximation to the wave equation (Shearer 2019). The eikonal equation directly relates wave velocity and propagation directions to the spatial gradients of a traveltime surface (Lin and Ritzwoller 2011). Following Lin et al. (2009), a series of new applications in eikonal tomography soon followed (De Ridder 2011; Gouédard et al. 2012; De Ridder et al. 2015; Qiu et al. 2019).

In contrast to traditional surface wave tomography, eikonal tomography automatically accounts for ray bending and thereby provides a more accurate representation of wave propagation (Lin et al. 2009). However, classical eikonal tomography uses a generic interpolation algorithm to reconstruct the traveltime surfaces between stations in order to evaluate the eikonal equation that yields the velocity. This approach effectively biases the velocity depending on the particular algorithm used for interpolation. Linear interpolation enforces a homogeneous velocity between stations, whereas bicubic spline interpolation smoothens the traveltime surfaces reducing the resolution. One way to solve this problem is to introduce known physical constraints, similar to the introduction of the wave equation as PDE constraint in seismic waveform fitting (Van Leeuwen and Herrmann 2013; Van Leeuwen and Herrmann 2015; De Ridder and Maddison 2018; Shaiban et al. 2022). Here we argue that a physics based interpolation is easily achieved using neural networks (NNs) by formulating the surface wave tomography problem in a physics-informed neural network (PINN) framework.

Over the past decade, deep learning has been used as an efficient tool in many domains, such as image recognition (Krizhevsky et al. 2012), natural language processing (LeCun et al. 2015) and object detection (Pathak et al. 2018). In seismology, deep learning also found many applications to improve seismic data processing and imaging (Lim 2005; Zhang et al. 2014; Liu et al. 2018). Recently, an advanced deep learning framework, called physics-informed neural network (PINN), was proposed to solve PDEs, which imposes deep neural networks (DNNs) as universal function

approximators (Raissi et al. 2019). In contrast to classic DNNs, PINNs are better placed to combine data science and theory-based models. They leverage the mathematical descriptions of the physical process as constraints to the data driven deep learning approaches and determine the physical parameters during the training process. PINNs have successfully been included in other machine learning approaches, like transfer learning and meta-learning (Chakraborty 2021; Psaros et al. 2021), probabilistic PINNs (Grigo and Koutsourelakis 2019; Yang et al. 2021) and error analysis (Mishra and Molinaro 2020; Jiao et al. 2021). This idea is highly beneficial to seismic tomography for avoiding the iterative process required by the nonlinearity and can directly extract the predicted parameters (e.g., velocity) for the model. The PINN framework has already shown great potential in solving the seismic forward problem (Waheed et al. 2021; Smith et al. 2020; Moseley et al. 2020; Song et al. 2021) and seismic inverse problem (Song and Alkhalifah 2021). Waheed et al. (2021) suggested a PINN framework for exploration scale seismic tomography based on a factored eikonal equation, supported with synthetic examples.

Here we present a PINN-based algorithm for eikonal tomography and show the application in regional scale. The PINN framework is realized in SciAnn - a high-level deep learning library for physics-informed deep learning (Haghighat and Juanes 2021). We use the eikonal equation to define the loss function, which is used to describe the difference between predicted and true value, and train this NN to obtain the solution of the inversion by minimizing the loss function. In contrast to traditional eikonal tomography, this approach utilizes DNNs to optimize the field data and extract the traveltime and velocity during the optimization process. We will first formulate pinnET and then present the results of a field data trial using data from ChinArray II in the northeastern (NE) Tibetan plateau, see Figure 2.2(a).

2.2 Physics-informed Deep Learning for Eikonal Tomography

A classical NN is defined as a mathematical operation (a general function) that generates an output value given several input values. Training data is used to optimize the parameters of the NN such that the error between predicted output and true output (in the training data) is minimized (here using an L2 loss function):

$$\mathcal{L}(\theta_\tau) = \sum_{j}^{N_{src}} \sum_{i}^{N_{rcv}} \left| N_\tau(\theta_\tau; \mathbf{x}_{r;i}, \mathbf{x}_{s;j}) - \tau_{i,j} \right|^2, \tag{2.1}$$

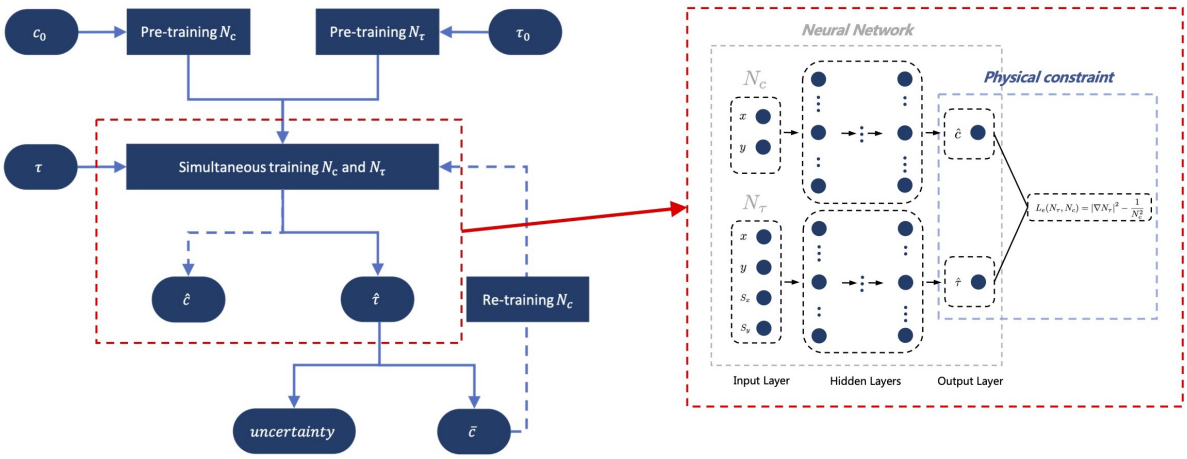


Figure 2.1: Workflow and PINN framework for traveltime eikonal tomography, where c_o and τ_0: initial phase velocity and traveltime; N_c and N_τ: neural networks of velocity and traveltime; τ: observed traveltime; $\hat{c}$ and $\hat{\tau}$: predicted phase velocity and traveltime from pinnET; $\bar{c}$: average phase velocity from all sources. In the PINN algorithm, the loss function L consists of two parts: the neural networks composed of N_c and N_τ used to minimize the misfit of the traveltime data providing the approximate solution and the physical constraint utilizing the approximate solution and adding the residual of physical equation. This physical constraint contains the eikonal equation and appropriate boundary conditions.

where i and j are the index of receivers and sources respectively, θ_τ are the hyperparameters of the traveltime NN N_τ, N_{src} represents the number of sources, N_{rcv} represents the number of receivers, $\mathbf{x}_{r;i} = (x_r, y_r)_i$ and $\mathbf{x}_{s;j} = (x_s, y_s)_j$ are the receiver and source locations of the spatial coordinates $\mathbf{x}$, $\tau_{i,j}$ represents the traveltime surfaces of seismic waves. $\mathcal{L}$ is the loss function.

Representing traveltimes by NNs does not guarantee that their output conforms to the physics of wave propagation. We add a physics constraint to the NN (forming a PINN). In contrast to classical NNs, PINNs ensure interpretability in NNs by combining data sets and physical constraints (Figure 2.1). In this study, we choose the eikonal equation as the governing physical law. The eikonal equation directly relates the local phase velocity to the local spatial gradients of the traveltime surface. In the high frequency approximation, the eikonal equation can be expressed in first order-hyperbolic form (e.g. (Wielandt 1993)):

$$|\nabla\tau(\mathbf{x})|^2 \approx \frac{1}{c^2(\mathbf{x})}, \tag{2.2}$$

where ∇ is the Laplace operator. The traveltime is expressed as a continuous scalar function $\tau(\mathbf{x})$, and $c(\mathbf{x})$ represents the local phase velocity on the surface of the Earth, at location $\mathbf{x} = (x, y)$. In our study, both the traveltimes and the phase velocity are expressed as outputs

of a NN, thus the physical constraint can be written as:

$$L_e(N_\tau, N_c) = |\nabla N_\tau|^2 - \frac{1}{N_c^2}, \tag{2.3}$$

where $N_\tau = N_\tau(\theta_\tau; \mathbf{x}_r, \mathbf{x}_s)$ is the traveltime at $\mathbf{x}_r$ from a source at $\mathbf{x}_s$, and $N_c = N_c(\theta_c; \mathbf{x}_r)$ is the phase velocity at x_r, representing the traveltime surfaces $\tau(\mathbf{x}_r, \mathbf{x}_s)$ and velocities $c(\mathbf{x}_r)$. θ_τ and θ_c are hyperparameters of traveltime and velocity NNs, respectively. Combining Equation 2.1 and 2.3, we define the PINN loss function:

$$\mathcal{L}(\theta_\tau, \theta_c) = \sum_{j}^{N_{src}} \sum_{i}^{N_{rcv}} \left[\left| N_\tau(\theta_\tau; \mathbf{x}_{r;i}, \mathbf{x}_{s;j}) - \tau_{i,j} \right|^2 + \epsilon_e \left| L_e(N_\tau(\theta_\tau; \mathbf{x}_{r;i}, \mathbf{x}_{s;j}), N_c(\theta_c; \mathbf{x}_{r;i})) \right|^2 \right], \tag{2.4}$$

where ϵ_e is a parameter that describes the relative weight of the eikonal constraint. Both terms of the loss function are evaluated at collocated points, coinciding with the observations locations of the traveltime data. If ϵ_e is chosen too large, the first estimate for the velocity in the eikonal equation constraint will completely control the outcome of the optimization. In that regime, the velocity NN will not be updated, and the velocity captured in the outputs of the traveltime NN will not deviate from the first estimate for the velocity. Alternatively, if ϵ_e is too small, the observed data will dominate this process and the results are evaluated without actually enforcing the physics constraint. So ϵ_e was chosen to be in between these two regimes, which we ascertained through tests was a rather narrow range. SciAnn (Haghighat and Juanes 2021) allows to specify physics constraints symbolically and handles them through automatic differentiation in the optimisation.

Here the training process is the only step needed to learn the correct physics. The training process aims to obtain the hyperparameters $\hat{\theta}_\tau$ and $\hat{\theta}_c$ by minimizing the loss function:

$$\arg\min_{\theta_\tau, \theta_c} \{\mathcal{L}(\theta_\tau, \theta_c)\} \rightarrow \hat{\theta}_\tau, \hat{\theta}_c. \tag{2.5}$$

In order to train a PINN, we would need the true velocity model of the Earth. The principle concept of pinnET is to update the velocity in the eikonal equation (physics) constraint, while training a NN that fits the observations.

We follow a training procedure that starts with individually pre-training the traveltime and velocity NNs using a starting velocity model, followed by iterated training in which the traveltime and velocity NNs may be updated simultaneously or sequentially (individually retraining the velocity NN), see Figure 2.1. The starting velocity c_0, for pre-training N_c, is calculated by dividing the sum of distances between all station pairs by the sum of observed traveltimes. The initial traveltimes τ_0, for pre-training N_τ, are calculated by dividing each distance between station pair by the initial velocity c_0. The traveltime NN is pre-trained using a loss function as in Equation 2.1, with $\tau_{i,j} = \tau_{0;i,j}$ for pre-training, and the velocity NN is pre-trained using:

$$\mathcal{L}(\theta_c) = \sum_{i}^{N_{grid}} \left[\left| N_c(\theta_c; \mathbf{x}_i) - c_i \right|^2, \right. \tag{2.6}$$

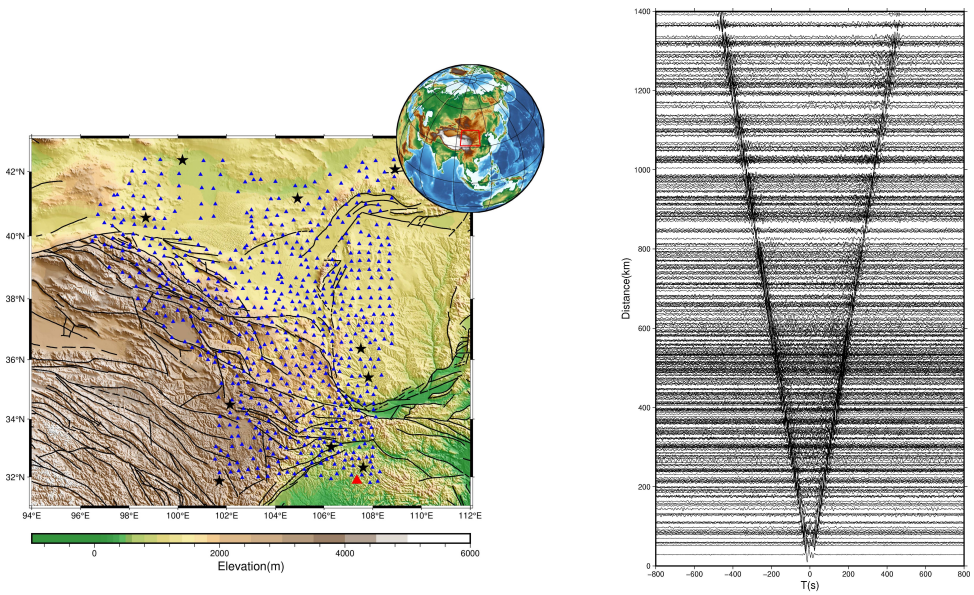
with $c_i = c_{0;i}$ for pre-training, where N_{grid} is the number of grid points. The pre-training process provides good initial weights for the NNs and results in a stable convergence during the process of iterated training. Simultaneous training aims to minimize the loss function in Equation 2.4. After this process, the velocity can be extracted in two different ways. The first method is by evaluating the velocity directly from the trained NNs, denoted $\hat{c}$:

$$\hat{c}(\mathbf{x}) = N_c(\hat{\theta}_c; \mathbf{x}), \tag{2.7}$$

where $\hat{c}$ is the trained velocities. The second method is by evaluating the traveltime NN on a fine regular grid using finite differences, yielding $\bar{c}$:

$$\bar{c}(\mathbf{x}) = \frac{1}{N_{src}} \sum_{j}^{N_{src}} \frac{1}{\left| \nabla N_\tau(\hat{\theta}_\tau; \mathbf{x}, \mathbf{x}_{s;j}) \right|}, \tag{2.8}$$

where $\bar{c}$ is the average velocity surfaces from all predicted traveltime surfaces, N_τ indicates the trained traveltime surfaces $\hat{\tau}$. In fact we employ a trimmed geometric mean (discarding the 10 percentile outliers at both extremes) and compute its corresponding standard deviation. The second method is relevant for sequential updates during which the velocity NN is updated explicitly by minimizing the loss function in Equation 2.6. We perform 2 outer iterations (with one sequential updating) to update the estimate for the velocity, and find that this strategy accelerates the overall convergence rate (see the Supplementary Material in Appendix A). An-



(a) (b)

Figure 2.2: (a) Overview of the station network in Tibet used in data example. Black lines indicate main faults, blue triangles indicate stations of the network and black stars indicate the selected source stations for the velocity determination. (b) An example of Z-Z component cross-correlations for station pairs at station 51511 (red triangle in (a)), the V-shaped arrivals are Rayleigh waves.

other advantage of using traveltime surfaces to predict the velocity is that it works robustly (parameter tuning is less tenuous) and comes with a measure of uncertainty (Figure 2.4(c)).

2.3 PINNs Eikonal Tomography on NE Tibetan Plateau

To demonstrate the suitability of our algorithm we apply it to a sample dataset of 25s period Rayleigh wave traveltimes collected from station of the ChinArray II network located on the NE Tibetan Plateau (Figure 2.2(a)). This dataset allows us to compare the pinnET approach with other eikonal tomography solutions, traditional surface wave tomography and ambient noise surface wave tomography results. The station distribution of the network is also well suited for our approach and the tectonic history of the region shows a variety of structures which allow us to test the resolution capabilities of our pinnET approach.

Since the collision of the Indian and Eurasian continents ($\sim$ 60 Ma), the Tibetan Plateau was elevated due to the N-S shortening of the crust (Yin and Harrison 2000). The physical processes that controlled the deformation and the mechanisms of crustal shortening remain subject for debate (Avouac and Tapponnier 1993; England and Molnar 1997; Clark and Royden 2000; Tapponnier et al. 2001; Yang et al. 2012; Guo and Chen 2017; Hao et al. 2021). The ChinArray II

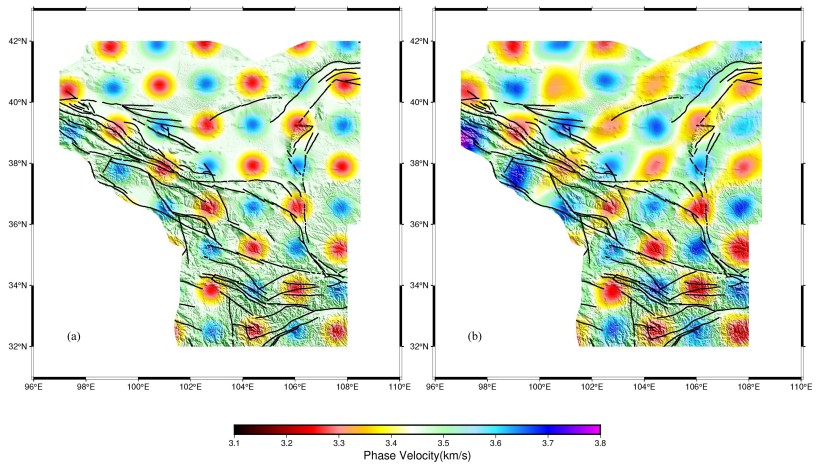


Figure 2.3: (a) Input synthetic velocity model of checkerboard resolution test for pinnET; (b) Retrieved Rayleigh wave phase velocity at periods of 25 s.

dataset has also been used to derive two-station Rayleigh wave tomography (Li et al. 2017), joint receiver functions and Rayleigh wave tomography (Wang et al. 2017), beamforming Rayleigh wave tomography (Wang et al. 2020) and Rayleigh wave eikonal tomography (Hao et al. 2021) allowing us to compare our results with more established methods. The traveltime data at each source station are obtained based on the seismic ambient noise cross-correlations between all station pairs and arrival time picking. Here we choose Rayleigh wave data at 25 s period as a proof-of-concept dataset to reveal phase velocity structure beneath NE Tibetan Plateau. In contrast to standard ambient noise or eikonal tomography we restrict the number of sources used in our approach to 10 sources recorded at all 676 stations of the network compared to all source-receiver combinations common in other methods. We randomly choose 10 sources to provide a good coverage of the model space. The training points are selected as the cross-correlated traveltime between those and other stations but we removed the points one kilometer around the sources to avoid singularity around them. The physics constraint is only enforced at the spatial locations of all the training traveltime data.

Our PINNs algorithm for traveltime eikonal tomography comprises two parts: the traveltime NN that is used to approximate the traveltime $\hat{t}_{s_i}$ and the velocity NN which aims to extract the velocity $\hat{c}(\mathbf{x})$. There are 10 hidden layers in both NN, but for the traveltime NN, each layer contains 20 neurons while only 5 neurons in each layer of velocity NNs. The size and numbers of layers in the NN where tuned to be able to represent all 10 traveltime surfaces with sufficient accuracy as not to reduce uncertainty in the final velocity. The '*arctan*' function is

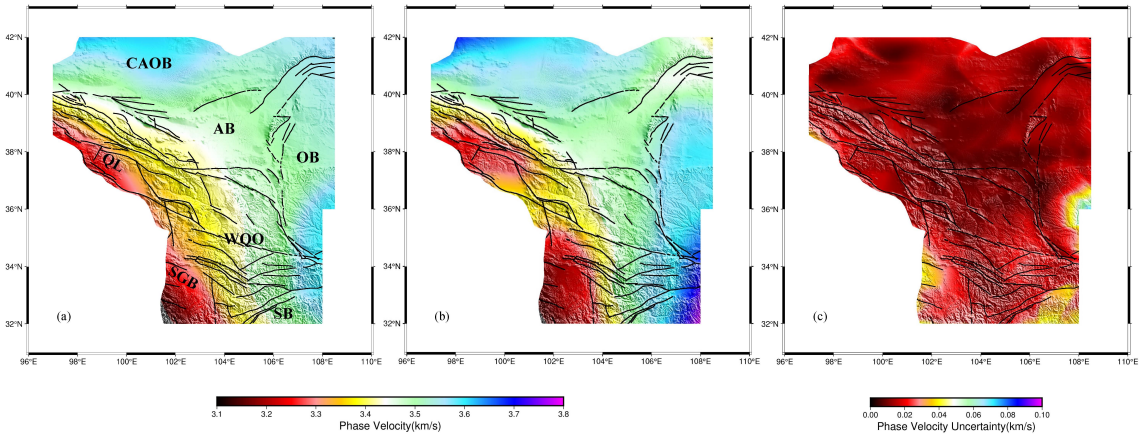


Figure 2.4: (a) The predicted 25s phase velocity map beneath NE Tibetan Plateau using pinnET (Block and structure are as follows: AB = Alxa Block; CAOB = Central Asian Orogenic Belt; OB = Ordos Block; QL = Qilian; SB = Sichuan basin; SGB = Songpan-Ganzi Block; WQO = Western Qinling Orogen); (b) The 25s phase velocity structure generated by conventional eikonal tomography; (c) The 25s phase velocity uncertainty map.

used as the activation function for the hidden layers. The optimizer is defined as the '*adam*' optimizer (Kingma and Ba 2014), which is computationally efficient and widely used in deep learning. We divide all samples into a batch size of 676 and train the networks for 2000 epochs. The learning rate for all samples is 0.0002. These hyperparameters were all chosen based on systematic synthetic tests.

We use checkerboard tests to evaluate the resolution of our results (Figure 2.3). The synthetic velocity model has a background velocity of 3.45 km/s on 8 $\times$ 8 anomalies. The magnitudes of the velocity perturbations are 2 % and these anomalies have a maximum radius of 75 km distributed in the latitude and longitude direction, respectively. The distance between alternating low and high velocity anomalies is therefore 150 km (Figure 2.3(a)). We choose the same number and distribution of 676 stations with 10 sources (Figure 2.2(a)) for the checkerboard resolution tests as in the recorded data. Gaussian noise with a mean of 0.1 s and a standard deviation of 0.01 has been added to the traveltime data to simulate noise in the observed data. Figure 2.3(b) shows the recovered velocity structure by PINNs eikonal tomography. Most anomalies can be well recovered when the ray coverage is sufficient. Less well sampled structure at the edge of station networks is still acceptably resolved. We observe some lateral smearing especially in SW-NE direction in the north of the network which could be related to the choice of the source distribution. The checkerboard tests show that our phase velocity results are reliable with these parameters at 25 s.

We apply our pinnET approach to recorded ambient noise data from the Tibetan Plateau (Figure 2.2(a)). The convergence rate of the different training processes is shown by the loss curves in Figure A.1 in the Supplementary Material. Using 10 sources to all stations and averaging the velocities from the solution for different sources (Equation 2.8), we resolve the phase velocity structure (Figure 2.4(a)) beneath NE Tibetan Plateau. Comparing with solutions derived from other Rayleigh wave tomography approaches, we find very good agreement. We show a direct comparison with the results at 25 s by ambient noise eikonal tomography. The phase velocities we find in the region are within $\pm$ 0.1 km/s of other methods. We find good agreement with the lower velocities beneath the Songpang-Ganzi block, Qilian and Western Qinling Orogen. At the southeastern and northwestern edges of the network we find lower velocities beneath the Sichuan Basin and the Central Asian Orogenic Belt. This could be because bicubic spline interpolation (used in conventional eikonal tomography) flattens the interpolated traveltime surfaces near the edges of the data coverage. Boundaries of all features are in good agreement with other studies (Zhong et al. 2017; Wang et al. 2020). Figure 2.4(c) shows the Rayleigh wave local phase velocity uncertainty. The uncertainty of most area across the map is less than 0.02 km/s. Large uncertainties occur in a small part of the southeastern Ordos Block, the western Songpan-Ganzi Block and the eastern Sichuan basin, these are most likely due to data coverage in the random selection of 10 sources.

The observed difference between the phase velocity maps resulting from conventional eikonal tomography result and pinnET could be considered significant (in some areas in the order of 100 m/s). The resolution of pinnET (in general) depends on the observation station spacing, the trade off between the measurement errors in the data and the number of traveltime sets (source instances) used, and on the hyperparameters of the NN, and parameters controlling the training process. The method is certainly memory efficient because compressing the traveltimes as outputs to a NN is a concept akin to compressed sensing. We use much less data to achieve the similar imaging with a benefit of including the physics constraint while reconstructing the traveltime surfaces.

We restricted ourselves to an inversion at 25 s period. Data at other periods can be inverted using pinnET in the same way leading to wide spectrum phase velocity maps that can be inverted from surface wave tomography using traditional approaches or neural networks.

2.4 Conclusions

We present a novel method for traveltime eikonal tomography using PINNs and apply it to recorded field data. The method leverages neural networks as universal function approximators and utilizes the estimated medium properties in the eikonal equation which are treated as underlying physical laws. Reconstruction inversions based on NN function approximators are memory efficient. The hyperparameters of the NNs should be carefully selected to ensure the appropriate resolution of pinnET, e.g. through a checkerboard test. The final algorithm is an eikonal tomography that uses physics consistent interpolation while reconstructing (interpolating) the traveltimes. The reconstruction inversion mitigates errors in the original traveltime measurements and the number of traveltime sets used in the training can be significantly reduced.

We applied the PINNs eikonal tomography on seismic data recorded by ChinArray II installed over the NE Tibetan Plateau. We extracted Rayleigh wave phase velocities at 25 s and associated uncertainties using only 10 sources. The results compares well to the velocity structure obtained by the conventional eikonal tomography using all 676 sources of the seismic networks.

Open Research Section

The traveltime data sets used in this study and the Rayleigh wave phase velocity models can be downloaded at `https://doi.org/10.5281/zenodo.7223219`.

Acknowledgments

We acknowledge China Seismic Array Data Management Center at Institute of Geophysics, China Earthquake Administration for providing ambient noise data used in this study. We thank all the people who participated in the field deployment of seismic arrays for the ChinArray II project. This work is supported by the National Science Foundation of China (Grants 41890814 and U1901602). YP.C. is also supported by a split-site PhD program of SUSTech and the University of Leeds. Z.G. is also supported by Key Special Project for Introduced Talents Team of the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0210).

References

Avouac, Jean-Philippe and Paul Tapponnier (1993). "Kinematic model of active deformation in central Asia". In: *Geophysical Research Letters* 20.10, pp. 895–898.

Chakraborty, Souvik (2021). "Transfer learning based multi-fidelity physics informed deep neural network". In: *Journal of Computational Physics* 426, p. 109942.

Clark, Marin Kristen and Leigh Handy Royden (2000). "Topographic ooze: Building the eastern margin of Tibet by lower crustal flow". In: *Geology* 28.8, pp. 703–706.

De Ridder, SA (2011). "Ambient seismic noise tomography for exploration seismology at Valhall". In: *AGU Fall Meeting Abstracts*. Vol. 2011, S31F–04.

De Ridder, SAL, BL Biondi, and D Nichols (2015). "Elliptical-anisotropic eikonal phase velocity tomography". In: *Geophysical Research Letters* 42.3, pp. 758–764.

De Ridder, SAL and JR Maddison (2018). "Full wavefield inversion of ambient seismic noise". In: *Geophysical Journal International* 215.2, pp. 1215–1230.

England, Philip and Peter Molnar (1997). "Active deformation of Asia: From kinematics to dynamics". In: *Science* 278.5338, pp. 647–650.

Gouédard, Pierre, Huajian Yao, Fabian Ernst, and Robert D van der Hilst (2012). "Surface wave eikonal tomography in heterogeneous media using exploration data". In: *Geophysical Journal International* 191.2, pp. 781–788.

Grigo, Constantin and Phaedon-Stelios Koutsourelakis (2019). "A physics-aware, probabilistic machine learning framework for coarse-graining high-dimensional systems in the Small Data regime". In: *Journal of Computational Physics* 397, p. 108842.

Guo, Zhen and Yongshun John Chen (2017). "Mountain building at northeastern boundary of Tibetan Plateau and craton reworking at Ordos block from joint inversion of ambient noise tomography and receiver functions". In: *Earth and Planetary Science Letters* 463, pp. 232–242.

Haghighat, Ehsan and Ruben Juanes (2021). "Sciann: A keras/tensorflow wrapper for scientific computations and physics-informed deep learning using artificial neural networks". In: *Computer Methods in Applied Mechanics and Engineering* 373, p. 113552.

Hao, Shijie, Zhouchuan Huang, Cunrui Han, Liangshu Wang, Mingjie Xu, Ning Mi, and Dayong Yu (2021). "Layered crustal azimuthal anisotropy beneath the northeastern Tibetan Plateau revealed by Rayleigh-wave Eikonal tomography". In: *Earth and Planetary Science Letters* 563, p. 116891.

Jiao, Yuling, Yanming Lai, Yisu Lo, Yang Wang, and Yunfei Yang (2021). "Error analysis of deep Ritz methods for elliptic equations". In: *arXiv preprint arXiv:2107.14478*.

Kingma, Diederik P and Jimmy Ba (2014). "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980*.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton (2012). "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems* 25, pp. 1097–1105.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). "Deep learning". In: *nature* 521.7553, pp. 436–444.

Li, Yonghua, Jiatie Pan, Qingju Wu, and Zhifeng Ding (2017). "Lithospheric structure beneath the northeastern Tibetan Plateau and the western Sino-Korea Craton revealed by Rayleigh wave tomography". In: *Geophysical Journal International* 210.2, pp. 570–584.

Lim, Jong-Se (2005). "Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea". In: *Journal of Petroleum Science and Engineering* 49.3-4, pp. 182–192.

Lin, Fan-Chi and Michael H Ritzwoller (2011). "Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure". In: *Geophysical Journal International* 186.3, pp. 1104–1120.

Lin, Fan-Chi, Michael H Ritzwoller, and Roel Snieder (2009). "Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array". In: *Geophysical Journal International* 177.3, pp. 1091–1110.

Liu, Dawei, Wei Wang, Wenchao Chen, Xiaokai Wang, Yanhui Zhou, and Zhensheng Shi (2018). "Random noise suppression in seismic data: What can deep learning do?" In: *SEG Technical Program Expanded Abstracts 2018*. Society of Exploration Geophysicists, pp. 2016–2020.

Mishra, Siddhartha and Roberto Molinaro (2020). "Estimates on the generalization error of Physics Informed Neural Networks (PINNs) for approximating a class of inverse problems for PDEs". In: *arXiv preprint arXiv:2007.01138*.

Moseley, Ben, Andrew Markham, and Tarje Nissen-Meyer (2020). "Solving the wave equation with physics-informed deep learning". In: *arXiv preprint arXiv:2006.11894*.

Pathak, Ajeet Ram, Manjusha Pandey, and Siddharth Rautaray (2018). "Application of deep learning for object detection". In: *Procedia computer science* 132, pp. 1706–1717.

Psaros, Apostolos F, Kenji Kawaguchi, and George Em Karniadakis (2021). "Meta-learning PINN loss functions". In: *arXiv preprint arXiv:2107.05544*.

Qiu, Hongrui, Fan-Chi Lin, and Yehuda Ben-Zion (2019). "Eikonal tomography of the Southern California plate boundary region". In: *Journal of Geophysical Research: Solid Earth* 124.9, pp. 9755–9779.

Raissi, Maziar, Paris Perdikaris, and George E Karniadakis (2019). "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In: *Journal of Computational Physics* 378, pp. 686–707.

Shaiban, Ali, SAL De Ridder, and Andrew Curtis (2022). "Wavefield reconstruction and wave equation inversion for seismic surface waves". In: *Geophysical Journal International* 229.3, pp. 1870–1880.

Shearer, Peter M (2019). *Introduction to seismology*. Cambridge university press.

Smith, Jonathan D, Kamyar Azizzadenesheli, and Zachary E Ross (2020). "Eikonet: Solving the eikonal equation with deep neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 59.12, pp. 10685–10696.

Song, Chao, Tariq Alkhalifah, and Umair Bin Waheed (2021). "Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks". In: *Geophysical Journal International* 225.2, pp. 846–859.

Song, Chao and Tariq A Alkhalifah (2021). "Wavefield reconstruction inversion via physics-informed neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–12.

Tapponnier, Paul, Xu Zhiqin, Francoise Roger, Bertrand Meyer, Nicolas Arnaud, Gérard Wittlinger, and Yang Jingsui (2001). "Oblique stepwise rise and growth of the Tibet Plateau". In: *science* 294.5547, pp. 1671–1677.

Van Leeuwen, Tristan and Felix J Herrmann (2013). "Mitigating local minima in full-waveform inversion by expanding the search space". In: *Geophysical Journal International* 195.1, pp. 661–667.

– (2015). "A penalty method for PDE-constrained optimization in inverse problems". In: *Inverse Problems* 32.1, p. 015007.

Waheed, Umair bin, Ehsan Haghighat, Tariq Alkhalifah, Chao Song, and Qi Hao (2021). "PINNeik: Eikonal solution using physics-informed neural networks". In: *Computers & Geosciences* 155, p. 104833.

Waheed, Umair bin, Tariq Alkhalifah, Ehsan Haghighat, Chao Song, and Jean Virieux (2021). "PINNtomo: Seismic tomography using physics-informed neural networks". In: *arXiv preprint arXiv:2104.01588*.

Wang, Kaiming, Laiyu Lu, Valérie Maupin, Zhifeng Ding, Chen Zheng, and Shijun Zhong (2020). "Surface wave tomography of northeastern Tibetan Plateau using beamforming of seismic noise at a dense array". In: *Journal of Geophysical Research: Solid Earth* 125.4, e2019JB018416.

Wang, Xingchen, Yonghua Li, Zhifeng Ding, Lupei Zhu, Chunyong Wang, Xuewei Bao, and Yan Wu (2017). "Three-dimensional lithospheric S wave velocity model of the NE Tibetan Plateau and western North China Craton". In: *Journal of Geophysical Research: Solid Earth* 122.8, pp. 6703–6720.

Wielandt, E (1993). "Propagation and structural interpretation of non-plane waves". In: *Geophysical Journal International* 113.1, pp. 45–53.

