

UNIVERSITY OF SÃO PAULO - BRAZIL
INSTITUTE OF ASTRONOMY, GEOPHYSICS, AND
ATMOSPHERIC SCIENCES

Caio Ciardelli

Adjoint Tomography of South America based on 3D
Spectral-Element Seismic Wave Simulations

Tomografia Adjunta da América do Sul baseada em
Simulações 3D de Ondas Sísmicas por Elementos Espectrais

Ph.D. Dissertation
Department of Geophysics

June 14, 2021

CAIO CIARDELLI

Adjoint Tomography of South America based on 3D Spectral-Element
Seismic Wave Simulations

Thesis presented to the Department of
Geophysics at the University of São
Paulo to obtain the title of Doctor in
Seismology. Advisor: Prof. Marcelo
Assumpção. Co-Advisors: Ebru Boz-
dağ and Suzan van der Lee

June 14, 2021

To my beloved country.

»... The goal of seismology is to investigate the Earth's interior and to proceed at the point where the geologist stops, having in modern seismographs a sort of binoculars with which the greatest depths may be observed ...«¹

Andrija Mohorovičić

¹Herak (2010)

Acknowledgements

I thank God for giving me the best opportunities that anyone could have, in addition to all the help, strength, and determination necessary to accomplish my duty, and I thank my parents for their love, encouragement, and unconditional support throughout my life.

I thank Professors Marcelo Assumpção, Suzan van der Lee, Ebru Bozdağ, Carlos Moreno Chaves, and Marcelo Bianchi not only for their assistance in the execution of my doctoral project but also for all the encouragement, guidance, and motivation throughout these years. I also thank all my friends, relatives, professors, and everyone else who somehow contributed to my formation, including this university and its faculty, management, and administration, without which none of this would be possible.

I thank FAPESP (Grants: 2013/24215-6, 2016/03120-5, 2018/04918-6, and 2018/04917-0) for providing financial support for this research and María Laura Rosa, from the National University of La Plata, for providing the waveforms and receiver response files for station LPA.

This research was supported in part through the computational resources and staff contributions provided for the Quest high-performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

Contents

Acknowledgements	v
Abstract	viii
Resumo	x
1 Introduction	1
2 The Wave Equation in Elastodynamics	4
3 Spectral-Element Method	7
3.1 Spectral-Element Modeling	7
3.1.1 Weak Form of the Wave Equation	8
3.1.2 Partition of the Medium into Elements	9
3.1.3 Lagrange Polynomials	9
3.1.4 Assembly and Solution of the System of Equations	13
4 Geophysical Inversion	20
4.1 Inverting Seismic Velocity	20
4.1.1 Wave Equation using Operators	20
4.1.2 Local Minima	21
4.1.3 Computation of the Gradient	25
4.1.4 Steepest Descent Method	26
4.1.5 Newton's Method	27
4.1.6 Conjugated Gradient Method	29
5 Adjoint-state Method	31
5.1 Adjoint Method in Seismic Tomography	33
6 Sensitivity Kernels	37

7	Previous Studies of South America	68
7.1	Amazonian Craton	71
7.2	São Francisco Craton	72
7.3	Paraná Basin	73
7.3.1	Cratonic Nucleus	73
7.3.2	"Paraná Plume"	73
7.3.3	Rio de la Plata Craton	74
7.4	Parnaíba Basin	74
7.5	Andes	74
8	Adjoint Tomography	79
9	Model Visualization and Sharing	132
10	Conclusion	152
A	Book Cover	162

Abstract

Adjoint tomography, a full-waveform inversion technique based on 3D wave simulations, is now commonly used in earthquake seismology, drawing on advances in computational power and numerical methods. In this study, we use 3D spectral element continental-scale seismic wave simulations (Komatitsch and Tromp, 2002a,b) and 112 earthquakes recorded by 1311 seismic stations to construct an adjoint waveform tomography model of South America. The thesis begins with a review of the wave equation in elastodynamics followed by an introductory explanation of the spectral-element method (Schubert, 2003; Igel, 2017). We also revisit the inversion problem in geophysics and the adjoint-state method (Plessix, 2006) in an intuitive way. We proceed with a simplified explanation of the finite-frequency theory (Dahlen et al., 2000) and a review of previous tomographic studies in South America. To carry out our adjoint tomography, we detect and remove noisy & problematic data using our multi-stage algorithm before the time-window selection, reducing the likelihood of discarding useful data or assimilating bad-quality waveforms in inversions. Our misfit function is a complex-exponentiated instantaneous phase (Yuan et al., 2020), which optimizes the information extracted from each time series without the need for short-time windows. We performed 23 iterations, gradually increasing the frequency content of the data to prevent local minima from hampering the convergence. Our final model (SAAM23, **S**outh **A**merican **A**djoint **M**odel, iteration **23**) shows a $\sim 50\%$ decrease in the misfit. We further assessed the improvement by using cross-correlation measurements using 53 earthquakes that were not included in the adjoint inversion. In the long wavelengths, the model is compatible with previous studies, such as Van der Lee et al. (2001), Feng et al. (2007), Celli et al. (2020), and Lei et al. (2020). The Nazca Slab is well imaged and is shown to be continuous in the 300-500 km depth following the Peruvian flat-slab segment. Beneath northern South America, the slab crosses the mantle transition zone and plunges into the lower mantle. In the central and southern part of South America, the slab appear to flatten near the 650 km discontinuity, before plunging into the lower mantle. In the stable platform, both exposed cratons (Amazonian and São Francisco), as well as covered cratonic blocks (Paranapanema and Parnaíba, beneath the intracratonic

Paraná and Parnaíba Basins, respectively), show high velocities at lithospheric depths. The seismic lithosphere/asthenosphere boundary (LAB) was measured at the depth with the steepest negative velocity gradient. Good agreement was found between this seismic LAB and the values obtained by *S*-wave receiver functions. In the Amazonian Craton, both positive lithospheric *S*-wave velocity anomalies, as well as LAB depth, increase with the average age of the geochronological provinces. On the other hand, no high-velocity anomalies were found beneath the Rio de La Plata Craton. The thesis ends with the presentation of SphGLLTools, an open-source toolbox we designed to allow easy and practical visualization of tomographic models defined on spectral-element meshes using either direct interpolation or the flexible expansion using spherical harmonics while taking advantage of GMT6 (Wessel et al., 2019) to create high-quality images. We also lead the reader through a comprehensible yet intuitive explanation of the theory and concepts used by the routines.

Keywords: Wave equation. Spectral elements. Numerical optimization. Adjoint method. Finite-frequency theory. Full-waveform inversion. Visualization. B-splines. Spherical harmonics.

Resumo

A tomografia adjunta, uma técnica de inversão de forma de onda completa baseada em simulações de ondas 3D, agora é comumente usada em sismologia graças aos avanços no poder computacional e nos métodos numéricos. Neste estudo, usamos simulações de onda sísmicas com o método dos elementos espectrais 3D em escala continental (Komatitsch and Tromp, 2002a,b) e 112 terremotos registrados por 1311 estações sísmográficas para construir um modelo de tomografia de forma de onda da América do Sul. A tese começa com uma revisão da equação de onda em elastodinâmica seguida por uma explicação introdutória do método dos elementos espectrais (Schubert, 2003; Igel, 2017). Também revisitamos o problema da inversão em geofísica e o método adjunto (Plessix, 2006) de forma intuitiva. Prosseguimos com uma explicação simplificada da teoria da frequência finita (Dahlen et al., 2000) e uma revisão de estudos tomográficos anteriores na América do Sul. Para realizar nossa tomografia, detectamos e removemos dados ruidosos e problemáticos usando nosso algoritmo de múltiplos estágios antes da seleção das janelas temporais, reduzindo a probabilidade de descartar dados úteis ou assimilar formas de onda de má qualidade em inversões. Nossa função objetivo usa a exponencial complexa da fase instantânea (Yuan et al., 2020), que otimiza as informações extraídas de cada série temporal sem a necessidade de janelas de curtas. Realizamos 23 iterações, aumentando gradualmente o conteúdo da frequência dos dados para evitar que mínimos locais atrapalhassem a convergência. Nosso modelo final (**SAAM23, South American Adjoint Model, iteration 23**) mostra uma redução de $\sim 50\%$ no resíduo total. Também mensuramos a melhora através de correlação cruzada usando 53 terremotos que não foram incluídos na inversão. Nos longos comprimentos de onda, o modelo é compatível com estudos anteriores, como Van der Lee et al. (2001), Feng et al. (2007), Celli et al. (2020) e Lei et al. (2020). A Placa de Nazca é bem imageada e aparece contínua nas profundidades de 300-500 km seguindo o segmento de placa horizontalizada sob o Peru. Abaixo da região norte da América do Sul, a placa cruza a zona de transição e mergulha no manto inferior. Na parte central e sul da América do Sul, a placa se horizontaliza perto da descontinuidade de 650 km, antes de mergulhar no manto inferior. Na plataforma estável, tanto os cratons expostos (Amazônico

e São Francisco), quanto os blocos cratônicos cobertos (Paranapanema e Parnaíba, sob as bacias intracratônicas do Paraná e Parnaíba, respectivamente), apresentam altas velocidades em profundidades litosféricas. A descontinuidade sísmica que separa a litosfera da astenosfera (DLA) foi estimada através da profundidade do gradiente de velocidade negativo mais acentuado. Uma boa concordância foi encontrada entre a DLA sísmica e os valores obtidos pelas funções do receptor de onda S . No Cráton Amazônico, tanto as anomalias positivas da velocidade da onda S na litosfera, quanto a profundidade da DLA, aumentam com a idade média das províncias geocronológicas. Por outro lado, nenhuma anomalia de alta velocidade foi encontrada abaixo do Cráton do Rio de La Plata. A tese termina com a apresentação do SphGLLTools, um conjunto de rotinas de código aberto que projetamos para permitir a visualização fácil e prática de modelos tomográficos definidos em malhas de elementos espectrais usando interpolação direta ou uma expansão versátil em harmônicos esféricos, usando o GMT6 (Wessel et al., 2019) para criar imagens de alta qualidade. Também conduzimos o leitor através de uma explicação completa, porém intuitiva, da teoria e dos conceitos usados pelas rotinas.

Palavras-chave: Equação de onda. Elementos espectrais. Otimização numérica. Método adjunto. Teoria da frequência finita. Inversão de forma de onda completa. Visualização. B-splines. Harmônicos esféricos.

Chapter 1

Introduction

Whenever a seismic event, such as an earthquake, an explosion, or an asteroid impact occurs, seismic waves propagate in all directions through the Earth's interior, traveling at thousands of kilometers per hour, until they eventually emerge somewhere on the surface of the planet. Seismographic stations can record the ground displacements in the form of seismograms, which contain valuable information both of the seismic source and the Earth's interior. Seismologists can use that information to image the Earth's inner structure, a technique known as *seismic tomography*.

Keith Bullen, in collaboration with Harold Jeffreys, prepared travel time tables for the propagation of seismic waves within the Earth. The Jeffreys-Bullen tables were published in 1940 (Jeffreys and E., 1940), later than the discovery of major inner structures of the Earth, such as the core (Oldham, 1906), the Moho discontinuity (Mohorovičić, 1909), and the inner core (Lehmann, 1936). Herrin et al. (1968) estimate corrections to the Jeffreys-Bullen *P*-wave travel times and azimuthally dependent station corrections using 400 earthquakes.

The works of Aki and Lee (1976); Aki et al. (1977), Sengupta and Toksöz (1977), and Dziewonski et al. (1977) initiated the computational era of three-dimensional tomographic studies. Dziewonski and Anderson (1981) created a 1D reference Earth model, known as PREM (Preliminary Reference Earth Model), using normal mode data, travel time observations, and mass and moment of inertia information. Kennett and Engdahl (1991) constructed IASP91 using high-frequency body-wave travel time data only. A few years later, Kennett et al. (1995) released AK135, an improved version of IASP91.

Since then, many 3D Earth models (e.g. Ritsema et al., 1999; Grand, 2002; Montelli et al., 2006; Houser et al., 2008; Ritsema et al., 2011; Schaeffer and Lebedev, 2013; French and Romanowicz, 2014; Koelemeijer et al., 2016; Durand et al., 2017; Lu et al., 2019) using multiple techniques and various types of data have been developed,

such as body-wave travel times, surface-wave dispersion, Earth’s normal modes, mass and moment of inertia, and waveform inversion. With the continuous accumulation of knowledge, progressively larger and higher-quality seismic database, and increasing computational power, more advanced techniques emerge, such as full-waveform inversion (FWI). FWI allows for the extraction of much more information from the data than the former travel-time methods. Rather than using ray theory, a high-frequency regime approximation, FWI models the entire wavefield, taking into account finite-frequency effects (Dahlen et al., 2000).

Back in 2009, when Tape et al. (2009) carried out an adjoint waveform tomography of the Southern California crust, the first 3D regional application of FWI using spectral elements (Patera, 1984) and the adjoint method (Tarantola, 1984), FWI was so new that most of the state-of-the-art technology needed to carry it out was still undergoing vigorous development. Now, in 2021, FWI is becoming widely adopted and, in a matter of years, it may be the standard method in seismology to image the Earth’s interior. Most of that change came with the fast increase in computational power achieved in the last decades. Twenty years ago, FWI was simply infeasible. Accompanying the new possibilities opened by the modern HPC, intensive research in the field resulted in major theoretical advances followed by the development of highly efficient computational packages, such as SPEC-FEM3D_GLOBE (Komatitsch and Tromp, 2002a,b).

Despite all these advances, gaining a deep understanding of the theory that underpins FWI, which is paramount to any scientist who wishes to further develop the method or simply make the most of it, remains a big challenge given its elevated theoretical complexity, not to speak of the other countless challenges of the computational implementation itself. For that reason, most of the effort put into this project was not directed to carry out the tomography itself, but to understand the theory required to do so.

Chapter 2 gives a brief introduction to the wave equation in elastodynamics, defining the coordinate system used in the following chapters. Chapter 3 explains the basics of the spectral-element method (SEM), the numerical method implemented by SPEC-FEM3D_GLOBE, which uses a particular set of basis functions and coordinates transformation to solve the wave equation in its weak form. It combines high accuracy

and efficiency with the flexibility required to adapt to complex structures, naturally implementing free boundaries even in the presence of realistic topography. Chapter 4 shows an overview of some inversion techniques and gradient-based optimization methods. It also describes challenges of the inverse problem and solutions to overcome them. Chapter 5 explains the adjoint method, the most efficient way of calculating the gradient needed to update the velocity models in FWI. The adjoint method is the reason why FWI is also known as adjoint tomography. Chapter 6 is a paper in final preparation to be submitted to the *European Journal of Physics*, explaining the finite-frequency theory in a different way, prioritizing didactic over formalism. Chapter 7 is a review of previous tomography studies of South America, briefly explaining the methodology of each one and presenting a summary of the results grouped by some important geological domains of the continent. The results are also compared with this study. Chapter 8 is a paper submitted to the *Journal of Geophysical Research: Solid Earth*, showing the methodology and the results of the adjoint tomography of South America. Chapter 9 is a paper submitted to *Computers & Geosciences*, whose content is a package of routines to process, analyze, share, and visualize large tomographic models created in SPECFEM3D_GLOBE. The paper also includes a detailed explanation of the theory used to create the routines. These routines were used in the project and by Plesa et al. (2020) to study Mars.

By-products of this project include a theoretical contribution establishing the connection between the instantaneous phase (Bozdağ et al., 2011) and the exponentiated instantaneous phase (Yuan et al., 2020) measurements and their adjoint sources. That contribution (Appendix A3 of Yuan et al. (2020)) resulted in the co-authorship of a publication in the *Geophysical Journal International* introducing a new objective function for adjoint tomography (Yuan et al., 2020). Additionally, a depth slice of the final model, SAAM23, became the book cover (Appendix A) of a book in honor of Prof. Umberto Giuseppe Cordani (Bartorelli et al., 2020).

Chapter 2

The Wave Equation in Elastodynamics

On the time scale in which seismic waves propagate, they perceive the Earth's interior (except for the outer core) as an elastic solid. For the sake of simplicity, we also assume that it is an isotropic and non-dissipative medium. The rocks allow the propagation of only two fundamental types of seismic waves: longitudinal and transverse. Other types are derived or composed of those two and originate from the introduction of boundary conditions or changes in the physical properties of the medium. Longitudinal waves are faster and, therefore, are known as P (primary), while transverse waves, which are slower, are called S (secondary) waves. Fig. 2.1 describes a Cartesian coordinate system we use to orient the medium. Let us limit ourselves to 1D and 2D propagation only.

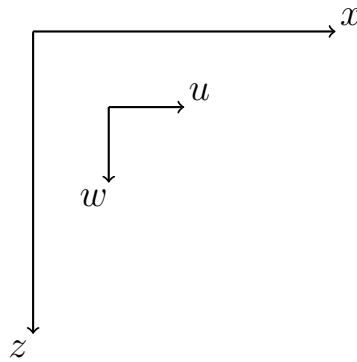


Figure 2.1: Cartesian coordinate system. The axes x and z refer to the direction of the wave propagation; the axes u and w describe the movement of the particles in the medium.

We define $\rho(x, z)$ as the density of the medium and $\lambda(x, z)$ and $\mu(x, z)$ as the elastic constants, known as Lamé coefficients. In the one-dimensional case, there is no movement in z , only in x . When the particles move in the u direction, the wave is longitudinal and described by the P -wave equation, given by 2.1:

$$\rho(x) \frac{\partial^2 u}{\partial t^2} = \frac{\partial[\lambda(x) + 2\mu(x)]}{\partial x} \frac{\partial u}{\partial x} + f_u(x, t), \quad (2.1)$$

where $f_u(x, t)$ is an external excitation force in the direction u . For a homogeneous medium, the P -wave velocity α is:

$$\alpha = \sqrt{\frac{\lambda + 2\mu}{\rho}}. \quad (2.2)$$

On the other hand, if the particles are moving in the transverse direction (w), the propagation will be described by the S -wave equation:

$$\rho(x) \frac{\partial^2 w}{\partial t^2} = \frac{\partial\mu(x)}{\partial x} \frac{\partial w}{\partial x} + f_w(x, t), \quad (2.3)$$

$f_w(x, t)$ being an external force in the direction w .

Similarly, the S -wave velocity β will be:

$$\beta = \sqrt{\frac{\mu}{\rho}}. \quad (2.4)$$

In the one-dimensional case, the P - and S -wave equations are independent. In two dimensions, however, they are coupled, and the propagation is described by the system of equations 2.5 (Moczo, 1998):

$$\begin{aligned} \rho(x, z) \frac{\partial^2 u}{\partial t^2} &= \frac{\partial[\lambda(x, z) + 2\mu(x, z)]}{\partial x} \frac{\partial u}{\partial x} + \frac{\partial\mu(x, z)}{\partial z} \frac{\partial u}{\partial z} \\ &\quad + \frac{\partial\lambda(x, z)}{\partial z} \frac{\partial w}{\partial x} + \frac{\partial\mu(x, z)}{\partial x} \frac{\partial w}{\partial z} + f_u(x, z, t), \\ \rho(x, z) \frac{\partial^2 w}{\partial t^2} &= \frac{\partial\mu(x, z)}{\partial x} \frac{\partial w}{\partial x} + \frac{\partial[\lambda(x, z) + 2\mu(x, z)]}{\partial z} \frac{\partial w}{\partial z} \\ &\quad + \frac{\partial\mu(x, z)}{\partial z} \frac{\partial u}{\partial x} + \frac{\partial\lambda(x, z)}{\partial x} \frac{\partial u}{\partial z} + f_w(x, z, t). \end{aligned} \quad (2.5)$$

When the medium is homogeneous, we can use vector notation to rewrite the wave equation in a more compact form (Müller, 2007). $\vec{d} = \vec{d}(x, z, t)$ the displacement vector and $\vec{f} = \vec{f}(x, z, t)$ the external force vector, 2.5 is written as:

$$\rho \frac{\partial^2 \vec{d}}{\partial t^2} = (\lambda + 2\mu) \nabla \nabla \cdot \vec{d} - \mu \nabla \times \nabla \times \vec{d} + \vec{f}. \quad (2.6)$$

Chapter 3

Spectral-Element Method

Firstly applied by Patera (1984) to fluid dynamics, the spectral-element method (SEM) is currently used in seismology for modeling seismic wave-propagation inside the Earth (Komatitsch and Tromp, 1999). As explained by Schubert (2003) and Igel (2017), the spectral-element method overcomes many of the challenges experienced by older methods. Finite-differences suffered from many difficulties to accurately implement free-surface boundary conditions in the case of realistic topography. The pseudospectral method is accurate, but could not deal with complex geometry easily and was hard to scale on parallel hardware due to global communication requirements. The classical finite-element method is flexible, works properly in the presence of realistic topography and complex geometry, and naturally implements free-surface boundary conditions. However, its implementation required the solution of large linear system of equations, rendering it inefficient and hard to implement in parallel.

By using Lagrange polynomials as basis functions inside the rectangular (in 2D) or hexahedral (in 3D) elements distributed on the Gauss-Lobatto-Legendre (GLL) collocation points, the spectral-element method combines the accuracy of the pseudospectral method with the flexibility of finite-elements implemented as an explicit scheme easily scaled up in parallel hardware, making it ideal for modeling seismic waves (Schubert, 2003; Igel, 2017).

3.1 Spectral-Element Modeling

For simplicity, let us carry out all the mathematical development of the method using Eq. 2.3. This derivation is mainly based on Schubert (2003) and Ciardelli (2015). We omit the two-dimensional formulation of the SEM because it is beyond the scope of this work. For more details, refer to Komatitsch and Tromp (1999), Komatitsch and Tromp (2002a,b), Schubert (2003), and Igel (2017).

3.1.1 Weak Form of the Wave Equation

The weak form of a differential equation, unlike its strong form (governing equations plus boundary conditions), states the conditions that the solution must satisfy in an integral sense, not the conditions at every point over the domain (Mokashi, 2014). To find the weak form of the wave equation, let us multiply both members of 2.3 by an auxiliary function $v(x)$ that is differentiable over Ω , the domain of $w(x, t)$ (the propose of doing that will be clear later), and then integrate both sides with respect to x :

$$\int_{\Omega} v(x) \rho(x) \frac{\partial^2 w}{\partial t^2} dx - \int_{\Omega} v(x) \frac{\partial[\mu(x) \frac{\partial w}{\partial x}]}{\partial x} dx = \int_{\Omega} v(x) f_w(x, t) dx. \quad (3.1)$$

We assume that there is no external force at work. That term could simulate the source that generates the elastic waves, generally consisting of a transient function that describes the moment rate release and the rupture evolution (Gibowicz, 2009). After a few seconds, that force disappears, and $f_w(x, t)$ vanishes from the equation. The way we specify the initial conditions in the absence of an external force is by setting non-null values for $w(x, 0)$ and $w'(x, 0)$. If we were using the external force to simulate the source, we could simply set $w(x, 0) = 0$ and $w(x, 0)' = 0$. Thus, imposing the condition of null external force 3.2:

$$f_w(x, t) = 0 \longrightarrow \int_{\Omega} v(x) f_w(x, t) dx = 0, \quad (3.2)$$

and integrating the second term of 3.1 by parts, Eq. 3.1 becomes:

$$\int_{\Omega} v(x) \rho(x) \frac{\partial^2 w}{\partial t^2} dx - v(x) \mu(x) \frac{\partial w}{\partial x} \Big|_{\Gamma} + \int_{\Omega} \frac{dv}{dx} \mu(x) \frac{\partial w}{\partial x} dx = 0. \quad (3.3)$$

The term $\mu(x) \frac{\partial w}{\partial x} \Big|_{\Gamma}$ corresponds to the tension σ_{Γ} at Γ , the border of Ω . Thus, assuming that this term is null, we naturally satisfy the free-boundary conditions:

$$\mu(x) \frac{\partial w}{\partial x} \Big|_{\Gamma} = \sigma_{\Gamma} = 0. \quad (3.4)$$

Thus, we find the weak form for Eq. 2.3:

$$\int_{\Omega} v(x) \rho(x) \frac{\partial^2 w}{\partial t^2} dx = - \int_{\Omega} \frac{dv}{dx} \mu(x) \frac{\partial w}{\partial x} dx. \quad (3.5)$$

In the 1D case, the boundaries are the two edges of Ω . For a wave traveling in a string, for example, the free-boundary conditions imply that the two ends of the string are free to move.

3.1.2 Partition of the Medium into Elements

One of the major advantages of SEM is the flexibility to adapt to complex shapes. This is done by partitioning the medium into smaller regions, called "elements". The elements can take almost any shape and, when combined, mold to different types of structure. The wave equation is then solved separately for each element, which is subsequently recombined to provide the complete solution. At the element level, Eq. 3.5 is:

$$\int_{\Omega_e} v(x) \rho(x) \frac{\partial^2 w}{\partial t^2} dx = - \int_{\Omega_e} \frac{dv}{dx} \mu(x) \frac{\partial w}{\partial x} dx \quad (3.6)$$

$$e = 1, 2, \dots, n_e.$$

In the unidimensional case, each element is just a line of arbitrary size, defined by the coordinates x_i of its ends:

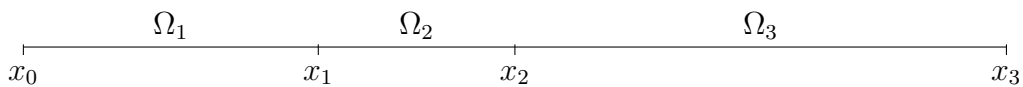


Figure 3.1: Example of partition for a 1D medium. Each element is defined in a subdomain Ω_e , delimited by its ends x_i and x_{i+1} . In this case, we exemplify a division into three elements of different sizes.

3.1.3 Lagrange Polynomials

To numerically solve 3.6, we need an accurate way to differentiate and integrate $w(x, t)$ inside Ω_e . For that, we discretize the function and then interpolate it with a polynomial that we use to calculate the derivatives and integrals.

Runge Phenomenon

We approximate a general function $f(\xi)$ with an n -th degree polynomial $I_n(\xi)$ using equally spaced points. One would expect that the larger the degree n , the more accurate the approximation of $f(\xi)$ would be. In fact, for "well-behaved" functions that happens:

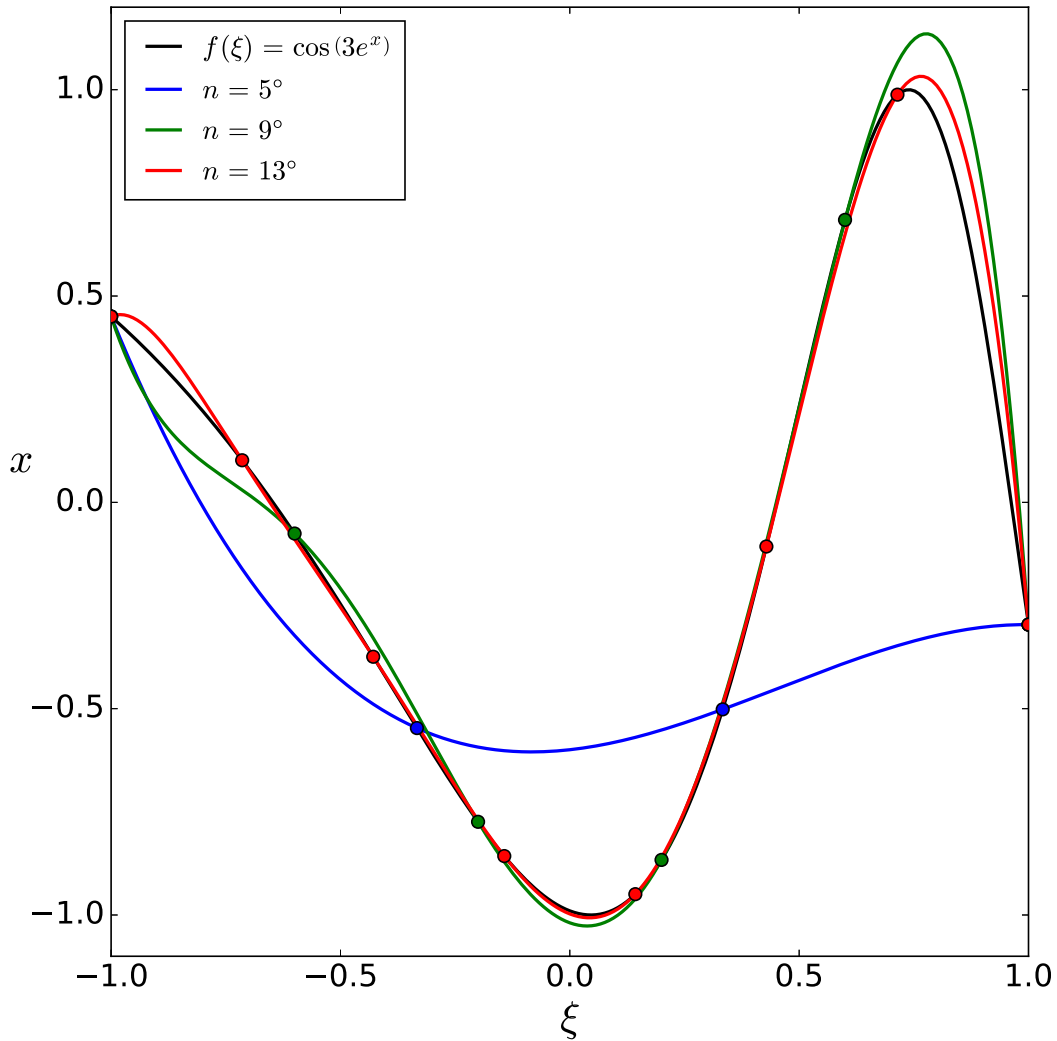


Figure 3.2: Interpolation of a function without Runge phenomenon in a regular grid. The black line represents function $f(\xi) = \cos(2e^\xi)$; the blue line is a polynomial of degree $n = 3$; the green line is a 5° interpolator; the red line is a 7° polynomial. Increasing the value of n causes the error in the approximation to decrease.

For this function, we see that making $n \rightarrow \infty$ implies that the error of the approximation $\max_{-1 \leq \xi \leq 1} |E(\xi)| = \max_{-1 \leq \xi \leq 1} |f(\xi) - I_n(\xi)| \rightarrow 0$. However, there are

other types of functions for which this does not occur. The Runge function, $R(\xi) = 1 / (1 + 25\xi^2)$, is an example:

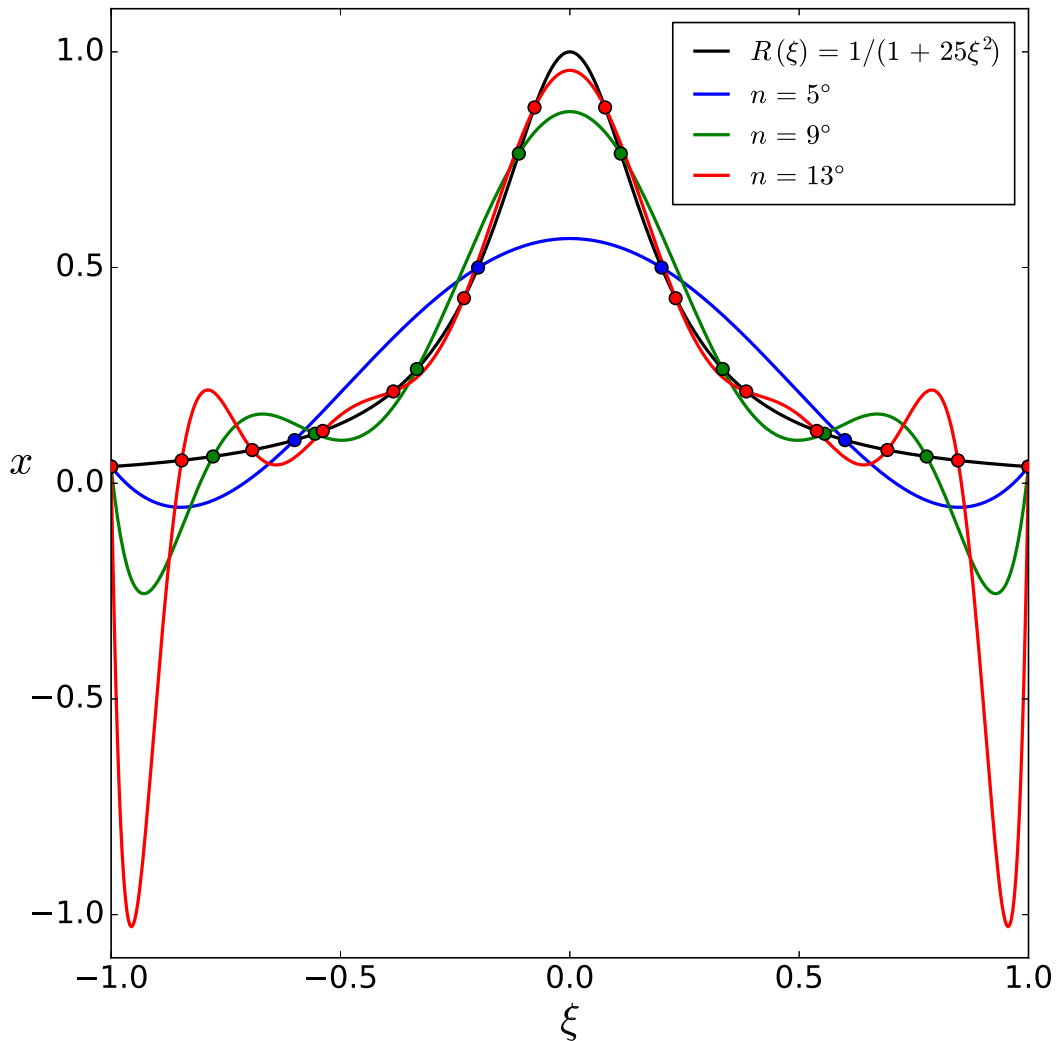


Figure 3.3: Interpolation of the Runge function in a regular grid. The black line represents the Runge function; the blue line is a polynomial of degree $n = 5$; the green line is a 9° interpolator; the red line is a 13° polynomial. Runge phenomenon causes the oscillation at the boundaries of the interval to increase rapidly with the polynomial degree.

Due to the Runge phenomenon, high-degree polynomial interpolation using a regular grid is almost always inappropriate.

Chebyshev Nodes

One solution to this problem is to replace the regular grid with another that concentrates more points at the ends of the interval, which is where convergence problems

usually occur. The best possible distribution is given by the so-called Chebyshev nodes (Epperson, 1987). For the standard interval $\Lambda \rightarrow [-1, 1]$, they are easily calculated by:

$$\xi_i = \cos \left[\frac{(2i + 1)\pi}{2(n + 1)} \right], \quad 0 \leq i \leq n. \quad (3.7)$$

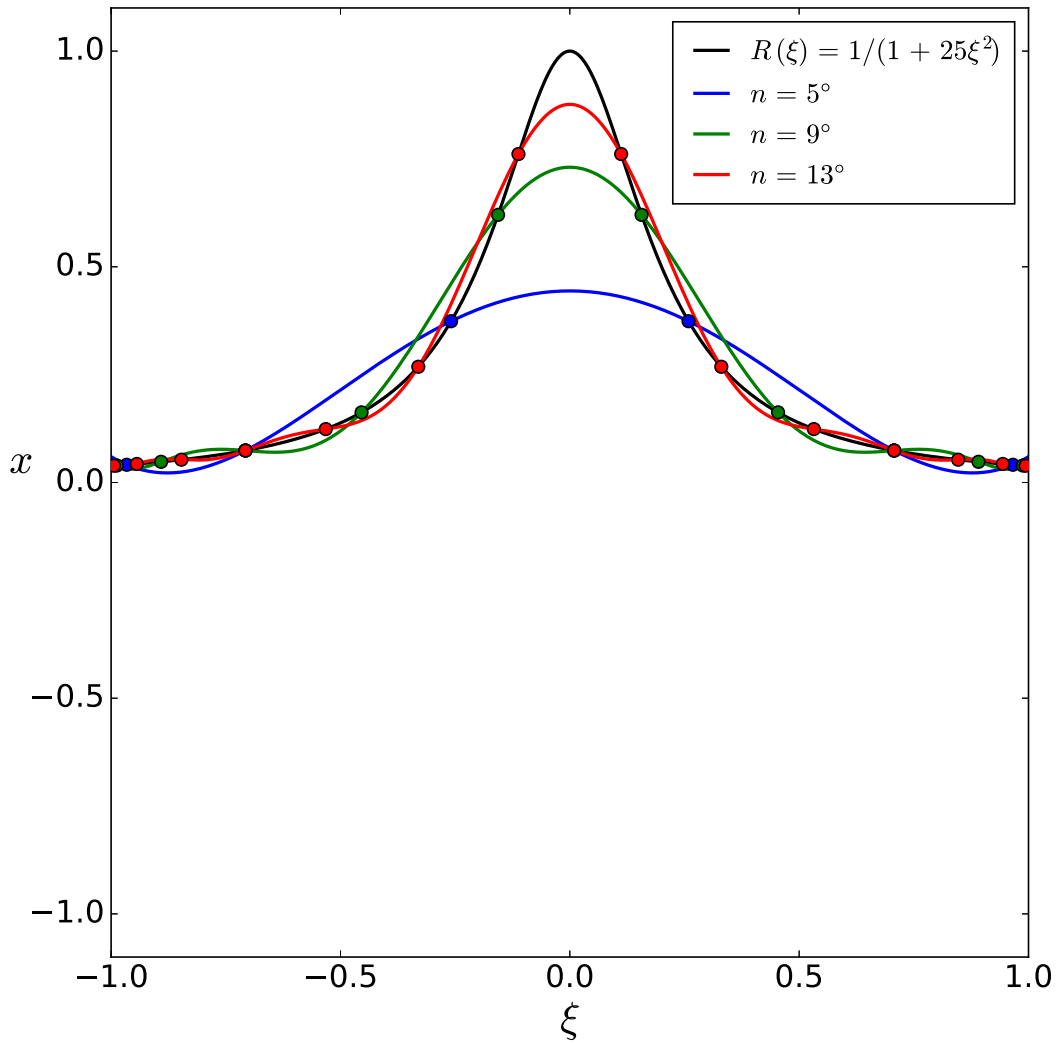


Figure 3.4: Interpolation of the Runge function using Chebyshev nodes. In black, we see the Runge function; in blue, an interpolating polynomial of degree $n = 5$; in green, a 9° interpolator; in red, a 13° polynomial. Using Chebyshev nodes instead of a regular grid causes $\max_{-1 \leq \xi \leq 1} |E(\xi)|$ is minimal and decreases by n .

As we see in Fig. 3.4, Runge phenomenon disappears using the values of ξ_i calculated with 3.7.

3.1.4 Assembly and Solution of the System of Equations

The first step to solve 3.6 is mapping each of the e_i elements from its subdomain $\Omega_i \rightarrow [x_i, x_{i+1}]$ to the standard interval $\Lambda \rightarrow [-1, 1]$, using a coordinate transformation. This is necessary to calculate the integrals using Gaussian quadrature. Thus, $w(x, t)$ can be approximated in Λ by a n -degree polynomial:

$$w_e(\xi) \approx \sum_{i=0}^n w_e(\xi_i) l_i(\xi), \quad \xi \in [-1, 1], \quad (3.8)$$

being $l_i(\xi)$ the i -th Lagrange polynomial at ξ .

Similarly, the first spatial derivative of $w_e(\xi)$ at a point $\xi_i \in \Lambda$ will be:

$$\frac{\partial w_e}{\partial \xi}(\xi_i) \approx \sum_{j=0}^n w_e(\xi_j) l'_j(\xi_i). \quad (3.9)$$

To approximate the integral, we use the so-called Gauss-Lobatto-Legendre (GLL) quadrature, which uses a set of points similar to Chebyshev's nodes, known as Gauss-Lobatto-Legendre points. Unlike Chebyshev nodes, GLL nodes include the two ends of the integration domain: -1 and 1 . By forcing them to be among the nodes, the GLL quadrature loses some accuracy compared to the Chebyshev points, as the number of degrees of freedom is reduced by two. However, it also simplifies the solution of the system of equations that will be shown later (Schubert, 2003). Thus, the calculation of the integral in Λ is:

$$\int_{-1}^1 w_e(\xi) d\xi \approx \sum_{i=0}^n w_i w_e(\xi_i). \quad (3.10)$$

The values w_i are referred to as "weights" in the quadrature (Casio Computer, 2015). Eq. 3.11 is the formula to calculate them:

$$w_i = \begin{cases} \frac{2}{n(n+1)} & \text{of } \xi_i = \pm 1 \\ \frac{2}{n(n+1)P_n(\xi_i)^2} & \text{if } \xi_i \neq \pm 1, \end{cases} \quad (3.11)$$

where $P_n(x)$ are the Legendre polynomials (Eq. 3.12):

$$P_n(\xi) = \frac{1}{2^n n!} \frac{d^n}{d\xi^n} (\xi^2 - 1)^n, \quad (3.12)$$

where the nodes ξ_i are the roots of $P'_n(x)$.

We now begin the mathematical manipulation that will lead to an accurate scheme to numerically solve the S -wave equation at an elementary level. We start by mapping, interpolating, and integrating the first term of 3.6:

$$\begin{aligned}
& \text{mapping in } \Lambda \downarrow \\
& \int_{\Omega_e} v(x) \rho(x) \ddot{w}(x) dx = \int_{\Lambda} \rho_e(\xi) v_e(\xi) \ddot{w}_e(\xi) \frac{dx_e}{d\xi} d\xi \\
& \text{interpolating } v_e(\xi) \text{ e } \ddot{w}_e(\xi) \downarrow \\
& \approx \int_{-1}^1 \rho_e(\xi) \left[\sum_{k=0}^n v_e(\xi_k) l_k(\xi) \right] \left[\sum_{i=0}^n \ddot{w}_e(\xi_i) l_i(\xi) \right] \frac{dx_e}{d\xi} d\xi \\
& \text{GLL quadrature } \downarrow \\
& \approx \sum_{j=0}^n \left\{ w_j \rho_e(\xi_j) \left[\sum_{k=0}^n v_e(\xi_k) l_k(\xi_j) \right] \left[\sum_{i=0}^n \ddot{w}_e(\xi_i) l_i(\xi_j) \right] \frac{dx_e}{d\xi} \right\}.
\end{aligned} \tag{3.13}$$

We use the Kronecker delta, which is given by:

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases} \tag{3.14}$$

From 3.14 we derive an useful property:

$$\sum_{j=0}^n f(\xi_j) \delta_{ij} = f(\xi_i). \tag{3.15}$$

rearranging the summations \downarrow

$$\int_{\Omega_e} v(x) \rho(x) \ddot{w}(x) dx \approx \sum_{i=0}^n \left\{ \ddot{w}_e(\xi_i) \left[\sum_{j=0}^n \sum_{k=0}^n w_j \rho_e(\xi_j) \delta_{ij} v_e(\xi_k) \delta_{kj} \right] \frac{dx_e}{d\xi} \right\}$$

applying 3.15 to the summation in k \downarrow

$$= \sum_{i=0}^n \left\{ \ddot{w}_e(\xi_i) \left[\sum_{j=0}^n w_j \rho_e(\xi_j) v_e(\xi_j) \delta_{ij} \right] \frac{dx_e}{d\xi} \right\}$$

applying 3.15 to the summation in j \downarrow

$$= \sum_{i=0}^n \ddot{w}_e(\xi_i) w_i \rho_e(\xi_i) v_e(\xi_i) \frac{dx_e}{d\xi}.$$

(3.16)

On the other hand, for the right-hand side of 3.6:

mapping in Λ \downarrow

$$\int_{\Omega_e} \frac{dv}{dx} \mu(x) \frac{\partial w}{\partial x} dx = \int_{\Lambda} \frac{dv_e}{d\xi} \mu_e(\xi) \frac{\partial w_e}{\partial x} \frac{dx_e}{d\xi} d\xi$$

applying the chain rule \downarrow

$$= \int_{\Lambda} \frac{dv_e}{d\xi} \frac{d\xi}{dx_e} \mu_e(\xi) \frac{\partial w_e}{\partial \xi} \frac{\cancel{\partial \xi}}{\cancel{\partial x_e}} \frac{\cancel{dx_e}}{\cancel{d\xi}} d\xi$$

interpolating $\frac{dv_e}{d\xi}$ and $\frac{\partial w_e}{\partial \xi}$ \downarrow

$$\approx \int_{-1}^1 \mu_e(\xi) \left[\sum_{k=0}^n v_e(\xi_k) l'_k(\xi) \right] \left[\sum_{i=0}^n w_e(\xi_i) l'_i(\xi) \right] \frac{d\xi}{dx_e} d\xi$$

GLL quadrature \downarrow

$$\approx \sum_{j=0}^n \left\{ w_j \mu_e(\xi_j) \left[\sum_{k=0}^n v_e(\xi_k) l'_k(\xi_j) \right] \left[\sum_{i=0}^n w_e(\xi_i) l'_i(\xi_j) \right] \frac{d\xi}{dx_e} \right\}$$

rearranging the summations \downarrow

$$= \sum_{i=0}^n \left\{ w_e(\xi_i) \left[\sum_{j=0}^n \sum_{k=0}^n w_j \mu_e(\xi_j) v_e(\xi_k) l'_i(\xi_j) l'_k(\xi_j) \right] \frac{d\xi}{dx_e} \right\}.$$

(3.17)

The formula for calculating $l'_i(\xi_j)$ and $l'_k(\xi_j)$ is given by 3.18 (El-Baghdady et al., 2015):

$$l'_j(\xi_i) = \begin{cases} -\frac{n(n+1)}{4} & \text{if } i = j = 0 \\ \frac{n(n+1)}{4} & \text{if } i = j = n \\ 0 & \text{if } 1 \leq i = j \leq n-1 \\ \frac{P_n(\xi_i)}{P_n(\xi_j)} \frac{1}{\xi_i - \xi_j} & \text{if } 0 \leq i \leq n, 0 \leq j \leq n \text{ and } i \neq j. \end{cases} \quad (3.18)$$

From 3.6, 3.16, and 3.17 we write:

$$\begin{aligned} \sum_{i=0}^n \ddot{w}_e(\xi_i) w_i \rho_e(\xi_i) v_e(\xi_i) \frac{dx_e}{d\xi} = \\ - \sum_{i=0}^n \left\{ w_e(\xi_i) \left[\sum_{j=0}^n \sum_{k=0}^n w_j \mu_e(\xi_j) v_e(\xi_k) l'_i(\xi_j) l'_k(\xi_j) \right] \frac{d\xi}{dx_e} \right\}. \end{aligned} \quad (3.19)$$

Our goal is to find an analytical expression to calculate each of the values of $\ddot{w}_e(\xi_i)$. So far, we have not made any assumptions about $v(x)$ other than that it is differentiable in Ω . As 3.19 is valid in the whole domain Ω , we can set the values of $v_e(\xi_i)$ as:

$$v_e(\xi_i) = \begin{cases} 1 & \text{se } i = 0 \\ 0 & \text{se } i \neq 0. \end{cases} \quad (3.20)$$

Thus, 3.19 reduces to:

$$\ddot{w}_e(\xi_0) w_0 \rho_e(\xi_0) \frac{dx_e}{d\xi} = - \sum_{i=0}^n \left\{ w_e(\xi_i) \left[\sum_{j=0}^n w_j \mu_e(\xi_j) l'_i(\xi_j) l'_0(\xi_j) \right] \frac{d\xi}{dx_e} \right\}. \quad (3.21)$$

or

$$\ddot{w}_e(\xi_0) = \frac{-1}{w_0 \rho_e(\xi_0)} \sum_{i=0}^n \left\{ w_e(\xi_i) \left[\sum_{j=0}^n w_j \mu_e(\xi_j) l'_i(\xi_j) l'_0(\xi_j) \right] \left(\frac{d\xi}{dx_e} \right)^2 \right\}. \quad (3.22)$$

By replacing Eq. 3.20 into 3.19 we extract the value of $\ddot{w}_e(\xi_0)$, given by Eq. 3.22. To extract $\ddot{w}_e(\xi_1)$, we could do something analogous, but replacing 3.20 by 3.23:

$$v_e(\xi_i) = \begin{cases} 1 & \text{se } i = 1 \\ 0 & \text{se } i \neq 1 \end{cases}. \quad (3.23)$$

We proceed in this way for each value in $[0, n]$, finding all the values of $\ddot{w}_e(\xi_i)$:

$$\ddot{w}_e(\xi_i) = \frac{-1}{w_i \rho_e(\xi_i)} \sum_{j=0}^n \left\{ w_e(\xi_j) \left[\sum_{k=0}^n w_k \mu_e(\xi_k) l'_j(\xi_k) l'_i(\xi_k) \right] \left(\frac{d\xi}{dx_e} \right)^2 \right\} \quad (3.24)$$

$$i = 0, 1, \dots, n.$$

Notice that we changed the indexes on the right-hand side of 3.22 to avoid confusion with the value of i . The finite-difference approximation for the second derivative of function $f(x)$ is given by 3.25:

$$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = f''(x) + O(h^2). \quad (3.25)$$

If we applied 3.25 to the left-hand side of the equation, we would find a recurrence relation to calculate $w_{i,l+1}$ as a function of $w_{i,l}$ and $w_{i,l-1}$. However, we have to remember that 3.24 is restricted to only one element Ω_e from our domain Ω . To find a general formula, which combines all the elements, it is more appropriate to rewrite 3.24 in matrix notation:

$$\ddot{w}_e D_e + w_e R_e = 0. \quad (3.26)$$

Because we performed the interpolation using GLL nodes, the matrix D_e is diagonal. That makes its inversion trivial and reduces the amount of memory used and the number of multiplications performed at each iteration. Subtracting the second term from both sides of the system, we write 3.26 in expanded form:

$$\underbrace{\begin{bmatrix} \ddot{w}_0 & \ddot{w}_1 & \dots & \ddot{w}_n \end{bmatrix}}_{\ddot{w}_e} \cdot \underbrace{\begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix}}_{D_e} = - \underbrace{\begin{bmatrix} w_0 & w_1 & \dots & w_n \end{bmatrix}}_{w_e} \cdot \underbrace{\begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nn} \end{bmatrix}}_{R_e}, \quad (3.27)$$

where

$$d_i = w_i \rho_e(\xi_i) \frac{dx_e}{d\xi} \quad \text{and} \quad r_{jk} = \left[\sum_{k=0}^n w_k \mu_e(\xi_k) l'_j(\xi_k) l'_i(\xi_k) \right] \frac{d\xi}{dx_e}. \quad (3.28)$$

To solve 3.28 at the elementary level, we multiply the two members by the inverse of the density matrix (also known as "mass matrix"):

$$\ddot{w}_e = -w_e R_e D_e^{-1}. \quad (3.29)$$

$$\begin{bmatrix} \ddot{w}_0 & \ddot{w}_1 & \dots & \ddot{w}_n \end{bmatrix} = - \begin{bmatrix} w_0 & w_1 & \dots & w_n \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nn} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{d_1} & 0 & \dots & 0 \\ 0 & \frac{1}{d_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{d_n} \end{bmatrix}.$$

To obtain the global system, we combine each of the elementary matrices D_e and R_e . In both cases, we know that the edge points are usually shared by two elements. Nonetheless, since integration is a linear operation, it is enough to add the contribution of each side. Assuming that the domain was partitioned into three elements, each with three nodes, we would have for the density matrix:

$$D = \begin{bmatrix} d_1^{(1)} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & d_2^{(1)} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & d_3^{(1)} + d_1^{(2)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_2^{(2)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & d_3^{(2)} + d_1^{(3)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & d_2^{(3)} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & d_3^{(3)} \end{bmatrix}. \quad (3.30)$$

On the other hand, for the stiffness matrix:

$$R = \begin{bmatrix} r_{11}^{(1)} & r_{12}^{(1)} & r_{13}^{(1)} & 0 & 0 & 0 & 0 \\ r_{21}^{(1)} & r_{22}^{(1)} & r_{23}^{(1)} & 0 & 0 & 0 & 0 \\ r_{31}^{(1)} & r_{32}^{(1)} & r_{33}^{(1)} + r_{11}^{(2)} & r_{12}^{(2)} & r_{13}^{(2)} & 0 & 0 \\ 0 & 0 & r_{21}^{(2)} & r_{22}^{(2)} & r_{23}^{(2)} & 0 & 0 \\ 0 & 0 & r_{31}^{(2)} & r_{32}^{(2)} & r_{33}^{(2)} + r_{11}^{(3)} & r_{12}^{(3)} & r_{13}^{(3)} \\ 0 & 0 & 0 & 0 & r_{21}^{(3)} & r_{22}^{(3)} & r_{23}^{(3)} \\ 0 & 0 & 0 & 0 & r_{31}^{(3)} & r_{32}^{(3)} & r_{33}^{(3)} \end{bmatrix}. \quad (3.31)$$

Both are square matrices with $N \times N$ cells, where:

$$N = n \times n_e + 1. \quad (3.32)$$

Therefore, to solve the global system, we write:

$$\ddot{w} = -w R D^{-1}. \quad (3.33)$$

Applying 3.25 to the first member:

$$\frac{w_{i,l+1} - 2w_{i,l} + w_{i,l-1}}{\Delta t^2} = -w R D^{-1}. \quad (3.34)$$

Isolating $w_{i,l+1}$, we find a recurrence relation for the global system:

$$w_{i,l+1} = -\Delta t^2 w R D^{-1} + 2w_{i,l} - w_{i,l-1}. \quad (3.35)$$

The recurrence relation 3.35 is an algorithm that allows us to compute the future state of the system $w_{i,l+1}$ using its present state $w_{i,l}$ and past state $w_{i,l-1}$. The future been completely determined by the present and the past is what characterizes any deterministic phenomenon.

This chapter was devoted to giving a brief introduction to the SEM using the 1D case. However, some of the main advantages of using spectral elements are much clearer in higher dimensions. The possibility of using coordinate transformations, for example, is what gives the SEM its flexibility to adapt to complex geometries. In 2D and 3D, we can use distorted elements to honor the surface topography and the internal structures properly without any efficiency loss. That occurs because all the elements are mapped to the standard interval and treated in the same way.

Chapter 4

Geophysical Inversion

Inversion is the mathematical process in which we combine observed data with numerical modeling and an optimization method to find a mathematical representation of a physical parameter. In our case, the seismic velocities and density inside the Earth.

4.1 Inverting Seismic Velocity

As briefly mentioned, there are several methods available to model the propagation of seismic waves. The next step is to use the difference between the observed and the synthetic data to update the velocity model, making it the best possible representation of the Earth's internal structure.

4.1.1 Wave Equation using Operators

First of all, let us introduce some formalisms that will be useful from now on, such as the concept of operators. The one-dimensional wave equation is:

$$[m(x) \partial_{tt} - \partial_{xx}] w(x, t) = f(x, t), \quad \text{where } m(x) = \frac{1}{c(x)^2}, \quad (4.1)$$

or even:

$$L(m) w(x, t) = f(x, t), \quad \text{where } L(m) = m(x) \partial_{tt} - \partial_{xx}. \quad (4.2)$$

In this case, $L(m)$ is the operator, dependent on the model $m(x) = \frac{1}{c(x)^2}$, that when applied to the wavefield $w(x, t)$, returns the seismic source $f(x, t)$. To solve Eq. 4.1, we need to find the inverse operator of $L(m)$, such that:

$$w(x, t) = L^{-1}(m) f(x, t). \quad (4.3)$$

The inverse operator of a differentiation operator is an integration operator. The solution to this and many other non-homogeneous differential equations is conveniently expressed using Green's functions.

4.1.2 Local Minima

The most trivial form of inversion consists of carrying out a global search, i.e., test numerous models and choose the one that provides the minimum residue. However, whether due to the insufficient number of observations or even the intrinsic non-uniqueness of the solutions, often, there are multiple models that fit the data equally well. In general, there will be a whole family of solutions that satisfactorily explain the data (Fig. 4.1). One of the advantages of global search methods is that they easily identify this non-uniqueness and reduce the likelihood that the optimization process will get stuck at a local minimum. Knowing the family of solutions allows us to have a good idea of how well-defined the model parameters are. One of the advantages of using the entire waveform in seismic tomography is that it substantially reduces the solutions space since all the information available in the seismogram is used (Fichtner, 2010). On the other hand, this methodology requires far more processing power than travel-time tomography, which makes the use of global optimization algorithms, especially in 3D simulations, infeasible. Even in our one-dimensional example, this type of optimization method would be too inefficient.

Suppose, for example, that we are sure that our starting model does not differ by more than 30% from the real Earth. Suppose also that we want to check velocity values with a resolution of 1%. In this case, for each coordinate x in our model $c(x)$, there are 61 possible values. In a simulation with $n = 1,000$ grid points, there would be a total of $61^{1,000}$ possible model combinations and, for each of these combinations, it would be necessary to simulate the entire wavefield to calculate the synthetic seismograms and compare their residue with the observed data. Needless to say, this is impossible even with the most powerful computer ever created. However, there are more efficient forms of global search.

Probabilistic algorithms can be used to check only a small fraction of the solutions representative of the whole set. Genetic algorithms, for example, search families

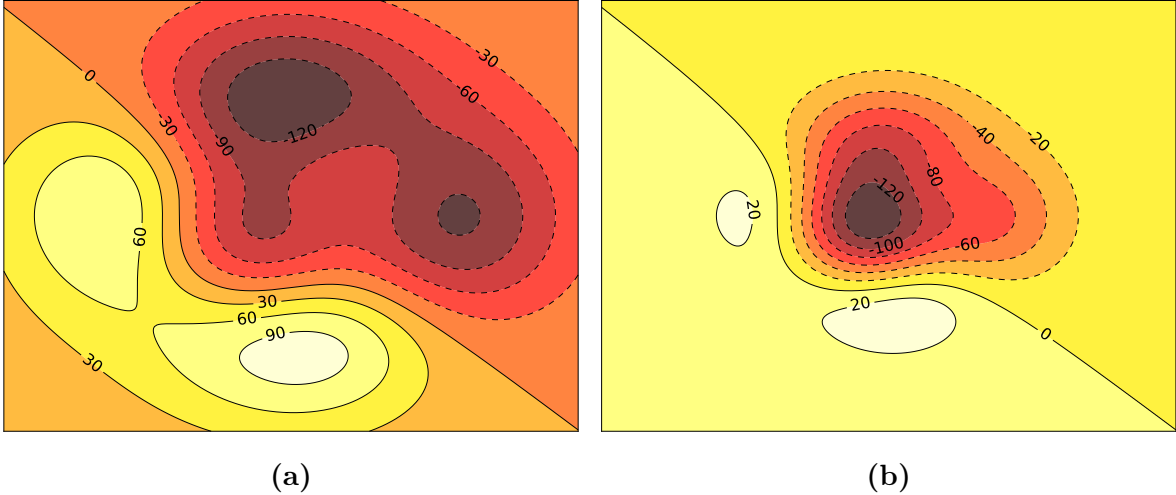


Figure 4.1: Contours of the RMS residual of a hypothetical problem. On the left, we see an objective function in which the solution is not well-constrained. If we move away from the minimum, the residue between the observed and the synthetic data increases slowly. On the right, another objective function with parameters better constrained. When we move away from the minimum, the residue grows faster than in the previous case.

of models that better explain the data and use them to generate new families that, on each generation, tend to get progressively closer to the solution. They converge much faster than brute force methods. However, even genetic algorithms are too slow for 3D tomography. In the three-dimensional case, our search space would be overly large: $(61^3)^{1,000} = 61^{3,000}$ and we still need to take into account that, in three dimensions, each simulation is far more expensive, taking hours even on supercomputers.

With the computational resources currently available, the only solution is using local search methods, in which optimization is performed using gradients to accelerate the convergence to the local minimum as much as possible. That means that if our starting model is not close enough to the global minimum, depending on the objective function, there is a reasonable chance that we would end up stuck into a local minimum (Fig. 4.2). Therefore, to carry out a full-waveform inversion (FWI), the first prerequisite is having a relatively good starting model, often created through travel-time tomography or by fitting surface-wave dispersion curves (Fichtner, 2010).

Another technique to reduce the risks of being stuck in a local minimum is smoothing the starting model and performing the first iterations using only the long wavelengths in the seismogram. Once we fit them, we gradually increase the frequency

content in the data to fit increasingly shorter wavelengths (Tape et al., 2009; Fichtner, 2010; Bozdağ et al., 2016). This technique, called the multistage approach, has the additional advantage of saving a lot of processing time in the first iterations as reducing the shortest wavelength of the synthetic seismogram rapidly increases the computational cost in 3D simulations. A reduction in the shortest period from 30 s to 15 s, for example, leads to a ~ 16 -fold increase in the computational cost. In our 1D simulations, the processing time increases quadratically with the minimum period of the simulation.