Yang, Liu, Xuhui Meng, and George Em Karniadakis (2021). "B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data". In: *Journal of Computational Physics* 425, p. 109913.

Yang, Yingjie, Michael H Ritzwoller, Yong Zheng, Weisen Shen, Anatoli L Levshin, and Zujun Xie (2012). "A synoptic view of the distribution and connectivity of the mid-crustal low velocity zone beneath Tibet". In: *Journal of Geophysical Research: Solid Earth* 117.B4.

Yin, An and T Mark Harrison (2000). "Geologic evolution of the Himalayan-Tibetan orogen". In: *Annual review of earth and planetary sciences* 28.1, pp. 211–280.

Zhang, Chiyuan, Charlie Frogner, M Araya-Polo, and D Hohl (2014). "Machine-learning based automated fault detection in seismic traces". In: *76th EAGE Conference and Exhibition 2014*. Vol. 2014. 1. European Association of Geoscientists & Engineers, pp. 1–5.

Zhong, Shi-Jun, Jian-Ping Wu, Li-Hua Fand, Wei-Lai Wang, Li-Ping Fan, and Huai-Fu Wang (2017). "Surface wave Eikonal tomography in and around the northeastern margin of the Tibetan plateau". In: *Chinese Journal of Geophysics* 60.6, pp. 2304–2314.

Chapter 3

Physics-informed neural networks for elliptical-anisotropy eikonal tomography: application to data from the northeastern Tibetan Plateau

Yunpeng Chen[1,2], Sjoerd A.L. de Ridder[1], Sebastian Rost[1], Zhen Guo[2], Xiaoyang Wu[2], Shilin Li[2]and Yongshun Chen[2]

[1]*School of Earth and Environment, University of Leeds, LS29JT, Leeds, UK*

[2]*Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

Citation: Chen, Y., de Ridder, S. A., Rost, S., Guo, Z., Wu, X., Li, S., & Chen, Y. (2023). Physics-Informed Neural Networks for Elliptical-Anisotropy Eikonal Tomography: Application to Data From the Northeastern Tibetan Plateau. *Journal of Geophysical Research: Solid Earth, 128*(12), e2023JB027378.

Abstract

We develop a novel approach for multi-frequency, elliptical-anisotropic eikonal tomography based on physics-informed neural networks (pinnEAET). This approach simultaneously estimates the medium properties controlling anisotropic Rayleigh waves and reconstructs the traveltimes. The physics constraints built into pinnEAET's neural network enable high-resolution results with limited inputs by inferring physically plausible models between data points. Even with a single source, pinnEAET can achieve stable convergence on key features where traditional methods lack resolution. We apply pinnEAET to ambient noise data from a dense seismic array (ChinArray-Himalaya II) in the northeastern Tibetan Plateau with only 20 quasi-randomly distributed stations as sources. Anisotropic phase velocity maps for Rayleigh waves in the period range from 10-40 s are obtained by training on observed traveltimes. Despite using only about 3% of the total stations as sources, our results show low uncertainties, good resolution and are consistent with results from conventional tomography.

Plain Language Summary

Anisotropy refers to the directional dependence of seismic wave velocities, which can arise from a variety of factors such as crystal alignment, stress fields, or fluid-filled cracks. Elliptical-anisotropic eikonal tomography is a variant of eikonal tomography that can be used to estimate medium properties and reconstructed traveltimes from ambient noise data. In this study, we propose a new algorithm to implement multi-frequency, elliptical-anisotropic eikonal tomography based on physics-informed neural networks (pinnEAET), which combine data-driven models with theory-based models that include physics constraints on the system. We apply this architecture to data from a dense seismic array deployed on the northeastern Tibetan Plateau. Our results can achieve at least the same resolution as traditional methods while requiring less traveltime data. This strategy can provide new insights into the seismic imaging in case of limited or noisy data.

3.1 Introduction

Seismic anisotropy is the characteristic that seismic wave velocity varies with propagation direction, providing essential structural constraints on geodynamic evolution of the Earth. For example, crustal and upper mantle deformation often leads to seismic anisotropy, either due to lattice preferred orientation of anisotropic constituent minerals, or caused by shaped preferred orientation of isotropic materials with distinct shapes (Crampin and Booth 1985; Nicolas and Christensen 1987; Silver 1996; Montagner and Guillot 2002; Long and Becker 2010). Different measurements have been proposed to reveal anisotropic structure such as shear wave splitting (Savage 1999; Li et al. 2011), surface wave tomography (Yao et al. 2010; Russell et al. 2019), body wave tomography (Zhao et al. 2016; Creasy et al. 2019) and receiver functions (Schulte-Pelkum and Mahan 2014; Zheng et al. 2021). Recovering seismic anisotropy using seismic tomography has key implications for understanding the dynamic processes of tectonic units.

In contrast to other methods, surface wave tomography has better vertical resolution of seismic anisotropy, making it a critical tool for measuring anisotropy in the crust and upper mantle (Simons et al. 2002; Ekström 2011; Becker et al. 2012). In case of sufficient azimuthal path coverage, phase velocity and azimuthal anisotropy can be derived either from earthquake data (Montagner 1986; Romanowicz 2002; Yao et al. 2010) or ambient noise cross-correlations (Wapenaar et al. 2010; Ritzwoller et al. 2011). Compared with earthquake-based surface wave tomography, ambient noise tomography enables substantial additional coverage at short periods (Shapiro et al. 2005). Numerous array-based tomographic methods have been developed to derive anisotropy from ambient noise, such as eikonal and Helmholtz tomography (Lin et al. 2009; Lin and Ritzwoller 2011), wave gradiometry (De Ridder and Curtis 2017; Cao et al. 2020) and beamforming (Wu et al. 2023; Soergel et al. 2023).

Eikonal tomography is a surface wave tomography method that relates phase front tracking to local propagation direction dependent wave speed (Lin et al. 2009). Conventional eikonal tomography estimates azimuthal anisotropy by fitting a parametric function to phase velocities from different azimuths, obtained using an isotropic eikonal equation. De Ridder et al. (2015) proposed an elliptical-anisotropic eikonal tomography which employs ellipse parameters to characterize anisotropic velocity structures with an anisotropic eikonal equation. This method enables explicit regularization of the medium parameters and retrieves azimuthal anisotropic

velocity robustly (De Ridder et al. 2015). However, the interpolation schemes in these eikonal tomography methods reduce the resolution since it smooths the phase traveltime surfaces beyond the level of the inter-station spacing.

In recent years, deep learning has been applied to diverse aspects of geoscience, such as geology (Kim and Yun 2021; Ho et al. 2023), geomorphology (Li et al. 2020), geochemistry (Luo et al. 2020) and geophysics (Yang and Ma 2019). Particularly, a broad range of methods using deep neural networks (DNNs) in seismology are proposed to leverage the increasing amount of observed data for modeling, prediction, detection and classification (Yu and Ma 2021). These applications of utilizing deep learning to solve seismic problems include but are not limited to arrival time picking (Ross et al. 2018), seismic data processing (Zhu et al. 2019), earthquake location (Mousavi et al. 2020) and seismic imaging (Araya-Polo et al. 2018). As a data-driven method, classical DNNs have been proven to overcome some limitations in traditional seismic methods. But in these applications, the training data often imply some prior knowledge that is ignored in classical deep learning methods. This purely data-driven training is more susceptible to observation errors and leads to poor generalization of predicting seismic wave propagation outside of the span of the training data. Combining data-driven and physics-based models is a promising way for seismic machine learning.

Physics-informed neural networks (PINNs), a new kind of deep learning framework, were proposed by Raissi et al. (2019) to solve forward and inverse problems of partial differential equations (PDEs). PINNs utilize the capability of automatic differentiation which is widely used in training DNN and it also allows adding underlying physical laws to the loss function to combine the data-based model and theory-based model during the training process (Baydin et al. 2018; Karpatne et al. 2017). PINNs have gradually become a research highlight of scientific machine learning within various fields, such as fluid mechanics (Raissi et al. 2020), material science (Fang and Zhan 2019), biomedicine (Kissas et al. 2020) and power systems (Misyris et al. 2020). During these applications, PINNs can train accurate and generalized models even with limited data by automatically incorporating physics constraints of the system.

Significantly, PINNs have already shown great potential in seismological applications. For forward problems, PINNs have been applied to the eikonal equation for traveltime calculation in isotropic and anisotropic media (Waheed et al. 2020; Smith et al. 2020; Waheed et al. 2021; Taufik et al. 2022) and directly simulate wave equation solutions for acoustic and elastic wave

propagation (Alkhalifah et al. 2020; Karimpouli and Tahmasebi 2020; Moseley et al. 2020a; Moseley et al. 2020b; Song et al. 2021; Song et al. 2022; Song and Wang 2023). For inverse problems, PINNs have been proposed for exploration-scale seismic tomography with the factored eikonal equation (Waheed et al. 2021; Gou et al. 2023) and wavefield reconstruction inversion (Song and Alkhalifah 2021). A PINN algorithm has also been developed for full waveform inversion, as demonstrated through various synthetic case studies (Rasht-Behesht et al. 2022). Recently, Chen et al. (2022) presented the first application of PINNs to field seismic data for eikonal tomography, demonstrating their feasibility on real datasets and at scale. However, PINN-based tomography methods, especially for anisotropic models of field data, remains an open area for continued progress.

Here we propose a PINN based elliptical-anisotropy eikonal tomography (pinnEAET) to retrieve Rayleigh wave phase velocity and azimuthal anisotropy for multiple frequencies simultaneously. Two kinds of neural networks (NNs) are used to represent azimuthal dependence of phase velocities related matrices and phase traveltime surfaces (Figure 3.1). We apply our algorithm to field data gained from a dense network of stations on the northeastern Tibetan Plateau and discuss the advantages and challenges of the proposed approach.

3.2 Methods

In this section, we will introduce the fundamental aspects of elliptical-anisotropic eikonal equation, then describe incorporating this PDE into a PINN framework, and finally share the idea of estimating the Rayleigh wave anisotropic phase velocity from the trained networks.

3.2.1 Elliptical-anisotropic eikonal equation

In the wave number domain, the phase velocity $c = c(\phi, \boldsymbol{x}, \omega)$ at location $\boldsymbol{x} = (x, y)$ and frequency ω, where ϕ is the direction of wave propagation, exhibiting elliptical-anisotropic azimuthal anisotropy can be defined as (De Ridder et al. 2015):

$$c^2(\phi) = c_f^2 cos^2(\phi - \alpha) + c_s^2 sin^2(\phi - \alpha), \tag{3.1}$$

where $c_f = c_f(\boldsymbol{x}, \omega)$ and $c_s = c_s(\boldsymbol{x}, \omega)$ are the fast and slow velocities, respectively. $\alpha = \alpha(\boldsymbol{x}, \omega)$ is the azimuth of the fast direction. Equation 3.1 describes an ellipse with the major and minor

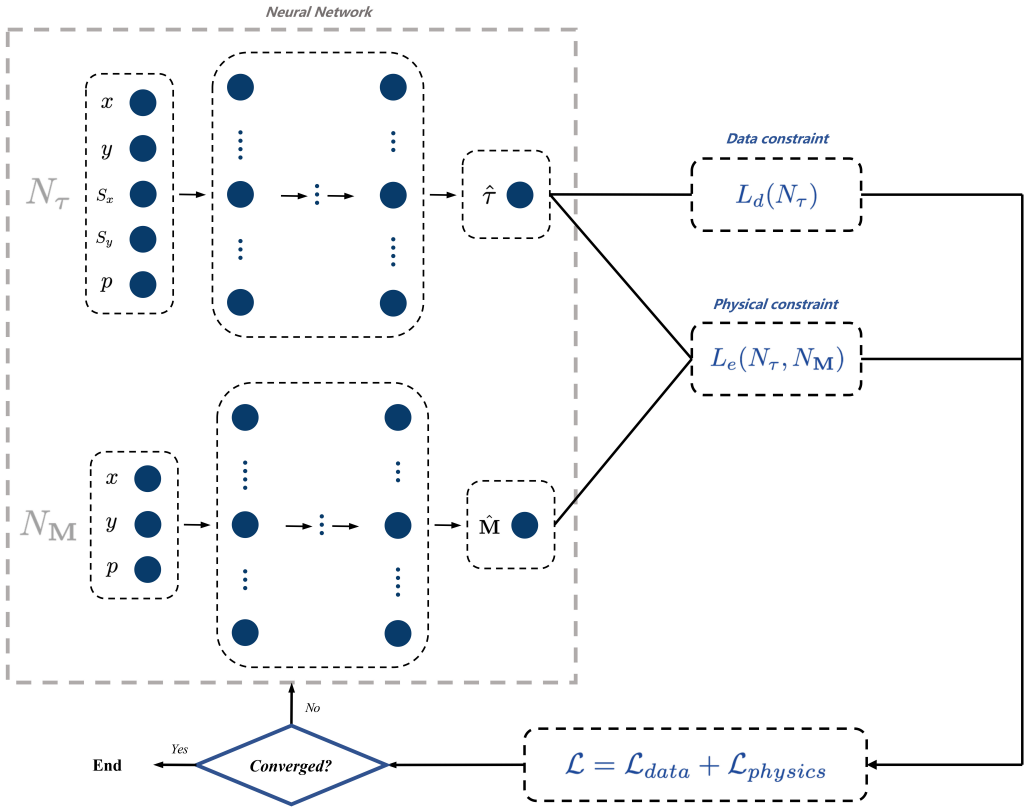


Figure 3.1: Schematic of PINN framework for elliptical-anisotropic eikonal tomography, where N_τ and $N_{\boldsymbol{M}}$ are the traveltime and medium property NNs, x, y and s_x, s_y are spatial location coordinates of receivers and sources respectively, p contains discrete distinct periods for the multi-period solutions, $\hat{\tau}$ and $\hat{\boldsymbol{M}}$ are the outputs of trained traveltime and medium properties, $L_d(N_\tau)$ represents the observed Rayleigh wave traveltime data constraint, $L_e(N_\tau, N_{\boldsymbol{M}})$ represents the elliptical-anisotropic eikonal equation constraint. The neural network part on the left is used to approximate the solution, while the data and physics constraints part on the right is used to optimize the network parameters. The training process ends when the model is considered converged.

radii corresponding to c_f and c_s.

To derive an eikonal equation for elliptical anisotropy, the scalar wavefield is transformed from the wave number domain to the spatial domain and the phase traveltime $\tau = \tau(\boldsymbol{x}, \boldsymbol{x}_s, \omega)$ for a given source location $\boldsymbol{x}_s = (x_s, y_s)$ at a specific frequency ω (also known as the linear phase) is introduced (Aki and Richards 2002). In the high-frequency approximation, we can find the elliptical-anisotropic eikonal equation (De Ridder et al. 2015):

$$1 = \begin{bmatrix} \partial_x \tau & \partial_y \tau \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \partial_x \tau \\ \partial_y \tau \end{bmatrix}. \tag{3.2}$$

The matrix elements are respectively expressed as:

$$M_{11} = M_{11}(\boldsymbol{x}, \omega) = (c_f^2 - c_s^2)sin^2(\alpha) + c_s^2, \tag{3.3}$$

$$M_{12} = M_{21} = M_{12}(\boldsymbol{x}, \omega) = (c_f^2 - c_s^2)sin(\alpha)cos(\alpha), \tag{3.4}$$

$$M_{22} = M_{22}(\boldsymbol{x}, \omega) = (c_f^2 - c_s^2)cos^2(\alpha) + c_s^2. \tag{3.5}$$

Then the problem of solving the phase velocities and azimuthal anisotropy converts to computing the eigenvalues (c_f^2 and c_s^2, representing the fast and slow velocities) and eigenvectors (indicating the fast and slow directions) of the matrix $\boldsymbol{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$. For the 2×2 matrices $\boldsymbol{M}$, there is an explicit algebraic solution of the eigenvalues and eigenvectors (see Appendix B.1). In this way, the fast and slow velocity can be computed as:

$$c_f = c_f(\boldsymbol{x}, \omega) = \sqrt{\frac{(M_{11} + M_{22} + \sqrt{(M_{11} + M_{22})^2 - 4(M_{11}M_{22} - M_{12}^2)})}{2}}, \tag{3.6}$$

$$c_s = c_s(\boldsymbol{x}, \omega) = \sqrt{\frac{(M_{11} + M_{22} - \sqrt{(M_{11} + M_{22})^2 - 4(M_{11}M_{22} - M_{12}^2)})}{2}}, \tag{3.7}$$

and if M_{12} is not zero, the fast direction of anisotropy can be defined as:

$$\alpha(\boldsymbol{x}, \omega) = \arctan(\frac{c_f - M_{22}}{M_{12}}). \tag{3.8}$$

Note that this is just one specific way to represent the fast and slow velocity and the azimuthal angle. Consequently, the isotropic component, which is the radius of a circle with an area equal to the velocity ellipse (Weisstein 2014), can be calculated as:

$$c_0 = c_0(\boldsymbol{x}, \omega) = \sqrt{c_f c_s}. \tag{3.9}$$

Alternatively, we could have used the average between the fast and slow velocity $c_0 = (c_f + c_s)/2$. The anisotropic component can be characterized as either eccentricity, $e = e(\boldsymbol{x}, \omega) = \sqrt{1 - \frac{c_s^2}{c_f^2}}$, flattening factor, $f = f(\boldsymbol{x}, \omega) = 1 - \frac{c_s}{c_f}$ (Weisstein 2014), or equivalent amplitude (Smith and Dahlen 1973):

$$A = A(\boldsymbol{x}, \omega) = \frac{c_f - c_s}{2} \tag{3.10}$$

Here we use equivalent amplitude A to represent the strength of anisotropy commonly used in other studies (Lin et al. 2009; Hao et al. 2021; Kästle et al. 2022).

3.2.2 Elliptical-anisotropic eikonal tomography using PINNs

In order to approach this problem, we first consider a feed-forward NN (see Figure 3.1), which consists of L layers with network parameters $\boldsymbol{W}$ and $\boldsymbol{b}$:

$$\boldsymbol{z}^l = \sigma(\boldsymbol{W}^l \cdot \boldsymbol{z}^{l-1} + \boldsymbol{b}^l), \quad l = 1, \ldots, L, \tag{3.11}$$

where L is the maximum number of layers with l being the layer. $\boldsymbol{z}$ refers to the affine transformations between the layers of NN, when $l = 1$, $\boldsymbol{z}^0$ represents inputs of the model; when $l = L$, $\boldsymbol{z}^L = N(\boldsymbol{W}^1, \ldots, \boldsymbol{W}^L, \boldsymbol{b}^1, \ldots, \boldsymbol{b}^L; \boldsymbol{z}^0)$, where N is the NN, represents the outputs of the model and other $\boldsymbol{z}^l$, $\quad l = 2, \ldots, L-1$ represent the neurons in hidden layers. σ is the activation function that is usually nonlinear. $\boldsymbol{W}^l$ and b^l are the weights and the biases of each layer l, respectively. Training a feed-forward NN involves passing data through interconnected layers that learn to map features in a hierarchical fashion, with backward-propagated weight adjustments.

We take the traveltime $\tau = \tau(\boldsymbol{x}_r, \boldsymbol{x}_s, \omega)$, where $\boldsymbol{x}_r, \boldsymbol{x}_s$ are receiver and source locations, as the field value of the NN architecture. Whereas the trainable parameters ($\boldsymbol{W}$ and $\boldsymbol{b}$) are represented with θ, the chosen parameters such as number of layers, neurons and type of activation functions are unique for each NN and denoted with the subscript to N. Then a transformation of Equation 3.11 can be used to represent an approximate solution for phase traveltime surfaces $\tau = \tau(\boldsymbol{x}, \boldsymbol{x}_s, \omega)$ in the eikonal equation:

$$\tau(\boldsymbol{x}, \boldsymbol{x}_s, \omega) = N_\tau(\theta_\tau; \boldsymbol{x}, \boldsymbol{x}_s, \omega) = N_\tau(\boldsymbol{W}^1, \ldots, \boldsymbol{W}^L, \boldsymbol{b}^1, \ldots, \boldsymbol{b}^L; \boldsymbol{z}^0) \tag{3.12}$$

where θ_τ indicates the trainable parameters of the traveltime NN, N_τ. In this case, the inputs in Equation 3.11 are $\boldsymbol{z}^0 = (\boldsymbol{x}_r, \boldsymbol{x}_s, \omega)$ for specific coordinates and the outputs are $\boldsymbol{z}^L = (\boldsymbol{x}, \boldsymbol{x}_s, \omega)$ for generic coordinates. The main idea of training the NN is to find the appropriate weights and biases to minimize the errors between the observations and predictions, where the sum of errors defines the loss function. For a specific frequency ω, we use the mean squared error (MSE) with an L2 norm to define the loss function only including a difference between predictions and observations:

$$\mathcal{L}_d(\theta_\tau) = \sum_{j}^{N_{src}} \sum_{i}^{N_{rcv}} \left| N_\tau(\theta_\tau; \boldsymbol{x}_{r;i}, \boldsymbol{x}_{s;j}, \omega) - \tau_{i,j} \right|^2, \tag{3.13}$$

where $\mathcal{L}_d$ is the loss function of the data constraint, i and j are the index of receivers and sources, respectively, N_{src} is the number of sources, N_{rcv} is the number of receivers, $\boldsymbol{x}_{r;i} = (x_r, y_r)_i$ and $\boldsymbol{x}_{s;j} = (x_s, y_s)_j$ are the receiver and source locations, $\tau_{i,j}$ represents the observed traveltime data.

In classical NN architecture, the phase traveltime surfaces of seismic waves are denoted as in Equation 3.13. This data-constrained loss can learn features based on observed data but has no knowledge of the underlying physical principles. Unlike this purely data dependent training, PINNs enable NNs to naturally meet the physical laws by modifying the loss function and modulating the training phase. In this study, the PINN framework is realized in SciAnn, a Keras and Tensorflow wrapper designed with physics-informed deep learning (Haghighat and Juanes 2021). Our model incorporates Equation 3.2 to govern the Rayleigh wave propagation, and the physics-constrained loss function at a single frequency can be written as:

$$L_e(N_\tau, N_{\boldsymbol{M}}) = N_{M11}\partial_x^2 N_\tau + 2N_{M12}\partial_x\partial_y N_\tau + N_{M22}\partial_y^2 N_\tau - 1, \tag{3.14}$$

where L_e is the loss function of the elliptical-anisotropic eikonal constraint, $N_{\boldsymbol{M}} = N_{\boldsymbol{M}}(\theta_{\boldsymbol{M}}; \boldsymbol{x}_r, \omega) = [N_{M11}(\theta_{M11}; \boldsymbol{x}_r, \omega), N_{M12}(\theta_{M12}; \boldsymbol{x}_r, \omega), N_{M22}(\theta_{M22}; \boldsymbol{x}_r, \omega)]$ specify the NNs of medium property matrices $\boldsymbol{M}$ in Equation 3.2, $N_\tau = N_\tau(\theta_\tau; \boldsymbol{x}_r, \boldsymbol{x}_s, \omega)$ is the traveltime at $\boldsymbol{x}_r$ from the source loca-

tion $\boldsymbol{x}_s$. θ_τ and $\theta_{\boldsymbol{M}}$ indicate parameters of traveltime NN, N_τ and medium property NNs, $N_{\boldsymbol{M}}$, respectively. Please note that here we use three NNs to represent the medium property matrices described by Equations 3.3, 3.4, and 3.5, but they share the same NN parameters. To enhance the robustness and efficiency of our methodology, we also explore the use of a single medium property NN for representing the three medium property matrices (M_{11}, M_{12}, M_{22}). Our findings reveal that although a single large network led to a reduction in training time, it also introduced unexpected artefacts, particularly noticeable at the corners of velocity maps (Figure B.3 in Supplementary Material). These artefacts are likely due to the lower sampling at the edge of our study areas. Detailed results and further discussion can be found in the Supplementary Material.

Equation 3.13 and 3.14 demonstrate single frequency loss functions. In practice, surface waves at different frequencies have depth-dependent sensitivity expressed as sensitivity kernels. So multi-frequency solutions are required for Rayleigh wave tomography to fully resolve the medium. Here we extend the NNs of the traveltime and medium property matrix to contain a period term as input so that the results of different periods can be obtained by training the network only once. For the case of simultaneous optimisation of anisotropic velocities at k frequencies, the multi-frequency total loss function that consists of data and physics constraints is expressed as:

$$\begin{aligned}\mathcal{L}(\theta_\tau, \theta_{\boldsymbol{M}}) = \sum_{k}^{N_{freq}} \sum_{j}^{N_{src}} \sum_{i}^{N_{rcv}} \Big[\Big| N_\tau(\theta_\tau; \boldsymbol{x}_{r;i}, \boldsymbol{x}_{s;j}, \omega_k) - \tau_{i,j,k} \Big|^2 \\ + \epsilon_e \Big| L_e(N_\tau(\theta_\tau; \boldsymbol{x}_{r;i}, \boldsymbol{x}_{s;j}, \omega_k), N_{\boldsymbol{M}}(\theta_{\boldsymbol{M}}; \boldsymbol{x}_{r;i}, \omega_k)) \Big|^2 \Big],\end{aligned} \tag{3.15}$$

where k and N_{freq} are the frequency index and maximum number of frequencies, ω_k represents the selected frequency k, $\tau_{i,j,k}$ is the phase traveltime surface at frequency k. ϵ_e is a weight factor set to normalize the different terms of data and physics constraints. The value of ϵ_e has been determined through synthetic tests and can be chosen within a narrow range: large enough to allow velocity model updates but small enough to enforce the physics constraints. The collocation points to evaluate the physics constraints coincide with the available data locations, although other choices are possible. In this network, the inputs are the spatial coordinates of observed source and receiver locations and given period values. For displaying

purposes, the trained NNs are evaluated on a chosen set of points forming a grid over a chosen extent. The network must be designed to have sufficient capability to represent the diversity and complexity in a set of phase traveltime surfaces over anisotropic velocity structures. By including physics constraints as loss terms, pinnEAET is fundamentally a joint state and parameter space approach.

The objective of the elliptical-anisotropy eikonal tomography is to estimate velocity structure and azimuthal anisotropy. The fast and slow velocity and the azimuth information can be directly evaluated from the trained networks $N_{\boldsymbol{M}}$. Meanwhile, the reconstructed phase traveltime surfaces $\hat{\tau}$ can be obtained by training the networks N_τ. With the loss term of elliptical-anisotropic eikonal equation, these trained phase traveltime surfaces can capture the constrained velocity information by calculating their gradients. The updated network parameters are generated by minimizing the MSE loss function:

$$\arg\min_{\theta_\tau, \theta_{\boldsymbol{M}}} \{\mathcal{L}(\theta_\tau, \theta_{\boldsymbol{M}})\} \rightarrow \hat{\theta}_\tau, \hat{\theta}_{\boldsymbol{M}}. \tag{3.16}$$

where $\hat{\theta}_\tau$ specifies the parameters of the trained network N_τ while $\hat{\theta}_{\boldsymbol{M}}$ specifies the parameters of the trained network $N_{\boldsymbol{M}}$.

Once the networks are trained based on the observed data points, we can directly evaluate the phase traveltime surfaces and three medium property matrices (Equation 3.3, 3.4 and 3.5) at regular sampling grid points. The final Rayleigh wave phase velocity solution can be extracted in two ways based on the trained networks. One way is to solve the eigenvalue and eigenvector problem of the inverted medium properties. The inverted Rayleigh wave phase velocity is shown in Equation 3.9. The orientation of the anisotropic fast axes and the magnitude of anisotropy can be computed as Equation 3.8 and 3.10. Employing this method we can directly obtain the isotropic phase velocity and anisotropic amplitude and directions. An alternative way to extract the velocity information from the trained networks is to calculate the gradient of the trained phase traveltime surfaces:

$$\bar{c}(\boldsymbol{x}, \omega) = \frac{1}{N_{src}} \sum_{j}^{N_{src}} \frac{1}{\left| \nabla N_\tau(\hat{\theta}_\tau; \boldsymbol{x}, \boldsymbol{x}_{s;j}, \omega_k) \right|}, \tag{3.17}$$

where $\bar{c}$ is the average velocity surface from all predicted phase traveltime surfaces, N_τ indicates

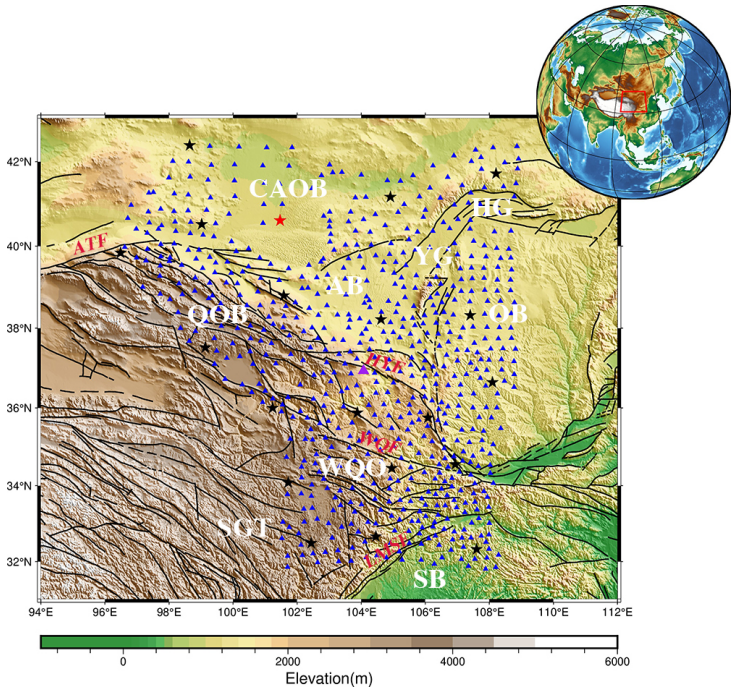


Figure 3.2: Dense seismic array geometry deployed on the northeastern Tibetan Plateau. Black lines indicate the main faults, blue triangles indicate stations from ChinArray-Himalaya II and stars indicate the selected source stations used in this study. The main faults and tectonic blocks are labeled as: ATF = Altyn-Tagh Fault; HYF = Haiyuan Fault; LMSF = Longmenshan Fault; WQF = West-Qinling Fault; AB = Alxa Block; CAOB = Central Asian Orogenic Belt; HTB = Hetao Graben; OB = Ordos Block; QOB = Qilian Orogenic Belt; SB = Sichuan Basin; SGT = Songpan-Ganzi Terrane; WQO = Western Qinling Orogen; YCG = Yinchuan Graben.

the trained phase traveltime surfaces $\hat{\tau}$. In order to reduce or even eliminate the influence of singular values near the sources, we take a trimmed geometric mean to exclude the 10 percentile outliers of the traveltime gradient datasets (Chen et al. 2022). We also calculate the associated standard deviation between the phase velocities from individual virtual sources and the average phase velocity $\bar{c}$ at each frequency and spatial location. This process results in phase velocity standard deviations at each location, providing estimates of standard uncertainties in our model. Both Equation 3.9 and 3.17 can illustrate the isotropic phase velocity of Rayleigh waves and their results are quite similar. Given that gradient calculations may introduce errors, here we choose Equation 3.9, 3.8 and 3.10 as isotropic and anisotropic components. The trained phase traveltime surfaces $\hat{\tau}$ and average phase velocities in Equation 3.17 are used to measure the uncertainties.

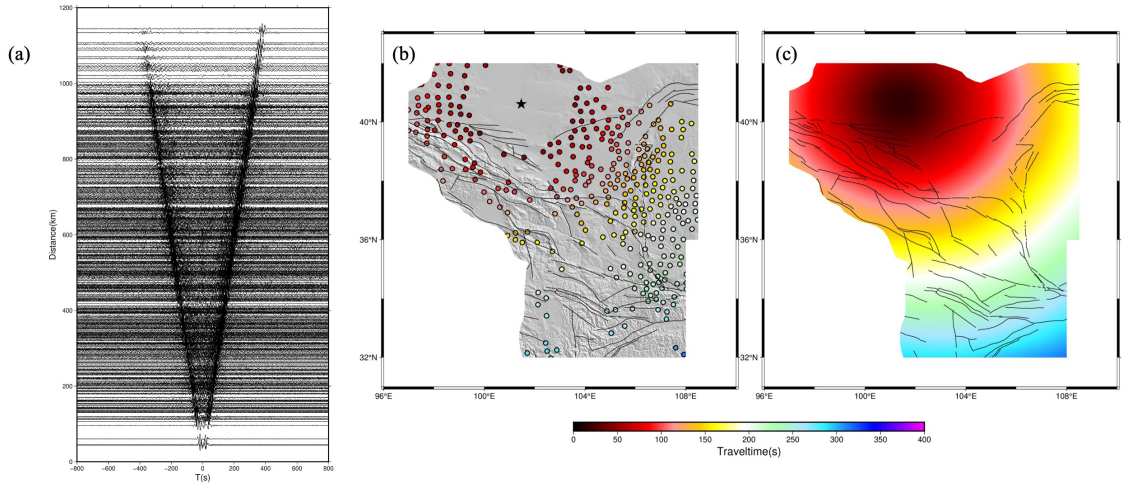


Figure 3.3: (a) The stacked waveform of Z-Z component cross-correlations for station pairs at station WT045 (red star in Figure 3.2) for a period range of 10-40 s, the V-shaped arrivals indicate the signals of Rayleigh wave; (b) The 25 s Rayleigh wave phase traveltime measured from cross-correlation centred on station WT045 shown as a star; (c) The 25 s Rayleigh wave phase traveltime surfaces predicted at station WT045 using pinnEAET.

3.3 Results

3.3.1 Training data

In this section, we apply our algorithm to ambient noise cross-correlation multi-frequency Rayleigh wave phase traveltime data recorded by the dense seismic array ChinArray-Himalaya II (Figure 3.2). Since 50 Ma, the continuous collision of India and Eurasia continents has led to high topography and crustal thickening within the Tibetan Plateau interior that primarily affects the tectonics in Asia (Molnar and Tapponnier 1975; Yin and Harrison 2000). The northeastern Tibetan Plateau plays a key role in studying the geodynamic mechanism of the crustal thickening and deformation of the Tibetan Plateau. Several hypotheses for the topographic elevation have been proposed (e.g. crustal flow (Royden et al. 1997; Clark and Royden 2000); oblique subduction of lithospheric mantle and strike-slip extrusion (Tapponnier et al. 2001); thin viscous sheet model (England and Houseman 1986)) to explain the mechanism of continental deformation, but the precise source of deformation remains enigmatic and subject to debate.

For the ChinArray-Himalaya II array, a total of 676 stations were deployed between 2013 and 2016 for about 2.5 years of continuous recordings with a station spacing of 40 to 70 km. Due to the good coverage of the northeastern Tibetan Plateau, these data enable us to test the capability of the proposed approach. In addition, the ChinArray-Himalaya II dataset has also been used to study the Rayleigh wave phase velocity and anisotropy using various methods,

for example, two-station Rayleigh wave tomography (Li et al. 2017), joint receiver functions and Rayleigh wave tomography (Wang et al. 2017), beamforming Rayleigh wave tomography (Wang et al. 2020) and Rayleigh wave eikonal tomography (Hao et al. 2021). These results enable comparison of the performance of the pinnEAET approach.

The training phase traveltime data are generated following the seismic ambient noise data processing steps as in Bensen et al. (2007) and Lin et al. (2008): the Z component of raw data are processed by resampling, removing the instrument response, filtering and applying temporal and spectral normalization. Traveltime calculation involves cross-correlating signals across station pairs and measuring the phase delay (or traveltime) of the resulting waveform. Figure 3.3 (a) shows an example of Z-Z component cross-correlations for station pairs associated with station WT045, filtered in the period range 10-40 s. Under ideal conditions, ambient noise cross-correlations should yield empirical Green's functions in both causal and anti-causal time windows (Lobkis and Weaver 2001; Wapenaar 2004). In Figure 3.3 (a), the energy of the fundamental Rayleigh wave can be clearly captured for both causal and anti-causal time windows. Figure 3.3 (b) shows the measured Rayleigh wave phase traveltime data centred on station WT045 at 25 s period. After picking phase traveltimes, 20 source stations are quasi-randomly selected among all 676 stations to achieve good coverage of the geographic area (Figure B.1 in Supplementary Material). The training targets are determined by the selected 20 phase traveltime surfaces. Figure 3.3 (c) shows an example of the 25 s Rayleigh wave phase traveltime surface evaluated from the trained N_τ at the effective source station WT045. It is worth emphasizing that receivers located within two wavelengths radius of each source are removed, because traveltime measurements collected at distances shorter than 1-2 wavelengths are unreliable (Lin et al. 2009).

3.3.2 Resolution tests

Before applying our approach to field data, we perform checkerboard tests to assess the performance of the pinnEAET and tune the NN design. As shown in Figure 3.4 (a)-(d), the isotropic synthetic checkerboard model assumes a constant background velocity defined by subtracting the observed frequency-dependent traveltime from the inter-station distances. Then $\pm 2\%$ of alternating velocity perturbations are added to the initial model to build the synthetic phase velocity model for checkerboard tests. There are 8×8 anomalies at each period with a maximum radius of 75 km for each anomaly. These alternating low and high velocity patterns are

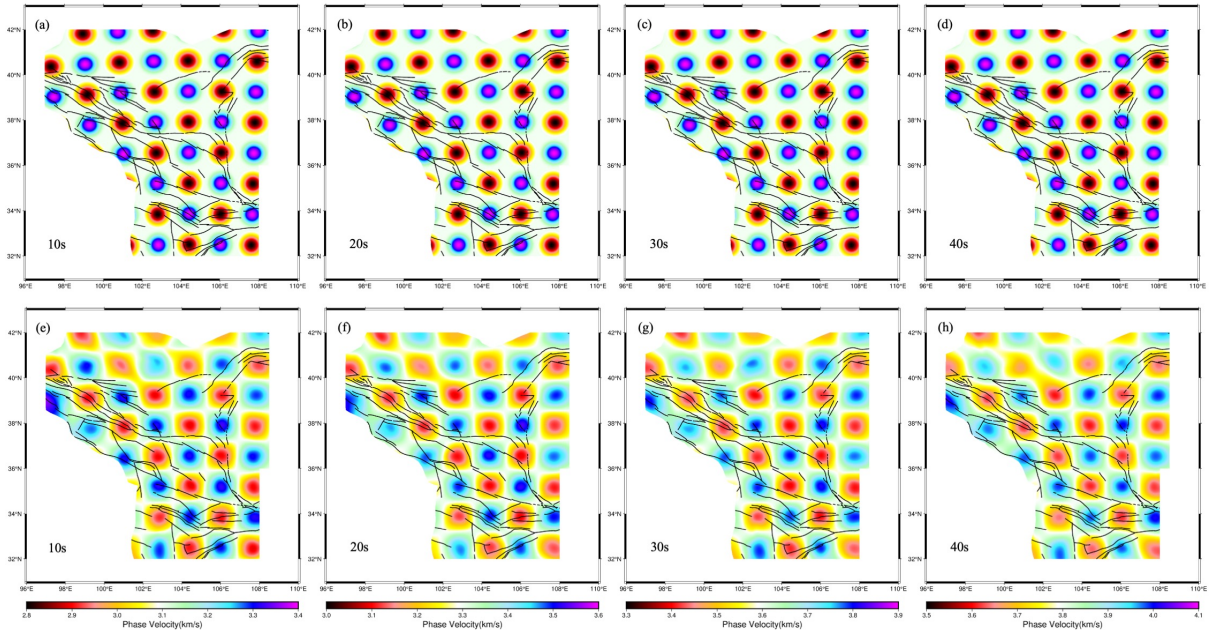


Figure 3.4: (a)-(d) Input synthetic Rayleigh wave phase velocity model of checkerboard resolution tests at periods of 10, 20, 30 and 40 s; (e)-(h) Corresponding retrieved velocity slices using pinnEAET.

distributed at 150 km intervals in latitude and longitude direction, respectively. The network parameters used in the checkerboard tests are consistent with the parameters of the subsequent field data training. The synthetic traveltimes of the checkerboard model are calculated using the fast marching method based on the eikonal equation (Sethian 1999; Treister and Haber 2016), and 0.1 s Gaussian random noise with a standard deviation of 0.01 is added to the synthetic traveltime data to simulate the noise level in observed data.

Figure 3.4 (e)-(h) shows the cross-section view of the retrieved Rayleigh wave phase velocity maps at 10, 20, 30 and 40 s for the resolution tests. It can be seen that the inversion recovers the pattern of velocity variation well. The resolution of less sampled structures at the periphery of the station network is still satisfactory. The northern part of the network exhibits lateral blurring in the SW-NE direction and the edges of the velocity anomalies appear indistinct. These may be linked to the selection of the source distribution.

3.3.3 Phase velocities and azimuth anisotropy beneath northeastern Tibetan Plateau

As discussed in the Methods section, the pinnEAET algorithm consists of two different structures of networks (see Figure 3.1) with a total of four fully connected feed-forward NNs: one

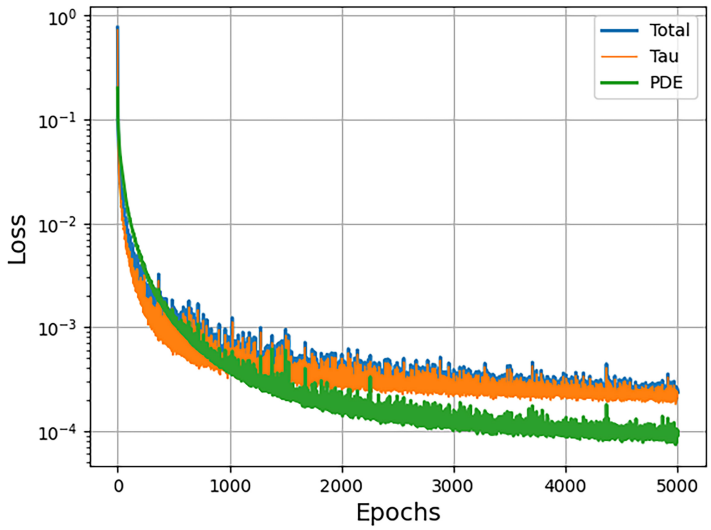


Figure 3.5: History of convergence of the total loss function (Total), the data constraint (Tau) and physics constraint (PDE) in Equation 3.15 for pinnEAET in the northeastern Tibet

traveltime NN N_τ and three medium property NNs $N_{\boldsymbol{M}}$. In addition to a parameter of target period k, N_τ training requires spatial locations and corresponding observed traveltimes, whereas $N_{\boldsymbol{M}}$ needs only spatial coordinates. Both networks are set to have 20 hidden layers. There are 80 neurons in each layer for N_τ while we use 60 neurons for all $N_{\boldsymbol{M}}$. The number of neurons and hidden layers in the NNs were chosen based on the checkerboard tests and adjusted to accurately represent all phase traveltime surfaces, without compromising the uncertainty in the final anisotropic velocity and azimuthal anisotropy. The network was trained using an Adam optimizer over 5000 epochs (Kingma and Ba 2014), with a learning rate of 0.001. These training parameters were chosen based on systematic synthetic tests for accurately and efficiently extracting maximum details from field data. Figure 3.5 shows the convergence processes of three loss functions: data constraints, physics constraints as well as total loss. These terms converge quickly and uniformly to below 10^{-3}, indicating that the model has a fast learning speed and the training process exhibits strong stability.

The results of analyzing Rayleigh wave phase velocity and azimuthal anisotropy over periods of 10, 20, 30 and 40 s in northeastern Tibet are shown in Figure 3.6. Figure 3.7 shows the corresponding uncertainties estimated from the traveltime NNs (Equation 3.17). The uncertainties for all periods are below 50 m/s and below 20 m/s in most areas. Compared to the central study area, the uncertainties tend to be higher along the boundaries. The higher uncertainty in the Hetao Graben at 10 s is likely due to the lower signal-to-noise ratio at shorter periods in that region. Although the uncertainties increase slightly with period, the overall reliability

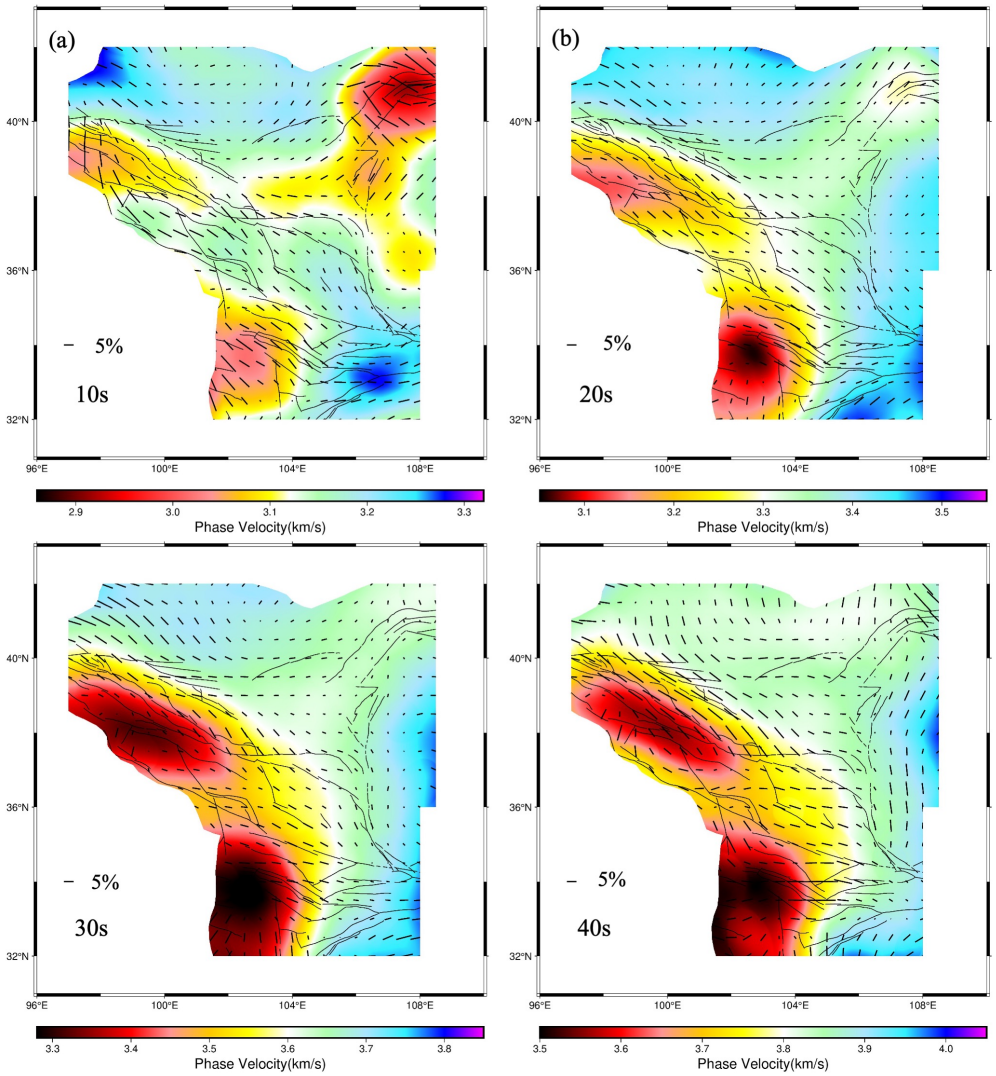


Figure 3.6: (a)-(d) Azimuthal anisotropic phase velocity at periods of 10, 20, 30 and 40 s beneath northeastern Tibetan Plateau using pinnEAET. Black bold vectors indicate the strength and fast propagation direction of anisotropy, black thin lines indicate main faults.

remains high.

For azimuthally dependent phase velocity, as a whole the distribution of low or high velocity zones and anisotropic fast directions seem significantly controlled by the fault strikes and block boundaries. Prominent low velocity zones (LVZs) exist at almost all periods along the southwestern boundary of the study area, corresponding to the Qilian Orogenic Belt and the Songpan-Ganzi Terrane, and relatively high velocities appear in the northern and southeastern regions, corresponding to the Central Asian Orogenic Belt, Western Qinling Orogen and Sichuan Basin. The fast directions of azimuthal anisotropy changes slowly at each period.

The 10 s azimuthal anisotropic phase velocity is sensitive to the structure of the shallow sed-

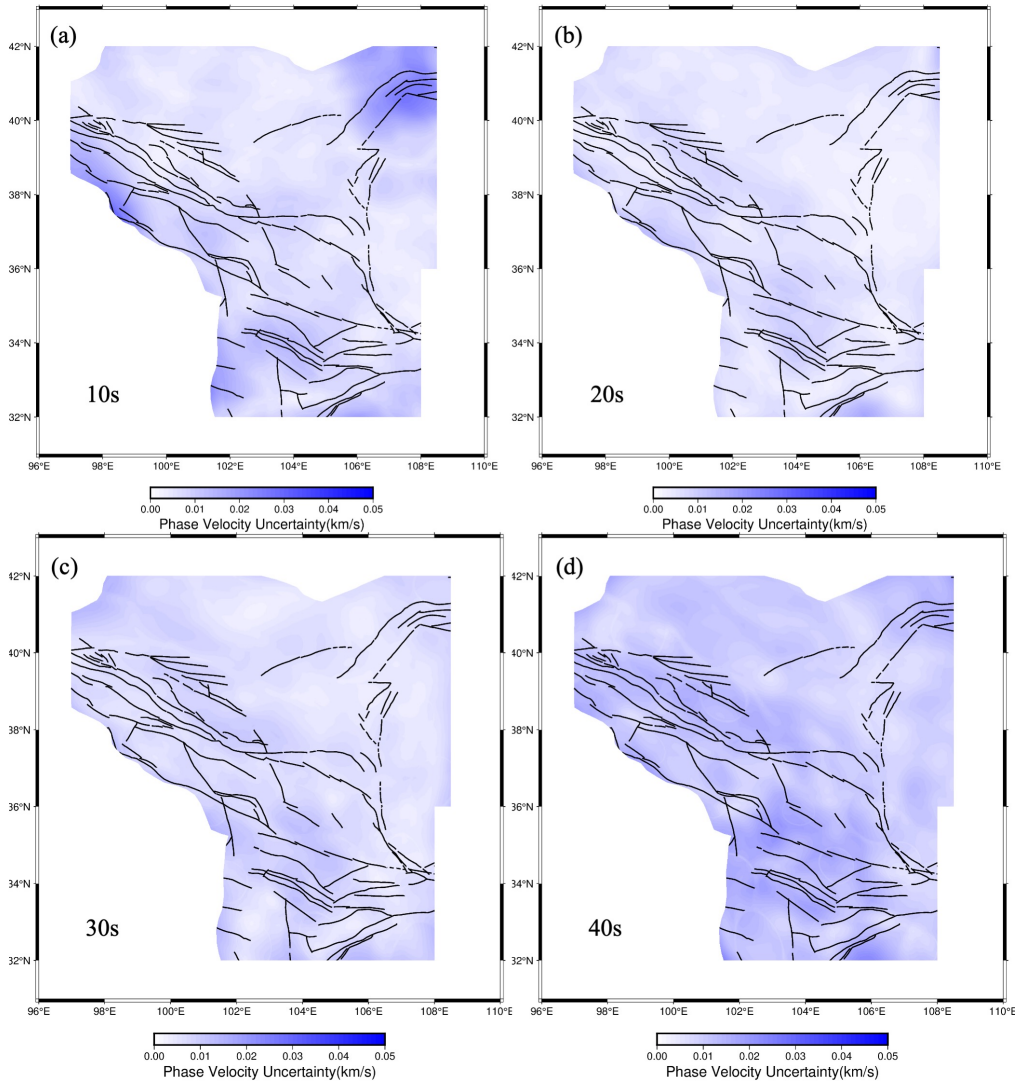


Figure 3.7: (a)-(d) Uncertainty map of the Rayleigh wave phase velocity at periods of 10, 20, 30 and 40 s beneath northeastern Tibetan Plateau using pinnEAET.

imentary layer and the upper crust. Unlike other periods, it is primarily distinguished by the presence of LVZs in the northeast and relatively high velocity zone (HVZ) in the southeast of the study area. These LVZs are closely associated with the sedimentary basins in Hetao Graben, Yinchuan Graben, Alxa Block and Ordos Block, while the relatively HVZ may be linked to the exposure of metamorphic rocks and granites in the east of the Western Qinling orogen and the northern Sichuan Basin (Wang et al. 2020). The anisotropic fast direction is related to fault strikes and block boundaries in most study areas. The fast direction in the southwest regions of Qilian Orogenic Belt, Western Qinling Orogen and Songpan-Ganzi Terrane is mainly in NE-SW direction. The fast directions at 10 s beneath the Songpan-Ganzi Terrane are not consistent with longer periods, supporting the idea that crustal flows are predominantly found in the middle and lower crust (Gao et al. 2020; Li et al. 2022).

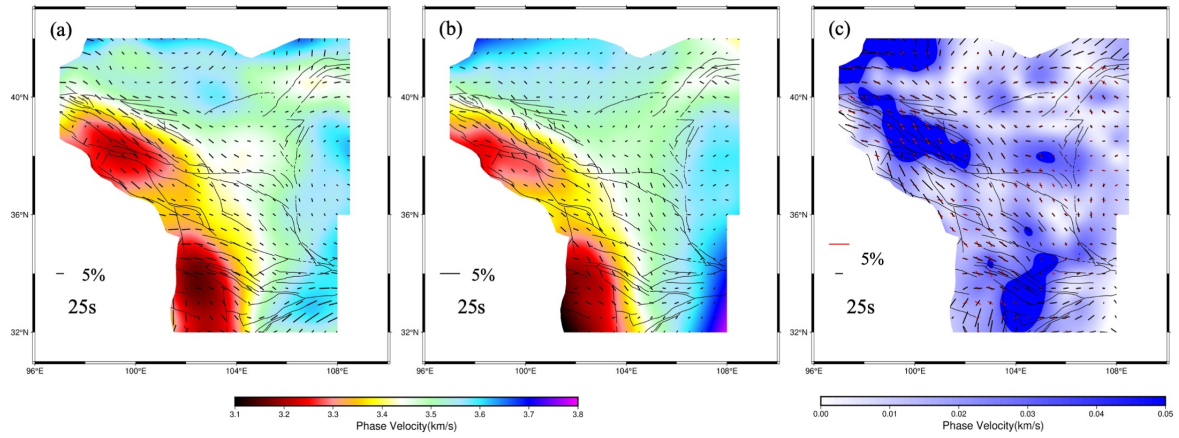


Figure 3.8: (a) The 25 s azimuthal anisotropic phase velocity beneath northeastern Tibetan Plateau using pinnEAET; (b) The 25 s azimuthal anisotropic phase velocity generated by conventional eikonal tomography; (c) The difference between 25 s azimuthal anisotropic phase velocity (background) from pinnEAET (black lines) and eikonal tomography (red lines).

The 20 s azimuthal anisotropic phase velocity mainly represents structures of the mid-crust. Relative to the velocity distribution at 10 s, the low velocity anomalies in Yinchuan Graben, Alxa block and Ordos Block disappear, but there is still a smaller LVZ in the Hetao Graben. The northwestern (Central Asian Orogenic Belt) and southeastern parts (Sichuan Basin) are characterized by relatively high velocity anomalies without clear boundaries. In the southwest corner of the study area the fast directions resolved by periods of 10 s and 20 s are similar. It shows a tendency of near clockwise rotation along the margin of the northeastern Tibetan Plateau. The anisotropic strength is relatively weak in the Ordos Block.

The 30-40 s azimuthal anisotropic phase velocity illustrates the lower crustal and part of upper mantle structures. The main features in these periods are significant LVZs in the southwest and high velocity in the northeast. As the period increases, the velocity difference between the Tibetan Plateau and the surrounding areas progressively becomes more pronounced. The anisotropic strength increases noticeably in areas such as the Central Asian Orogenic Belt and Qilian Orogenic belt, which is possibly related to the inhomogeneity of the Moho (Zhang et al. 2020).

3.4 Discussion

Figure 3.8 shows a comparison of the azimuthal anisotropic velocities obtained by pinnEAET and traditional anisotropic eikonal tomography. We find very good agreement in the Rayleigh

wave phase velocity structures in most regions with differences of under 50 m/s. Two major differences between the results exist in the northwest corner of the study area as well as part of the Qilian Orogenic Belt along the margin of the northeastern Tibetan Plateau where the result of conventional eikonal tomography shows significant high velocity zones that are not visible in other results. These could be spurious anomalies caused by less path coverage at the boundary. A similar discrepancy can be found in a small region in the centre (105°E, 37°N) of the study area. To facilitate the analysis, we also compare the results of azimuthal double beamforming tomography (Wu et al. 2023) (Figure B.2 in Supplementary Material) and no differences of more than 50 m/s were found in these areas except for a very small area at the edge of the northeast corner, which also confirms the reliability of our method.

In the case of anisotropy, the pinnEAET directly calculates an equivalent amplitude of anisotropy using Equation 3.10 whereas traditional eikonal tomography fits a function of phase velocity variations over azimuth. Due to the different strategies, we only compare the azimuth of the fast propagation direction of those results. Both of them exhibit a trend of clockwise rotation in the fast direction along the margin of northeastern Tibetan Plateau. They show similar distribution of azimuthal anisotropy inside the study area, except that our anisotropic results illustrate the NWW-SEE-dominated fast direction in the Western Qinling Orogen rather than the NNW-SSE-dominated fast direction obtained by the conventional eikonal tomography. It is worth pointing out that our fast direction results have a better correspondence with the strikes of the West-Qinling Fault. In the centre of the study area, south of the Alxa block, the fast directions of anisotropy of the two are not consistent, but our results are in good agreement with those of azimuthal double beamforming tomography (Figure B.2 (c) in Supplementary Material). The anisotropy results for both differ considerably at the boundaries of the study area, especially in the northeast corner (Hetao Graben). Combined with supplemental Figure B.2 (c), the anisotropy results of all three methods differ in this region, which may be due to the fact that the path distribution on the boundary is relatively poor from some directions.

We also compare our anisotropic phase velocities with past studies in the region to assess the validity of the pinnEAET to the alternative approaches employed, such as teleseismic data based surface wave tomography (Li et al. 2017; Hao et al. 2021) and ambient noise beamforming tomography (Wang et al. 2020). These phase velocity results corroborate the velocity distribution of the southwest low velocities and northeast high velocities in the study area. The anisotropy

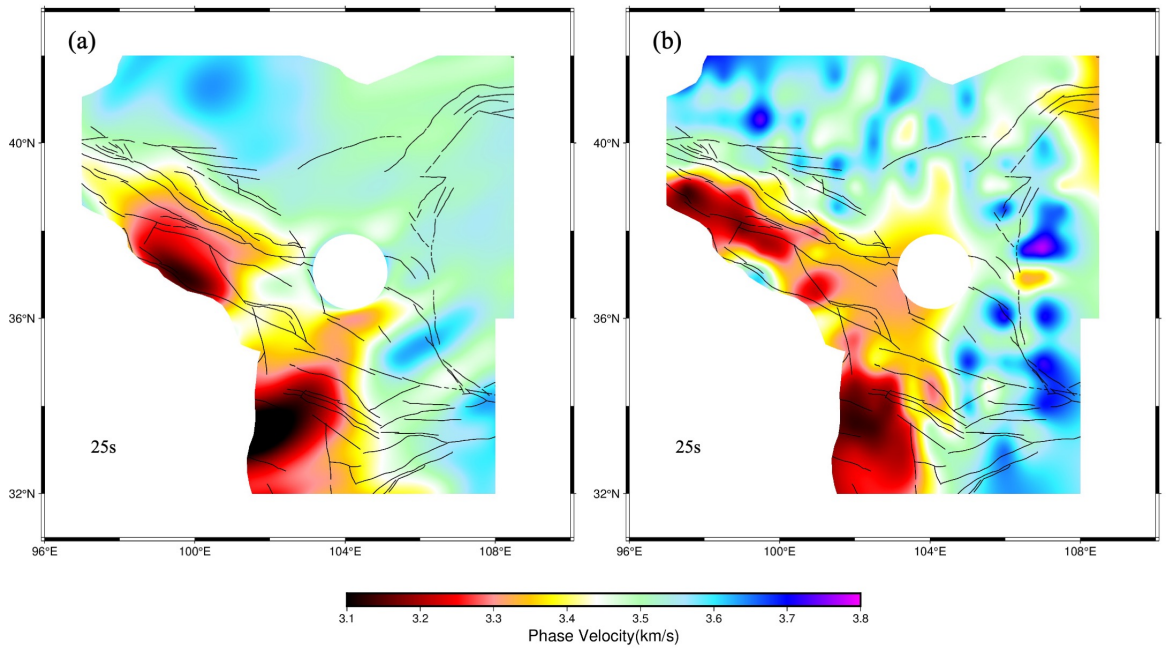


Figure 3.9: (a) The 25 s Rayleigh wave phase velocity beneath northeastern Tibetan Plateau using pinnEAET with only one source station (purple triangle in Figure 3.2); (b) The 25 s phase velocity generated by conventional eikonal tomography with same datasets.

results are also overall similar compared to those of Hao et al. (2021), only differing in small regions at certain periods. For example, for the 25 s period, our result shows a NW-SE fast direction in the northeast of Songpan-Ganzi Terrane (104°E, 33°N), while their result shows a near N-S direction. This is probably caused by the difference in the generation mechanism and information content between ambient noise data and teleseismic surface wave data. Combining the above comparisons, our results are consistent with those of other methods, confirming the viability of our pinnEAET approach.

To verify the accuracy and reliability of the pinnEAET for small datasets, we choose only one station as the virtual source to invert for the Rayleigh wave phase velocity at 25 s. The NN hyperparameters are kept the same as those used in the field data training with multiple sources. Traditional ambient noise eikonal tomography is applied to the same data, utilizing the same source for direct comparison of phase velocity models and evaluation of methodology resolution. The velocity results in the vicinity of the source are unreliable due to the ambiguity between overlapping causal and anti-causal empirical Green's functions close to the source. Thus the data points with a region of 1.6° around the source location are removed from both results.

Figure 3.9(a) illustrates the phase velocity result for a single source recovered by pinnEAET

while Figure 3.9(b) demonstrates that retrieved by conventional eikonal tomography. Both results show the main low velocity zones along the Qilian Orogenic Belt, Western Qinling Orogen and Songpan-Ganzi Terrane. The result from pinnEAET is closer to that of others utilizing all stations. It well recovers the majority of velocity structures except for some areas where path coverage is extremely limited. The result using conventional eikonal tomography displays many spurious velocity anomaly patterns and obvious smearing phenomena. It is clear that the results of pinnEAET show a more accurate solution compared to these earlier studies. The performance test gives support to the ability of our approach especially when the available data are limited.

3.5 Conclusion

We have developed a deep learning method for anisotropic ambient noise tomography based on physics-informed neural networks. By integrating physics constraints into the architecture, these models can effectively capture the underlying physics of the elliptical-anisotropic eikonal tomography problem. This method is designed to achieve high accuracy and robust results even with limited data and noisy input. We apply this approach to the Rayleigh wave correlation traveltime data recorded by a dense seismic array deployed on the northeastern Tibetan Plateau. The algorithm enables the simultaneous training of multiple periods of traveltime by setting the period as training data. The results can be achieved with comparable resolution as conventional methods using only 20 source stations, minimizing necessary data and memory requirements. We also show that our method still performs well when using extremely limited data and resolves structures that are more consistent with established results than those obtained from traditional eikonal tomography. Anisotropy fast directions can be well resolved and are in agreement with existing fault zone structure. Overall, we show that elliptical-anisotropy eikonal tomography based on physics-informed neural networks is a competitive alternative to traditional methods to extract anisotropic velocity information from Rayleigh wave data.

Open Research Section

The traveltime datasets from ambient noise cross-correlations used in this study and the Rayleigh wave azimuthal anisotropic phase velocity models can be downloaded at `https://doi.org/10.5281/zenodo.8088610`.

Acknowledgments

We thank all people who participated in field deployment of the ChinArray-Himalaya II. We thank the China Seismic Array Data Management Center at the Institute of Geophysics, China Earthquake Administration for providing the ambient noise data used in this study. We are very grateful to the anonymous reviewers whose insightful comments and constructive suggestions greatly contributed to the improvement of this manuscript. This work is supported by the National Science Foundation of China (Grants 41890814 and U1901602). YP.C. is also supported by a split-site PhD program of SUSTech and the University of Leeds.

References

Aki, Keiiti and Paul G Richards (2002). *Quantitative seismology*.

Alkhalifah, Tariq, Chao Song, U Bin Waheed, and Qi Hao (2020). "Wavefield solutions from machine learned functions that approximately satisfy the wave equation". In: *EAGE 2020 annual conference & exhibition online*. Vol. 2020. 1. European Association of Geoscientists & Engineers, pp. 1–5.

Araya-Polo, Mauricio, Joseph Jennings, Amir Adler, and Taylor Dahlke (2018). "Deep-learning tomography". In: *The Leading Edge* 37.1, pp. 58–66.

Baydin, Atilim Gunes, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind (2018). "Automatic differentiation in machine learning: a survey". In: *Journal of Machine Learning Research* 18, pp. 1–43.

Becker, Th W, Sergei Lebedev, and MD Long (2012). "On the relationship between azimuthal anisotropy from shear wave splitting and surface wave tomography". In: *Journal of Geophysical Research: Solid Earth* 117.B1.

Bensen, GD, MH Ritzwoller, MP Barmin, A Lin Levshin, Feifan Lin, MP Moschetti, NM Shapiro, and Yanyan Yang (2007). "Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements". In: *Geophysical Journal International* 169.3, pp. 1239–1260.

Cao, Feihuang, Chuntao Liang, Lu Zhou, and Jieshou Zhu (2020). "Seismic Azimuthal anisotropy for the southeastern Tibetan plateau extracted by wave gradiometry analysis". In: *Journal of Geophysical Research: Solid Earth* 125.5, e2019JB018395.

Chen, Yunpeng, Sjoerd AL de Ridder, Sebastian Rost, Zhen Guo, Xiaoyang Wu, and Yongshun Chen (2022). "Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the northeastern margin of the Tibetan Plateau". In: *Geophysical Research Letters* 49.21, e2022GL099053.

Clark, Marin Kristen and Leigh Handy Royden (2000). "Topographic ooze: Building the eastern margin of Tibet by lower crustal flow". In: *Geology* 28.8, pp. 703–706.

Crampin, Stuart and David C Booth (1985). "Shear-wave polarizations near the North Anatolian Fault–II. Interpretation in terms of crack-induced anisotropy". In: *Geophysical Journal International* 83.1, pp. 75–92.

Creasy, Neala, Angelo Pisconti, Maureen D Long, Christine Thomas, and James Wookey (2019). "Constraining lowermost mantle anisotropy with body waves: a synthetic modelling study". In: *Geophysical Journal International* 217.2, pp. 766–783.

De Ridder, SAL, BL Biondi, and D Nichols (2015). "Elliptical-anisotropic eikonal phase velocity tomography". In: *Geophysical Research Letters* 42.3, pp. 758–764.

De Ridder, SAL and Andrew Curtis (2017). "Seismic gradiometry using ambient seismic noise in an anisotropic Earth". In: *Geophysical Journal International* 209.2, pp. 1168–1179.

Ekström, Göran (2011). "A global model of Love and Rayleigh surface wave dispersion and anisotropy, 25-250 s". In: *Geophysical Journal International* 187.3, pp. 1668–1686.

England, Philip and Gregory Houseman (1986). "Finite strain calculations of continental deformation: 2. Comparison with the India-Asia collision zone". In: *Journal of Geophysical Research: Solid Earth* 91.B3, pp. 3664–3676.

Fang, Zhiwei and Justin Zhan (2019). "Deep physical informed neural networks for metamaterial design". In: *IEEE Access* 8, pp. 24506–24513.

Gao, Yuan, Yutao Shi, and Qiong Wang (2020). "Seismic anisotropy in the southeastern margin of the Tibetan Plateau and its deep tectonic significances". In: *Chinese Journal of Geophysics* 63.3, pp. 802–816.

Gou, Rongxi, Yijie Zhang, Xueyu Zhu, and Jinghuai Gao (2023). "Bayesian physics-informed neural networks for the subsurface tomography based on the eikonal equation". In: *IEEE Transactions on Geoscience and Remote Sensing* 61, pp. 1–12.