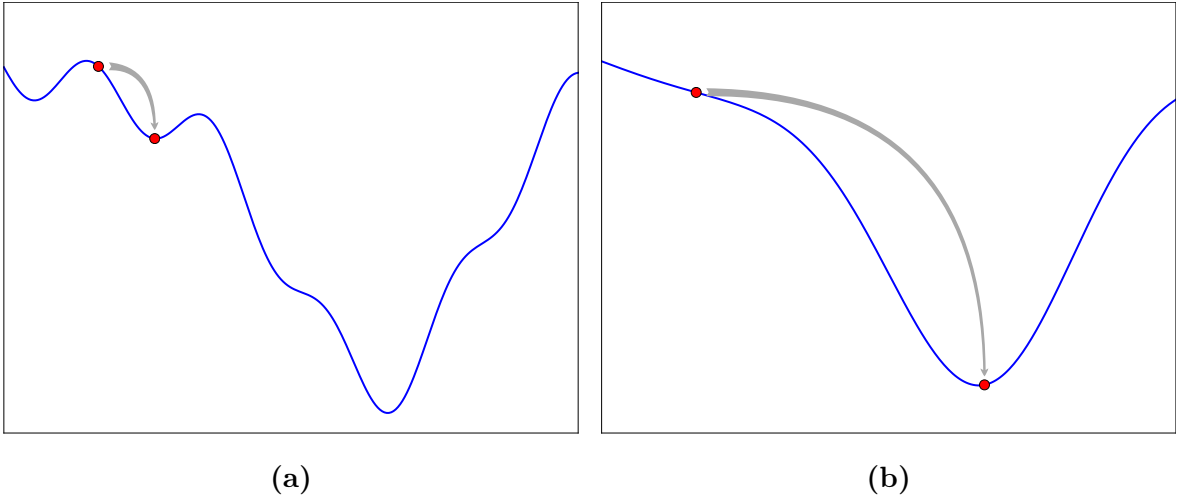


Figure 4.2: On the left, a profile of an objective function whose simulation of the propagation of the wavefield, performed including short-period data. As the starting model (red dot) was not close enough to the global minimum, convergence stopped at a local minimum. On the right, the same simulation performed using only the low frequencies of the seismogram. Once the solution has converged to the global minimum and the long periods have been adjusted, the highest frequencies are added to perform the following iterations.

Even when combining this strategy with a good starting model, additional measures should be taken to increase the chances of convergence. One is beginning the tomography using only the data that is already relatively well-fitted. As the solution converges, we gradually introduce the remaining data. It is also critical to remove as much compromised data as possible, as they may introduce errors and hamper the inversion.

Each seismographic station has three sensors to measure the E-W, N-S, and Z components of the movement. Ideally, the selection should take place independently

for each component, station, and earthquake. Besides, within each seismogram, we select only the segments that are well-adjusted and ignore the others. As the solution converges, the width of the time windows should increase so that, in the last iterations, we are using as much data as possible.

The fourth measure to improve convergence and reduce the number of local minima is related to the choice of the objective function. The simplest approach is measuring the difference between the observed and the synthetic data using the integral of the squared residual (Fichtner, 2010):

$$\chi(c) = \frac{1}{2} \int_0^T [d(t) - s(t, c)]^2 dt, \quad (4.4)$$

where $d(t)$ is the observed seismogram, $s(t, c)$ is the synthetics, which is function of the model c , and T is the total duration of the record.

In general, as several earthquakes are used in the tomography, recorded by many seismographic stations, each with three components, the overall misfit will be the sum of the misfit of each trace. Eq. 4.4 is simple, intuitive, and takes advantage of all waveform information. However, the square in the formula further accentuates the nonlinearity between the velocity model and the waveforms, aggravating the problem with local minima (Fichtner, 2010). Most of the nonlinearity, however, comes from amplitude measurements (Fichtner, 2010). That is why it is a common practice to use objective functions that use only phase information. For now, let us make things easier and use the objective function given by 4.4.

The idea of gradient-based optimization methods is quite simple: The gradient always points towards the maximum increase in the misfit. Therefore, if we want to find the minimum misfit, we should move in the opposite direction. But how much should we move? The gradient tells us the direction but not the distance. We have to remember that the derivative is a local measurement. Therefore, if we move too far from the starting point, the gradient may no longer be pointing in the right direction and has to be recalculated. On the other hand, if we use a small step, the convergence may take too long. Thus, we must find a balance between stability and the convergence rate. One way of solving that problem is performing a linear search. A linear search consists of trying a few different steps and computing the corresponding misfits. Then,

we fit a parabola (or a higher-degree polynomial if we have more than three points) and use its minimum as the ideal step.

4.1.3 Computation of the Gradient

Now that we know how to use the gradient of the objective function to optimize the model, we must find a way to calculate the gradient of function 4.4 with respect to a perturbation $\delta c(x)$ in the model. The velocity model is a discretized function $c = c(x_i)$, for $i \in [1, 2 \dots n]$, therefore, function $\chi(c)$ is multidimensional, having n variables. The gradient is a vector whose components are the partial derivatives of $\chi(c)$ with respect to each of the n model parameters (Eq. 4.5):

$$\frac{\partial \chi}{\partial c}(c) = \begin{bmatrix} \frac{\partial \chi}{\partial c_1} \\ \frac{\partial \chi}{\partial c_2} \\ \vdots \\ \frac{\partial \chi}{\partial c_n} \end{bmatrix}. \quad (4.5)$$

The simplest way to calculate partial derivatives is, again, using finite differences. The central approximation for the partial derivative of the function calculated at the point is given by:

$$\frac{\partial \chi}{\partial c_i}(c_i) \approx \frac{\chi(c_i + \epsilon) - \chi(c_i - \epsilon)}{2\epsilon}, \quad (4.6)$$

where ϵ is the perturbation in the parameter c_i , corresponding to the velocity in the position $x = x_i = x[i]$ of the model.

In other words, the process to calculate the gradient is to perturb each point in the velocity model separately. Each time a perturbation is applied, we recompute the synthetic seismogram $s(t, c)$ and the value of $\chi(c)$. Then, we remove the perturbation and apply an equal perturbation in the next parameter of the model, repeating the previous process. Altogether, this must be done $2n$ times, twice (one positive and one negative perturbation) for each model value. Using $\chi(c_i + \epsilon)$ and $\chi(c_i - \epsilon)$, we apply Eq. 4.6 to each index i and calculate the gradient. Therefore, by using finite differences, we limit the number of simulations required in each iteration to twice

the number of model parameters, which is nothing but the number of points used in discretization (in this case, $n = 2,000$). If twenty iterations are needed for the solution to converge, we would have a total of 40,000 simulations to reach the final model. This number is far less than the simulations required by the brute force algorithm. For our 1D tomography, this amount is completely feasible, and the whole process, from the starting model to the final solution, takes a few hours on a personal computer. However, for a 3D tomography in which each simulation takes hours, this is not fast enough. Later, we shall see a far more efficient way of calculating the gradient and reducing the number of iterations needed.

One of the problems in Eq. 4.6 is to determine an appropriate value for ϵ . Unlike Δx and Δt , which are linked to the maximum frequency of the seismic source and the maximum velocity of the model, the choice of ϵ is not very obvious. In principle, it should be as small as possible, since the closer to zero, the more accurate the derivative should be. If it were infinitesimal, the finite differences would become the derivative itself. However, the reason for using finite differences is precisely the fact that infinitesimal values do not exist inside computers. Hence, if ϵ is too small, the numerical error can become more significant than the perturbation itself, and we would end up losing accuracy rather than increasing it.

4.1.4 Steepest Descent Method

In the steepest descent method, in each iteration, we update the model using the gradient multiplied by a negative constant. For a faster convergence, that constant (step) can be estimated using linear search, as explained previously. Let us look at a simple example using a paraboloid (Fig. 4.3):

Even using linear search, the convergence took sixteen iterations. Looking at Fig. 4.3, we see what the problem is. The gradient is always perpendicular to the contour lines, pointing towards the direction in which, locally, the function decreases most quickly (remembering that we are going in the negative direction of the gradient). However, that direction does not necessarily point to the minimum of the function. Each time we recalculate the gradient, it will point in a direction perpendicular to the previous one, resulting in slow convergence.

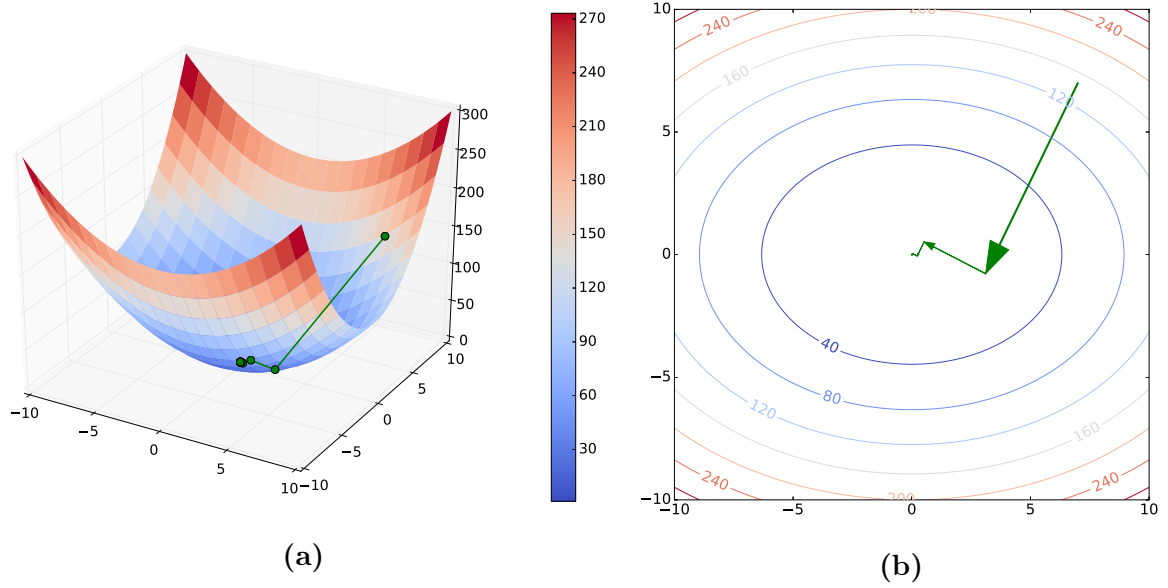


Figure 4.3: Left: 3D view of convergence to the minimum using the steepest descent method with a step estimated by a linear search for a paraboloid with axes of different sizes. Right: same process represented in 2D with contour lines.

4.1.5 Newton's Method

As mentioned before, if the objective function is analytic, we can expand it using the Taylor series. In the neighborhood of c , we can approximate the misfit function by (Komatitsch and Tromp, 1999):

$$\chi(c + \delta c) \approx \chi(c) + \frac{\partial \chi}{\partial c}(c) \delta c + \frac{1}{2} \frac{\partial^2 \chi}{\partial c^2}(c) \delta c^2. \quad (4.7)$$

Our goal is to find the perturbation δc that we must add to the model c so that $\chi(c + \delta c)$ is minimum. To do that, we differentiate 4.7 with respect to δc and equal the result to zero:

$$\frac{\partial \chi}{\partial \delta c}(c + \delta c) \approx \frac{\partial \chi}{\partial c}(c) + \frac{\partial^2 \chi}{\partial c^2}(c) \delta c = 0. \quad (4.8)$$

Therefore:

$$\delta c = -H(c)^{-1} g(c), \quad (4.9)$$

where $g(c)$ is the gradient of $\chi(c)$:

$$g(c) = \frac{\partial \chi}{\partial c}(c), \quad (4.10)$$

and $H(c)$ is the Hessian:

$$H(c) = \frac{\partial^2 \chi}{\partial c^2}(c). \quad (4.11)$$

$\chi(c)$ is a multidimensional function, $g(c)$ is a vector, and $H(c)$ is a matrix, given by:

$$\frac{\partial^2 \chi}{\partial c^2}(c) = \begin{bmatrix} \frac{\partial^2 \chi}{\partial c_1^2}(c) & \frac{\partial^2 \chi}{\partial c_1 \partial c_2}(c) & \cdots & \frac{\partial^2 \chi}{\partial c_1 \partial c_n}(c) \\ \frac{\partial^2 \chi}{\partial c_2 \partial c_1}(c) & \frac{\partial^2 \chi}{\partial c_2^2}(c) & \cdots & \frac{\partial^2 \chi}{\partial c_2 \partial c_n}(c) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \chi}{\partial c_n \partial c_1}(c) & \frac{\partial^2 \chi}{\partial c_n \partial c_2}(c) & \cdots & \frac{\partial^2 \chi}{\partial c_n^2}(c) \end{bmatrix}. \quad (4.12)$$

By using the second derivative, Newton's method takes the curvature into account and avoids that problem, ensuring a much faster convergence than the steepest descent method. In this case, as the function is a paraboloid, the convergence occurs with a single iteration (Fig. 4.4).

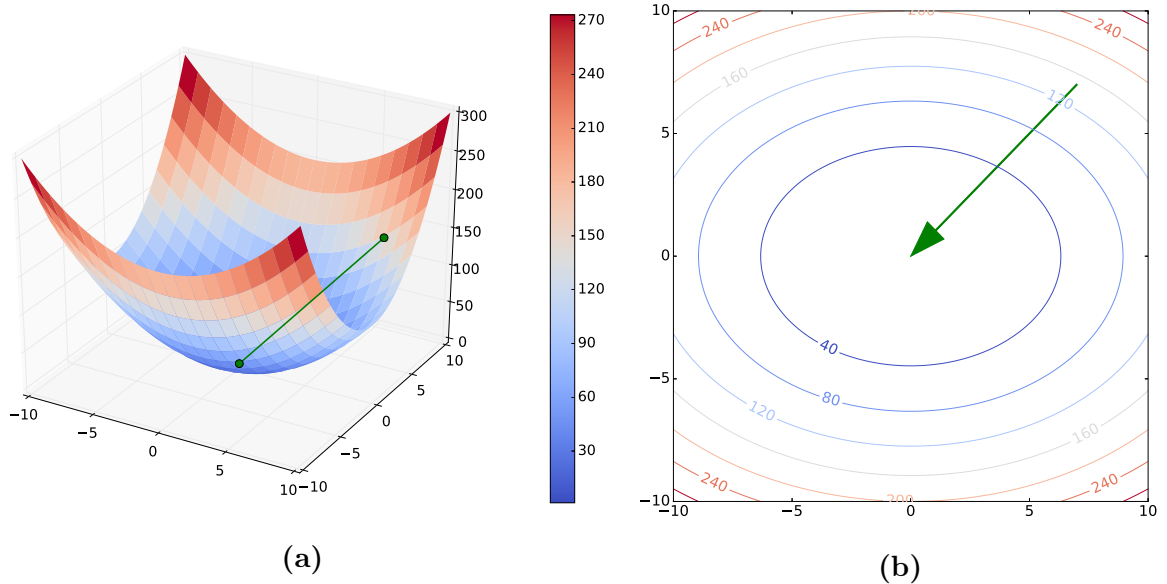


Figure 4.4: Left: 3D view of convergence to the minimum by Newton's method. Right: same process represented in 2D with contour lines.

Another advantage of Newton's method is that, according to Eq. 4.9, multiplication by the inverse of the Hessian matrix not only rotates the gradient in the direction of the minimum but also adjusts its length to the exact size. Thus, unlike the steepest descent method, we know both the direction and how much we should move. Therefore, in Newton's method, there is no need for a linear search to determine the step size. In 3D tomography, the Hessian matrix is too big to store in memory and virtually impossible to invert (Fichtner, 2010). However, as most of the energy is concentrated in the main diagonal, what is usually done is replacing the Hessian with an approximation of the diagonal terms (Luo et al., 2013). In the 1D case, it is possible to calculate the true Hessian. To do that, we can use approximations involving finite differences again. For derivatives along the main diagonal, we use the approximation:

$$\frac{\partial^2 \chi}{\partial c_i^2}(c) \approx \frac{\chi(c_i + \epsilon) - 2\chi(c_i) + \chi(c_i - \epsilon)}{\epsilon^2}, \quad (4.13)$$

and, for the other terms of the matrix, we have:

$$\frac{\partial^2 \chi}{\partial c_i \partial c_j}(c) \approx \frac{\chi_{++} - \chi_{+-} - \chi_{-+} + \chi_{--}}{4\epsilon^2}, \quad (4.14)$$

where

$$\begin{aligned} \chi_{++} &= \chi(c_i + \epsilon, c_j + \epsilon), \\ \chi_{+-} &= \chi(c_i + \epsilon, c_j - \epsilon), \\ \chi_{-+} &= \chi(c_i - \epsilon, c_j + \epsilon), \\ \chi_{--} &= \chi(c_i - \epsilon, c_j - \epsilon). \end{aligned} \quad (4.15)$$

Even in one dimension, when the number of samples is large, the processing time quickly becomes prohibitive.

4.1.6 Conjugated Gradient Method

The convergence of the steepest descent method is slow (Fig. 4.3), among other reasons, because the alternation of 90° in each new iteration ends up "undoing" part of the work done in the previous iteration. However, when we calculate a new gradient, we could use the direction of the previous gradient to avoid that. In other words, we calculate the

update direction by making a linear combination of two consecutive gradients. That is the idea behind the conjugate gradient method.

A detailed derivation of the method are found in Press et al. (2007), Menke (2012), and Schuster (2017). The basic idea is to perform the first update using the steepest descent method and, from the second iteration on, calculate the new directions using both the current and previous gradient information. The ideal steps can be determined using linear search. We see the result in Fig. 4.5. Although it required two iterations with linear search, the convergence showed a significant improvement compared to the steepest descent method, without the need to calculate and invert the Hessian as in Newton’s method.

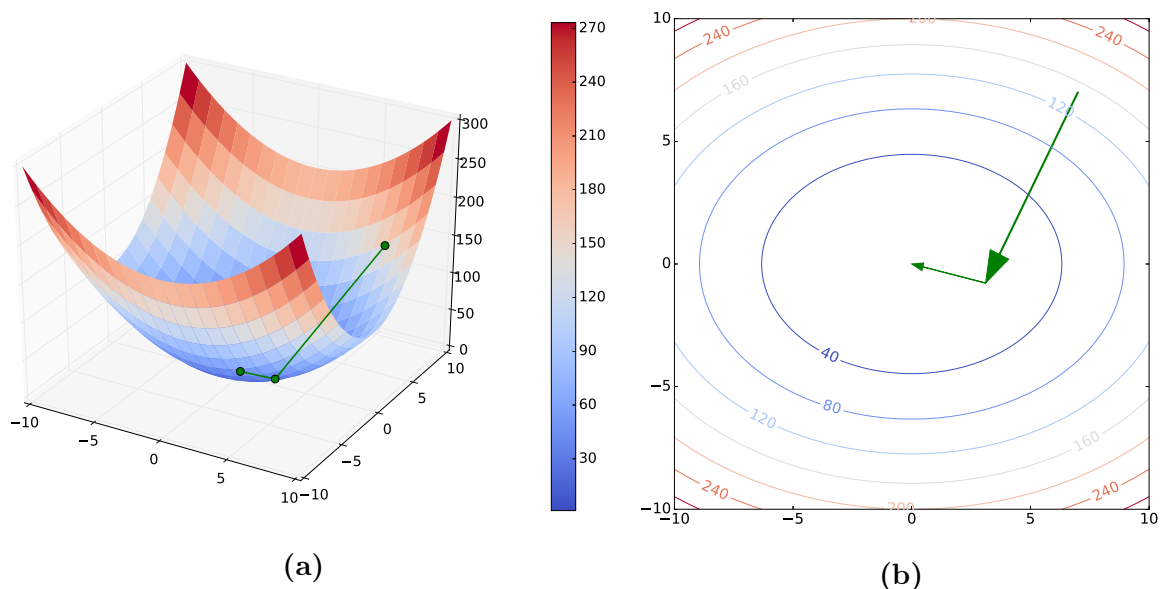


Figure 4.5: Left: 3D view of convergence to the minimum using the conjugate gradient method. Right: same process represented in 2D with contour lines.

The conjugate gradient method with preconditioning by the pseudo-Hessian is one way to optimize the model parameters in FWI (Chapter 8). Preconditioning is the process in which we divide the gradient by the modulus of the pseudo-Hessian. Given that the modulus of the pseudo-Hessian is related to the data coverage (Luo et al., 2013), the preconditioning balances the gradient and helps to compensate for any unevenness in the data distribution, speeding up the converge.

Chapter 5

Adjoint-state Method

In the last chapter, we saw many ways to speed up convergence in optimization problems, reducing the number of iterations needed. However, a bottleneck remains in the computation of the gradient. In the one-dimensional case, we can calculate it in a few minutes using finite differences, but in three dimensions, it would take years even on a supercomputer. Furthermore, as mentioned, determining an ideal value for ϵ is not trivial. To apply FWI to real problems, we need a more efficient way of calculating the gradient. The adjoint-state method (or simply "adjoint method") allows us to do just that.

The adjoint method is an efficient way of calculating the gradient in certain categories of problems, including the one we are studying (Plessix, 2006). It reduces the number of simulations from $2n$ to 2, a performance gain of 10^3 times, in our example, and 10^9 times, in the three-dimensional case. In other words, it limits the number of simulations required to only two, regardless of the number of parameters of the problem. The idea of how this is done is better understood by thinking in terms of matrices, the discrete equivalents of linear operators. Let an $n \times n$ square matrix A be multiplying an $n \times 1$ matrix u , resulting in an $n \times 1$ matrix f :

$$Au = f. \tag{5.1}$$

Another $n \times n$ square matrix B is multiplying an $n \times 1$ matrix v , resulting in an $n \times 1$ matrix g :

$$Bv = g. \tag{5.2}$$

For now, we do not assign any physical meaning to the variables in question. Let us think only in terms of vector and matrix multiplications only. The question that we are interested in answering is what should be the relationship between matrices A

and B so that the dot product between f and v is equal to the dot product between u and g ? In matrix notation:

$$f^T v = u^T g \implies (Au)^T v = u^T Bv, \quad (5.3)$$

where

$$f^T v = \sum_{i=1}^n f_i v_i, \quad u^T g = \sum_{i=1}^n u_i g_i, \quad Au = \sum_{j=1}^n A_{ij} u_j, \quad \text{and} \quad Bv = \sum_{j=1}^n B_{ij} v_j. \quad (5.4)$$

By combining 5.3 and 5.4, we find:

$$\sum_{i=1}^n \sum_{j=1}^n A_{ij} u_j v_i = \sum_{i=1}^n \sum_{j=1}^n B_{ij} u_i v_j. \quad (5.5)$$

It is easy to see that Eq. 5.5 is satisfied when $B_{ij} = A_{ji}$, i.e. when B is the transpose of A :

$$B = A^T. \quad (5.6)$$

Similarly, in the continuous space of functions, given a linear operator L applied to a function u , and a second linear operator M applied to a function q , we wish to find the relationship between L and M such that:

$$\langle Lu \cdot q \rangle = \langle u \cdot Mq \rangle. \quad (5.7)$$

The operator M that satisfies 5.7 is called the adjoint operator of L :

$$M = L^\dagger. \quad (5.8)$$

Following the example of (Giles and Pierce, 2000), suppose we want to calculate the result of $f^T v$. We could calculate the result of that multiplication as $u^T g$, where:

$$A^T v = g. \quad (5.9)$$

That is, instead of solving the system of equations 5.1 and then calculate the product $f^T v$, we could solve the system 5.9 and calculate the product $u^T g$. At first glance, there is no gain in that change. In fact, for a single f and a single g , the

computational cost of calculating $f^T v$ and $u^T g$ is exactly the same. However, what if we wanted to calculate that result for n values of f and m values of g ? In this case, we could choose between n operations using the original formulation of Eq. 5.1, or m operations using the alternative form of the problem given by 5.9. When the dimension of the system is large, the cost of the dot products becomes much less than that of solving the linear system, and the alternative problem is much faster than the original whenever $n \gg m$. The alternative approach, known as *dual problem*, requires the use of the transposed matrix A^T . When dealing with operators, the equivalent of transposing an operator L is the so-called adjoint operator L^\dagger .

5.1 Adjoint Method in Seismic Tomography

In our case, the goal is to find the gradient of the objective function given by Eq. 4.4, which is:

$$\frac{\partial \chi}{\partial c}(c) = - \int_0^T [d(t) - s(t, c)] \frac{\partial s}{\partial c}(t, c) dt, \quad (5.10)$$

where the derivative $\frac{\partial s}{\partial c}(t, c)$ is the most expensive term to calculate in 5.10.

Since the synthetic seismogram $s(t, c)$ is just the wavefield $w(x, t)$ sampled by a Dirac delta at the position of the receiver, we can think in terms of $w(x, t)$ instead of working with $s(x, t)$. Following Demanet (2015), when applying the chain rule, we write the gradient as the following inner product:

$$\frac{\delta \chi}{\delta m} = \left\langle \frac{\delta \chi}{\delta w} \cdot \frac{\delta w}{\delta m} \right\rangle, \quad (5.11)$$

where we replaced the derivative with respect to $c(x)$ by the derivative in $m(x) = \frac{1}{c(x)^2}$ and changed the notation of the differentials to facilitate the derivation. Again, the costly operation is to calculate $\frac{\partial w}{\partial m}$, and we use the adjoint method to get rid of that term. We begin by differentiating 4.1 with respect to m :

$$\frac{\partial L}{\partial m} w + L \frac{\partial w}{\partial m} = 0 \implies L \frac{\partial w}{\partial m} = - \frac{\partial L}{\partial m} w. \quad (5.12)$$

The term $\frac{\partial L}{\partial m}$ is the derivative of the operator $L = [m(x) \partial_{tt} - \partial_{xx}]$ with respect to $m(x)$, which is nothing but the operator ∂_{tt} . Thus:

$$L \frac{\partial w}{\partial m} = - \frac{\partial^2 w}{\partial t^2}. \quad (5.13)$$

Calculating the second time derivative of the wavefield $w(x, t)$ is simpler and more efficient than calculating $\frac{\partial w}{\partial m}$. The only problem is that Eq. 5.13 gives us a simple way of calculating $L \frac{\partial w}{\partial m}$, not $\frac{\partial w}{\partial m}$. Inverting the operator L is not an option in this case, but the properties of the adjoint operator give us the solution.

Be $q(x, t)$ the solution of:

$$L^\dagger(m) q = \frac{\delta \chi}{\delta w}, \quad (5.14)$$

where L^\dagger is the adjoint operator of $L = [m(x) \partial_{tt} - \partial_{xx}]$. While L is an operator that describes a wave propagating from the past to the future, L^\dagger describes a wave traveling from the future to the past. In this case, we rewrite 5.11 as:

$$\frac{\delta \chi}{\delta m} = \left\langle L^\dagger(m) q \cdot \frac{\delta w}{\delta m} \right\rangle = \left\langle q \cdot L \frac{\delta w}{\delta m} \right\rangle = \left\langle -q \cdot \frac{\partial^2 w}{\partial t^2} \right\rangle = - \int_0^T q \frac{\partial^2 w}{\partial t^2} dt, \quad (5.15)$$

where function $q(x, t)$ is called adjoint wavefield.

The above derivation is not rigorous and does not reveal all the nuances of the problem, but it does help us understand the general idea. For more details, the reader should refer to Fichtner (2010) and Demanet (2015). The physical interpretation is more intuitive than the mathematical derivation, though. We begin solving the forward problem using the starting model $m = \frac{1}{c^2}$, which allows us to calculate the wavefield $w(x, t)$ and the synthetic seismogram $s(t, c)$. The function $w(x, t)$ must be stored in memory for later use in calculating the gradient (Eq. 5.16):

$$\frac{\delta \chi}{\delta m}(m) = - \int_0^T q(x, t) \frac{\partial^2 w}{\partial t^2}(x, t) dt. \quad (5.16)$$

Then, we calculate the so-called adjoint source $f^\dagger(x_r, t)$ (Eq. 5.17), which is nothing but the time-reversed residue between the observed and the synthetic seismograms. Reverting the adjoint source and performing the simulation with the operator



Figure 5.1: The source (red star) is located at x_s and the receiver (yellow triangle) is located at x_r .

L is equivalent to not reversing it and running the simulation with the adjoint operator L^\dagger . This allows us to use the same code for forward and the adjoint simulations. We just have to remember to reverse the adjoint field $q(x, t)$ in time before applying Eq. 5.16.

$$f^\dagger(x_r, t) = -[d(T - t) - s(T - t, m)] \delta(x - x_r), \quad (5.17)$$

remembering that T is the total duration of the seismogram.

An important remark is that while the seismic source of the forward problem is at $x = x_s$, the adjoint source must be positioned at the receiver's location $x = x_r$ (Fig. 5.1). That is, we calculate the forward wavefield using the source located at $x = x_s$ and the propagation occurs in the normal direction of time. The adjoint field, on the other hand, is calculated by injecting the residue (difference between the observed and the synthetic data) at the receiver's location $x = x_r$ and running the simulation backward in time (or by injecting the time-reversed residue and running the simulation forward in time). The correlation between the second time derivative of the forward field and the adjoint field is what gives us the gradient in Eq. 5.16. To update $c(x)$ using this gradient, we use Eq. 5.18:

$$c_{k+1}(x) = \sqrt{\frac{1}{\frac{1}{c_k(x)^2} + \alpha \frac{\delta\chi}{\delta m}(m)}}, \quad (5.18)$$

where α is the step.

We can also use the adjoint method to calculate Hessian (Fichtner, 2010; Demanet, 2015). Nonetheless, as mentioned, the most common approach is calculating only an approximation for its diagonal, according to Luo et al. (2013) and given by Eq. 5.19:

$$\frac{\delta\chi^2}{\delta^2m}(m) \approx - \int_0^T \frac{\partial^2 q}{\partial t^2}(x, t) \frac{\partial^2 w}{\partial t^2}(x, t) dt. \quad (5.19)$$

The pseudo-Hessian resembles the shape of the gradient but with the reversed sign. Since it contains negative and null values, it requires a water level and smoothing before being applied as a preconditioner.

Chapter 6

Sensitivity Kernels

In Chapter 4, we have seen that the sensitivity of the chosen objective function (misfit measurement) to a perturbation in each model parameter directly determines the gradient required to update the model. Chapter 5 shows how the adjoint state method allows for efficient calculation of that gradient, accelerating the computational time needed to calculate it by many orders of magnitude, making FWI a feasible option for carrying out tomographic studies.

In the infinite-frequency regime, the sensitivity for a given source-receiver pair is all along the ray path connecting them, found by using Snell's Law. Nonetheless, since seismic waves always carry a finite amount of energy, their frequency can never be infinity. Taking finite-frequency effects into account requires a more complex theory that many students (myself included) struggle to understand. This chapter contains a manuscript about to be submitted to the *European Journal of Physics*, whose purpose is addressing that problem, providing a simplified explanation of the finite-frequency theory and the so-called *sensitivity kernels*, which is a fundamental part of the FWI theory.

Sensitivity Kernels in Seismic Wave Propagation: A Simplified Explanation for the Banana-doughnut Paradox

Caio Ciardelli

Department of Geophysics

Institute of Astronomy, Geophysics and Atmospheric Sciences

University of São Paulo - Brazil

June 10, 2021

Abstract

Ray theory, a high-frequency approximation to describe wave propagation, has been a cornerstone in seismology for over a hundred years. Despite its simplicity and wide range of applications, some limitations combined with the ever-increasing computational power motivated the development of finite-frequency theory, a better model to describe how the Earth's inner structure affects seismic waves. Finite-frequency theory has matured a lot in the last decades, and it is now widely applied in many geophysical problems. However, most students and even some experienced researchers face difficulties understanding it. An appropriate theoretical comprehension is paramount to make the most out of the methods a theory underpins, avoid pushing it beyond its limits, and further develop it. With that problem in mind, this paper shows a simplified formulation of the sensitivity kernels, which are the generalization of rays in the finite-frequency regime. The resultant model, despite its limitations, correctly predicts the main features of finite-frequency theory, including the zero sensitivity in the middle of the travel-time kernels, known as the *banana-doughnut paradox*, shedding new light on that intriguing phenomenon. A Colab Notebook implementing the main formulas accompanies the paper, allowing readers to interact and play with the results.

1 Introduction

The increasing levels of specialization and automation are two crucial factors that have allowed both the ever-growing accumulation of knowledge and the technological progress that have brought

countless benefits to humankind. In the last decades, science has experienced remarkable theoretical, experimental, and computational advances, and seismology is no exception. That enormous progress brought about the development of a wide range of methods, computational codes, and instruments which are now available to scientists. The development of high-level computational languages and the availability of many sophisticated computational packages made many activities very easy and productive, allowing us to perform complex tasks without the need for a deep theoretical understanding of them. That has obvious advantages but also some drawbacks.

Despite all the advances and undeniable benefits of computational developments and automation, many scientists fail to make the most out of them. Often, the reason is that they do not have a deep understanding of the methodologies, techniques, and codes they are using. To be always *reinventing the wheel* would be a waste of time. However, it could be useful to devote more time to understand the inner workings of the theories and methods on which we base our researches. That could not only lead to higher-quality science but also increase efficiency in the long run. Frequently, avoiding mistakes is less time-consuming than fixing them.

In geosciences, for example, one could mention the example of seismic tomography, a technique that uses seismic waves to image the Earth's interior. A variety of phenomena can create seismic waves, including earthquakes, asteroid impacts, and artificial explosions. Large earthquakes are the main source of seismic data for carrying out seismic tomography on a global scale. Provided that one knows the location and origin time of the earthquakes and that a network of seismographic stations is available, it is possible to use the record of the seismic waves at the receivers to infer the composition of the Earth's interior. The more earthquakes and stations spread over the surface of the planet, the better the result.

Historically, seismologists have been carrying out that process by using a velocity model, which is a simplified mathematical representation of the Earth. Using the velocity model, one computes the theoretical travel times of the seismic waves and compares them with the observed values recorded at the stations. If the predicted values match the observed ones, we know our model is a reasonable representation of the planet's interior. Otherwise, we iteratively update the model until the difference between them is minimum. Aki and Lee (1976); Aki et al. (1977) used teleseismic P-wave travel time residuals to determine the three-dimensional seismic structure of the lithosphere. Sengupta and Toksöz (1977) computed a 3D velocity model for the whole mantle combining P, S, and some PcP and ScS travel time measurements. Dziewonski et al. (1977) imaged the top 1,100 km of the mantle using P-wave residuals. Since then, many 1D reference Earth models. Since then, many many 1D reference Earth models (e.g., Dziewonski and Anderson, 1981; Kennet, 1991; Kennett et al., 1995) and 3D Earth models (e.g., Ritsema et al., 1999; Grand, 2002; Montelli et al., 2006; Houser et al., 2008; Ritsema et al., 2011; Schaeffer and Lebedev, 2013; French and Romanowicz, 2014; Koelemeijer et al., 2016; Durand et al., 2017; Lu et al., 2019; Lei, Wenjie and Ruan, Youyi and Bozdağ, Ebru and Peter, Daniel and Lefebvre, Matthieu and Komatitsch, Dimitri and Tromp, Jeroen and Hill, Judith and Podhorszki, Norbert and Pugmire, David, 2020) using multiple techniques and various types of data have been developed, such as body-wave travel times, surface-wave dispersion, Earth's normal modes, mass and moment of inertia, and waveform inversion.

In the early years of the tomographic techniques, the limited computational power required low-cost computational methods that used seismic ray theory. Seismic ray theory is an extremely

important method in seismology, consisting of a high-frequency approximation for the wave field, borrowed from optics, in which elementary waves, such as P, S, reflected, refracted, and transmitted waves, propagate along rays that can be handled independently (Cerveny, 2005). Ray theory allows for easy and efficient computation of travel times and amplitudes of seismic waves. However, it only works properly in smooth media, in which the dominant wavelength is considerably smaller than the dimensions of the inhomogeneities (Arora et al., 2011). It also fails to take scattering effects into account. The limitations of the seismic ray theory led to the development of the so-called finite-frequency theory in seismology. In that new theory, the high-frequency approximation is relaxed, and volumetric structures named *Fresnel volumes* (Brokešová, 2006), *Fréchet kernels* (Tarantola, 1987) or *sensitivity kernels* (Marquering et al., 1998) replace the infinitesimal seismic rays. The term "sensitivity", in this case, refers to the fact that these structures describe how waves "sense" the Earth's interior as they travel through it.

2 Sensitivity Kernels

Whenever one uses ray theory in seismic tomography, the implicit assumption is that any velocity anomaly not crossed by the ray will not affect the amplitude and the travel time of the recorded signal at the receiver since all the energy is assumed to travel along the ray. Another way of saying this is that all the amplitude and travel-time sensitivity lies along the ray path. Sensitivity kernels having a non-zero volume shows that, in the finite-frequency regime, structures outside the ray path can affect the seismograms. The sensitivity kernels have a volume proportional to the prevailing wavelength of the seismic waves. As one would expect, the shorter the wavelength, the thinner the kernels. As the wavelength approaches zero, they collapse into geometrical rays with infinitesimal width, thus recovering ray theory. According to Snieder (1999) and Hung et al. (2001), the kernels width scale with $\sqrt{\lambda L}$, where λ is the wavelength and L is the distance between the source and the receiver.

Besides the wavelength, the kernels also depend on the kind of wave (P, S, Pp, ScS, Rayleigh, Love, etc.) and the measurement. Amplitude measurements issue solid kernels whose shape resembles a banana. The maximum sensitivity occurs in the middle of the volume, where ray theory predicts the ray path. However, as first realized by Woodward (1992) and later by Marquering et al. (1999), travel-time measurements give origin to what seems to be a paradoxical structure: a hollow banana. That result, at first sight, seems so strange because it implies that right at the ray path, where one would expect the sensitivity to be maximum, the influence on the travel time of the waves is zero. Following Marquering et al. (1999), Bozdağ et al. (2011) and using Specfem 3D Globe (Komatitsch and Tromp, 2002a,b) we can compute very accurate amplitude and travel-time kernels (Fig. 1). The sensitivity kernels are also known as Fréchet kernels because they are a kind of Fréchet derivative. They show how the amplitude or the travel time varies with respect to an infinitesimal perturbation in a model parameter. We see in Fig. 1 that all the kernels are predominantly negative. That is because increasing the propagation velocity of the seismic waves has the double effect of reducing both the amplitudes and the travel times.

Ishimaru (1978) and Keiiti and Richards (1980) discuss the effects of scattering in wave propagation and show cross-sections of the sensitivity kernels, derived from Fresnel's wave theory.

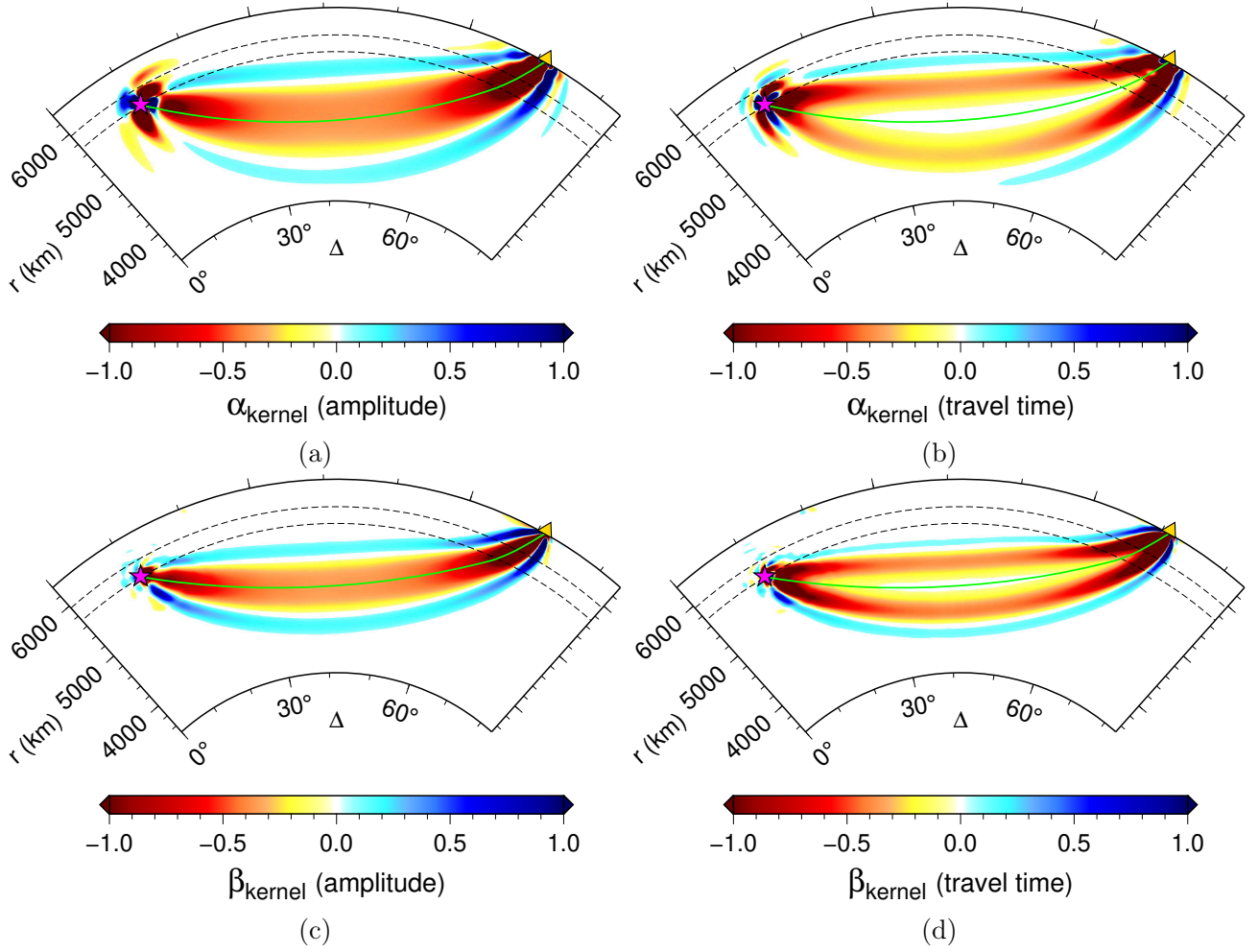


Figure 1: a) P-wave amplitude kernel; b) P-wave travel-time kernel; c) S-wave amplitude kernel; d) S-wave travel-time kernel. The amplitude kernels connecting the source (magenta star) and the receiver (yellow triangle) are solid, while the travel-time kernels are hollow, with zero sensitivity at the ray path (green line). The first and the second Fresnel zones are clear in all figures. The S-wave kernels (β) are narrower than the P-wave kernels (α) because the P-wave velocity is approximately $\sqrt{3}$ times larger than the S-wave velocity, resulting in a shorter wavelength. The velocity model is the isotropic version of PREM and the black dashed lines denote the 410 and 660 km seismic velocity discontinuities. All kernels are normalized.

Groenenboom and Snieder (1995) shows that the properties of the direct wave are determined by a weighted average over the first Fresnel zone for strongly scattering media. That result, in combination with the fact that in practical tomographic inversions, one always imposes some level of smoothness through model regularization, explains why in some cases ray theory gives correct results even for media where short-length-scale perturbations, smaller than the dominant wavelength, are present (Snieder and Lomax, 1996). Marquering et al. (1999) and Hung et al. (2001) explained why the travel-time kernels are hollow using scattering and a finite-frequency phenomenon known as *wavefront healing*. Spetzler and Snieder (2004) makes an intuitive review of the finite-frequency theory, remarking the limitations of ray theory and showing the connection

between the sensitivity kernels and the Kirchhoff integral. The authors highlight that the advantage of the sensitivity kernels is that the integration is performed over a volume in the medium, whereas the Kirchhoff integral is defined on a surface only.

Despite these excellent papers, many phenomena related to finite-frequency effects remain mysterious for many people working in the field. As it might be the case that navigating through a more accessible approach could lead to a better understanding of the described phenomenon, the creation of a kernel model that is both simple and representative of its main features could, hopefully, help shed some light on the topic. Developing and exploring such a model is the subject of the rest of this paper.

3 Basic Principles of Wave Propagation

To create a model, it is convenient to review two principles that apply to undulatory phenomena: Huygens' principle (also known as Huygens-Fresnel principle) and Fermat's principle.

3.1 Huygens' Principle

Huygens' principle states that every point on a wavefront is also a source of spherical wavelets that propagate in all directions. The result of the superposition and interference of all these wavelets is the wavefront itself. An intriguing question concerning the Huygens' principle then arises: *If every point on the wavefront creates a new wave that propagates in all directions, then why each wavefront does not create two new wavefronts: one moving outwards, away from the source, and another traveling inwards, towards it? Why do we only observe the outwards propagating wavefront?* Huygens' principle itself does not explain this fact, nevertheless, one can demonstrate that only the outwards propagating waves can exist by using the laws of conservation of energy and momentum.

As mentioned by Robinson and Clark (2017), that problem was later solved by Fresnel and Kirchhoff with a more advanced formulation of Huygens' principle, expressed in the Fresnel–Kirchhoff integral theorem. Explaining that theorem is beyond the scope of this paper, however, in short, it conveys that, although the new wavelets are spherical, their amplitude varies with the direction so that the amplitude of the resultant inward wavefront becomes zero. An equivalent way of stating this is that the inward propagating wavelets cancel each other out, resulting in a null inward wavefront. An alternative way of stating Huygens' principle to avoid that confusion is that each point on the wavefront is a source of hemispherical wavelets whose wavefronts point away from the source.

3.1.1 Huygens' Principle and Diffraction

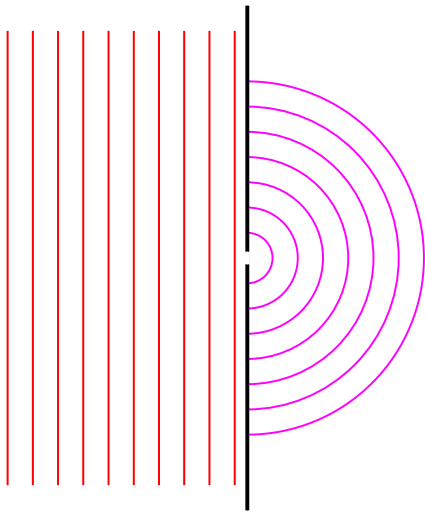
An easy way to visualize Huygens' principle is by observing diffraction phenomena. Suppose a plane wave is propagating from left to right and, eventually, the wavefront hits a wall with a single slit (Fig. 2). The wall will block the waves created on almost all points on the wavefront, except those located right at the slit position. With a single slit, only a spherical wavefront emerges on the other side of the wall. However, what happens when we add more slits to the experiment? With two or more slits, an interference pattern emerges. If we add many slits, the waves will interfere so that the resultant wavefront resembles the plane wave. With infinitely many slits, which is the equivalent of having no wall, we perfectly recover the plane wavefront.

A remark about Huygens' principle is that it does not apply when the number of spatial dimensions is even. In that case, the Green's function that solves the corresponding wave equation has a tail, which implies that the information on the wavefront is not enough to predict its future state. Under those circumstances, one must also consider all the information in the space between the source and the wavefront (Dai and Stojkovic, 2013).

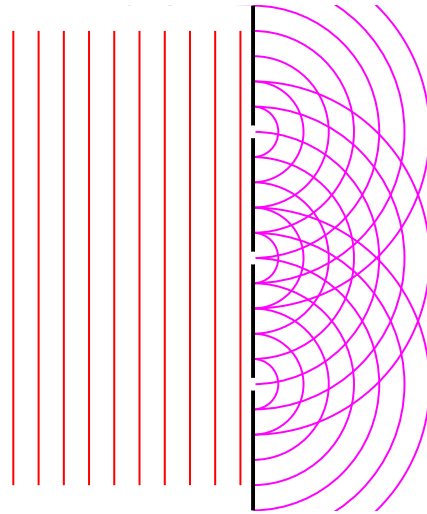
3.2 Fermat's Principle

According to Fermat's principle, a beam of light traveling from a point A to a point B always takes a path that is stationary with respect to small perturbations in the trajectory (Fig. 3). This principle is also commonly referred to as the principle of least time because the path of minimum time is the most common kind of stationary path. Nevertheless, maximum time paths or any other kind of stationary path are also possible according to Fermat's principle.

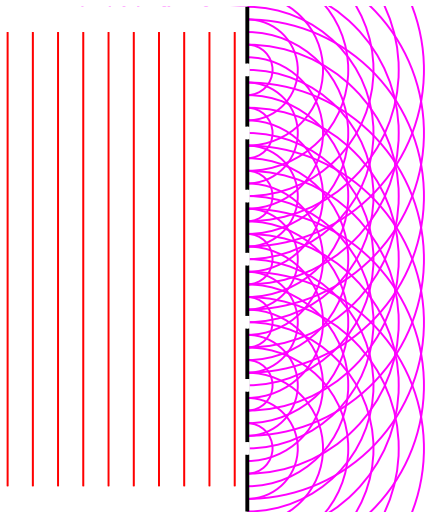
However, how do waves "know" which one is the path of the least time? In Feynman (2006), Richard Feynman explains why a photon bouncing off a mirror reflects with identical incidence and emergence angles (the path of the least time). He does so by using the quantum description of light, which allows the photon to simultaneously take all possible paths from the source to the receiver. He also shows that only the paths next to the one of the least time contribute to the resultant recorded signal. The key for understanding this is to realize that the travel time of a stationary path is invariant under small perturbations in the trajectory (Fig. 3). As a result, all paths close to that of the least time will arrive at the receiver in phase and constructively interfere. Because all the remaining paths will result in waves with randomly distributed phases, they will cancel each other out (Fig. 4). This provides some insight regarding the dependence of the kernels width on the wavelength. The larger the wavelength, the more the travel time can change before the waves get out of phase and destructively interfere. In the infinite-frequency regime, the wavelength is zero. In this case, a tiny difference in the travel time is enough to misalign the waves so that only the stationary path survives, recovering ray theory.



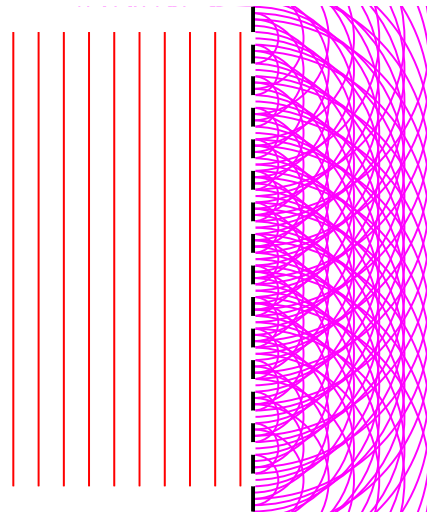
(a) One slit.



(b) Three slits.



(c) Seven slits.



(d) Fifteen slits.

Figure 2: Diffraction experiment using a different number of slits. As the number of slits tends to infinity, the interference pattern recovers the plane wave.

4 Model Formulation

We begin our simplified model of a sensitivity kernel by defining a convenient spherical coordinates system, shown in Fig. 5.

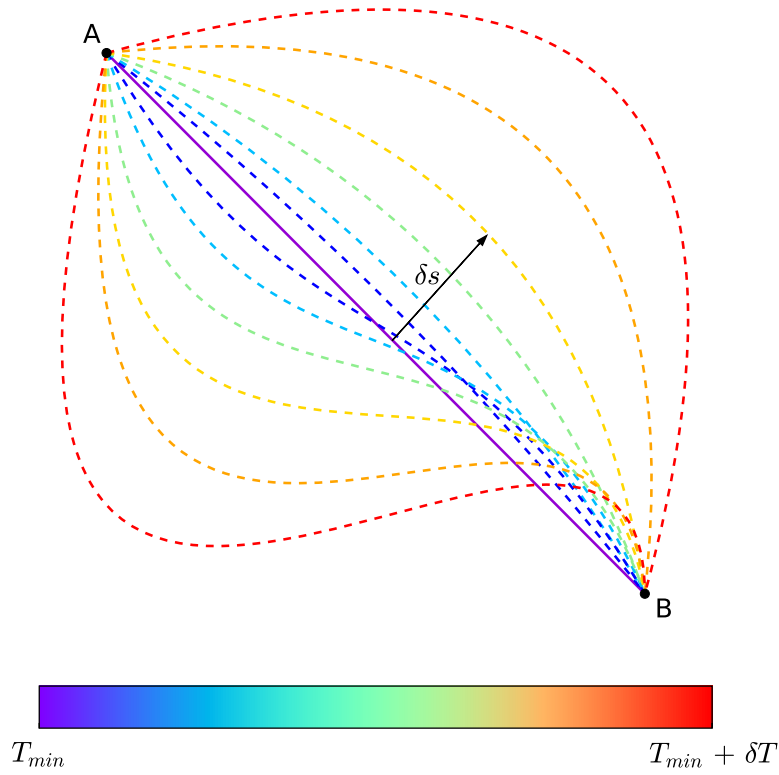


Figure 3: Illustration of Fermat's principle. In a homogeneous medium, the path of the least time for a wave created at point **A** and detected by a receiver at point **B** is a straight line (purple line). All the trajectories that are close to the least time trajectory (blue lines) have travel times very similar to the minimum. As the perturbation on the path gets larger (red lines), the travel time also increases.

A surface element in this coordinate system is given by:

$$dS = r^2 \sin \theta d\theta d\phi. \quad (1)$$

We can calculate the surface area of a sphere S_S with radius r by integrating over the domain $S = \{0 \leq \theta \leq \pi \text{ and } 0 \leq \phi \leq 2\pi\}$:

$$S_S = \int_S dS = \int_0^{2\pi} \int_0^\pi r^2 \sin \theta d\theta d\phi = r^2 \int_0^{2\pi} d\phi \left(-\cos \theta \Big|_0^\pi \right) = 4\pi r^2. \quad (2)$$

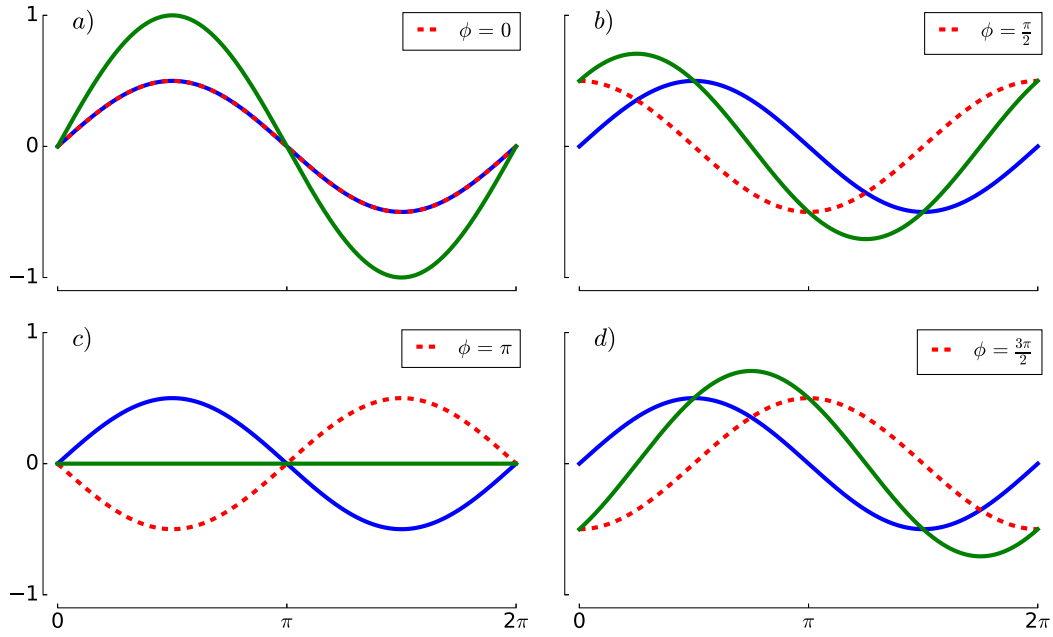


Figure 4: The interference can be constructive or destructive. *a)* When the phase shift (ϕ) between the two signals (blue line and dashed red line) is zero, there is complete construction (green line). *b)* When $\phi = \frac{\pi}{2}$, we have partial construction. *c)* When $\phi = \pi$, there is complete destruction. *d)* When $\phi = \frac{3\pi}{2}$, there is partial construction again.

4.1 Inverse-square Law

As a next step, we place a light source at the origin of the coordinate system. The total energy output per unit of time is E . We define the energy density on the wavefront (ε_S) by the energy output divided by the area of the wavefront, which is the surface of a sphere. From that result, it is easy to derive the inverse-square law. The inverse-square law says that the energy density on the wavefront in a three-dimensional space is inversely proportional to the distance from the source squared, as shown in Eq. 3 and in Fig. 6.

$$\varepsilon_S = \frac{E}{4\pi r^2} \quad \text{and} \quad \frac{\varepsilon_{S_2}}{\varepsilon_{S_1}} = \frac{dS_1}{dS_2} = \frac{r_1^2}{r_2^2} \quad (3)$$

In case we want to compute the energy contained in the spherical wavefront (E_S), all we need to do is integrate the energy density element $dE = \varepsilon dS$ over the surface S (see equation 4). The result, as expected, is exactly the total energy output of the source E .

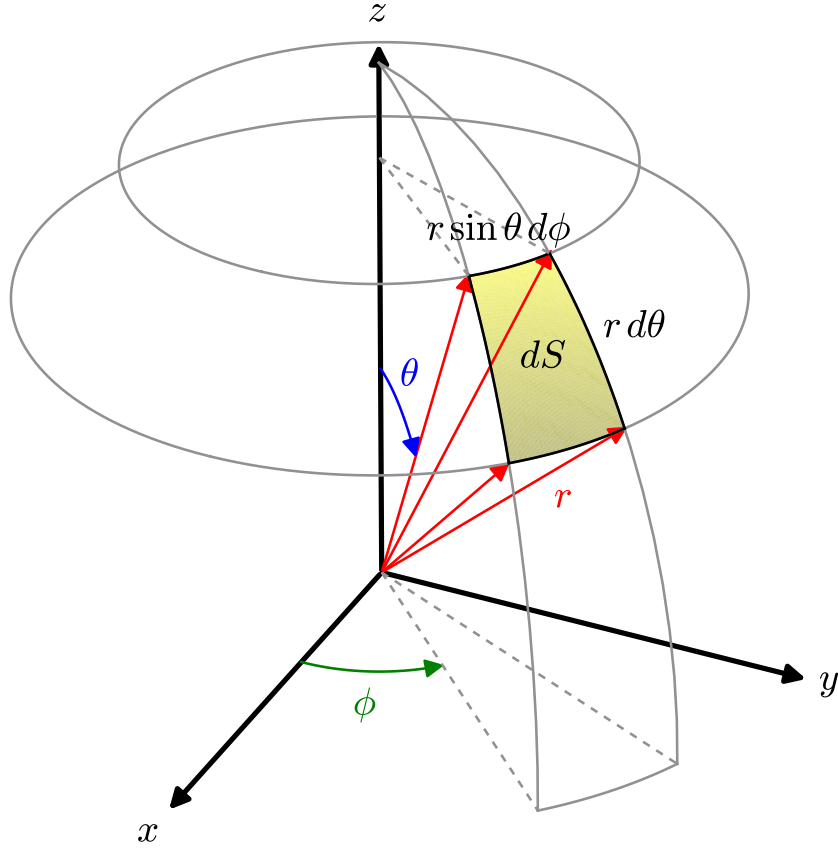


Figure 5: System of spherical coordinates in which r is the radius (red arrows), θ is the colatitude (blue curved arrow), and ϕ in the longitude (green curved arrow). The yellow region represents an infinitesimal element of area dS .

$$\begin{aligned}
 E_S &= \int_S dE = \int_S \varepsilon_S dS = \int_0^{2\pi} \int_0^\pi \varepsilon_S r^2 \sin \theta d\theta d\phi \\
 &= \frac{E}{4\pi r^2} \int_0^{2\pi} \int_0^\pi r^2 \sin \theta d\theta d\phi \\
 &= \frac{E}{4\pi r^2} \cancel{4\pi r^2} = E
 \end{aligned} \tag{4}$$

4.2 Energy Incident on a Virtual Plane

We proceed by setting a virtual plane at a distance r from the energy source, and defining a new coordinate system, as depicted in Fig. 7. In that case, how much energy per unity time does the plane receive? In other words, if that plane was a 100% efficient solar panel, how much energy would it generate?