Haghighat, Ehsan and Ruben Juanes (2021). "Sciann: A keras/tensorflow wrapper for scientific computations and physics-informed deep learning using artificial neural networks". In: *Computer Methods in Applied Mechanics and Engineering* 373, p. 113552.

Hao, Shijie, Zhouchuan Huang, Cunrui Han, Liangshu Wang, Mingjie Xu, Ning Mi, and Dayong Yu (2021). "Layered crustal azimuthal anisotropy beneath the northeastern Tibetan Plateau revealed by Rayleigh-wave Eikonal tomography". In: *Earth and Planetary Science Letters* 563, p. 116891.

Ho, Madison, Sidhant Idgunji, Jonathan L Payne, and Ardiansyah Koeshidayatullah (2023). "Hierarchical multi-label taxonomic classification of carbonate skeletal grains with deep learning". In: *Sedimentary Geology* 443, p. 106298.

Karimpouli, Sadegh and Pejman Tahmasebi (2020). "Physics informed machine learning: Seismic wave equation". In: *Geoscience Frontiers* 11.6, pp. 1993–2001.

Karpatne, Anuj, Gowtham Atluri, James H Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar (2017). "Theory-guided data science: A new paradigm for scientific discovery from data". In: *IEEE Transactions on knowledge and data engineering* 29.10, pp. 2318–2331.

Kästle, Emanuel David, Irene Molinari, Lapo Boschi, Edi Kissling, and AlpArray Working Group (2022). "Azimuthal anisotropy from eikonal tomography: example from ambient-noise measurements in the AlpArray network". In: *Geophysical Journal International* 229.1, pp. 151–170.

Kim, Yejin and Tae Sup Yun (2021). "How to classify sand types: A deep learning approach". In: *Engineering geology* 288, p. 106142.

Kingma, Diederik P and Jimmy Ba (2014). "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980*.

Kissas, Georgios, Yibo Yang, Eileen Hwuang, Walter R Witschey, John A Detre, and Paris Perdikaris (2020). "Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks". In: *Computer Methods in Applied Mechanics and Engineering* 358, p. 112623.

Li, Shilin, Zhen Guo, Yong Yu, Xiaoyang Wu, and Yongshun John Chen (2022). "Imaging the Northeastern crustal boundary of the Tibetan Plateau with radial anisotropy". In: *Geophysical Research Letters* 49.23, e2022GL100672.

Li, Sijin, Liyang Xiong, Guoan Tang, and Josef Strobl (2020). "Deep learning-based approach for landform classification from integrated data sources of digital elevation model and imagery". In: *Geomorphology* 354, p. 107045.

Li, Yonghua, Jiatie Pan, Qingju Wu, and Zhifeng Ding (2017). "Lithospheric structure beneath the northeastern Tibetan Plateau and the western Sino-Korea Craton revealed by Rayleigh wave tomography". In: *Geophysical Journal International* 210.2, pp. 570–584.

Li, Yonghua, Qingju Wu, Fengxue Zhang, Qiangqiang Feng, and Ruiqing Zhang (2011). "Seismic anisotropy of the Northeastern Tibetan Plateau from shear wave splitting analysis". In: *Earth and Planetary Science Letters* 304.1-2, pp. 147–157.

Lin, Fan-Chi, Morgan P Moschetti, and Michael H Ritzwoller (2008). "Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps". In: *Geophysical Journal International* 173.1, pp. 281–298.

Lin, Fan-Chi and Michael H Ritzwoller (2011). "Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure". In: *Geophysical Journal International* 186.3, pp. 1104–1120.

Lin, Fan-Chi, Michael H Ritzwoller, and Roel Snieder (2009). "Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array". In: *Geophysical Journal International* 177.3, pp. 1091–1110.

Lobkis, Oleg I and Richard L Weaver (2001). "On the emergence of the Green's function in the correlations of a diffuse field". In: *Journal of the Acoustical Society of America* 110.6, 3011–3017".

Long, Maureen D and Thorsten W Becker (2010). "Mantle dynamics and seismic anisotropy". In: *Earth and Planetary Science Letters* 297.3-4, pp. 341–354.

Luo, Zijing, Yihui Xiong, and Renguang Zuo (2020). "Recognition of geochemical anomalies using a deep variational autoencoder network". In: *Applied Geochemistry* 122, p. 104710.

Misyris, George S, Andreas Venzke, and Spyros Chatzivasileiadis (2020). "Physics-informed neural networks for power systems". In: *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, pp. 1–5.

Molnar, Peter and Paul Tapponnier (1975). "Cenozoic Tectonics of Asia: Effects of a Continental Collision: Features of recent continental tectonics in Asia can be interpreted as results of the India-Eurasia collision". In: *science* 189.4201, pp. 419–426.

Montagner, Jean-Paul and Laurent Guillot (2002). "Seismic anisotropy and global geodynamics". In: *Reviews in mineralogy and geochemistry* 51.1, pp. 353–385.

Montagner, JP (1986). "Regional three-dimensional structures using long-period surface waves". In: *Ann. Geophys* 4.B3, pp. 283–294.

Moseley, Ben, Andrew Markham, and Tarje Nissen-Meyer (2020a). "Solving the wave equation with physics-informed deep learning". In: *arXiv preprint arXiv:2006.11894*.

Moseley, Ben, Tarje Nissen-Meyer, and Andrew Markham (2020b). "Deep learning for fast simulation of seismic waves in complex media". In: *Solid Earth* 11.4, pp. 1527–1549.

Mousavi, S Mostafa, William L Ellsworth, Weiqiang Zhu, Lindsay Y Chuang, and Gregory C Beroza (2020). "Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking". In: *Nature Communications* 11.1, pp. 1–12.

Nicolas, Adolphe and Nikolas I Christensen (1987). "Formation of anisotropy in upper mantle peridotites-A review". In: *Composition, structure and dynamics of the lithosphere-asthenosphere system* 16, pp. 111–123.

Raissi, Maziar, Paris Perdikaris, and George E Karniadakis (2019). "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In: *Journal of Computational Physics* 378, pp. 686–707.

Raissi, Maziar, Alireza Yazdani, and George Em Karniadakis (2020). "Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations". In: *Science* 367.6481, pp. 1026–1030.

Rasht-Behesht, Majid, Christian Huber, Khemraj Shukla, and George Em Karniadakis (2022). "Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions". In: *Journal of Geophysical Research: Solid Earth* 127.5, e2021JB023120.

Ritzwoller, Michael H, Fan-Chi Lin, and Weisen Shen (2011). "Ambient noise tomography with a large seismic array". In: *Comptes Rendus Geoscience* 343.8-9, pp. 558–570.

Romanowicz, Barbara (2002). "Inversion of surface waves: a review". In: *International Geophysics Series* 81.A, pp. 149–174.

Ross, Zachary E, Men-Andrin Meier, and Egill Hauksson (2018). "P wave arrival picking and first-motion polarity determination with deep learning". In: *Journal of Geophysical Research: Solid Earth* 123.6, pp. 5120–5129.

Royden, Leigh H, B Clark Burchfiel, Robert W King, Erchie Wang, Zhiliang Chen, Feng Shen, and Yuping Liu (1997). "Surface deformation and lower crustal flow in eastern Tibet". In: *science* 276.5313, pp. 788–790.

Russell, Joshua B, James B Gaherty, Pei-Ying Patty Lin, Daniel Lizarralde, John A Collins, Greg Hirth, and Rob L Evans (2019). "High-resolution constraints on Pacific upper mantle petrofabric inferred from surface-wave anisotropy". In: *Journal of Geophysical Research: Solid Earth* 124.1, pp. 631–657.

Savage, MK (1999). "Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting?" In: *Reviews of Geophysics* 37.1, pp. 65–106.

Schulte-Pelkum, Vera and Kevin H Mahan (2014). "A method for mapping crustal deformation and anisotropy with receiver functions and first results from USArray". In: *Earth and Planetary Science Letters* 402, pp. 221–233.

Sethian, James A (1999). "Fast marching methods". In: *SIAM review* 41.2, pp. 199–235.

Shapiro, Nikolai M, Michel Campillo, Laurent Stehly, and Michael H Ritzwoller (2005). "High-resolution surface-wave tomography from ambient seismic noise". In: *Science* 307.5715, pp. 1615–1618.

Silver, Paul G (1996). "Seismic anisotropy beneath the continents: Probing the depths of geology". In: *Annual review of earth and planetary sciences* 24, pp. 385–432.

Simons, Frederik J, Rob D Van Der Hilst, Jean-Paul Montagner, and Alet Zielhuis (2002). "Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle". In: *Geophysical Journal International* 151.3, pp. 738–754.

Smith, Jonathan D, Kamyar Azizzadenesheli, and Zachary E Ross (2020). "Eikonet: Solving the eikonal equation with deep neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 59.12, pp. 10685–10696.

Smith, Martin L and FA Dahlen (1973). "The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium". In: *Journal of Geophysical Research* 78.17, pp. 3321–3333.

Soergel, Dorian, Helle A Pedersen, Thomas Bodin, Anne Paul, and Laurent Stehly (2023). "Bayesian analysis of azimuthal anisotropy in the Alpine lithosphere from beamforming of ambient noise cross-correlations". In: *Geophysical Journal International* 232.1, pp. 429–450.

Song, Chao, Tariq Alkhalifah, and UB Waheed (2022). "A versatile framework to solve the Helmholtz equation using physics-informed neural networks". In: *Geophysical Journal International* 228.3, pp. 1750–1762.

Song, Chao, Tariq Alkhalifah, and Umair Bin Waheed (2021). "Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks". In: *Geophysical Journal International* 225.2, pp. 846–859.

Song, Chao and Tariq A Alkhalifah (2021). "Wavefield reconstruction inversion via physics-informed neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–12.

Song, Chao and Yanghua Wang (2023). "Simulating seismic multifrequency wavefields with the Fourier feature physics-informed neural network". In: *Geophysical Journal International* 232.3, pp. 1503–1514.

Tapponnier, Paul, Xu Zhiqin, Francoise Roger, Bertrand Meyer, Nicolas Arnaud, Gérard Wittlinger, and Yang Jingsui (2001). "Oblique stepwise rise and growth of the Tibet Plateau". In: *science* 294.5547, pp. 1671–1677.

Taufik, Mohammad Hasyim, Umair bin Waheed, and Tariq A Alkhalifah (2022). "Upwind, no more: Flexible traveltime solutions using physics-informed neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–12.

Treister, Eran and Eldad Haber (2016). "A fast marching algorithm for the factored eikonal equation". In: *Journal of Computational physics* 324, pp. 210–225.

Waheed, Umair bin, Ehsan Haghighat, Tariq Alkhalifah, Chao Song, and Qi Hao (2021). "PINNeik: Eikonal solution using physics-informed neural networks". In: *Computers & Geosciences* 155, p. 104833.

Waheed, Umair bin, Tariq Alkhalifah, Ehsan Haghighat, Chao Song, and Jean Virieux (2021). "PINNtomo: Seismic tomography using physics-informed neural networks". In: *arXiv preprint arXiv:2104.01588*.

Waheed, Umair bin, Ehsan Haghighat, and Tariq Alkhalifah (2020). "Anisotropic eikonal solution using physics-informed neural networks". In: *SEG international exposition and annual meeting*. SEG, D041S100R003.

Wang, Kaiming, Laiyu Lu, Valérie Maupin, Zhifeng Ding, Chen Zheng, and Shijun Zhong (2020). "Surface wave tomography of northeastern Tibetan Plateau using beamforming of seismic noise at a dense array". In: *Journal of Geophysical Research: Solid Earth* 125.4, e2019JB018416.

Wang, Xingchen, Yonghua Li, Zhifeng Ding, Lupei Zhu, Chunyong Wang, Xuewei Bao, and Yan Wu (2017). "Three-dimensional lithospheric S wave velocity model of the NE Tibetan Plateau and western North China Craton". In: *Journal of Geophysical Research: Solid Earth* 122.8, pp. 6703–6720.

Wapenaar, Kees (2004). "Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation". In: *Physical review letters* 93.25, p. 254301.

Wapenaar, Kees, Deyan Draganov, Roel Snieder, Xander Campman, and Arie Verdel (2010). "Tutorial on seismic interferometry: Part 1—Basic principles and applications". In: *Geophysics* 75.5, 75A195–75A209.

Weisstein, Eric W (2014). *Ellipse. From MathWorld–A Wolfram Web Resource.* URL: `https://mathworld.wolfram.com/Ellipse.html`.

Wu, Xiaoyang, Zhen Guo, Shilin Li, Yong Yu, Qipeng Bai, and Yongshun John Chen (2023). "Seismic azimuthal anisotropy of northeastern Tibetan Plateau from ambient noise double beamforming tomography: Implications for crustal deformation". In: *Journal of Geophysical Research: Solid Earth* 128.6, e2022JB026109.

Yang, Fangshu and Jianwei Ma (2019). "Deep-learning inversion: A next-generation seismic velocity model building method". In: *Geophysics* 84.4, R583–R599.

Yao, Huajian, Robert D Van Der Hilst, and Jean-Paul Montagner (2010). "Heterogeneity and anisotropy of the lithosphere of SE Tibet from surface wave array tomography". In: *Journal of Geophysical Research: Solid Earth* 115.B12.

Yin, An and T Mark Harrison (2000). "Geologic evolution of the Himalayan-Tibetan orogen". In: *Annual review of earth and planetary sciences* 28.1, pp. 211–280.

Yu, Siwei and Jianwei Ma (2021). "Deep learning for geophysics: Current and future trends". In: *Reviews of Geophysics* 59.3, e2021RG000742.

Zhang, Chen, Zhen Guo, and Yongshun John Chen (2020). "Lithospheric thickening controls the ongoing growth of northeastern Tibetan Plateau: Evidence from P and S receiver functions". In: *Geophysical Research Letters* 47.15, e2020GL088972.

Zhao, Dapeng, Sheng Yu, and Xin Liu (2016). "Seismic anisotropy tomography: New insight into subduction dynamics". In: *Gondwana Research* 33, pp. 24–43.

Zheng, Tuo, Stephen S Gao, Zhifeng Ding, Kelly H Liu, Lijun Chang, Xiaoping Fan, Fansheng Kong, and Youqiang Yu (2021). "Crustal azimuthal anisotropy and deformation beneath the northeastern Tibetan Plateau and adjacent areas: Insights from receiver function analysis". In: *Tectonophysics* 816, p. 229014.

Zhu, Weiqiang, S Mostafa Mousavi, and Gregory C Beroza (2019). "Seismic signal denoising and decomposition using deep neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 57.11, pp. 9476–9488.

Chapter 4

Teleseismic Rayleigh wave tomography in northeastern Tibet using physics-informed neural networks

Yunpeng Chen[1,2], Sjoerd A.L. de Ridder[1], Sebastian Rost[1], Zhen Guo[2]and Yongshun Chen[2]

[1] *School of Earth and Environment, University of Leeds, LS29JT, Leeds, UK*

[2] *Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

Abstract

In this study, we present a novel approach to teleseismic Rayleigh wave tomography using physics-informed neural networks (PINNs), termed pinnTET, to investigate the crustal and upper mantle structure beneath the northeastern Tibetan Plateau. The method employs two neural networks, one for modeling phase velocities and another for traveltime surfaces, while incorporating physical constraints from the eikonal equation. We apply this approach to data from the dense seismic array deployed, analyzing Rayleigh wave phase velocities at periods between 20-80 s. Our results reveal significant lateral heterogeneity in velocity structure that correlates well with major tectonic features. The method achieves reliable resolution using only approximately 20% of the observational data compared to traditional approaches. Comparison with ambient noise tomography results shows good consistency in velocity anomaly patterns at overlapping periods, while highlighting the combination of earthquake and ambient noise data provides a systematic understanding of phase velocity structure across multiple periods. The study demonstrates that PINN-based teleseismic eikonal tomography offers a memory efficient and robust approach for investigating deep Earth structure, particularly beneficial in regions with sparse or unevenly distributed data.

Plain Language Summary

Teleseismic surface wave tomography is a vital tool for studying the Earth's interior using seismic records from earthquakes. In this study, we develop a new method to obtain subsurface images by combining deep learning with physics principles. Traditional methods require large amounts of earthquake data to produce clear images, but our approach, which we call pinnTET, can achieve similar quality results with much less data. We applied this method to the northeastern Tibet. Our results reveal the velocity variations in the crust and upper mantle, which help us understand this region's tectonic evolution and dynamics. When comparing our results with those from other methods, we find a good consistency about the structure beneath Tibet. This method could be particularly useful in regions where earthquake data is limited, helping us better understand Earth's structure in previously challenging areas.

4.1 Introduction

Unraveling the structure of the crust and upper mantle is crucial for understanding a wide range of the Earth's geodynamic processes, including plate tectonics, mantle convection, and earthquake generation. Among various geophysical techniques, seismic surface wave tomography has emerged as a pivotal tool for obtaining the subsurface velocity structure of the crust and upper mantle (Woodhouse and Dziewonski 1984; Anderson 1987; Forsyth et al. 2005; Sabra et al. 2005; Shapiro et al. 2005; Nishida et al. 2009). Surface waves are particularly sensitive to shear wave velocity variations within the crust and upper mantle, due to their dispersive nature and depth-dependent sampling properties. These properties allow surface waves with different periods to provide valuable constraints on the velocity structure at different depths, enabling higher vertical resolution compared to body wave tomography. Surface wave tomography can generate phase and/or group velocity models using either earthquake data (Zhang et al. 2011) or ambient noise (Li et al. 2014). Teleseismic surface wave tomography leverages the long-period surface waves generated by global earthquakes. As these waves propagate through the Earth's interior, they carry information about the subsurface velocity structure over extended distances. In contrast to ambient noise tomography, which relies on shorter-period surface waves, teleseismic tomography can reveal deeper structures into the upper mantle due to the longer wavelengths involved, making it a powerful tool for investigating deeper geological features and geodynamic processes.

With the development of dense arrays, a large number of high-resolution imaging methods have been developed, including the two-station method (Sato 1955; Yao et al. 2005; Yao et al. 2006), two-plane-wave surface wave tomography (Forsyth et al. 1998; Forsyth et al. 2005; Li and Li 2015), and eikonal and Helmholtz tomography (Lin et al. 2009; Lin and Ritzwoller 2011; Jin and Gaherty 2015), which are widely used in analyzing surface wave signals generated by earthquakes. Among these methods, eikonal tomography has emerged as a particularly effective approach. The main steps in eikonal tomography involve wavefront tracking and gradient calculation based on the eikonal equation, enabling it to directly obtain surface wave phase/group velocity without conventional forward and inverse processes (Lin et al. 2009). The eikonal tomography method avoids the errors introduced by manual parameter setting during the inverse processes, and the calculation is relatively straightforward. However, a fundamental limitation of conventional eikonal tomography method is the reliance on simple interpolation schemes for

reconstructing continuous traveltime surfaces. These interpolation methods can introduce significant errors and uncertainties, particularly when station spacing is irregular or sparse, thereby limiting the method's accuracy and resolution (Li et al. 2023).

The application of deep learning has brought new development opportunities to seismology. Deep neural networks (DNNs), composed of multiple layers of interconnected neurons, are able to learn to represent and extract complex patterns from massive amounts of seismic data. For instance, DNNs have been widely used in seismic arrival time picking (Zhu and Beroza 2019), earthquake and noise classification (Seydoux et al. 2020) and earthquake location (Zhang et al. 2020). Beyond these applications, the powerful nonlinear fitting and regression abilities of DNNs have led to substantial advancements in seismic tomography, enhancing both efficiency and precision (Araya-Polo et al. 2018; Bianco et al. 2019; Yang and Ma 2019; Fu et al. 2021; Liu et al. 2021; Jo and Ha 2023; Muller et al. 2023). A key feature of DNNs is their use of automatic differentiation, which eliminates the need for discretization (Baydin et al. 2018). Moreover, deep learning offers significant advantages in its ability to operate without relying on mesh grids. Neural networks (NNs) function as mesh-free methods, capable of evaluating on new data points regardless of their distribution after training (Grossmann et al. 2024). These capabilities enable DNNs to handle the partial differential equations (PDEs) in tomographic problems more effectively. Compared to traditional grid-based methods, deep learning provides a more adaptable and robust approach to solving the eikonal equation.

Physics-informed neural networks (PINNs) represent a novel framework for solving various PDEs using deep learning techniques (Raissi et al. 2019). This approach combines the strengths of data-driven and theory-based modeling, enabling the training of accurate and generalized models even with limited data. It offers several advantages over purely data-driven models, including improved robustness to observation errors and enhanced generalization beyond the training data. In seismology, PINNs have demonstrated remarkable potential in various applications, such as wave equation simulations (Karimpouli and Tahmasebi 2020; Moseley et al. 2020; Waheed et al. 2021; Song et al. 2021; Huang and Alkhalifah 2022a; Huang and Alkhalifah 2022b; Song and Wang 2023; Wu et al. 2023; Chai et al. 2024; Ren et al. 2024), traveltime tomography (Waheed et al. 2021; Agata et al. 2023; Gou et al. 2023; Chen et al. 2022; Chen et al. 2023; Song et al. 2024), wavefield reconstruction inversion (Song and Alkhalifah 2021) and full waveform inversion (Rasht-Behesht et al. 2022; Yang and Ma 2023; Lu et al. 2024), which

include forward and inverse problems in isotropic and anisotropic media and in time-domain and frequency-domain. During these applications, the framework and architecture of PINNs have also been improved using Bayesian theory (Agata et al. 2023; Gou et al. 2023), Fourier feature (Song and Wang 2023; Chai et al. 2024), recurrent neural network (Lu et al. 2024) and generative adversarial network (Yang and Ma 2023). Particularly, Chen et al. (2022) and Chen et al. (2023) were the first to apply PINNs to dense arrays using ambient noise eikonal tomography for both the isotropic and anisotropic case, demonstrating their practicality on large-scale, real-world datasets. However, PINN-based tomography in analyzing long-period surface waves and earthquake data in complex tectonic regions like Tibet remains unexplored.

In this study, we introduce a PINN-based teleseismic eikonal tomography method (pinnTET) to simultaneously retrieve Rayleigh wave phase velocities across multiple frequencies. Our approach employs two types of NNs to model phase velocities and traveltime surfaces. We apply this teleseismic eikonal tomography technique to traveltimes collected from a dense network of stations on the northeastern Tibetan Plateau, highlighting the benefits and challenges of the proposed method.

4.2 Data

The seismic records were collected from ChinArray-Himalaya II, a dense seismic array consisting of 676 stations deployed across the northeastern Tibetan Plateau and its surrounding regions. These stations were deployed between September 2013 and March 2016, with inter-station spacing from 40 to 70 km (Figure 4.1). The earthquake events were selected based on the following criteria: epicentral distances between 20° and 160°, source depths less than 50 km, and magnitude greater than Ms 5.8. Data preprocessing of the ZZ-component included cutting into 1-day segments, resampling to 1 Hz and removing mean value, trends and instrument responses. Following these criteria and preprocessing steps, we identified 119 teleseismic events suitable for surface wave tomography analysis. The spatial distribution of these earthquakes (Figure 4.2) shows generally good backazimuthal coverage, though with a higher concentration of events from the northeast directions and relatively fewer from the northwest. It should be noted that this pinnTET was accomplished utilizing only approximately 20% of the total seismic events, with the number of events varying across different periods, as detailed in Figure C.1 of the Supplementary Material.

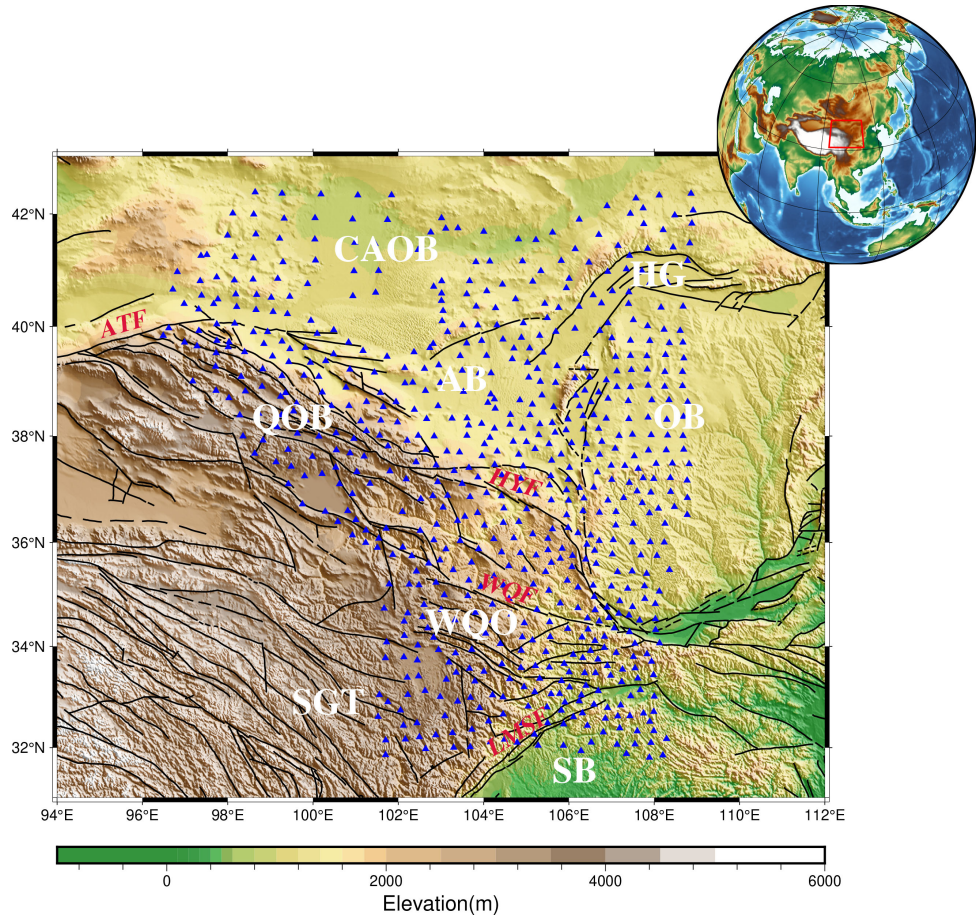


Figure 4.1: Seismic stations geometry and geological background in northeastern Tibetan Plateau. Black lines indicate the main faults, blue triangles indicate stations. The main faults are labeled as: ATF = Altyn-Tagh Fault; HYF = Haiyuan Fault; KLF = Kunlun Fault; LMSF = Longmenshan Fault; WQF = West-Qinling Fault. The main tectonic blocks are labeled as AB = Alxa Block; CAOB = Central Asian Orogenic Belt; HG = Hetao Graben; OB = Ordos Block; QOB = Qilian Orogenic Belt; SB = Sichuan Basin; SGT = Songpan-Ganzi Terrane; WQO = Western Qinling Orogen.

Gee and Jordan (1992) proposed a generalized seismological data functional (GSDF) that employs cross-correlations between observed and synthetic waveforms to extract phase and amplitude from individual seismograms. A five-parameter Gaussian wavelet was then applied to the narrow-band filtered cross-correlation functions to extract a series of dispersion parameters. This method has been widely used to construct models of Earth's crust and upper mantle (Gaherty and Jordan 1995; Gaherty et al. 1996; Gaherty 2001; Gaherty 2004; Chen et al. 2007). Jin and Gaherty (2015) extended the GSDF method by applying cross-correlation functions between neighboring stations to derive dispersion parameters from observed waveforms rather than synthetic waveforms, and then applied it to multi-channel analysis. We use this automated surface wave measuring system (ASWMS) to measure phase delays between all nearby station pairs. The five-parameter Gaussian wavelet used in this approach is defined as (Gee and Jordan

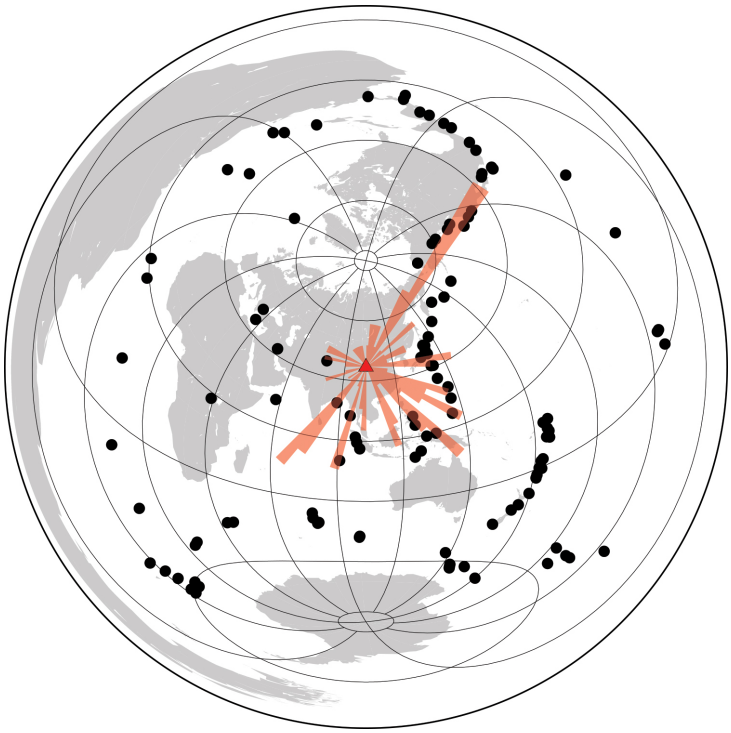


Figure 4.2: Distribution of the earthquakes after data processing. Black dots indicate teleseismic events, red triangle indicates the center of the seismic array, and the rose map indicates the backazimuth distribution of those events.

1992):

$$F_i * W_c C(t) \approx A G_a[\sigma(t - t_g)] cos[w(t - t_p)], \tag{4.1}$$

where $F_i * W_c C(t)$ represents the filtered correlograms obtained by convolving each narrow-band filter F_i (centered at frequency ω_i) with the windowed cross-correlogram $W_c C(t)$, W_c is a window function, $C(t) = S_1 * W_s S_2$ represents the cross-correlation function between record S_1 and windowed record $W_s S_2$, G_a is the Gaussian function $e^{-x^2/2}$, five parameters are A, σ, ω, t_g and t_p, which respectively represent the amplitude factor, the half-bandwidth, the center frequency, the frequency-dependent group delay time and the phase delay time. These parameters are estimated by fitting Equation 4.1 to the filtered and windowed cross-correlogram using a weighted least-squares method. The phase delay t_p is further corrected for cycle-skipping and bias caused by the window function W_s (Gee and Jordan 1992; Jin and Gaherty 2015):

$$\delta\tau_p = t_p - \widetilde{t}_p, \tag{4.2}$$

where $\delta\tau_p$ is the bias-corrected phase delay time, $\widetilde{t}_p$ is obtained by fitting the five-parameter Gaussian wavelet to the filtered, windowed cross-correlation function $\widetilde{C}(t) = S_2 * W_s S_2$ between record S_2 and windowed record $W_s S_2$. We calculated the phase delay time between a given station and several nearby stations, with the inter-station distance usually within a radius of 200 km.

4.3 Methodology

4.3.1 Physics-informed neural network

PINNs leverage the powerful capabilities of DNNs as universal function approximators to solve problems governed by physical laws (Hornik et al. 1989). The fundamental principle of PINNs is to embed PDEs as soft constraints in the training process, allowing the network to learn solutions that naturally satisfy the underlying physical constraints. This is achieved by adding a PDE-based penalty term to the network's loss function (Raissi et al. 2019). Through automatic differentiation, the network learns to approximate functions that map input coordinates to solutions while respecting the governing physical equations.

PINNs provide a flexible framework for solving diverse types of PDEs, generally expressed as (Raissi et al. 2017):

$$u_t + \mathcal{F}[u] = 0, \quad x \in \Omega, \quad t \in [0, T], \tag{4.3}$$

where $u = u(\mathbf{x}, t)$ represents the solution to be determined, $\mathbf{x}$ and t denote spatial and temporal coordinates, respectively, u_t represents the partial derivative of u with respect to t, and Ω is a subset of $\mathbb{R}^d$. The differential operator $\mathcal{F}[\cdot]$ encapsulates the mathematical structure of the PDE, which can accommodate both linear and nonlinear systems, as well as time-dependent and time-independent problems. Given initial and boundary conditions, the physical field $u(\mathbf{x}, t)$ can be solved over the spatiotemporal domain.

The PINN loss function typically consists of two terms: one term ensuring adherence to the governing equations, and another term enforcing consistency with observed measurements. The total loss function in PINNs is expressed as:

$$\mathcal{L} = \sum_{i}^{N_d} ||N_u(\theta; \mathbf{x}_i, t_i) - u(\mathbf{x}_i, t_i)||^2 + \sum_{j}^{N_p} ||L(N_u(\theta; \mathbf{x}_j, t_j))||^2, \tag{4.4}$$

where N_d is the total number of data points, $N_u(\theta; \mathbf{x}_i, t_i)$ refers to the NN's prediction at points $(\mathbf{x}_i, t_i)$, θ denotes the network's parameters that are specific to each NN, $u(\mathbf{x}_i, t_i)$ represents the observed data value, N_p is the number of collocation points for enforcing physical laws, L represents the residual constructed based on Equation 4.3. The first term (denoted the data term) on the right side of Equation 4.4 measures the discrepancy between the network's output and the known values and the second term (denoted the physical term) measures how well the network satisfies the PDE residual at collocation points. This combined loss guides the NN to find solutions that satisfy both the data and the governing equations, forming a robust framework for solving complex physics-based problems.

4.3.2 Teleseismic eikonal tomography using PINNs

In seismic tomography problems we are interested in solving for both the traveltimes and the medium parameters (e.g. velocity) governing the traveltimes through a PDE. Figure 4.3 presents a schematic diagram of PINN architecture for teleseismic eikonal tomography. The pinnTET is composed of two main networks (N_τ and N_c) working in conjunction with physical constraints, ensuring that the predicted velocity (N_c) and traveltime fields (N_τ) are consistent with wave propagation physics while fitting the observed data. In this section, we work with frequency-dependent phase velocities and traveltimes, where the spatial variations of phase velocity are mapped to the observed phase traveltimes through the eikonal equation. The automatic differentiation capability enables efficient computation of derivatives needed for the physical constraints, making the system both physically accurate and computationally tractable.

To further classify this PINN-based approach, it is important to understand the eikonal equation, a first-order PDE linking the surface wave traveltime with local velocity. Traditionally, seismic tomography involves solving a global inverse problem where observed traveltimes are related to velocity structure through ray path integrals. This means that each traveltime observation provides constraints on phase velocities along the entire ray path, rather than at specific points. While this global approach provides robust large-scale structural information, it may have limited resolution for features smaller than the wavelength of seismic waves. In contrast, eikonal tomography directly determines local phase velocities from the spatial gradients

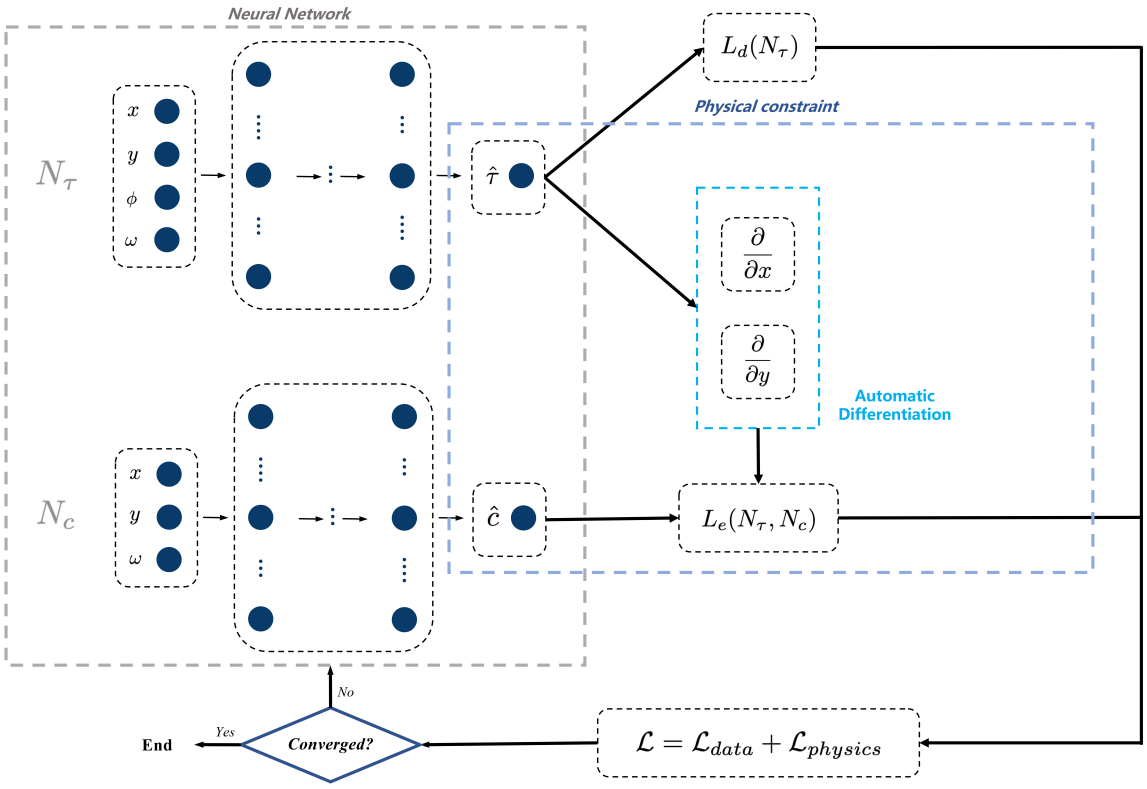


Figure 4.3: Schematic of PINN for teleseismic eikonal tomography. N_τ and N_c are the traveltime and velocity NNs, respectively, x, y represent the spatial coordinates of seismic receivers in the study area, ϕ denotes the measured backazimuth, ω represents the angular frequency of the Rayleigh waves being analyzed, $\hat{\tau}$ and $\hat{c}$ are the predicted traveltime and velocity field from the trained N_τ and N_c. The total loss function ($\mathcal{L}$) consists of two main terms: $\mathcal{L} = \mathcal{L}_{data} + \mathcal{L}_{physics}$, where $\mathcal{L}_{data}$ ($L_d(N\tau)$) represents the data misfit between predicted and observed Rayleigh wave traveltimes, $\mathcal{L}_{physics}$ ($L_e(N\tau, N_c)$) enforces the eikonal equation constraint. The automatic differentiation of NNs calculates spatial derivatives $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$, which relates the traveltime gradient to the local velocity structure. The neural network part on the left is used to approximate the solution, while the data and physical constraints part on the right is used to optimize the network parameters N_θ. The training process should terminate when the model is considered converged.

of traveltime surfaces, avoiding the need for conventional forward and inverse processes (Lin et al. 2009). At each frequency, the eikonal equation provides a direct relationship between the traveltime and the local phase velocity:

$$\frac{1}{c^2(\mathbf{x})} = |\nabla\tau(\mathbf{x})|^2 , \tag{4.5}$$

where $\tau(\mathbf{x})$ represents the phase traveltime from an effective source to a receiver at location $\mathbf{x} = (x, y)$, while $c(\mathbf{x})$ denotes the local phase velocity at the same receiver location. Note that both τ and c are frequency-dependent parameters due to the dispersive nature of surface

waves. The eikonal equation provides a first-order approximation of wave propagation. Based on Equation 4.5, the physical term in the loss function is defined as:

$$L_e(N_\tau, N_c) = |\nabla N_\tau|^2 - \frac{1}{N_c^2}, \tag{4.6}$$

where L_e is the loss function of the eikonal constraint, $N_\tau = N_\tau(\mathbf{x}, \phi, \omega)$ is the NN-based estimates of traveltime at locations $\mathbf{x}$ and angular frequency ω with local backazimuth ϕ, respectively, $N_c = N_c(\mathbf{x}, \omega)$ represents phase velocities evaluated at $\mathbf{x}$ for a specific frequency. Notably, here we used measured backazimuth instead of using source locations or theoretical backazimuth and epicenter distance indicate effective source information (ϕ in Figure 4.3). Using measured backazimuth allows us to eliminate the effects of uncertainties in source location and depth on the results.

Here we use two NNs to model phase velocities and traveltime surfaces. The overall loss function used to train these networks simultaneously has the form:

$$\begin{aligned}\mathcal{L}(\theta_\tau, \theta_c) = \sum_{k}^{N_{freq}} \sum_{j}^{N_{src}} \sum_{i}^{N_{rcv}} \Bigg[\Big| N_\tau(\theta_\tau; \mathbf{x}_i, \phi_j, \omega_k) - \tau_{i,j,k} \Big|^2 \\ + \epsilon_e \Big| L_e\Big(N_\tau(\theta_\tau; \mathbf{x}_i, \phi_j, \omega_k), N_c(\theta_c; \mathbf{x}_i, \omega_k) \Big) \Big|^2 \Bigg],\end{aligned} \tag{4.7}$$

where N_{freq}, N_{src} and N_{rcv} are the maximum number of frequencies, sources and receivers, and i, j, k are the index of those value, respectively. $\tau_{i,j,k}$ is phase traveltime at locations $\mathbf{x}_i = (x_i, y_i)$ and frequency ω_k from direction ϕ_j. ϵ_e is the weight factor for data term and physical term. The loss function design adopted a multi-term combination approach, including data constraint terms (ensuring prediction fits with observational data) and physical constraint terms (guaranteeing solutions satisfy the eikonal equation). The weight factor was determined based on synthetic tests to ensure each term has a reasonable weight during the training, avoiding overfitting for any term.

The training process is to find proper value for θ_τ and θ_c by minimizing the the loss function in Equation 4.7:

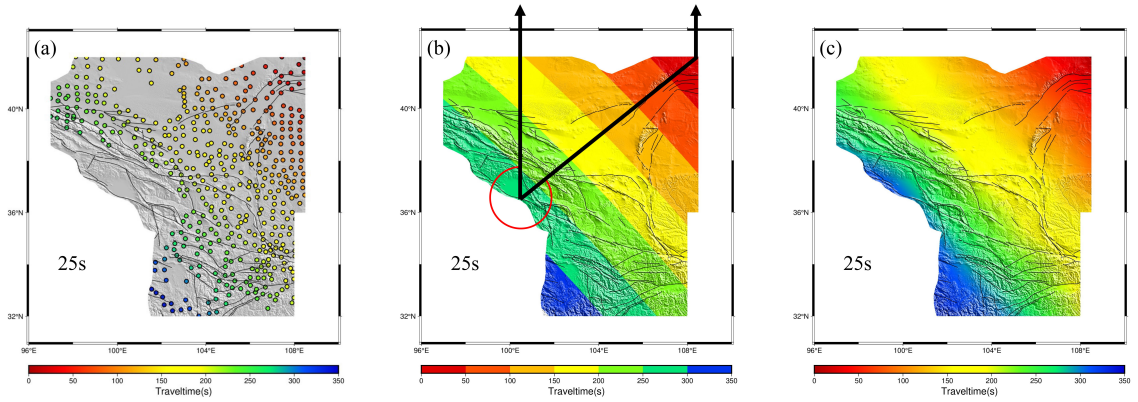


Figure 4.4: (a) Observed Rayleigh wave phase delay time for seismic event 201502131859 measured using the ASWMS. (b) Schematic illustration of the measured backazimuth calculation. The colored surface represents a 2-D linear interpolation of the observed traveltimes shown in (a), the arrow indicates geographic north, the red curve represents the measured backazimtuh that is the clockwise angle between the normal vector to the observed traveltime surface and geographic north. (c) Predicted traveltime surface for seismic event 201502131859, evaluated from the trained N_τ network using the pinnTET method.

$$\arg\min_{\theta_\tau,\theta_c}\{\mathcal{L}(\theta_\tau,\theta_c)\} \rightarrow \hat{\theta}_\tau,\hat{\theta}_c. \tag{4.8}$$

After training, the phase velocities across all frequencies can be direly evaluated from the trained velocity NN N_c. We can also evaluate traveltime surfaces from the trained traveltime NN N_τ.

4.4 Results

4.4.1 Training inputs and targets

We applied the pinnTET method to teleseismic recordings collected from the dense seismic array deployed in the northeastern Tibetan Plateau to obtain Rayleigh wave phase velocities for periods of 20-80 s. As shown in Figure 4.3, the inputs to both NNs include the spatial coordinates (x, y) of stations, measured backazimuths ϕ, and frequencies ω. The training targets are observed traveltime (Figure 4.4 (a)) and the residual of the eikonal equation. The NNs are trained to fit the observed traveltimes while simultaneously driving the residual of the physical equation towards zero. The spatial coordinates of stations were converted from latitude and longitude coordinates to the Cartesian coordinate system using Universal Transverse Mercator (UTM) projection. As shown in Figure 4.4 (b), the backazimuths were determined from the phase delay times by fitting the 2-D linear traveltime surfaces to these delays and calculating the normal vectors to the fitted planes. The backazimuth is defined as the clockwise

angle between geographic north and the traveltime surface normal direction, where the normal direction is measured from the station location to the point of minimum traveltime (source location), giving the effective source direction relative to the receivers without requiring precise source locations. This use of backazimuth, rather than direct source locations, offers several advantages. Since teleseismic earthquakes occur at great distances from the stations, projecting source and receiver locations from latitude/longitude to Cartesian coordinates can introduce errors due to Earth's curvature. The backazimuth, being derived from the direction of wave arrival at the local array, is less sensitive to these distance-related inaccuracies. Furthermore, the method of fitting 2-D linear traveltime surfaces and calculating their normal vectors provides a robust way to capture the wavefront's direction, thereby offering a local constraint for phase velocity determination. Figure 4.5 displays the distribution of the measured backazimuth data along with the distribution of the observed traveltime data. The backazimuth spans nearly the full range of possible angles (0-360 degrees), though the distribution is uneven. The distribution is characterized by several distinct peaks and gaps, with the most prominent peak occurring around 140-150 degrees, reaching nearly 7000 counts. The traveltime distribution shows a more regular, roughly bell-shaped pattern with clear concentration between approximately 50 to 250 s. In addition, the period values are discrete numbers ranging from 20 s to 80 s. All training inputs underwent normalization processing to ensure their values are on the same order of magnitude. This normalization is crucial to prevent errors that could arise from the significant scale differences between input parameters. Without normalization, inputs with larger numerical ranges (like spatial coordinates) might dominate the NN's learning process over inputs with smaller ranges (like periods), leading to biased or unstable training results.

4.4.2 Neural network design

PinnTET employs two separate NNs (N_τ and N_c) to learn the traveltime and the phase velocity respectively. Each network is designed as a deep fully-connected NN with specific architecture to handle the complexity of seismic wave propagation. The input layers for each network are designed to fulfill specific functions. N_τ takes 4-D inputs (x, y, ϕ, p), while N_c takes 3-D inputs (x, y, p), as phase velocity is independent of propagation direction. The design of the PINN architecture follows a "step-by-step design, joint optimization" strategy. First, we need to design the structures of the traveltime network (N_τ) since this network has direct observational data as evaluation criteria. To determine the optimal network architectures, we conducted

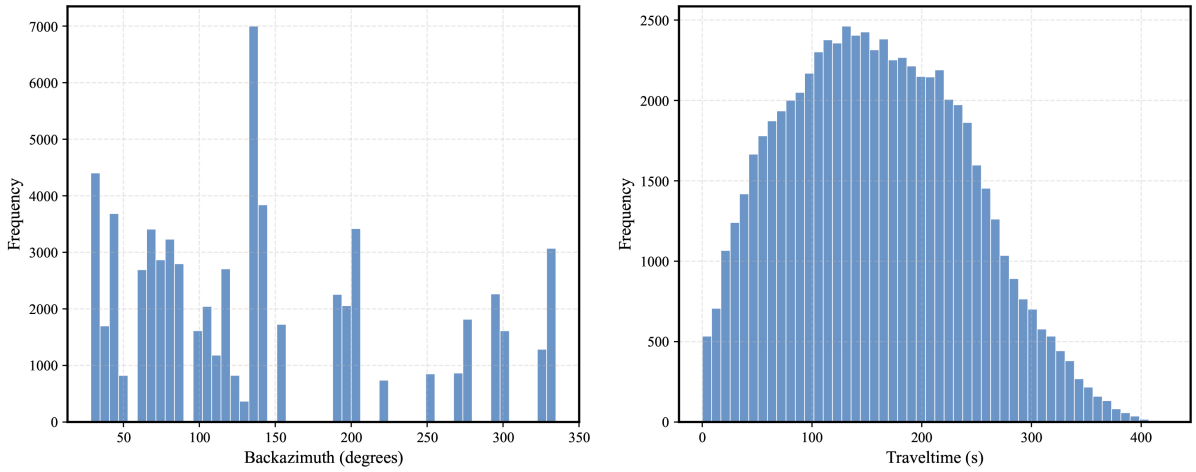


Figure 4.5: (a) Distribution of the measured backazimuth. (b) Distribution of the observed traveltime.

several experiments evaluating the trade-off between prediction accuracy and computational efficiency. The analysis focused on both the model performance and practical implementation considerations. We used the root mean square error (RMSE) to represent the discrepancy between the observations and the NN's predictions:

$$RMSE_\tau = \sqrt{\frac{1}{N_{rcv}} \sum_{i=1}^{N_{rcv}} (\tau_i^{pred} - \tau_i^{obs})^2}, \tag{4.9}$$

The computational time is the training time for the entire dataset. We tested various network configurations with combinations of different numbers of layers and neurons. Specifically, we had configurations with 5, 10, 20 and 30 layers respectively. For each layer configuration, we also tested networks with 5, 10, 20, 30, 40, 50 and 60 neurons. As Shown in Figure 4.6, there is a clear trade-off between the number of layers and neurons in the network and the computation time required. Generally, as the number of layers and neurons increases, the RMSE decreases, indicating better performance. However, this improvement in performance comes at the cost of increased computation time. For example, the RMSE of 5-layer networks decrease rapidly with 5, 10, 20, 30, 40, 50 and 60 neurons respectively, while the RMSE of 30-layer networks have lower RMSE values but much longer computation times with the same range of neurons. Based on these tests, we determined that the best trade-off for N_τ contains 10 hidden layers with 50 neurons in each layer (red circle in Figure 4.7). This configuration achieves a good balance between prediction accuracy (RMSE ≈ 0.2) and computational efficiency (computation time $\approx 5000\ s$), beyond which adding more neurons or layers yields diminishing returns in

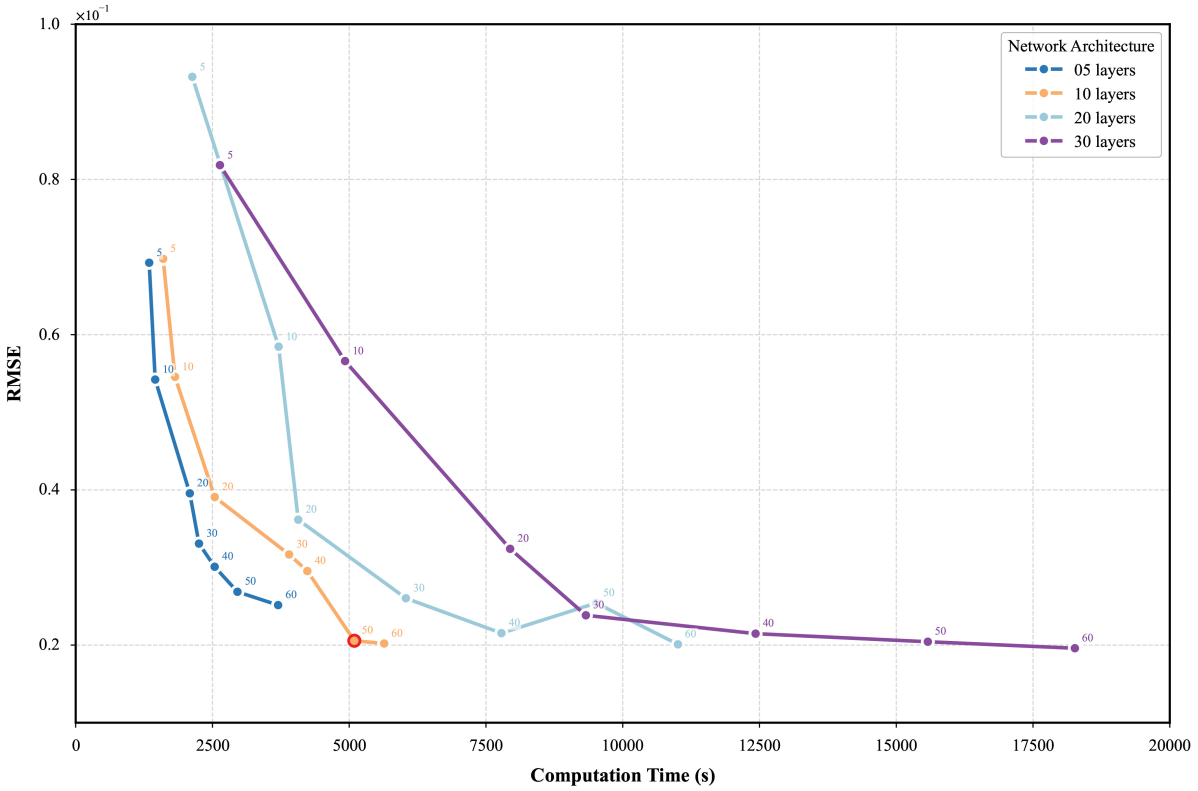


Figure 4.6: Trade-off between traveltime prediction errors and computational cost. The different colors and curves indicate the different number of layers, and the different points and labels indicate the different number of neurons in each layer. The red circle highlights an optimal configuration for traveltime NNs with 10 layers and 50 neurons per layer.

performance improvement. After determining the network structures of N_τ, we proceed to design the structure of the velocity network (N_c). Since N_c has no direct observational data, its evaluation mainly relies on the eikonal equation residual and the prediction accuracy of the entire system. By fixing the optimized structures of N_τ, testing different N_c configurations, observing the convergence of the eikonal equation residual and the overall system performance, we finally determine a 10-layer network with 50 neurons per layer as optimal NN design. This step-by-step design approach not only simplifies the optimization process but also ensures the physical consistency and numerical stability of the entire system.

4.4.3 Training inputs

All other network parameters were determined through systematic synthetic tests. We selected the Adam optimizer due to its ability to effectively handle non-convex optimization problems and excellent performance in processing large-scale data (Kingma and Ba 2014). The learning rate was set to 0.0001, which demonstrated a good balance in synthetic tests. This value

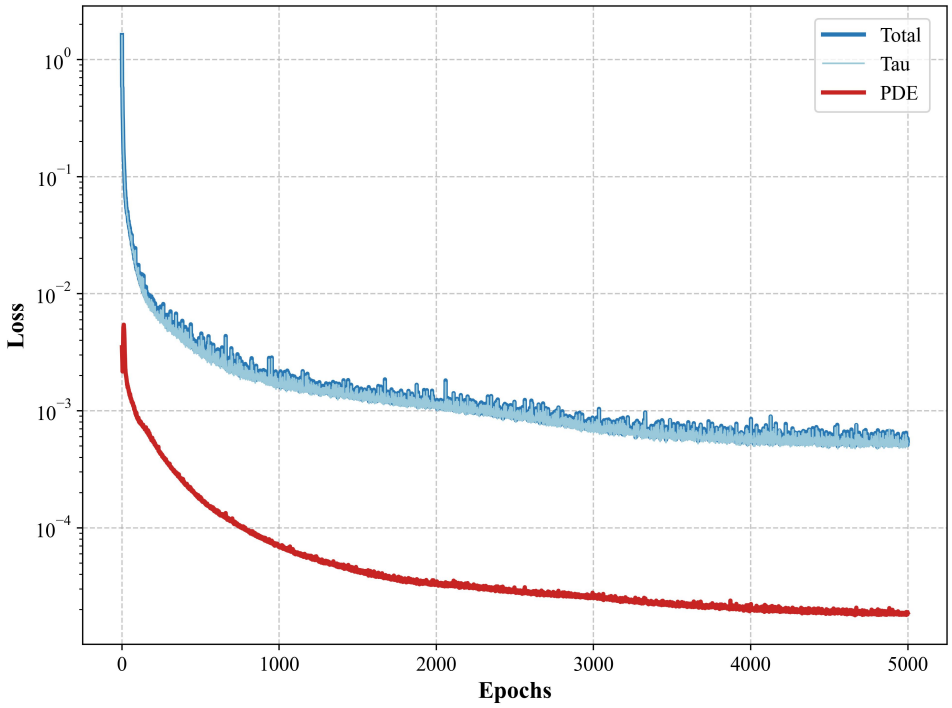


Figure 4.7: Convergence of the total loss function (Total), the data constraint (Tau) and physical constraint (PDE) in Equation 4.7.

is neither too large to cause training instability nor too small to result in slow convergence. The total number of training epochs was set to 5000, determined based on the convergence characteristics of the loss functions. The weights for physical constraints (ϵ_e in Equation 4.7) were set to increase linearly with period, since data at longer periods typically have lower signal-to-noise ratios and are more susceptible to measurement uncertainties. We used arctangent as the activation function, which determines the nonlinear approximation effect of the NN. These parameter selections not only ensure the model's ability to accurately and efficiently extract maximum details from field data but also guarantee the stability and reliability of the training process. Figure 4.7 shows the convergence processes of data constraints, physical constraints, and total loss. The total and data terms converge quickly and uniformly to below 10^{-3} and the physical term converge to below 10^{-4}, indicating fast learning speed and strong training stability.

4.4.4 Phase velocity maps in northeastern Tibet

The Rayleigh wave phase velocity maps across periods of 20-80 s are shown in Figure 4.8, with corresponding velocity uncertainties presented in Figure C.2 of the Supplementary Material. Figure 4.8 indicates the distribution of low velocity zones (LVZs) and high velocity zones (HVZs)

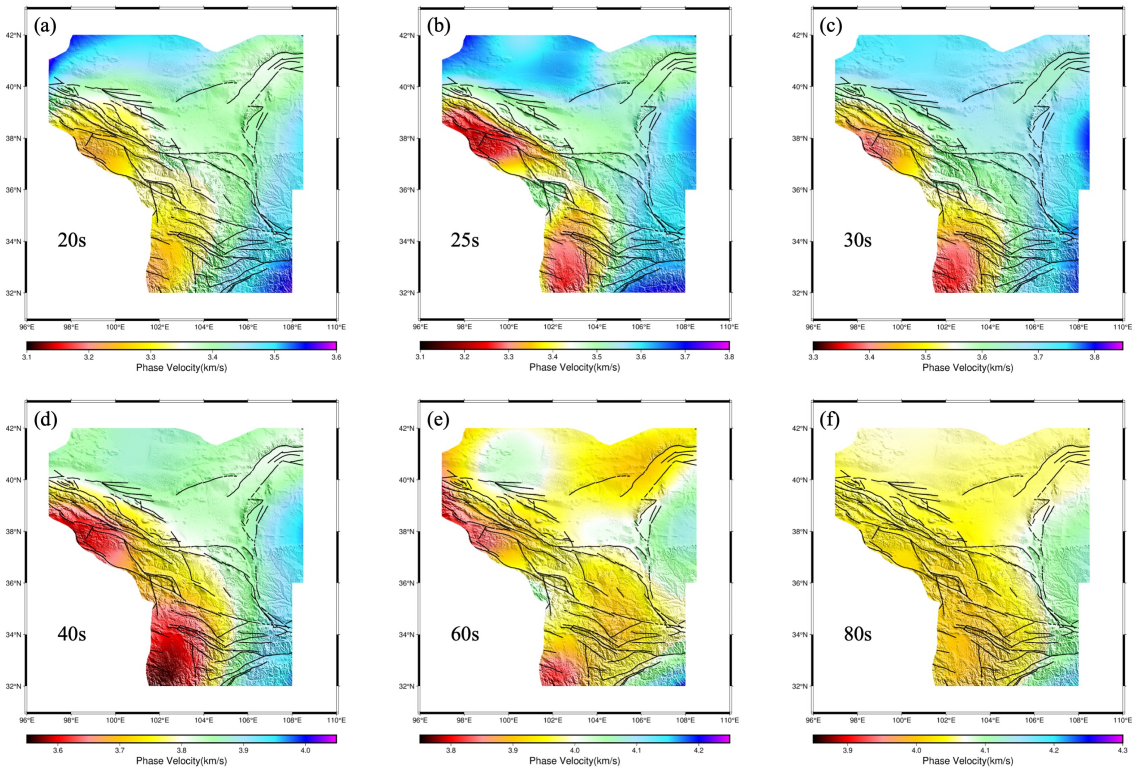


Figure 4.8: (a-f) Rayleigh wave phase velocities at periods of 20, 25, 30, 40, 60, and 80 s beneath northeastern Tibetan Plateau obtained from pinnTET.

are controlled by the strike of faults and the boundaries of blocks, with significant lateral heterogeneity. It can be seen that the phase velocity patterns in different geological units vary with different periods. The phase velocities of different periods are sensitive to the S-wave velocities at different depths with roughly about one-third of its wavelength.

The phase velocities of Rayleigh waves with periods of 20-25 s are sensitive to the S-wave velocities at a depth of 20-40 km, mainly reflecting the velocity variations in the middle crust within the study area. While the actual depth sensitivity is described by continuous kernels rather than discrete depth ranges, this simplified interpretation helps identify the main structural features. Along the southwest boundary, prominent LVZs exist in the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane. From 20 s to 25 s, both the range and strength of these LVZs increase. The Hetao Graben also shows a relatively low-velocity anomaly. HVZs appear in the northern and eastern boundaries and southeastern corner of the area, corresponding to the southern central Asian Orogenic Belt, northern Alxa Block, Ordos Block, and central Qinling Orogen (the north of Sichuan Basin).

The phase velocities of Rayleigh waves with periods of 30-40 s are sensitive to the S-wave

velocities at a depth of 30-70 km, mainly reflecting the velocity variations in the lower crust and the uppermost mantle within the study area. Significant LVZs remain in the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane, while the isolated LVZ in the Hetao Graben disappears. The range of HVZs in the central Qinling Orogen shrinks. The high-velocity anomalies in the southern Ordos Block and central Qinling Orogen form a continuous feature. The high-velocities in the central Asian Orogenic Belt and Alxa Block weaken as the period increases, showing generally uniform medium-high velocity characteristics.

The phase velocities of Rayleigh waves with periods of 60-80 s are sensitive to the S-wave velocities at a depth of 70-120 km, mainly reflecting the velocity variations in the upper mantle within the study area. The lateral heterogeneities of the phase velocity in these blocks are further reduced. Overall, the Ordos Block and central Qilian Orogen show relatively high velocities, while the rest of the areas show relatively low velocities. The HVZs and LVZs are separated by the faults distributed around the blocks. A relatively HVZ appears in the northwestern corner of the study area at 60 s, which may be related to the sparse station distribution in this region. The LVZs along the southwestern boundary of this region almost merge together, with relative low velocity remaining in the Songpan-Ganzi Terrane at 80 s. Phase velocities in the central Alxa Block change from a relatively higher velocity compared to its surroundings to a similar velocity value from 60 s to 80 s.

4.5 Discussion

In this paper, we obtained surface wave phase velocities at periods of 20-80 s using PINN-based teleseismic eikonal tomography, achieving reliable resolution with only approximately 20% of the observational data. In this section, we first compare our results with those from traditional eikonal tomography. Figure 4.9 and C.3 show the phase velocity results obtained by both methods at different periods and their differences. Overall, the phase velocity differences in most regions are within 0.05 km/s, with velocity differences in the northwestern and southeastern corners of the study area likely caused by lower ray coverage. As shown in Figure 4.9 (a) (b) and (c), at 20 s period, both methods observe LVZs in the western Qilian Orogen, western Qinling Orogen, Songpan-Ganzi Terrane, and Hetao Graben, as well as HVZs in the Ordos Block and central Qinling Orogen. However, differences exceeding 0.05 km/s exist at the study area's margins between the two methods. The traditional eikonal tomography results show

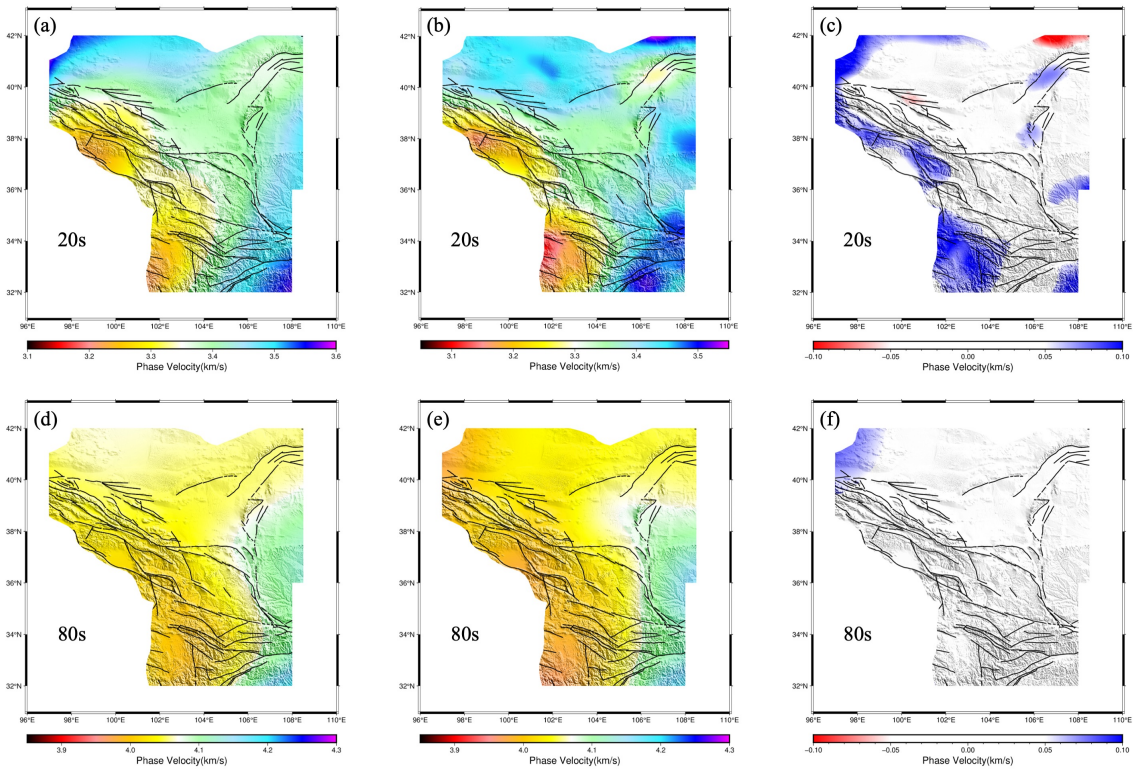


Figure 4.9: (a)(d) Rayleigh wave phase velocity at periods of 20 and 80 s beneath northeastern Tibetan Plateau obtained from pinnTET. (b)(e) Rayleigh wave phase velocity at periods of 20 and 80 s beneath northeastern Tibetan Plateau obtained from conventional eikonal tomography. (a)(d) Difference of Rayleigh wave phase velocity between two methods at periods of 20 and 80 s.

lower velocity anomalies, and the high-velocity connection between the southern Ordos Block and central Qinling Orogen observed in traditional results is not present in the pinnTET results, possibly due to insufficient event numbers and uneven event distribution at 20 s (Figure C.1 (a)). For phase velocity results at periods greater than 20 s, the differences between the two methods are generally less than 0.05 km/s, particularly at 30, 40, and 80 s (Figure 4.9 and C.3). As shown in Figure 4.9 (d) (e) and (f), at the 80 s period, lateral heterogeneity in velocity distribution significantly weakens, but the pattern of low velocities in the southwest and high velocities in the east remains observable. The HVZs in the Ordos Block and central Qinling Orogen show clear correlation with fault distributions.

To verify the reliability of our findings and gain a deeper understanding of the velocity structure characteristics beneath the northeastern Tibetan Plateau, we conducted a detailed comparison between the Rayleigh wave phase velocity results obtained using the pinnTET and the previously published results based on ambient noise data in the same study area (Chen et al. 2023).