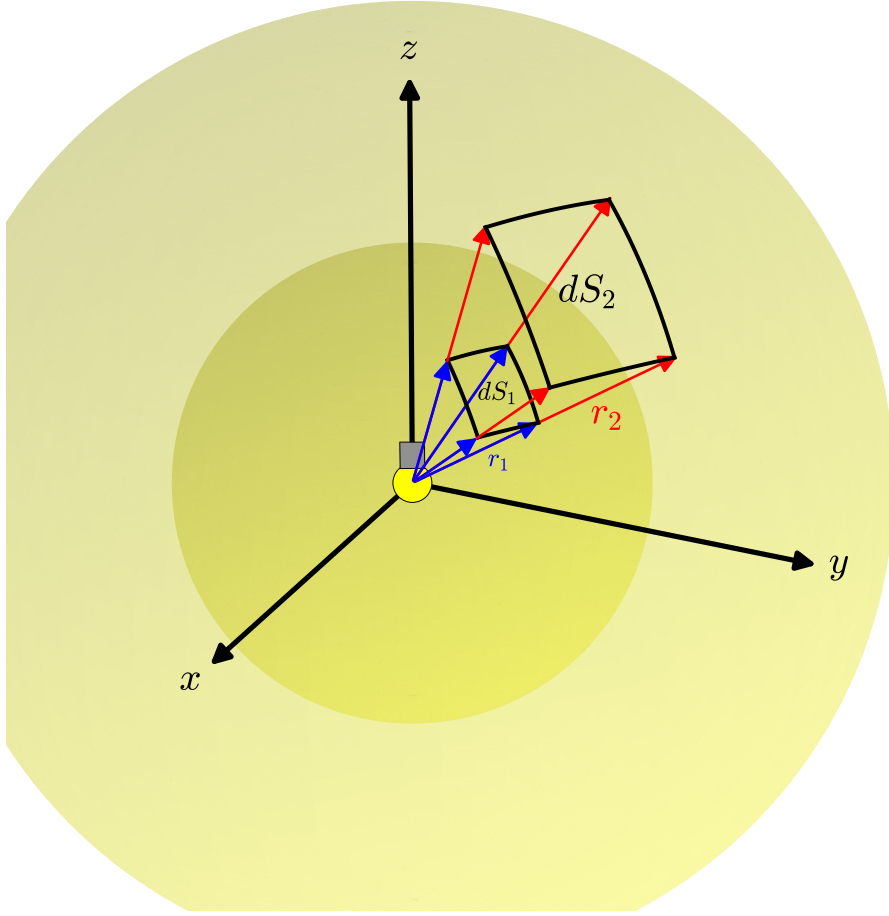


Figure 6: We represent energy source by a light bulb at the origin of the coordinate system. The energy density ratio between the two wavefronts is equal to the inverse ratio of their radii squared, as derived in equation 3.

Using the parametrization described in Fig. 7, we can calculate the total energy reaching the virtual plane by integrating all the light rings from $s = 0$ to infinity. The first thing we need is the energy density on each ring (ε_D). Unlike the previous case, the energy density on the surface of the plane is not constant. It not only decays with the square of the distance d but also depends on the incidence angle θ (Lambert's cosine law). Just as sunlight during winter spreads over a larger area due to a higher angle of incidence, the energy density on the disc decays with the cosine of θ . Thus, combining the geometrical spreading with the effect of the oblique incidence, we find:

$$\varepsilon_D(d) = \frac{E}{4\pi d^2} \cos \theta, \text{ where } \cos \theta = \frac{r}{\sqrt{r^2 + s^2}}. \quad (5)$$

We can rewrite Eq. 5 in terms of r and s :

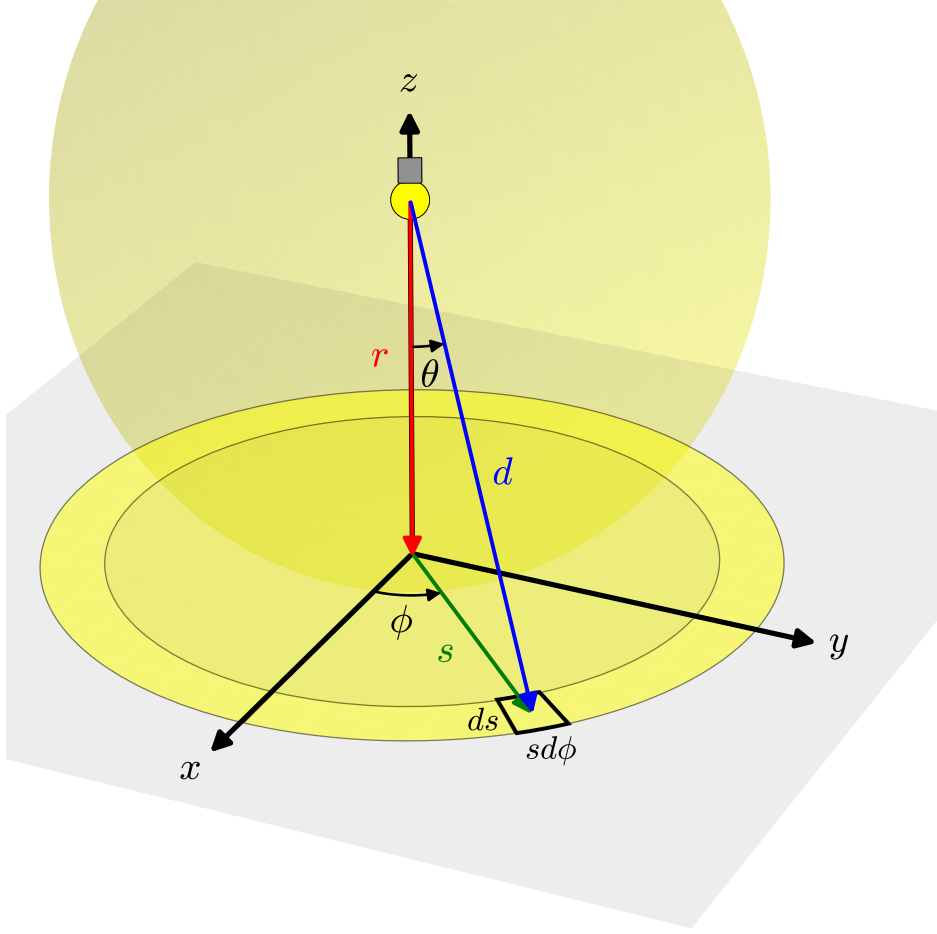


Figure 7: Schematic describing a plane located at a distance r from the light source. The energy first arrives at the closest point on the plane, corresponding to $s = 0$. As time progresses, more and more light reaches the plane as a series of concentric rings of width ds .

$$\varepsilon_D(r, s) = \frac{E}{4\pi(r^2 + s^2)} \frac{r}{\sqrt{r^2 + s^2}}. \quad (6)$$

Therefore, the total energy reaching the plane E_D is:

$$\begin{aligned} E_D &= \int_S dE = \int_S \varepsilon_D(r, s) dS = \int_0^{2\pi} \int_0^\infty \frac{E}{4\pi(r^2 + s^2)} \frac{rs}{\sqrt{r^2 + s^2}} ds d\phi \\ &= \frac{Er}{4\pi} \int_0^{2\pi} d\phi \int_0^\infty \frac{s}{(r^2 + s^2)^{\frac{3}{2}}} ds \\ &= \frac{Er}{4\pi} 2\pi \left(\frac{-1}{\sqrt{r^2 + s^2}} \Big|_0^\infty \right) \\ &= \frac{E}{2}. \end{aligned} \quad (7)$$

The result given by 7 makes complete sense. Since it is an infinite virtual plane, one would expect that exactly half of the energy output would eventually reach it.

4.3 Energy Density at the Receiver Using Huygens' Principle

With the previous results at hand, we can calculate the energy density at a receiver located at a distance $L = 2r$ from the source, as depicted in Fig. 8. All we need to do is apply Huygens' principle on the surface of the plane.

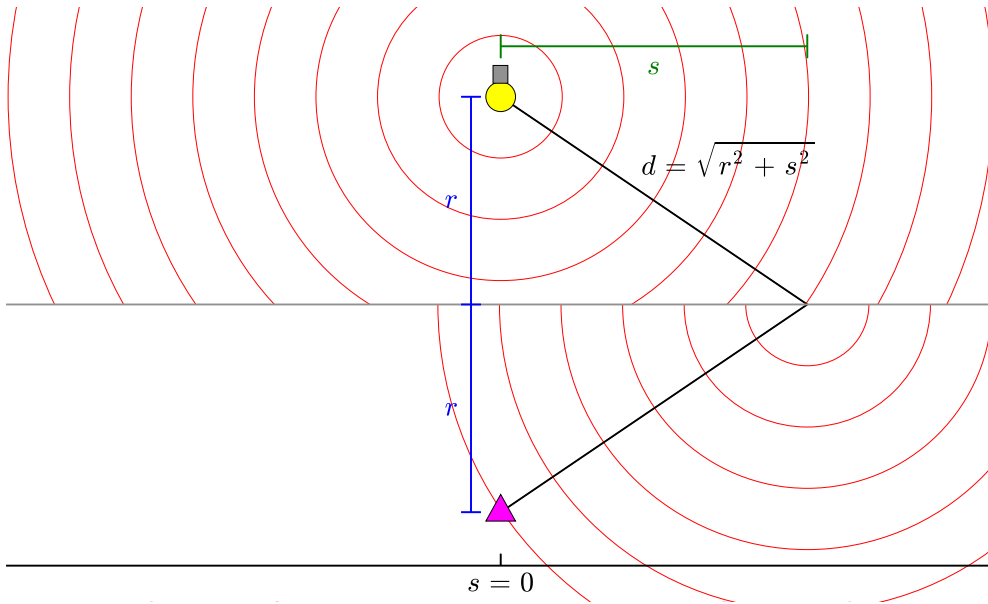


Figure 8: Schematic showing a receiver (magenta triangle) located on the other side of the virtual plane surface, diametrically opposite to the light source. The red lines represent the wavefronts.

As shown in Fig. 8, each point on the wavefront reaching the plane can be considered a new source of hemispherical waves (waves propagating back to the source do not exist, as previously pointed out). Thus, one can evaluate the "energy density density" at the receiver (ϵ_R) from the energy density on the plane (ϵ_D) by taking into account the geometrical spreading and the incidence angle once again:

$$\epsilon_R(r, s) = \frac{\epsilon_D(r, s)}{2\pi(r^2 + s^2)} \frac{r}{\sqrt{r^2 + s^2}} = \frac{E}{8\pi^2(r^2 + s^2)^2} \frac{r^2 s}{r^2 + s^2}. \quad (8)$$

Notice that we are assuming the hemispherical wavefronts of the Huygens' principle, which emerge when the original spherical waves reach the plane. Hence, the factor in the denominator of equation 8 is 2π instead of 4π . Therefore, to compute the energy density at the receiver, we must integrate the contributions of the entire plane:

$$\begin{aligned}
\varepsilon_R(r) &= \int_S dE = \int_S \epsilon_R(r, s) dS = \int_0^{2\pi} \int_0^\infty \epsilon_R(r, s) ds s d\phi \\
&= \int_0^{2\pi} \int_0^\infty \frac{E}{8\pi^2(r^2 + s^2)^2} \frac{r^2 s}{r^2 + s^2} ds d\phi \\
&= \frac{Er^2}{8\pi^2} \int_0^{2\pi} d\phi \int_0^\infty \frac{s}{(r^2 + s^2)^3} ds \\
&= \frac{Er^2}{8\pi^2} 2\pi \left[\frac{-1}{4(r^2 + s^2)^2} \Big|_0^\infty \right] \\
&= \frac{E}{4\pi(2r)^2}.
\end{aligned} \tag{9}$$

Equation 9 retrieves the result from 3, in which the energy density at a receiver located at a distance $L = 2r$ from the source is just the energy output of the source divided by the surface area of the sphere of radius L . We can also evaluate the energy density as a function of the integration radius s :

$$\begin{aligned}
\varepsilon_R(r, s) &= \int_0^{2\pi} \int_0^s \epsilon_R(r, \xi) d\xi u d\phi = \int_0^{2\pi} \int_0^s \frac{E}{8\pi^2(r^2 + \xi^2)^2} \frac{r^2 \xi}{r^2 + \xi^2} d\xi d\phi \\
&= \frac{Er^2}{8\pi^2} 2\pi \left[\frac{-1}{4(r^2 + \xi^2)^2} \Big|_0^s \right] \\
&= \frac{E}{4\pi(2r)^2} \left[1 - \frac{r^4}{(r^2 + s^2)^2} \right].
\end{aligned} \tag{10}$$

4.4 A Simplified Model of an Amplitude Sensitivity Kernel

We know that $\varepsilon \propto A^2$ where A is the amplitude of the wave (Igel, 2017). Therefore, from equation 10 we can derive:

$$A_R(r, s) = \frac{A_0}{4\pi(2r)} \sqrt{1 - \frac{r^4}{(r^2 + s^2)^2}}. \tag{11}$$

It is easy to see that when $s \rightarrow \infty$, Eq. 11 reduces to the far-field approximation for the amplitude. In the far field, the amplitude at the receiver decays linearly with the distance from the source $L = 2r$:

$$A_R(r) = \frac{A_0}{4\pi(2r)}. \tag{12}$$

By differentiating Eq. 11 and setting $A_0 = 4\pi$ (for simplicity), we find the expression that relates the differentials dA_R and ds . Eq. 13 tells us how much each ring of light on the surface S contributes to the amplitude recorded at the receiver:

$$dA_R = \frac{r^3 ds}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}}. \quad (13)$$

So far, we have taken into account the effects of the geometrical spreading, oblique incidence, and constructive interference in the amplitudes, but we have completely ignored the destructive interference. As previously mentioned, destructive interference is crucial to explain why the waves take only stationary paths. To include it in our model, we can insert a sine function into the differential:

$$dF = \frac{r^3 \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} ds, \quad (14)$$

where f is the frequency of the source and c is the velocity of propagation.

We assume a monochromatic source for simplicity, but since any pulse can be decomposed as a sum of sine waves using the Fourier transform, it is easy to see how that formulation could be generalized for a multi-frequency light source. The travel time between the source and the receiver

$$t_T(r, s) = \frac{2\sqrt{r^2 + s^2}}{c}, \quad (15)$$

introduces the phase factor

$$\phi(r, s) = \frac{-4\pi f \sqrt{r^2 + s^2}}{c}. \quad (16)$$

The sine function in equation 14 has the effect of halving amplitude and increasing the rate with which it decays with the distance, as it neutralizes the rays that take a longer path. It also makes the amplitude increase with the square root of the wavelength $\lambda = \frac{c}{f}$. To compensate

these side effects, one can introduce the correction factor $2\sqrt{\frac{r}{\lambda}}$ to renormalize the differential

$$dF^* = \frac{2}{\sqrt{\lambda}} \frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} ds. \quad (17)$$

By integrating from 0 to ∞ we find the resultant wave:

$$F^*(r, t) = \int_0^\infty dF^* = \frac{2}{\sqrt{\lambda}} \int_0^\infty \frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} ds. \quad (18)$$

Although we cannot express the integral function $F^*(r, t)$ (Eq. 18) as a combination of elementary functions, we know that, for given values of r , it consists of a weighted sum of sine functions with different phases but constant frequency. That necessarily results in another sine function with the same frequency. Knowing this, we can extract the amplitude of the resultant sinusoidal wave and eliminate its dependence on t by integrating function $F^*(r, t)$ squared from 0 to T , where $T = \frac{1}{f}$, and taking the square root of the result multiplied by $\frac{2}{T}$:

$$\begin{aligned} A_{R^*}(r) &= \sqrt{\frac{2}{T} \int_0^T [F^*(r, t)]^2 dt} \\ &= 2f \sqrt{\frac{2}{c}} \sqrt{\int_0^T \left[\int_0^\infty \frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} ds \right]^2 dt}. \end{aligned} \quad (19)$$

We can evaluate the integral above numerically, fixing $c = 1$ for simplicity. Fig. 9 shows the effect of the wavelength in the approximation by varying the values of f and Fig. 10 shows that our model is an excellent approximation for the far field whenever $2r \geq \lambda$. Now, let us explore how the amplitude values vary with s :

$$A_{R^*}(r, s) = 2f \sqrt{\frac{2}{c}} \sqrt{\int_0^T \left[\int_0^s \frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right) \right]}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}} d\xi \right]^2 dt}. \quad (20)$$

We are interested in the partial derivative of 20 with respect to s , which is the approximation for the amplitude kernel $K_A(r, s)$ using our model. Thus:

$$K_A(r, s) = \frac{\partial A_{R^*}}{\partial s}(r, s) = 2f \sqrt{\frac{2}{c}} \frac{n(r, s)}{m(r, s)}, \quad (21)$$

where

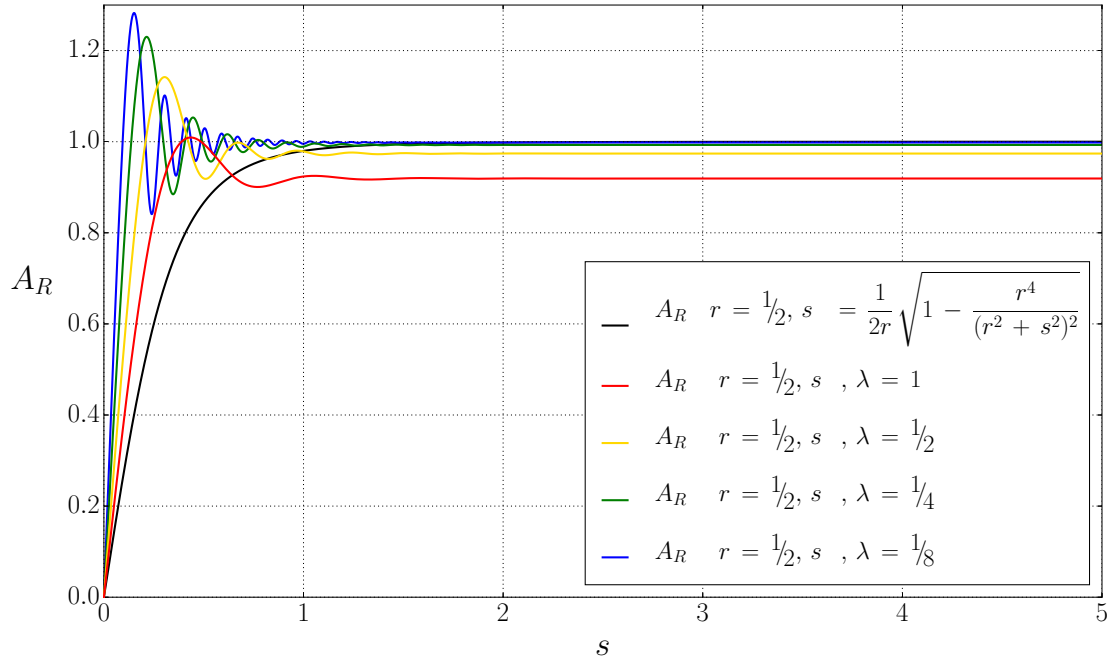


Figure 9: Amplitude as a function of s for different wavelengths. We set $r = 1/2$.

$$n(r, s) = \int_0^T \underbrace{\frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}}}_{\alpha(r, s, t)} \left\{ \int_0^s \underbrace{\frac{\sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right) \right]}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}}}_{\beta(r, s, t)} d\xi \right\} dt$$

and

$$m(r, s) = \sqrt{\int_0^T \left\{ \int_0^s \frac{\sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right) \right]}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}} d\xi \right\}^2 dt} = \frac{1}{2f} \sqrt{\frac{c}{2}} r^{-\frac{7}{2}} A_{R^*}(r, s). \quad (22)$$

Equation 21 is not defined for $s = 0$. A reasonable solution is replacing $K_A(r, 0)$ by the limit of $K_A(r, s)$ when s approaches zero (see appendix A for the complete derivation):

$$K_A(r, s) \Big|_{s=0} = \lim_{s \rightarrow 0} \frac{\partial A_{R^*}}{\partial s}(r, s) = \lim_{s \rightarrow 0} E(r, s) = \sqrt{\frac{2}{\lambda r^3}}, \quad (23)$$

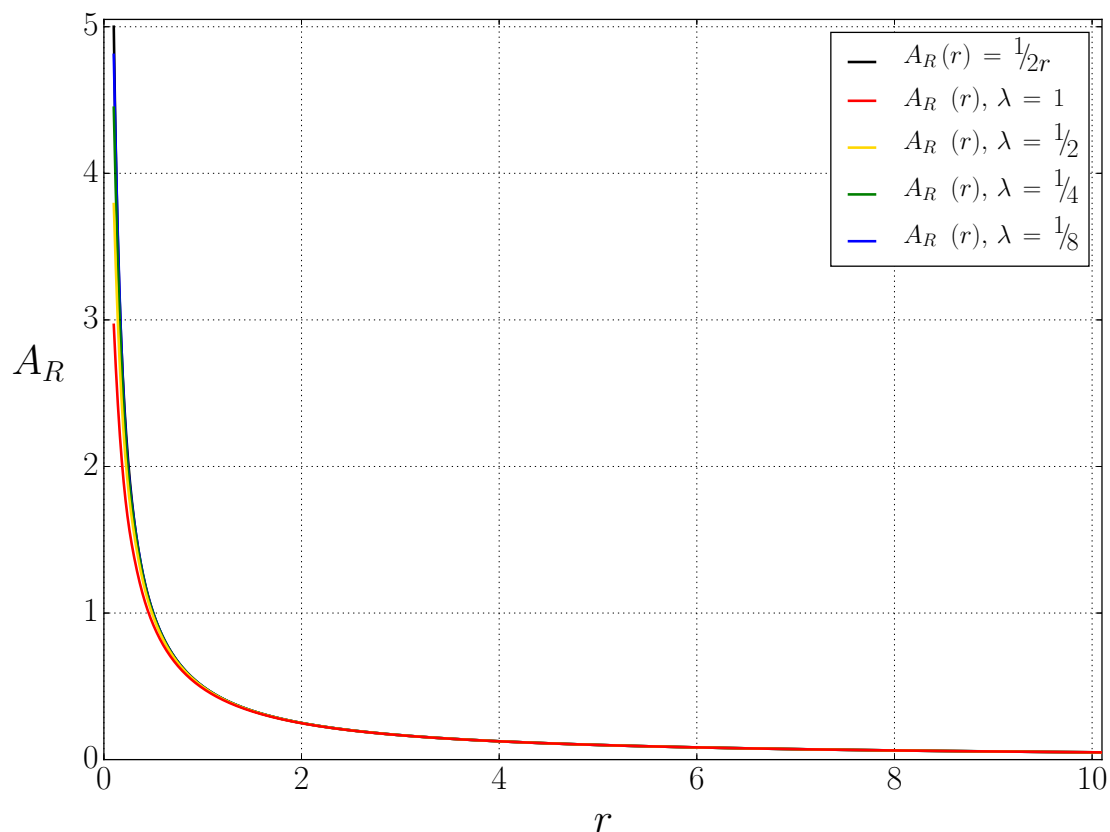


Figure 10: Amplitude decay with r for different values of λ . The shorter the wavelength, the more the decay rate approaches the linear decay expected in the far-field regime $A_R(r)$.

where

$$E(r, s) = \lim_{s \rightarrow 0} \frac{2r^{\frac{7}{2}}}{(r^2 + s^2)^2 \sqrt{\lambda(2r^2 + s^2)}}. \quad (24)$$

Therefore:

$$K_A(r, s) = \begin{cases} \sqrt{\frac{2}{\lambda r^3}} & \text{if } s = 0 \\ 2f \sqrt{\frac{2}{c}} \frac{n(r, s)}{m(r, s)} & \text{otherwise.} \end{cases} \quad (25)$$

The fact that all paths next to that of the least time have almost identical travel times implies that, for small values of s , the waves will arrive at the receiver with nearly the same phase and constructively interfere, rapidly increasing the amplitude of the detected signal. As s

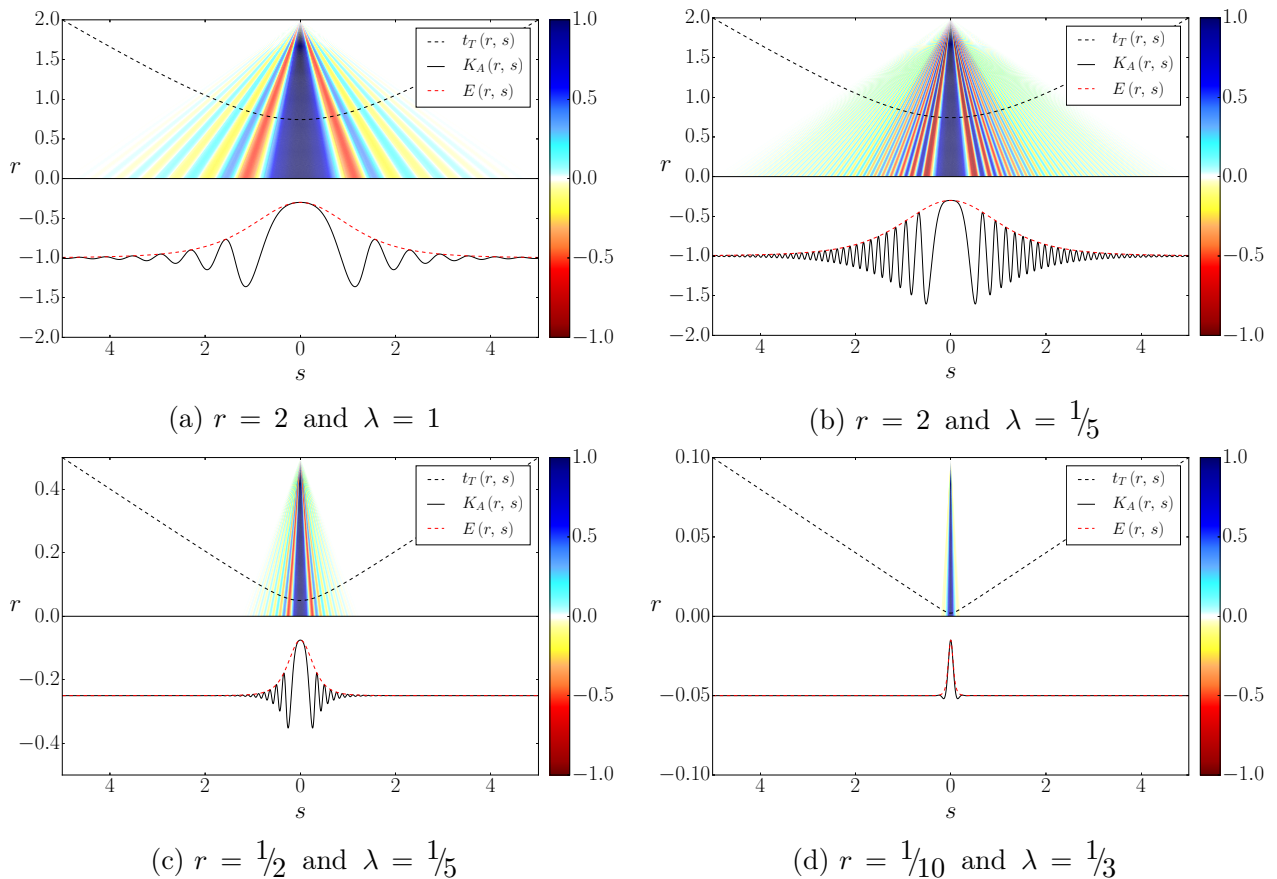


Figure 11: Upper half of the amplitude kernels for different values of r and λ . The bottom half is just the mirrored version of the upper one. The black dashed line shows the travel-time curve predicted using ray theory. The red dashed line shows function $E(r, s)$, which is the envelope of the amplitude sensitivity kernel $K_A(r, s)$. The color pallet describes the normalized values of $K_A(r, s)$, highlighting the Fresnel zones. We see that the kernels' width increases with $\sqrt{\lambda L}$.

increases, the waves begin to cancel each other out, resulting in the negative values of $K_A(r, s)$ in Fig. 11. From then on, we alternate between regions of constructive and destructive interference, known as Fresnel zones. However, the main contribution to the amplitude at the receiver, by far, comes from the first Fresnel zone, located around the ray path. That is the reason why the maximum amplitude sensitivity is at the ray path.

4.5 A Simplified Model of a Travel-time Sensitivity Kernel

So far, we have been studying the effect of the ray paths on the amplitude of the wave detected at the receiver. Now we continue to evaluate how much each path contributes to the travel time of the resultant wave $F^*(r, t)$. The phase is related to the travel time by the opposite of the angular velocity $-\omega = -2\pi f$. Equations 15 and 16 give the travel time and phase of each ray path. However, we want the travel time of the detected wave, resultant from the interference of all possible paths. Remembering again that, for a given r , 18 is a sinusoidal wave, one can extract

its phase information by fixing the value of $t = t_i$, normalizing $F^*(r, t_i)$ using function $A_R^*(r)$, and taking the arcsine of the result. For t_i , it is convenient to use:

$$t_i = \left(\frac{2r}{\lambda} \bmod 1 \right) T. \quad (26)$$

Whenever the distance between the source and the receiver is a multiple of the wavelength, t_i will be zero. Otherwise, t_i will be a fraction of the the period T corresponding to the remainder of the integer division of the distance $2r$ by the wavelength λ . That guarantees that we are always evaluating $F^*(r, t)$ at the beginning of the cycle, regardless of the distance from the receiver. On the other hand, after the normalization by $A_R^*(r)$, $F^*(r, t_i)$ will be constrained on the interval $[-1, 1]$, which is the domain of the arcsine function. Now, we convert the phase information to travel time by dividing the result by $-\omega$ and adding $\frac{2r}{c}$, which is the minimum travel time from the source to the receiver. Putting everything together, we find the expression for the travel time $T(r)$ to be:

$$T_T(r) = \frac{2r}{c} - \frac{1}{2\pi f} \arcsin \left[\frac{F^*(r, t_i)}{A_R^*(r)} \right]. \quad (27)$$

Analogously to 20, we can also define T_T as a function of s (Fig. 12):

$$T_T(r, s) = \frac{2r}{c} - \frac{1}{2\pi f} \arcsin \left[\frac{F^*(r, s, t_i)}{A_R^*(r, s)} \right], \quad (28)$$

where

$$F^*(r, s, t) = \frac{2}{\sqrt{\lambda}} \int_0^s \frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + \xi^2}}{c} \right) \right]}{(r^2 + \xi^2)^2 \sqrt{2r^2 + \xi^2}} d\xi. \quad (29)$$

Equation 28 is not defined at $s = 0$ as the factor $\frac{F^*(r, s, t_i)}{A_R^*(r, s)}$ results in the indeterminate form $\frac{0}{0}$. Once again, we replace the value of $T_T(r, 0)$ by adopting the limit of $T_T(r, s)$ for $s \rightarrow 0$ (see the complete derivation in appendix B). Therefore:

$$T_T(r, s) = \begin{cases} \frac{2r}{c} & \text{if } s = 0 \\ \frac{2r}{c} - \frac{1}{2\pi f} \arcsin \left[\frac{F^*(r, s, t_i)}{A_R^*(r, s)} \right] & \text{otherwise.} \end{cases} \quad (30)$$

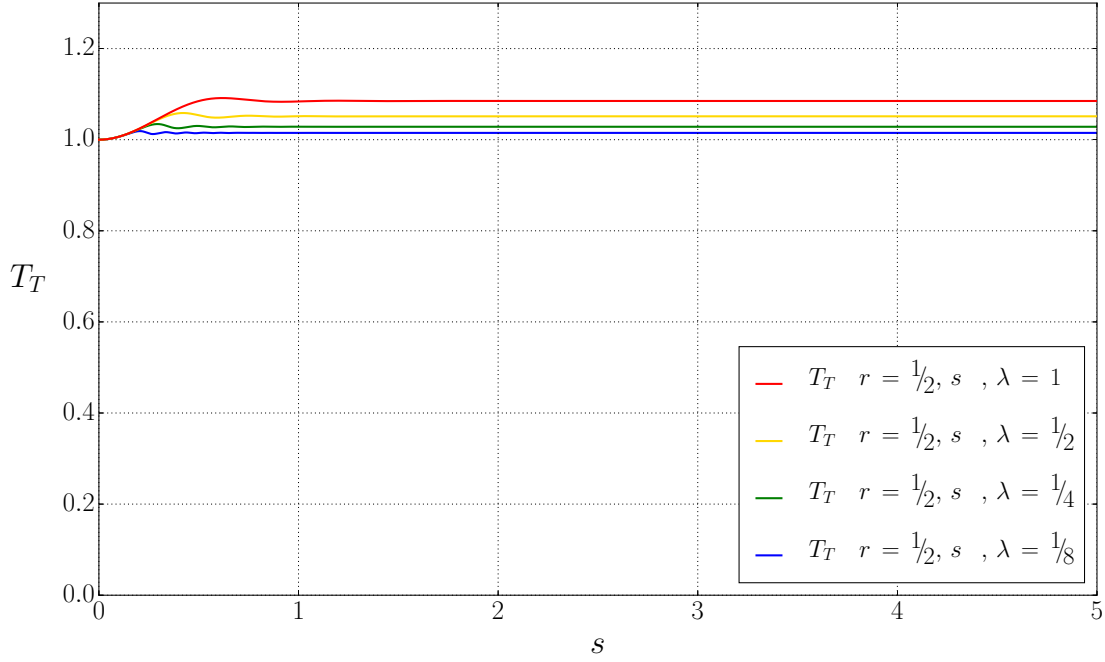


Figure 12: Travel time as a function of s for different wavelengths. We set $r = \frac{1}{2}$.

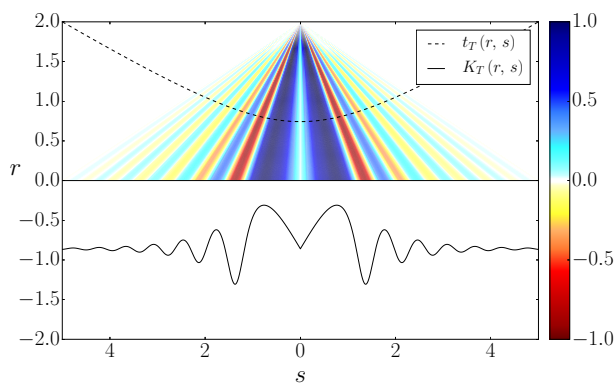
The travel-time kernel $K_T(r, s)$ is given by the partial derivative of 12 with respect to s :

$$\begin{aligned}
 \frac{\partial T_T}{\partial s}(r, s) &= -\frac{1}{2\pi f} \frac{1}{\sqrt{1 - \left[\frac{F^*(r, s, t_i)}{A_R^*(r, s)} \right]^2}} \frac{\partial}{\partial s} \left[\frac{F^*(r, s, t_i)}{A_R^*(r, s)} \right] \\
 &= \frac{1}{2\pi f} \frac{F^*(r, s, t_i) \frac{\partial A_R^*}{\partial s}(r, s) - \frac{\partial F^*}{\partial s}(r, s, t_i) A_R^*(r, s)}{A_R^*(r, s)^2 \sqrt{1 - \left[\frac{F^*(r, s, t_i)}{A_R^*(r, s)} \right]^2}}
 \end{aligned} \tag{31}$$

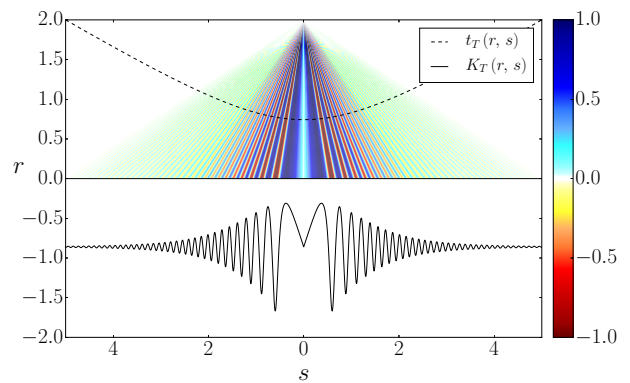
Once again, we found an expression that is not defined for $s = 0$. However, in Fig. 12 it is easy to see that as s approaches zero, the derivative also approaches zero. That can be formally proved by computing the limit of 31 for $s \rightarrow 0$, as we did for the former expressions. Nevertheless, the complete derivation is too extensive and falls outside the scope of this paper. Thus:

$$K_T(r, s) = \begin{cases} 0 & \text{if } s = 0 \\ \frac{1}{2\pi f} \frac{F^*(r, s, t_i) \frac{\partial A_R^*}{\partial s}(r, s) - \frac{\partial F^*}{\partial s}(r, s, t_i) A_R^*(r, s)}{A_R^*(r, s)^2 \sqrt{1 - \left[\frac{F^*(r, s, t_i)}{A_R^*(r, s)} \right]^2}} & \text{otherwise,} \end{cases} \quad (32)$$

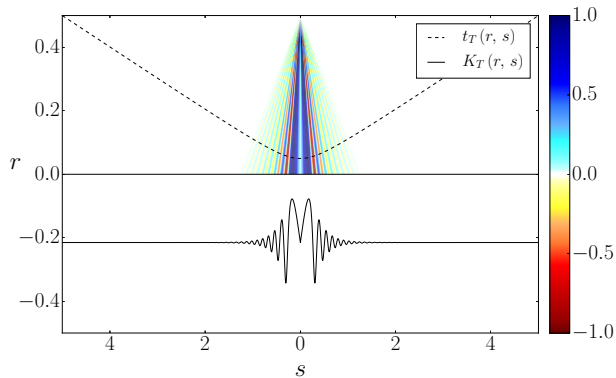
where t_i is given by Eq. 26.



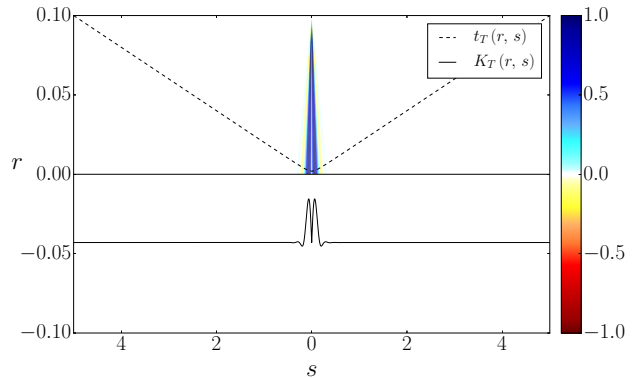
(a) $r = 2$ and $\lambda = 1$



(b) $r = 2$ and $\lambda = 1/5$



(c) $r = 1/2$ and $\lambda = 1/5$



(d) $r = 1/10$ and $\lambda = 1/3$

Figure 13: Upper half of the travel-time kernels $K_T(r, s)$ for different values of r and λ . The meaning of the symbols and colors is the same as the one in Fig. 11, with the difference that there is no equation describing the envelope. In the middle of the kernels, we observe a region of zero sensitivity.

In the previous section, the fact that the paths next to the stationary one had nearly identical travel times, making them arrive in phase, explained why the amplitude sensitivity grows rapidly for small values of s . However, for the very same reason, the travel-time sensitivity in this region is minuscule. When all waves arrive aligned, each new wave does not change the phase of the resultant wave, explaining the zero sensitivity at the ray path. As s increases, the new

waves begin to arrive with a progressively larger phase delay. As a result, we enter the destructive interference zone, gradually reducing the amplitude and moving the resultant wave backward. In that region, the travel-time sensitivity reaches its maximum. As s increases even further, though, the phase difference between the resultant wave and the new arriving waves becomes larger than π . When that happens, the new waves begin to pull the resultant wave back to its original phase, originating the negative values of $K_T(r, s)$. From then on, the cycle repeats over and over again, originating the higher-order Fresnel zones.

It is important to realize that the larger the phase difference between each new wave and the existing resultant (up to π), and the smaller the amplitude difference between them, the larger the effect that a new wave will have on the travel time of the detected signal. It is the combined effect of these two mechanisms that makes the sensitivity become zero at the ray path and maximum in the region next (but not very close) to it.

5 Colab Notebook

A Colab notebook implementing the most important formulas is available at: https://colab.research.google.com/drive/1Hmr_xaKFLxRQYPil7jgoJtnuVhVJ_5fZ#scrollTo=Yce7bSluqz86

6 Discussions

The exact computation of the sensitivity kernels would require taking into account **all possible paths** from source to receiver, even the craziest ones. In our derivation, we considered only paths that can be decomposed into two straight lines of equal length, which is far more restricted than an accurate computation would require. It suffices, however, that this simple approach provided the pivotal conditions required to mimic a more realistic scenario and divert from the straight-line path predicted by ray theory, which would render our model useless. Despite its simplicity, it is remarkable how our models are well representative of real kernels, correctly predicting the Fresnel zones and the rate with which their width increases with $\sqrt{\lambda L}$.

The fact that the amplitude decays linearly with the distance came as no surprise since we imposed this during the derivation of the equations. However, even *Quantum Field Theory* (QFT) uses normalization tricks to ensure that the Feynman path integrals give accurate predictions for the quantum property of particles instead of diverging to infinity (Feynman et al., 2010).

Our model could be made progressively more accurate by adding more virtual planes to the formulation. If instead of one, we use two parallel planes, for example, then the rays would have much more degrees of freedom to bend. With infinitely many planes spanning the entire space, we would be considering all possible paths.

7 Conclusion

Despite its limitations, our formulation correctly predicts the most relevant features of sensitivity kernels, including the fact that travel-time kernels are hollow. The step-by-step derivation and relatively easy equations should be understandable by an undergraduate student with a reasonable knowledge of classical physics and calculus. This work provides new insight into the banana-doughnut paradox, helping Earth scientists understand the finite-frequency theory, increasingly used in seismology and seismic exploration.

Acknowledgements

I thank FAPESP (Grants: 2016/03120-5 and 2018/04918-6) for providing financial support for this research. I also thank an anonymous reviewer for the excellent suggestion for improving the text.

References

- Aki, K., Christoffersson, A., and Husebye, E. S. (1977). Determination of the three-dimensional seismic structure of the lithosphere. *Journal of Geophysical Research*, 82(2):277–296.
- Aki, K. and Lee, W. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model. *Journal of Geophysical research*, 81(23):4381–4399.
- Arora, K., Cazenave, A., Engdahl, E. R., Kind, R., Manglik, A., Roy, S., Sain, K., and Uyeda, S. (2011). *Encyclopedia of solid earth geophysics*. Springer Science & Business Media.
- Bozdağ, E., Trampert, J., and Tromp, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, 185(2):845–870.
- Brokešová, J. (2006). Asymptotic Ray Method in Seismology: A Tutorial.
- Cerveny, V. (2005). *Seismic ray theory*. Cambridge university press.
- Dai, D.-C. and Stojkovic, D. (2013). Origin of the tail in Green’s functions in odd-dimensional space-times. *The European Physical Journal Plus*, 128(10):1–8.
- Durand, S., Debayle, E., Ricard, Y., Zaroли, C., and Lambotte, S. (2017). Confirmation of a change in the global shear velocity pattern at around 1000 km depth. *Geophysical Journal International*, 211(3):1628–1639.
- Dziewonski, A. M. and Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the earth and planetary interiors*, 25(4):297–356.
- Dziewonski, A. M., Hager, B. H., and O’Connell, R. J. (1977). Large-scale heterogeneities in the lower mantle. *Journal of Geophysical Research*, 82(2):239–255.
- Feynman, R. P. (2006). *QED: The strange theory of light and matter*, volume 90. Princeton University Press.
- Feynman, R. P., Hibbs, A. R., and Styer, D. F. (2010). *Quantum mechanics and path integrals*. Courier Corporation.
- French, S. and Romanowicz, B. (2014). Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. *Geophysical Journal International*, 199(3):1303–1327.
- Grand, S. P. (2002). Mantle shear-wave tomography and the fate of subducted slabs. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 360(1800):2475–2491.
- Groenenboom, J. and Snieder, R. (1995). Attenuation, dispersion, and anisotropy by multiple scattering of transmitted waves through distributions of scatterers. *The Journal of the Acoustical Society of America*, 98(6):3482–3492.

- Houser, C., Masters, G., Shearer, P., and Laske, G. (2008). Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms. *Geophysical Journal International*, 174(1):195–212.
- Hung, S.-H., Dahlen, F., and Nolet, G. (2001). Wavefront healing: a banana–doughnut perspective. *Geophysical Journal International*, 146(2):289–312.
- Igel, H. (2017). *Computational seismology: a practical introduction*. Oxford University Press.
- Ishimaru, A. (1978). *Wave propagation and scattering in random media*, volume 2. Academic press New York.
- Keiiti, A. and Richards, P. G. (1980). Quantitative seismology: Theory and methods.
- Kennet, B. (1991). IASPEI 1991 seismological tables. *Terra Nova*, 3(2):122–122.
- Kennett, B. L., Engdahl, E., and Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1):108–124.
- Koelemeijer, P., Ritsema, J., Deuss, A., and Van Heijst, H.-J. (2016). SP12RTS: a degree-12 model of shear-and compressional-wave velocity for Earth’s mantle. *Geophysical Journal International*, 204(2):1024–1039.
- Komatitsch, D. and Tromp, J. (2002a). Spectral-element simulations of global seismic wave propagation-I. Validation. *Geophysical Journal International*, 149(2):390–412.
- Komatitsch, D. and Tromp, J. (2002b). Spectral-Element Simulations of Global Seismic Wave Propagation-II. Three-Dimensional Models, Oceans, Rotation and Self-Gravitation. *Geophysical Journal International*, 150(1):303–318.
- Lei, Wenjie and Ruan, Youyi and Bozdağ, Ebru and Peter, Daniel and Lefebvre, Matthieu and Komatitsch, Dimitri and Tromp, Jeroen and Hill, Judith and Podhorszki, Norbert and Pugmire, David (2020). Global adjoint tomography—model glad-m25. *Geophysical Journal International*, 223(1):1–21.
- Lu, C., Grand, S. P., Lai, H., and Garnero, E. J. (2019). TX2019slab: A new P and S tomography model incorporating subducting slabs. *Journal of Geophysical Research: Solid Earth*, 124(11):11549–11567.
- Marquering, H., Dahlen, F., and Nolet, G. (1999). Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox. *Geophysical Journal International*, 137(3):805–815.
- Marquering, H., Nolet, G., and Dahlen, F. (1998). Three-dimensional waveform sensitivity kernels. *Geophysical Journal International*, 132(3):521–534.
- Montelli, R., Nolet, G., Dahlen, F., and Masters, G. (2006). A catalogue of deep mantle plumes: New results from finite-frequency tomography. *Geochemistry, Geophysics, Geosystems*, 7(11).

- Ritsema, J., Deuss, a. A., Van Heijst, H., and Woodhouse, J. (2011). S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements. *Geophysical Journal International*, 184(3):1223–1236.
- Ritsema, J., van Heijst, H. J., and Woodhouse, J. H. (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science*, 286(5446):1925–1928.
- Robinson, E. A. and Clark, D. (2017). *Basic Geophysics*. Society of Exploration Geophysicists.
- Schaeffer, A. and Lebedev, S. (2013). Global shear speed structure of the upper mantle and transition zone. *Geophysical Journal International*, 194(1):417–449.
- Sengupta, M. K. and Toksöz, M. N. (1977). Three dimensional model of seismic velocity variation in the Earth’s mantle. *Geophysical Research Letters*, 3(2):84–86.
- Snieder, R. (1999). Imaging and averaging in complex media. In *Diffuse waves in complex media*, pages 405–454. Springer.
- Snieder, R. and Lomax, A. (1996). Wavefield smoothing and the effect of rough velocity perturbations on arrival times and amplitudes. *Geophysical Journal International*, 125(3):796–812.
- Spetzler, J. and Snieder, R. (2004). The Fresnel volume and transmitted waves. *Geophysics*, 69(3):653–663.
- Tarantola, A. (1987). *Inverse problem theory and methods for model parameter estimation*. Elsevier.
- Woodward, M. J. (1992). Wave-equation tomography. *Geophysics*, 57(1):15–26.

Appendix A Amplitude kernel at the ray path

Here we show the complete derivation to find a representative value for the amplitude kernel when $s = 0$, which corresponds to the ray path.

We know that:

$$K_A(r, s) \Big|_{s=0} = \lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) = 2f \sqrt{\frac{2}{c}} \lim_{s \rightarrow 0} \frac{n(r, s)}{m(r, s)} = \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{n(r, s)}{A_R^*(r, s)}. \quad (33)$$

If we substitute $s = 0$ into equation 33, we get an indeterminate form $\frac{0}{0}$. Thus, that limit qualifies for the usage of the L'Hôpital's rule:

$$\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) = \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{\frac{\partial n}{\partial s}(r, s)}{\frac{\partial A_R^*}{\partial s}(r, s)} = \frac{8f^2 r^{\frac{7}{2}}}{c} \frac{\lim_{s \rightarrow 0} \frac{\partial n}{\partial s}(r, s)}{\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s)}. \quad (34)$$

By multiplying both members of 34 by $\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s)$, we find:

$$\left[\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s}(r, s) \right]^2 = \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{\partial n}{\partial s}(r, s), \quad (35)$$

where

$$\frac{\partial n}{\partial s}(r, s) = \frac{\partial}{\partial s} \int_0^T \alpha(r, s, t) \beta(r, s, t) dt = \int_0^T \left[\frac{\partial \alpha}{\partial s} \beta(r, s, t) + \alpha(r, s, t) \frac{\partial \beta}{\partial s} \right] dt. \quad (36)$$

By the addition rule for limits:

$$\lim_{s \rightarrow 0} \frac{\partial n}{\partial s}(r, s) = \int_0^T \left\{ \lim_{s \rightarrow 0} \left[\frac{\partial \alpha}{\partial s} \beta(r, s, t) \right] + \lim_{s \rightarrow 0} \left[\alpha(r, s, t) \frac{\partial \beta}{\partial s} \right] \right\} dt. \quad (37)$$

From equation 22, it is easy to see that $s \rightarrow 0 \implies \beta(r, s, t) \rightarrow 0$ and, albeit less obvious, $s \rightarrow 0 \implies \frac{\partial \alpha}{\partial s} \rightarrow 0$ too. Therefore, the first limit of 37 vanishes.

For the second term, remembering that $\int_0^s f(\xi) d\xi = f(s)$:

$$\begin{aligned}
\alpha(r, s, t) \frac{\partial \beta}{\partial s} &= \int_0^T \frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} \left\{ \frac{\sin \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}} \right\} dt \\
&= \int_0^T \frac{r^{\frac{7}{2}} \sin^2 \left[2\pi f \left(t - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^4 (2r^2 + s^2)} dt \\
&= \frac{r^{\frac{7}{2}} \left[\sin \left(\frac{8\pi f \sqrt{r^2 + s^2} - 4\pi c}{c} \right) - \sin \left(\frac{8\pi f \sqrt{r^2 + s^2}}{c} \right) + 4\pi \right]}{8\pi f (r^2 + s^2)^4 (2r^2 + s^2)} \\
&= \frac{r^{\frac{7}{2}}}{2f (r^2 + s^2)^4 (2r^2 + s^2)}.
\end{aligned} \tag{38}$$

Substituting the result of 38 into equation 35:

$$\begin{aligned}
\left[\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s} (r, s) \right]^2 &= \frac{8f^2 r^{\frac{7}{2}}}{c} \lim_{s \rightarrow 0} \frac{r^{\frac{7}{2}}}{2f (r^2 + s^2)^4 (2r^2 + s^2)} \\
&= \frac{4f}{c} \lim_{s \rightarrow 0} \frac{r^7}{(r^2 + s^2)^4 (2r^2 + s^2)}.
\end{aligned} \tag{39}$$

Taking the square root of both sides of 39:

$$\lim_{s \rightarrow 0} \frac{\partial A_R^*}{\partial s} (r, s) = \lim_{s \rightarrow 0} \frac{2r^{\frac{7}{2}}}{\underbrace{(r^2 + s^2)^2 \sqrt{\lambda(2r^2 + s^2)}}_{E(r, s)}} \implies K_A(r, s) \Big|_{s=0} = \sqrt{\frac{2}{\lambda r^3}}. \tag{40}$$

Appendix B Travel time function at the ray path

In this appendix, we show how to calculate the limit of $T_T(r, s)$ as s approaches zero, used to replace the singularity at $s = 0$.

Again, we begin by applying L'Hôpital's rule:

$$\begin{aligned}
\lim_{s \rightarrow 0} \frac{F^*(r, s, t_i)}{A_{R^*}(r, s)} &= \frac{\lim_{s \rightarrow 0} \frac{\partial F^*}{\partial s}(r, s, t_i)}{\lim_{s \rightarrow 0} \frac{\partial A_{R^*}}{\partial s}(r, s)} = \frac{\frac{2}{\sqrt{\lambda}} \lim_{s \rightarrow 0} \frac{r^{\frac{7}{2}} \sin \left[2\pi f \left(t_i - \frac{2\sqrt{r^2 + s^2}}{c} \right) \right]}{(r^2 + s^2)^2 \sqrt{2r^2 + s^2}}}{\sqrt{\frac{2}{\lambda r^3}}} \\
&= \frac{\cancel{r^5 \sqrt{2}} \sin \left[2\pi f \left(t_i - \frac{2r}{c} \right) \right]}{\cancel{r^5 \sqrt{2}}} \\
&= \sin \left\{ 2\pi f \left[\left(\frac{2r}{\lambda} \bmod 1 \right) T - \frac{2r}{c} \right] \right\} \\
&= \sin \left\{ 2\pi \left[\left(\frac{2r}{\lambda} \bmod 1 \right) - \frac{2r}{\lambda} \right] \right\}.
\end{aligned} \tag{41}$$

Noticing that $\frac{2r}{\lambda} \bmod 1 = \frac{2r}{\lambda} - \left\lfloor \frac{2r}{\lambda} \right\rfloor$, where $\left\lfloor \cdot \right\rfloor$ denotes a floor division, we have:

$$\begin{aligned}
\lim_{s \rightarrow 0} \frac{F^*(r, s, t_i)}{A_{R^*}(r, s)} &= \sin \left[2\pi \left(\frac{2r}{\lambda} - \left\lfloor \frac{2r}{\lambda} \right\rfloor - \frac{2r}{\lambda} \right) \right] \\
&= -\sin(2\pi k), \quad k = 0, 1, 2, \dots, \infty \implies \lim_{s \rightarrow 0} \frac{F^*(r, s, t_i)}{A_{R^*}(r, s)} = 0.
\end{aligned} \tag{42}$$

Chapter 7

Previous Studies of South America

In this chapter, we review some important studies of South America. We begin with a short description of the methodology of each study, followed by a summary of the results grouped by some important geological domains of the continent. The results are also compared with SAAM23 (**S**outh **A**merican **A**djoint **M**odel iteration **23**), the model resulting from the tomography carried out in this study (see the next chapter).

Van der Lee et al. (2001) interactively fitted the waveforms of S and surface wave trains in more than 500 seismograms from 90 earthquakes and 60 different broadband stations from portable and permanent networks in South America to image the three-dimensional S -velocity structure beneath central and western parts of the continent. The constraints on upper mantle S -velocity structure provided by the waveform fits have been combined with independent estimates of depth to the Moho discontinuity, yielding model SA99 (Fig. 7.1).

Heintz et al. (2005) inverted the waveform of 5850 Rayleigh wave seismograms to create a three-dimensional SV-wave velocity model for the upper mantle beneath South America and the surrounding oceans. The dense path coverage and the use of higher modes to supplement the fundamental mode of surface waves allowed to constrain seismic heterogeneities with horizontal wavelengths of a few hundred kilometres in the uppermost 400 km of the mantle (Fig. 7.2).

Feng et al. (2007) simultaneously inverted using partitioned waveform inversion (PWI) regional S and Rayleigh waveforms and fundamental mode Rayleigh wave group velocities, using 5700 Rayleigh wave group velocity dispersion curves and 1537 regional wave trains with paths principally passing through the South American continent. The joint inversion of this data set provided a three-dimensional upper mantle S velocity model and a Moho depth model for South America (Fig. 7.3).

Rocha et al. (2019) carried out a teleseismic P -wave tomography using 339 stations to record 4,989 events for P and PKIKP phases, during the years 1992-2017,

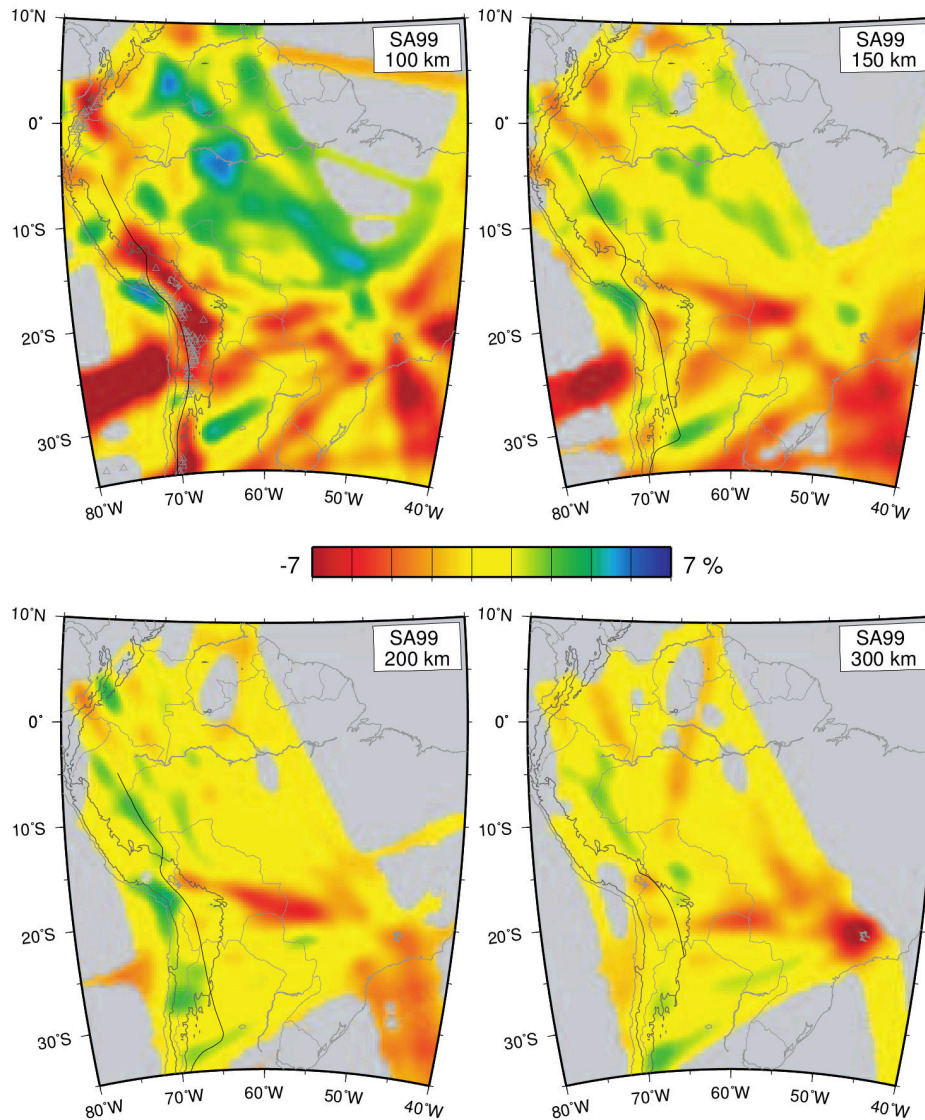


Figure 7.1: Van der Lee et al. (2001) - SA99 at depths of 100 150 200, and 300 km. The 2-km elevation and seismicity (Hasegawa and Sacks, 1981; Cahill and Isacks, 1992) are contoured with fat grey and black lines, respectively. Historical and active volcanoes are represented by light grey triangles (James, 1999). Areas to which our data have negligible sensitivity are made grey. The color scale is saturated. Resolution varies laterally.

complementing the database with a temporary network composed of 34 stations to improve the coverage in that region. The resulting model shows the velocity anomalies for the upper mantle beneath the Pantanal, Paraná, and Chaco-Paraná basins (Fig. 7.4).

Celli et al. (2020) used waveform fits of over 1.2 million vertical-component seismograms obtained with the automated multimode inversion of surface, S and multiple S waves to create a tomographic model for the crust, upper mantle and transition zone

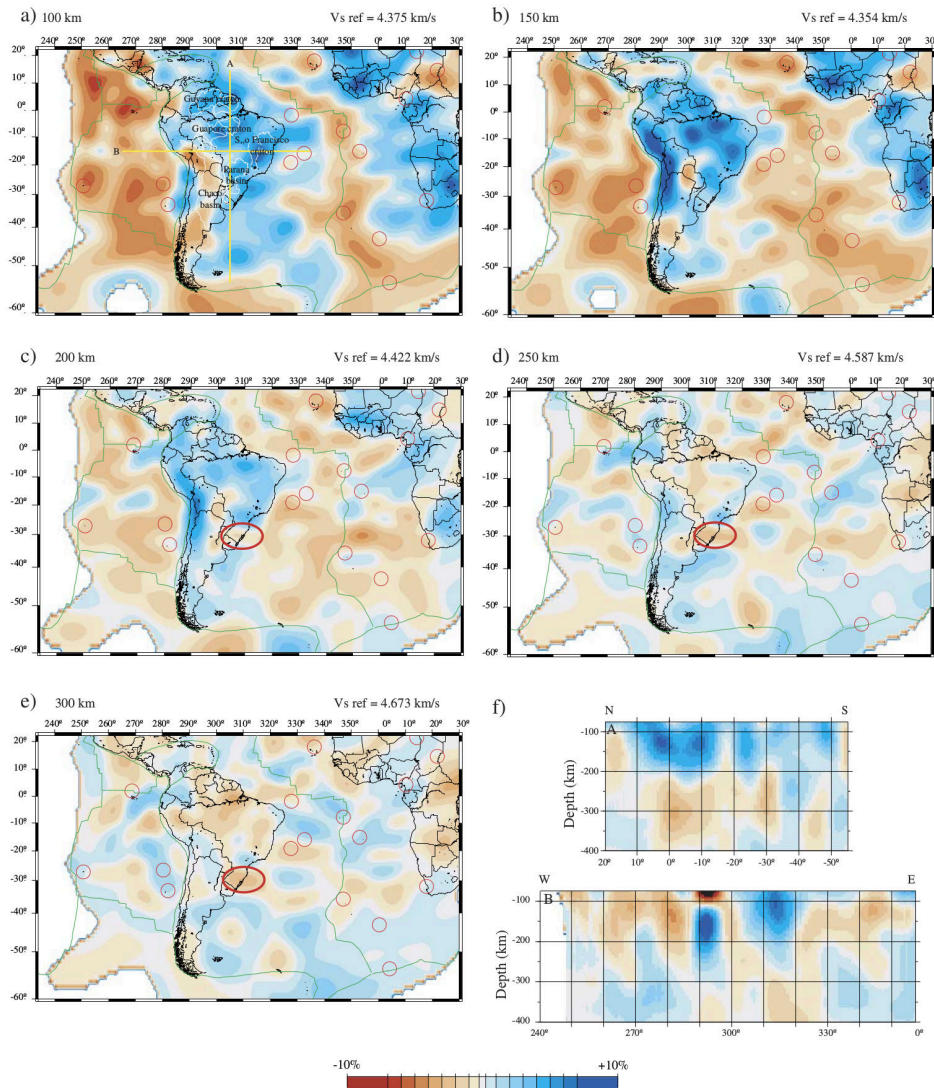


Figure 7.2: Heintz et al. (2005) - SV wave heterogeneity maps at 100 km (a), 150 km (b), 200 km (c), 250 km (d) and 300 km (e) depth with two cross sections (f) at 305°E longitude (A) and 20°S latitude (B). For each map, the hotspots are plotted as open circles, and the tectonic plates are delimited by green lines. Regions in white correspond to areas with a lack of resolution. The red circle on figures c, d and e, outlines the low velocity anomaly imaged under the Paraná large igneous province.

beneath the South Atlantic. Each waveform provides a set of linear equations constraining perturbations with respect to a 3-D reference model within an approximate sensitivity volume. All these equations were combined into a large system and solved to compute the *S*- and *P*-wave velocities, and the anisotropy within the crust, upper mantle, and uppermost lower mantle (Fig. 7.5).

7.1 Amazonian Craton

Van der Lee et al. (2001) observed a high-velocity lithosphere beneath the western Guyana and Central Brazil shields, indicating that both shields are underlain by cratonic lithosphere of the Precambrian Amazonian Craton. Feng et al. (2007) also found high-velocity lithosphere beneath the Amazonian Craton (AC) with an average thickness of around 160 km, with the eastern Archean part being faster and thicker (~ 200 km) than the western portion of the Craton, generally consistent with the ages of the geochronological provinces (Tassinari and Macambira, 1999). The Transbrasiliano lineament shows up as a belt of lower velocities at 100-200 km depths separating the Amazonian and the São Francisco cratons (Feng et al., 2007). In our model for South America (SAAM23), we observe the same correlation between higher velocities and older lithosphere.

Model SA2019 (Celli et al., 2020) shows a high S -wave-velocity anomaly (8%) consistent with the AC. Under large parts of the Guyana and Central Brazil shields, corresponding to the older geochronological provinces of the craton (Carajás, Central Amazonian, Transamazonian, and Tapajós-Parima), the authors found high-velocity roots, while the younger Rio Negro, Rondônia-Juruena, and Sunsás provinces are underlain by lower velocities. An area of pronounced low velocities, spanning the Pantanal-Chaco basin and most of the Rondônia-Juruena and southern Sunsás provinces, shows up at 200-330 km depth (Celli et al., 2020). The Rio Negro province in the north also lacks a lithospheric root, suggesting either that the youngest, western portion of the craton never developed a thick lithospheric root or that its root has been eroded (Celli et al., 2020).

Unlike the other studies (Van der Lee et al., 2001; Feng et al., 2007; Celli et al., 2020), Heintz et al. (2005) observed in their velocity model for South America an elongated domain of moderately high velocities located beneath the Amazon and Solimões basins and extending down to 100 km depth. This domain separates two high-velocity anomalies, corresponding to the Guyana and the Central Brazil shields, suggesting that the rifting episode responsible for the formation of the Amazon basin has involved a significant portion of the lithosphere (Heintz et al., 2005). SAAM23 also shows a clear separation between the two shields as a thinned lithosphere along the eastern

and central parts of the Amazon basin. However, that separation becomes much less pronounced in the western part of the craton (Rondonian-San Ignácio province).

7.2 São Francisco Craton

The high S -velocity anomaly in model SA99 (Van der Lee et al., 2001) located at the Southern portion of the São Francisco Craton (SFC) suggests a lithosphere-asthenosphere boundary (LAB) with 100-150 km depth, depending on the criteria adopted to define it. Assuming that these high S -velocity anomalies are caused predominantly by lower temperatures (Jordan, 1988) but also by a depleted composition, the relatively weak high S -velocity anomaly in this region implies that the root of the SFC is not as cold or as depleted as the higher velocity cratonic root of the AC (Van der Lee et al., 2001). Heintz et al. (2005) observed high shear-wave velocities down to 200 km depth beneath the Amazonian and the São Francisco cratons. Both AC and SFC have been interpreted as having depleted lithospheric mantle (Finger et al., 2021).

As in the AC, Feng et al. (2007) found an average thickness of 160 km for the high-velocity lithosphere. High shear velocities extending down to about 150-200 km depth delimit the Archean block in the southern part of the SFC (Feng et al., 2007). High velocities at 100 to 150 km depth extend from the SFC westward beneath the Brasília Belt, and the tomography also shows a thicker lithospheric root beneath the southern part of the craton, compared to its northern part (Feng et al., 2007).

The SFC shows up as a high-velocity anomaly as strong as the one corresponding to the AC in SA2019 (Celli et al., 2020) and in SAAM23. Both SA2019 and SAAM23 show a weak positive anomaly at lithospheric depths underlying the Borborema Province, located north of the SFC, indicating the absence of a thick lithosphere, in agreement with recent P -wave tomography of Simões Neto et al. (2019), who also found lower velocities in the region (Celli et al., 2020). All models (Van der Lee et al., 2001; Heintz et al., 2005; Feng et al., 2007; Celli et al., 2020), as well as SAAM23 show a clear separation between the high-velocity anomalies associated with the Amazonian and the São Francisco cratons, consistent with the existence of a Neoproterozoic belt between the two blocks, marked by the Transbrasiliano Lineament.

7.3 Paraná Basin

7.3.1 Cratonic Nucleus

All studies, including this one, show a high-velocity anomaly underlying the Paraná Basin and coinciding with the proposed location of the Paranapanema Block (PpB) (Cordani, 1984; Mantovani and Brito Neves, 2005). Heintz et al. (2005) found the high velocities down to 200 km depth, which may correspond to cratonic lithosphere underlying the Paraná Basin. That velocity anomaly extends northeastward, beneath the SFC, so that the two structures are indistinguishable (Heintz et al., 2005). Feng et al. (2007) observed high velocities to 100-150 km beneath the Paraná Basin, consistent with the existence of a Proterozoic cratonic nucleus. The results indicate that any rifting episode during the evolution of the basin was not strong enough to affect the thick lithosphere on a regional scale detectable by the surface wave tomography (Feng et al., 2007). In the NE portion of the Paraná Basin, an average to low-velocity lithosphere indicates that the limits of the PpB do not reach that far or that the original cratonic basement was affected by Late Cretaceous intraplate magmatism (Rocha et al., 2019). SAAM23 shows the PpB separated from the SFC. Affonso et al. (2021) also observed the separation between the PpB and the SFC as a decreased high-velocity anomaly.

7.3.2 "Paraná Plume"

Using the travel times of teleseismic P and S waves, VanDecar et al. (1995) found evidence for a fossil plume in the deep upper mantle beneath the Paraná Basin as a cylindrical low-velocity anomaly. The existence of the cylindrical anomaly was confirmed by Schimmel et al. (2003).

Van der Lee et al. (2001), Feng et al. (2007), Rocha et al. (2019), Celli et al. (2020), and this study also observed a low-velocity anomaly in the upper mantle beneath the Paraná Basin. Heintz et al. (2005) failed to recover the low velocity structure at the same location as VanDecar et al. (1995) and Schimmel et al. (2003), most likely due to a lack of horizontal resolution. Van der Lee et al. (2001) position the anomaly in the eastern portion of the basin, (Feng et al., 2007) in the north, Rocha et al. (2019) in the northeast, and Celli et al. (2020) in the southern part.

7.3.3 Rio de la Plata Craton

None of the studies, including ours, observed high-velocity anomalies consistent with the location of the Rio de la Plata Craton proposed by Rapela et al. (2011) or Oyhan-tçabal et al. (2011).

7.4 Parnaíba Basin

Feng et al. (2007) found high S velocities down to 100-150 km beneath the Parnaíba Basin in NE Brazil, where a cratonic nucleus (Parnaíba Block) has also been proposed (Cordani, 1984). Celli et al. (2020) also imaged the cratonic root of the Paranaíba basin as a high-velocity anomaly under the northern half of the basin. At 110 km depth, the Parnaíba Block appears connecting the Amazonian and the São Francisco cratons. However, lower velocities under the Araguaia and Borborema fold belts at 80 and 150 km depth suggest that the Paranaíba cratonic nucleus is an independent body, separated from the neighbouring cratons, at the limits of the resolution of SA2019 (Celli et al., 2020). This result is in agreement with our findings, as SAAM23 clearly shows the Paranaíba Block as an independent cratonic nucleus, even though its shape may not be as inferred from other geophysical studies.

7.5 Andes

Van der Lee et al. (2001), Feng et al. (2007), and this study found a mantle wedge with extremely low S -velocities below the thick crust along the volcanically active parts of the central Andes. The bottom of the low velocity mantle wedge is immediately adjacent to the Wadati-Benioff zone and thus to the subducting oceanic crust (Van der Lee et al., 2001). These low velocities in the mantle wedge were explained by a 2.5-4% partial melt, which could occur at temperatures near 1,200° in the presence of realistic amounts (0.15 wt%) of volatiles released from the subducting oceanic crust (Van der Lee et al., 2001).

Feng et al. (2007) observed that flat subduction slabs along the Andes are characterized by moderately high S velocity and the absence of a low-velocity mantle wedge. Celli et al. (2020) imaged the Peruvian and Bucaramanga flat slab in the north of the

Andes. In the central Andes, Celli et al. (2020) imaged the Nazca Slab and the Pampean slab gap, and also imaged a low-velocity conduit rising from the deep upper mantle under the Central Volcanic Province. In the southern Cordillera, low velocities of the subducted Chile Rise dominate the lithosphere, creating a slab window (Celli et al., 2020). Very low shear-wave velocity was also found in SAAM23 correlated with this "slab window", as well as in most of the Patagonian Domain.

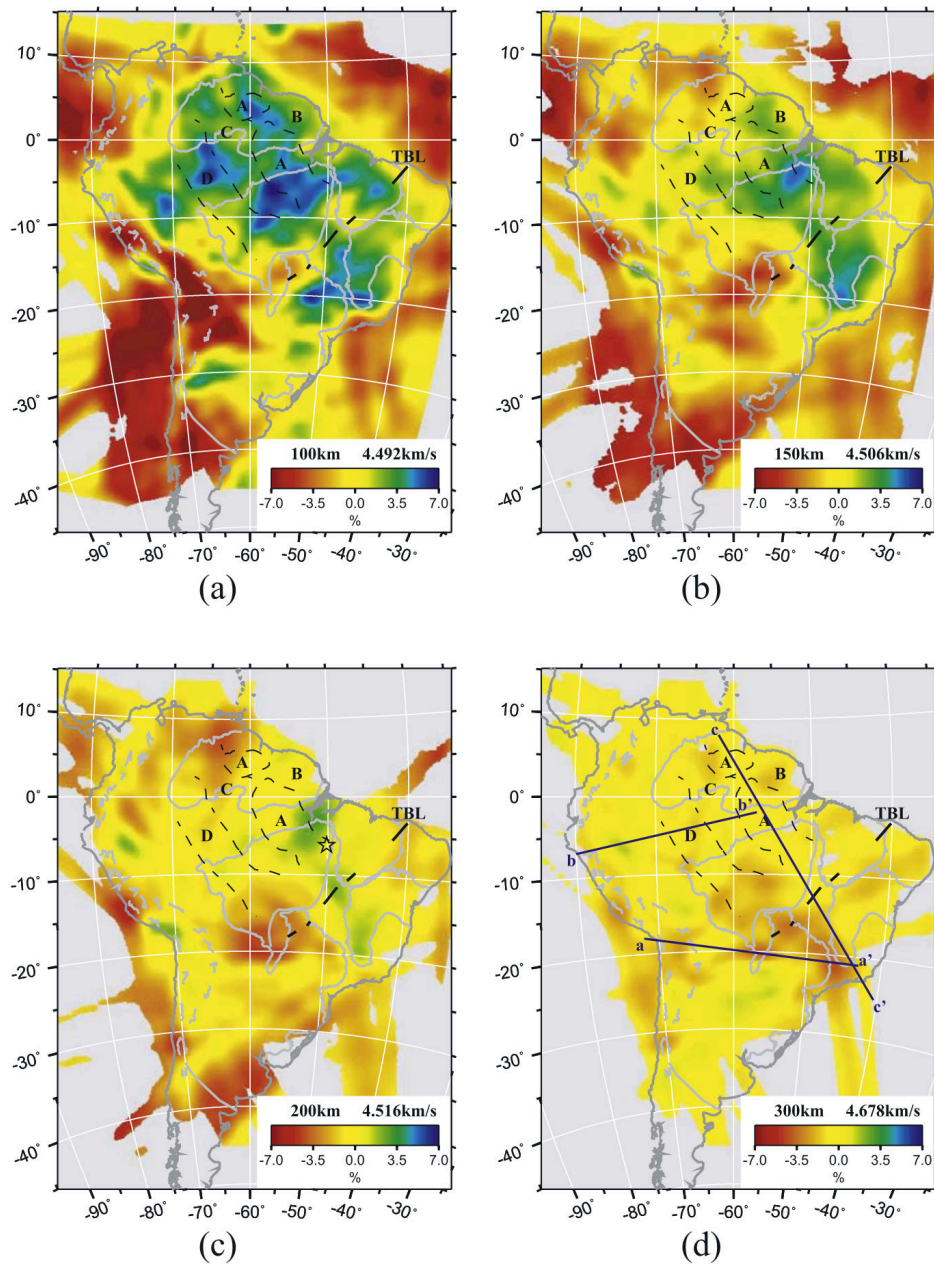


Figure 7.3: Feng et al. (2007) - Horizontal slices of the obtained S velocity model at 100, 150, 200, and 300 km. The S velocities are mapped as deviations from those in IASP91. The dashed domains indicate the geochronological provinces of the Amazonian Craton, where "A" is equivalent to the Central amazonian province (oldest domain), "B" to Maroni-Itacaiúnas province, "C" to Ventuari-Tapajós province, and "D" to Rio Negro-Juruena province (youngest domain). Rondonian-San Ignácio e Sunsás-Aguapei provinces are too small to be shown here. TBL denotes the Transbrasiliano lineament. The star on the 200 km map indicates the location where the oldest rocks (2.9 - 3.0 Ga granitoids/greenstones) found in the Amazonian Craton.

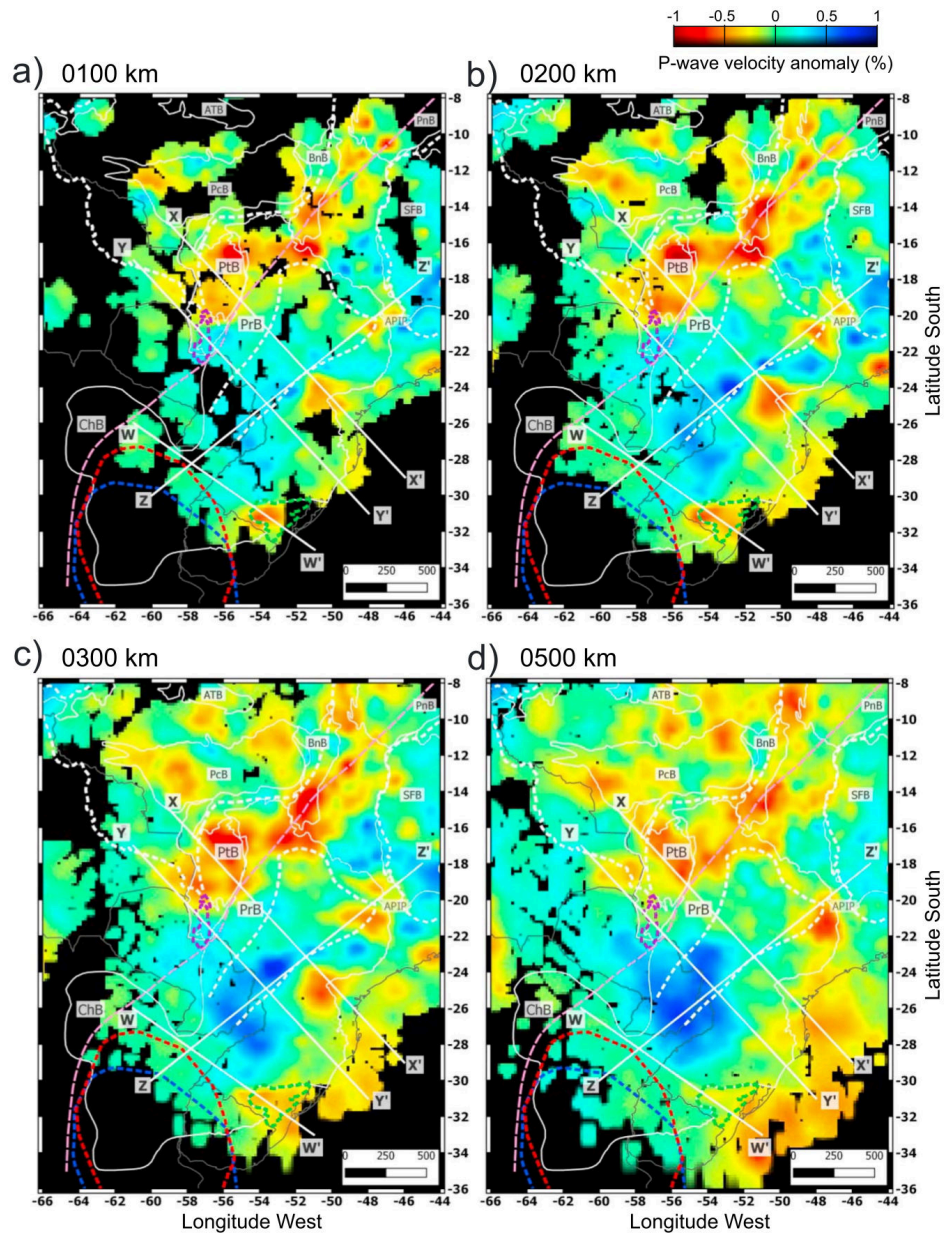


Figure 7.4: Rocha et al. (2019) - Horizontal tomographic images for selected depths in the upper mantle. The white squares are the stations. The solid white contours are the limits of the main sedimentary basins, and the dashed lines are the limits of cratons (Amazonian, S. Francisco, and the Paranapanema Block). APiP is the Alto Paranaíba Igneous Province. The thin solid, black lines are country boundaries.

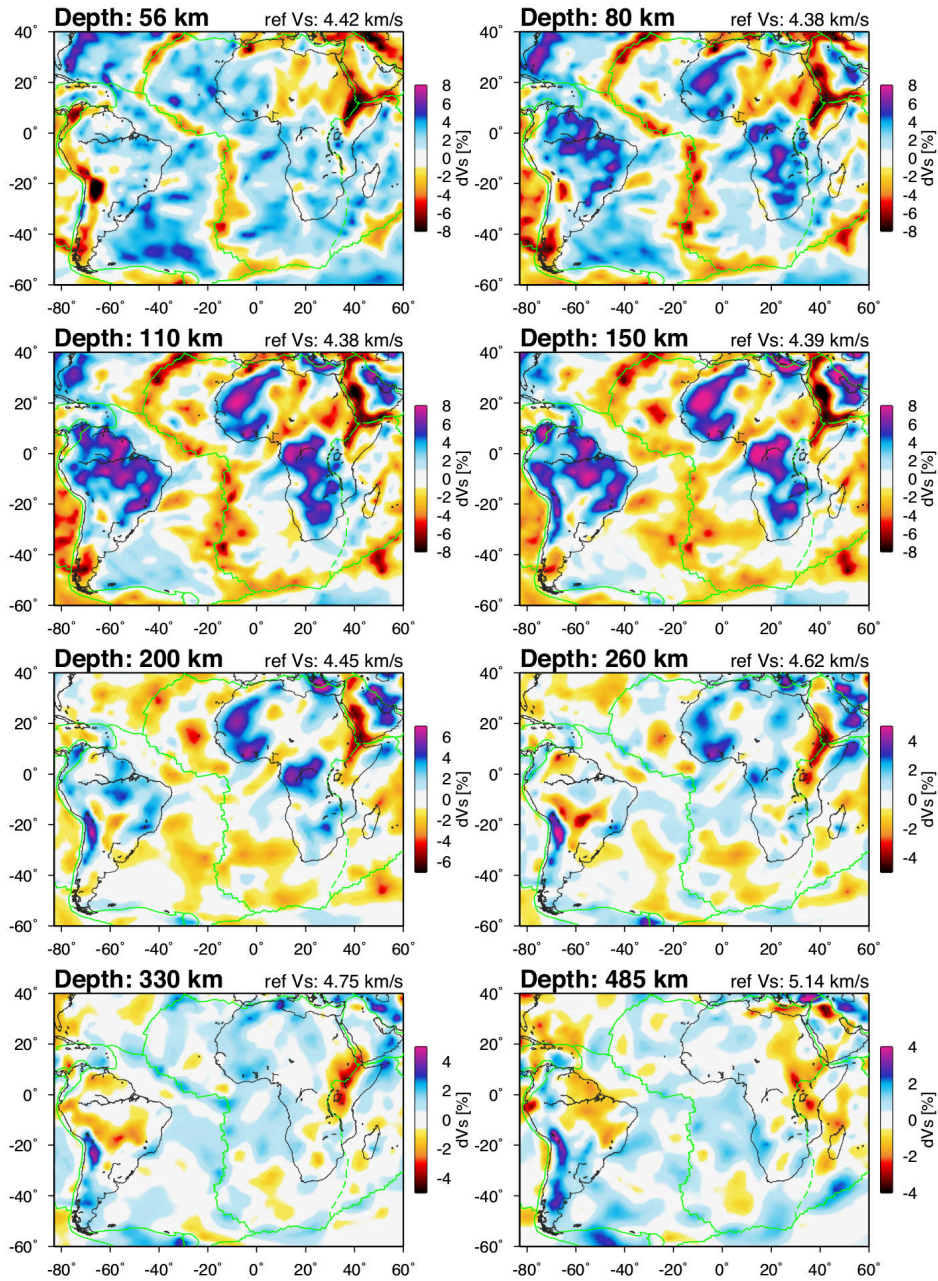


Figure 7.5: Celli et al. (2020) - Map views of SA2019 *S*-wave velocity variations at eight depths (56, 80, 110, 150, 200, 260, 330, 485). Velocity perturbations are plotted in percent with respect to the reference at depth (marked on the top right of each map). Major plate boundaries are plotted as green lines (solid lines: verified; dotted lines: proposed).

Chapter 8

Adjoint Tomography

The previous chapters were devoted to explaining the theory behind FWI and reviewing other tomographic studies of South America. This chapter includes a manuscript, submitted to *JGR: Solid Earth*, that contains the main subject of this thesis, which is the full-waveform inversion carried out to create an adjoint tomographic model of South America. In that paper, we make a concise yet comprehensive description of the geology of South America and a briefing of the historical development of the adjoint method in seismology. We also explain the importance of the choice of the misfit measurement, how to carry out accurate 3D seismic-wave simulations, and how to use them in conjunction with the adjoint method to perform an FWI. We estimate the quality of our resulting model using four different approaches, compare it with many other studies, and tectonically interpret the vertically polarized shear seismic velocity anomalies.

Adjoint Tomography of South America based on 3D Spectral-Element Seismic Wave Simulations

Caio Ciardelli¹, Marcelo Assumpção¹, Ebru Bozdağ², and Suzan van der Lee³

¹Department of Geophysics, Institute of Astronomy, Geophysics and Atmospheric Sciences, University of São Paulo

²Department of Geophysics, Colorado School of Mines

³Department of Earth and Planetary Sciences Weinberg College of Arts and Sciences, Northwestern University

Key Points:

- A new model of South America mantle down to $\sim 1,500$ km was obtained with full-waveform tomography
- Cratons and hidden cratonic nuclei were well resolved; in the Amazonian craton, higher-velocity anomalies correlate with the older provinces
- The Nazca Slab plunges into the lower mantle in northern South America and flattens towards the south near the 650 km discontinuity

Corresponding author: Caio Ciardelli, caio.ciardelli@gmail.com

Abstract

We use 3D spectral-element seismic wave simulations and data from 112 earthquakes recorded by 1311 seismic stations to construct an adjoint waveform tomography model of South America. We performed 23 iterations using exponentiated instantaneous phase measurements, which optimize the cycle skip problem and the information extracted from each time series. Our final model (SAAM23, South American Adjoint Model - iteration 23) shows a 50% decrease in the misfit relative to its 3D starting model. We assessed SAAM23 by cross-correlation traveltimes measurements using 53 earthquakes not included in the adjoint inversion and point-spread function tests. The Nazca Slab is well imaged and is shown to be continuous in the 300-500 km depth range following the Peruvian flat-slab segment. Beneath northern South America, the slab crosses the mantle transition zone and continues into the lower mantle. In the central and southern part of South America, the slab appears to flatten near the 650 km discontinuity before continuing into the lower mantle. In the stable Precambrian platform, both exposed cratons (Amazonian and São Francisco), as well as covered cratonic blocks beneath the intracratonic Paraná and Parnaíba basins (Paranapanema and Parnaíba, respectively), show high velocities at lithospheric depths. The seismic Lithosphere/Asthenosphere boundary (LAB) agrees well with that obtained by *S*-wave receiver functions. In the Amazonian Craton, the positive lithospheric *S*-wave velocity anomalies and LAB depth increase with the average age of the geochronological provinces. No high-velocity anomalies were found beneath the Río de la Plata Craton.

Plain Language Summary

We developed a new model of mantle seismic velocities beneath the South American Plate with full-waveform inversion by calculating synthetic seismograms that match observations from 112 earthquakes recorded by 1311 stations. This model (SAAM23, South American Adjoint Model - iteration 23) was validated with an independent set of observations. The model shows the Nazca Slab crossing the transition zone and plunging directly into the lower mantle beneath the 650 km discontinuity. In the southern part of the continent, the Nazca Slab flattens and remains close to the bottom of the transition zone. The oldest continental regions (the exposed Amazon and São Francisco cratons, as well as the cratonic blocks buried beneath the Paraná and Parnaíba Basins) have high velocities in the upper mantle. The boundary between the lithosphere (the rigid upper portion of the mantle) and the asthenosphere (the more ductile region below), called LAB, estimated by SAAM23 agrees well with other studies. In the Amazonian Craton, both the upper mantle velocities as well as the LAB depth increase with the average age of the geochronological provinces. However, no high velocities were found beneath the Río de la Plata Craton.

1 Introduction

South America is separated into three major tectonic domains: the South American Platform, the Patagonian Platform, and the Andes (Almeida et al., 2000). The South American Platform, located in the eastern part of the continent, can be further broken down into the Amazonian Domain and the Brasiliano Domain, both composed of Archean-Proterozoic cratons covered by Phanerozoic sediments (Brito Neves & Fuck, 2014). The Patagonian Platform locates in the southern portion of South America and is described by V. A. Ramos (2004) as an allochthonous terrane amalgamated to Gondwana in the Early Permian. The Andes, to the west of the two platforms, consists of the Northern, Central, and Southern cordilleras (V. Ramos, 1999). Despite the high seismic activity in the western boundary of the continent due to the subduction of the Nazca Plate, which is responsible for the very existence of the Andean Cordillera, a long-lived dearth of seismographic stations, especially in the South American Platform, has limited the resolution in the regional tomographic studies of South America. Over the last decade, the deployment of the Brazilian Seismic

Network (RSBR) significantly improved seismic data coverage in Brazil, helping to increase the resolution of tomographic images of the continent.

The first continental-scale tomographic study of South America was constructed by Van der Lee et al. (2001) by fitting S -wave and multimode surface-wave waveforms to image the 3D S -wave velocity structure beneath the central and western parts of the continent. This study revealed the architecture and outlines of the South American cratons and basins, continuous along-Andes variations in slab dip and mantle wedge structure, and estimated the degree of melting in the prominent mantle wedge of the central part of the Central Andes. Heintz et al. (2005) inverted the waveforms of Rayleigh waves to create a 3D vertically-polarized shear-wave velocity (β_v) and Moho depth model for the upper mantle beneath South America and the surrounding oceans. Feng et al. (2007) performed a partitioned waveform inversion (PWI) using both regional S and Rayleigh waveforms together with fundamental-mode Rayleigh-wave group wave velocity to construct a 3D upper-mantle β and a Moho depth models for the region. M. Rocha et al. (2019) carried out a teleseismic P -wave tomography, and the resulting model shows the velocity anomalies for the upper mantle beneath the Pantanal, Paraná, and Chaco-Paraná basins. Celli et al. (2020) used waveform fits of vertical-component seismograms obtained with the automated-multimode inversion of surface, S , and multiply-reflected S waves for a tomographic model for the crust, upper mantle, and transition zone beneath the South Atlantic region, including South America. Common findings among these studies include a clear separation between the high-velocity anomalies associated with the Amazonian and the São Francisco cratons and the existence of a high-velocity lithosphere under the Paraná Basin.

The resolution of seismic-tomographic models is directly controlled by data coverage and the chosen forward and inverse theory used in tomography. Recognizing their importance to further improve the resolution of tomographic models, finite-frequency effects (e.g., Dahlen et al., 2000) have been taken into account both in body-wave (e.g., Montelli et al., 2004) and surface-wave (e.g., Zhou et al., 2006) tomographic studies computing Fréchet kernels in 1D reference models. Finite-frequency kernels for 3D background models must be computed numerically. Following the increase in computational power and advances in numerical methods to solve the 3D wave equation (Komatitsch & Vilotte, 1998; Komatitsch & Tromp, 1999, 2002b, 2002a) adjoint waveform inversions, a full-waveform inversion (FWI) technique, have become feasible in earthquake seismology from crustal to global scales. Introduced by Tarantola (1984), the adjoint waveform inversion takes advantage of the full complexity of the seismic wavefield both in the computation of synthetic seismograms and Fréchet kernels to iteratively update seismic-tomographic models (e.g., Tromp et al., 2005; Plessix, 2006; Fichtner et al., 2006; Virieux & Operto, 2009). 3D waveform simulations allow for maximizing the information extracted from each seismic record with appropriately defined misfit functions based on the difference between observed and simulated (synthetic) data (e.g., Brossier et al., 2010; Bozdağ et al., 2011). A review of the method, discussing the tomographic resolution of the mantle models, and highlighting the most recent theoretical and computational advances in direct modeling methods may be found in Q. Liu and Gu (2012).

The application of adjoint tomography on regional and global scales using 3D wave simulations began a little more than a decade ago with first examples from imaging the Southern Californian crust (Tape et al., 2009) and the upper-mantle structure beneath Australia (Fichtner et al., 2009). Other successful applications of the method followed at the regional and continental scales are: Zhu et al. (2012), Rickers et al. (2013), Fichtner et al. (2013), M. Chen et al. (2015), and Zhu et al. (2017). More recently, the method has also become feasible at the global scale (e.g., Bozdağ et al., 2016; Lei et al., 2020). Colli et al. (2013) carried out the first adjoint tomography of the upper mantle in the South Atlantic region, imaging the Amazonian and the São Francisco cratons as separate blocks in South America. Gao et al. (2021) used adjoint inversions to image the Central Andes, confirming low-velocities resulting from a hydrated mantle wedge and finding evidence for

the intense crustal partial melting and underthrusting of the Brazilian Shield beneath the southern Altiplano and delamination beneath the southern Puna.

The choice of the misfit function is a key step that controls the success and the convergence of FWI (e.g., Modrak & Tromp, 2016). It is common to split seismic traces into smaller measurement windows using automated-window selection algorithms (e.g., Maggi et al., 2009; Lee & Chen, 2013; Y. Chen et al., 2017) to select high-quality portions of seismograms as well as to maximize the information extracted from each seismic trace. The latter is mainly a concern for time-domain cross-correlation (CC) (Luo & Schuster, 1991; Tanimoto, 1995; Marquering et al., 1999; Dahlen et al., 2000; Tromp et al., 2005) and frequency-dependent multitaper traveltime (MP) (Zhou et al., 2004, 2005; Tape et al., 2009) measurements where each seismic phase should be selected by measurement windows to avoid the dominance of high-amplitude signals in the misfit. Furthermore, the smaller amplitude scattered waves, which provide valuable constraints on the structure of the medium they propagate through, are generally suppressed in measurement windows (e.g., Rickers et al., 2012). Unlike CC based misfit functions, time-frequency misfits (e.g., Kristeková et al., 2009; Fichtner et al., 2009) or instantaneous phase and amplitude misfits (e.g., Bozdağ et al., 2011) can be applied to wavetrains without the need for multiple measurement windows. Rickers et al. (2013) used an instantaneous-phase misfit to image the Icelandic plume to overcome potential wavefront-healing problems (e.g., Nolet & Dahlen, 2000). In a synthetic study, Yuan et al. (2020) proposed a variant of the instantaneous phase misfit of Bozdağ et al. (2011), called exponentiated phase (EP) misfit, to overcome potential cycle skip problems of phase measurements.

In this study, taking advantage of the improved ray coverage provided by the RSBR and the EP misfit (Yuan et al., 2020), we construct a high-resolution adjoint waveform model for the South American continent and surrounding oceans to provide improved constraints to have better insight into the tectonics and geological processes of the region. In the following section, we give background information on the geology and tectonics as well as previous tomographic studies of South America. In Section 3, we describe our adjoint waveform inversion workflow. Section 4 presents the assessment of the model quality based on the misfit evolution, an independent data set that was not used in inversions and resolution tests. We discuss our observations and findings in Section 5 followed by our conclusions in Section 6.

2 Geology and Previous Studies of South America

South American geology includes stable Archean cratons & sediment-covered Proterozoic platforms, Paleozoic orogenies, Mesozoic, Cenozoic, and ongoing deformation related to mountain building & volcanism in the Andes. Within typical plate-tectonic reference frames, the South American plate moves westwards more rapidly than the adjacent plates while overriding the Nazca Plate and helping to grow the Atlantic Ocean (DeMets et al., 2010). Geologically, the South American continent is composed of three major tectonic domains: the South American Platform, the Patagonian Platform, and the Andes (Almeida et al., 2000; Hasui et al., 2012). In this section, we give a general overview of some of the most important tectonic and geological structures in South America (Fig. 1).

2.1 South American Platform

Almeida et al. (2000) define the South American Platform as the stable continental portion of the South American plate that was not affected by the Phanerozoic orogenic zones of the Andes and the Caribbean. The South American Platform is separated into the Amazonian Domain (AD) and the Brasiliano Domain (BD) by the Transbrasiliano Lineament (TBL) and the Araguaia Lineament (AL) (Fig. 1), two long linear fault zones (Brito Neves & Fuck, 2014), and is composed mainly of Archean rocks overlain by Phanerozoic sedimentary cover that developed through six main tectono-sedimentary sequences ranging from the Pale-

oproterozoic to the Cambrian (Almeida et al., 2000; Hasui et al., 2012). Consolidation of all components of the platform occurred when the orogenic belts developed during the Brasiliano cycle stabilized, around 440 Ma (Upper Ordovician), and became part of the Gondwana supercontinent (Hasui et al., 2012). The three main sets of orogenic events that amalgamated the platform are: 1) Trans-Amazonian (Paleoproterozoic), 2) Late Mesoproterozoic, and 3) Brasiliano/Pan African (Almeida et al., 2000). The Brasiliano cycle consolidated the youngest mobile belts of the platform basement and is the main responsible for the overall distribution of the cratonic nuclei and fold belts (Almeida et al., 2000).

2.1.1 Amazonian Domain

In the AD, the Amazonian Craton (AC) (Fig. 1) comprises over four million km^2 and is thereby one of the largest regions on the planet from the Archean and Proterozoic periods (Coutinho, 2008). The northern portion of the craton is known as the Guyana Shield, and the southern portion is named the Central Brazil Shield. Separating the two shields, are the Phanerozoic Solimões and the Amazon basins. The Andean sedimentary basins cover the northwestern and southwestern portions of the craton (M. L. Vasquez, 2006).

Some authors such as Teixeira et al. (1989), C. C. Tassinari and Macambira (1999), and Cordani et al. (2016) separate the craton into six geochronological provinces: Central Amazonian (>2.3 Ga), Maroni-Itacaiúnas (2.2 to 1.95 Ga), Ventuari-Tapajós (1.95 to 1.8 Ga), Rio Negro-Juruena (1.8 to 1.55 Ga), Rondonian-San Ignácio (1.5 to 1.3 Ga), and Sunsas-Aguapeí (1.25 to 1.0 Ga). According to the geochronological model proposed by M. L. Vasquez (2006) for the formation of the AC, the process initiated with an Archean protocraton formed by a merger of three microcontinents (Imataca, Carajás, and Xingu-Iricoumé) via orogenic belts of the Maroni-Itacaiúnas province. Later on, between 1,95 and 1,45 Ga, successive magmatic arcs (Ventuari-Tapajós, Rio Negro-Juruena, and part of the Rondonian-San Ignácio province) accreted onto the western edge of that new continent. Eventually, the continental collision between the AC and Laurentia (located further west) formed the younger orogenic belts of the Rondonian-San Ignácio and Sunsas-Aguapeí provinces (M. L. Vasquez, 2006), as well as the Grenville Orogen on the eastern margin of Laurentia (Hynes & Rivers, 2010).

Van der Lee et al. (2001) observed a high-velocity lithosphere beneath the western Guyana and Central Brazil shields, indicating that both shields are underlain by cratonic lithosphere of the Precambrian AC. Feng et al. (2007) also found high-velocity lithosphere beneath the AC with an average thickness of around 160 km, with the eastern Archean part being faster and thicker (~ 200 km) than the western portion of the Craton, generally consistent with the ages of the geochronological provinces (C. C. Tassinari & Macambira, 1999). The TBL shows up as a belt of lower velocities at 100-200 km depths separating the Amazonian and the São Francisco cratons (Feng et al., 2007). In our model for South America (SAAM23), we observe the same correlation between higher velocities and older lithosphere.

Model SA2019 by Celli et al. (2020) shows an elevated S -wave velocity anomaly of 8% beneath the AC. Under large parts of the Guyana and Central Brazil shields, corresponding to the older geochronological provinces of the craton (Carajás, Central Amazonian, Transamazonian, and Tapajós-Parima). The authors found high-velocity roots while the younger Rio Negro, Rondônia-Juruena, and Sunsás provinces are underlain by lower velocities. An area of pronounced low velocities between depths of 200 and 330 km and comprising the Pantanal-Chaco Basin and most of the Rondônia-Juruena and southern Sunsás provinces was imaged by Celli et al. (2020). The Rio Negro province in the north also lacks a cool lithospheric root (Celli et al., 2020).

Heintz et al. (2005) reported an elongated domain, marked by moderately high velocities, that separates the high-velocity lithosphere of the Guyana and the Central Brazil shields, which is not observed in the studies, for instance, by Van der Lee et al. (2001), Feng et al. (2007) and Celli et al. (2020). This domain follows the Solimões and Amazon

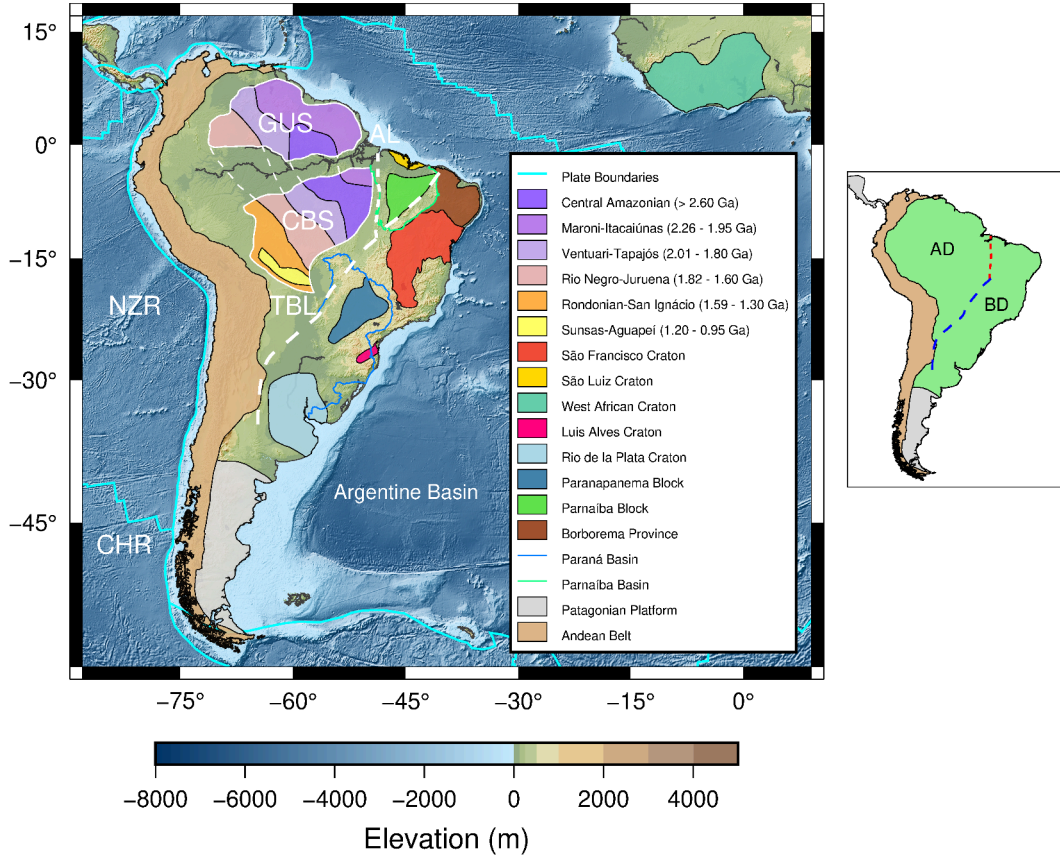


Figure 1. Major geological structures in South America. The Andean Belt (to the west) appears in light brown and the Patagonian Platform (to the south) appears in light gray. The cratons are represented by all the other colored polygons. The Amazonian Craton is composed by the Guyana Shield (GUS) to the north and by the Central Brazil Shield (CBS) to the south. We also identify the six geochronological provinces (Central Amazonian, Maroni-Itacaiúnas, Ventuari-Tapajós, Rio Negro-Juruena, Rondonian-San Ignacio, and Sunsas-Aguapeí) using different shades and colors. The darkblue polygon represents the Paranapanema Block and the lightgreen one the Parnaíba Block. The map also includes the Paraná (blue contour) and the Parnaíba (green contour) basins, the AL, and the TBL. The Nazca Ridge (NZR) and the Chile Ridge (CHR) are also shown to the west of the subduction zone. Inset map: South American Platform (light green), Patagonian Platform (light gray), and Andean Belt (light brown). The AD is separated from the BD by the TBL in the southern and central parts of Brazil, and by the BD in the northern portion of Brazil.

basins. Our model, SAAM23, also shows a clear separation between the two shields as a thinned lithosphere along with the eastern and central parts of the Amazon Basin. However, the separation is less pronounced in the western part (Solimões Basin) of the craton (Rondonian-San Ignácio province).

2.1.2 *Brasiliano Domain*

The BD is characterized by the eastern portion of the South American Platform in which Archean and Early to Middle Proterozoic rocks were reworked during the Brasiliano tectono-magmatic cycle (Moura & Gaudette, 1993). It consists of a collage of many cratonic nuclei: São Luiz, São Francisco, Paranapanema, Luis Alves, Río de la Plata, and, possibly, Paranaíba, overlaid by Phanerozoic cover (Brito Neves & Fuck, 2014).

The most intensively studied portion of the Precambrian nucleus of the South American Plate is the São Francisco Craton (SFC) and its margins (Fig. 1). The craton contains rocks from the PaleoArchean to the Cenozoic, including several sedimentary deposits (Heilbron et al., 2016). The SFC is surrounded by Brasiliano/Panafrican fold belts. Archean and Paleoproterozoic basement is exposed in the northeastern and southern portions of the craton (Engler, 2009). Most of the central region is covered by Precambrian and Phanerozoic units (Alkmim, 2004), combined with parts of the crystalline basement that were tectonically reworked between 1.8 and 1.2 Ga. A rift-thrust belt also developed during that period and was filled with Paleo to Mesoproterozoic sediments (Espinhaço Supergroup). Neoproterozoic sediments (Bambu Group) also cover extensive portions of the basement (Engler, 2009). The resistance offered by the São Francisco/Congo Craton and São Luis/West African Craton during the final breakup of South America from Africa around 130-90 Ma ago may explain the large magmatic manifestations in the Paraná and Parnaíba sedimentary basins (Filho et al., 2000).

Heintz et al. (2005) observed high shear-wave velocities down to 200 km depth beneath the Amazonian and the SFC. Both AC and SFC have been interpreted as having depleted lithospheric mantle (Finger et al., 2021). As for the AC, Feng et al. (2007) found an average thickness of 160 km for the high-velocity lithosphere of the SFC. High shear velocities extending down to about 150-200 km depth delimit the Archean block in the southern part of the SFC (Feng et al., 2007). High velocities at 100 to 150 km depth extend from the SFC westward beneath the Brasília Belt, and the tomography also shows a thicker lithospheric root beneath the southern part of the craton, compared with its northern part (Feng et al., 2007).

The SFC shows up as a high-velocity anomaly as strong as the one corresponding to the AC in models SA2019 (Celli et al., 2020) and SAAM23. The models of Van der Lee et al. (2001); Heintz et al. (2005); Feng et al. (2007); Celli et al. (2020) and our model, SAAM23, show a clear separation between the high-velocity anomalies associated with the Amazonian and the São Francisco cratons, consistent with the existence of a Neoproterozoic belt between the two blocks, marked by the TBL. North of the SFC, the thin lithosphere under the Borborema Province appears as a weak positive anomaly in the *P*-wave tomography of Simões Neto et al. (2019) as well as in models SA2019 and SAAM23. At the beginning of the Neoproterozoic (900 Ma), during the breakup of Rodinia, the Borborema Province was an ocean separating the São Luís and West-African cratons from the São Francisco and Congo cratons, as the Goiano (or Climene) Ocean separated the Amazonian from the São Francisco-Paraná, and the Adamastor Ocean separated São Francisco-Paraná from the Congo-Kalahari (Hasui et al., 2012).

The Paranapanema Block (PpB) is an igneous block beneath the central part of the Paraná Basin (Fig. 1). Based on a few geological and many indirect (geological, geochemical, and isotopic) data, M. S. Mantovani and Brito Neves (2009) concludes that the PpB is predominantly of granitic nature (orthogneisses). Geotectonically, it is a segment of con-

tinental lithosphere that initially behaved as the upper plate during a subduction process (M. S. Mantovani & Brito Neves, 2009).

All aforementioned tomographic studies, including this one, show a high-velocity anomaly underlying the Paraná Basin and coinciding with the proposed location of the PpB (Cordani, 1984; M. Mantovani & Brito Neves, 2005). Heintz et al. (2005) found the high velocities down to 200 km depth, while Feng et al. (2007) observed high velocities to 100-150 km beneath the Paraná Basin, both consistent with the existence of a Proterozoic cratonic nucleus. In the NE portion of the Paraná Basin, an average to low-velocity lithosphere indicates that the limits of the PpB do not reach that far or that the original cratonic basement was affected by Late Cretaceous intraplate magmatism (M. Rocha et al., 2019). Both Feng et al. (2007) and Affonso et al. (2021) imaged a separation between the PpB and the SFC in the form of a decreased high-velocity anomaly. Using the travel times of teleseismic *P* and *S* waves, VanDecar et al. (1995) imaged a cylindrical low-velocity anomaly in the deep upper mantle beneath the Paraná Basin, which was interpreted as a fossil plume. The existence of the cylindrical anomaly was confirmed by Schimmel et al. (2003). Van der Lee et al. (2001), Feng et al. (2007), M. Rocha et al. (2019), Celli et al. (2020), and this study also observed a low-velocity anomaly in the upper mantle beneath the Paraná Basin, albeit at different locations. Van der Lee et al. (2001) position the anomaly in the eastern portion of the basin, Feng et al. (2007) in the north, M. Rocha et al. (2019) in the northeast, and Celli et al. (2020) in the southern part.

The Parnaíba Block (PnB) is bounded in the east by the NE-SW oriented TBL, which alongside other Brasiliano shear zones, represents the major structural trend of the Borborema Province (Fig. 1) to the east and also controlled old troughs and Paleozoic depocenters (Castro et al., 2014). Feng et al. (2007) found high *S* velocities down to 100-150 km beneath the Parnaíba Basin in NE Brazil, where a cratonic nucleus (PnB) has also been proposed (Cordani, 1984). Celli et al. (2020) also imaged the cratonic root of the Parnaíba Basin as a high-velocity anomaly under the northern half of the basin.

The São Luiz Craton (SLC), located in the northeastern coast of Brazil (Fig. 1), is a small fragment of the West African Craton (Fig. 1) that remained in South America after the Pangea breakup in the Cretaceous (Castro et al., 2014). The Luis Alves Craton (Fig. 1), on the other hand, is a small block, having 275 km in length and 77 km in width, in the southeastern region of South America. It is composed mainly of Archean to Paleoproterozoic granulitic rocks (Betiolo et al., 2018). The closure of the Adamastor Ocean with the orogenic collage between PpB, Curitiba Terrane, and Luis Alves Craton must have occurred in the 600 to 570 Ma interval (Passarelli et al., 2018).

According to Rapela et al. (2007), the Río de la Plata Craton (RPC), the southernmost craton of the American continent (Fig. 1), is dominated by juvenile Palaeoproterozoic rocks but with some Archean sectors. The majority of the reconstructions place it at the core of southwestern Gondwana, bounded by the Sierras Pampeanas and the Precordillera Terrane on the proto-Pacific margin and the Cuchilla Dionisio Terrane on the proto-Atlantic margin. None of the cited studies, including ours, observed high-velocity anomalies consistent with the location of the RPC proposed by Rapela et al. (2011) or Oyhantçabal et al. (2011).

2.2 Patagonian Platform

The Patagonian Platform corresponds to the southern portion of the South American continent (Fig. 1) and evolved tectonically differently than the South American Platform, being involved in orogenic processes during the Phanerozoic, as can be seen by its younger age and present shape, dimensions and position among three active margins and one passive margin (Almeida et al., 2000). V. A. Ramos (2004) proposed that Patagonia is an allochthonous terrane incorporated into Gondwana in the Early Permian.

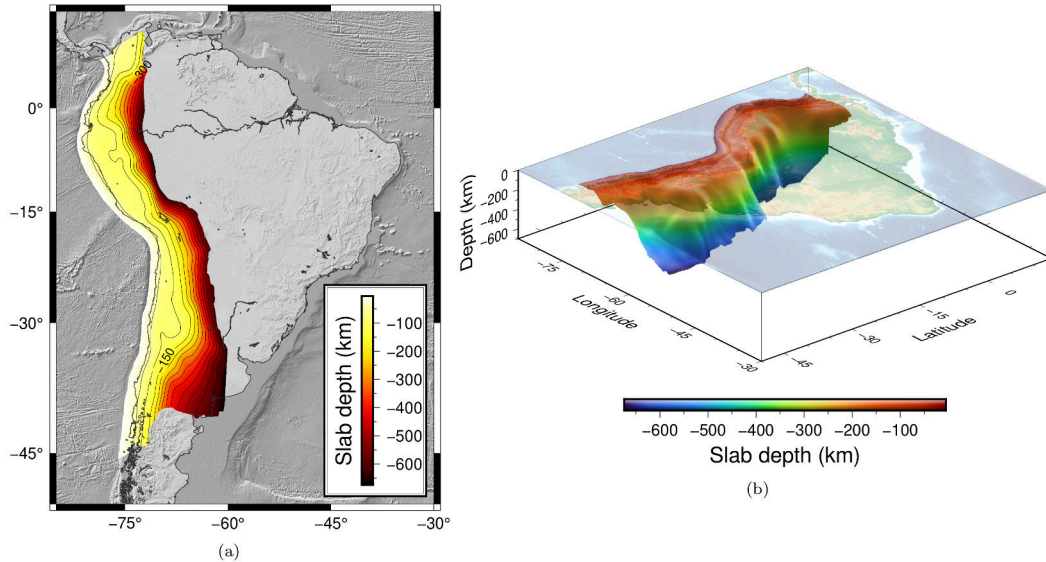


Figure 2. a) Two-dimensional representation of the Nazca Plate based on the Slab2 model. The model reaches 675 km depth in some regions. b) Same as a) but now represented in three dimensions. In both figures, it is clear that the subducting Nazca plate is more horizontalized beneath Peru (parallels 15°S to 0°) and in the southern portion of the trench (30°S).

2.3 Andean Belt and Nazca Plate

The Andes is the second-highest orogenic belt on Earth. It has been generated by Cenozoic tectonic shortening of the South American Plate’s western margin, which overrides the subducting Nazca Plate (Sobolev & Babeyko, 2005). Within the Andean Belt (Fig. 1), there are several cordilleras, sierras, plateaux, basins, and valleys. (Cordani et al., 2000). A review of Andean evolution is given by V. Ramos (1999). The Nazca Plate originated from the breakup of the Farallon Plate into the Nazca and the Cocos plates during the late Oligocene, around 23 million years ago (Walther & Flueh, 2002; Lonsdale, 2005). Y.-W. Chen et al. (2019) suggest that the Nazca-Farallon Plate subduction began in the late Cretaceous period at the northern Andes and propagated to the southern portion of the Andes by the early Cenozoic.

The convergence of the Nazca and the South American plates begot one of the most seismically active subduction zones on the planet, caused the uplift of the Andes, and created a chain of arc volcanoes along most of the mountain range. Lim et al. (2018) imaged the Nazca Slab in southern Peru using a combination of relocated earthquakes and double-difference tomography. They concluded that a weaker slab pull helps maintain the flat plate geometry near the subducting Nazca Ridge (NZR) as its negative buoyancy force alone would not be enough. The lower velocities in the flat slab found in their tomography suggest that, by supplying fluid into the system (Kim & Clayton, 2015), the subduction of the NZR retards the phase transition to a denser mineral, such as eclogite, further contributing to the positive buoyancy force that supports the slab in a horizontal position (Lim et al., 2018).

The seismicity and volcanic activity along the Andes vary spatio-temporally, following variations in slab dip, which is potentially influenced by variations in plate age upon subduction, spreading rates, convergence angles, the presence of extinct spreading ridges (e.g., the Easter Seamount Chain, the NZR, and Carnegie Ridge) on the subducting slab, or other complex geological processes (James, 1971; Ray et al., 2012; Hayes et al., 2015; V. Ramos, 1999). The majority of the large earthquakes in South America are confined to

shallow depths (0-70 km), resulting from crustal deformation and interplate slip (Hayes et al., 2015), including the 1960 9.5 Mw Valdivia earthquake, the largest seismic event ever recorded (Satake & Atwater, 2007). At the intermediate depth range (70 to 300 km), large earthquakes occur within the slab due to internal deformation and dehydration, and generally occur beneath northern Chile and southwestern Bolivia (Hayes et al., 2015). They also cluster beneath northern Peru and southern Ecuador, but with less intensity (Hayes et al., 2015). Earthquakes in South America do not occur between 300 and 500 km but are observed between 500 and 660 km within the Nazca Plate, concentrated into two zones: one that runs beneath the Peru-Brazil border and another that extends from central Bolivia to central Argentina (Hayes et al., 2015).

Van der Lee et al. (2001), Feng et al. (2007), and this study found a mantle wedge with extremely low S -velocities below the thick crust along with the volcanically active parts of the central Andes. The bottom of the low-velocity mantle wedge is immediately adjacent to the Wadati-Benioff zone and the subducting oceanic crust (Van der Lee et al., 2001). These low velocities in the mantle wedge can be explained by a 2.5-4% partial melt generated at typical mantle temperatures near 1,200 °C in the presence of realistic amounts (0.15 wt%) of water released from the subducting oceanic crust (Van der Lee et al., 2001). In areas where subduction occurs at relatively flat dip angles, the slab is imaged as moderately high S velocities, and a low-velocity mantle wedge is absent (Feng et al., 2007). Celli et al. (2020) imaged the Peruvian and Bucaramanga flat slab in the north of the Andes and the Nazca Slab and the Pampean slab gap in the central Andes. In the southern Cordillera, the subduction of the Chile Ridge, which separates the Nazca Slab from the Antarctic Plate, has opened up a gap in the subducted oceanic lithosphere. The existence of this slab window was first studied and imaged by Russo, Gallego, et al. (2010) and Russo, VanDecar, et al. (2010) using P -wave tomography and shear wave splitting. Fig. 2 shows the Nazca Plate three-dimensional geometry according to the Slab2 model of Hayes et al. (2018), which was constructed by combining active-source seismic data interpretations, receiver functions, local and regional seismicity catalogs, and seismic tomography models.

3 Adjoint Waveform Tomography

The goal of this study is to provide new constraints for the geological and tectonic processes in South America by constructing a new P - and S -wave model based on adjoint waveform tomography. At this stage, we exclude the amplitude information from waveforms to simplify the inverse problem and focus on the elastic structure by using the phase information only where we make phase measurements based on observed and computed synthetic waveforms. In this study, we use the SPEC-FEM3D_GLOBE package (Komatitsch & Tromp, 2002b, 2002a) to perform numerical simulations both for forward wavefield, to compute synthetic seismograms, and adjoint wavefield, to compute the gradient of our misfit function in the adjoint waveform inversion. The adjoint waveform inversion workflow is based on three basic stages (numerical simulations, pre-processing and post-processing), which closely follows e.g., Zhu et al. (2012); Bozdağ et al. (2016) except the pre-processing part (Fig. 3). The pre-processing stage involves data processing, selecting measurement windows, and the computation of adjoint sources. Unlike Zhu et al. (2015); Bozdağ et al. (2016); Lei et al. (2020), who used frequency-dependent cross-correlation based traveltime and amplitude misfits, we use an instantaneous phase measurement (Bozdağ et al., 2011), EP misfit (Yuan et al., 2020) for our phase measurements. For the selection of measurement windows, we used our Python-based window-selection algorithm, PyWinEPAdjoint (Ciardelli, 2021), instead of FLEXWIN (Maggi et al., 2009), which was used in Zhu et al. (2015); Bozdağ et al. (2016); Lei et al. (2020). The post-processing stage is where we update the model iteratively based on a conjugate-gradient optimization method (Nocedal & Wright, 2006). This section first gives information about our starting model and forward and adjoint simulations, then pre- and post-processing stages.

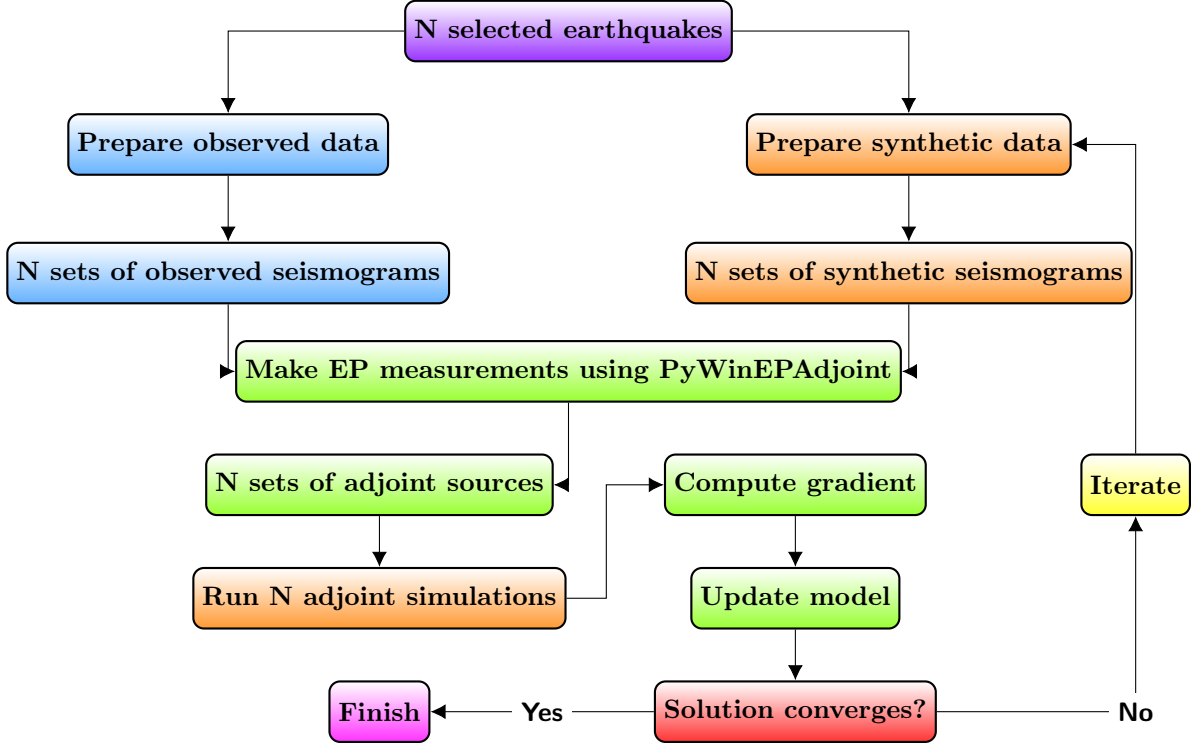


Figure 3. Adjoint tomography workflow. The preparation of the observed data stage comprises both the data request as well as the pre-processing: detrend, taper, remove response, filter, and downsample. The preparation of the synthetic data consists in running the mesher and the N forward simulations using the CMT source files. The measurement stage includes the selection of windows, the misfit measurements, and the computation of the adjoint sources. This is the stage in which our workflow differentiates from others (e.g., Bozdağ et al., 2016; Lei et al., 2020) as we use our own window selection algorithm, designed to create large windows, which are adequate for EP measurements (Yuan et al., 2020). The window selection and the creation of the adjoint sources were carried out using PyWinEPAdjoint. The computation of gradient comprises the summation of all kernels, smoothing, and preconditioning. Lastly, the model update stage includes the determination of the step length and the model update using the conjugate gradient method.