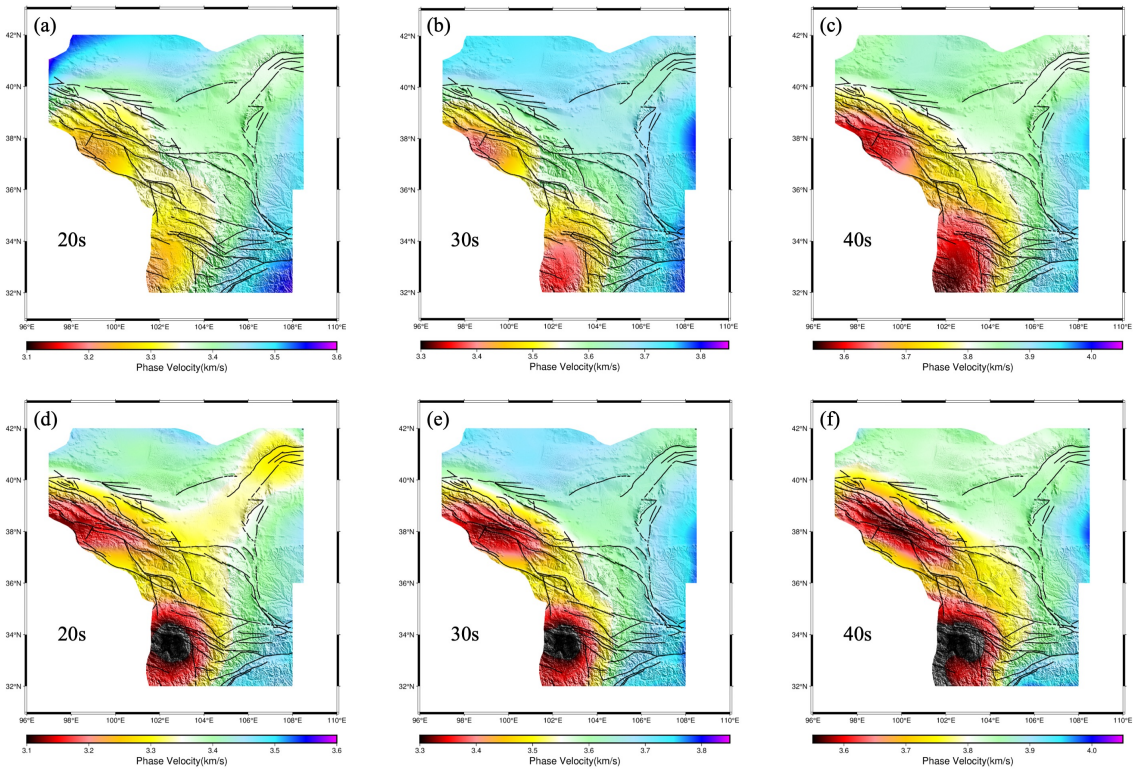


Figure 4.10: (a)-(c) Rayleigh wave phase velocity at periods of 20, 30 and 40 *s* beneath northeastern Tibetan Plateau obtained from pinnTET. (d)-(f) Rayleigh wave phase velocity at periods of 20, 30 and 40 *s* beneath northeastern Tibetan Plateau obtained from pinn-based ambient noise tomography.

In general, teleseismic surface wave tomography is capable of revealing the deeper structure in the lower crust and upper mantle based on surface wave at longer periods generated by earthquakes, whereas ambient noise tomography has the ability to leverage the surface wave at shorter periods, exhibiting higher resolution in the shallow crustal structures. Here we compared results of Rayleigh wave phase velocities at 20, 30 and 40 *s* period, which represents an overlapping interval where both methods exhibit reasonable resolution. Figure 4.10 shows phase velocity maps obtained through PINN-based eikonal tomography using both teleseismic data and ambient noise data from the same dense array. The absolute velocity differences are shown in Figure C.4. At all three periods, both methods observe distinct LVZs along the southwestern boundary of the study area, corresponding to the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane, while relatively HVZs are observed in the eastern to southeastern parts of the study area, corresponding to the Ordos Block and central Qinling Orogen. However, the amplitude of velocity anomalies, especially the low-velocity anomalies, is larger in the ambient noise tomography results compared to the teleseismic surface wave

tomography results. The areas where velocity differences exceed 0.05 km/s gradually decrease with increasing period (Figure C.4), reflecting different constraining capabilities of the two data types at different periods.

At 20 s period, the teleseismic surface wave tomography results (Figure 4.10 (a)) show weaker LVZs in the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane, as well as in the Hetao Graben, compared to the ambient noise tomography results (Figure 4.10 (d)). These LVZs along the southwestern boundary is separated from the LVZ in the Hetao Graben, while in the ambient noise tomography results, the low-velocity anomalies form a continuous feature from southwest to northeast across the Alxa Block. At 30 s, the teleseismic surface wave tomography results (Figure 4.10 (b)) show intensified low-velocity anomalies along the southwestern boundary, though still smaller than the ambient noise tomography results in these area (Figure 4.10 (e)). However, stronger HVZs can be found from the Ordos Block to the central Qinling Orogen compared to the ambient noise tomography results. At 40 s, the velocity differences between the two methods in the western Qinling Orogen and Songpan-Ganzi Terrane are mostly less than 0.05 km/s (Figure C.4 (c)), but significant differences in the low-velocity anomaly amplitude still exist in the western Qilian Orogen. Additionally, the ambient noise tomography results (Figure 4.10 (f)) show a higher velocity anomaly in the Ordos Block than in Figure 4.10 (c). Overall, the two datasets demonstrate good consistency in the distribution of velocity anomalies across major tectonic units. The ambient noise tomography results shows stronger low-velocity anomalies, particularly at shorter periods (20-30 s). This difference likely reflects varying depth sensitivity between the two methods, with ambient noise having better resolution of shallow structures.

By combining ambient noise tomography results at short periods with teleseismic surface wave tomography results, we can develop a systematic understanding of the lithospheric structure in the northeastern margin of the Tibetan Plateau. In the sedimentary layer and upper crust (Chen et al. 2023), the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane exhibit low-velocity anomalies. Notably, distinct LVZs are present in the Hetao Graben, Yinchuan Graben, central Alxa Block, and central Ordos Block, particularly in the Hetao Graben, which is typically attributed to loose basin sediments (Fu and Xiao 2020). In the middle and lower crust, the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane manifest as extensive LVZs, while the Hetao Graben maintains relatively low velocities. The Ordos Block

and central Qinling Orogen display prominent HVZs. These patterns reflect significant tectonic differences in the middle and lower crust between the northeastern margin of the Tibetan Plateau and its surrounding blocks. The LVZs typically indicate lower mechanical strength of the medium, making these regions more susceptible to deformation under tectonic stress. In upper mantle, the relative contrast between low and high-velocity anomalies diminishes. The LVZs extend across all areas except the Ordos Block and central Qinling Orogen. The Songpan-Ganzi Terrane exhibits particularly distinct relative low velocities, potentially associated with asthenospheric upwelling following the delamination of thick lithospheric root (Li et al. 2017).

4.6 Conclusion

This research demonstrates the successful application of physics-informed neural networks to teleseismic Rayleigh wave tomography in the northeastern Tibetan Plateau. The pinnTET method provides several key advantages over existing methods: it requires significantly less observational data while maintaining resolution quality, naturally incorporates physical constraints through the eikonal equation, and shows strong consistency with traditional tomographic methods. The phase velocity maps reveal detailed structural variations that correlate well with known tectonic features in crust to upper mantle. Low-velocity zones in the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane contrast with high-velocity zones in the Ordos Block and central Qinling Orogen, reflecting significant tectonic differences in the region. The integration of deep learning with physical principles opens new possibilities for surface wave tomography, particularly in regions with limited data coverage. By combining teleseismic surface wave data with ambient noise data using PINN-based eikonal tomography, we can achieve a more systematic understanding of the velocity structure across the northeastern Tibetan Plateau.

Acknowledgments

We thank all people who participated in field deployment of the ChinArray-Himalaya II. We thank the China Seismic Array Data Management Center at the Institute of Geophysics, China Earthquake Administration for providing the seismic records used in this study. YP.C. is supported by a split-site PhD program of SUSTech and the University of Leeds.

References

Agata, Ryoichiro, Kazuya Shiraishi, and Gou Fujie (2023). "Bayesian seismic tomography based on velocity-space Stein variational gradient descent for physics-informed neural network". In: *IEEE Transactions on Geoscience and Remote Sensing*.

Anderson, Don L (1987). "Global mapping of the uppermantle by surface wave tomography". In: *Composition, Structure and Dynamics of the Lithosphere-Asthenosphere System* 16, pp. 89–97.

Araya-Polo, Mauricio, Joseph Jennings, Amir Adler, and Taylor Dahlke (2018). "Deep-learning tomography". In: *The Leading Edge* 37.1, pp. 58–66.

Baydin, Atilim Gunes, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind (2018). "Automatic differentiation in machine learning: a survey". In: *Journal of Machine Learning Research* 18, pp. 1–43.

Bianco, Michael J, Peter Gerstoft, Kim B Olsen, and Fan-Chi Lin (2019). "High-resolution seismic tomography of Long Beach, CA using machine learning". In: *Scientific reports* 9.1, p. 14987.

Chai, Xintao, Zhiyuan Gu, Hang Long, Shaoyong Liu, Taihui Yang, Lei Wang, Fenglin Zhan, Xiaodong Sun, and Wenjun Cao (2024). "Modeling multisource multifrequency acoustic wavefields by a multiscale Fourier feature physics-informed neural network with adaptive activation functions". In: *Geophysics* 89.3, T79–T94.

Chen, Po, Li Zhao, and Thomas H Jordan (2007). "Full 3D tomography for the crustal structure of the Los Angeles region". In: *Bulletin of the Seismological Society of America* 97.4, pp. 1094–1120.

Chen, Yunpeng, Sjoerd AL de Ridder, Sebastian Rost, Zhen Guo, Xiaoyang Wu, and Yongshun Chen (2022). "Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the northeastern margin of the Tibetan Plateau". In: *Geophysical Research Letters* 49.21, e2022GL099053.

Chen, Yunpeng, Sjoerd AL de Ridder, Sebastian Rost, Zhen Guo, Xiaoyang Wu, Shilin Li, and Yongshun Chen (2023). "Physics-Informed Neural Networks for Elliptical-Anisotropy Eikonal Tomography: Application to Data From the Northeastern Tibetan Plateau". In: *Journal of Geophysical Research: Solid Earth* 128.12, e2023JB027378.

Forsyth, Donald W, Aibing Li, A Levander, and G Nolet (2005). "Array analysis of two-dimensional variations in surface wave phase velocity and azimuthal anisotropy in the presence of multipathing interference". In: *Geophysical Monograph-American Geophysical Union* 157, p. 81.

Forsyth, Donald W, Spahr C Webb, LeRoy M Dorman, and Yang Shen (1998). "Phase velocities of Rayleigh waves in the MELT experiment on the East Pacific Rise". In: *Science* 280.5367, pp. 1235–1238.

Fu, Lei, Lei Pan, Qingbo Ma, Sheng Dong, and Xiaofei Chen (2021). "Retrieving S-wave velocity from surface wave multimode dispersion curves with DispINet". In: *Journal of Applied Geophysics* 193, p. 104430.

Fu, Yuanyuan and Zhuo Xiao (2020). "Ambient noise tomography of Rayleigh and Love wave in Northeast Tibetan plateau and adjacent regions". In: *Chinese Journal of Geophysics* 63.3, pp. 860–870.

Gaherty, James B (2001). "Seismic evidence for hotspot-induced buoyant flow beneath the Reykjanes Ridge". In: *Science* 293.5535, pp. 1645–1647.

– (2004). "A surface wave analysis of seismic anisotropy beneath eastern North America". In: *Geophysical Journal International* 158.3, pp. 1053–1066.

Gaherty, James B and Thomas H Jordan (1995). "Lehmann discontinuity as the base of an anisotropic layer beneath continents". In: *Science* 268.5216, pp. 1468–1471.

Gaherty, James B, Thomas H Jordan, and Lind S Gee (1996). "Seismic structure of the upper mantle in a central Pacific corridor". In: *Journal of Geophysical Research: Solid Earth* 101.B10, pp. 22291–22309.

Gee, Lind S and Thomas H Jordan (1992). "Generalized seismological data functionals". In: *Geophysical Journal International* 111.2, pp. 363–390.

Gou, Rongxi, Yijie Zhang, Xueyu Zhu, and Jinghuai Gao (2023). "Bayesian physics-informed neural networks for the subsurface tomography based on the eikonal equation". In: *IEEE Transactions on Geoscience and Remote Sensing* 61, pp. 1–12.

Grossmann, Tamara G, Urszula Julia Komorowska, Jonas Latz, and Carola-Bibiane Schönlieb (2024). "Can physics-informed neural networks beat the finite element method?" In: *IMA Journal of Applied Mathematics*, hxae011.

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White (1989). "Multilayer feedforward networks are universal approximators". In: *Neural networks* 2.5, pp. 359–366.

Huang, Xinquan and Tariq Alkhalifah (2022a). "PINNup: Robust neural network wavefield solutions using frequency upscaling and neuron splitting". In: *Journal of Geophysical Research: Solid Earth* 127.6, e2021JB023703.

– (2022b). "Single Reference Frequency Loss for Multi-frequency Wavefield Representation using Physics-Informed Neural Networks". In: *IEEE Geoscience and Remote Sensing Letters*.

Jin, Ge and James B Gaherty (2015). "Surface wave phase-velocity tomography based on multichannel cross-correlation". In: *Geophysical Journal International* 201.3, pp. 1383–1398.

Jo, Jun Hyeon and Wansoo Ha (2023). "Seismic Traveltime Tomography Using Deep Learning". In: *IEEE Transactions on Geoscience and Remote Sensing*.

Karimpouli, Sadegh and Pejman Tahmasebi (2020). "Physics informed machine learning: Seismic wave equation". In: *Geoscience Frontiers* 11.6, pp. 1993–2001.

Kingma, Diederik P and Jimmy Ba (2014). "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980*.

Li, Aibing and Lun Li (2015). "Love wave tomography in southern Africa from a two-plane-wave inversion method". In: *Geophysical Journal International* 202.2, pp. 1005–1020.

Li, Hongyi, Yang Shen, Zhongxian Huang, Xinfu Li, Meng Gong, Danian Shi, Eric Sandvol, and Aibing Li (2014). "The distribution of the mid-to-lower crustal low-velocity zone beneath the northeastern Tibetan Plateau revealed from ambient noise tomography". In: *Journal of Geophysical Research: Solid Earth* 119.3, pp. 1954–1970.

Li, Lun, Chen Cai, Yuanyuan Fu, and Hongjian Fang (2023). "Multiple surface wave tomography methods and their applications to the Tibetan Plateau". In: *Reviews of Geophysics and Planetary Physics* 54.2, pp. 174–196.

Li, Yonghua, Jiatie Pan, Qingju Wu, and Zhifeng Ding (2017). "Lithospheric structure beneath the northeastern Tibetan Plateau and the western Sino-Korea Craton revealed by Rayleigh wave tomography". In: *Geophysical Journal International* 210.2, pp. 570–584.

Lin, Fan-Chi and Michael H Ritzwoller (2011). "Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure". In: *Geophysical Journal International* 186.3, pp. 1104–1120.

Lin, Fan-Chi, Michael H Ritzwoller, and Roel Snieder (2009). "Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array". In: *Geophysical Journal International* 177.3, pp. 1091–1110.

Liu, Bin, Senlin Yang, Yuxiao Ren, Xinji Xu, Peng Jiang, and Yangkang Chen (2021). "Deep-learning seismic full-waveform inversion for realistic structural models". In: *Geophysics* 86.1, R31–R44.

Lu, Cai, Yunchen Wang, Xuyang Zou, Jingjing Zong, and Qin Su (2024). "Elastic Full-Waveform Inversion via Physics-Informed Recurrent Neural Network". In: *IEEE Transactions on Geoscience and Remote Sensing*.

Moseley, Ben, Andrew Markham, and Tarje Nissen-Meyer (2020). "Solving the wave equation with physics-informed deep learning". In: *arXiv preprint arXiv:2006.11894*.

Muller, Ana PO, Clecio R Bom, Jessé C Costa, Matheus Klatt, Elisangela L Faria, Bruno dos Santos Silva, Marcelo P de Albuquerque, and Marcio P de Albuquerque (2023). "Deep-tomography: iterative velocity model building with deep learning". In: *Geophysical Journal International* 232.2, pp. 975–989.

Nishida, Kiwamu, Jean-Paul Montagner, and Hitoshi Kawakatsu (2009). "Global surface wave tomography using seismic hum". In: *Science* 326.5949, pp. 112–112.

Raissi, Maziar, Paris Perdikaris, and George E Karniadakis (2019). "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In: *Journal of Computational Physics* 378, pp. 686–707.

Raissi, Maziar, Paris Perdikaris, and George Em Karniadakis (2017). "Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations". In: *arXiv preprint arXiv:1711.10561*.

Rasht-Behesht, Majid, Christian Huber, Khemraj Shukla, and George Em Karniadakis (2022). "Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions". In: *Journal of Geophysical Research: Solid Earth* 127.5, e2021JB023120.

Ren, Pu, Chengping Rao, Su Chen, Jian-Xun Wang, Hao Sun, and Yang Liu (2024). "SeismicNet: Physics-informed neural networks for seismic wave modeling in semi-infinite domain". In: *Computer Physics Communications* 295, p. 109010.

Sabra, Karim G, Peter Gerstoft, Philippe Roux, WA Kuperman, and Michael C Fehler (2005). "Surface wave tomography from microseisms in Southern California". In: *Geophysical Research Letters* 32.14.

Sato, Yasuo (1955). "Analysis of dispersed surface waves by means of Fourier transform I". In: *Bull. Earthq. Res. Inst* 33, pp. 33–48.

Seydoux, Léonard, Randall Balestriero, Piero Poli, Maarten de Hoop, Michel Campillo, and Richard Baraniuk (2020). "Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning". In: *Nature communications* 11.1, p. 3972.

Shapiro, Nikolai M, Michel Campillo, Laurent Stehly, and Michael H Ritzwoller (2005). "High-resolution surface-wave tomography from ambient seismic noise". In: *Science* 307.5715, pp. 1615–1618.

Song, Chao, Tariq Alkhalifah, and Umair Bin Waheed (2021). "Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks". In: *Geophysical Journal International* 225.2, pp. 846–859.

Song, Chao and Tariq A Alkhalifah (2021). "Wavefield reconstruction inversion via physics-informed neural networks". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–12.

Song, Chao, Hang Geng, Umair bin Waheed, and Cai Liu (2024). "PINNPStomo: Simultaneous P-and S-wave seismic traveltime tomography using physics-informed neural networks with a new factored eikonal equation". In: *arXiv preprint arXiv:2407.16439*.

Song, Chao and Yanghua Wang (2023). "Simulating seismic multifrequency wavefields with the Fourier feature physics-informed neural network". In: *Geophysical Journal International* 232.3, pp. 1503–1514.

Waheed, Umair bin, Ehsan Haghighat, Tariq Alkhalifah, Chao Song, and Qi Hao (2021). "PINNeik: Eikonal solution using physics-informed neural networks". In: *Computers & Geosciences* 155, p. 104833.

Waheed, Umair bin, Tariq Alkhalifah, Ehsan Haghighat, Chao Song, and Jean Virieux (2021). "PINNtomo: Seismic tomography using physics-informed neural networks". In: *arXiv preprint arXiv:2104.01588*.

Woodhouse, John H and Adam M Dziewonski (1984). "Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms". In: *Journal of Geophysical Research: Solid Earth* 89.B7, pp. 5953–5986.

Wu, Yanqi, Hossein S Aghamiry, Stephane Operto, and Jianwei Ma (2023). "Helmholtz-equation solution in nonsmooth media by a physics-informed neural network incorporating quadratic terms and a perfectly matching layer condition". In: *Geophysics* 88.4, T185–T202.

Yang, Fangshu and Jianwei Ma (2019). "Deep-learning inversion: A next-generation seismic velocity model building method". In: *Geophysics* 84.4, R583–R599.

Yang, Fangshu and Jianwei Ma (2023). "FWIGAN: Full-Waveform Inversion via a Physics-Informed Generative Adversarial Network". In: *Journal of Geophysical Research: Solid Earth* 128.4, e2022JB025493.

Yao, Huajian, Robert D van Der Hilst, and Maarten V De Hoop (2006). "Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—I. Phase velocity maps". In: *Geophysical Journal International* 166.2, pp. 732–744.

Yao, Huajian, Guoming Xu, Liangbao Zhu, and Xiang Xiao (2005). "Mantle structure from inter-station Rayleigh wave dispersion and its tectonic implication in western China and neighboring regions". In: *Physics of the Earth and Planetary Interiors* 148.1, pp. 39–54.

Zhang, Qie, Eric Sandvol, James Ni, Yingjie Yang, and Yongshun John Chen (2011). "Rayleigh wave tomography of the northeastern margin of the Tibetan Plateau". In: *Earth and Planetary Science Letters* 304.1-2, pp. 103–112.

Zhang, Xiong, Jie Zhang, Congcong Yuan, Sen Liu, Zhibo Chen, and Weiping Li (2020). "Locating induced earthquakes with a network of seismic stations in Oklahoma via a deep learning method". In: *Scientific reports* 10.1, p. 1941.

Zhu, Weiqiang and Gregory C Beroza (2019). "PhaseNet: a deep-neural-network-based seismic arrival-time picking method". In: *Geophysical Journal International* 216.1, pp. 261–273.

Chapter 5

Discussion

5.1 Summary of work

5.1.1 Summary of research contributions

This thesis presents a comprehensive development and validation of physics-informed neural networks (PINNs) for surface wave tomography, progressing from a basic implementation (isotropic ambient noise tomography) to more complex applications (anisotropic ambient noise tomography and teleseismic earthquake tomography). Through examples from the northeastern Tibetan Plateau, it demonstrates how PINNs can advance seismic tomography technology and better resolve the lithospheric structure. My research began with the development of PINN-based eikonal tomography (pinnET) (chapter 2, Chen et al. (2022)), where I established the foundational framework by integrating the eikonal equation into neural networks (NNs). This novel approach demonstrated that PINNs could effectively reconstruct traveltime surfaces while handling measurement noise, achieving comparable phase velocity results with less data to conventional methods. The success of this initial implementation validated the potential of physics-informed deep learning approaches in seismic tomography. Building on this foundation, I expanded my research to develop PINN-based elliptical-anisotropic eikonal tomography (pinnEAET) (chapter 3, Chen et al. (2023)), addressing the more complex challenge of seismic anisotropy. This extension enabled simultaneous estimation of medium properties across multiple frequencies, providing crucial insights into crustal structure. This method has ability to achieve stable results with only 3% of available stations as effective sources, representing a significant advancement in memory efficiency and practical applicability. In the final phase, I extended the

methodology to PINN-based teleseismic eikonal tomography (pinnTET) (chapter 4), demonstrating the framework's flexibility across different data types. Using only 20% of the data required by traditional methods, PinnTET enabled the investigation of deeper structures at longer periods while maintaining resolution quality. The integration of both earthquake and ambient noise data provided a more comprehensive understanding of velocity structures across multiple periods, from 10 s to 80 s. Overall, my research maintained focus on methodological innovation and practical implementation. The successful application to the northeastern Tibetan Plateau not only validated PINN-based methods through uncertainty quantification and comparison with established results but also provided valuable insights into this geologically complex region. I paid particular attention to optimizing NN hyperparameters and validating results against conventional methods, ensuring the reliability of my findings and establishing a foundation for future applications in seismology and other geophysical disciplines. A consistent achievement throughout my work has been the ability to obtain high-quality results with substantially reduced data volumes that makes these methods particularly valuable for regions with sparse or uneven data coverage.

5.1.2 Addressing the scientific gaps

1. This research successfully addresses the limitations of traditional eikonal tomography methods that rely on generic interpolation algorithms. The developed series of PINN-based methods (pinnET, pinnEAET, and pinnTET) embed the eikonal equation directly into NNs, ensuring that traveltime field reconstruction adheres to wave propagation physics rather than relying on physically inconsistent interpolation methods. This physics-consistent approach inherently eliminates velocity biases and artifacts caused by interpolation. My research demonstrates that PINN methods excel with sparse or irregularly distributed data, requiring much less data volume needed by traditional methods to achieve comparable quality results. Furthermore, the continuous representation provided by PINNs offers significant advantages over traditional grid-based methods, allowing researchers to evaluate NN outputs at any spatial location and frequency within the study domain, providing great convenience for subsequent analysis.

2. Regarding the limitations of PINNs in seismological applications, this research provides viable solutions through multiple improvements for practical applications. I implemented reliable data uncertainty assessment mechanisms, addressing stability issues of PINN methods. PINN-based surface wave tomography provides a unified framework that naturally integrates forward

modeling and inversion processes, simultaneously optimizing traveltime prediction and medium property reconstruction by incorporating both data fitting and physical constraints into the loss function. By applying these methods to field data from northeastern Tibet, I demonstrated the effectiveness of PINN methods in handling complex geological conditions, extending beyond previous studies limited to synthetic or idealized cases. The revealed velocity structure and anisotropy patterns exhibit remarkable consistency with the region's principal tectonic features, enhancing our understanding of lithospheric architecture within the study area.

5.1.3 Evaluation of aims and objectives

The two primary aims of this thesis were achieved through the systematic development of three complementary methodologies (pinnET, pinnEAET, and pinnTET). These methods advanced from isotropic to anisotropic, ambient noise to teleseismic earthquake applications, while demonstrating significant practical advantages in data efficiency and resolution capabilities when applied to the recorded dataset. The successful realization of these aims was accomplished through meeting five specific technical objectives, detailed below:

Reshape eikonal tomography as a PINNs training problem where the NNs predict the medium properties and the traveltime observations.

I designed a PINN framework for eikonal tomography that comprises NNs with complementary functions. The traveltime NN (N_τ) takes 4-D (or 5-D) inputs, including station coordinates (x, y), source information (either source coordinates S_x, S_y or backazimuth ϕ), and period p (enabling multi-frequency training). The velocity NN (N_c, or medium property NN $N_{\mathbf{M}}$) operates with 3-D input, consisting of station coordinates (x, y) and period p. The eikonal equation was incorporated as a physical constraint in the loss function, where the training process simultaneously fits the traveltime surfaces and minimizes the eikonal equation residual. This physical constraint ensures wave propagation principles are involved during the inversion. After training, phase velocities can be directly evaluated through the trained velocity network, providing an efficient way to extract physically plausible velocity models.

Quantify and analyze the uncertainty in the estimated phase velocity models.

I conducted uncertainty analysis by utilizing the traveltime NN predictions to calculate average velocities across different sources, and the uncertainty was quantified through standard deviations from these average velocities. The uncertainty map serves as a reasonable metric relative

to the average velocity map from the traveltime NN, but it is only a proxy for the velocity values directly inverted by the velocity NN. The uncertainty map adopted in this thesis is a statistical measure of the average velocity distribution, not the model uncertainty directly output by the velocity NN. While it can indicate the local dispersion of data fitting, it is essential to note its distinction from model uncertainty in the strict sense. Overall, the PINN-based ambient noise tomography results generally showed lower uncertainties compared to PINN-based teleseismic surface wave tomography results, demonstrating the inherent challenges in resolving deeper structures with longer-period teleseismic surface waves. For ambient noise data (10-40 s), the uncertainty maps show consistently low values (below 0.02 km/s) across most of the study area. The distribution of uncertainties remains relatively uniform from 10-30 s, with slightly higher uncertainties (up to 0.03 km/s) appearing around the edges of the network. At 40 s, while maintaining generally low uncertainties in the central region, there is an increase in uncertainty values in the northeastern part of the study area (reaching about 0.04 km/s). Teleseismic data (20-80s) exhibits more complex spatial patterns with period-dependent variations. At shorter periods (20-30 s), uncertainties remain relatively low (below 0.05 km/s) and uniformly distributed. However, as period increases, uncertainties become larger and more spatially variable. For 60 s and 80 s, significantly higher uncertainties (reaching about 0.1 km/s) are observed in the western and southwestern regions of the study area, particularly in Qilian Orogen and Songpan-Ganzi Terrane. This spatial pattern likely reflects both the structural complexity of these regions and the variations in data quality and volume.

Develop an approach for choosing the hyperparameters of the NNs to ensure their ability of accurately representing the effect of complex geology on phase velocity and traveltimes.

I optimized the hyperparameters following a strategy that emphasized independent optimization of each parameter while maintaining others constant, enabling clear assessment of the effects of individual parameters on model performance. For instance, the size of the traveltime NN was determined by evaluating the trade-off between prediction accuracy and computational cost. After establishing the optimal traveltime network configuration, the velocity network architecture was determined by testing different sizes and assessing their performance while keeping the traveltime network fixed. Similar progressive optimization from simple to complex configurations was applied to other hyperparameters until the trained networks demonstrated the

ability to represent complex geological effects. The effectiveness of the hyperparameter selection was evaluated through multiple criteria including prediction accuracy between observed and predicted data, convergence behavior of the loss function, resolution capability through checkerboard tests, and uncertainty estimates. This comprehensive evaluation framework ensured that the final network configuration achieved both physical consistency and computational efficiency while maintaining the ability to resolve complex geological structures.

Adapt the setup of the NNs and validate PINN-based surface wave tomography for phase velocity tomography across range of periods using using both ambient noise and teleseismic data.

I adapted the NN structures for different data types and volumes across pinnET (single-frequency ambient noise data), pinnEAET (multi-frequency ambient noise data), and pinnTET (multi-frequency teleseismic earthquake data). To accommodate multi-frequency training, I modified both the traveltime and medium property networks by adding period as an additional input term, enabling direct evaluation of traveltimes and velocities at any desired period. The network architectures were then scaled appropriately based on the volume of observational data for each application. For different data types, I implemented distinct network design strategies. The batch size selection followed a unified principle of approximating one complete traveltime surface, but was applied slightly different according to data characteristics. For ambient noise cross-correlations, I set the batch size equal to the number of stations, while for teleseismic data, I determined batch sizes by dividing the total number of traveltimes by the total number of earthquakes across all periods. The treatment of the eikonal constraint weight (ϵ_e) also varied between data types. For ambient noise data, where data quality remains relatively consistent across periods, I applied a constant ϵ_e value. However, for teleseismic data, where quality generally decreases with increasing period, I implemented a period-dependent linear increase in ϵ_e. This gradual increase in the physical constraint's influence helps compensate for declining data quality at longer periods while maintaining solution stability. These adaptations enabled successful application of the methods across different data types and frequency ranges, as validated through the consistency of results in checkerboard resolution tests and agreement with conventional methods.

Compare the surface wave phase velocity maps obtained using PINN-based surface wave tomography to traditional eikonal and beamforming tomography techniques.

I compared the results from PINN-based surface wave tomography with those from traditional methods. Specifically, the phase velocity results from pinnET and pinnTET were compared with conventional eikonal tomography results, showing differences generally less than 0.05 km/s in well-sampled regions. For pinnEAET, I conducted more extensive comparisons, evaluating both phase velocities and azimuthal anisotropies against results from conventional eikonal tomography and double beamforming tomography. In regions with good ray coverage, all methods showed consistent velocity patterns and fast directions. Moreover, I conducted a specific comparison using data from just a single source to evaluate the performance of different methods. The results demonstrated that PINN-based surface wave tomography maintained reliable imaging capabilities even with such limited data, producing results similar to those obtained from sufficient data coverage. In contrast, conventional eikonal tomography showed many artifacts and smearing effects under the same limited data conditions. This difference highlights the power of PINNs in stabilizing the inversion and extracting maximum information from limited observations. These comparisons validate the reliability of PINN-based approaches while highlighting their advantages in data efficiency. The consistent results across different methods and data types provide strong support for the robustness of PINN-based surface wave tomography.

5.2 Methodological evaluation

5.2.1 Comparative advantages

The PINN-based surface wave tomography methods (pinnET, pinnEAET, and pinnTET) developed in this thesis demonstrate several significant advantages over traditional tomography methods, as evidenced by the results across different applications.

Most notably, these methods achieve comparable results using only approximately 20% or even less of the data typically needed for conventional tomography. For ambient noise tomography, pinnET produced comparable resolution using just 10 source stations compared to traditional approaches that would use all available stations (676), while pinnEAET achieved stable anisotropic results with only 20 sources. This efficiency extends to teleseismic applications, where pinnTET generated robust velocity models using approximately one-fifth of the available earthquake data. These substantial reductions in data requirements across multiple tomographic applications demonstrate the potential of PINN-based surface wave tomography in dramatically improving data efficiency while preserving resolution quality.

Furthermore, the PINN-based eikonal tomography provides natural regularization through the eikonal equation constraint, improving the reduction in resolution caused by linear interpolation in traditional eikonal tomography. This physics-guided interpolation provides physically plausible models in regions with sparse or unevenly distributed data, as shown by checkerboard resolution tests and uncertainty analyses. This ability is particularly validated in the single-source experiment, where PINN-based approach produced velocity models closely resembling those obtained with sufficient data coverage, while traditional eikonal tomography showed significant artifacts under the same limited data conditions.

Moreover, PINN frameworks demonstrate flexibility in both multi-frequency processing and result evaluation. PINN-based approaches can achieve simultaneous inversion across multiple frequencies simplified by incorporating period terms directly into the NNs. This joint processing ensures more coherent velocity structures across different depth ranges, as demonstrated by Rayleigh wave phase velocity results spanning periods from 10 to 80 s. Once trained, the outputs of NNs can be evaluated at any spatial location and frequency within the study domain without additional interpolation or computation. The continuous representation of PINNs stands in contrast to traditional grid-based methods, which are limited to fixed sampling points. This ability proves particularly valuable for subsequent analysis, allowing researchers to examine structural variations at any desired scale without the computational burden of reprocessing data.

Another advantage is that PINN-based surface wave tomography provides a unified framework that naturally integrates forward modeling and inversion processes. Unlike traditional methods that typically separate forward simulation and inverse problem into different steps, these approaches simultaneously optimizes the predicted wavefield (forward problem) and recovers the medium properties (inverse problem) through the physics-constrained training process. This integration is achieved by incorporating both data fitting and physical constraints into the loss function, where the traveltime NNs learn to predict wave propagation while the velocity NNs reconstruct the medium properties. This simultaneous optimization not only simplifies implementation but also enhances solution stability, as the physical constraints guide both the forward and inverse components throughout the training process.