3.1 Starting Model and Model Parametrization

Our starting model M00 is a combination of the 3D mantle model S362ANI (Kustowski et al., 2008) and the 3D global crustal model CRUST1.0 (Laske et al., 2012). S362ANI is transversely isotropic in the uppermost mantle and results from a compilation of surface-wave phase speeds, long-period waveforms, and body-wave travel times. S362ANI also involves the topography of the 410 and 650 km discontinuities in the upper mantle. CRUST1.0 is a 1×1 degree resolution global crustal model that uses crustal thickness data from active-source seismic and receiver function studies and incorporates an updated version of the global sediments, which is the successor of model CRUST2.0 (Bassin, 2000) with a 2×2 degree resolution. Following the starting model S362ANI+CRUST1.0, in our parameterization, we keep transverse-isotropy in the upper mantle where the rest of the model, including the crust, is kept isotropic. A transversely isotropic model may be described using five elastic parameters: A , C , L , N , and F following the notation of Love (1927). A and C can be obtained by measuring the P -wave velocity perpendicular and parallel to the axis of symmetry:

$$A = \rho \alpha_h^2 \quad \text{and} \quad C = \rho \alpha_v^2, \quad (1)$$

where ρ is the density, α_h is the P -wave velocity polarized in the horizontal direction, and α_v the P -wave velocity polarized in the vertical one. Similarly, the relation between N and L with the horizontally (β_h) and vertically (β_v) polarized shear-wave velocities may be written as:

$$N = \rho \beta_h^2 \quad \text{and} \quad L = \rho \beta_v^2. \quad (2)$$

The last constant, F , is a function of the velocities at intermediate incidence angles and can be conveniently introduced through the dimensionless parameter $\eta = F/(A - 2L)$.

Following Zhu et al. (2012) and Bozdağ et al. (2016), we reduce the dependency of P and S wavespeeds to each other through the shear modulus by ignoring P -wave anisotropy and inverting for the bulk sound speed c in our inversions. Therefore, we have five parameters in our inversion: density (ρ), bulk-sound speed ($c = \sqrt{\kappa/\rho}$ where κ is the bulk modulus), vertically-polarized shear wavespeed ($\beta_v = \sqrt{L/\rho}$), horizontally-polarized shear wavespeed ($\beta_h = \sqrt{N/\rho}$), and the dimensionless parameter η .

The trade-off between density and wavespeeds is a well-known problem. Some researchers address it by attempting to simultaneously invert wavespeeds and density (e.g., Beller et al., 2018; Blom et al., 2020). To minimize the trade-off between parameters, we preferred to scale the density from the updated shear wavespeed model at each iteration similar to Zhu et al. (2012) and Bozdağ et al. (2016). Following Montagner and Anderson (1989), we use the following relation to scale density:

$$\delta \ln \rho = 0.33 \delta \ln \beta, \quad (3)$$

where β is the isotropic shear wavespeed estimated via the Voigt average (Babuska & Cara, 1991):

$$\beta = \sqrt{\frac{2\beta_v^2 + \beta_h^2}{3}}. \quad (4)$$

Finally, the gradient of our misfit function, which is discussed in Section 3.3.3, may be written as:

$$\delta\chi = \int_V (K_c \delta \ln c + K_{\beta_v} \delta \ln \beta_v + K_{\beta_h} \delta \ln \beta_h + K_\eta \delta \ln \eta) dV, \quad (5)$$

where K_c , K_{β_v} , K_{β_h} , and K_η are the Fréchet derivatives of c , β_v , β_h and η parameters and the logarithmic terms are the associated model perturbations.

3.2 Forward Simulations

In our simulations, we used SPEC-FEM3D_GLOBE (Komatitsch & Tromp, 2002b, 2002a), a 3D global seismic wave simulation package for forward simulations to compute synthetic waveforms. In SPEC-FEM3D_GLOBE, the globe consists of six chunks. Our study region fits in one chunk, capturing a 90×90 degree area. All seismic sources and stations are located within our computational domain, and simulations are performed for each earthquake where the cost of simulations is independent of the number of seismic stations. The mesh used to sample the background model honors the first-order discontinuities (Komatitsch & Tromp, 2002b) as well as the topographic variations on the upper-mantle discontinuities. Moho is also honored whenever the crustal thickness is less than 15 km and more than 35 km (Tromp et al., 2010). We used 3 and 1 spectral elements to sample the continental and the oceanic crusts in our study region. During the simulations, the topography and bathymetry of ETOPO1 (Amante & Eakins, 2009) are superimposed onto the crustal model. In addition, Earth’s rotation, self-gravitation (Cowling approximation) (Cowling, 1941), the ocean load, and 1D attenuation were taken into account. Both the Earth’s rotation and the self-gravitation effects become important only at long periods (> 100 s) (Komatitsch & Tromp, 2002a), which are above the maximum period we used in this study. Nevertheless, since taking these effects into account has little effect on the computational time of the simulations, we decided to include them, ensuring that our synthetic seismograms would be as accurate as possible. For the attenuation effects, we used the 1D Q model QL6 (Durek & Ekström, 1996) of our starting mantle model S362ANI (Kustowski et al., 2008), which is fixed during the inversion. During the first 15 iterations, we used a mesh with a resolution of $\text{NEX} = 128$ (NEX is the number of spectral elements on each side of our mesh at the surface) that can resolve periods down to ~ 34 s. During the next six iterations, we increased the mesh resolution to $\text{NEX} = 192$, which is capable of resolving periods down to ~ 23 s. For the last two iterations, we used $\text{NEX} = 256$, where the minimum resolvable period in our synthetics is ~ 17 s.

3.3 Pre-processing stage

3.3.1 Data selection

In this study, we used seismic data from 112 earthquakes (Fig. 4a) from the global CMT (Centroid Moment Tensor) catalog (Dziewonski et al., 1981; Ekström et al., 2012) between April 1994 and April 2018, and with the moment magnitudes (Mw) between 6 and 7. This Mw range is selected to maximize signal-to-noise ratios and avoid finite source effects in our numerical simulations considering the minimum period of our simulations. For each earthquake we computed synthetic seismograms recorded by the seismic stations shown in Fig. 4a by the SPEC-FEM3D_GLOBE (Komatitsch & Tromp, 2002b, 2002a) package. We downloaded the observed data from both permanent and temporary networks, including stations from USP, IRIS, and GFZ (see the complete list in Section 7). The events were selected by checking the waveform similarity between observed and synthetic waveforms. Since we perform numerical simulations for each earthquake, we tried to optimize the cost of simulations by selecting events with high-quality data while trying to preserve a good azimuthal coverage. We also used as many stations as possible, given that the number of receivers does not change the computational cost of simulations. At this stage, we performed a visual inspection to eliminate problematic data. Before we performed a quality check, observed and synthetic waveforms went through classical data processing

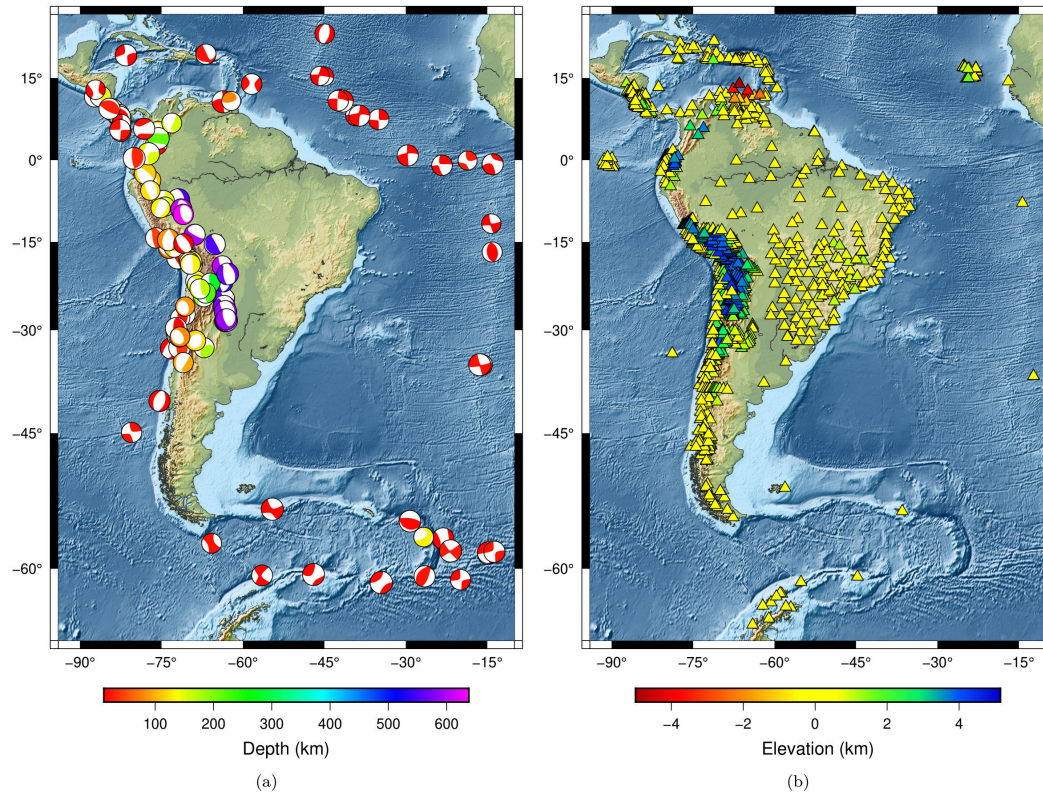


Figure 4. a) The 112 selected earthquakes. The size of the beachballs is related to the magnitude and the color indicates the depth. b) The 1311 selected stations. The color denotes the stations elevation. The mesher does not use those elevations, but places all stations at the surface, according to the topography values in ETOPO1. Thus, in the end, the elevation is taken into account.

steps, i.e., removing the mean and trend, tapering, and deconvolving the instrument response to displacement from observed data. The same data processing steps and filters were applied to observed and synthetic seismograms.

After computing the synthetic seismograms, we visually compared them to observed seismograms. We observed that some receivers had orientation issues, such as swapped signal polarity and swapped components (EW becomes NS and vice-versa). In such cases, we either corrected the metadata information whenever we could or discarded the data.

Another issue we considered was the waveforms where the amplitudes of observed and synthetic data showed large discrepancies (Fig. 5). Although our misfit function (Section 3.3.3) only measures phase differences, if the observed and the synthetic seismograms have large amplitude anomalies, this may suggest an issue with sensor installation or site or with the metadata in the response files. Hence, our automated time-window selection algorithm (Section 3.3.2) is set to reject those stations. Most of the time, the rejected stations were outliers located in densely covered regions, such as Chile. We visually monitored the data from receivers that are not part of any cluster of stations. After this visual inspection and manual cleaning, we kept data from 1311 stations. Fig. 4 shows the distribution of all the sources and receivers.

3.3.2 Measurement Window Selection

We further checked the quality of waveforms and select measurement windows using a Python-based automated window selection algorithm, PyWinEPAdjoint, that we developed specifically for this study. PyWinEPAdjoint utilizes fewer selection parameters than FLEXWIN (Maggi et al., 2009), but has extra parameters designed for quality control purposes, such as eliminating problematic data before the adjoint inversion (Section 7). PyWinEPAdjoint uses a 4-stage algorithm, in which the first three steps detect and reject noisy components (Fig. 5a). The last one carries out the time-window selection based on waveform and amplitude similarity. Before selecting the measurement windows, we process observed and synthetic data in three-period bands. Fig. 5 displays the three-period bands we used in this study: 17 to 45 s (Fig. 5b), 30 to 60 s (Fig. 5c), and 45 to 100 s (Fig. 5d).

3.3.3 Measurements and Computation of Adjoint Sources

The last stage of the pre-processing is extracting measurements from each selected time window, based on a chosen misfit function, and compute associated adjoint sources. In this study, we chose the EP misfit (Yuan et al., 2020), which is an extension to the instantaneous phase measurement by Bozdağ et al. (2011). The reasons for this choice are: 1) We try to linearize the inverse problem first focusing on the elastic model, which is more linearly related to the phase information; 2) EP measurements can be applied to wavetrains, which do not require as many time windows as cross-correlation measurements need to highlight the small amplitude phases in the total gradient; 3) EP better deals with the cycle skip problem of phase measurements by tapering phase shifts larger than $\pi/4$ (Yuan et al., 2020).

Following Yuan et al. (2020), we define the EP based on the observed $\tilde{d}(t)$ and synthetic $\tilde{s}(t)$ analytic signals normalized by their envelopes:

$$\chi_{EP} = \frac{1}{2} \sum_{s,r} \int_0^T \left\| \tilde{d}_i(t) - \tilde{s}_i(t) \right\|^2 dt, \quad (6)$$

where $\tilde{d}_i = e^{i\phi(t)}$ and $\tilde{s}_i = e^{i\phi_s(t, \mathbf{m})}$ are the normalized analytical signals for observed and synthetic data, respectively. ϕ and ϕ_s are the instantaneous phase of observed and synthetic data, respectively. s and r indices indicate the number of sources and receivers. The associated adjoint source for the EP misfit then becomes:

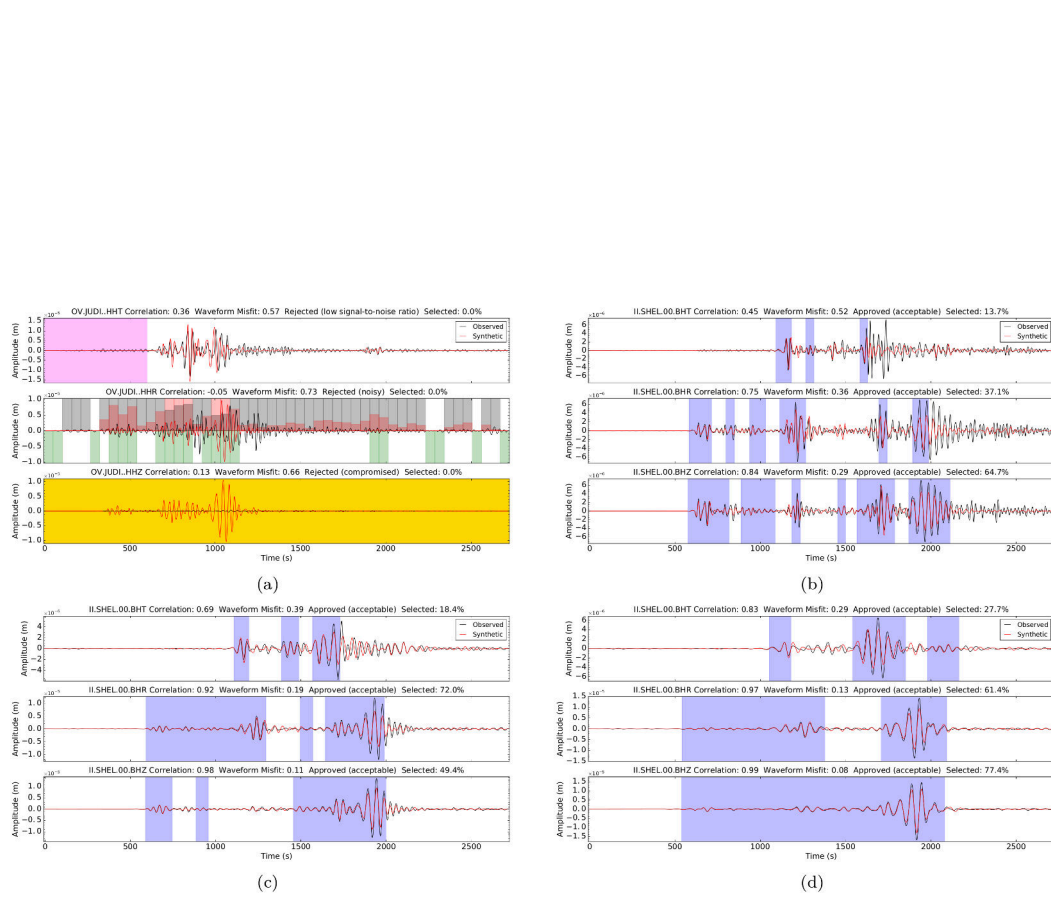


Figure 5. a) Automated detection and removal of bad data. Compromised or noisy seismograms are identified by: 1) Highly different overall amplitudes (Z-component of Fig. a), marked in yellow); 2) Large amplitude signal before P -wave arrival (T-component of Fig. a). The trace segment used to calculate the signal-to-noise ratio is marked in magenta); 3) Very different amplitudes in more than 20% of the record duration (R-component of Fig. a). The algorithm splits the trace into 50 windows and, for each of them, the average value of signal module is computed both for the observed and the synthetic seismograms. If the ratio between these two numbers exceeds 0.35 the window is rejected. If more than 10 windows were rejected, the whole component is discarded. Windows accepted are marked in green. Red and gray bars show the amplitude ratios. b) Selected windows in the 17-45 s period band. c) Selected windows in the 30-60 s period band. d) Selected windows in the 40-100 s period band.

$$\begin{aligned}
f_{EP}^\dagger = \sum_r \left[\Delta I(t) \frac{s(t) \mathcal{H}[s(t)]}{E_s^3(t)} - \Delta R(t) \frac{\{\mathcal{H}[s(t)]\}^2}{E_s^3(t)} \right. \\
\left. + \mathcal{H} \left\{ \Delta I(t) \frac{s^2(t)}{E_s^3(t)} - \Delta R(t) \frac{s(t) \mathcal{H}[s(t)]}{E_s^3(t)} \right\} \right] \delta(x - \mathbf{x}_r), \quad (7)
\end{aligned}$$

where \mathcal{H} denotes the Hilbert transform and \mathbf{x}_r is the location of the adjoint source, which is the same as that of the station that recorded the signal. $\Delta R(t)$ stands for differences in real part and $\Delta I(t)$ stands for the difference in the imaginary part of the analytic signal (Yuan et al., 2020).

To balance the uneven distribution of earthquakes and seismic stations we introduced a geographical weighting to our misfit function, following (Ruan et al., 2019). For each receiver r_i , a weight w_i is calculated according to Eq. 8:

$$w_i = \left[\sum_{j=1}^N e^{-\left(\frac{D_{ij}}{\Delta}\right)^2} \right]^{-1}, \quad (8)$$

where D_{ij} are the distances between two stations for all the N pairs that include r_i and all the other receivers. A reference distance Δ is used so that the condition number of the matrix given by 8 is not too large. In practice, it is determined through a linear search so that its value is around 35% of the maximum (Fig. 6b). We used the same scheme to balance the source distribution (Fig. 6a). Both for sources and receivers, D_{ij} was measured in 3D.

3.4 Adjoint Simulations

Initiated by the adjoint sources computed based on the EP misfit, the adjoint simulations were numerically performed with the SPEC3D_GLOBE package. Similar to forward simulations, topography/bathymetry, the ocean load, gravity, rotation, ellipticity, and attenuation were all considered in adjoint simulations. We carried out the adjoint simulations to compute the Fréchet derivatives (Tromp et al., 2005) where the result of each simulation per earthquake gives the summation of Fréchet kernels, which lead to event kernels (Tape et al., 2007). Since the first iteration, we took into account the full attenuation in the adjoint simulations (Komatitsch et al., 2016) to ensure accurate sensitivity kernels.

3.5 Post-processing Stage

Once we computed event kernels for 112 earthquakes for each iteration, we first summed them up to obtain the gradient of the misfit function (Eq. 6) for each model parameter and the pseudo-Hessian kernel (Eq. 9). The gradients were multiplied by the source weights. During the summation, we masked the sources to mitigate the formation of numerical artifacts.

We smoothed the gradients to minimize the numerical noise and help balance the imperfect data coverage further (Bozdağ et al., 2016). Although we filtered adjoint sources with the same band-pass filter applied to seismograms before running the adjoint simulations, we smoothed the gradients of all the parameters using a Gaussian function in the vertical and horizontal directions (Zhu et al., 2015) to remove remaining short-wavelength signals. The amount of smoothing depends both on the shortest-resolvable wavelength and the data coverage. Hence, the crust and the upper mantle, in which the wavelength is shorter and the data coverage is better, require less smoothing than the transition zone and the lower mantle. In the initial iterations, we used a horizontal smoothing with a radius of 150 km

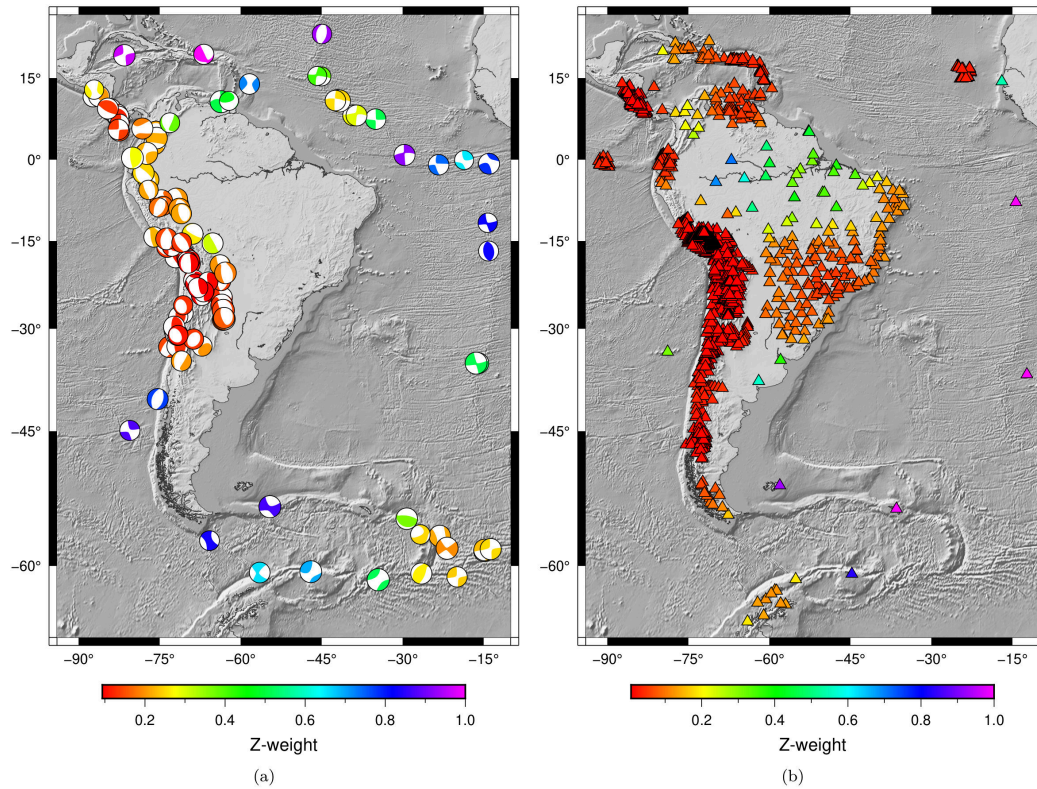


Figure 6. a) Source weights are applied to each earthquake to balance the geographical distribution by attributing larger weights to more isolated events. b) Same as a) but for stations. The receiver weights are applied to the adjoint sources (Section 3.3.3) whereas the sources weights are applied to the event kernels (Sections 3.4 and 3.5).

and a vertical smoothing with a radius of 15 km for the crust and the upper mantle. These values gradually increased by depth where they were 200 km and 100 km, respectively, at the bottom of the lower mantle. Smoothing was also gradually decreased as the iterations proceeded, progressively including finer details into the model. In the final iterations, we reduced the horizontal smoothing radius to 80 km and the vertical smoothing radius to 5 km for the crust and the upper mantle, and to 120 km and 60 km, respectively, at the bottom of the lower mantle.

We then weighted the gradients by the inverse of the pseudo-Hessian kernel to speed up the convergence rate of iterations. Ideally, we would use the complete Hessian matrix in the optimization process. However, since that would be prohibitively expensive (Fichtner, 2010), we replaced it by the so-called pseudo-Hessian $P(\mathbf{x})$, which approximates the diagonal terms of the Hessian matrix and can be computed using the second temporal derivatives of the forward and adjoint displacement wave fields $s(\mathbf{x}, t)$ and $s^\dagger(\mathbf{x}, t)$ (Luo et al., 2013):

$$P(\mathbf{x}) = \sum_{e=1}^E \int_0^T \partial_t^2 s(\mathbf{x}, t) \cdot \partial_t^2 s^\dagger(\mathbf{x}, T - t) dt. \quad (9)$$

The reason of choosing the pseudo-Hessian as a preconditioner is that it resembles data coverage (Luo et al., 2013), which helps balance the gradient further while suppressing high-amplitude values at source and receiver locations.

3.5.1 Model Update

During the first iteration, we use the resulting preconditioned gradients to update the starting model using the steepest descent method (Debye, 1909). From the second iteration on, we moved to the conjugate gradient method (Fletcher & Reeves, 1964) to speed up the convergence. We can keep using the conjugate gradient method as long as we keep the measurement windows fixed or until the gradients lose conjugacy (Fichtner, 2010). Both the steepest descent and the conjugate gradient methods provide the direction in which we should update the model d_i to have the fastest convergence. We then performed a line search Tape et al. (2007) to determine the step length α for the model update at each iteration. In this study, the line search consists of running forward simulations for a subset of 40 earthquakes for various values of α , typically ranging from 0 to 3.5% perturbations in the search direction d_i . Using these values (at least four, in our case), we fitted a polynomial and used its minimum as the step size. Then, we updated the model parameters using Eq. 10:

$$\ln\left(\frac{m_{i+1}}{m_i}\right) = \alpha d_i. \quad (10)$$

Pratt et al. (1998) showed that Newton-like methods generally provide faster convergence rates than conjugate gradient in seismic waveform inversion. L-BFGS, a limited-memory quasi-Newton method, is useful for solving large problems whose Hessian matrices cannot be computed at a reasonable cost or are too dense to be manipulated easily (Nocedal & Wright, 2006). In future studies, we consider switching to L-BFGS to speed up our inversions and save computational time.

3.6 Computational Requirements

As explained in Section 3.3.3, using the phase information only in the misfit function is one way to mitigate the non-linearity of the inverse problem. The multi-scale approach, which consists of fitting the long-wavelength data first and then gradually moving on to shorter wavelengths, is another common method to mitigate the problem since fewer local minima exist at larger scales (e.g., Bunks et al., 1995; Van der Lee & Nolet, 1997). In our

Source inversions	1 event (forward)	1 event (adjoint)	112 events (forward + adjoint)
Type I inversion CPU-h, NEX = 128, $T_{min} \sim 34$ s	~ 13	~ 43	$\sim 6,212$
Type II inversion CPU-h, NEX = 192, $T_{min} \sim 23$ s	~ 27	~ 81	$\sim 12,066$
Type III inversion CPU-h, NEX = 256, $T_{min} \sim 17$ s	~ 56	~ 149	$\sim 22,938$
15 \times Type I + 6 \times Type II + 2 \times Type III CPU-h, Total (23 iterations)	~ 469	$\sim 1,429$	$\sim 211,452$

Table 1. Computational time required to carry out all iterations. In the first 15 iterations, we used a 128-resolution mesh. In the following 6, we increased the resolution to NEX = 192 spectral elements. Finally, in the last 2 iterations, we used 256 elements. We neglected time used in pre- and post-processing stages as it is small compared with the resources needed to run the forward and the adjoint simulations. T_{min} indicates the minimum period of the synthetic seismograms resolved by the simulations. NEX is the number of spectral elements on each side of the computational domain at the surface.

workflow, we carried out 15 low-resolution simulations, capable of modeling periods down to ~ 34 s, six simulations accurate down to ~ 23 s, and two simulations that can model periods down to ~ 17 s. Table 1 summarizes the computational resources consumed by all iterations, including the forward and the adjoint simulations.

4 Assessment of the Model Quality

In this section, we assessed the quality of our model by: 1) Monitoring the misfit evolution; 2) Using cross-correlation traveltime measurements for an independent set of earthquakes not used in the inversion to check the improvement in waveforms by component and period band; 3) Performing Point Spread Function (PSF) tests (Fichtner & Trampert, 2011). In addition, we computed mean correlation coefficients by source and receiver to estimate the improvement in waveform fitting per event and per station.

4.1 Misfit Evolution

We use the misfit evolution as a metric to assess if and how much the model improves after each iteration. In the beginning, especially in the first iteration, generally, a large reduction in the misfit is expected. As the inversion progresses, the misfit gradually decreases after each iteration until the decrease becomes negligible, signaling that we are near a minimum of our objective function. Fig. 7 summarizes our results after 23 iterations for the total misfit and misfit reduction in each measurement category (period band and component).

We used a single period band during the first five iterations, including signals between 50 and 100 s. We chose 50 s as the minimum period based on our starting model, which is accurate to explain down to 50 s surface waves. The maximum period was set to 100 s because that was the upper period band we had to use during the receiver response removal stage to avoid amplification of low-frequency noise in the data. After reaching a nearly 20% reduction in the overall misfit, we recomputed the time windows. Due to the higher similarity between the observed and the synthetic seismograms, the algorithm can select more data, enlarging and merging many windows. That inclusion of additional data explains the increase in the misfit from iteration 4 to 5. Using the new windows, we carried out two additional iterations (5 and 6). Because the reduction of residuals progressively slowed down, we decided to reduce the lower period of our band from 50 to 40 s, including higher

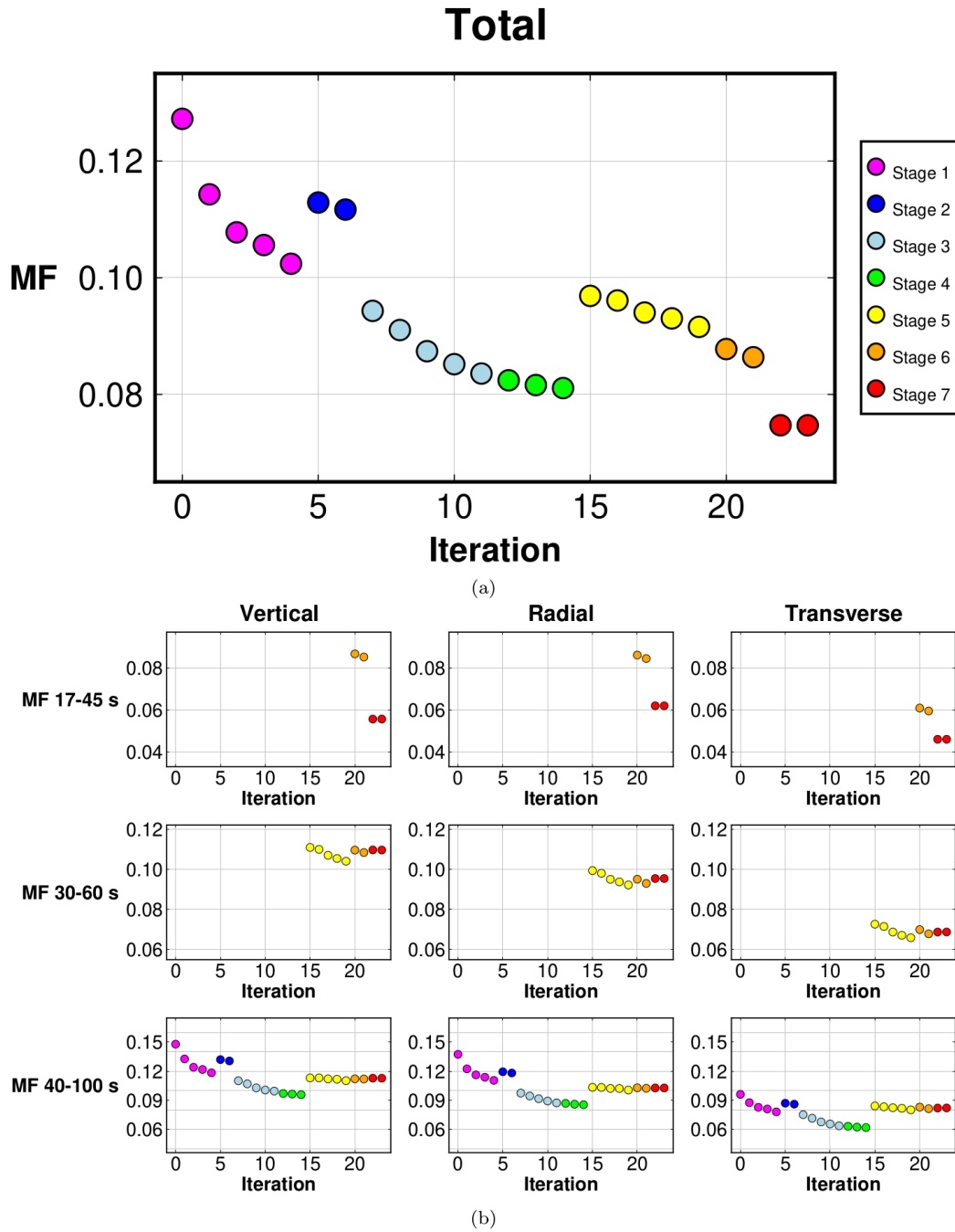


Figure 7. Upper large figure a): Overall misfit drop after each iteration. Different colors denote different time windows. Usually, an increase in misfit from one stage to another is caused by larger windows including more data or by the addition of another period band. On the other hand, decreases in misfit between stages are a combination of model improvement with reduction of data selected. Lower small figures b): Misfit drop for each category separately. The reason why the misfits are smaller in the transverse component is because less data is selected in that component, reducing its contribution to the overall misfit. In all cases, the misfit is dimensionless.

frequency signals. We performed five more iterations with the newly selected measurement windows in which we constantly observe a decrease in misfit.

At iteration 11, we noticed short-wavelength artifacts in the lower mantle, indicating that we needed to smooth the gradients below the transition zone further. We first filtered the lower mantle by using the same smoothing applied to the gradients. The filtering process increased the misfit by a negligible amount ($\sim 0.1\%$), as our measurements have little sensitivity to the lower-mantle structure. No sharp change in the misfit occurred from iteration 11 to 12.

After performing three more iterations, the misfit nearly stagnated. At that point, to be able to reduce the minimum period of measurements, we needed to increase the resolution of our simulations by interpolating the $NEX = 128$ resolution mesh to a $NEX = 192$. The higher resolution can accurately simulate wave propagation down to a minimum period of ~ 23 s, allowing the introduction of the second-period band, ranging from 30 to 60 s. Then, to compute the gradient using multiple frequency bands, we computed the adjoint sources for each band and summed them up with appropriate weightings to balance the different window lengths of each period band.

The higher-frequency data reduced the waveform similarity between observed and synthetic seismograms, where the window-selection algorithm selects more data in the lower-frequency bands. To ensure that each frequency band had an equal contribution to the overall misfit, we weighted the adjoint sources of each period band by a constant inversely proportional to the amount of data selected before summing them up. We kept the windows fixed from iteration 15 to 19, entering another cycle of continuous misfit reduction. At iteration 20, we added a third frequency band, ranging from 23 to 45 s. Then, we recomputed the windows and carried out two additional iterations.

Finally, at iteration 22, we interpolated the mesh once more to increase its resolution from $NEX = 196$ to $NEX = 256$. Then, we decreased the lower period of measurements from 23 to 17 s. Six seconds of period reduction represents a $\sim 26\%$ increase in frequency. At that period range, such an increase had an important impact on the waveforms' complexity, causing a considerable reduction in the amount of data selected, observed as a large drop in the misfit from iteration 21 to 22. We recomputed the windows the last time and carried out two more iterations, but the misfit remained nearly unchanged, so we ended the inversion at iteration 23.

Since we changed the windows used to measure the misfit, the stages of the misfit evolution cannot be directly compared to one another. Thus, we cannot estimate the overall misfit drop after the 23 iterations directly from Fig. 7. The proper way of doing this is running a $NEX = 256$ resolution forward simulation with M00, creating synthetic seismograms accurate down to ~ 17 s, and applying the same windows used in stage 7 (Fig. 7) to measure the misfit for the three frequency bands. The resulting misfit is fairly comparable with the misfit of the last iteration and shows an overall reduction of 50% from M00 to M23.

4.2 Cross-correlation Traveltime Residuals

The misfit reduction attributes a single value for each period band and component. We measure the overall reduction in the phase residual for all the waveforms, but we do not know the behavior of the residuals individually in each segment of seismograms. Another limitation is that we measure the misfit in the same data set used in the inversion. A more robust and informative metric would assess the model improvement based on an independent data set that was not used in the inversion. To this end, we computed the misfit based on cross-correlation traveltime measurements where we re-selected smaller measurement windows appropriate for cross-correlation measurements. To compute the histograms, we downloaded the waveforms of all the events with $M_w = 6 - 7$ from the GCMT catalog in the region of interest that occurred between August 14, 2018, and November 3, 2020 (53 events

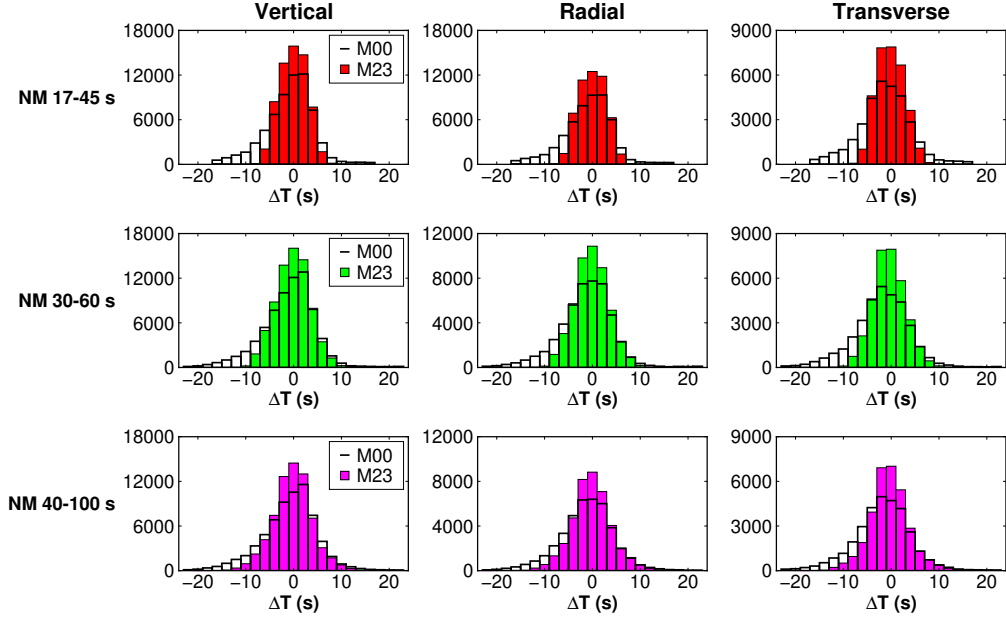


Figure 8. Comparison between the travel-time residuals histograms of M00 (black line) and M23 (colored bars) grouped by component and period band.

altogether), a period not used in the inversion. Then, we ran $NEX = 256$ resolution forward simulations to generate the corresponding synthetic seismograms both for the starting model (M00) and for the final model (M23). Fig. 8 shows all the residuals grouped by component and period band. The histograms show that even using an independent set of data, we observe a substantial improvement from M00 to M23 based on cross-correlation traveltimes.

The most obvious improvement after 23 iterations is the narrower histograms for all the nine groups (Fig. 8) nicely centered around zero. Our observations based on cross-correlation measurements are also important to show the robustness of EP measurements which, to the best of our knowledge, is used in a 3D adjoint tomography study with real data for the first time.

4.2.1 Correlation by Source and Receiver

Another useful analysis to assess the improvement of the final model over the initial one is comparing the mean correlation coefficient per event and per station for the original data set (the same used in the inversion). However, unlike the previous subsections, this time, we do not use any windows. For each trace that was accepted for measurement (i.e., not rejected by the windows selector), we calculate the correlation coefficient between the observed and synthetic as the dot product between the normalized traces filtered from 40 to 100 s, without excluding any segment of the seismogram and without shifting any of the traces. That coefficient is a measure of the waveform similarity between the two traces.

By computing the mean of the coefficients, we evaluate the overall waveform similarity for each source or receiver. Since we are not subdividing the seismograms into multiple disconnected windows, the correlation coefficient is dominated by the largest amplitudes in them, which are the surface waves. Despite this limitation, this analysis may be useful to identify outliers (events or stations) and to assess the model quality and improvement per region. Fig. 9 shows that for both sources and receivers, there is an improvement in waveform similarity after 23 iterations. For example, as shown in Fig. 9b, we notice that

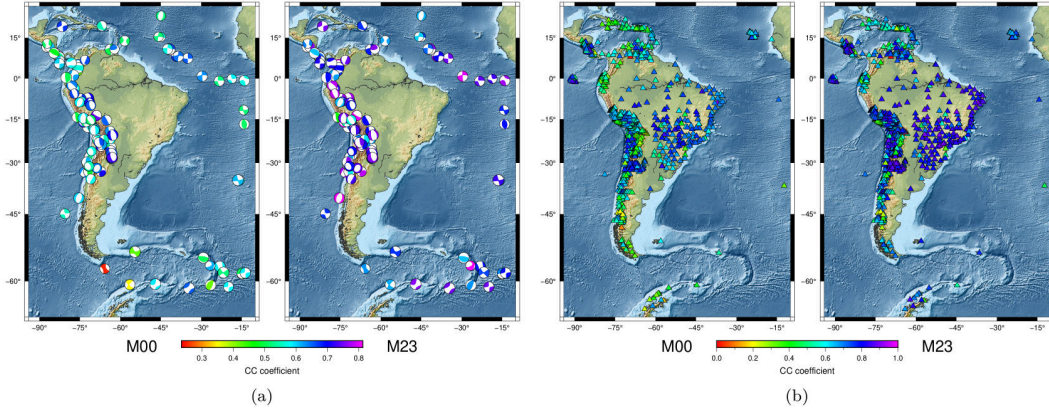


Figure 9. a) Mean correlation of all the 1311 stations per event for M00 and M23. The overall correlation coefficients for M00 and M23 are 0.58 and 0.70, respectively. b) Mean correlation of all the 112 events per station for M00 and M23. The overall correlation coefficient for M00 and M23 are 0.55 and 0.65, respectively.

the model better explains waveform data in Brazil, the central Andes, and oceanic regions. Less impressive fits are achieved in the southern portion of South America and especially in the Caribbean, which are both tectonically highly complex regions.

4.3 Resolution Tests

The inversion of synthetic data generated for checkerboard or tectonic-structure models are commonly used in linearized seismic tomography to have an idea about the resolution power of data used in inversions (e.g., Grand, 1987; Inoue et al., 1990; Fukao et al., 1992; Zelt, 1998; Feng et al., 2007; Celli et al., 2020). Despite the relevance of those tests, they are infeasible in adjoint tomography, as they cost as much as the actual iterative inversions. Fichtner and Trampert (2011) introduced 'Point Spread Functions' (PSF) into adjoint inversions as a way of directly estimating the model resolution at a given spot using a finite-difference approximation to calculate the local action of the Hessian $H \cdot \delta m$ without the need of the actual Hessian H (Eq. 11):

$$H \cdot \delta m \approx g(\tilde{m} + \delta m) - g(\tilde{m}), \quad (11)$$

where H represents the Hessian, δm is the localized Gaussian model perturbation with respect to the optimal model \tilde{m} , $g(\tilde{m})$ is the gradient at \tilde{m} , and $g(\tilde{m} + \delta m)$ is the gradient at $\tilde{m} + \delta m$. The principle underneath the method is that in the vicinity of the optimum model, the Hessian describes the convexity of the objective function, providing a direct measurement of the resolution and trade-offs in the misfit caused by δm (Fichtner & Trampert, 2011).

Fig. 10 shows resolution tests for four regions we selected based on data coverage and tectonic and geological features: the middle of the Central Brazil Shield (Figs. 10a and 10b) and the PpB (Figs. 10c and 10d) at 110 km depth, the Nazca Plate beneath the 19°S (Figs. 10e and 10f) at 1,100 km depth, and south Brazil (Figs. 10g and 10h) at 2,500 km depth. The two Gaussian perturbations at 110 km depth have a diameter of 100 km. The perturbation at 1,100 km is 130 km in diameter and the perturbation at 2,500 km is 150 km. All four anomalies are spherical. We clearly see that the deeper we introduce the perturbation, the more blurring is observed in the recovery results. The amount of blurring is a direct estimate of the expected resolution.

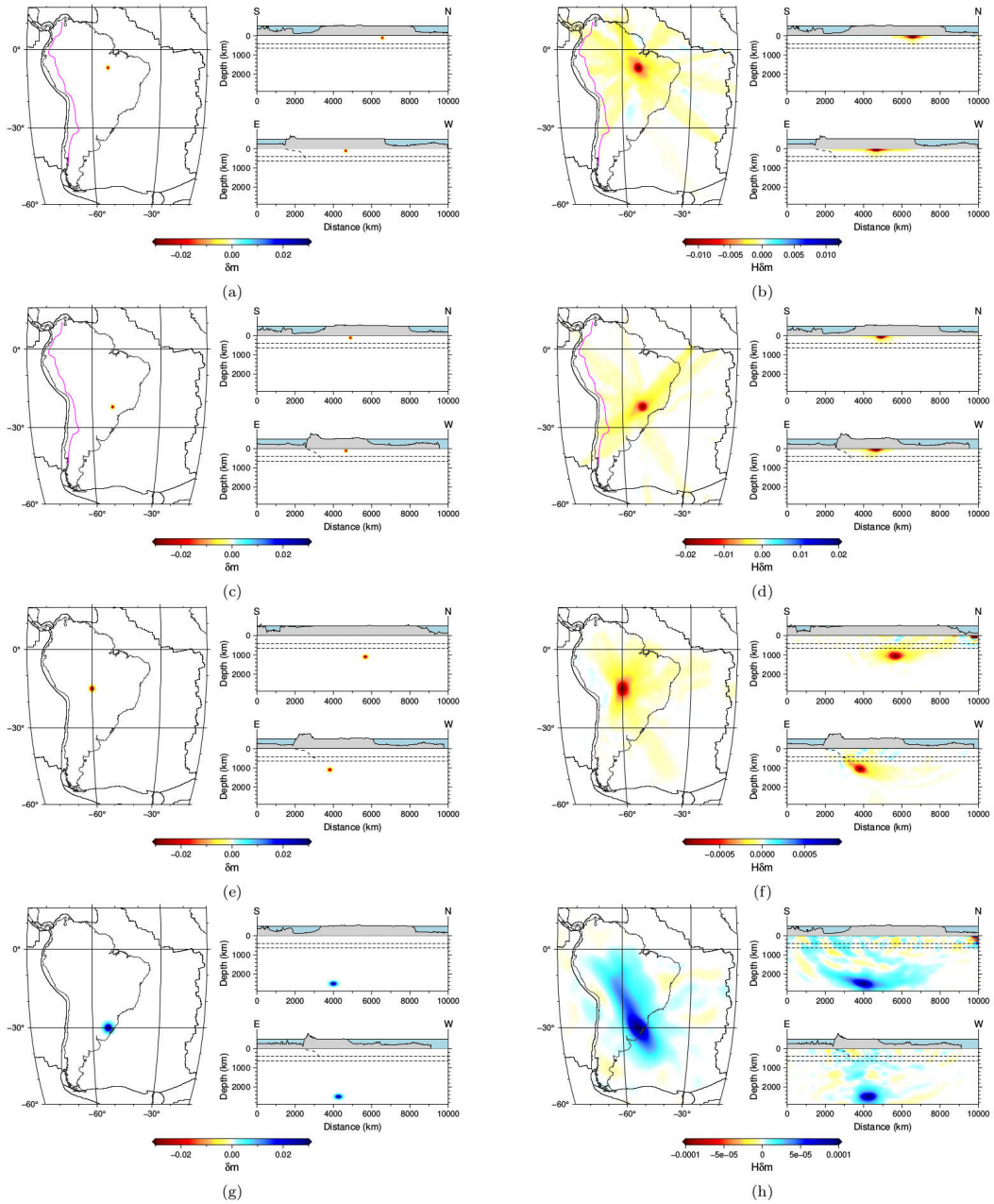


Figure 10. Resolution analysis for four points in the model. Figures to left (a, c, e, and g) show the perturbations δm , and figures to the right (b, d, f, and h) show the corresponding action of the Hessian $H \cdot \delta m$. The widths of the Gaussian perturbations were adjusted roughly to match the minimum wavelength expected to be resolved at each depth, according to the velocities and the minimum resolved periods, ranging from 100 to 150 km. The perturbations are 0.03 in magnitude, positive at locations with negative velocity anomalies, and negative in the opposite case.

According to the tests, the vertical resolution of the model at 110 km depth ranges between 150 to 200 km. The resolution decreases to nearly 400 to 200 km at 1,100 km depth and 700 to 900 km at 2,500 km depth. The horizontal resolution corresponding to these same depths are, respectively: 600 to 1,000 km, 1,100 to 1,400 km, and 1,500 to 3,000 km.

5 Results

In this section, we interpret our adjoint waveform model SAAM23 and how it compares with previous studies.

5.1 Overview

Fig. 11 shows M00 (S362ANI + CRUST1.0), our starting model, and the final model M23 (SAAM23). In SAAM23, we identify structures such as the Mid-Atlantic Ridge (although the resolution at the ridge, likely, is not great and may potentially be biased due to data coverage), including the Romanche Fracture Zone, a low-velocity anomaly beneath the South Sandwich Islands subduction zone, and beneath the spreading center connecting the Cocos and the Nazca plates to the north of the Galapagos Islands. In the South Atlantic region, the contour of the Argentine Basin is imaged as a high-velocity ring circumventing the sedimentary basin. At 30°S, a distinct high-velocity anomaly represents a shallow-dipping Nazca Plate (Pampean flat slab). Strong high-velocity anomalies beneath the Paranapanema and the Parnaíba blocks and the SFC represent the cool mantle lithosphere. Similar to the previous tomographic studies, the high-velocity lithospheric mantle is not imaged beneath the RPC. The Borborema Province appears as a low-velocity anomaly in the far northeast of Brazil. Finally, high-velocity anomalies are imaged beneath the cratonic nuclei in Central Brazil and the Guyana shields of the AC. These anomalies weaken to the southwest, towards the younger lithosphere of the Rondonian-San Ignacio and Sunsas-Aguapei geochronological provinces.

5.2 The Deep Nazca Slab

The subducting Nazca lithospheric slab is imaged as a narrow high-velocity belt parallel to the Andean continental margin (Fig. 12). Here we show some of the main features of the subducted Nazca plate as imaged in SAAM23 and compare them with previous models (e.g., Montelli et al., 2006; Li et al., 2008; Simmons et al., 2012; Obayashi et al., 2013; Lei et al., 2020; Portner et al., 2020; Rodríguez et al., 2021; Celli et al., 2020; Gao et al., 2021).

5.2.1 Slab Continuity Near Peru

The continuity of the flat slab segment beneath Peru into the deeper upper mantle has not been consistently imaged. Some global models such as UU-P07 (Amaru, 2007), MITP08 (Li et al., 2008) and Detox-P1 (Hosseini et al., 2020) show a gap in the high velocity belt in the depth range 300 to 500 km (Fig. 14). The South American regional model of Celli et al. (2020) images the flat slab beneath Peru but not between 250-400 km depth. Seismicity in the Benioff zone is not continuous: earthquakes occur down to ~250 km, and then near 660 km, with no activity in between. This could lead to an interpretation of the absence of the Nazca Slab in that region. However, other global models such as PRI-P05 (Montelli et al., 2006) and G3Dv3 (Simmons et al., 2012) show a continuity of the high-velocity slab in that region, albeit with lower amplitude. GLAD-M25 (Lei et al., 2020) also shows high velocity anomalies in the 300-400 km depth range, despite the lower resolution (Figs. 12 and 13). The regional model SAM5-P-2019 of Portner et al. (2020) shows alternating high and low velocities in that region. Moreover, James and Snoke (1990) inferred a continuous slab based on an observed seismic phase that most likely reflected off of the slab surface in the purported gap.

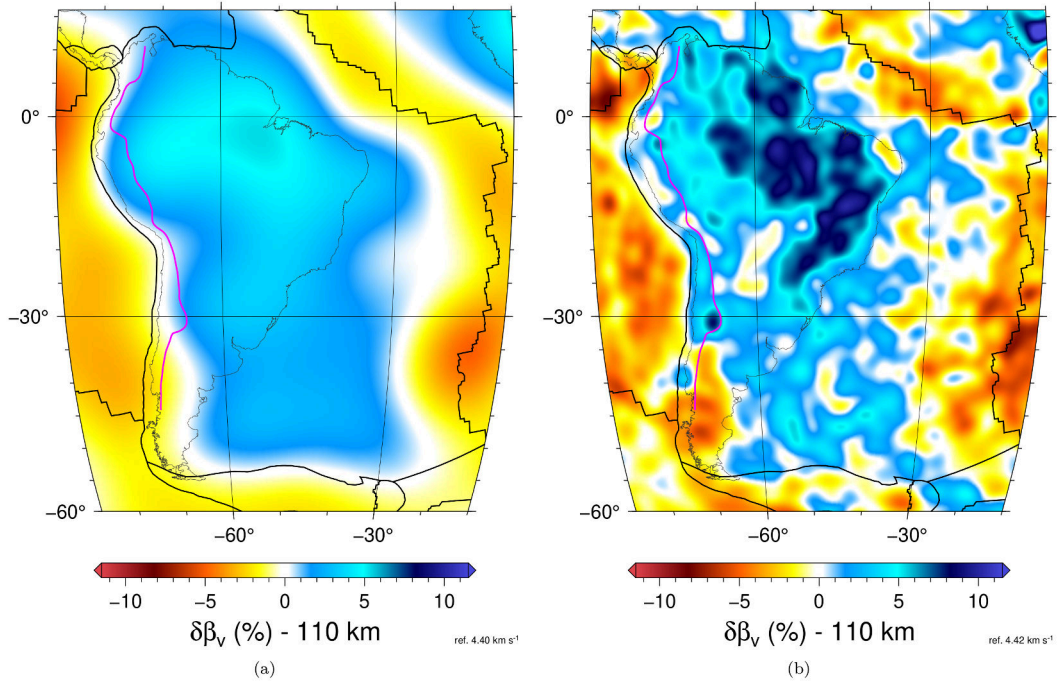


Figure 11. a) β_v anomalies at 110 km depth for the starting model (M00). The magenta line shows the top contour of the Nazca Slab according to Slab2. b) Same as Fig. a), but now showing the result after 23 iterations (M23).

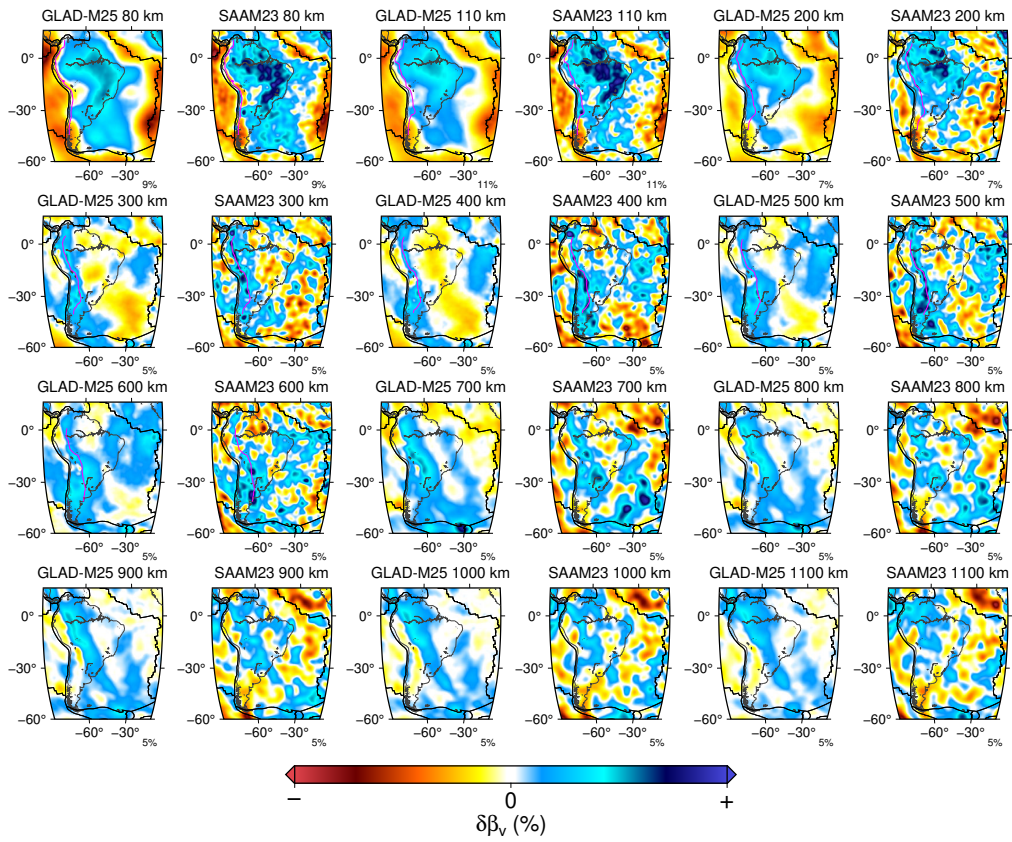


Figure 12. Depth slices of GLAD-M25 (Lei et al., 2020) compared with SAAM23.

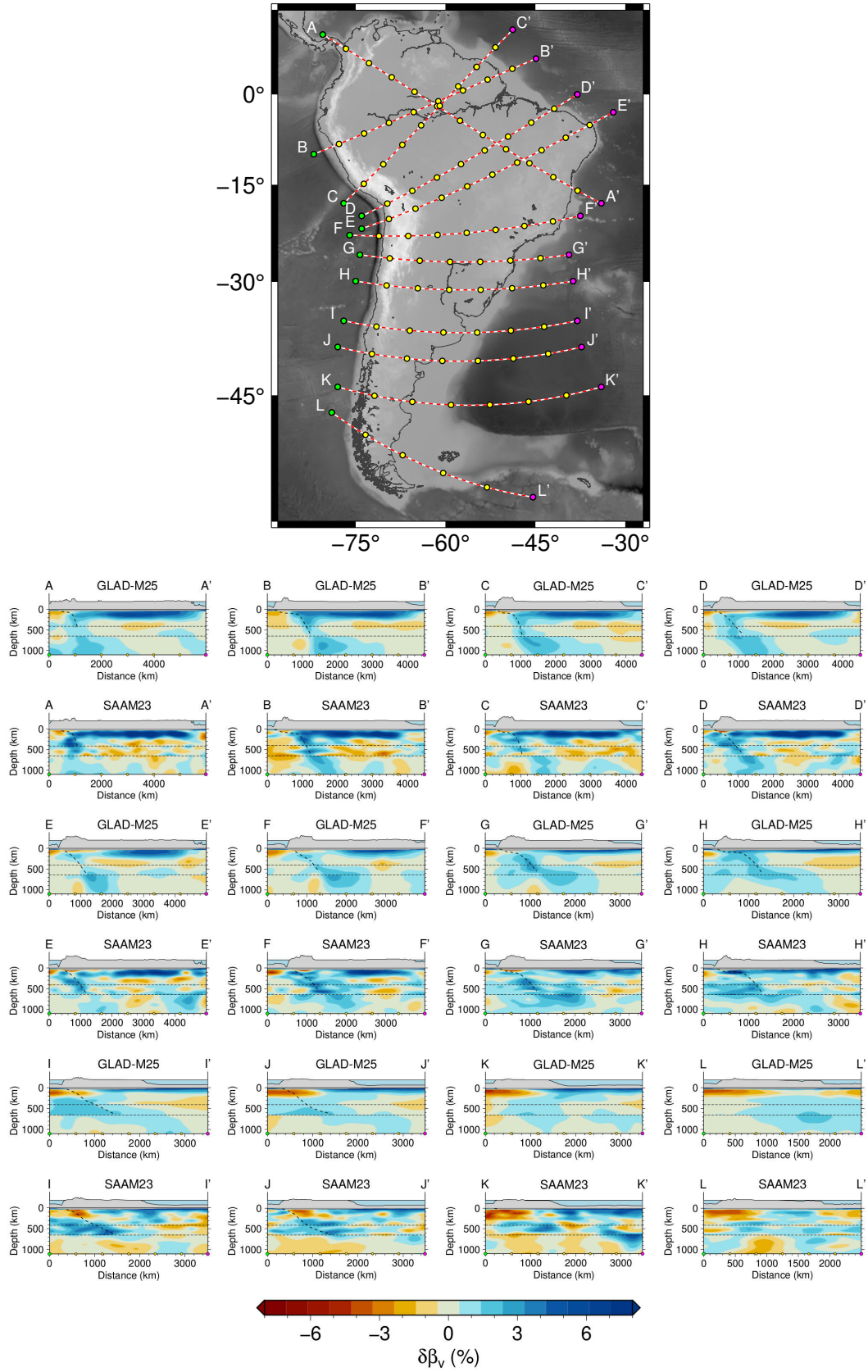


Figure 13. Vertical cross-sections of GLAD-M25 compared with SAAM23. The black dashed line is the Slab2 model. The two horizontal dashed lines indicate the 410 and 650 km discontinuities.