5.2.2 Practical implementation consideration

The implementation of PINNs to field data presents several challenges that require careful consideration. Here, I will analyze the implementation strategies and their feasibility for PINN-based surface wave tomography methods.

The initial step involves data preprocessing and training set preparation. The data preprocessing follows conventional seismological practices, usually including mean removal, trends removal, resampling, filtering, and instrument response removal. After obtaining observed traveltimes through cross-correlations, selecting an appropriate training dataset volume becomes crucial. A common challenge in PINN implementation is the lack of guidelines regarding optimal data quantity requirements. To address this challenge, I conducted experiments with varying data volumes to empirically determine the optimal amount of training data. For example, through empirical analysis with increasing source numbers (as demonstrated in Figure B.2, Appendix B), I observed that the average phase velocity stabilized when the number of sources reached approximately 20 for pinnEAET, with minimal variation upon further increases. It is important to note that the optimal data quantity requires problem-specific evaluation. To ensure training stability, all data undergo normalization to similar magnitude ranges, preventing larger-valued parameters from dominating the learning process over smaller-valued ones.

The network architecture was designed through a a step-by-step strategy, starting with simple architectures and gradually increasing complexity while monitoring performance improvements. For the traveltime NNs, I evaluated the trade-off between prediction accuracy (e.g. measured by RMSE between predicted and observed traveltimes) and computational cost (e.g. training time) as this network was direct observational data as evaluation criteria (illustrated in Figure 4.6, chapter 4). The NN architecture should provide the optimal balance, achieving RMSE below target threshold while maintaining reasonable training times. After establishing the size of traveltime NNs, I systematically increased the number of layers and neurons in the velocity NNs until performance improvements reached a stable level. In addition, PINN implementations require particular attention to activation function selection due to partial differential equations (PDEs) operations, as activation functions determine the NN's nonlinear approximation characteristics. The arctangent function was selected as the activation function since its smooth derivatives better matched wave equation physics compared to alternatives like ReLU (Cuomo et al. 2022).

The training optimization framework required systematic tests of multiple critical hyperparameters. The learning rate was determined through a decay strategy, starting with a relatively large value (0.001) and gradually decreasing based on loss function convergence. The batch size was determined through the principle of preserving complete traveltime surfaces in each batch. Specifically, I configured the batch size to match the number of measurements that constitute a single traveltime surface for each source. This approach ensures that each batch contains the full spatial sampling of wave propagation, maintaining the physical coherence of the data during training. The physical constraint weighting parameter ϵ_e required particularly careful tuning within a narrow effective range. It needs to be small enough to allow velocity updates but big enough to maintain physical constraints. The selection of ϵ_e should also account for both the quantity and quality of training data. While ambient noise applications performed well with constant weights, teleseismic data required period-dependent scaling (chosen linear increase with period as the simplest approach) to compensate for declining signal quality at longer periods. This adaptation proved essential for maintaining solution stability.

To ensure robust parameter selection, each hyperparameter was tuned independently while maintaining sufficient training time and keeping other conditions constant, allowing for clear comparisons of their individual effects on model performance. This methodical optimization strategy ensured both data efficiency and solution accuracy while maintaining physical consistency.

5.2.3 Critical assessment of physics-informed regularization

While the presented PINN-based approaches have demonstrated promising results, it is essential to critically examine the fundamental value added by the physics-informed framework compared to conventional methods. A key question emerges: given that the traveltime network (N_τ) could theoretically be differentiated in space to obtain a velocity field consistent with the eikonal equation, what genuine advantage does the additional velocity network (N_c) and the physics-informed constraint provide?

In principle, one could directly compute gradients of the traveltime surfaces from N_τ to derive velocities without employing a separate velocity network. However, this direct approach frequently leads to unstable solutions, particularly when dealing with field data containing measurement noise. The differentiation operation amplifies noise in the traveltime surface,

potentially resulting in unphysical velocity artifacts. The introduction of N_c and the eikonal constraint serves as a form of regularization, imposing smoothness on the solution space.

When comparing with regularization in conventional eikonal tomography, PINNs rely on the eikonal equation as a soft constraint within a high-dimensional NN, whose implicit regularization is less interpretable. The advantage of PINNs lies in their ability to adaptively balance data fitting and physics, especially in regions with sparse data. However, the subjective choices in NN's design (e.g., NN size, activation functions) may unintentionally bias solutions. The strong performance in sparse data scenarios, as demonstrated in the single-source experiments where PINN-based approaches outperformed conventional eikonal tomography, suggests that the PINNs can indeed extract more information from limited data than traditional approaches. This may be attributed to the NN's capacity to learn spatially varying regularization patterns that adapt to the underlying physics rather than applying uniform smoothing constraints.

Future research should consider architectures with more physical interpretability, such as radial basis function networks which have clearer spatial localization properties. These approaches might bridge the gap between traditional regularization methods and NN flexibility, providing a more transparent framework for understanding how the regularization influences the final solution.

5.2.4 Current limitations

Through the integration of data and physical constraints, PINN-based surface wave tomography has demonstrated significant advantages and helped reduce training data requirements. However, several key limitations remain in its real-world applications.

First, each method requires specific design considerations for each application scenario, with performance heavily dependent on careful parameter tuning. As demonstrated across the thesis chapters, different applications - whether handling multiple frequencies (chapters 3 and 4), incorporating anisotropic parameters (chapter 3), or adapting to different data types (ambient noise in chapters 2 and 3 versus teleseismic data in chapters 4) - demand substantial experimentation for network design and parameter adjustment. Specific parameter settings that work well for one problem may fail for another, and incorrect parameter settings can lead to local optimum solution, requiring repeated testing and manual adjustments. While design principles have been discussed in previous chapters, automated parameter optimization methods remain a

critical area for future research, as current parameter tuning still relies heavily on comparisons with traditional approaches to validate and optimize performance.

Second, while PINN-based approaches (pinnET, pinnEAET, and pinnTET) have significantly reduced data requirements in field applications, they haven't advantages for demonstrated computational efficiency. Traditional methods often require less processing time even with larger input datasets compared to PINNs. As data volumes and training parameters increase, the global approximation nature of PINNs requires processing large datasets and complex computational graphs in each iteration, leading to high computational resource requirements and slow computational speed. Furthermore, balancing physical constraints with data fitting in sophisticated loss functions can create convergence conflicts, further increasing computational resource demands. In resource-constrained environments, this limitation can make it impractical to achieve real-time processing or rapid calculations, potentially limiting the method's applicability in time-sensitive scenarios or large-scale applications.

Third, although incorporating physical constraints enhances model generalization and interpretability, accurately modeling subsurface velocity structures at arbitrary locations remains challenging. NNs cannot truly learn the underlying physical processes for different datasets, and the trained models only reflect solutions under specific data and physical constraints. While the networks can predict velocities at arbitrary positions, accuracy significantly decreases when evaluation points fall outside the training data distribution. This limitation is especially apparent in areas without stations, particularly near the boundaries of the study area, where uncertainties increase significantly. Consequently, any change in study area or physical constraints may require complete system retraining.

Finally, the implementation of physical constraints remains controversial. Theoretically, the competition between physical loss and other loss terms in the optimization objective may lead to solutions that prioritize data fitting over physical consistency. This balance becomes particularly critical when dealing with noisy or incomplete data, as the physical constraints and noise may contain inconsistent information. While I have attempted to address this through adjustments of the eikonal constraint weight (ϵ_e) for different applications, the fundamental question of how effectively physical principles are imposed through loss function regularization requires further investigation.

5.3 Geological implications for northeastern Tibetan Plateau

Based on seismic ambient noise data (10-40 s) and teleseismic earthquakes (20-80 s), PINN-based surface wave tomography was applied to measure the Rayleigh wave phase velocity and azimuthal anisotropy beneath the northeastern Tibetan Plateau. Figure 5.1 presents a 3-D representation of multi-period Rayleigh wave phase velocity and azimuthal anisotropy structure in the study area, where the Z-axis from top to bottom indirectly reflects the velocity structure from shallow to deep depths. The results reveal significant lateral heterogeneity in the study area, showing good correlation between geological features and subsurface velocity structures. In the crust and upper mantle, the velocity structure shows generally low velocities within the Tibetan Plateau transitioning to high velocities in surrounding regions. In the crust, the azimuthal anisotropy varies gradually across the region, with fast directions near the Tibetan Plateau showing a clockwise rotation trend along the northeastern margin of the plateau.

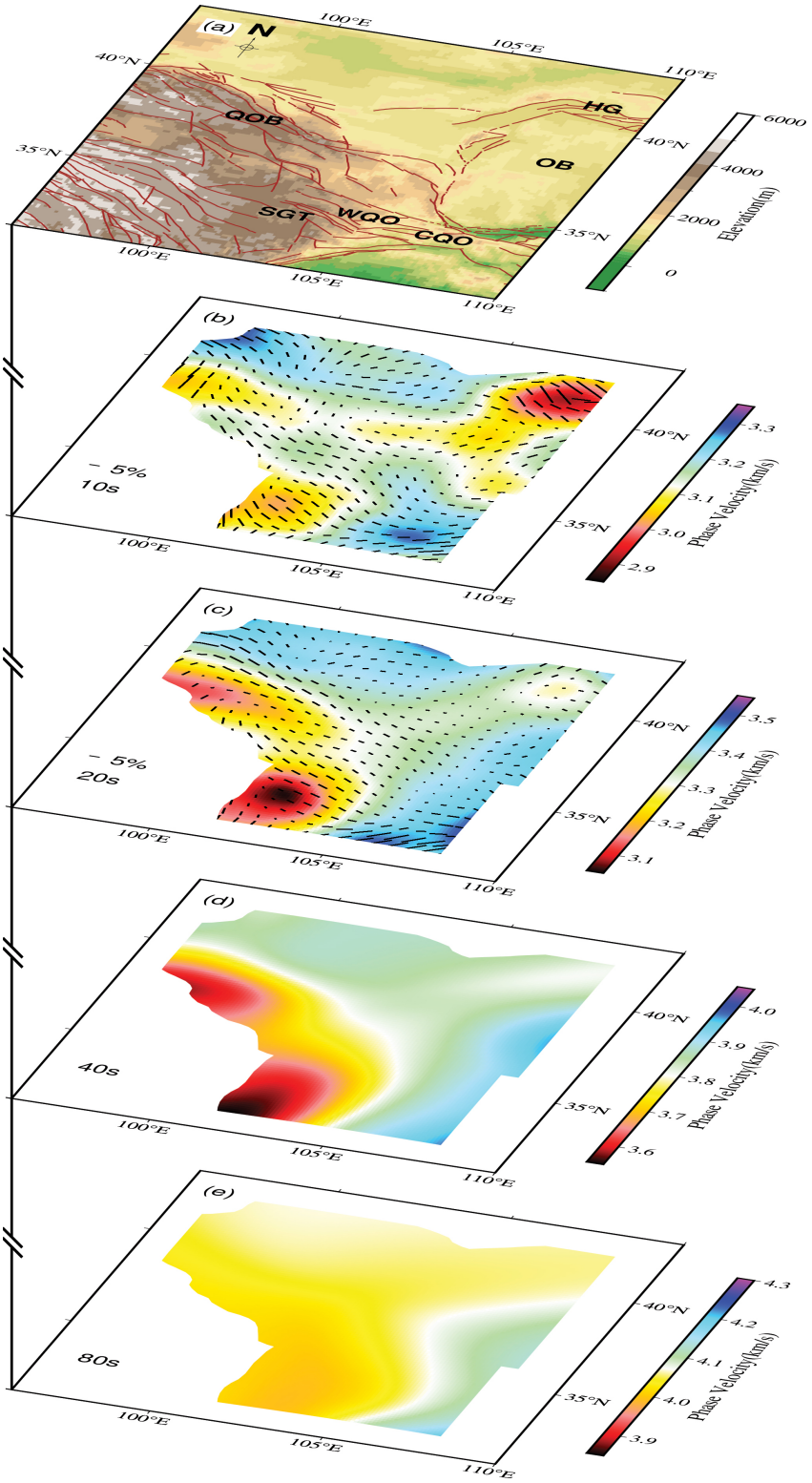


Figure 5.1: 3-D representation of Rayleigh wave phase velocity structure at 10, 20, 30 and 40 s and azimuthal anisotropy at 10 and 20 s. (a) Topography and geological units of northeastern Tibetan Plateau. Red lines indicate the main faults. The main tectonic blocks are labeled as AB = Alxa Block; CQO = Central Qinling Orogen; HG = Hetao Graben; OB = Ordos Block; QOB = Qilian Orogenic Belt; SGT = Songpan-Ganzi Terrane; WQO = Western Qinling Orogen. (b) Azimuthal anisotropic phase velocity at period of 10 s obtained from pinnEAET. (c) Azimuthal anisotropic phase velocity at period of 20 s obtained from pinnEAET. (d) Isotropic phase velocity at period of 40 s obtained from pinnTET. (d) Isotropic phase velocity at period of 80 s obtained from pinnTET.

The western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane exhibit prominent low-velocity zones (LVZs) across all periods, which is highly consistent with previous surface wave studies using different methods (Li et al. 2017; Zhong et al. 2017; Wang et al. 2020; Hao et al. 2021; Wu et al. 2023). These low-velocity anomalies indicate that the crustal and upper mantle media in this region have relatively low mechanical strength and are prone to deformation under tectonic stress, corresponding to the fast uplift of the northeastern Tibetan Plateau (Zhong et al. 2017). However, interpretations of these crustal LVZs remain controversial. For instance, Li et al. (2014) suggest that the LVZ beneath the western Qilian Orogen might be related to crustal thickening, while Jiang et al. (2014) propose that the mid-crustal LVZs result from partial melting. Our results show that the Songpan-Ganzi Terrane has lower velocities and stronger azimuthal anisotropy in the middle crust compared to the western Qilian Orogen. These findings may support different origins for the LVZs in these two tectonic units: the LVZ beneath the Songpan-Ganzi Terrane might be caused by crustal partial melting, while the LVZ beneath the western Qilian Orogen might be attributed to fluid infiltration (Li et al. 2017). Clark et al. (2005) propose that there exists a channel flow with low viscosity in the middle-lower crust within the Tibetan Plateau, flowing outward around the Sichuan Basin. The LVZ in the middle-lower crust beneath the Songpan-Ganzi Terrane, combined with strong positive radial anisotropy (Li et al. 2022), can serve as evidence for the existence of crustal flow in this region, though the extent of this crustal flow requires further discussion. The LVZs in the upper mantle beneath the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane indicate weak lithosphere and the presence of mantle flow in these regions.

The velocity characteristics of the regions surrounding the northeastern Tibet differ from the LVZs within the plateau. The central Qinling Orogen exhibits a high-velocity zone (HVZ) across almost all periods, particularly the section between the Sichuan Basin and Ordos Block shows relatively high velocities throughout the crust, suggesting that crustal flow may not have passed through this area or is in an early developmental stage of crustal flow (Bao et al. 2013). The Ordos Block shows distinct low velocities at 10 s, likely related to its thick sedimentary layers (Sun et al. 2010). At 20 s and longer periods, the lithosphere in this region displays prominent high-velocity characteristics accompanied by weak crustal azimuthal anisotropy, with fast directions well correlated with the fault strikes along the block boundaries, reflecting its rigid characteristics as a stable Cratonic block. However, near its western boundary, the phase velocities are lower than within the block and show significant lateral heterogeneity, suggesting that the

stable Ordos Block may have been affected by tectonic activities and undergone deformation at its western margin (Zhong et al. 2017). The Hetao Graben also exhibits significant low-velocity anomalies in the shallow crust, associated with approximately 10 km thick sedimentary layers (Wang et al. 2017). This LVZ persists up to 25 s period, which Ge et al. (2011) attributed to partial melting. However, since the low-velocity anomaly values in this region are not as low as those in the LVZs of the western Qilian Orogen, western Qinling Orogen, and Songpan-Ganzi Terrane, our results may support that this LVZ from mafic magmatic underplating (Wang et al. 2014).

5.4 Future work

5.4.1 Applying physical constraints on finer grids

Current PINN implementations enforce physical constraints only at traveltime grid nodes (station locations). However, wave propagation physics requires continuous validity across the domain. Should physical constraints be applied to a finer grid than the traveltime grid?

From a theoretical perspective, applying physical constraints to a finer grid offers potential advantages. When dealing with complex geological structures, a finer grid can capture more subtle geological variations. For example, in the presence of small-scale faults or velocity anomalies within the study area, applying physical constraints only at station locations might overlook the influence of these small-scale geological features on wave propagation. Applying physical constraints on a finer grid allows for a more accurate description of wave propagation paths and velocity variations in the medium, thereby improving the precision and resolution of tomographic results. For instance, in simulating wave propagation, a finer grid enables more precise calculations of wavefront shapes and propagation directions, resulting in velocity models that better reflect actual geological conditions.

However, in practical applications, applying physical constraints to a finer grid also presents several challenges. First, computational costs increase significantly. A finer grid means more nodes and computational units, requiring the NN to process larger datasets and more complex computational graphs during training. Each iteration involves calculating physical constraints and loss functions at more nodes, substantially prolonging training time and demanding significantly more computational resources. For large-scale seismic data processing, this increase in

computational cost may render the method impractical. Second, determining an appropriate fine grid scale is a challenge. Different geological regions and research objectives may require different grid scales. If the grid is too fine, it may introduce unnecessary complexity; if it is not fine enough, the advantages of a finer grid may not be fully realized. This requires an in-depth understanding of the geological characteristics of the study area and extensive experimentation to determine the optimal grid scale.

In summary, whether to apply physical constraints to a finer grid requires balancing potential advantages with practical challenges. Future research could explore compromise approaches, such as applying physical constraints on finer grids in key areas or regions with significant impact on results while maintaining the current grid scale in other areas to balance computational costs and imaging accuracy. Simultaneously, developing more efficient computational algorithms and optimization strategies to reduce computational costs and avoid overfitting will be key to making this approach feasible.

5.4.2 Helmholtz tomography using PINNs

The success of PINN-based surface wave tomography opens opportunities to incorporate multiple physical parameters in seismic tomography. While this study demonstrates the effectiveness of eikonal tomography, including additional constraints such as amplitude information through Helmholtz tomography could further improve resolution and accuracy. The eikonal equation provides a first-order approximation of wave propagation. However, it neglects amplitude information, which may introduce errors in the phase velocity measurement when the velocity anomalies are smaller than the wavelength or multi-path propagation of waves occurs (Li et al. 2023). The Helmholtz equation, a second order non-linear PDE, further correct the influence of amplitude on the basis of the eikonal equation (Lin and Ritzwoller 2011):

$$\frac{1}{c^2(\mathbf{x})} = |\nabla\tau(\mathbf{x})|^2 - \frac{\nabla^2 A(\mathbf{x})}{A(\mathbf{x})\omega^2}, \tag{5.1}$$

where $\tau(\mathbf{x})$ represents the traveltime from an effective source to a receiver at location $\mathbf{x} = (x, y)$, while $c(\mathbf{x})$ denotes the local phase velocity at the surface for the same receiver location, $A(\mathbf{x})$ represents the amplitude at position $\mathbf{x}$, ω is the angular frequency. By incorporating this correction, the Helmholtz equation improves the accuracy of phase velocity estimates especially

at longer periods.

Based on Equation 5.1, the physical term in the loss function is defined as:

$$L_H(N_\tau, N_A, N_c) = |\nabla N_\tau|^2 - \frac{1}{N_c^2} - \frac{\nabla^2 N_A}{N_A \omega^2}, \tag{5.2}$$

where L_H is the loss function of the Helmholtz constraint, $N_\tau = N_\tau(\mathbf{x}, \phi, \omega)$ and $N_A = N_A(\mathbf{x}, \phi, \omega)$ are the NN-based estimates of traveltime and amplitude at locations $\mathbf{x}$ and angular frequency ω with local backazimuth ϕ, respectively, N_c represents phase velocities evaluated at $\mathbf{x}$ for a specific frequency.

Three NNs are needed to model phase velocities, traveltime surfaces and amplitudes. The overall loss function used to train these three networks simultaneously has the form:

$$\begin{aligned}\mathcal{L}(\theta_\tau, \theta_A, \theta_c) = \sum_k^{N_{freq}} \sum_j^{N_{src}} \sum_i^{N_{rcv}} \Bigg[& \Big| N_\tau(\theta_\tau; \mathbf{x}_i, \phi_j, \omega_k) - \tau_{i,j,k} \Big|^2 \\ & + \epsilon_A \Big| N_A(\theta_A; \mathbf{x}_i, \phi_j, \omega_k) - A_{i,j,k} \Big|^2 \\ & + \epsilon_H \Big| L_H\Big(N_\tau(\theta_\tau; \mathbf{x}_i, \phi_j, \omega_k), N_A(\theta_A; \mathbf{x}_i, \phi_j, \omega_k), N_c(\theta_c; \mathbf{x}_i, \omega_k) \Big) \Big|^2 \Bigg],\end{aligned} \tag{5.3}$$

where N_{freq}, N_{src} and N_{rcv} are the maximum number of frequencies, sources and receivers, and i, j, k are the index of those value, respectively. $\tau_{i,j,k}$ and $A_{i,j,k}$ are phase travetime and amplitude at locations $\mathbf{x}_i = (x_i, y_i)$ and frequency ω_k from direction ϕ_j. ϵ_A and ϵ_H are weight factors for different data terms and physical term. The structural phase velocity from PINN-based Helmholtz tomography can be obtained by training the PINNs with appropriate hyperparameters to minimize Equation 5.3.

This multi-physics approach could be further extended to full waveform inversion, though this would require careful consideration of the increased complexity in both the physics and the optimization problem.

5.4.3 Physics constrained dispersion curve inversion

While this study achieved Rayleigh wave phase velocity imaging in the northeastern Tibetan Plateau, converting these phase velocities to depth-dependent S-wave velocity structure would provide more direct insights into the region's crustal and upper mantle structure. Since PINN-based approaches have demonstrated reliable performance in solving seismic inverse problems, developing physics-constrained dispersion curve inversion could effectively accomplish this goal.

Aki and Richards (2002) introduce a matrix method for analyzing plane waves in homogeneous media. If we define the stress-displacement vector for Rayleigh wave as $(r_1, r_2, R_3, R4)$, where r_1, r_2 are scalar displacement functions, $\sigma_{rr} = iR_3$ and $R_4 = \sigma_{rz}$. Under the assumptions of a traction-free boundary condition at the free surface and vanishing displacement as the bottom, the Rayleigh wave eigenproblem is defined as:

$$\begin{bmatrix} 0 & -\frac{\lambda}{\lambda+2\mu}\frac{d}{dz} & \frac{1}{\lambda+2\mu} & 0 \\ \frac{d}{dz} & 0 & 0 & -\frac{1}{\mu} \\ \rho\omega^2 & 0 & 0 & \frac{d}{dz} \\ 0 & -\rho\omega^2 - \frac{d}{dz}(\frac{4\mu(\lambda+\mu)}{\lambda+2\mu}\frac{d}{dz}) & -\frac{d}{dz}\frac{\lambda}{\lambda+2\mu} & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ R_3 \\ R_4 \end{bmatrix} = k \begin{bmatrix} r_1 \\ r_2 \\ R_3 \\ R_4 \end{bmatrix} \tag{5.4}$$

where z is the depth and k is the wavenumber. The elastic medium is characterized by its density ρ and Lame parameters λ and μ. $r_1 = r_2 = 0$ at the bottom and $\frac{d}{dz}\left((\lambda+2\mu)\frac{dr_2}{dz}\right) + \lambda\frac{dkr_1}{dz} = 0$ and $R_4 = 0$ at the free surface. This equation can be converted to a finite-dimensional generalized matrix eigenvalue problem at $N+1$ collocation points within each layer using a collocation method. Then the Equation (5.4) can be discretized as (Denolle et al. 2012):

$$\begin{bmatrix} 0 & -\frac{\lambda}{\lambda+2\mu}\mathbf{D} & \frac{1}{\lambda+2\mu} & 0 \\ \mathbf{D} & 0 & 0 & -\frac{1}{\mu} \\ \rho\omega^2 & 0 & 0 & \mathbf{D} \\ 0 & -\rho\omega^2 - \mathbf{D}(\frac{4\mu(\lambda+\mu)}{\lambda+2\mu}\mathbf{D}) & -\mathbf{D}\frac{\lambda}{\lambda+2\mu} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_0^R \\ . \\ . \\ . \\ \mathbf{u}_{4N+3}^R \end{bmatrix} = k \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{u}_0^R \\ . \\ . \\ . \\ \mathbf{u}_{4N+3}^R \end{bmatrix} \tag{5.5}$$

where $\mathbf{D}$ is a difference operator of size $(N+1)\times(N+1)$ used to approximate d/dz. $\mathbf{I}$ denotes the

identity matrix of identical size to $\mathbf{D}$. The vector $\mathbf{u}^R$ encompasses the components $\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}_3, \mathbf{R}_4$ sampled at the collocation points. λ, μ and ρ are diagonal matrices of the depth-dependent shear modulus and density, respectively, evaluated at the collocation points.

Equation 5.5 takes the form of $\mathbf{A}(\mathbf{m})\mathbf{v} = k\mathbf{I}\mathbf{v}$ at each frequency, where $\mathbf{m}$ contains medium properties λ, μ and ρ, $\mathbf{v}$ contains all Rayleigh-wave displacement and stress vectors. If observed phase velocities $\mathbf{c}_{obs}$ (replacing wave number k) are available for specified N_f frequencies, the cost function can be defined as:

$$C(\mathbf{c}, \mathbf{v}, \mathbf{m}) \quad = \quad \sum_{i}^{N_f} (c_i - c_{obs,i})^2 + \sum_{i}^{N_f} \left\| ([\mathbf{A}_i(\mathbf{m}) - c_i\mathbf{I}]\mathbf{v}_i) \right\|_2^2 \tag{5.6}$$

where $\mathbf{c} = \left[c_1, c_2, \cdots c_{N_f}\right]^T$ represents the phase velocities at different frequencies from 1 to N_f, and $\mathbf{m} = [\mu_1, \mu_2, \cdots, \mu_{N_l}]^T$ represents the medium properties of different layers from 1 to N_l in the layered medium. Given the observed phase velocities within the corresponding frequency range, the elastic parameters of the subsurface medium can be obtained by optimizing Equation 5.6, from which the shear wave velocity can be further calculated.

By extending physics-constrained optimization approach to dispersion curve inversion, the benefits of incorporating physical principles could be maintained while resolving vertical velocity variations. This would enable better understanding of crustal and upper mantle structure in complex tectonic regions, such as the northeastern Tibetan Plateau, while ensuring solutions remain physically plausible through explicit physical constraints.

5.5 Conclusions

In this thesis, I presented three surface wave tomography methods using physics-informed neural networks (PINNs) and demonstrated their feasibility through field data applications in the northeastern Tibet. Through the development of PINN-based eikonal tomography (pinnET), PINN-based elliptical-anisotropic eikonal tomography (pinnEAET), and PINN-based teleseismic eikonal tomography (pinnTET), this research shows a progression in handling various seismic tomography problems, including isotropic and anisotropic velocity structure modeling for both ambient noise and teleseismic surface waves. These works establish that PINNs can suc-

cessfully solve real-world seismic tomography problems while adhering to underlying physical principles.

The PINN-based approach offers several distinct advantages over traditional methods. Most importantly, it achieves comparable resolution using only approximately 20% or even less of conventionally required data, significantly improving data efficiency without compromising solution quality. The framework naturally incorporates physical constraints through the eikonal equation, providing physics-guided interpolation that improves results in regions with sparse or uneven data coverage. Furthermore, the ability to simultaneously process multiple frequencies through direct integration of period terms into the neural networks (NNs) ensures more coherent velocity structures across different depth ranges. However, implementing PINNs for seismic problems also presents important challenges. The methods require careful parameter tuning and network architecture design for each specific application, with performance heavily dependent on these choices. While data requirements are reduced, computational demands remain significant due to the global approximation nature of NNs and the complexity of balancing physical constraints with data fitting in the loss function. Additionally, while the networks can predict velocities at arbitrary positions, accuracy decreases notably when evaluation points fall outside the training data distribution.

Through extensive testing on field data from the northeastern Tibetan Plateau, this research provides practical guidelines for implementing PINN-based seismic tomography methods. The successful resolution of complex geological structures, validated through uncertainty analysis and comparison with conventional methods, demonstrates that these approaches can reliably image Earth structure when properly implemented. The revealed velocity structure and anisotropy patterns show good correlation with major tectonic features, such as the prominent low-velocity zones beneath the western Qilian Orogen and Songpan-Ganzi Terrane, which may indicate different deformation mechanisms. The presence of a high-velocity zone in the central Qinling Orogen suggests that the mid-to-lower crustal flow from the Songpan-Ganzi Terrane currently lacks a channel between the Ordos Block and Sichuan Basin. These findings contribute to our understanding of the crustal deformation and deep processes in the study area. The reduction in data requirements is particularly valuable for regions where dense data coverage is impractical or impossible to achieve. While meeting the primary objectives of the research, I also provide several promising avenues for future development. These include extension to

multi-physics problems like Helmholtz tomography to incorporate amplitude information and physics-constrained dispersion curve inversion for S-wave velocity structure inversion. The methodologies developed here provide a solid foundation for future developments in seismic tomography using physics-constrained approaches.

While PINN-based approaches may not replace traditional seismic methods, they represent a valuable addition to the seismologist's toolbox, particularly for scenarios with limited data availability. The ability to combine data-driven learning with physical principles offers promising directions for advancing not just surface wave tomography, but potentially other geophysical applications as well. Future work should focus on addressing the identified limitations, particularly in computational efficiency and automated parameter selection, to make these methods more accessible to the broader geophysical community.

References

Aki, Keiiti and Paul G Richards (2002). *Quantitative seismology*.

Bao, Xuewei, Xiaodong Song, Mingjie Xu, Liangshu Wang, Xiaoxiao Sun, Ning Mi, Dayong Yu, and Hua Li (2013). "Crust and upper mantle structure of the North China Craton and the NE Tibetan Plateau and its tectonic implications". In: *Earth and Planetary Science Letters* 369, pp. 129–137.

Chen, Yunpeng, Sjoerd AL de Ridder, Sebastian Rost, Zhen Guo, Xiaoyang Wu, and Yongshun Chen (2022). "Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the northeastern margin of the Tibetan Plateau". In: *Geophysical Research Letters* 49.21, e2022GL099053.

Chen, Yunpeng, Sjoerd AL de Ridder, Sebastian Rost, Zhen Guo, Xiaoyang Wu, Shilin Li, and Yongshun Chen (2023). "Physics-Informed Neural Networks for Elliptical-Anisotropy Eikonal Tomography: Application to Data From the Northeastern Tibetan Plateau". In: *Journal of Geophysical Research: Solid Earth* 128.12, e2023JB027378.

Clark, Marin Kristen, MA House, LH Royden, KX Whipple, BC Burchfiel, X Zhang, and W Tang (2005). "Late Cenozoic uplift of southeastern Tibet". In: *Geology* 33.6, pp. 525–528.

Cuomo, Salvatore, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli (2022). "Scientific machine learning through physics–informed neural networks: Where we are and what's next". In: *Journal of Scientific Computing* 92.3, p. 88.

Denolle, Marine A, Eric M Dunham, and Gregory C Beroza (2012). "Solving the surface-wave eigenproblem with Chebyshev spectral collocation". In: *Bulletin of the Seismological Society of America* 102.3, pp. 1214–1223.

Ge, Can, Yong Zheng, and Xiong Xiong (2011). "Study of crustal thickness and Poisson ratio of the North China Craton". In: *AGU Fall Meeting Abstracts*. Vol. 2011, S11B–2207.

Hao, Shijie, Zhouchuan Huang, Cunrui Han, Liangshu Wang, Mingjie Xu, Ning Mi, and Dayong Yu (2021). "Layered crustal azimuthal anisotropy beneath the northeastern Tibetan Plateau

revealed by Rayleigh-wave Eikonal tomography". In: *Earth and Planetary Science Letters* 563, p. 116891.

Jiang, Chengxin, Yingjie Yang, and Yong Zheng (2014). "Penetration of mid-crustal low velocity zone across the Kunlun Fault in the NE Tibetan Plateau revealed by ambient noise tomography". In: *Earth and Planetary Science Letters* 406, pp. 81–92.

Li, Hongyi, Yang Shen, Zhongxian Huang, Xinfu Li, Meng Gong, Danian Shi, Eric Sandvol, and Aibing Li (2014). "The distribution of the mid-to-lower crustal low-velocity zone beneath the northeastern Tibetan Plateau revealed from ambient noise tomography". In: *Journal of Geophysical Research: Solid Earth* 119.3, pp. 1954–1970.

Li, Lun, Chen Cai, Yuanyuan Fu, and Hongjian Fang (2023). "Multiple surface wave tomography methods and their applications to the Tibetan Plateau". In: *Reviews of Geophysics and Planetary Physics* 54.2, pp. 174–196.

Li, Shilin, Zhen Guo, Yong Yu, Xiaoyang Wu, and Yongshun John Chen (2022). "Imaging the Northeastern crustal boundary of the Tibetan Plateau with radial anisotropy". In: *Geophysical Research Letters* 49.23, e2022GL100672.

Li, Yonghua, Jiatie Pan, Qingju Wu, and Zhifeng Ding (2017). "Lithospheric structure beneath the northeastern Tibetan Plateau and the western Sino-Korea Craton revealed by Rayleigh wave tomography". In: *Geophysical Journal International* 210.2, pp. 570–584.

Lin, Fan-Chi and Michael H Ritzwoller (2011). "Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure". In: *Geophysical Journal International* 186.3, pp. 1104–1120.

Sun, Xinlei, Xiaodong Song, Sihua Zheng, Yingjie Yang, and Michael H Ritzwoller (2010). "Three dimensional shear wave velocity structure of the crust and upper mantle beneath China from ambient noise surface wave tomography". In: *Earthquake Science* 23, pp. 449–463.

Wang, Chun-Yong, E Sandvol, L Zhu, Hai Lou, Zhixiang Yao, and Xinghua Luo (2014). "Lateral variation of crustal structure in the Ordos block and surrounding regions, North China, and its tectonic implications". In: *Earth and Planetary Science Letters* 387, pp. 198–211.