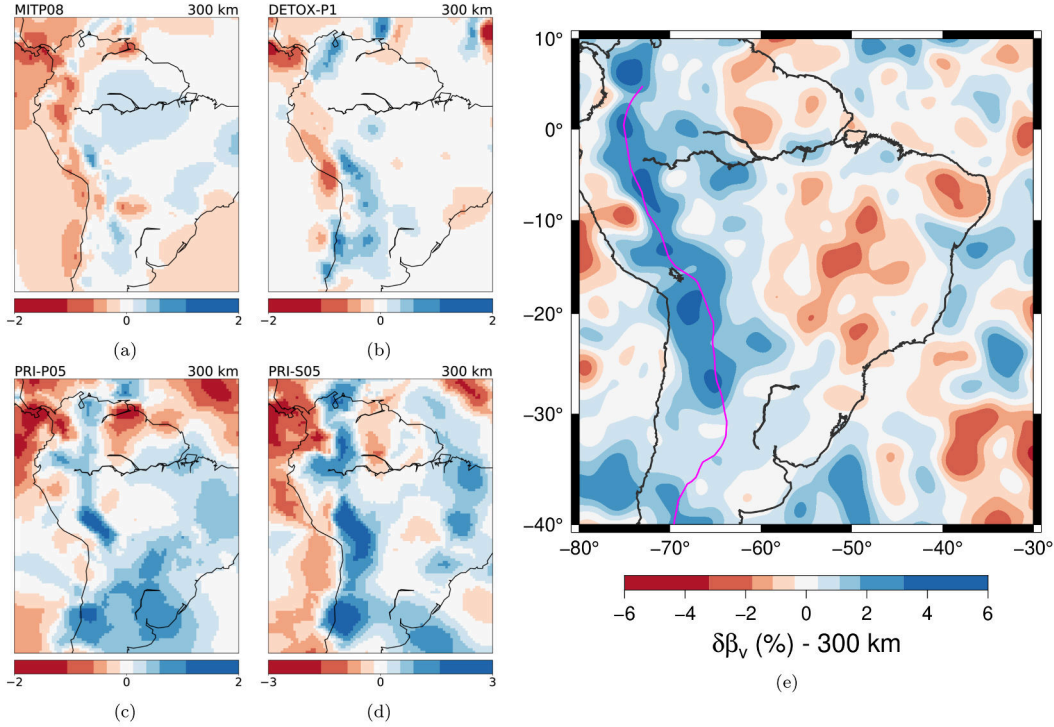


Figure 14. The Nazca Slab East of Peru (from about 6°S to 3°N) is not imaged in some models like MITP08 (Li et al., 2008) (Fig. 14a)) and DETOX-P1 (Hosseini et al., 2020) (Fig. 14b), but was seen in the models PRI-P05 (Fig. 14c), and S05 (Fig. 14d) (Montelli et al., 2006). Our model SAAM23 (Fig. 14e) shows clear high velocity anomalies confirming a continuity of the Nazca Slab in that region. The maps of the four previous models were plotted with the Oxford SubMachine tool (Hosseini et al., 2018).

SAAM23 confirms the previous findings of GLAD-M25 (Lei et al., 2020) showing a continuous slab in that region and no evidence for a slab gap or absence (Fig. 14). However, the slab high-velocity anomaly following the Peruvian flat segment has lower amplitudes, at 400 km depth, compared with the Bolivian and Argentinian sections to the south and with the North Andean segment to the North. Model Detox-P3 (including ISC PP arrival time picks) also shows low amplitude high velocities in that region. This part of the slab with reduced anomaly amplitudes may be related to a plateau or other anomalous part of the Nazca Plate that subducted during the Neogene. This diminished anomaly amplitude, together with a generally poor station coverage in that region, might explain the absence of the slab image in some tomographic models.

5.2.2 The Nazca Slab in the Lower Mantle

North of about 20°S, several tomographic models image the Nazca Slab crossing the transition zone and continuing into the lower mantle, such as the global models LLNL-3Dv3 (Simmons et al., 2012), GAP-4 (Obayashi et al., 2013), SPani-P,S (Tesoniero et al., 2015), DETOX-P1 (Hosseini et al., 2020), as well as the regional models SAM-P-2019 and SAM-S-2020 (Portner et al., 2020; Rodríguez et al., 2021), as shown in Fig. 15. Our model, SAAM23, confirms this slab behavior, including its apparent flattening below a depth of 1,000 km (Fig. 15 - profile D).

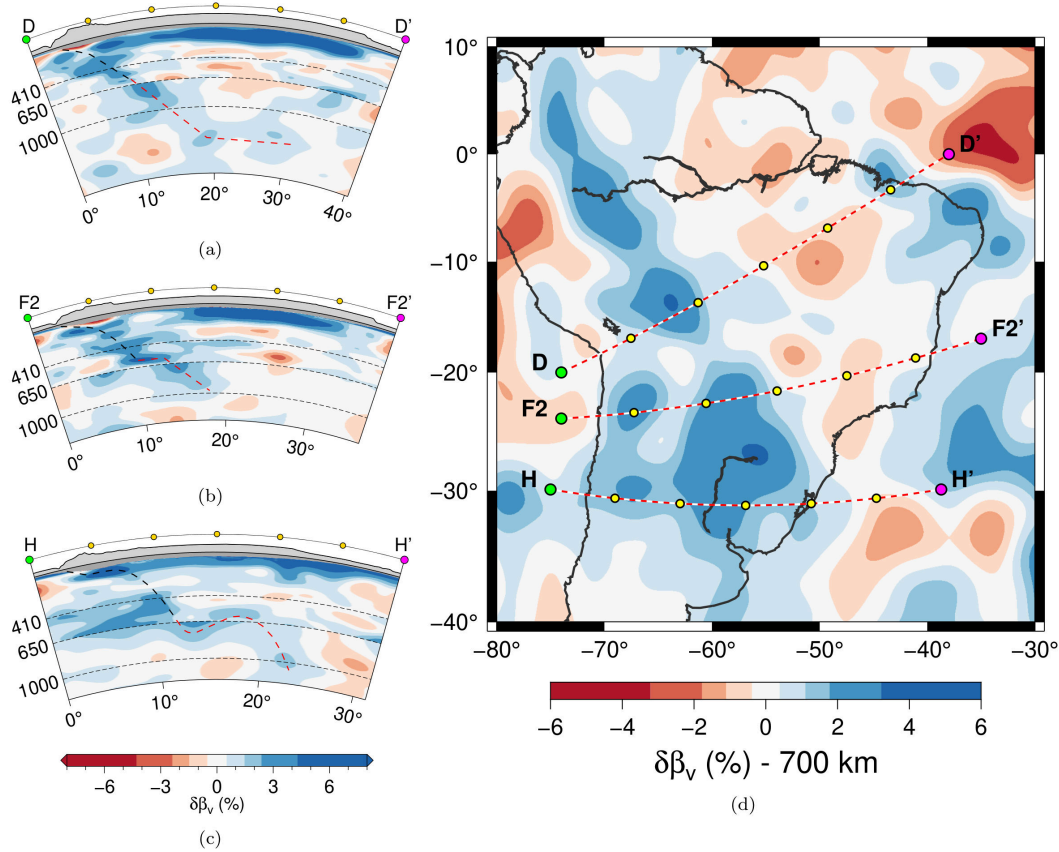


Figure 15. Changes in the Nazca Slab geometry from north to south. North of $\sim 20^\circ\text{S}$ the slab plunges across the transition zone directly into the lower mantle, as seen in profile D. South of $\sim 20^\circ\text{S}$ the slab flattens near the bottom of the transition zone for a while before plunging into the lower mantle (profiles F2 and H). The black dashed line is the Slab2 model after the Benioff zone; the red dashed line is a suggested continuation of the slab further down.

However, near about 20°S , some tomographic models showed the Nazca Slab to flatten at the bottom of the transition zone, just above the 660 km discontinuity, such as LLNL-G3Dv3 (Simmons et al., 2012) and GAP-P4, whereas models SAM-P2019 (Portner et al., 2020) and DETOX-P1 (Hosseini et al., 2020) show a straighter continuation of the slab into the lower mantle. The waveform modeling inversion of Celli et al. (2020) appears to be in agreement with the former models, as their results also suggest a slab crossing the transition zone beneath the Amazon and Central Brazil and a trend of stagnation near 660 km depth further south. SAAM23 (Fig. 15, profile F2) shows intermediate behavior: the slab seems to flatten at the bottom of the transition zone slightly, then resumes a straighter trajectory into the lower mantle.

Further South, around 30°S and in the transition zone, the Nazca Slab has been imaged mainly as a relatively flat high-velocity anomaly along and near the 650 km discontinuity, suggesting either a diminished slab pull (stagnation) just above or just below the transition zone or reflects a change in convergence rate or westwards trench migration (Rodríguez et al., 2021). Our model (Fig. 15, profile H) shows this deep, flat slab segment to be around 1,000 km long, and it seems to be descending into the lower mantle in two separate parts. A very similar image was obtained by Rodríguez et al. (2021) (their Fig. 8b, our Fig. 16a). Rodríguez et al. (2021) interpret this behavior as the result of spatio-temporal

changes in the rate of westwards trench migration rather than as a change in slab pull with depth. This behavior favors the hypothesis that the absolute motion of slabs is mainly vertical downwards, and the apparent dip of the slab at shallow depths is the result of the westwards trench migration, as the absolute South American motion is towards the NW.

Moving even further southwards, around 46°S, a region of very low shear-wave velocity that spans most of the Patagonian Domain in SAAM23 (Fig. 11) and previously observed in other tomographic studies, such as Feng et al. (2007) and Celli et al. (2020), correlates with the "slab window" proposed by Russo, Gallego, et al. (2010) and Russo, VanDecar, et al. (2010).

5.2.3 The "Paraná" Plume

A vertical "cylindrical" low-velocity anomaly beneath the Paraná Basin, first detected by VanDecar et al. (1995), was confirmed and extended down to ~800 km by Schimmel et al. (2003); M. P. Rocha et al. (2011). The initial interpretation of the anomaly as a thermal one was challenged by K. H. Liu et al. (2003) and Bianchi et al. (2020) who did not observe any effect on the depths of the 410 and 650-km discontinuities, based on receiver functions, that could be attributed to a thermal anomaly. Portner et al. (2020); Rodríguez et al. (2021) showed a low-velocity anomaly in the 150 to 1,200 km depth range beneath the Paraná Basin, which appears to connect to the Nazca Slab in the top of the lower mantle and transition zone (Fig. 16a). SAAM23 also shows low velocities from 250 km down to the bottom of the transition zone, consistent with previous models (Fig. 16b). The origin of this low-velocity anomaly is not known, but Rodríguez et al. (2021) suggested that the it may represent a current mantle upwelling induced by the continued motion of the leading edge of the Nazca Slab or by the accumulation of hydrous minerals from a second phase of deep dehydration in the slab, as hypothesized by van der Lee et al. (2008).

5.3 The Lithosphere of the Stable Continental Interior

High velocities in the subcontinental lithosphere (SCL) are imaged beneath all cratons in Brazil down to about 200 km (Fig. 12), especially in the exposed Amazonian and São Francisco cratons (Fig. 17). Higher velocities are observed on the eastern half of the AC, consistent with older radiometric ages (e.g., C. C. Tassinari & Macambira, 1999; J. O. S. Santos et al., 2000; M. Vasquez et al., 2008). In the SFC, high velocities are also observed to the west of its surface limit, similar to the surface wave tomography of Feng et al. (2007) and the teleseismic *P*-wave tomography of Schimmel et al. (2003) and M. P. Rocha et al. (2011). This westward extension of the SFC, at depth, is consistent with the assumed larger Neoproterozoic São Francisco paleocontinent (M. P. Rocha et al., 2019).

Smaller cratonic fragments such as São Luís near the equatorial coast, and Luís Alves near the SE coast, are too small to be resolved by our tomographic model. The anomalies related to the São Luís Craton may be overlapping with those of the PnB. The Luís Alves craton is located in a region that suffered large extensional deformation during the Atlantic rifting, related to a plume impact, and may no longer have a deep root. However, Fig. 17 suggests a possible continuation of the Luís Alves Craton westwards beneath the Paraná Basin, as also inferred from teleseismic *P*-wave tomography (Affonso et al., 2021).

Two large intracratonic basins (Parnaíba in the north and Paraná in the SE) are thought to be overly old cratonic blocks, inferred from radiometric dates of a few samples drilled from the basement (Cordani, 1984; Cordani et al., 2009), as well as from models of geological evolution (Almeida et al., 2000; Brito Neves & Fuck, 2014). These former cratonic nuclei, named Parnaíba and Paranapanema blocks, have been tentatively delineated by geophysical mapping (e.g., M. Mantovani et al., 2005; Affonso et al., 2021). SAAM23 (Fig. 17) clearly shows high SCL velocities in the two blocks.

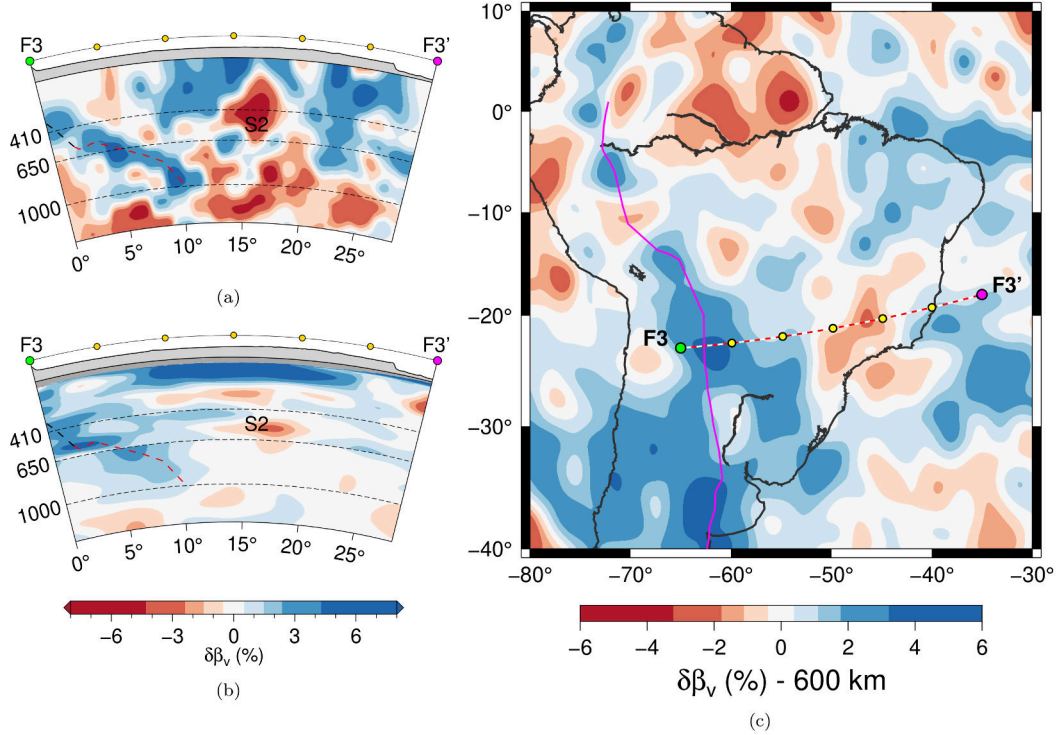


Figure 16. The low-velocity "Paraná Plume" imaged by a) the S -wave teleseismic tomography of Rodríguez et al. (2021), and b) the SAAM23 model. c) maps of β_v anomalies at 600 km depth with the two profiles. The two dashed lines in the vertical sections are the 410 and 650 km discontinuities.

SAAM23 clearly shows the Paranaíba Block as an independent cratonic nucleus, separated from the Amazonian and the São Francisco cratons, even though its shape may not be as inferred from other geophysical studies. This result is in agreement with Celli et al. (2020), whose SA2019 model also shows the PnB as an independent body.

The Paranapanema, on the other hand, is clearly separated from the surrounding fold belts and quite consistent with the gravity-inferred limits of M. Mantovani et al. (2005). These results indicate that any rifting episode in the Paraná Basin during the evolution of the basin was not strong enough to affect the thick lithosphere on a regional scale.

The RPC is mainly covered beneath the Paraná and Chaco basins and different boundaries have been proposed (e.g., Oyhantçabal et al., 2011; Rapela et al., 2011). However, none of the studied tomographic models show high SCL velocities for this cratonic area, despite the use of different tomographic techniques: Feng et al. (2007); Rosa et al. (2016); M. P. Rocha et al. (2019); Affonso et al. (2021). SAAM23 also shows no high velocities in the RPC (Figs. 17a,b).

5.3.1 The Lithosphere/Asthenosphere Boundary

High SCL velocities are commonly used as a proxy for lithospheric depth, especially in tomographic models based on teleseismic body waves as vertical smearing usually prevents a robust determination of lithospheric depth. Our full-waveform fitting method incorporates both body and surface wave measurements, weakening the trade-off between lithospheric velocity anomaly amplitude and lithospheric thickness. Therefore, we here infer the bottom of the SCL, that is, the lithosphere-asthenosphere boundary (LAB), by measuring the depth

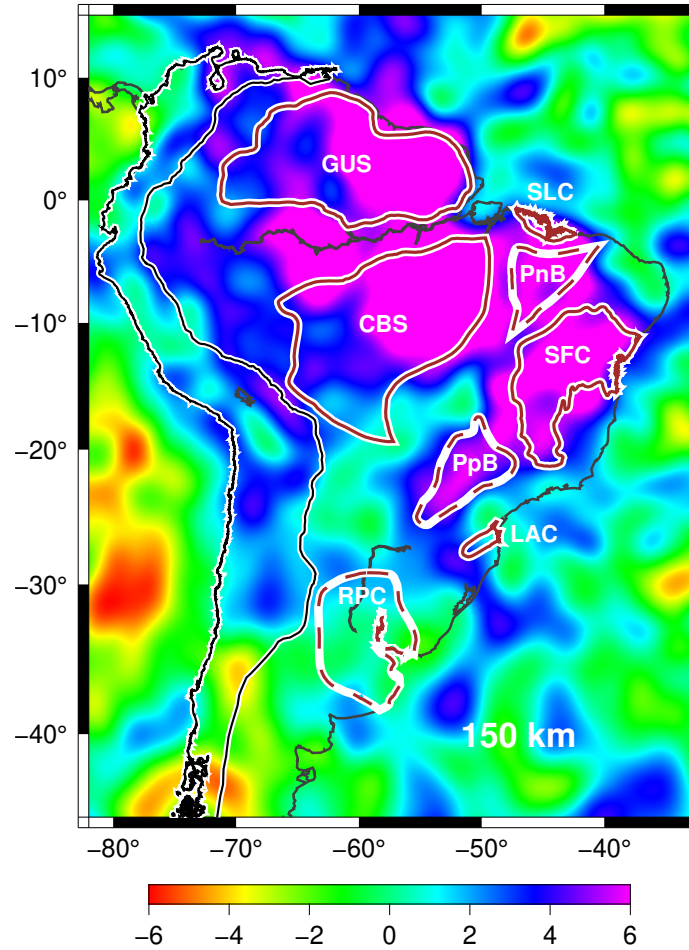


Figure 17. Lithospheric β_v anomaly at 150 km depth and the cratonic units in South America. Exposed cratons: GUS= Guyana Shield, CBS= Central Brazil Shield (part of the AC); SFC = São Francisco; SLC = São Luís; LAC = Luis Alves cratons. Mainly covered cratonic blocks: PnB= Parnaíba Block; PpB = Parapanema Block; RPC = Río de la Plata Craton. The black solid line delimits the Andean Belt according to CPRM.

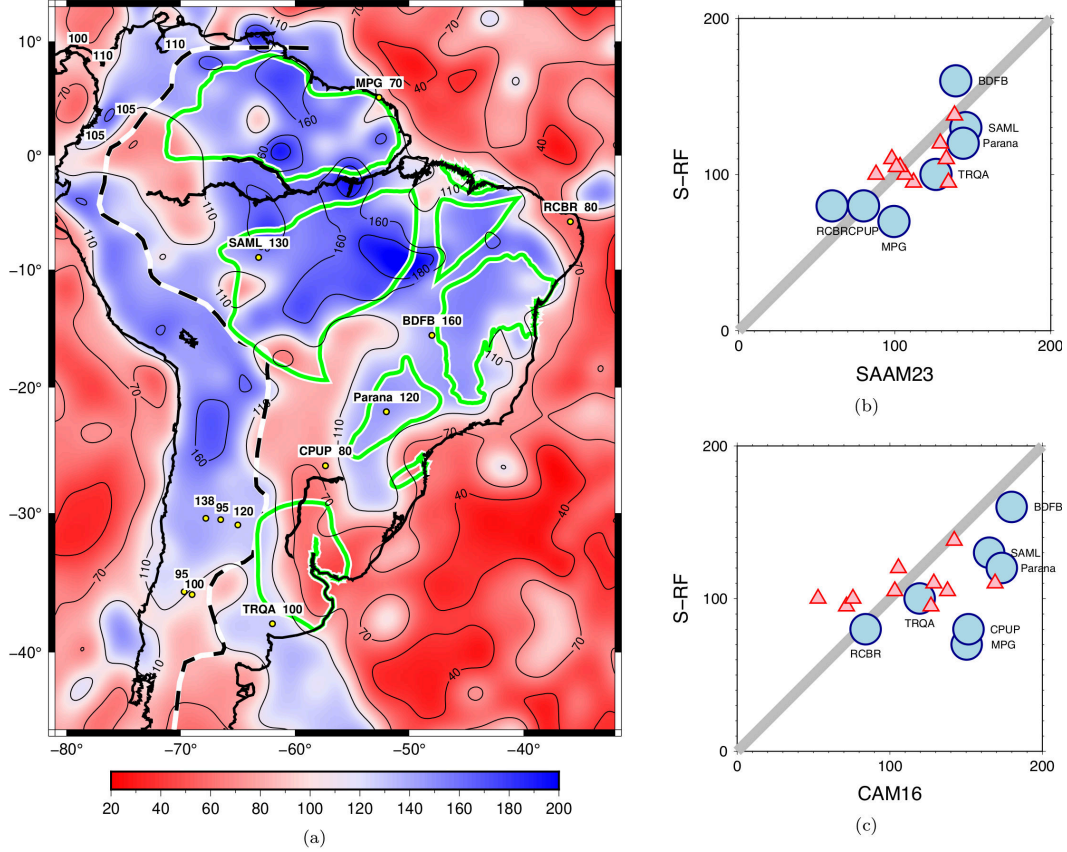


Figure 18. The seismic Lithosphere-Asthenosphere Boundary (LAB) based on the largest negative gradient of the β_v vertical profiles. a) LAB contours; labelled bold numbers are LAB depths from S -wave receiver functions (Heit, Sodoudi, et al., 2007); smaller bold numbers are LAB depths from Andean stations (Heit, Sodoudi, et al., 2007; Blanco et al., 2017). Red lines are the cratons and cratonic blocks. The black dashed line is the western limit of the stable continental region (SCR), as defined by Johnston et al. (1994). b) Comparison of the LAB depth from SAAM23 with those of S-RF: blue circles with labels are SCR stations (Heit, Sodoudi, et al., 2007) and triangles are Andean stations (Heit et al., 2008; Blanco et al., 2017). c) Comparison of the S-RF LAB with the mechanical LAB depths from the CAM16 model (Priestley et al., 2018), based on geochemical and geotherm criteria, which gives about 40 km deeper LAB than our seismic estimates, especially for SCR regions.

of the largest negative velocity gradient with depth in each vertical profile, below the Moho (Van der Lee, 2002). We took the depth of the most negative gradient at every cell ($1^\circ \times 1^\circ$), and applied a Gaussian smoothing filter with 100 km half-width to remove short-wavelength oscillations in the tomography model. Fig. 18a shows the LAB map, which shows good agreement with the values obtained from S -wave receiver function analysis (Fig. 18b). We used the LAB depths from the stable continental region (Heit, Sodoudi, et al., 2007), from sub-Andean stations (Heit et al., 2008), and from regional station averages in Colombia (Blanco et al., 2017). The mean difference between our tomographic LAB and those from S -wave receiver functions was only about 7 km.

Using the global surface-wave tomography model CAM16, Priestley et al. (2018) estimated the LAB depth using petrological models of the upper mantle and β_v -derived tem-

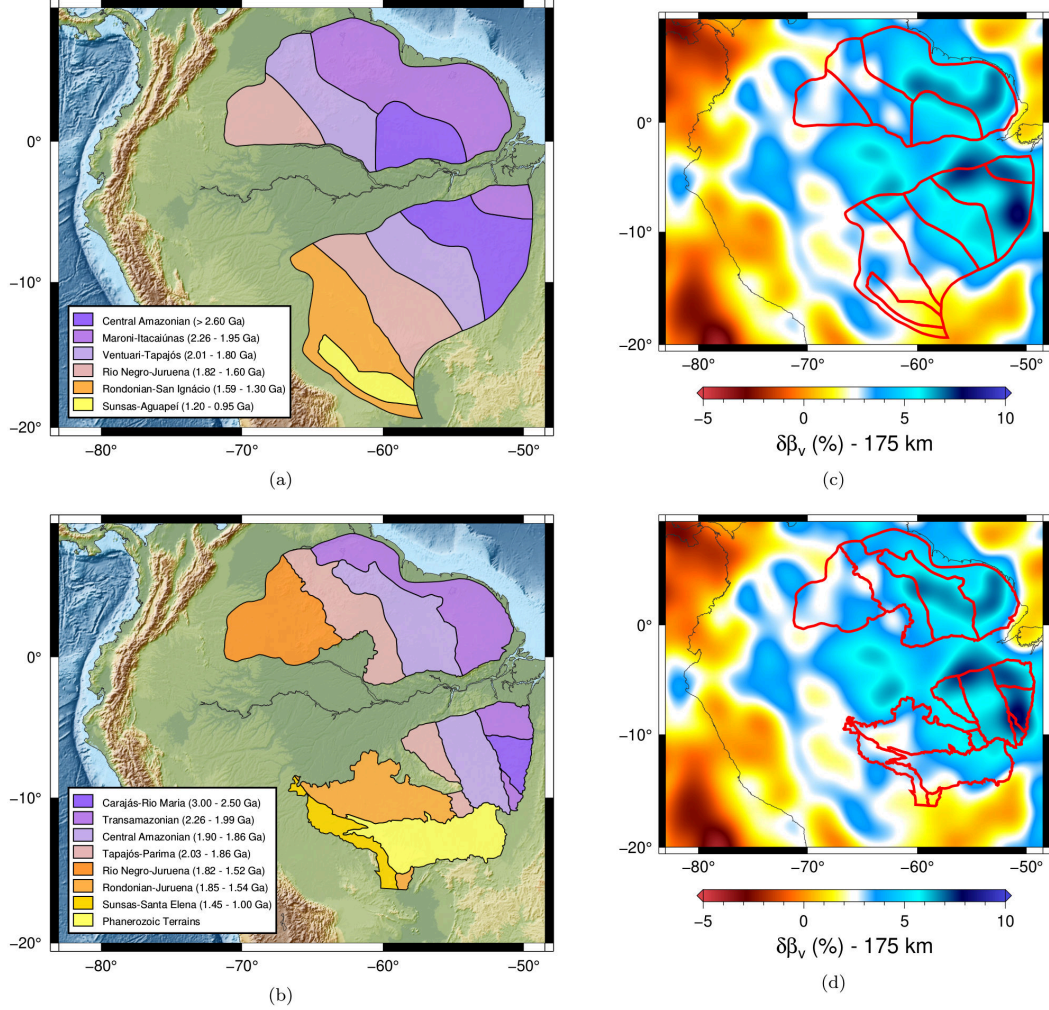


Figure 19. a) AC geochronological provinces of Model 1 (Teixeira et al., 2019); b) β_v anomaly at 175 km depth with the contours of model 1; c) Amazonian geochronological Model 2 (J. d. Santos, 2003; M. Vasquez et al., 2008); d) β_v anomaly at 175 km depth with the contours of model 2.

perature profile to identify the base of the lithosphere as the transition from the conducting to advecting geotherm. This LAB definition gives values about 40 km deeper in the stable continental region than those of the *S*-wave receiver functions (Fig. 18c). In the Andes, there seems to be a better agreement of the LAB mechanical definition with those of the *S*-RF, on average, but the scatter is larger.

In the stable continental region (East of the dashed line in Fig. 18a) the thickest lithosphere is found in the AC. A N-S belt of the thin lithosphere (near 70 km) is seen from the RPC, through the Chaco Basin, and reaching the Pantanal Basin. The northeastern corner of Brazil (Borborema fold province) also has a very thin lithosphere, consistent with the generally SCL low velocities observed by teleseismic *P*-wave tomography north of 7.5°S (Simões Neto et al., 2019) interpreted as due to lateral mantle flow and intraplate volcanism. We observe a thinner lithosphere in the Patagonian Platform, which can be explained by its younger thermal history. The regions of thin lithosphere (less than 100 km) in Fig. 18 are also consistent with the higher temperatures (larger than 1,200 °C) found by the modeling of Finger et al. (2021).

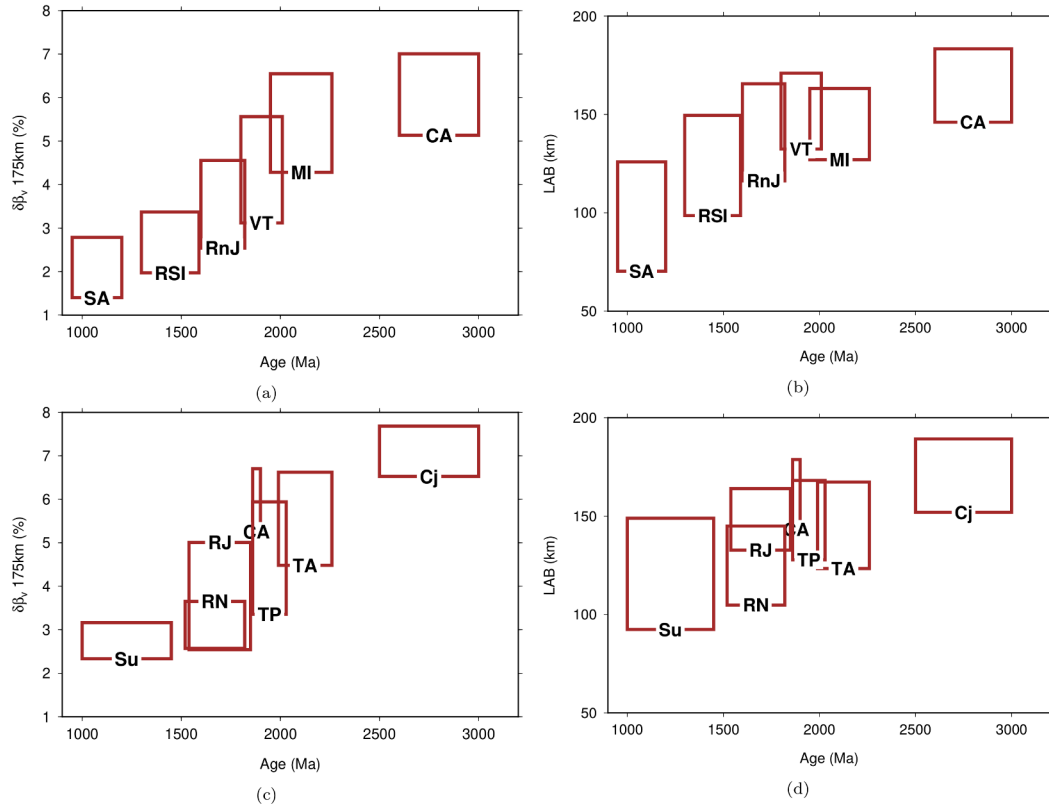


Figure 20. Correlation between age range of the Amazonian geochronological provinces (horizontal axis) and the average β_v anomaly at 175 km depth (vertical axis; plots a,c) and the LAB depth (b,d). The height of each box corresponds to the average anomaly ± 1 standard deviation. a,b Province Model-1 Teixeira et al. (2019); c,d Province Model-2 (J. d. Santos, 2003; M. Vasquez et al., 2008).

5.3.2 Lithospheric Thickness and Geochronological Provinces in the Amazonian Craton

The evolution of the AC has been explained by an initial Archean proto-craton and a series of accretionary terranes added by successive collision zones and orogenies, each terrain named as a "geochronological province". Different models for the limits of these provinces have been proposed. Based on increasing availability of radiometric dates, the model of C. C. Tassinari and Macambira (1999) and C. C. G. Tassinari and Macambira (2004) has evolved through Cordani et al. (2016), Teixeira et al. (2019), Macambira et al. (2020) and Frederico Faleiros, *personal communication*, 2021, and is presented in Fig. 19a. We call this proposal "Model-1". Another set of models for the limits of geochronological provinces, which we call "Model-2", has been proposed by J. O. S. Santos et al. (2000) and J. d. Santos (2003) and updated by M. Vasquez et al. (2008), as shown in Fig. 19b. Costa et al. (2020) used this Model-2 to interpret the results of their teleseismic P -wave tomography.

Initially, we compared the average S -wave anomaly of each province, at 175 km depth (as a proxy for lithospheric thickness), with its age range, to test if older provinces tend to have thicker (colder) lithosphere. Despite the non-uniform S -wave anomaly within each province, a remarkable correlation of positive seismic anomalies with increasing age was found for both geochronological models (Fig. 20a,c). This trend of increasing velocity anomalies with age is also seen in SA2019 (Celli et al., 2020).

We also used the direct LAB depth (Fig. 18a), instead of β_v anomaly, and a similar conclusion was found: older provinces tend to have thicker lithosphere in the AC, on average (Figs 20b,d). Our seismic tomography model cannot indicate the preferred geochronological model, as both show similarly good correlation between age range and lithospheric average depth.

Lithospheric blocks cool very little after the first ~ 500 Ma years. Between 1 Ga and 2.8 Ga, the expected decrease in temperature is only ~ 50 °C (Porter et al., 2019). For 175 km depth, a difference of 50 °C should correspond to 0.4-1.1% decrease in S -wave velocity (Goes et al., 2000). Therefore, the temperature difference due to cooling alone cannot explain the difference of $\delta\beta_v$ of 4% (Fig. 20) between the oldest and youngest blocks in the AC. It is well known that the old cratonic lithosphere contributes to continental buoyancy due to its relatively low-density (Jordan, 1978; Forte & Perry, 2000), a result from iron depletion (Goes & van der Lee, 2002; Griffin et al., 1999). One possible contribution to the large S -wave variation observed in Fig. 20, could be that older cratons are more depleted in iron than younger ones consistent with Griffin et al. (1999). Iron depletion simultaneously increases β and decreases ρ , thereby potentially contributing to the observed S -wave anomaly, but likely not sufficient to explain a 4% velocity anomaly (Goes et al., 2000). An alternative explanation of this trend in cratonic S -velocity anomaly and lithospheric thickness (Fig. 20) is that the SCL of the AC was relatively recently modified by upper mantle geodynamics, such as sub-lithospheric convection or interaction with a mobile and deformable asthenosphere that is also facilitating South America's rapid westward plate motion. For example, sub-lithospheric mantle flow could heat the lithosphere of the geochronological provinces of the AC from the West, which would explain the trend in Fig. 20 as heating from West to east rather than cooling from young to old. Finger et al. (2021) advanced a similar hypothesis after combining new seismic data with gravity data and mineral physics constraints to develop self-consistent models of temperature, composition, and density of the South American lithospheric and sub-lithospheric upper mantle. They found temperature differences from W to E ranging from 200 to 300 °C, which is consistent with the LAB depth variations observed in Figs. 18 and 20. At the northwestern AC, temperatures ranging between 1,200 and above 1,400 °C for depths of 150 and 200 km, respectively, indicate that the cratonic root was eroded, possibly by upwelling of hot mantle material under the Guyana highlands (Finger et al., 2021), consistent with our observed trend of lithospheric thinning towards the West.

Besides the E-W variations, the AC may also show differences between the two shields on both sides of the Amazon Basin. Our SAAM23 model shows reduced high-velocity anomalies beneath the Amazon Basin (Fig. 11) as in Heintz et al. (2005), suggesting that the rifting episode responsible for the formation of the Amazon Basin has involved a significant portion of the lithosphere.

6 Conclusions

In this study, we constructed an adjoint waveform model, SAAM23, based on 3D spectral-element simulations of wave propagation for the South American continent using data from 112 earthquakes recorded by 1311 seismic stations. The model results from 23 conjugate gradient iterations where we assimilated three-component data, including both body and surface waves in the inversion. During the construction of the model, we used the exponentiated phase misfit proposed by Yuan et al. (2020), a variant of the instantaneous phase misfit (Bozdağ et al., 2011), to better take the scattered waves into account while effectively mitigating the cycle skip problem of phase measurements.

At long wavelengths, SAAM23 is compatible with previous seismic-tomographic studies. Well resolved is the lithosphere of cratonic nuclei in the stable platform (Amazonian and São Francisco) as well as that of the covered Paranapanema and Parnaíba cratonic blocks. The Nazca Slab is imaged into the lower mantle and is shown to be continuous in the 300-500 km depth range down-dip from the Peruvian flat-slab segment. The slab crosses the mantle transition zone and dips into the lower mantle beneath the northern portion of South America. In the central and southern part of South America, the slab somewhat flattens near the 650 km discontinuity before straightening and continuing into the lower mantle.

Lithospheric thicknesses estimated from published *S*-wave receiver functions agree well with our Lithosphere-Asthenosphere boundary (LAB) estimated using the most significant negative velocity gradient with depth in SAAM23. Our LAB depths agree remarkably well with the temperature model of Finger et al. (2021). We additionally found a strong correlation between the average age of the geochronological provinces in the Amazonian Craton and the positive lithospheric β_v anomalies of our model. Despite different formation ages, the fact that these formation ages are all Precambrian rules out that these differences in lithospheric structure are the result of simple cooling histories. Instead, we suggest that different degrees of iron depletion account for some of the inter-provincial differences in β_v and that westwards thinning of the lithosphere might be related to vigorous asthenospheric flow and its interaction with spatio-temporally changing mantle wedge behavior on the western margin of the continent. Like previous tomographic studies, we found no high-velocity anomalies beneath the Río de la Plata Craton. Our continental-scale, high-resolution, deeply-extending, full-waveform based, new tomographic model, SAAM23, provides a comprehensive and robust basis for further studies of the evolution of the South American Plate.

7 Data Availability Statement

We downloaded the waveforms using ObsPy (Beyreuther et al., 2010) including stations from the networks: 2B (Heit, Yuan, et al., 2007), 8G (Meltzer & Beck, 2016), 9A (GEOFON Program (GFZ-Potsdam, 2007), AI (di Oceanografia e di Geofisica Sperimentale, 1992), AY (of Mines & (Haiti), 2010), BL (Institute of Astronomy & USP, 1988), BR (Universidade de Brasília, 1995), C (de Chile, 1991), C1 (de Chile, 2013), CM (Colombiano, 1993), CU ((ASL)/USGS, 2006), CW (for Seismological Research (CENAI Cuba), 1998), CX (for Geosciences & des Sciences de L'Univers-Centre National de la Recherche CNRS-INSU, 2006), CY (Government, 2006), DR (Domingo, 1998), EC (Ecuador), 2002), G (de Physique du Globe de Paris (IPGP) & et Observatoire des Sciences de la Terre de Strasbourg (EOST), 1982), GE (Centre, 1993), GL (de Physique du Globe de Paris, 1950), GT ((ASL)/USGS, 1993), II (Oceanography, 1986), IQ (Cesca et al., 2009), IU ((ASL)/USGS, 1988), JM (of the West Indies Mona (Jamaica), 1985), MQ (de Physique du Globe de Paris, 1935),

NA (KNMI, 2006), NB (do Rio Grande do Norte (UFRN Brazil), 2006), NU (de Estudios Territoriales (INETER), 1975), ON (Observatório Nacional, 2011), OV (Protti, 1984), PA (Baru, 2000), PR (Rico, 1986), TC (de Costa Rica, 2017), TO (MASE, 2007), VE (Fundación Venezolana De Investigaciones Sismológicas (FUNVISIS), 2000), WC (Curacao, 2006), WI (de Physique du Globe de Paris (IPGP), 2008), X1 (of Liverpool, 2007), X6 (Eric Sandvol, 2007), XB (Wiens, 1997), XC (Institute of Astronomy & USP, 2016), XE (Paul Silver, 1994), XH (Zandt, 1996), XJ (of Cambridge (UK) Earth Sciences, 2004), XN (Levander, 2008), XP (Michael West, 2010), XS (Vilotte & RESIF, 2011), XT (Vernon et al., 2003), XY (Schwartz, 1999), Y4 (Institute of Astronomy & USP, 2013), YC (T. W. Susan Beck, 2000), YH (of Oregon, 1999), YJ (Russo, 2004), YO (Geoffrey A. Abers, 2003), YS (Pritchard, 2009), YW (of Bristol (UK), 2002), YZ (Susan Y. Schwartz, 2009), ZA (GEOFON Program (GFZ-Potsdam, 1994), ZB (GEOFON Program (GFZ-Potsdam, 1997), ZC (Pulliam, 2013), ZD (Lara Wagner, 2010), ZE (Haberland et al., 1996), ZG (G. Z. Susan Beck, 2010), ZL (G. Z. Susan Beck, 2007), ZP (GEOFON Program (GFZ-Potsdam, 1999), and ZQ (GEOFON Program (GFZ-Potsdam, 2004).

The window selection and EP measurements were carried out using PyWinEPAdjoint (Ciardelli, 2021), available at <https://doi.org/10.5281/zenodo.4920792>. Maps were created using the Oxford SubMachine tool (Hosseini et al., 2018) and GMT6 (Wessel et al., 2019).

Acknowledgements

We thank FAPESP (Grants: 2013/24215-6, 2016/03120-5, 2018/04918-6, and 2018/04917-0) and CNPq (Grant: 30.1284/2017-2) for providing financial support for this research. We also thank María Laura Rosa, from the National University of La Plata, for providing the waveforms and receiver response files for station LPA. This research was supported in part through the computational resources and staff contributions provided for the Quest high performance computing facility at Northwestern University which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

References

- Affonso, G., Rocha, M., Costa, I., Assumpção, M., Fuck, R., Albuquerque, D., ... Beck, S. (2021). Lithospheric Architecture of the Paranapanema Block and Adjacent Nuclei Using Multiple-Frequency P-Wave Seismic Tomography. *Journal of Geophysical Research: Solid Earth*, 126(4), e2020JB021183. Retrieved from <https://doi.org/10.1029/2020JB021183>
- Alkmim, F. d. (2004). O que faz de um cráton um cráton? O Cráton do São Francisco e as revelações almeidianas ao delimitá-lo. *Geologia do continente Sul-Americano: evolução da obra de Fernando Flávio Marques de Almeida*, 17–34. Retrieved from <https://geologia.ufc.br/>
- Almeida, F. F. M. d., Brito Neves, B. B. d., & Carneiro, C. D. R. (2000). The origin and evolution of the South American Platform. *Earth-Science Reviews*, 50(1-2), 77–111. Retrieved from [https://doi.org/10.1016/S0012-8252\(99\)00072-0](https://doi.org/10.1016/S0012-8252(99)00072-0)
- Amante, C., & Eakins, B. W. (2009). *ETOPO1 arc-minute global relief model: procedures, data sources and analysis*. Retrieved from <https://doi.org/10.7289/V5C8276M>
- Amaru, M. (2007). *Global travel time tomography with 3-D reference models* (Vol. 274). Utrecht University.
- (ASL)/USGS, A. S. L. (1988). *Global Seismograph Network (GSN - IRIS/USGS)*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/iu>
- (ASL)/USGS, A. S. L. (1993). *Global Telemetered Seismograph Network (USAF/USGS)*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/iu>

- doi.org/10.7914/sn/gt
 (ASL)/USGS, A. S. L. (2006). *Caribbean USGS Network*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/cu>
- Babuska, V., & Cara, M. (1991). *Seismic anisotropy in the Earth* (Vol. 10). Springer Science & Business Media. Retrieved from <https://www.springer.com/gp/book/9780792313212>
- Baru, R. S. V. (2000). *ChiriNet*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/pa>
- Bassin, C. (2000). The Current Limits of resolution for surface wave tomography in North America. *Eos, Transactions American Geophysical Union*, *81*, F897. Retrieved from <https://igppweb.ucsd.edu/~%7Egabi/crust1/laske-egu13-crust1.pdf>
- Beller, S., Monteiller, V., Operto, S., Nolet, G., Paul, A., & Zhao, L. (2018). Lithospheric architecture of the South-Western Alps revealed by multiparameter teleseismic full-waveform inversion. *Geophysical Journal International*, *212*(2), 1369–1388. Retrieved from <https://doi.org/10.1093/gji/ggx216>
- Betiollo, L. M., Iglesias, C. M. d. F., & Rocha, P. G. (2018). *Área de Relevante Interesse Mineral, Integração geológica-geofísica e recursos minerais do Cráton Luis Alves*. CPRM. Retrieved from <http://rigeo.cprm.gov.br/handle/doc/16802>
- Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, *81*(3), 530–533. Retrieved from <https://doi.org/10.1785/gssrl.81.3.530>
- Bianchi, M. B., Assumpcao, M., Kock, C., & Beck, S. L. (2020). *Effect of the Cold Nazca Slab on the Depth of the 660 km Discontinuity in South America*.
- Blanco, J. F., Vargas, C. A., & Monsalve, G. (2017). Lithospheric thickness estimation beneath Northwestern South America from an S-wave receiver function analysis. *Geochemistry, Geophysics, Geosystems*, *18*(4), 1376–1387. Retrieved from <https://doi.org/10.1002/2016GC006785>
- Blom, N., Gokhberg, A., & Fichtner, A. (2020). Seismic waveform tomography of the central and eastern Mediterranean upper mantle. *Solid Earth*, *11*(2), 669–690. Retrieved from <https://doi.org/10.5194/se-11-669-2020>
- Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., ... Pugmire, D. (2016). Global adjoint tomography: first-generation model. *Geophysical Journal International*, *207*(3), 1739–1766. Retrieved from <https://doi.org/10.1093/gji/ggw356>
- Bozdağ, E., Trampert, J., & Tromp, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, *185*(2), 845–870. Retrieved from <https://doi.org/10.1111/j.1365-246X.2011.04970.x>
- Brito Neves, B. B. d., & Fuck, R. A. (2014). The basement of the South American platform: Half Laurentian (N-NW)+ half Gondwanan (E-SE) domains. *Precambrian Research*, *244*, 75–86. Retrieved from <https://doi.org/10.1016/j.precamres.2013.09.020>
- Brossier, R., Operto, S., & Virieux, J. (2010). Which Data Residual Norm for Robust Elastic Frequency-Domain Full Waveform Inversion? *Geophysics*, *75*(3), R37-R46. Retrieved from <https://doi.org/10.1190/1.3379323>
- Bunks, C., Saleck, F. M., Zaleski, S., & Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, *60*(5), 1457–1473. Retrieved from <https://doi.org/10.1190/1.1443880>
- Castro, D. L. d., Fuck, R. A., Phillips, J. D., Vidotti, R. M., Bezerra, F. H., & Dantas, E. L. (2014). Crustal structure beneath the Paleozoic Parnaíba Basin revealed by airborne gravity and magnetic data, Brazil. *Tectonophysics*, *614*, 128–145. Retrieved from <https://doi.org/10.1016/j.tecto.2013.12.009>
- Celli, N., Lebedev, S., Schaeffer, A., Ravenna, M., & Gaina, C. (2020). The upper mantle beneath the South Atlantic Ocean, South America and Africa from waveform tomography with massive data sets. *Geophysical Journal International*, *221*(1), 178–204. Retrieved from <https://doi.org/10.1093/gji/ggz574>

- Centre, G. D. (1993). *GEOFON Seismic Network*. Deutsches GeoForschungsZentrum GFZ. Retrieved from <https://doi.org/10.14470/tr560404>
- Cesca, S., Sobiesiak, M., Tassara, A., Olcay, M., Günther, E., Mikulla, S., & Dahm, T. (2009). *The Iquique Local Network and PicArray*. GFZ Data Services. Retrieved from <https://doi.org/10.14470/vd070092>
- Chen, M., Niu, F., Liu, Q., Tromp, J., & Zheng, X. (2015). Multiparameter adjoint tomography of the crust and upper mantle beneath East Asia: 1. Model construction and comparisons. *Journal of Geophysical Research Solid Earth*, *120*(3), 1762–1786. Retrieved from <http://doi.org/10.1002/2014JB011638>
- Chen, Y., Hill, J., Lei, W., Lefebvre, M., Tromp, J., Bozdog, E., & Komatitsch, D. (2017). Automated time-window selection based on machine learning for full-waveform inversion. In *Seg technical program expanded abstracts 2017* (pp. 1604–1609). Society of Exploration Geophysicists. Retrieved from <https://doi.org/10.1190/segam2017-17734162.1>
- Chen, Y.-W., Wu, J., & Suppe, J. (2019). Southward propagation of Nazca subduction along the Andes. *Nature*, *565*(7740), 441–447. Retrieved from <https://doi.org/10.1038/s41586-018-0860-1>
- Ciardelli, C. (2021, jun). *caiociardelli/pywinepadjoint: v1.0.0-alpha*. Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.4920792>
- Colli, L., Fichtner, A., & Bunge, H.-P. (2013). Full waveform tomography of the upper mantle in the South Atlantic region: Imaging a westward fluxing shallow asthenosphere? *Tectonophysics*, *604*, 26–40. Retrieved from <https://doi.org/10.1016/j.tecto.2013.06.015>
- Colombiano, S. G. (1993). *Red Sismologica Nacional de Colombia*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/cm>
- Cordani, U. G. (1984). Estudo preliminar de integração do Pré-Cambriano com os eventos tectônicos das bacias sedimentares brasileiras. *Bol. Ciênc. Técn. Petról.*, *15*, 1–70. Retrieved from <https://repositorio.usp.br/item/001464085>
- Cordani, U. G., Brito Neves, B., & Thomaz Filho, A. (2009). Estudo preliminar de integração do Pré-Cambriano com os eventos tectônicos das bacias sedimentares brasileiras (Atualização). *Bol. Geociênc. Petrobras, Rio de Janeiro*, *17*(1), 205–219. Retrieved from <https://repositorio.usp.br/item/002153382>
- Cordani, U. G., Milani, E. J., Thomaz Filho, A., & Campos, D. d. A. (2000). *Tectonic Evolution of South America*. Retrieved from <http://rigeo.cprm.gov.br/jspui/handle/doc/19419>
- Cordani, U. G., Sato, K., Sproessner, W., & Fernandes, F. S. (2016). U-Pb zircon ages of rocks from the Amazonas Territory of Colombia and their bearing on the tectonic history of the NW sector of the Amazonian Craton. *Brazilian Journal of Geology*, *46*, 5–35. Retrieved from <http://dx.doi.org/10.1590/2317-4889201620150012>
- Coutinho, M. G. d. N. (2008). *Província Mineral do Tapajós: geologia, metalogenia e mapa previsional para ouro em SIG*. CPRM. Retrieved from <http://rigeo.cprm.gov.br/jspui/handle/doc/1208>
- Cowling, T. G. (1941). The non-radial oscillations of polytropic stars. *Monthly Notices of the Royal Astronomical Society*, *101*, 367. Retrieved from <https://doi.org/10.1093/mnras/101.8.367>
- Curacao, M. D. (2006). *Curacao Seismic Network*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/WC/>
- Dahlen, F., Hung, S.-H., & Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes-I. Theory. *Geophysical Journal International*, *141*(1), 157–174. Retrieved from <https://doi.org/10.1046/j.1365-246X.2000.00070.x>
- Debye, P. (1909). Näherungsformeln für die Zylinderfunktionen für große Werte des Arguments und unbeschränkt veränderliche Werte des Index. *Mathematische Annalen*, *67*(4), 535–558. Retrieved from <https://doi.org/10.1007/%2FBF01450097>
- de Chile, U. (1991). *Chilean National Seismic Network*. International Federation of Digital

- Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/C/>
- de Chile, U. (2013). *Red Sismologica Nacional*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/c1>
- de Costa Rica, U. (2017). *Red Sismológica Nacional de Costa Rica (RSN: UCR-ICE)*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.15517/TC>
- de Estudios Territoriales (INETER), I. N. (1975). *Nicaraguan Seismic Network*. Author. Retrieved from <https://doi.org/10.7914/sn/nu>
- DeMets, C., Gordon, R. G., & Argus, D. F. (2010). Geologically current plate motions. *Geophysical Journal International*, 181(1), 1-80. Retrieved from 10.1111/j.1365-246X.2009.04491.x
- de Physique du Globe de Paris, I. (1935). *Martinique Seismic and Volcano Observatory Network (OVSM)*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.ipgp.fr/fr/ovsm>
- de Physique du Globe de Paris, I. (1950). *Guadeloupe Seismic and Volcano Observatory Network (OVSG)*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.ipgp.fr/ovsg>
- de Physique du Globe de Paris (IPGP), I. (2008). *GNSS, seismic broadband and strong motion permanent networks in West Indies*. Institut de physique du globe de Paris (IPGP), Université de Paris. Retrieved from <https://doi.org/10.18715/antilles.wi>
- de Physique du Globe de Paris (IPGP), I., & et Observatoire des Sciences de la Terre de Strasbourg (EOST), E. (1982). *GEOSCOPE, French Global Network of broad band seismic stations*. Institut de physique du globe de Paris (IPGP), Université de Paris. Retrieved from <https://doi.org/10.18715/geoscope.g>
- di Oceanografia e di Geofisica Sperimentale, I. N. (1992). *Antarctic Seismographic Argentinean Italian Network - OGS*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/ai>
- Domingo, N. S. C. O. A. U. O. S. (1998). *CNS-UASD*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/dr>
- do Rio Grande do Norte (UFRN Brazil), U. F. (2006). *Northeastern Brazil UFRN (NE Brazil)*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.sistemas.ufrn.br/>
- Durek, J. J., & Ekström, G. (1996). A radial model of anelasticity consistent with long-period surface-wave attenuation. *Bulletin of the Seismological Society of America*, 86(1A), 144-158. Retrieved from <http://bssa.geoscienceworld.org/content/ssabull/86/1A/144.full.pdf>
- Dziewonski, A., Chou, T.-A., & Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research: Solid Earth*, 86(B4), 2825-2852. Retrieved from <https://doi.org/10.1029/JB086iB04p02825>
- Ecuador), I. G. E. P. N. I.-E. (2002). *Ecuador Seismic Network*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/EC/>
- Ekström, G., Nettles, M., & Dziewoński, A. (2012). The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200, 1-9. Retrieved from <https://doi.org/10.1016/j.pepi.2012.04.002>
- Engler, A. (2009). The geology of South America. *GEOLOGY-Volume IV*, 373. Retrieved from <http://www.eolss.net/Sample- Chapters/C01/E6-15-07-06.pdf>
- Eric Sandvol, L. B. (2007). *SLIP - Seismic Lithospheric Imaging of the Puna Plateau*. International Federation of Digital Seismograph Networks. Retrieved from https://doi.org/10.7914/sn/x6_2007
- Feng, M., Van der Lee, S., & Assumpção, M. (2007). Upper mantle structure of South

- America from joint inversion of waveforms and fundamental mode group velocities of Rayleigh waves. *Journal of Geophysical Research: Solid Earth*, 112(B4). Retrieved from <https://doi.org/10.1029/2006JB004449>
- Fichtner, A. (2010). *Full seismic waveform modelling and inversion*. Springer Science & Business Media. Retrieved from <https://www.springer.com/gp/book/9783642158063>
- Fichtner, A., Bunge, H.-P., & Igel, H. (2006). The adjoint method in seismology: I. Theory. *Physics of the Earth and Planetary Interiors*, 157(1-2), 86–104. Retrieved from <https://doi.org/10.1016/j.pepi.2006.03.016>
- Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P. (2009). Full Seismic Waveform Tomography for Upper-Mantle Structure in the Australasian Region Using Adjoint Methods. *Geophysical Journal International*, 179(3), 1703–1725. Retrieved from <https://doi.org/10.1111/j.1365-246X.2009.04368.x>
- Fichtner, A., Saygin, E., Taymaz, T., Cupillard, P., Capdeville, Y., & Trampert, J. (2013). The Deep Structure of the North Anatolian Fault Zone. *Earth and Planetary Science Letters*, 373, 109–117. Retrieved from <http://doi.org/10.1016/j.epsl.2013.04.027>
- Fichtner, A., & Trampert, J. (2011). Resolution analysis in full waveform inversion. *Geophysical Journal International*, 187(3), 1604–1624. Retrieved from <https://doi.org/10.1111/j.1365-246X.2011.05218.x>
- Filho, A. T., Mizusaki, A. M., Milani, E. J., & Cesero, P. D. (2000). Rifting and magmatism associated with the south America and Africa breakup. *Brazilian Journal of Geology*, 30, 017–019. Retrieved from DOI:10.5327/RBG.V30I1.750
- Finger, N.-P., Kaban, M., Tesauro, M., Haeger, C., Mooney, W., & Thomas, M. (2021). A Thermo-Compositional Model of the Cratonic Lithosphere of South America. *Geochemistry, Geophysics, Geosystems*, 22(4), e2020GC009307. Retrieved from <https://doi.org/10.1029/2020GC009307>
- Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *The computer journal*, 7(2), 149–154. Retrieved from <https://doi.org/10.1093/comjnl/7.2.149>
- for Geosciences, G. G. R. C., & des Sciences de L’Univers-Centre National de la Recherche CNRS-INSU, I. (2006). *IPOC Seismic Network*. Integrated Plate boundary Observatory Chile - IPOC. Retrieved from <https://doi.org/10.14470/pk615318>
- for Seismological Research (CENAIIS Cuba), N. C. (1998). *Servicio Sismológico Nacional de Cuba*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/cw>
- Forte, A. M., & Perry, H. C. (2000). Geodynamic evidence for a chemically depleted continental tectosphere. *Science*, 290(5498), 1940–1944. Retrieved from <http://doi.org/10.1126/science.290.5498.1940>
- Fukao, Y., Obayashi, M., Inoue, H., & Nenbai, M. (1992). Subducting slabs stagnant in the mantle transition zone. *Journal of Geophysical Research: Solid Earth*, 97(B4), 4809–4822. Retrieved from <https://doi.org/10.1029/91JB02749>
- Fundación Venezolana De Investigaciones Sismológicas (FUNVISIS), C. (2000). *Red Sismológica Satelital Nacional*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/ve>
- Gao, Y., Tilmann, F., van Herwaarden, D.-P., Thrastarson, S., Fichtner, A., Heit, B., ... Schurr, B. D. (2021). Full Waveform Inversion beneath the Central Andes: Insight into the dehydration of the Nazca slab and delamination of the back-arc lithosphere. *Earth and Space Science Open Archive ESSOAr*. Retrieved from <http://doi.org/10.1002/essoar.10506371.1>
- Geoffrey A. Abers, K. M. F. (2003). *Tomography Under Costa Rica and Nicaragua*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/yo2003>
- GEOFON Program (GFZ-Potsdam, G. (1994). *PISCO94 PS (PISCO94)*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/>

- networks/detail/ZA_1994/
- GEOFON Program (GFZ-Potsdam, G. (1997). *PUNA97*. International Federation of Digital Seismograph Networks. Retrieved from http://www.fdsn.org/networks/detail/ZB_1997/
- GEOFON Program (GFZ-Potsdam, G. (1999). *ISSA*. International Federation of Digital Seismograph Networks. Retrieved from http://www.fdsn.org/networks/detail/ZP_1999/
- GEOFON Program (GFZ-Potsdam, G. (2004). *Cerro Blanco*. International Federation of Digital Seismograph Networks. Retrieved from http://www.fdsn.org/networks/detail/ZQ_2004/
- GEOFON Program (GFZ-Potsdam, G. (2007). *COBO*. International Federation of Digital Seismograph Networks. Retrieved from http://www.fdsn.org/networks/detail/9A_2007/
- Goes, S., Govers, R., & Vacher, P. (2000). Shallow mantle temperatures under Europe from P and S wave tomography. *Journal of Geophysical Research: Solid Earth*, 105(B5), 11153–11169. Retrieved from <https://doi.org/10.1029/1999JB900300>
- Goes, S., & van der Lee, S. (2002). Thermal structure of the North American uppermost mantle inferred from seismic tomography. *Journal of Geophysical Research: Solid Earth*, 107(B3), ETG–2. Retrieved from <https://doi.org/10.1029/2000JB000049>
- Government, C. I. (2006). *Cayman Islands*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/CY/>
- Grand, S. P. (1987). Tomographic inversion for shear velocity beneath the North American plate. *Journal of Geophysical Research: Solid Earth*, 92(B13), 14065–14090. Retrieved from <https://doi.org/10.1029/JB092iB13p14065>
- Griffin, W., O'Reilly, S. Y., & Ryan, C. (1999). *The composition and origin of subcontinental lithospheric mantle*. Geochemical Society, Houston, Geochem. Soc. Spec. Publ. Retrieved from <http://hdl.handle.net/102.100.100/215008?index=1>
- Haberland, C., Rietbrock, A., Asch, G., & Chong, G. (1996). *The ANCORP Seismic Network*. GFZ Data Services. Retrieved from <https://doi.org/10.14470/mr6441682066>
- Hasui, Y., Carneiro, C. D. R., de Almeida, F. F. M., & Bartorelli, A. (2012). *Geologia do Brasil*. Beca São Paulo.
- Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. *Science*, 362(6410), 58–61. Retrieved from <https://doi.org/10.5066/F7PV6JNV>
- Hayes, G. P., Smoczyk, G. M., Benz, H. M., Villaseñor, A., & Furlong, K. P. (2015). Seismicity of the Earth 1900–2013: Seismotectonics of South America (Nazca Plate Region). *U.S. Geological Survey Open-File Report 2015-1031-E, 1 sheet, scale 1:14,000,000*. Retrieved from <https://dx.doi.org/10.3133/ofr20151031E>
- Heilbron, M., Cordani, U. G., & Alkmim, F. F. (2016). *São Francisco Craton, Eastern Brazil: Tectonic Genealogy of a Miniature Continent*. Springer. Retrieved from <https://www.springer.com/gp/book/9783319017143>
- Heintz, M., Debayle, E., & Vauchez, A. (2005). Upper mantle structure of the South American continent and neighboring oceans from surface wave tomography. *Tectonophysics*, 406(1-2), 115–139. Retrieved from <https://doi.org/10.1016/j.tecto.2005.05.006>
- Heit, B., Sodoudi, F., Yuan, X., Bianchi, M., & Kind, R. (2007). An S receiver function analysis of the lithospheric structure in South America. *Geophysical Research Letters*, 34(14). Retrieved from <https://doi.org/10.1029/2007GL030317>
- Heit, B., Yuan, X., Bianchi, M., Sodoudi, F., & Kind, R. (2008). Crustal thickness estimation beneath the southern central Andes at 30 S and 36 S from S wave receiver function analysis. *Geophysical Journal International*, 174(1), 249–254. Retrieved from <https://doi.org/10.1111/j.1365-246X.2008.03780.x>
- Heit, B., Yuan, X., Kind, R., & Asch, G. (2007). *Lithospheric Dynamics in the Southernmost Andean Plateau (PUDEL)*. Deutsches GeoForschungsZentrum GFZ. Retrieved from

- <https://doi.org/10.14470/7o092361>
- Hosseini, K., Matthews, K. J., Sigloch, K., Shephard, G. E., Domeier, M., & Tsekhmistrenko, M. (2018). SubMachine: Web-based tools for exploring seismic tomography and other models of Earth's deep interior. *Geochemistry, Geophysics, Geosystems*, *19*(5), 1464–1483. Retrieved from <https://doi.org/10.1029/2018GC007431>
- Hosseini, K., Sigloch, K., Tsekhmistrenko, M., Zaheri, A., Nissen-Meyer, T., & Igel, H. (2020). Global mantle structure from multifrequency tomography using P, PP and P-diffracted waves. *Geophysical Journal International*, *220*(1), 96–141. Retrieved from <https://doi.org/10.1093/gji/ggz394>
- Hynes, A., & Rivers, T. (2010). Protracted continental collision-Evidence from the Grenville orogen. *Canadian Journal of Earth Sciences*, *47*(5), 591–620. Retrieved from <https://doi.org/10.1139/E10-003>
- Inoue, H., Fukao, Y., Tanabe, K., & Ogata, Y. (1990). Whole mantle P-wave travel time tomography. *Physics of the Earth and Planetary Interiors*, *59*(4), 294–328. Retrieved from [https://doi.org/10.1016/0031-9201\(90\)90236-Q](https://doi.org/10.1016/0031-9201(90)90236-Q)
- Institute of Astronomy, G., & USP, A. S. (1988). *Brazilian Lithospheric Seismic Project (BLSP)*. International Federation of Digital Seismograph Networks. Retrieved from <http://moho.iag.usp.br/>
- Institute of Astronomy, G., & USP, A. S. (2013). *Brazilian Temporary Seismographic Experiments*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.moho.iag.usp.br/>
- Institute of Astronomy, G., & USP, A. S. (2016). *Pantanal, Chaco and Paraná (PCPB) structural studies network (PCPB)*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.moho.iag.usp.br/>
- James, D. E. (1971). Plate tectonic model for the evolution of the Central Andes. *Geological Society of America Bulletin*, *82*(12), 3325–3346. Retrieved from [https://doi.org/10.1130/0016-7606\(1971\)82\[3325:PTMFTF\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1971)82[3325:PTMFTF]2.0.CO;2)
- James, D. E., & Snoke, J. A. (1990). Seismic evidence for continuity of the deep slab beneath central and eastern Peru. *Journal of Geophysical Research: Solid Earth*, *95*(B4), 4989–5001. Retrieved from <https://doi.org/10.1029/JB095iB04p04989>
- Johnston, A., Coppersmith, K., Kanter, L., & Cornell, C. (1994). The stable continental region earthquake data-base, in *The Earthquakes of Stable Continental Regions: Assessment of Large Earthquake Potential*. *EPRI Rpt. TR-102261*.
- Jordan, T. H. (1978). Composition and development of the continental tectosphere. *Nature*, *274*(5671), 544–548.
- Kim, Y., & Clayton, R. W. (2015). Seismic properties of the Nazca oceanic crust in southern Peruvian subduction system. *Earth and Planetary Science Letters*, *429*, 110–121. Retrieved from <https://doi.org/10.1016/j.epsl.2015.07.055>
- KNMI. (2006). *Caribbean Netherlands Seismic Network*. Royal Netherlands Meteorological Institute (KNMI). Retrieved from <https://doi.org/10.21944/dffa7a3f-7e3a-3b33-a436-516a01b6af3f>
- Komatitsch, D., & Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical journal international*, *139*(3), 806–822. Retrieved from <https://doi.org/10.1046/j.1365-246x.1999.00967.x>
- Komatitsch, D., & Tromp, J. (2002a, July). Spectral-Element Simulations of Global Seismic Wave Propagation-II. Three-Dimensional Models, Oceans, Rotation and Self-Gravitation. *Geophysical Journal International*, *150*(1), 303–318. Retrieved from <https://doi.org/10.1046/j.1365-246X.2002.01716.x>
- Komatitsch, D., & Tromp, J. (2002b). Spectral-element simulations of global seismic wave propagation-I. Validation. *Geophysical Journal International*, *149*(2), 390–412. Retrieved from <https://doi.org/10.1046/j.1365-246X.2002.01653.x>
- Komatitsch, D., & Vilotte, J.-P. (1998). The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bulletin of the seismological society of America*, *88*(2), 368–392.
- Komatitsch, D., Xie, Z., Bozdağ, E., Sales de Andrade, E., Peter, D., Liu, Q., & Tromp, J.