Wang, Kaiming, Laiyu Lu, Valérie Maupin, Zhifeng Ding, Chen Zheng, and Shijun Zhong (2020). "Surface wave tomography of northeastern Tibetan Plateau using beamforming of seismic noise at a dense array". In: *Journal of Geophysical Research: Solid Earth* 125.4, e2019JB018416.

Wang, Xingchen, Yonghua Li, Zhifeng Ding, Lupei Zhu, Chunyong Wang, Xuewei Bao, and Yan Wu (2017). "Three-dimensional lithospheric S wave velocity model of the NE Tibetan Plateau and western North China Craton". In: *Journal of Geophysical Research: Solid Earth* 122.8, pp. 6703–6720.

Wu, Xiaoyang, Zhen Guo, Shilin Li, Yong Yu, Qipeng Bai, and Yongshun John Chen (2023). "Seismic azimuthal anisotropy of northeastern Tibetan Plateau from ambient noise double beamforming tomography: Implications for crustal deformation". In: *Journal of Geophysical Research: Solid Earth* 128.6, e2022JB026109.

Zhong, Shi-Jun, Jian-Ping Wu, Li-Hua Fand, Wei-Lai Wang, Li-Ping Fan, and Huai-Fu Wang (2017). "Surface wave Eikonal tomography in and around the northeastern margin of the Tibetan plateau". In: *Chinese Journal of Geophysics* 60.6, pp. 2304–2314.

Appendix A

Supplementary material for Chapter 2

Contents of this appendix

1. Figure A.1

A.1 Introduction

This supporting information contains the figure with the loss curves for the different training processes. The loss terms for the traveltime data and PDE constraint (the two terms in Equation 2.4) run for a total of 4000 epochs (2000 in each outer iteration), while the loss function for accelerated updating of the velocity neural network (Equation 2.6) ran for 2000 epochs (this occurs at the start of the second outer iteration). Having retrained the velocity neural network, the PDE loss term drops steeply.

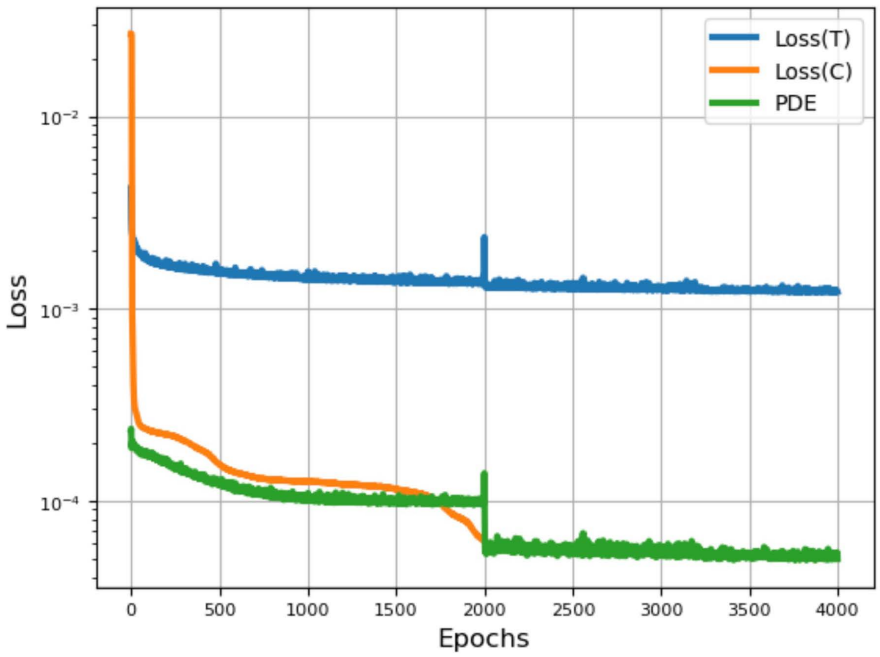


Figure A.1: Evolution of the different loss terms for the final training.

Appendix B

Supplementary material for Chapter 3

Contents of this appendix

1. Figures B.1 to B.3

B.1 Introduction

The first part in Appendix B is the derivation of the eigenvalues of matrix $\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$.

The eigenvalues λ of $\mathbf{M}$ satisfy the characteristic equation:

$$\det(\mathbf{M} - \lambda\mathbf{I}) = 0, \tag{B.1}$$

where $\mathbf{I}$ is the identity matrix. Expanding the determinant:

$$(M_{11} - \lambda)(M_{22} - \lambda) - M_{12}M_{21} = 0, \tag{B.2}$$

Since $M_{12} = M_{21}$, this simplifies to:

$$\lambda^2 - (M_{11} + M_{22})\lambda + (M_{11}M_{22} - M_{12}^2) = 0, \tag{B.3}$$

The quadric equation in λ has solutions:

$$\lambda = \frac{M_{11} + M_{22} \pm \sqrt{(M_{11} + M_{22})^2 - 4(M_{11}M_{22} - M_{12}^2)}}{2}, \tag{B.4}$$

According to the matrix elements in Equations 3.3 to 3.5, the phase velocities c_f and c_s are the square roots of the eigenvalues λ. Thus, the eigenvalues of $\mathbf{M}$ are c_f^2 and c_s^2. For each eigenvalue λ, solve $(\mathbf{M} - \lambda\mathbf{I})\mathbf{v} = 0$ (e.g. fast wave $\lambda = c_f^2$):

$$\begin{bmatrix} M_{11} - c_f^2 & M_{12} \\ M_{12} & M_{22} - c_f^2 \end{bmatrix} \begin{bmatrix} v_{f1} \\ v_{f2} \end{bmatrix} = 0 \tag{B.5}$$

Looking at the first row and rearranging to solve for the ratio of components:

$$\frac{v_{f2}}{v_{f1}} = -\frac{M_{11} - c_f^2}{M_{12}} = -\tan\alpha, \tag{B.6}$$

Thus, the fast eigenvector is parallel to the anisotropy symmetry axis α. Similarly:

$$\frac{v_{s2}}{v_{s1}} = -\frac{M_{11} - c_s^2}{M_{12}} = \cot\alpha. \tag{B.7}$$

The slow eigenvector is orthogonal to $\mathbf{v}_f$. Therefore, the the eigenvalues and eigenvectors of $\mathbf{M}$ indicates the fast, slow and the principle directions of the anisotropy.

This supporting information also contains 3 figures. Figure B.1 shows how the 25 s Rayleigh wave phase velocity model varies with the number of source stations. Models were generated using 1, 5, 10, 20, 30, 40 and 50 effective source stations among the total 676 stations. When fewer than 20 sources are used, the average phase velocity changes more sharply as the number of sources increases. This phenomenon illustrates non-unique solutions and the resolution limitations of the 25 s Rayleigh wave phase velocity for sparse input data. Though computationally cheaper, models with less than 20 sources poorly approximated phase velocities and lacked detail. While the source number increases from 20 to 50 sources, the average phase velocity stabilizes around 3.46 km/s. The change in mean absolute error is more intuitive with the error dropping rapidly from about 0.09 at 1 sources to below 0.02 for 20 sources, then remaining below 0.02 from 20 to 50 sources. Overall, the resolution and accuracy of the trained phase velocity model gain rapidly with increasing source number and tend to level off after 20 sources. We select 20 stations as effective sources in this study to incorporate sufficient sampling and limit computational expense.

Figure B.2 compares the azimuthal anisotropic velocities obtained by pinnEAET and azimuthal double beamforming tomography (Wu et al. 2023), respectively. The phase velocities correspond closely across most regions of the study area, diverging primarily in small areas at the northeast and south boundaries and within parts of the Central Asian Orogenic Belt. However, our results are consistent with traditional eikonal tomography (Figure 3.8 (c)) in those areas. Centrally, azimuthal anisotropy orientations show close agreement between the two models. Discrepancies near boundaries likely arise from lacking ray paths in certain directions at the boundary. The fast direction derived from the conventional eikonal tomography and azimuthal double beamforming tomography also differ on the boundary (Figure 3.8 (b) and Figure B.2 (b)). Overall, pinnEAET azimuthal anisotropic phase velocities match azimuthal double beamforming results well where data coverage is sufficient but diverge near the boundaries, where limited azimuthal coverage likely degrades model resolution and accuracy in both studies.

Figure B.3 shows the results of Rayleigh wave phase velocity and azimuthal anisotropy over periods of 10, 20, 30, and 40 s in northeastern Tibet obtained from simultaneously training with the same traveltime NN and a single medium property NN. Employing a single network is estimated to save approximately 20 % of the training time to achieve a similar convergence level compared to the use of three networks, while maintaining consistent training conditions in all other aspects. In the map of inverted phase velocities (Figure B.3), we observe a noticeable high velocity anomaly at the southeast corner of the array across all periods, which is not present in the results using three networks for the medium properties (Figure 3.6). Additionally, as shown in Figure B.3 (b)-(d), the phase velocities exhibit a more extensive area of significant low velocity anomalies along the southwestern boundary in the Western Qinling Orogen compared to the results in Figure 3.6 (b)-(d). These differences may indicate that a single network is more influenced by data coverage. Regarding azimuthal anisotropy, the results from both network architectures are highly similar. Minor differences can be observed within small areas for certain periods. For example, comparing the azimuthal anisotropy in Figure 3.6 (d) and Figure B.3 (d), anisotropic results using three networks illustrate a distinct NW-SE fast direction at 40 s at the northeastern corner in the study area beneath the Hetao Graben, which is absent in the results from a single network. Overall, using a single medium property NN can improve training efficiency, but it also introduces some unexpected artefacts possibly due to unintentionally introduced cross-dependencies. Further optimization of the methodology is necessary to strike a balance between efficiency and accuracy.

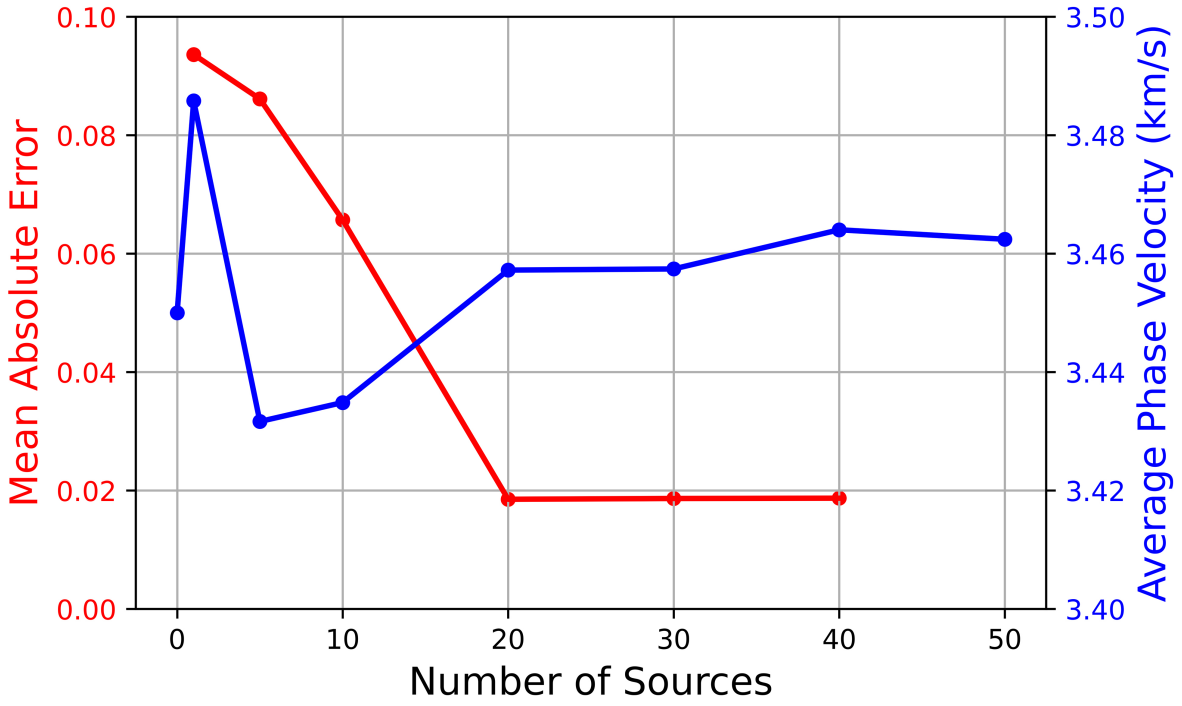


Figure B.1: Variation of the 25 s Rayleigh wave phase velocity with 1 to 50 source stations. Blue stars and line indicate the variation of average phase velocity, starting at a background velocity of 3.45 km/s; red stars and line indicate the variation of mean absolute error calculated between the phase velocity models generated from input data for each two adjacent source numbers.

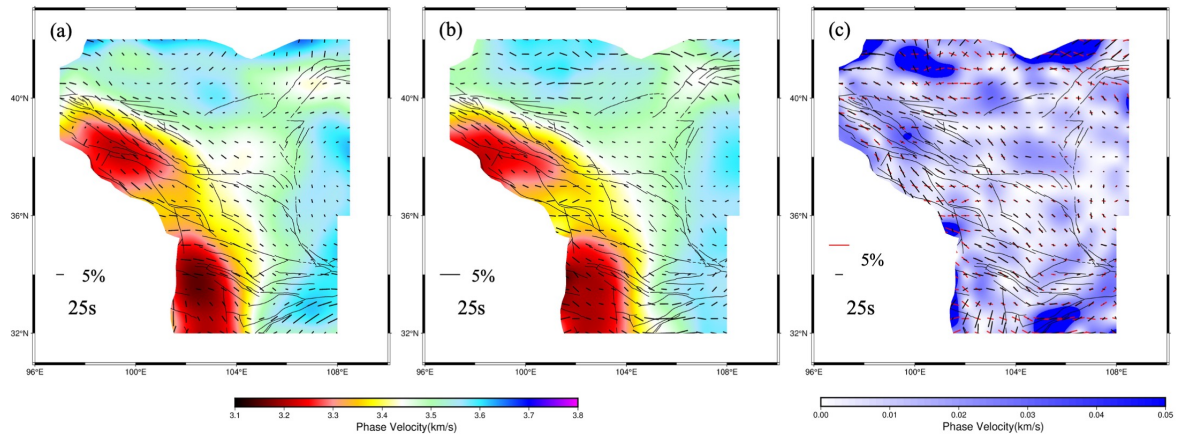


Figure B.2: (a) The 25 s azimuthal anisotropic phase velocity beneath the northeastern Tibetan Plateau using pinnEAET; (b) The 25 s azimuthal anisotropic phase velocity generated by azimuthal double beamforming tomography (Wu et al. 2023); (c) The difference between 25 s azimuthal anisotropic phase velocity (background) from pinnEAET (black lines) and azimuthal double beamforming tomography (red lines).

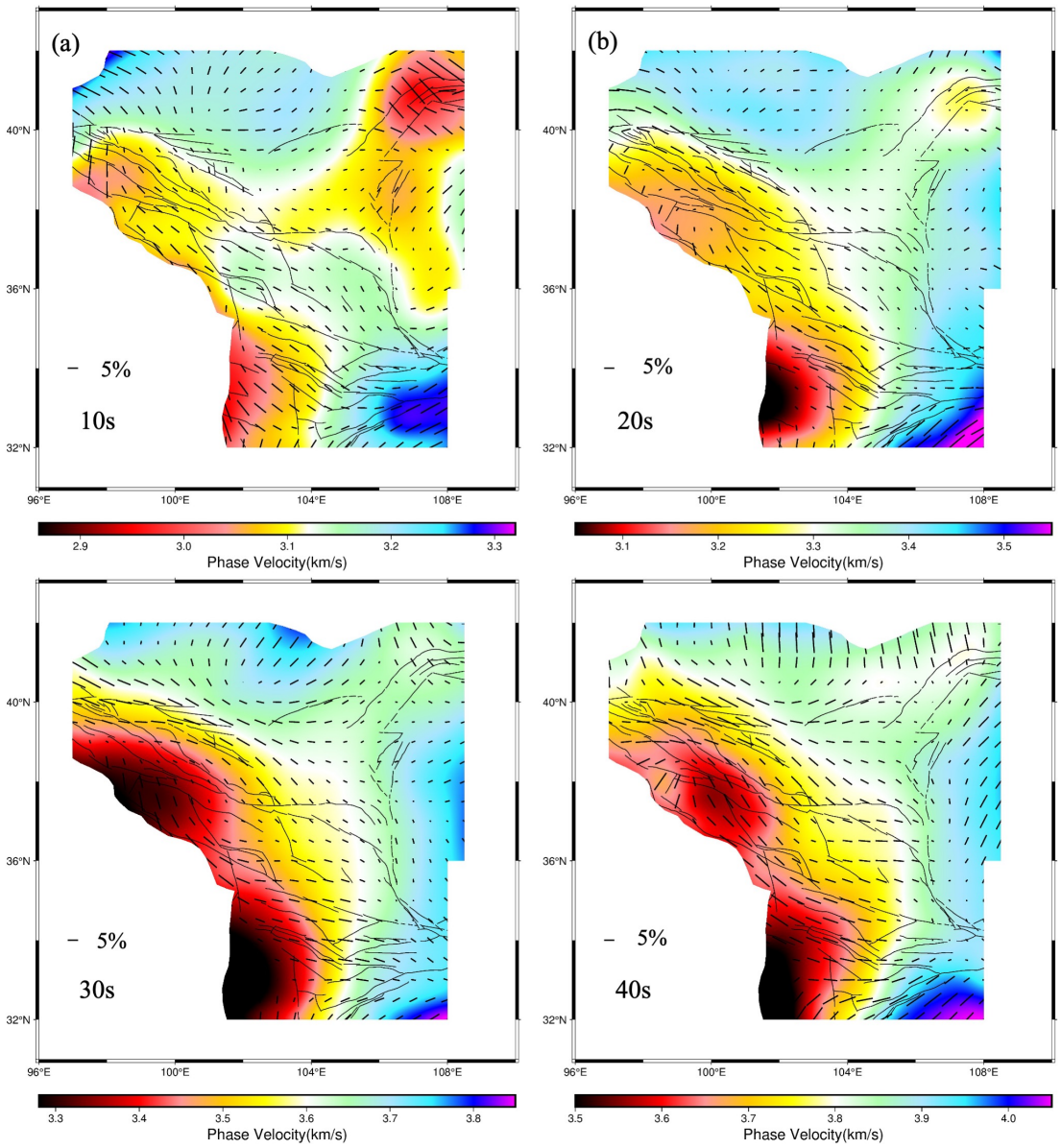


Figure B.3: (a)-(d) Azimuthal anisotropic phase velocity at periods of 10, 20, 30 and 40 s beneath northeastern Tibetan Plateau using pinnEAET with a single medium property NN. Black bold vectors indicate the strength and fast propagation direction of anisotropy, black thin lines indicate main faults.

Appendix C

Supplementary material for Chapter 4

Contents of this a

1. Figures C.1 to C.4

C.1 Introduction

This supporting information contains 4 figures.

Figure C.1 shows the number and distribution of earthquakes selected in this study at different periods of 20, 25, 30, 40, 60, and 80 s, with 11, 15, 30, 25, 25, and 25 seismic events used, respectively. The event selection criteria required that at least 400 stations recorded each seismic event, and the events should have a relatively uniform backazimuth distribution. From Figure C.1, we can see that the distribution of seismic events at 20 s and 25 s is relatively sparse, with 20 s seismic events showing gaps in the clockwise range from southwest to northeast, specifically between backazimuth 200 to 360 degrees and 0 to 40 degrees, indicating insufficient teleseismic data at shorter periods. The seismic events at other periods show a more uniform distribution.

Figure C.2 shows the map of Rayleigh wave phase velocity uncertainties for periods from 20 to 80 s. Overall, the velocity uncertainty shows a clear increasing trend with increasing period. Between 20 and 30 s, the velocity uncertainties are generally less than 0.04 km/s and most value concentrate in the lower uncertainty range. At 40 s, the velocity uncertainties mainly concentrate between 0.04-0.08 km/s. For 60 s and 80 s, the velocity uncertainties primarily fall between 0.06-0.12 km/s, showing a significantly wider distribution range of uncertainties. This gradual broadening of the distribution range indicates larger velocity measurement uncertainties at longer periods, suggesting a decrease in measurement precision as the period increases.

Figure C.3 compares our results with those from traditional eikonal tomography at periods of 25, 30, 40, and, 60 s, as a supplement to Figure 4.9. In these periods, the difference in Rayleigh wave phase velocity between both methods is generally within 0.05 km/s, showing consistency in the recovery of velocity distribution. At 25 s, the velocity difference is located in the southwestern margin of the study area between 32° N and 36° N. Our results show LVZs in western Qilian Orogen and western Qinlin Orogen, Songpan-Ganzi Terrane are separated by a relative HVZ. At 30 and 40 s, both methods show consistent results. From 30 to 40 s, the low-velocity anomalies in western Qilian Orogen, western Qinlin Orogen, and Songpan-Ganzi Terrane increase, while the high-velocity anomalies in southern central Asian Orogenic Belt and northern Alxa Block decrease. The relative variation of high-velocities in Ordos Block and central Qinling Orogen also decreases. At 60 s, differences exceeding 0.05 km/s are observed in the southwestern boundary

of the the study area between 33° N and 37° N, as well as between 37° N and 40° N. The traditional eikonal tomography results show a continuous relative LVZ along the southwestern boundary and low velocities only appear in Songpan-Ganzi Terrane, while our results show two LVZs in western Qilian Orogen and Songpan-Ganzi Terrane. Another velocity difference exists as a isolated roughly circular region between 39° N-42° N and 98° E-101° E, which may be related to poor ray path coverages.

Figure C.4 shows the difference between phase velocity structures obtained using PINN-based teleseismic surface wave tomography and ambient noise tomography, as a supplement to Figure 4.10. The difference in velocity results obtained from PINN-based eikonal tomography using the two datasets is mainly reflected in the strength of low-velocity anomalies. This difference gradually decreases with increasing period. At 20 s, the ambient noise tomography results show lower velocity values in western Qilian Orogen, western Qinling Orogen, Songpan-Ganzi Terrane, and Hetao Basin. At 30 s, these differences narrow to regions in western Qilian Orogen, Songpan-Ganzi Terrane, and part of western Qinling Orogen. At 40 s, the differences essentially remain only in western Qilian Orogen. This indicates that different types of data enable varying constraints on phase velocity results at different periods.

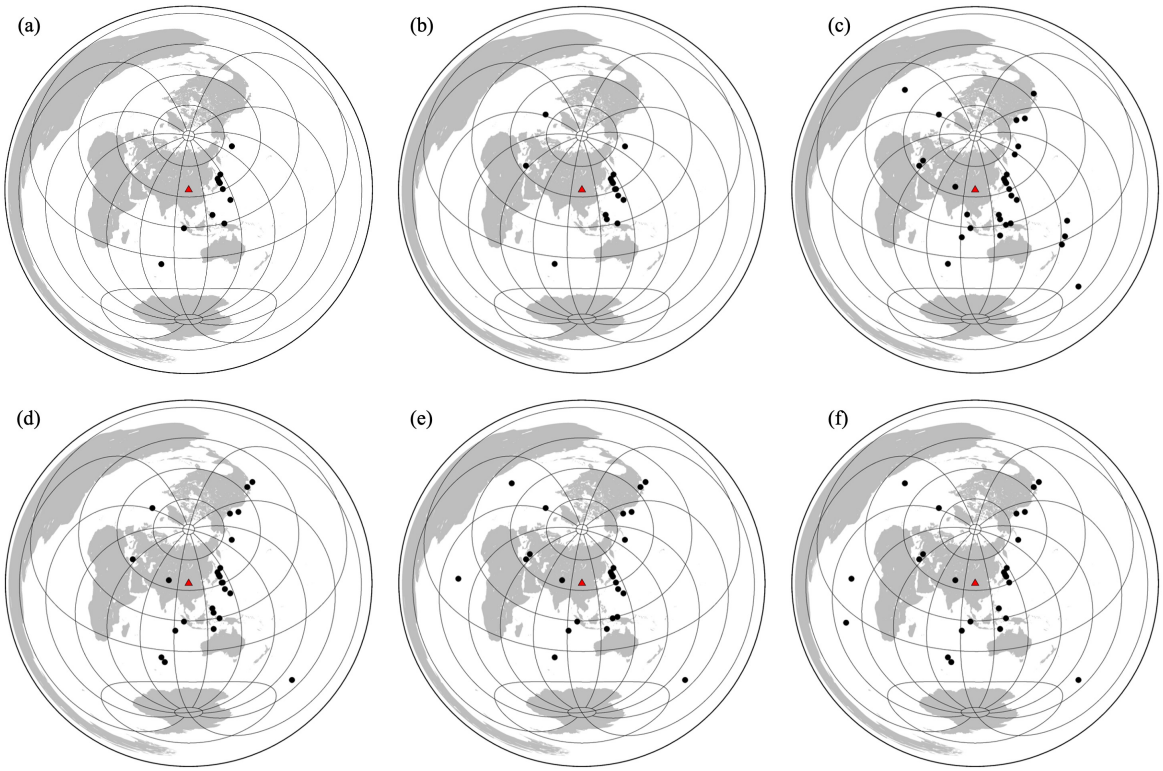


Figure C.1: Distribution of the earthquakes utilized in the analysis for periods across 20, 25, 30, 40, 60, and 80 s. Black dots indicate teleseismic events, red triangle indicates the center of the seismic array.

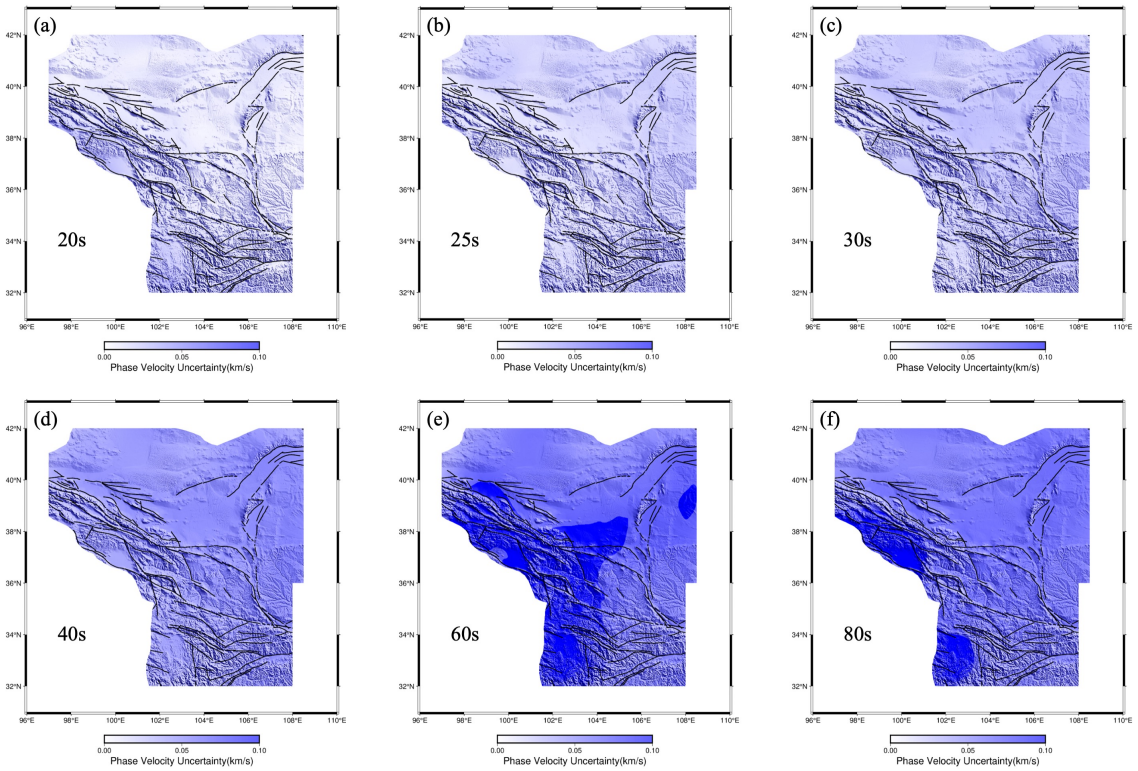


Figure C.2: (a-d) Uncertainty map of the Rayleigh wave phase velocity at periods of 20, 25, 30, 40, 60, and 80 s beneath northeastern Tibetan Plateau using pinnTET.

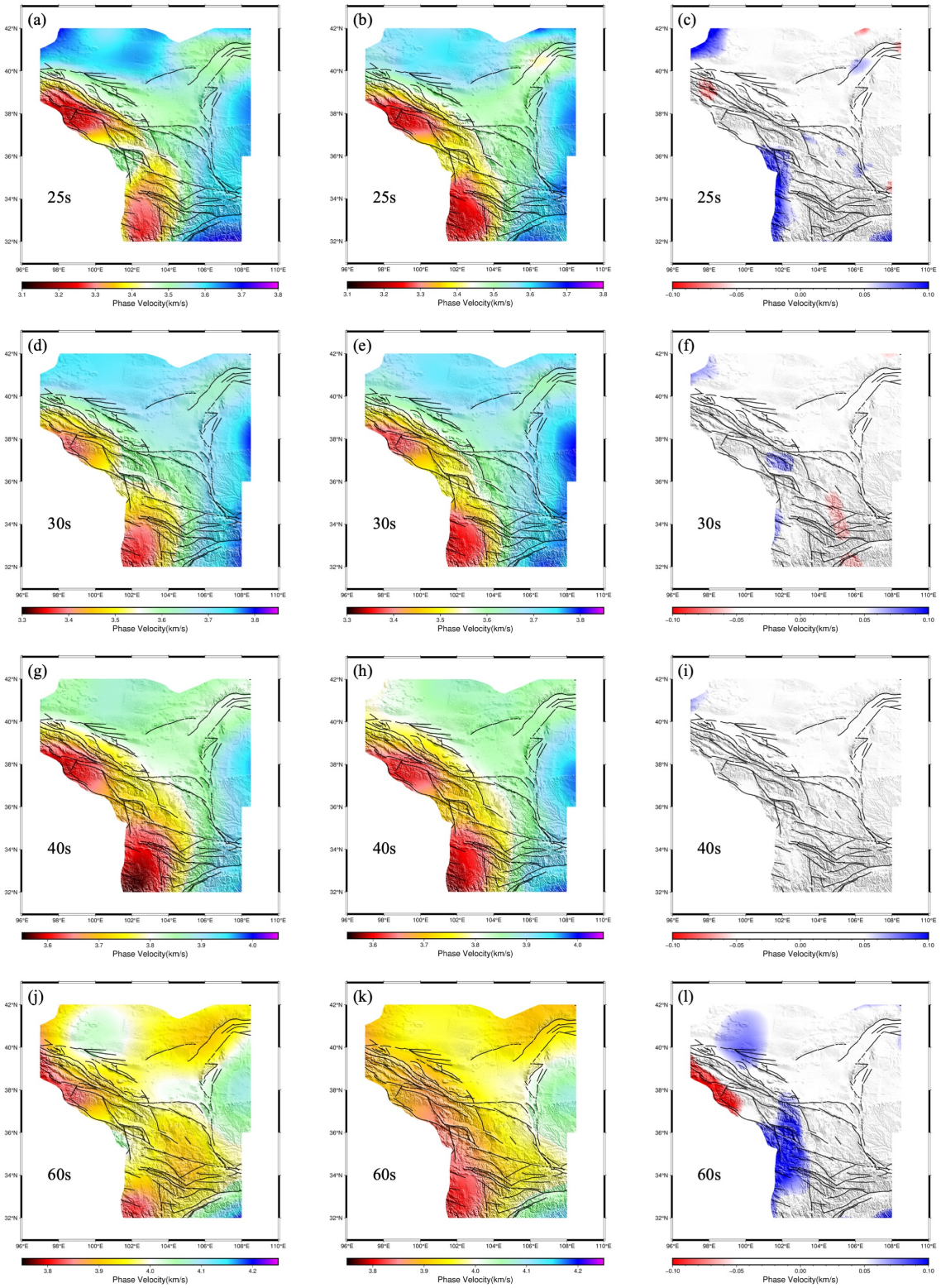


Figure C.3: (a-j) Rayleigh wave phase velocity at periods of 25, 30, 40, and 60 s beneath northeastern Tibetan Plateau obtained from pinnTET. (b-k) Rayleigh wave phase velocity at periods of 25, 30, 40, and 60 s beneath northeastern Tibetan Plateau obtained from conventional eikonal tomography. (c-i) Difference of Rayleigh wave phase velocity between two methods at periods of 25, 30, 40, and 60 s.

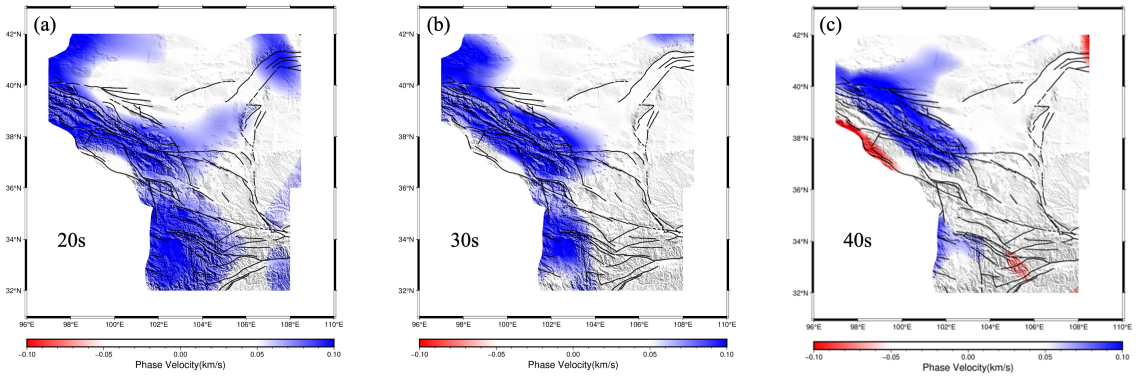


Figure C.4: (a-c) Difference of Rayleigh wave phase velocity beneath northeastern Tibetan Plateau at periods of 20, 30, and 40 s between PINN-based teleseimic surface wave tomography and ambient noise tomography.