- (2016). Anelastic sensitivity kernels with parsimonious storage for adjoint tomography and full waveform inversion. *Geophysical Journal International*, *206*(3), 1467–1478. Retrieved from <https://doi.org/10.1093/gji/ggw224>
- Kristeková, M., Kristek, J., & Moczo, P. (2009). Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of time signals. *Geophysical Journal International*, *178*(2), 813–825. Retrieved from <https://doi.org/10.1111/j.1365-246X.2009.04177.x>
- Kustowski, B., Ekström, G., & Dziewoński, A. (2008). Anisotropic shear-wave velocity structure of the Earth's mantle: A global model. *Journal of Geophysical Research: Solid Earth*, *113*(B6). Retrieved from <https://doi.org/10.1029/2007JB005169>
- Lara Wagner, S. B. (2010). *Peru Lithosphere and Slab Experiment*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/zd2010>
- Laske, G., Masters, G., Ma, Z., & Pasyanos, M. (2012). CRUST1.0: An updated global model of Earth's crust. *Geophys Res Abs*, *14*, 3743. Retrieved from <https://doi.org/10.17611/DP/emccrust10>
- Lee, E.-J., & Chen, P. (2013). Automating seismic waveform analysis for full 3-D waveform inversions. *Geophysical Journal International*, *194*(1), 572–589. Retrieved from <https://doi.org/10.1093/gji/ggt124>
- Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., ... Pugmire, D. (2020). Global adjoint tomography-model GLAD-M25. *Geophysical Journal International*, *223*(1), 1–21. Retrieved from <https://doi.org/10.1093/gji/ggaa253>
- Levander, A. (2008). *Bolivar: Western Venezuela*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/xn2008>
- Li, C., Van Der Hilst, R. D., Engdahl, E. R., & Burdick, S. (2008). A new global model for P wave speed variations in Earth's mantle. *Geochemistry, Geophysics, Geosystems*, *9*(5). Retrieved from <https://doi.org/10.1029/2007GC001806>
- Lim, H., Kim, Y., Clayton, R. W., & Thurber, C. H. (2018). Seismicity and structure of Nazca Plate subduction zone in southern Peru. *Earth and Planetary Science Letters*, *498*, 334–347. Retrieved from <https://doi.org/10.1016/j.epsl.2018.07.014>
- Liu, K. H., Gao, S. S., Silver, P. G., & Zhang, Y. (2003). Mantle layering across central South America. *Journal of Geophysical Research: Solid Earth*, *108*(B11). Retrieved from <https://doi.org/10.1029/2002JB002208>
- Liu, Q., & Gu, Y. (2012). Seismic imaging: From classical to adjoint tomography. *Tectonophysics*, *566*, 31–66. Retrieved from <https://doi.org/10.1016/j.tecto.2012.07.006>
- Lonsdale, P. (2005). Creation of the Cocos and Nazca plates by fission of the Farallon plate. *Tectonophysics*, *404*(3-4), 237–264. Retrieved from <https://doi.org/10.1016/j.tecto.2005.05.011>
- Love, A. (1927). *The mathematical theory of elasticity*. Cambridge University Press.
- Luo, Y., Modrak, R., & Tromp, J. (2013). Strategies in adjoint tomography. In *Handbook of geomathematics: Second edition* (pp. 1943–2001). Springer Berlin Heidelberg. Retrieved from https://doi.org/10.1007/978-3-642-27793-1_96-2
- Luo, Y., & Schuster, G. T. (1991). Wave-equation travelttime inversion. *Geophysics*, *56*(5), 645–653. Retrieved from <https://doi.org/10.1190/1.1443081>
- Macambira, M. J. B., Teixeira, W., Vasquez, M. L., Bartorelli, A., Teixeira, W., & Brito Neves, B. B. d. (2020). *O Cráton Amazônico e suas províncias geocronológicas: o legado de Umberto Cordani*. In *"Geocronologia e Evolução Tectônica do Continente Sul-Americano: a contribuição de Umberto Giuseppe Cordani* (Vol. 1). Solaris Edições Culturais. Retrieved from <https://repositorio.usp.br/item/003009798>
- Maggi, A., Tape, C., Chen, M., Chao, D., & Tromp, J. (2009). An automated time-window selection algorithm for seismic tomography. *Geophysical Journal International*, *178*(1), 257–281. Retrieved from <https://doi.org/10.1111/j.1365-246X.2009.04099.x>
- Mantovani, M., & Brito Neves, B. d. (2005). The Paranapanema lithospheric block: Its im-

- portance for Proterozoic (Rodinia, Gondwana) supercontinent theories. *Gondwana Research*, 8(3), 303–315. Retrieved from [https://doi.org/10.1016/S1342-937X\(05\)71137-0](https://doi.org/10.1016/S1342-937X(05)71137-0)
- Mantovani, M., Quintas, M., Shukowsky, W., & Brito Neves, B. d. (2005). Delimitation of the Paranapanema Proterozoic block: a geophysical contribution. *Episodes-Newsletter of the International Union of Geological Sciences*, 28(1), 18–22. Retrieved from [https://doi.org/10.1016/S1342-937X\(05\)71137-0](https://doi.org/10.1016/S1342-937X(05)71137-0)
- Mantovani, M. S., & Brito Neves, B. B. d. (2009). The Paranapanema lithospheric block: its nature and role in the accretion of Gondwana. *Developments in Precambrian Geology*, 16, 257–272. Retrieved from [https://doi.org/10.1016/S0166-2635\(09\)01619-3](https://doi.org/10.1016/S0166-2635(09)01619-3)
- Marquering, H., Dahlen, F., & Nolet, G. (1999). Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox. *Geophysical Journal International*, 137(3), 805–815. Retrieved from <https://doi.org/10.1046/j.1365-246x.1999.00837.x>
- MASE. (2007). *Meso America Subduction Experiment*. Caltech. Retrieved from <https://doi.org/10.7909/c3rn35sp>
- Meltzer, A., & Beck, S. (2016). *2016 Pedernales Earthquake Aftershock Deployment Ecuador*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/8g2016>
- Michael West, D. C. (2010). *Investigating the relationship between pluton growth and volcanism at two active intrusions in the central Andes*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/xp2010>
- Modrak, R., & Tromp, J. (2016). Seismic Waveform Inversion Best Practices: Regional, Global and Exploration Test Cases. *Geophysical Journal International*, 206(3), 1864–1889. Retrieved from <https://doi.org/10.1093/gji/ggw202>
- Montagner, J.-P., & Anderson, D. L. (1989). Petrological constraints on seismic anisotropy. *Physics of the earth and planetary interiors*, 54(1-2), 82–105. Retrieved from [https://doi.org/10.1016/0031-9201\(89\)90189-1](https://doi.org/10.1016/0031-9201(89)90189-1)
- Montelli, R., Nolet, G., Dahlen, F., & Masters, G. (2006). A catalogue of deep mantle plumes: New results from finite-frequency tomography. *Geochemistry, Geophysics, Geosystems*, 7(11). Retrieved from <https://doi.org/10.1029/2006GC001248>
- Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E. R., & Hung, S.-H. (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303(5656), 338–343. Retrieved from <https://doi.org/10.1126/science.1092485>
- Moura, C. A., & Gaudette, H. (1993). Evidence of Brasiliano/Panafrican deformation in the Araguaia belt: implication for Gondwana evolution. *Rev. Bras. Geoc.*, 23, 117–123. Retrieved from <https://doi.org/10.25249/0375-7536.1993232117123>
- Nocedal, J., & Wright, S. J. (2006). *Numerical Optimization* (2nd ed.). Springer.
- Nolet, G., & Dahlen, F. A. (2000). Wave Front Healing and the Evolution of Seismic Delay Times. *Journal of Geophysical Research: Solid Earth*, 105(B8), 19043–19054. Retrieved from <https://doi.org/10.1029/2000JB900161>
- Obayashi, M., Yoshimitsu, J., Nolet, G., Fukao, Y., Shiobara, H., Sugioka, H., ... Gao, Y. (2013). Finite frequency whole mantle P wave tomography: Improvement of subducted slab images. *Geophysical Research Letters*, 40(21), 5652–5657. Retrieved from <https://doi.org/10.1002/2013GL057401>
- Observatório Nacional, R. D. J. (2011). *Rede Sismográfica do Sul e do Sudeste (RSIS)*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/on>
- Oceanography, S. I. O. (1986). *IRIS/IDA Seismic Network*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/ii>
- of Bristol (UK), U. (2002). *Cape Verde Mantle Structure*. International Federation of Digital Seismograph Networks. Retrieved from http://www.fdsn.org/networks/detail/YW_2002/
- of Cambridge (UK) Earth Sciences, U. (2004). *Seismic Experiment in the Aisen Region*

- of Chile (*SEARCH*). International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/XJ\2004/>
- of Liverpool, U. (2007). *Aysen Chile Aftershock Deployment (ACAD)*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/X1\2007/>
- of Mines, B., & (Haiti), E. (2010). *Haitian Seismic Network*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/AY/>
- of Oregon, U. (1999). *Imaging Galapagos Upwelling and Neotectonics (IGUANA)*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.fdsn.org/networks/detail/YH\1999/>
- of the West Indies Mona (Jamaica), U. (1985). *Jamaica Seismograph Network*. International Federation of Digital Seismograph Networks. Retrieved from <http://www.mona.uwi.edu/earthquake/jamseisnet.htm>
- Oyhantçabal, P., Siegesmund, S., & Wemmer, K. (2011). The Río de la Plata Craton: a review of units, boundaries, ages and isotopic signature. *International Journal of Earth Sciences*, *100*(2-3), 201–220. Retrieved from <https://doi.org/10.1007/s00531-010-0580-8>
- Passarelli, C. R., Basei, M. A. S., Siga, O., & Harara, O. M. M. (2018). The Luis Alves and Curitiba terranes: continental fragments in the Adamastor Ocean. In *Geology of southwest gondwana* (pp. 189–215). Springer. Retrieved from <https://doi.org/10.1007/978-3-319-68920-3\8>
- Paul Silver, S. B. (1994). *Broadband Study of the Altiplano and Central Andes*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/xs1994>
- Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, *167*(2), 495–503. Retrieved from <https://doi.org/10.1111/j.1365-246X.2006.02978.x>
- Porter, R. C., van der Lee, S., & Whitmeyer, S. J. (2019). Synthesizing EarthScope data to constrain the thermal evolution of the continental US lithosphere. *Geosphere*, *15*(6), 1722–1737. Retrieved from <https://doi.org/10.1130/GES02000.1>
- Portner, D. E., Rodríguez, E. E., Beck, S., Zandt, G., Scire, A., Rocha, M. P., ... others (2020). Detailed structure of the subducted Nazca slab into the lower mantle derived from continent-scale teleseismic P wave tomography. *Journal of Geophysical Research: Solid Earth*, *125*(5), e2019JB017884. Retrieved from <https://doi.org/10.1029/2019JB017884>
- Pratt, R. G., Shin, C., & Hick, G. (1998). Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion. *Geophysical journal international*, *133*(2), 341–362. Retrieved from <https://doi.org/10.1046/j.1365-246X.1998.00498.x>
- Priestley, K., McKenzie, D., & Ho, T. (2018). A Lithosphere–Asthenosphere Boundary–a Global Model Derived from Multimode Surface–Wave Tomography and Petrology. *Lithospheric discontinuities*, 111–123. Retrieved from <https://doi.org/10.1002/9781119249740.ch6>
- Pritchard, M. (2009). *The life cycle of Andean volcanoes: Combining space-based and field studies*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/ys2009>
- Protti, M. (1984). *Observatorio Vulcanológico y Sismológico de Costa Rica*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/ov>
- Pulliam, J. (2013). *Greater Antilles Seismic Program*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/zc2013>
- Ramos, V. (1999). Plate tectonic setting of the Andean Cordillera. *Episodes*, *22*, 183–190. Retrieved from <https://doi.org/10.18814/epiiugs/1999/v22i3/005>
- Ramos, V. A. (2004). La Plataforma Patagónica y sus relaciones con la Plataforma Brasileira.

- Mantesso-Neto, V., Bartorelli, A., Carneiro, CDR, and Brito-Neves, BB, *Geologia do Continente Sul-Americano-Evolução da Obra de Fernando Flávio Marques de Almeida: Beca Produções Culturais Limitada [Brazil]*, 673.
- Rapela, C. W., Fanning, C. M., Casquet, C., Pankhurst, R. J., Spalletti, L., Poiré, D., & Baldo, E. G. (2011). The Rio de la Plata craton and the adjoining Pan-African/brasiliano terranes: their origins and incorporation into south-west Gondwana. *Gondwana Research*, 20(4), 673–690. Retrieved from <https://doi.org/10.1016/j.gr.2011.05.001>
- Rapela, C. W., Pankhurst, R. J., Casquet, C., Fanning, C., Baldo, E. G., González-Casado, J. M., ... Dahlquist, J. (2007). The Río de la Plata craton and the assembly of SW Gondwana. *Earth-Science Reviews*, 83(1-2), 49–82. Retrieved from <https://doi.org/10.1016/j.earscirev.2007.03.004>
- Ray, J. S., Mahoney, J. J., Duncan, R. A., Ray, J., Wessel, P., & Naar, D. F. (2012). Chronology and geochemistry of lavas from the Nazca Ridge and Easter Seamount Chain: an 30 Myr hotspot record. *Journal of Petrology*, 53(7), 1417–1448. Retrieved from <https://doi.org/10.1093/petrology/egs021>
- Rickers, F., Fichtner, A., & Trampert, J. (2012). Imaging Mantle Plumes with Instantaneous Phase Measurements of Diffracted Waves. *Geophysical Journal International*, 190(1), 650–664. Retrieved from <http://doi.org/10.1111/j.1365-246X.2012.05515.x>
- Rickers, F., Fichtner, A., & Trampert, J. (2013). The Iceland-Jan Mayen Plume System and Its Impact on Mantle Dynamics in the North Atlantic Region: Evidence from Full-Waveform Inversion. *Earth and Planetary Science Letters*, 367, 39–51. Retrieved from <http://doi.org/10.1016/j.epsl.2013.02.022>
- Rico, U. O. P. (1986). *Puerto Rico Seismic Network and Puerto Rico Strong Motion Program (PRSN and PRSMP)*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/pr>
- Rocha, M., Assumpção, M., Affonso, G., Azevedo, P., & Bianchi, M. (2019). Teleseismic P wave tomography beneath the Pantanal, Paraná, and Chaco-Paraná Basins, SE South America: Delimiting lithospheric blocks of the SW Gondwana assemblage. *Journal of Geophysical Research: Solid Earth*, 124(7), 7120–7137. Retrieved from <https://doi.org/10.1029/2018JB016807>
- Rocha, M. P., Azevedo, P. A. d., Assumpção, M., Pedrosa-Soares, A. C., Fuck, R., & Von Huelsen, M. G. (2019). Delimiting the Neoproterozoic São Francisco Palecontinental Block with P-wave travelttime tomography. *Geophysical Journal International*, 219(1), 633–644. Retrieved from <https://doi.org/10.1093/gji/ggz323>
- Rocha, M. P., Schimmel, M., & Assumpcao, M. (2011). Upper-mantle seismic structure beneath SE and Central Brazil from P-and S-wave regional travelttime tomography. *Geophysical Journal International*, 184(1), 268–286. Retrieved from <https://doi.org/10.1111/j.1365-246X.2010.04831.x>
- Rodríguez, E. E., Portner, D. E., Beck, S. L., Rocha, M. P., Bianchi, M. B., Assumpção, M., ... Lynner, C. (2021). Mantle dynamics of the Andean Subduction Zone from continent-scale teleseismic S-wave tomography. *Geophysical Journal International*, 224(3), 1553–1571. Retrieved from <https://doi.org/10.1093/gji/ggaa536>
- Rosa, M. L., Collaço, B., Assumpção, M., Sabbione, N., & Sánchez, G. (2016). Thin crust beneath the Chaco-Paraná Basin by surface-wave tomography. *Journal of South American Earth Sciences*, 66, 1–14. Retrieved from <https://doi.org/10.1016/j.jsames.2015.11.010>
- Ruan, Y., Lei, W., Modrak, R., Örsvuran, R., Bozdağ, E., & Tromp, J. (2019). Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example. *Geophysical Journal International*, 219(2), 1225–1236. Retrieved from <https://doi.org/10.1093/gji/ggz356>
- Russo, R. (2004). *Studies of crust and upper mantle structure, mantle flow and geodynamics of the Chile Ridge subduction zone*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/yj2004>
- Russo, R., Gallego, A., Comte, D., Mocanu, V., Murdie, R., & VanDecar, J. (2010). Source-

- side shear wave splitting and upper mantle flow in the Chile Ridge subduction region. *Geology*, *38*(8), 707–710. Retrieved from <https://doi.org/10.1130/G30920.1>
- Russo, R., VanDecar, J. C., Comte, D., Mocanu, V. I., Gallego, A., & Murdie, R. E. (2010). Subduction of the Chile Ridge: Upper mantle structure and flow. *Gsa Today*, *20*(9), 4–10. Retrieved from <http://doi.org/10.1130/GSATG61A.1>
- Santos, J. d. (2003). Geotectônica dos escudos das Guianas e Brasil-Central. In *Geologia, tectônica e recursos minerais do brasil* (Vol. 4, pp. 169–226). CPRM Brasília.
- Santos, J. O. S., Hartmann, L. A., Gaudette, H. E., Groves, D. I., Mcnaughton, N. J., & Fletcher, I. R. (2000). A new understanding of the provinces of the Amazon Craton based on integration of field mapping and U-Pb and Sm-Nd geochronology. *Gondwana Research*, *3*(4), 453–488. Retrieved from [https://doi.org/10.1016/S1342-937X\(05\)70755-3](https://doi.org/10.1016/S1342-937X(05)70755-3)
- Satake, K., & Atwater, B. F. (2007). Long-term perspectives on giant earthquakes and tsunamis at subduction zones. *Annu. Rev. Earth Planet. Sci.*, *35*, 349–374. Retrieved from <https://doi.org/10.1146/annurev.earth.35.031306.140302>
- Schimmel, M., Assumpção, M., & VanDecar, J. (2003). Seismic velocity anomalies beneath SE Brazil from P and S wave travel time inversions. *Journal of Geophysical Research: Solid Earth*, *108*(B4). Retrieved from <http://dx.doi.org/10.129/2001JB000187>
- Schwartz, S. (1999). *Imaging the Seismogenic Zone with Geodesy and Seismology*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/xy1999>
- Simmons, N. A., Myers, S. C., Johannesson, G., & Matzel, E. (2012). LLNL-G3Dv3: Global P wave tomography model for improved regional and teleseismic travel time prediction. *Journal of Geophysical Research: Solid Earth*, *117*(B10). Retrieved from <https://doi.org/10.1029/2012JB009525>
- Simões Neto, F., Juliã, J., & Schimmel, M. (2019). Upper-mantle structure of the Borborema Province, NE Brazil, from P-wave tomography: implications for rheology and volcanism. *Geophysical Journal International*, *216*(1), 231–250. Retrieved from <https://doi.org/10.1093/gji/ggy421>
- Sobolev, S. V., & Babeyko, A. Y. (2005). What drives orogeny in the Andes? *Geology*, *33*(8), 617–620. Retrieved from <https://doi.org/10.1130/G21557.1>
- Susan Beck, G. Z. (2007). *Lithospheric Structure and Deformation of the Flat Slab Region of Argentina*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/z12007>
- Susan Beck, G. Z. (2010). *Central Andean Uplift and the Geodynamics of the High Topography*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/zg2010>
- Susan Beck, T. W. (2000). *Slab Geometry in the Southern Andes*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/yc2000>
- Susan Y. Schwartz, A. N. (2009). *Nicoya Seismogenic Zone*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/yz2009>
- Tanimoto, T. (1995). Formalism for travelttime inversion with finite frequency effects. *Geophysical Journal International*, *121*(1), 103–110. Retrieved from <https://doi.org/10.1111/j.1365-246X.1995.tb03513.x>
- Tape, C., Liu, Q., Maggi, A., & Tromp, J. (2009). Adjoint tomography of the southern California crust. *Science*, *325*(5943), 988–992. Retrieved from <https://science.sciencemag.org/content/325/5943/988>
- Tape, C., Liu, Q., & Tromp, J. (2007). Finite-frequency tomography using adjoint methods-Methodology and examples using membrane surface waves. *Geophysical Journal International*, *168*(3), 1105–1129. Retrieved from <https://doi.org/10.1111/j.1365-246X.2006.03191.x>
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, *49*(8), 1259–1266. Retrieved from <https://doi.org/10.1190/1.1441754>

- Tassinari, C. C., & Macambira, M. J. (1999). Geochronological provinces of the Amazonian Craton. *Episodes-News magazine of the International Union of Geological Sciences*, 22(3), 174–182. Retrieved from <https://doi.org/10.18814/epiiugs/1999/v22i3/004>
- Tassinari, C. C. G., & Macambira, M. (2004). A evolução tectônica do Cráton Amazônico. *Geologia do continente sul-americano: evolução da obra de Fernando Flávio Marques de Almeida*, 28, 471–485.
- Teixeira, W., Reis, N., Bettencourt, J., Klein, E., & Oliveira, D. (2019). Intraplate Proterozoic magmatism in the Amazonian Craton reviewed: geochronology, crustal tectonics and global barcode matches. In *Dyke swarms of the world: A modern perspective* (pp. 111–154). Springer. Retrieved from https://doi.org/10.1007/978-981-13-1666-1_4
- Teixeira, W., Tassinari, C. C. G., Cordani, U. G., & Kawashita, K. (1989). A review of the geochronology of the Amazonian Craton: tectonic implications. *Precambrian Research*, 42(3-4), 213–227. Retrieved from [https://doi.org/10.1016/0301-9268\(89\)90012-0](https://doi.org/10.1016/0301-9268(89)90012-0)
- Tesoniero, A., Auer, L., Boschi, L., & Cammarano, F. (2015). Hydration of marginal basins and compositional variations within the continental lithospheric mantle inferred from a new global model of shear and compressional velocity. *Journal of Geophysical Research: Solid Earth*, 120(11), 7789–7813. Retrieved from <https://doi.org/10.1002/2015JB012026>
- Tromp, J., Komatitsch, D., Hjörleifsdóttir, V., Liu, Q., Zhu, H., Peter, D., ... Hutko, A. (2010). Near Real-Time Simulations of Global CMT Earthquakes. *Geophysical Journal International*, 183(1), 381–389. Retrieved from <https://doi.org/10.1111/j.1365-246X.2010.04734.x>
- Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160(1), 195–216. Retrieved from <https://doi.org/10.1111/j.1365-246X.2004.02453.x>
- Universidade de Brasília. (1995). *University of Brasilia Seismic Network*. International Federation of Digital Seismograph Networks. Retrieved from <http://obsis.unb.br/portalsis/>
- VanDecar, J. C., James, D. E., & Assumpção, M. (1995). Seismic evidence for a fossil mantle plume beneath South America and implications for plate driving forces. *Nature*, 378(6552), 25–31. Retrieved from <https://doi.org/10.1038/378025a0>
- Van der Lee, S. (2002). High-resolution estimates of lithospheric thickness from Missouri to Massachusetts, USA. *Earth and Planetary Science Letters*, 203(1), 15–23. Retrieved from [https://doi.org/10.1016/S0012-821X\(02\)00846-4](https://doi.org/10.1016/S0012-821X(02)00846-4)
- Van der Lee, S., James, D., & Silver, P. (2001). Upper mantle S velocity structure of central and western South America. *Journal of Geophysical Research: Solid Earth*, 106(B12), 30821–30834. Retrieved from <https://doi.org/10.1029/2001JB000338>
- Van der Lee, S., & Nolet, G. (1997). Upper mantle S velocity structure of North America. *Journal of Geophysical Research: Solid Earth*, 102(B10), 22815–22838. Retrieved from <https://doi.org/10.1029/97JB01168>
- van der Lee, S., Regenauer-Lieb, K., & Yuen, D. A. (2008). The role of water in connecting past and future episodes of subduction. *Earth and Planetary Science Letters*, 273(1-2), 15–27. Retrieved from <https://doi.org/10.1016/j.epsl.2008.04.041>
- Vasquez, M., Rosa-Costa, L. d., Silva, C., & Klein, E. (2008). Compartimentação tectônica. *Geologia e Recursos Minerais do Estado do Pará: Sistema de Informações Geográficas-SIG: Texto explicativo dos mapas geológico e tectônico e de recursos minerais do Estado do Pará. Escala, 1(1.000), 000*. Retrieved from <http://rigeo.cprm.gov.br/jspui/handle/doc/10443>
- Vasquez, M. L. (2006). *Geocronologia em zircão, monazita e granada e isótopos de Nd das associações litológicas da porção oeste do Domínio Bacajá*. Retrieved from <http://rigeo.cprm.gov.br/jspui/handle/doc/165>
- Vernon, F., Pavlis, G., Levander, A., & Wallace, T. (2003). *Crust-Mantle Interactions*

- during Continental Growth and High-Pressure Rock Exhumation at an Oblique Arc-Continent Collision Zone: SE Caribbean Margin. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/xt2003>
- Viriotte, J.-P., & RESIF. (2011). *Seismic network XS:CHILE MAULE aftershock temporary experiment (RESIF-SISMOB)*. RESIF - Réseau Sismologique et géodésique Français. Retrieved from <https://doi.org/10.15778/resif.xs2010>
- Virieux, J., & Operto, S. (2009). An Overview of Full-Waveform Inversion in Exploration Geophysics. *Geophysics*, *74*(6), WCC1-WCC26. Retrieved from <https://doi.org/10.1190/1.3238367>
- Walther, C., & Flueh, E. (2002). Remnant of the ancient Farallon Plate breakup: A low-velocity body in the lower oceanic crust off Nicoya Peninsula, Costa Rica-evidence from wide-angle seismics. *Geophysical research letters*, *29*(19), 45–1. Retrieved from <https://doi.org/10.1029/2002GL015026>
- Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556–5564. Retrieved from <https://doi.org/10.1029/2019GC008515>
- Wiens, D. (1997). *A Broadband study of the tectonics and structure of the Antarctic Peninsula and Scotia Sea Regions*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/xb1997>
- Yuan, Y. O., Bozdağ, E., Ciardelli, C., Gao, F., & Simons, F. J. (2020). The exponentiated phase measurement, and objective-function hybridization for adjoint waveform tomography. *Geophysical Journal International*, *221*(2), 1145–1164. Retrieved from <https://doi.org/10.1093/gji/ggaa063>
- Zandt, G. (1996). *Altiplano-Puna Volcanic Complex Seismic Experiment*. International Federation of Digital Seismograph Networks. Retrieved from <https://doi.org/10.7914/sn/xh1996>
- Zelt, C. A. (1998). Lateral velocity resolution from three-dimensional seismic refraction data. *Geophysical Journal International*, *135*(3), 1101–1112. Retrieved from <https://doi.org/10.1046/j.1365-246X.1998.00695.x>
- Zhou, Y., Dahlen, F., & Nolet, G. (2004). Three-dimensional sensitivity kernels for surface wave observables. *Geophysical Journal International*, *158*(1), 142–168. Retrieved from <https://doi.org/10.1111/j.1365-246X.2004.02324.x>
- Zhou, Y., Dahlen, F., Nolet, G., & Laske, G. (2005). Finite-frequency effects in global surface-wave tomography. *Geophysical Journal International*, *163*(3), 1087–1111. Retrieved from <https://doi.org/10.1111/j.1365-246X.2005.02780.x>
- Zhou, Y., Nolet, G., Dahlen, F., & Laske, G. (2006). Global upper-mantle structure from finite-frequency surface-wave tomography. *Journal of Geophysical Research: Solid Earth*, *111*(B4). Retrieved from <https://doi.org/10.1029/2005JB003677>
- Zhu, H., Bozdağ, E., Peter, D., & Tromp, J. (2012). Structure of the European Upper Mantle Revealed by Adjoint Tomography. *Nature Geoscience*, *5*(7), 493–498. Retrieved from <https://doi.org/10.1038/ngeo1501>
- Zhu, H., Bozdağ, E., & Tromp, J. (2015). Seismic structure of the European upper mantle based on adjoint tomography. *Geophysical Journal International*, *201*(1), 18–52. Retrieved from <https://doi.org/10.1093/gji/ggu492>
- Zhu, H., Komatitsch, D., & Tromp, J. (2017). Radial Anisotropy of the North American Upper Mantle Based on Adjoint Tomography with USArray. *Geophysical Journal International*, *211*(1), 349–377. Retrieved from <https://doi.org/10.1093/gji/ggx305>

Chapter 9

Model Visualization and Sharing

Visualization of tomographic models is crucial to allow their interpretation and the publication of the results. Sharing them is paramount to ensure science reproducibility and comparison with other tomographic models, fostering scientific collaboration and knowledge exchange. The adjoint models created using SPEC-FEM3D_GLOBE (Komatitsch and Tromp, 2002a,b) are defined in the so-called Gauss-Lobatto-Legendre mesh (Komatitsch and Tromp, 1999), which is convenient to perform accurate numerical simulations, but cumbersome when it comes down to visualization and sharing, given the large size and unusual format of the binary files. In this chapter, a manuscript, submitted to *Computers & Geosciences*, describes a set of routines designed to address that problem using mesh interpolation and a combination of spherical harmonics and B-splines.

SphGLLTools: A toolbox for visualization of large seismic model files based on 3D spectral-element meshes

Caio Ciardelli^a, Ebru Bozdağ^{b,*}, Daniel Peter^c and Suzan van der Lee^d

^a*Department of Geophysics, Institute of Astronomy, Geophysics and Atmospheric Sciences - University of São Paulo*

^b*Department of Geophysics, Colorado School of Mines*

^c*Division of Physical Sciences and Engineering, King Abdullah University of Science and Technology*

^d*Department of Earth and Planetary Sciences Weinberg College of Arts and Sciences Northwestern University*

ARTICLE INFO

Keywords:

adjoint tomography
spectral elements
B-splines
Fourier series
associated Legendre polynomials
spherical harmonics
Gaussian filtering
interpolation
visualization

ABSTRACT

Adjoint tomography, a full-waveform inversion technique based on 3D wave simulations, is now commonly used in earthquake seismology, drawing on advances in computational power and numerical methods. From global to reservoir scales, the resolution of seismic models can iteratively be updated by taking advantage of 3D numerical wave simulations to extract information from full seismic waveforms. Constructed seismic models, called adjoint models, are typically updated on the mesh used for numerical wave simulations. Thus the size of model files is controlled by the resolution of numerical simulations, which increases with increased numerical resolution. This is specifically a concern for recent global-scale full-waveform adjoint tomographic models where the size and format of numerical meshes pose challenges for model visualization, processing, analysis, interpretation, and sharing model files with other researchers. Here, we present SphGLLTools, an open-source toolbox that intends to diminish these challenges by expanding global adjoint waveform models onto spherical harmonic functions, which are widely used by the global seismology community. Our tools are designed for spectral-element meshes such as those used in recent global adjoint waveform models. We provide a set of routines to facilitate many commonplace tasks for model visualization and analysis, including spherical-harmonics expansion of models sampled on spectral-element meshes together with associated tools with an explicit aim for easy sharing, visualization, and interpretation of large-scale seismic model files. All the developed routines are accompanied by usage instructions and available through GitHub. For transparency, reproducibility, and educational purposes, we also include Colab notebooks, which provide an intuitive and comprehensive review of the principles and methods for the spectral-element meshes and spherical-harmonic expansion as well as other model analysis tools.

1. Introduction

Imaging Earth's deep interior has always been one of the major goals of geophysics. Seismic tomography is a powerful technique in which seismologists use data from multiple active or passive seismic sources (i.e., explosions, earthquakes, ambient noise, etc.) that generate seismic waves recorded by a set of seismic stations on the surface to create Earth models in terms of (an)elastic parameters of the medium. Improving the resolution of these models is crucial to have better insights into the inner structure, thermo-chemical composition, and geodynamical evolution of our planet.

Traditionally, seismic tomography is based on ray theory, a high-frequency approximation to describe wave propagation as in optics, which is valid in smooth media when the dominant seismic wavelength is less than the scale-length of heterogeneities (Wang and Dahlen, 1995; Spetzler et al., 2001; Arora et al., 2011). Ray theory has provided robust seismic models of Earth's interior within its validity regime, for instance, based on traveltimes of body waves, surface-wave phase and group velocities, normal-mode splitting functions, some S-wave and surface-wave waveforms, and the combination of all these data types to improve data coverage at the global scale (e.g., Chang et al.,

*Corresponding author

✉ caio.ciardelli@gmail.com (C. Ciardelli); bozdag@mines.edu (E. Bozdağ); daniel.peter@kaust.edu.sa (D. Peter); suzan@earth.northwestern.edu (S. van der Lee)

🌐 <https://ebruclsm.wordpress.com/> (E. Bozdağ); <https://danielpeter.github.io/> (D. Peter); <http://geophysics.earth.northwestern.edu/seismology/suzan/> (S. van der Lee)

ORCID(s): 0000-0001-5316-5886 (C. Ciardelli); 0000-0002-4269-3533 (E. Bozdağ); 0000-0002-3397-5379 (D. Peter); 0000-0003-1884-1185 (S. van der Lee)

2014; Ritsema et al., 1999, 2011; Kustowski et al., 2008; Moulik and Ekström, 2014). However, in the current resolution of seismic models finite-frequency effects have become important (Woodhouse and Girnius, 1982; Marquering et al., 1999) and have been addressed in many global mantle models (e.g., Li and Romanowicz, 1996; Montelli et al., 2004).

Ever-increasing computational power and advances in 3D (an)elastic numerical wave propagation solvers in recent years have made full-waveform inversion (FWI) techniques feasible for large-scale earthquake seismology problems from crustal to global scales to address the mitigation of seismic hazard in earthquake-prone regions, understanding the composition and dynamics of the mantle by imaging the lithosphere and the upper mantle, and the whole mantle down to the core-mantle boundary. First applications of FWI at the global scale combined full waveforms based on 3D wave spectral-element wave simulations with 2D asymptotic data sensitivity kernels (e.g., Lekić and Romanowicz, 2011; French and Romanowicz, 2014). In finite-frequency theory, data sensitivity kernels, also known as Frechét derivatives, replace the one-dimensional rays (Marquering et al., 1998). The adjoint method allows for the incorporation of 3D wave simulations in the seismic FWI problem. The theory was introduced to seismology by Tarantola (1984) where data sensitivity kernels of a chosen misfit function may be computed based on the interaction of 1) a forward wavefield, generated by a regular seismic source recorded by a regular seismic station, and 2) an adjoint wavefield, generated by a fictitious source based on the chosen misfit between data and synthetics at the location of the regular station and recorded by a fictitious station located at the location of the regular seismic source (Tromp et al., 2005; Fichtner et al., 2006a,b). Hence, to compute 3D volumetric data sensitivity kernels, so-called adjoint kernels, and iteratively update seismic models, only two numerical simulations, one for forward and one for adjoint wavefields, are required (Tromp et al., 2005).

Adjoint tomography has become a routine tool in earthquake seismology that there are now many applications at the crustal (e.g., Tape et al., 2009; Fichtner et al., 2013; Wang et al., 2018) and continental (e.g., Fichtner et al., 2009; Zhu et al., 2012; Chen et al., 2014; Zhu et al., 2015) scales. Finally, following the first demonstration by Afanasiev et al. (2016) first-generation global adjoint models to image the whole Earth's mantle simultaneously inverted with the crust published by Bozdağ et al. (2016); Lei et al. (2020). The main advantage of FWI is the ability to extract more information from each seismic waveforms in seismic inversions by modeling the full complexity of wave propagation both in forward and adjoint wavefields based on 3D numerical simulations. Furthermore, crust and mantle may simultaneously be updated in iterative inversions without the need for crustal corrections, which may cause biased interpretation of the upper mantle structure and anisotropy (Bozdağ and Trampert, 2008; Ferreira et al., 2010; Bozdağ et al., 2016).

FWI relies on high-accuracy seismic waveform simulations. The spectral-element method has been most efficient to simulate wave propagation at the global scale (e.g., Komatitsch and Tromp, 1999; Chaljub et al., 2003; Capdeville et al., 2003; Komatitsch and Tromp, 2002a,b). The spectral-element method is a higher-order finite element method which uses a special type basis functions called Gauss-Lobatto-Legendre (GLL) polynomials (Komatitsch and Tromp, 2002a). Simulations are performed based on a mesh sampled at every GLL point. Spectral elements are a special kind of irregular grid in which the points are defined upon a GLL basis and then mapped to the desired shape. Due to the chosen basis functions, the spectral-element method has the accuracy of pseudo-spectral methods and the flexibility of finite elements to mesh the computational domain (Komatitsch and Tromp, 2002a; Schubert, 2003). We achieve the accuracy and flexibility through a particular node distribution in which the grid points concentrate near the boundaries of the elements, which improves the interpolation for the same polynomial degree, allowing a high-accuracy integration (e.g., Komatitsch and Tromp, 1999) (also see the Colab on Numerical methods) such as trapezium and Simpson's rule. The flexibility of spectral elements allows for accurately superimpose 3D crustal models on top of 3D mantle models, including topography and bathymetry. The variable crustal thickness may be honored to improve computational accuracy and efficiency, especially for surface waves propagating under oceans, by stretching and distorting the spectral elements (Tromp et al., 2010). Mainly to reduce the computational cost homogeneization of the crust is also preferred in FWI applications (Capdeville and Marigo, 2007; French and Romanowicz, 2014; Fichtner et al., 2013).

In principle, the tomographic model discretization follows three main ideas: 1) adapt the parameterization to a priori information (e.g., finer discretization in the lithosphere, where we expect stronger heterogeneities than the rest of the model, increased spacing by depth, etc.), 2) in case of numerical simulations, adapt and optimize the mesh in the numerical solver (i.e., because of increase in wavelengths by depth the element size in mesh might be increased), and 3) adapt to the data coverage in the inversion grid (i.e., address sparsity of data in inversions). Thus, one would have three different grids, and the problem lies in interpolating between them (Kissling et al., 2001). Generally, adaptive local grids (e.g., Zhou, 1996; Boschi and Dziewonski, 2000) or spherical harmonics (e.g., Li and Romanowicz, 1996; Ritsema et al., 1999; Boschi and Ekström, 2002; Lebedev et al., 2005) are used for horizontal parameterization with polynomials and splines in the vertical direction. The choice of parameterization is one of the factors that might control

the resolution of tomographic models (e.g. Giardini et al., 1988; Snieder et al., 1991). Typically FWI models are sampled on the mesh used for numerical simulations. The main advantage is that sensitivity kernels, models, and all the generated mesh files have the same spatial sampling that models can be updated on the same mesh that is used for numerical simulations (Bozdağ et al., 2016) during iterative inversions. The drawback is that as the size of the computational domain increases, computational and storage requirements also increase. Alternative solutions for file sizes include data compressing methods (Boehm et al., 2016) where the trade-off between the resolution and the size of models should be taken into account.

Sampling model files on GLL points makes it impractical to visualize and investigate models and share them with other researchers because of their large size, especially at the scale of the globe. The size of model files increases as the resolution of simulations increase, and parallel computing is needed to extract cross-sections at desired depths or great circles. A common way to share model files is to provide downsampled model parameters in ASCII or binary formats where the resolution, specifically in the lithosphere, may be degraded. More importantly, spectral-element simulations require the actual model files sampled on a specific mesh. Users need to interpolate and resample the original model files to the desired simulation resolution and partition spectral elements to different numbers of parallel processes. All these complexities make visualization and interpretation of models as well as numerical simulations at different resolutions impractical and time-consuming for users.

There is no consensus on model data formats to facilitate easy share between researchers. There are various plotting engines and visualization tools available in which typically seismic models are converted to a specific data format in a model library. For instance, IRIS (Incorporated Research Institutions for Seismology) provides seismic tomographic models in NetCDF format (network Common Data Form) EMC visualization tools (Trabant et al., 2012; Hutko et al., 2017) and with the The New IRIS EMC 3D Visualizer (Bahavar and Trabant, 2015), visualization tools for model preview, facilities to extract model data/metadata and access to the contributed processing software and scripts. Downsampled versions of the recent global adjoint tomography models (Bozdağ et al., 2016; Lei et al., 2020) are provided through IRIS to overcome the file size issue, reducing the model resolutions in the lithosphere. Another example is the Oxford SubMachine (Hosseini et al., 2018), which is a collection of web-based tools for the interactive visualization, analysis, and quantitative comparison of global-scale, volumetric (3-D) data sets of the subsurface, with supporting tools for interacting with other, complementary models and data sets. Other tools, such as Postpischl et al. (2011) and Yamagishi et al. (2018) use the Keyhole Markup Language (KML) visualization layers (Burggraf, 2015) with the Google Earth geo-browser as a flexible platform that can substitute specialized graphical tools to perform qualitative visual data analyses and comparisons.

In this work, we provide a comprehensive set of routines for efficient visualization of large seismic models, not limited to, but such as those generated as a result of global FWI based on spectral-element wave simulations. Strictly speaking, our tools are designed for the first-generation global adjoint tomography models GLAD-M15 (Bozdağ et al., 2016) and GLAD-M25 (Lei et al., 2020) and future models which are constructed based on the simulations of the open-source SPECFEM3D_GLOBE package (Komatitsch and Tromp, 2002a,b). Our open-source toolbox, SphGLLTools, converts any global adjoint model sampled on a spectral-element mesh onto spherical harmonics at any degree. Spherical harmonic expansion of spectral-element models drastically reduces the size of files, allowing for effective visualization and investigation of seismic models on a laptop and sharing them much easier with other researchers. The toolbox also involves routines for extracting and plotting horizontal & vertical cross-sections and one-dimensional profiles at desired locations and computing global mean models and block models in an efficient way based on spherical-harmonic expansions. In addition, we provide essential routines for spherical harmonic expansions such as improved interpolation, and smoothing tools work on GLL meshes as well. SphGLLTools does not presume prior knowledge of particular numerical meshes or spherical harmonic functions and may therefore be of particular interest to students; users can go through each step and the pertinent theory by working through the accompanying Google Colab notebooks. Fourteen Colab notebooks provide interactive Python shells in which users can learn/review the theory and apply equations and algorithms, change parameters, and further experiment with the code. The entire software package (the toolbox) and accompanying Colab notebooks are open source and have a manual, including application examples.

The manuscript is designed as follows: In section 2 we give background information about spectral-element simulations and associated full-waveform inversion models. After giving the theory of spherical harmonic expansions in section 3.1 and showing how to combine them with B-splines to carry out volumetric expansions in sections 3.2.1 and 3.2.2, we describe how we expanded global adjoint models onto spherical harmonics by carefully accommodating crustal variations. We present the main capabilities of SphGLLTools in section 4 and an overall view of the toolbox in

section 5 followed by conclusions.

2. Global FWI models

In spectral-element simulations, seismic model parameters (i.e., P- and S-wave velocities, density, attenuation, etc.) are sampled on a GLL mesh. In large-scale simulations, such as those performed for global adjoint tomography, the mesh is partitioned into many slices and distributed over multiple cores for parallel processing. These slices are typically stored as binary files or using more modern data formats such as NetCDF (Unidata, 2016), HDF5 (Koranne, 2011) or ADIOS (Liu et al., 2014). In all formats, unless data compression or some downsampling is applied, the files are sampled at every GLL point, leading to dozens of gigabytes of disk space depending on the model resolution. For instance, the resolution of spectral-element simulations of the recent global adjoint tomography models (Bozdağ et al., 2016; Lei et al., 2020) is 17 s, which is the minimum resolvable period in numerical simulations ($NEX = 256$ where NEX is the number of spectral elements at the surface on one side of each of six chunks of the globe) where the model size for transversely isotropic parameterization is about 12 GB for model files and 35 GB for mesh files. When the minimum resolvable period of global simulations goes down to about 9 s ($NEX = 512$), the size of model and mesh files increases to about 78 GB and 230 GB, respectively.

One of the most common ways for visualization is performed by the free visualization software Paraview (Ahrens et al., 2005) after combining all slices into a single VTU (VTK Unstructured grid) file (Schroeder et al., 2006). Paraview allows for interactive 3D volumetric plots and the extraction of horizontal and vertical cross-sections. Sample cross-sections taken by Paraview on a unit sphere are shown in Fig. 1. Creating VTU or VTK files requires direct access to the original mesh files, where generating them with the full resolution of the mesh may require parallel jobs to overcome the memory problems. Typically, for global adjoint models, they are created for downsampled models, which can degrade the model resolution in plots. Paraview provides an efficient way of plotting models in 3D. However, 2D cross-sections require additional steps, especially when we want to change the default projection.

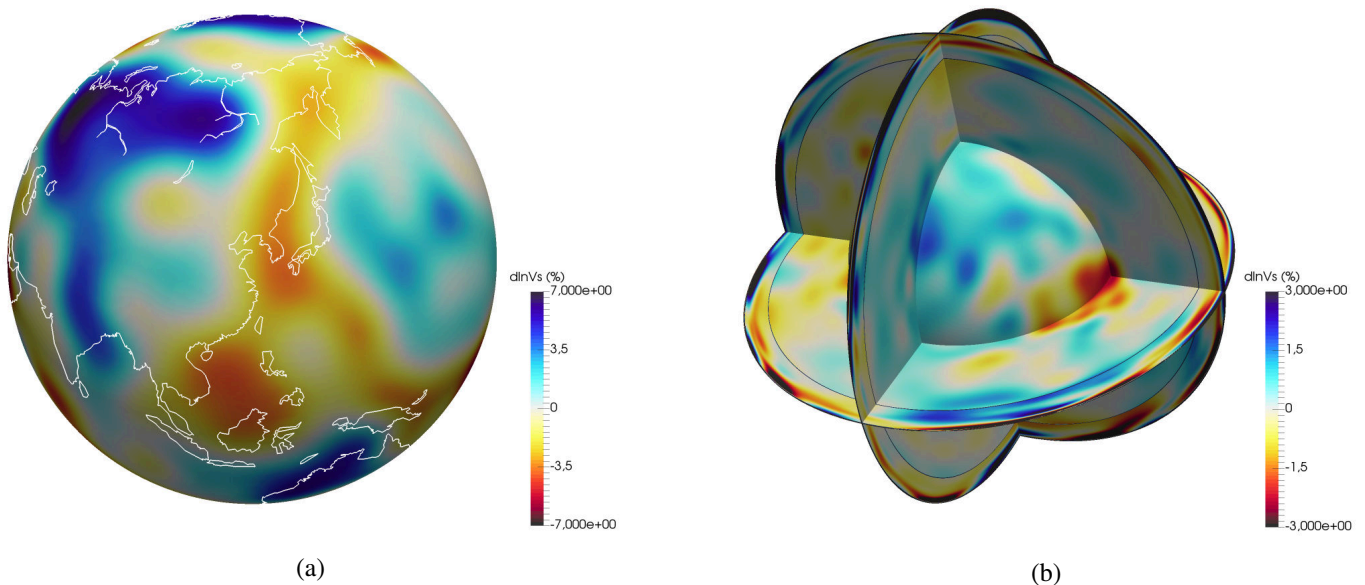


Figure 1: a) Horizontal cross-section of shear wavespeed perturbations from S362ANI (Kustowski et al., 2008) at 100 km depth. The perturbations are computed around S362ANI's mean model. The white lines show the boundaries of the continents. b) Same as a) but showing vertical cross-sections with a horizontal section at the core-mantle boundary.

For 2D cross-sections, in seismology, it is common to use GMT (Generic Mapping Tools) (Wessel et al., 2019) type of visualization packages. To this end, horizontal cross-sections at desired depths or vertical sections along desired paths need to be extracted from model files. This process requires running parallel jobs using the mesh files and the number of processors used for numerical simulations. On the other hand, it is common to parameterize ray-based global models based on spherical harmonics and splines (e.g., Ritsema et al., 1999, 2011; Chang et al., 2015) which are well known and frequently used by the seismology community. In this study, we expand global adjoint models sampled on spectral-element meshes onto spherical harmonics, which drastically reduces the model file size allowing for easy

sharing, visualization, and investigation of global adjoint tomography models. We use spherical harmonic expansion horizontally and B-splines for vertical parameterization, as we describe below. The expansion in B-splines + spherical harmonics of the whole mantle (from the core-mantle boundary to 80 km depth) of model GLAD-M15 (Bozdağ et al., 2016) occupies only ~29 MB. For GLAD-M25 (Lei et al., 2020), the size increases to ~55 MB, as a higher-degree expansion is required to represent the mantle adequately. To overcome the challenges of spherical harmonic expansion due to 3D crust we store the top 80 km as block models in binary format which slightly decreases the advantage of our approach. A block model with a $0.5^\circ \times 0.5^\circ$ horizontal and 1 km vertical resolution occupies ~537 MB when uncompressed and ~253 MB when compressed in the xz format.

3. Spherical-harmonic expansion of global adjoint models

Spherical harmonics are useful for expanding any continuous surfaces with spherical symmetry. For low-frequency signals, a relatively small set of coefficients can accurately represent the entire function. They also provide a concise way of analyzing the scale (or wavelength) content of the data. We can think of them as a generalization of the Fourier series in combination with the associated Legendre polynomials.

3.1. Spherical harmonics

It is common to use spherical harmonics in global geophysics. In seismology, some examples can be seen in ray-based seismic tomographic models (Ritsema et al., 1999, 2011; Chang et al., 2015) parametrized as a set of spherical shells expanded onto spherical harmonics. Following the convention of Dahlen and Tromp (1998), we use a normalized version of the associated Legendre polynomials in the following form:

$$P_{n,m}(\cos \theta) = \sqrt{(2 - \delta_m^0) \frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos \theta), \quad (1)$$

where δ_m^0 is the Kronecker delta,

$$\delta_m^0 = \begin{cases} 1 & \text{if } m = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

We can then expand a spherically symmetric function $f(\theta, \phi)$ into spherical harmonics,

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=0}^n (a_{m,n} \cos m\phi + b_{m,n} \sin m\phi) P_{n,m}(\cos \theta), \quad (3)$$

where

$$\begin{aligned} a_{m,n} &= \int_{-\pi}^{\pi} \int_0^{\pi} f(\theta, \phi) P_{n,m}(\cos \theta) \cos m\phi \sin \theta d\theta d\phi, \\ b_{m,n} &= \int_{-\pi}^{\pi} \int_0^{\pi} f(\theta, \phi) P_{n,m}(\cos \theta) \sin m\phi \sin \theta d\theta d\phi. \end{aligned} \quad (4)$$

3.2. Volumetric expansion

The chosen parameterization is one of the main factors controlling the resolution of tomographic models. Commonly used spherical harmonics introduced to mantle models by (Dziewonski et al., 1977) uses shells in the vertical direction. For instance, the global mantle model S20RTS (Ritsema et al., 1999) is parameterized horizontally in terms of spherical harmonics up to degree $N_{max} = 20$, and radially with a set of cubic splines. Similar parametrizations are used for mantle models SAW24B16 (Méglin and Romanowicz, 2000), S362D1 (Gu et al., 2001), and S40RTS (Ritsema et al., 2011), in which the maximum degree of the spherical harmonic shells is $N_{max} = 24$, $N_{max} = 18$, and $N_{max} = 40$,

respectively. While model S20A (Ekström and Dziewonski, 1998) uses spherical harmonics $N_{max} = 20$ horizontally, in the radial direction, two sets of Chebyshev polynomials parametrize the upper and the lower mantle separately.

In this section, we describe how we perform the volumetric expansion for global adjoint models. That approach differs from those used in other global models (section 3.1) because instead of expanding spherical shells and later connecting them using splines, we combine B-splines and spherical harmonics to create a 3D expansion capable of representing volumetric structures. We expand the current global adjoint models GLAD-M15 (Bozdağ et al., 2016) and GLAD-M25 (Lei et al., 2020) by combining spherical harmonics over θ , ϕ with cubic B-spline functions over r (Fig. 2).

3.2.1. B-splines + spherical harmonics

We can approximate a function $f(r, \theta, \phi)$ as,

$$f(r, \theta, \phi) \approx \sum_{s=0}^{N_s} \sum_{n=0}^{\infty} \sum_{m=0}^n R_s(r) (a_{m,n,s} \cos m\phi + b_{m,n,s} \sin m\phi) P_{n,m}(\cos \theta), \quad (5)$$

where

$$a_{m,n,s} = \sum_{s=0}^{N_s} A_{s',s}^{-1} \int_{-\pi}^{\pi} \int_0^{\pi} \int_{r_1}^{r_2} R_s(r) f(r, \theta, \phi) P_{n,m}(\cos \theta) \cos m\phi r^2 \sin \theta dr d\theta d\phi, \quad (6)$$

$$b_{m,n,s} = \sum_{s=0}^{N_s} A_{s',s}^{-1} \int_{-\pi}^{\pi} \int_0^{\pi} \int_{r_1}^{r_2} R_s(r) f(r, \theta, \phi) P_{n,m}(\cos \theta) \sin m\phi r^2 \sin \theta dr d\theta d\phi.$$

In eq. 6 $R_s(r) = B_{s,3}(r)$ is the radial basis composed of cubic B-splines, and $A_{s',s}$ is given by equation 7.

$$A_{s',s} = \int_{t_0}^{t_{N_s}} B_{s',n}(r) B_{s,n}(r) r^2 dr. \quad (7)$$

Fig. 2 shows an example using a test function given by eq. 8.

$$f(r, \theta, \phi) = \cos(2\pi r) \left(1 + \frac{\sin 6\theta \sin 6\phi}{10} \right), \quad \text{where } 0 \leq r \leq 1, 0 \leq \theta \leq \pi \text{ and } -\pi \leq \phi \leq \pi. \quad (8)$$

3.2.2. Expansion of global adjoint models

Our main motivation to expand global adjoint models sampled on spectral-element meshes is to facilitate easy and effective visualization, investigation (i.e., quick computation of seismic parameters such as radial anisotropy, power spectra, global mean of 3D model parameters, etc.) and sharing them with other researchers. The main challenge of expanding global adjoint models is that both B-splines and spherical harmonics cannot directly handle sharp first-order discontinuities in the mesh properly (i.e., Moho, upper-mantle discontinuities, core-mantle boundary). The most challenging one is the crust-mantle boundary (Mohorovičić discontinuity), as expected, which dramatically affects seismic wave propagation, specifically of surface waves. The Mohorovičić discontinuity can vary in depth from 7 km underneath oceans up to 70 km underneath continents (i.e., Himalayas, Andes) (e.g., Bassin et al., 2000; Laske et al., 2012) which causes significant oscillations in the spherical harmonic expansion. Classical ray-based seismic models do not suffer from this problem as they generally involve crustal corrections that mantle models start from about 24 km depth (Ritsema et al., 1999, 2011; Kustowski et al., 2008) whereas, in global adjoint models, crust and mantle are inverted simultaneously (Bozdağ et al., 2016). To avoid numerical instabilities, we preferred to represent the top 80 km of the Earth as a block model whose resolution is 0.5-degrees in the horizontal direction and 1 km in the vertical direction.

The spectral-element mesh in the SPEC-FEM3D_GLOBE package consists of three major zones: crust & mantle, fluid outer core, and the solid inner core. The mesh is partitioned into six chunks surrounding a cube of spectral elements

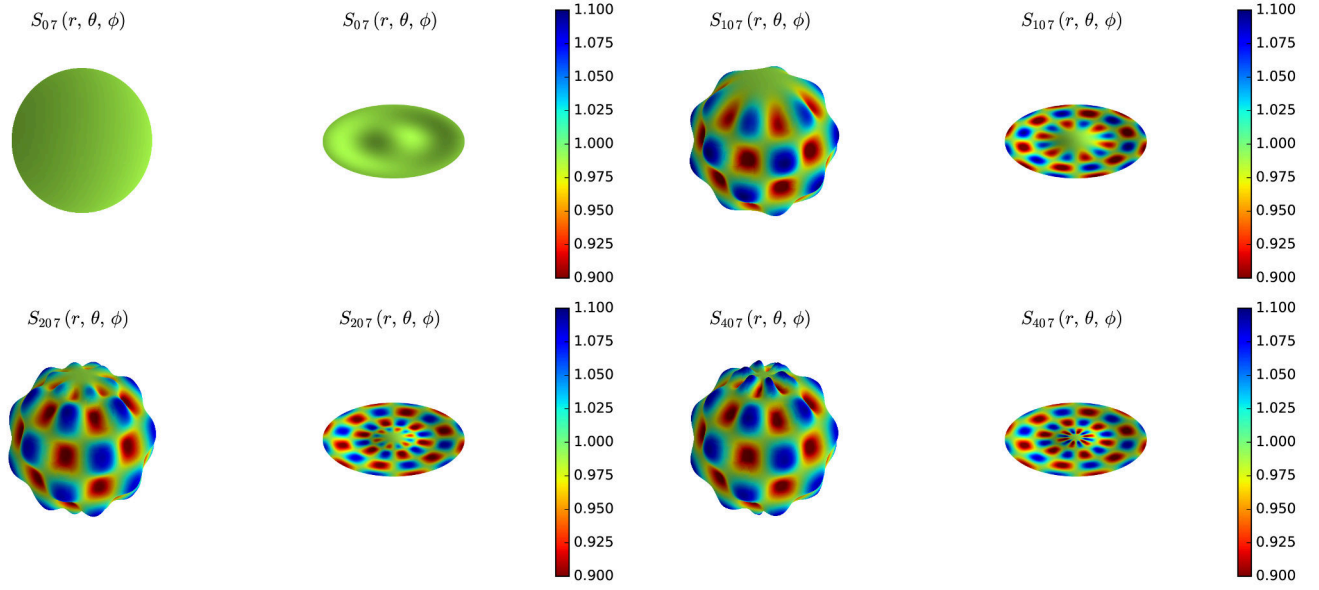


Figure 2: Example of volumetric expansion $\mathcal{S}_{n,7}(r, \theta, \phi)$ using eight cubic B-splines for degrees 0 (upper-left corner), 10 (upper-right corner), 20 (lower-left corner) and 40 (lower-right corner) of $f(r, \theta, \phi)$. On the left of each subfigure, we show the surface of the expansion. A horizontal cross-section of its interior appears on the right.

at the center of the planet. The central cube avoids the singularity at the origin of the spherical coordinate system (Chaljub, 2000). To optimize the computational cost, the size of spectral elements is doubled once below the Moho, a second time below the 650 km discontinuity, and a third time just above the inner core boundary (ICB) (Komatitsch and Tromp, 2002a). The mesh honors the first-order discontinuities of the PREM model (Dziewonski and Anderson, 1981) and the 3D Moho variations to better sample the crust in numerical simulations (Tromp et al., 2010). To create the block model, we first use interpolation to sample the mesh on a regular grid from 80 km depth (bottom of the second layer of spectral elements) to the desired elevation (e.g., 5 km above the sea level). Then, to recreate the model from the block, we use a trilinear interpolation, which does not suffer from the Gibbs phenomenon in the multiple zero-order discontinuities present within the crust.

For the rest of the mantle, instead of using a single set of cubic B-splines to represent the whole model, we expand the upper mantle, the transition zone, and the lower mantle onto spherical harmonics separately by splitting the mantle into three regions. While doing this process, we work on spherical meshes where we also suppress the internal topographies (i.e., 410 and 650 km topographies). To achieve a spherical mesh, we usually need to remesh the original mesh, as described in section 5.

We can calculate the spherical harmonic coefficients by integrating the GLL mesh through all the elements using the Gaussian quadrature:

$$\begin{aligned}
 a_{m,n,s} &= \sum_{s=0}^{N_s} A_{s',s}^{-1} \sum_{elem} \sum_{\alpha,\beta,\gamma} \omega_\alpha \omega_\beta \omega_\gamma R_s(r) f(r, \theta, \phi) J^{\alpha,\beta,\gamma} P_{n,m}(\cos \theta) \cos m\phi, \\
 b_{m,n,s} &= \sum_{s=0}^{N_s} A_{s',s}^{-1} \sum_{elem} \sum_{\alpha,\beta,\gamma} \omega_\alpha \omega_\beta \omega_\gamma R_s(r) f(r, \theta, \phi) J^{\alpha,\beta,\gamma} P_{n,m}(\cos \theta) \sin m\phi,
 \end{aligned} \tag{9}$$

where ω_α , ω_β and ω_γ are the GLL weights and $J^{\alpha,\beta,\gamma}$ is the Jacobian.

One of the advantages of carrying out independent expansions for the upper mantle, the transition zone, and the lower mantle is that we can use different degrees of spherical harmonics in each of these regions, depending on the resolution required. In the case of GLAD-M15, for example, when expanding vertically polarized shear-wavespeeds (V_{SV}), we used degree 100 for the upper mantle to fully capture the resolution of the GLL model, whereas degree 90 and degree 80 for the transition zone, and the lower mantle, respectively, was enough. For GLAD-M25, we used

Radial cubic B-splines

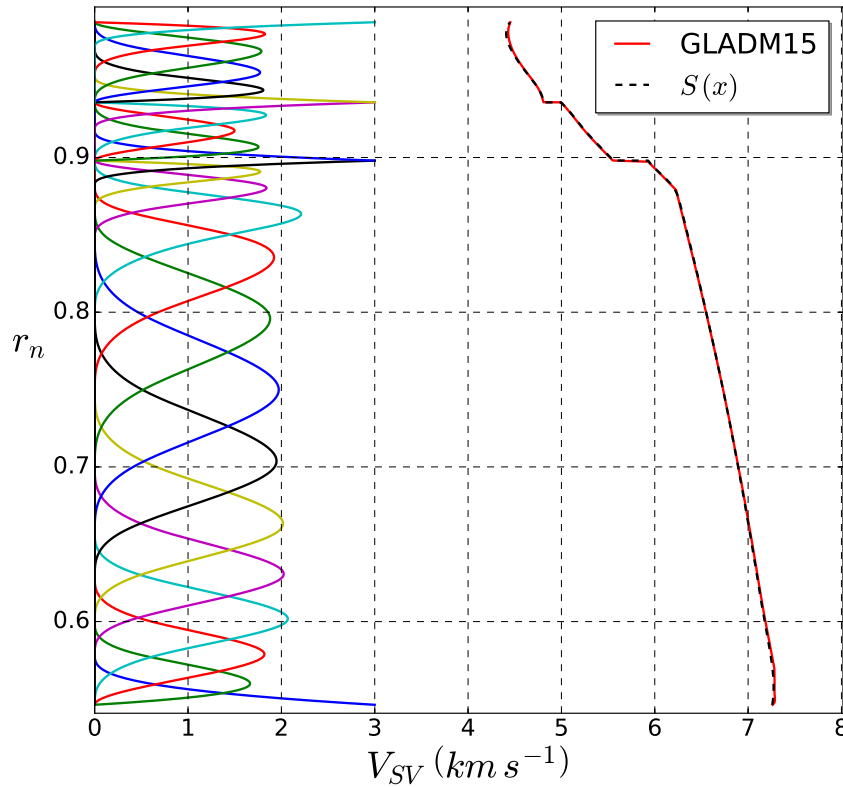


Figure 3: Twenty-five cubic B-splines used in the radial direction in the GLAD-M15 expansion. The mean V_{SV} as a function of depth appears in red. The black dashed line shows $S(x)$, the reconstruction of the model from the B-splines, and the coefficients.

degrees 155, 144, and 95 in each region, respectively. This speeds up expansions and avoids spurious oscillations in the deeper parts of the model, where a higher degree of expansion is not needed as the resolution decreases by depth.

4. Visualization

In addition to expanding spectral-element models onto spherical harmonics at desired degrees, SphGLLTools provides optimized C and python routines for multiple tasks, such as extracting horizontal and vertical cross-sections, one-dimensional profiles at desired points on the surface, and computing power spectra of the model. The package also has routines that can read the complex GLL mesh structure and interpolate it for spherical harmonic expansions and take cross-sections directly from the spectral-element mesh in ASCII format, which requires submission of parallel jobs. The routines can also calculate the model perturbations relative to the global mean of the chosen parameter (see Appendix B). The toolbox also includes Python scripts to compute additional parameters such as the bulk sound speed, transverse isotropy, V_P/V_S ratio, isotropic velocities or radial anisotropy, etc. In addition, we provide a set of Shell scripts for plotting, written in GMT6 (Wessel et al., 2019). Figs. 4 and 5 show examples of depth slices.

A great benefit of using spherical harmonics is that we can create cross-sections from the expansion with much less effort than using interpolation of the original GLL models. There is no need for parallelization and large memory. Any personal computer can do the job in a few minutes, or even seconds, depending on the maximum degree of the expansion. We can also vary the degree of the expansion to analyze the model at different wavelengths (Fig. 5). Model perturbations around the global mean can be computed directly using the C routines (Fig. 6). One-dimensional profiles as desired locations on the surface can also be easily taken without the need for parallel jobs (Fig. 7).

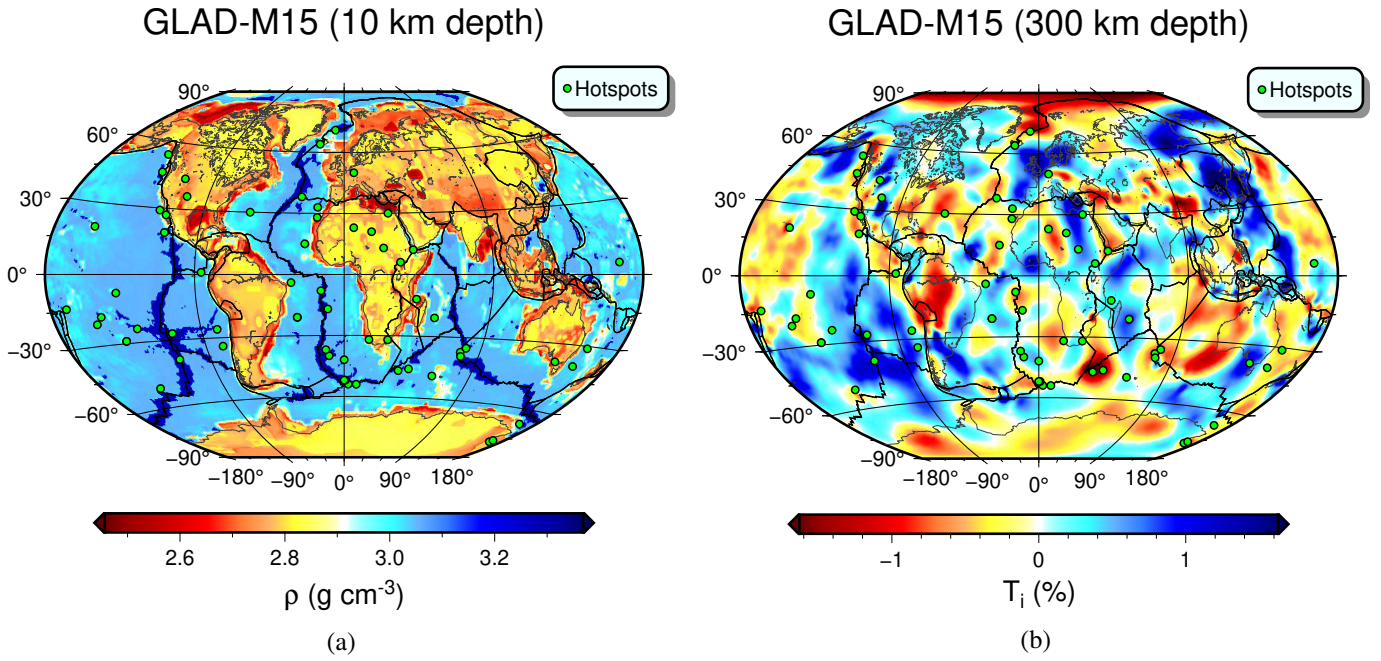


Figure 4: a) Horizontal cross-sections of density from GLAD-M15 (Bozdağ et al., 2016) at 10 km depth. The green dots represent the hot spots compiled by Don L. Anderson using a large number of sources, and the black lines denote the plate boundaries DeMets et al. (2010). b) Transverse isotropy of GLAD-M15 at 300 km depth (see Appendix A).

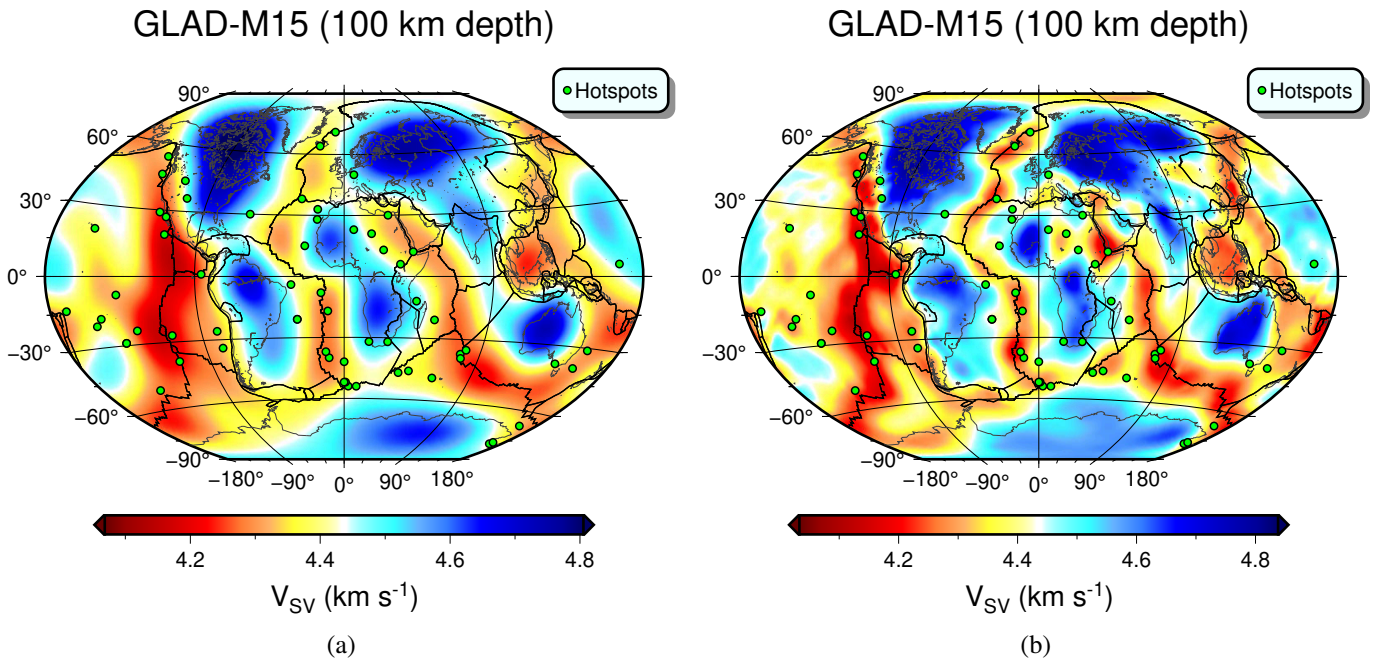


Figure 5: a) Horizontal cross-section of vertically polarized shear wavespeeds (V_{SV}) from GLAD-M15 at 100 km depth expanded up to degree 10. b) Same as a), but for spherical harmonic expansion up to degree 100.

4.0.1. Power spectra

Power spectra are useful for analyzing the spectral content of the model as a function of depth. Following Lei et al. (2020), we compute the normalized power per degree n via

$$\sigma_n^2 = \frac{1}{2n+1} \sum_{s=0}^{N_s} \sum_{s'=0}^{N_s} A_{s,s'} \sum_{m=0}^n (a_{m,n,s} a_{m,n,s'} + b_{m,n,s} b_{m,n,s'}), \quad (10)$$

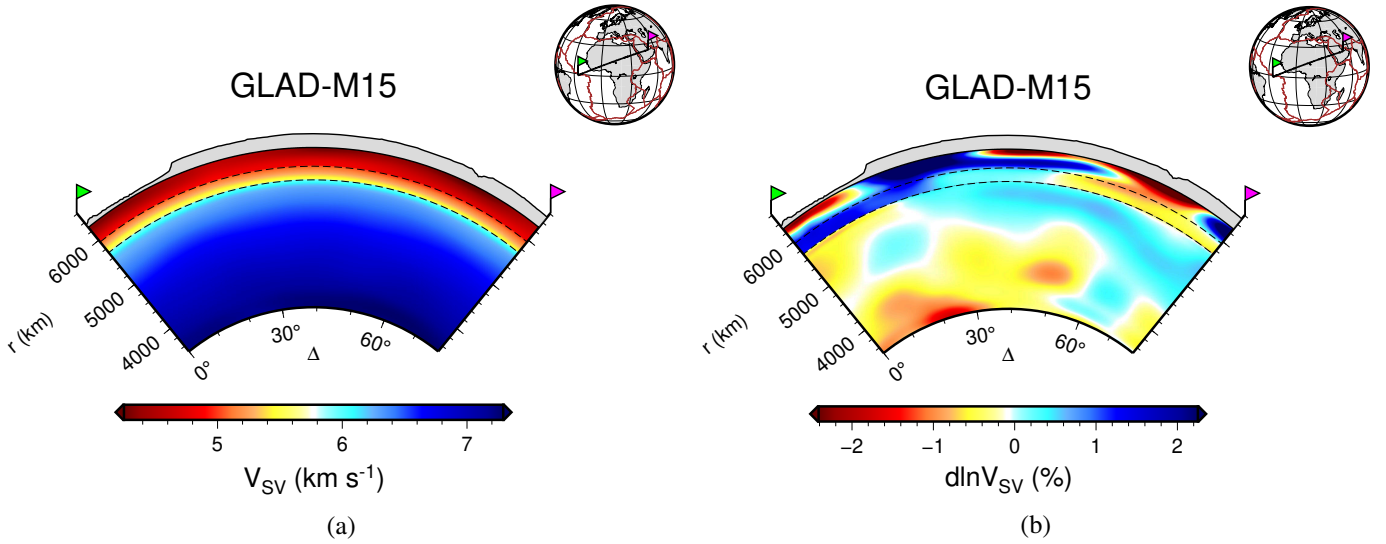


Figure 6: a) Vertical cross-section of vertically polarized shear wavespeeds (V_{SV}) from GLAD-M15 along a great circle path (the black line connecting the flags in the inset map) from the core-mantle boundary up to 80 km depth. b) Same as a), but showing model perturbations instead of the absolute values. In both figures, we expanded the model up to the maximum degrees in each zone: 100 for the upper mantle, 90 for the transition zone, and 80 for the lower mantle.

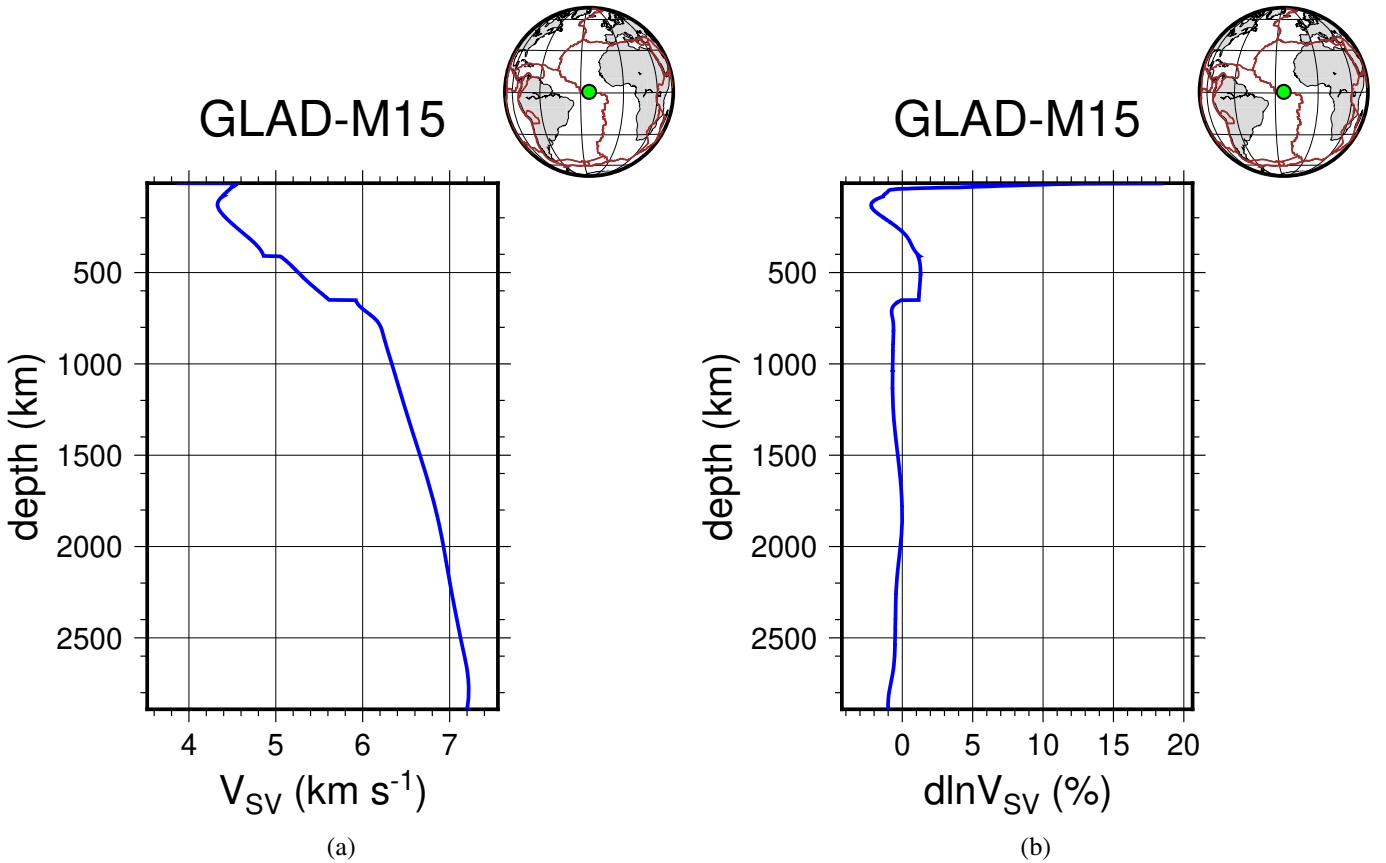


Figure 7: a) One-dimensional profile of vertically-polarized shear wavespeed (V_{SV}) from GLAD-M15 at a given location (green point in the inset map) from the core-mantle boundary up to 10 km depth. b) Same as a), but model perturbations instead of absolute values. In both figures, we expanded the model up to degree 40.

and the normalized power per degree as a function of radius is determined via

$$\Sigma_n^2(r) = \frac{1}{2} \sum_{s=0}^{N_s} \sum_{s'}^{N_s} R_s(r) R_{s'}(r) \sum_{m=0}^n (a_{m,n,s} a_{m,n,s'} + b_{m,n,s} b_{m,n,s'}) \quad (11)$$

Fig. 8 shows both kinds of power spectra computed up to degree 60 for three global mantle models; S362ANI (Kustowski et al., 2008), GLAD-M15 (Bozdağ et al., 2016), and GLAD-M25 (Lei et al., 2020), where S362ANI is a ray-based degree-20 global mantle model which is the starting model of GLAD-M15. GLAD-M15 is the result of 15 conjugate-gradient iterations, which is the starting model of GLAD-M25 constructed after 10 more L-BFGS iterations. The power spectra of these three models clearly show the evolution of the scale-length of heterogeneities in the mantle through 25 adjoint iterations.

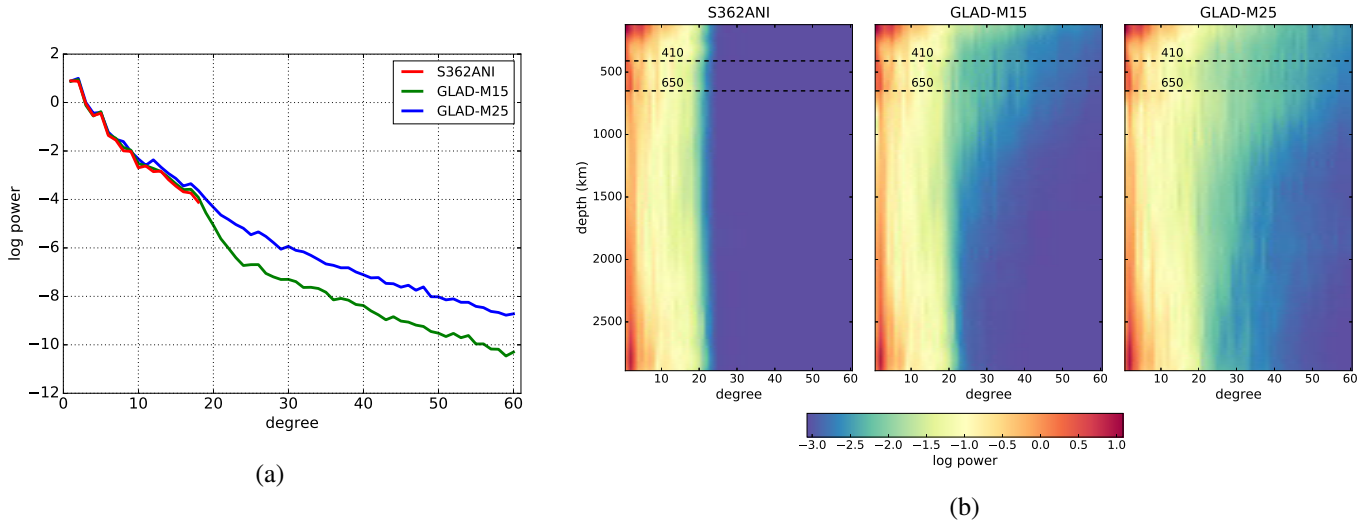


Figure 8: a) Normalized power per degree of vertically-polarized shear wavespeeds (V_{SV}) for S362ANI (red), GLAD-M15 (green), and GLAD-M25 (blue). b) Normalized power per degree as a function of depth of V_{SV} for S362ANI (left), GLAD-M15 (middle), and GLAD-M25 (right).

5. Overall view of the toolbox

Although the primary purpose of SphGLLTools is easy visualization and sharing large scale model files, we also included extra capabilities that help the user to deal with some common tasks that often occur in the workflow of creating or working with a GLL model, such as Gaussian smoothing (section 5.1) and interpolation (section 5.2). We explain the theory behind these algorithms for educational purposes and to ensure reproducibility. As explained previously, before expanding a spectral-element model onto spherical harmonics, one needs to interpolate and remesh it on a spherically symmetric mesh. Interpolation is also needed to create a block model to represent the top 80 km of global models. In addition to SphGLLTools visualization capabilities based on spherical harmonic models, the package also includes routines to take horizontal and vertical cross-sections from the spectral-element mesh based on a high-accuracy interpolation scheme we developed. These routines also assist spherical-harmonic expansions as they allow for a direct comparison using horizontal and vertical cross-sections and 1D vertical profiles between the original GLL model and the expanded version.

5.1. Gaussian smoothing in a GLL mesh

One can apply Gaussian smoothing to any kind of grid, including a Gauss-Lobatto-Legendre (GLL) mesh. Usually, it is essential to smooth the numerically computed gradient based on the adjoint method for a chosen misfit function to remove the numerical noise in the form of high-frequency artifacts (Komatitsch et al., 2016). Low-pass filtering of numerical noise is critical before computing the gradient for the model update. On the other hand, smoothing can also be applied to spectral-element models before interpolation if need be to avoid some potential numerical artifacts, which may also be useful to prevent aliasing when interpolating them to lower-resolution meshes. SphGLLTools has a routine for smoothing that can apply different filters in the horizontal and the vertical directions. This is important because, usually, deeper regions of the model require lower cut-off frequencies than the shallower ones. The algorithm ensures a smooth transition from one filter to another, thereby preventing numerical artifacts. It can also preserve the internal discontinuities if desired (Fig. 9).

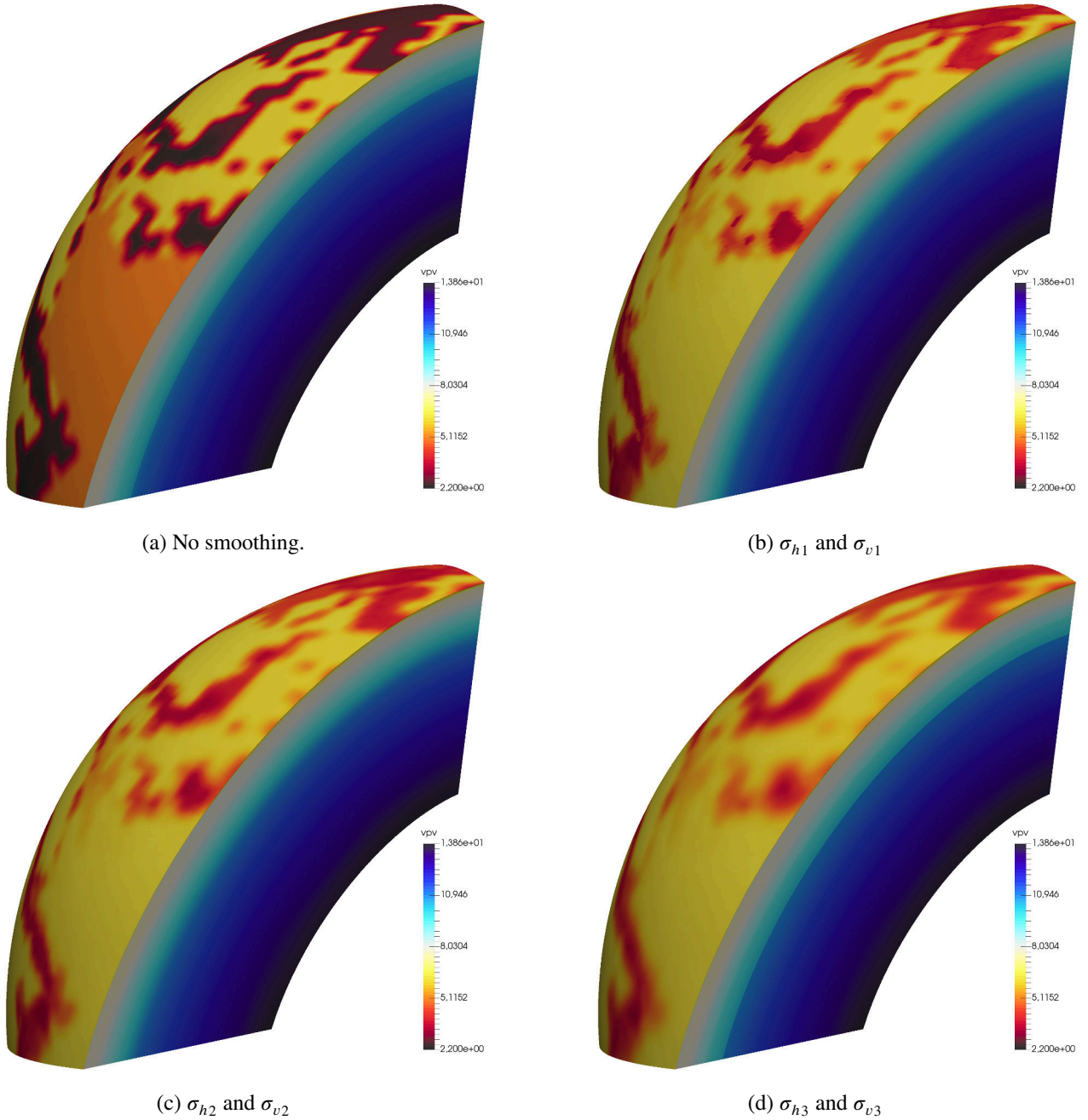


Figure 9: Model S362ANI (Kustowski et al., 2008) with different degrees of smoothing (see table 1). a) Original vertically-polarized compressional wavespeed (V_{PV}) values. b) V_{PV} values smoothed using parameters σ_{h1} and σ_{v1} . c) Smoothed using parameters σ_{h2} and σ_{v2} . d) Smoothed using parameters σ_{h3} and σ_{v3} but also preserving the 410 and 650 km discontinuities.

5.2. The spectral-element interpolation

The purpose of this section is to provide readers the theoretical basis behind the interpolation scheme we developed and implemented in the algorithms of SphGLLTools. We give an overall description of how interpolation is done in the 3D GLL mesh. The Colab notebooks provide extra details about various interpolation schemes and algorithms using 1D and 2D examples, which provide an easy way to visualize and understand 3D schemes. It suffices to understand the methods in lower dimensions since the extension to the 3D case is straightforward.

Interpolation is often needed to change the mesh resolution and the number of parallel processes. For performance reasons, SPECFEM3D_GLOBE partitions the mesh into many slices, one for each parallel process. The user chooses the number of slices and the mesh resolution when creating the mesh. However, once the mesh files are created, simulations must be run with the same number of cores and mesh resolution. Then one must remesh the spectral-element models to

Depth (km)	σ_{h1}	σ_{v1}	σ_{h2}	σ_{v2}	σ_{h3}	σ_{v3}
-10	5	5	50	5	100	5
100	10	10	50	10	100	10
200	20	20	50	20	100	20
300	30	30	50	30	100	30
410	40	40	50	40	110	40
520	50	50	50	50	120	50
650	60	60	60	60	130	60
1000	70	70	70	70	140	80
2000	80	80	80	80	150	100

Table 1

Smoothing profiles. The first column contains the depth of the top of each layer. σ_{hi} is the horizontal smoothing (half-width of the Gaussian kernel in the horizontal direction,) and σ_{vi} is the vertical smoothing (half-width in the vertical direction).

use different numbers of cores for parallel numerical simulations or change the resolution of the mesh and simulations. To this end, the original mesh should be interpolated and then re-partitioned and resampled on the desired mesh resolution (Fig. 10). Interpolation is also often needed to extract horizontal and vertical cross-sections and 1D mean profiles of parameters directly from the spectral-element mesh.

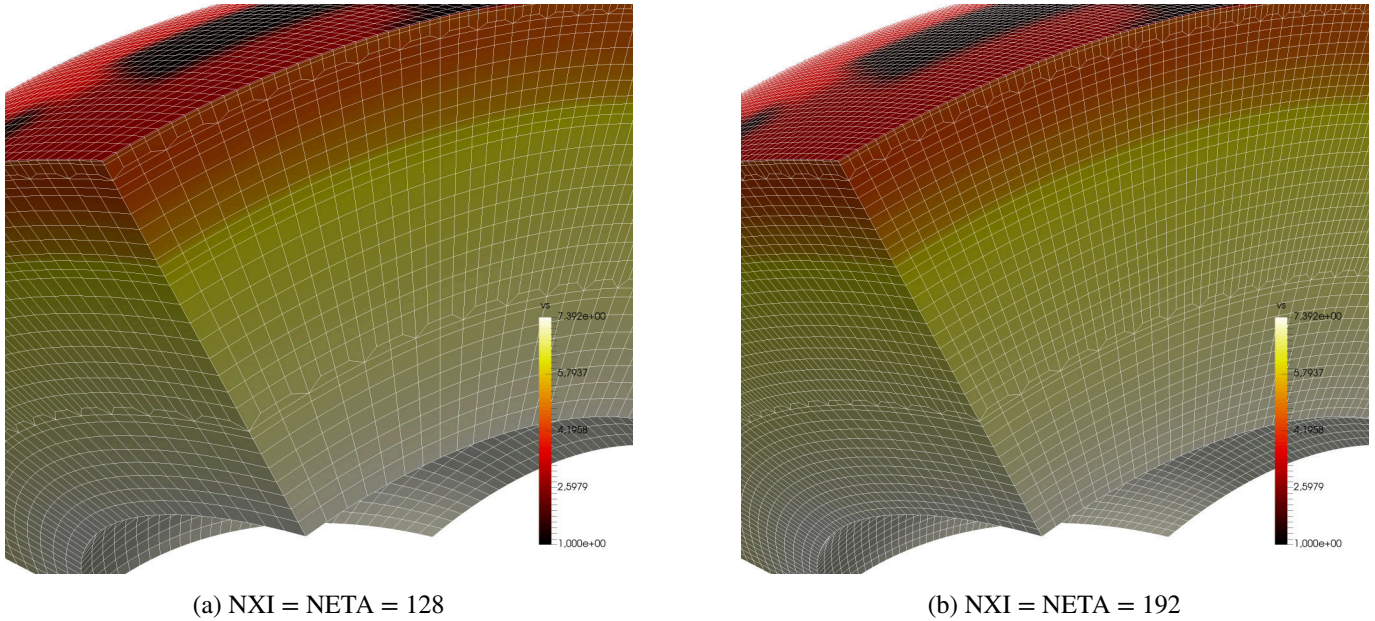


Figure 10: a) Low-resolution mesh with NEX=128, where NEX is the number of spectral elements at the surface on each side of one of six chunks that forms of the globe. b) Same as a) but for intermediate-resolution mesh with NEX=192. In both figures, the color palette denotes isotropic shear wavespeeds (V_S) in km/s, superimposed on white background.

The spectral-element method uses Lagrange interpolation which can also be used for our purposes. Nevertheless, since the elements are usually distorted and rotated, figuring out which element contains each point becomes challenging. That requires fitting polynomial surfaces to each face of all the elements, which is accomplished using a combination of multiple interpolation schemes in one, two, and three dimensions.

One of the main advantages of finite elements in the spectral-element method is that the GLL mesh can deform to accommodate the Earth's surface and internal topographies, ellipticity, and variable crustal thickness. Although each element has a unique shape, position, and orientation, they can all be mapped to the canonical basis $-1 \leq \xi \leq 1$, $-1 \leq \eta \leq 1$ and $-1 \leq \gamma \leq 1$ (where ξ , η , and γ are orthogonal spatial coordinates) and easily integrated using the Jacobian of the transformation (see Spectral-element interpolation). However, we can use Lagrange polynomials when all the dimensions of the elements are independent of each other, like in the canonical basis (ξ, η, γ) . That means spectral elements must be rectangles (in 2D) or regular hexahedrons (in 3D) with all the faces parallel to the planes

formed by all the three axes of the reference system. When we transform them to another basis such as (x, y, z) or even (r, θ, ϕ) , they usually undergo distortions, and the coordinates lose their linear independence, becoming coupled. Therefore, our interpolations are applied in the canonical domain.

Although Lagrange polynomials work just fine, for performance reasons, in our implementation, we assemble a linear system of equations to fit a 3D polynomial to the spectral element on the standard basis in which all elements are identical. Thus, once we have the coefficients, we can reuse them to interpolate any other element. In this paper, we call that approach *generic interpolation*. The shape functions are the map that transforms (ξ, η, γ) values from the standard basis to an (x, y, z) values on another basis (the corresponding coordinate in the deformed spectral element). When the map is linear, the inversion is trivial, and for a given set of (x_i, y_i, z_i) value, we can use the inverse map to calculate the corresponding $(\xi_i, \eta_i, \gamma_i)$ values. When the map is nonlinear, numerical optimization methods (such as Newton's method or other Householder's methods) can be used to reverse the mapping. Once we find $(\xi_i, \eta_i, \gamma_i)$, we can interpolate the values using either Lagrange polynomials or generic interpolation.

Because we define the Lagrange polynomials for the whole space, in theory, they should also work outside the spectral element. However, when there are regions in which the Jacobian becomes singular, we cannot always carry out the inverse mapping from (x, y, z) to (ξ, η, γ) . To overcome this problem, one should always check which spectral element encloses the point we want to interpolate before trying to invert the mapping using the respective Jacobian. To check whether that condition is fulfilled, we need to fit two-dimensional surfaces to each of the six faces of the element to accomplish the same goal. In SPEC3D_GLOBE, by default, each spectral element has 125 points $(5 \times 5 \times 5)$ and 27 nodes $(3 \times 3 \times 3)$. That leaves us with nine node points on each face. So, we need to fit quadratic surfaces in the following form:

$$\begin{aligned}
 P(x, y) = \sum_{i=0}^2 \sum_{j=0}^2 a_{ij} x^i y^j = a_{00} \\
 + a_{10} x + a_{01} y \\
 + a_{11} x y + a_{20} x^2 + a_{02} y^2 \\
 + a_{21} x^2 y + a_{12} x y^2 \\
 + a_{22} x^2 y^2 .
 \end{aligned} \tag{12}$$

Although the generic polynomial interpolation can handle a distorted grid, the accuracy of results may be affected depending on how uneven the node distribution is, for instance, on surfaces with extreme curvatures. However, it is possible to mitigate the problem if we choose the correct orientation of each face. We should parametrize a face whose normal vector predominantly points in the z direction in terms of $P(x, y)$. On the other hand, if the face points mainly in the y direction, we should parametrize it as $P(x, z)$. Analogously, we should use $P(y, z)$ for any surface primarily oriented in the x direction. Checking the orientation usually is not enough to guarantee the best fitting. In our implementation, we use a brute-force approach that consists of fitting surfaces for all three directions and picking the best one based on the norm of the coefficients, the mean surface curvature, and the residuals. When the algorithm fails to fit surface (Eq. 12) to the nine nodes, it automatically resorts to an alternative approach, which consists of fitting only the four corner nodes of the face using the following equation:

$$P(x, y) = a_{00} + a_{10} x + a_{01} y + a_{11} x y + a_{20} x^2 + a_{02} y^2 . \tag{13}$$

Since there are infinitely many surfaces in the above form that contain all four points, a reasonable idea would be to choose the one with minimum curvature. By imposing that the quadratic terms are as small as possible using Lagrange multipliers, when we have a regular grid, the two additional terms become zero, and we perfectly recover the bilinear interpolation. We call that approach *constrained interpolation*.

For our particular case, we can enhance the method by including the central node of the face with the four corner nodes. We end up with five points to fit with six unknowns. Again, we avoid the underdetermination by imposing a minimum-curvature surface. The penalty for using a lower-degree polynomial is that, sometimes, when a face is trying to accommodate sharp topographic variations (such as the Himalayas), the simplified surface cannot represent all the complexity properly. When that occurs, we observe a large residual between the surface and some of the points in

the face. To address those cases, we apply corrections to surfaces to reduce the residuals using a relatively simple yet robust interpolation scheme known as Inverse distance weighting (IDW). IDW is a method of interpolation that works in any number of dimensions and with any data set distribution. In general, it does not provide a high-quality result. However, the simplicity, robustness, and flexibility of IDW make it useful for applications in which we cannot apply other methods. Besides, unlike the nearest-neighbor algorithm, it provides a smooth interpolation.

6. Conclusion

We present SphGLLTools, a toolbox designed to facilitate easy visualization, processing, and analysis of tomographic models defined on spectral-element meshes sampled at every GLL point by expanding them onto spherical harmonics while keeping the top 80 km as a block model. We also lead the reader through a comprehensible yet intuitive explanation of the theory and concepts used by the routines. SphGLLTools allows for easy and practical visualization of spectral-element models using either direct interpolation or the flexible expansion using spherical harmonics while taking advantage of GMT6 to create high-quality images. The additional routines in the package that operates on spectral-element meshes in parallel for interpolation, remeshing, and smoothing are useful during the workflow that encompasses creating a new adjoint model and future studies using different resolutions computational resources.

The spherical-harmonic expansion of spectral-element models significantly reduces the size of GLL models by many orders of magnitude while also allow for a convenient way of analyzing, sharing, and exploring them. As an example, as mentioned in section 2, original GLAD-M15 and GLAD-M25 models sampled on their corresponding GLL meshes occupy 47 GB of disk space each (including model and mesh files). On the other hand, the expansions in B-splines + spherical harmonics of the whole mantle of GLAD-M15 and GLAD-M25 occupy only ~ 29 MB and ~ 55 MB, respectively. The high-resolution crustal models occupy ~ 537 MB each. Next, we will implement spherical-harmonic models directly into the numerical solver to run numerical simulations with the desired mesh resolution and number of processors avoiding intermediate steps like interpolation and partitioning of the mesh.

Acknowledgements

We gratefully acknowledge the discussions with Jeroen Tromp on the spherical harmonic expansion of global tomographic models sampled on spectral-element meshes. We thank FAPESP (Grants: 2016/03120-5, 2018/04918-6, and 2018/04917-0) for providing financial support for this research. This research was supported in part through the computational resources and staff contributions provided for the Quest high-performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology. EB is supported by her NSF-CAREER project with grant number 1945565. The code, including the documentation and the examples, is available at <https://github.com/caiociardelli/sphglltools.git>.

References

- Afanasiev, M., Peter, D., Sager, K., Simutè, S., Ermert, L., Krischer, L., Fichtner, A., 2016. Foundations for a Multiscale Collaborative Earth Model. *Geophysical Journal International* 204, 39–58. URL: <https://doi.org/10.1093/gji/ggv439>.
- Ahrens, J., Geveci, B., Law, C., 2005. Paraview: An end-user tool for large data visualization. *The visualization handbook* 717. URL: <https://doi.org/10.1016/B978-012387582-2/50038-1>.
- Arora, K., Cazenave, A., Engdahl, E.R., Kind, R., Manglik, A., Roy, S., Sain, K., Uyeda, S., 2011. *Encyclopedia of solid earth geophysics*. Springer Science & Business Media. URL: <https://link.springer.com/referencework/10.1007/978-90-481-8702-7>.
- Bahavar, M., Trabant, C., 2015. Data Services Products: EMC, A repository of Earth models. URL: <http://ds.iris.edu/ds/newsletter/vol17/17/437/the-new-iris-emc-3d-visualizer/>.
- Bassin, C., Laske, G., Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, in: EOS, p. 81.
- Boehm, C., Hanzich, M., de la Puente, J., Fichtner, A., 2016. Wavefield compression for adjoint methods in full-waveform inversion. *Geophysics* 81, R385–R397. URL: <https://doi.org/10.1190/geo2015-0653.1>.
- Boschi, L., Dziewonski, A., 2000. Whole Earth tomography from delay times of P, PcP and PKP phases: Lateral heterogeneities in the outer core and radial anisotropy in the mantle? *J. Geophys. Res.* 105, 13675–13696.
- Boschi, L., Ekström, G., 2002. New images of the earth's upper mantle from measurements of surface wave phase velocity anomalies. *J. Geophys. Res.* 107, 2059.
- Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., Pugmire, D., 2016. Global adjoint tomography: first-generation model. *Geophysical Journal International* 207, 1739–1766. URL: <https://doi.org/10.1093/gji/ggw356>.
- Bozdağ, E., Trampert, J., 2008. On Crustal Corrections in Surface Wave Tomography. *Geophysical Journal International* 172, 1066–1082. URL: <https://doi.org/10.1111/j.1365-246X.2007.03690.x>.
- Burggraf, D., 2015. OGC KML 2.3, Version 1.0. URL: <http://docs.openeospatial.org/is/12-007r2/12-007r2.html>.
- Capdeville, Y., Chaljub, E., Montagner, J.P., 2003. Coupling the spectral element method with a modal solution for elastic wave propagation in global earth models. *Geophysical Journal International* 152, 34–67. URL: <https://doi.org/10.1046/j.1365-246X.2003.01808.x>.
- Capdeville, Y., Marigo, J.J., 2007. Second order homogenization of the elastic wave equation for non-periodic layered media. *Geophysical Journal International*, 823–838 URL: <https://doi.org/10.1111/j.1365-246X.2007.03462.x>.
- Chaljub, E., 2000. Modélisation numérique de la propagation d'ondes sismiques en géométrie sphérique: application à la sismologie globale. Ph.D. thesis. Université Paris-Diderot-Paris VII.
- Chaljub, E., Capdeville, Y., Vilotte, J.P., 2003. Solving Elastodynamics in a Fluid-Solid Heterogeneous Sphere: A Parallel Spectral Element Approximation on Non-Conforming Grids. *Journal of Computational Physics* 187, 457–491. URL: [https://doi.org/10.1016/S0021-9991\(03\)00119-0](https://doi.org/10.1016/S0021-9991(03)00119-0).
- Chang, S.J., Ferreira, A.M., Ritsema, J., van Heijst, H.J., Woodhouse, J.H., 2015. Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations. *Journal of Geophysical Research: Solid Earth* 120, 4278–4300. URL: <https://doi.org/10.1002/2014JB011824>.
- Chang, S.J., Ferreira, A.M.G., Ritsema, J., van Heijst, H.J., Woodhouse, J.H., 2014. Global Radially Anisotropic Mantle Structure from Multiple Datasets: A Review, Current Challenges, and Outlook. *Tectonophysics* 617, 1–19. URL: <https://doi.org/10.1016/j.tecto.2014.01.033>.
- Chen, M., Huang, H., Yao, H., van der Hilst, R., Niu, F., 2014. Low wave speed zones in the crust beneath SE Tibet revealed by ambient noise adjoint tomography. *Geophysical Research Letters* 41, 334–340. URL: <https://doi.org/10.1002/2013GL058476>.
- Dahlen, F., Tromp, J., 1998. *Theoretical global seismology*. Princeton university press.
- DeMets, C., Gordon, R.G., Argus, D.F., 2010. Geologically current plate motions. *Geophysical Journal International* 181, 1–80. URL: <https://doi.org/10.1111/j.1365-246X.2009.04491.x>.
- Dziewonski, A.M., Anderson, D.L., 1981. Preliminary Reference Earth Model. *Physics of the Earth and Planetary Interiors* 25, 297–356. URL: [https://doi.org/10.1016/0031-9201\(81\)90046-7](https://doi.org/10.1016/0031-9201(81)90046-7).
- Dziewonski, A.M., Hager, B.H., O'Connell, R.J., 1977. Large-scale heterogeneities in the lower mantle. *Journal of Geophysical Research* 82, 239–255. URL: <https://doi.org/10.1029/JB082i002p00239>.
- Ekström, G., Dziewonski, A.M., 1998. The unique anisotropy of the Pacific upper mantle. *Nature* 394, 168–172. URL: <https://doi.org/10.1038/28148>.
- Ferreira, A., Woodhouse, J., Visser, K., Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography. *Journal of Geophysical Research: Solid Earth* 115. URL: <https://doi.org/10.1029/2009JB006716>.
- Fichtner, A., Bunge, H.P., Igel, H., 2006a. The Adjoint Method in Seismology-I. Theory. *Physics of the Earth and Planetary Interiors* 157, 86–104. URL: <https://doi.org/10.1016/j.pepi.2006.03.016>.
- Fichtner, A., Bunge, H.P., Igel, H., 2006b. The Adjoint Method in Seismology-II. Applications: Traveltimes and Sensitivity Functionals. *Physics of the Earth and Planetary Interiors* 157, 105–123. URL: <https://doi.org/10.1016/j.pepi.2006.03.018>.
- Fichtner, A., Kennett, B.L.N., Igel, H., Bunge, H.P., 2009. Full Seismic Waveform Tomography for Upper-Mantle Structure in the Australasian Region Using Adjoint Methods. *Geophysical Journal International* 179, 1703–1725. URL: <https://doi.org/10.1111/j.1365-246X.2009.04368.x>.
- Fichtner, A., Saygin, E., Taymaz, T., Cupillard, P., Capdeville, Y., Trampert, J., 2013. The Deep Structure of the North Anatolian Fault Zone. *Earth and Planetary Science Letters* 373, 109–117. URL: <https://doi.org/10.1016/j.epsl.2013.04.027>.
- French, S.W., Romanowicz, B.A., 2014. Whole-Mantle Radially Anisotropic Shear Velocity Structure from Spectral-Element Waveform Tomography. *Geophysical Journal International* 199, 1303–1327. URL: <https://doi.org/10.1093/gji/ggu334>.
- Giardini, D., Li, X.D., Woodhouse, J.H., 1988. Splitting functions of long-period normal modes of the Earth. *J. Geophys. Res.* 93, 13716–13742.
- Gu, Y.J., Dziewonski, A.M., Su, W., Ekström, G., 2001. Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities. *Journal of Geophysical Research: Solid Earth* 106, 11169–11199. URL: <https://doi.org/10.1029/2001JB000340>.

- Hosseini, K., Matthews, K.J., Sigloch, K., Shephard, G.E., Domeier, M., Tsekhmistrenko, M., 2018. SubMachine: Web-based tools for exploring seismic tomography and other models of Earth's deep interior. *Geochemistry, Geophysics, Geosystems* 19, 1464–1483. URL: <https://doi.org/10.1029/2018GC007431>.
- Hutko, A.R., Bahavar, M., Trabant, C., Weekly, R.T., Fossen, M.V., Ahern, T., 2017. Data products at the IRIS-DMC: Growth and usage. *Seismological Research Letters* 88, 892–903. URL: <https://doi.org/10.1785/0220160190>.
- Kissling, E., Husen, S., Haslinger, F., 2001. Model parametrization in seismic tomography: a choice of consequence for the solution quality. *Physics of the Earth and Planetary Interiors* 123, 89–101. URL: [https://doi.org/10.1016/S0031-9201\(00\)00203-X](https://doi.org/10.1016/S0031-9201(00)00203-X).
- Komatitsch, D., Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical Journal International* 139, 806–822. URL: <https://doi.org/10.1046/j.1365-246x.1999.00967.x>.
- Komatitsch, D., Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation-I. Validation. *Geophysical Journal International* 149, 390–412. URL: <https://doi.org/10.1046/j.1365-246x.2002.01653.x>.
- Komatitsch, D., Tromp, J., 2002b. Spectral-Element Simulations of Global Seismic Wave Propagation—II. Three-Dimensional Models, Oceans, Rotation and Self-Gravitation. *Geophysical Journal International* 150, 303–318. URL: <https://doi.org/10.1046/j.1365-246x.2002.01716.x>.
- Komatitsch, D., Xie, Z., Bozdağ, E., Sales de Andrade, E., Peter, D., Liu, Q., Tromp, J., 2016. Anelastic sensitivity kernels with parsimonious storage for adjoint tomography and full waveform inversion. *Geophysical Journal International* 206, 1467–1478. URL: <https://doi.org/10.1093/gji/ggw224>.
- Koranne, S., 2011. Hierarchical data format 5: HDF5, in: *Handbook of Open Source Tools*. Springer, pp. 191–200.
- Kustowski, B., Ekström, G., Dziewoński, A., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: A global model. *Journal of Geophysical Research: Solid Earth* 113. URL: <https://doi.org/10.1029/2007JB005169>.
- Laske, G., Masters, G., Ma, Z., Pasyanos, M., 2012. CRUST1.0: An updated global model of Earth's crust. *Geophys Res Abs* 14, 3743. URL: <https://doi.org/10.17611/DP/emccrust10>.
- Lebedev, S., Nolet, G., Meier, T., van der Hilst, R.D.G., 2005. Automated multimode inversion of surface and S waveforms. *Geophys. J. Int.* 162, 951–964.
- Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., Pugmire, D., 2020. Global adjoint tomography-model GLAD-M25. *Geophysical Journal International* 223, 1–21. URL: <https://doi.org/10.1093/gji/ggaa253>.
- Lekić, V., Romanowicz, B., 2011. Inferring Upper-Mantle Structure by Full Waveform Tomography with the Spectral Element Method. *Geophysical Journal International* 185, 799–831. URL: <https://doi.org/10.1111/j.1365-246x.2011.04969.x>.
- Li, X.D., Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory. *Journal of Geophysical Research: Solid Earth* 101, 22245–22272. URL: <https://doi.org/10.1029/96JB01306>.
- Liu, Q., Logan, J., Tian, Y., Abbasi, H., Podhorszki, N., Choi, J.Y., Klasky, S., Tchoua, R., Lofstead, J., Oldfield, R., et al., 2014. Hello ADIOS: the challenges and lessons of developing leadership class I/O frameworks. *Concurrency and Computation: Practice and Experience* 26, 1453–1473. URL: <https://doi.org/10.1002/cpe.3125>.
- Marquering, H., Dahlen, F.A., Nolet, G., 1999. Three-Dimensional Sensitivity Kernels for Finite-Frequency Traveltimes: The Banana-Doughnut Paradox. *Geophysical Journal International* 137, 805–815. URL: <https://doi.org/10.1046/j.1365-246x.1999.00837.x>.
- Marquering, H., Nolet, G., Dahlen, F., 1998. Three-dimensional waveform sensitivity kernels. *Geophysical Journal International* 132, 521–534. URL: <https://doi.org/10.1046/j.1365-246x.1998.00426.x>.
- Mégnin, C., Romanowicz, B., 2000. The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms. *Geophysical Journal International* 143, 709–728. URL: <https://doi.org/10.1046/j.1365-246x.2000.00298.x>.
- Montelli, R., Nolet, G., Dahlen, F.A., Masters, G., Engdahl, E.R., Hung, S.H., 2004. Finite-Frequency Tomography Reveals a Variety of Plumes in the Mantle. *Science* 303, 338–343. URL: <https://doi.org/10.1126/science.1092485>.
- Moulik, P., Ekström, G., 2014. An Anisotropic Shear Velocity Model of the Earth's Mantle Using Normal Modes, Body Waves, Surface Waves and Long-Period Waveforms. *Geophysical Journal International* 199, 1713–1738. URL: <https://doi.org/10.1093/gji/ggu356>.
- Postpischl, L., Danecek, P., Morelli, A., Pondrelli, S., 2011. Standardization of seismic tomographic models and earthquake focal mechanisms data sets based on web technologies, visualization with keyhole markup language. *Computers & geosciences* 37, 47–56. URL: <https://doi.org/10.1016/j.cageo.2010.05.006>.
- Ritsema, J., Deuss, a.A., Van Heijst, H., Woodhouse, J., 2011. S4ORTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements. *Geophysical Journal International* 184, 1223–1236. URL: <https://doi.org/10.1111/j.1365-246x.2010.04884.x>.
- Ritsema, J., van Heijst, H.J., Woodhouse, J.H., 1999. Complex Shear Wave Velocity Structure Imaged Beneath Africa and Iceland. *Science* 286, 1925–1928. URL: <https://doi.org/10.1126/science.286.5446.1925>.
- Schroeder, W., Martin, K., Lorensen, B., 2006. *The visualization toolkit*, 4th edn. Kitware, New York.
- Schubert, B., 2003. *The spectral-element method for seismic wave propagation: theory, implementation and comparison to finite difference methods*. Master's thesis, Ludwig-Maximilians-Universität München, unpublished.
- Snieder, R., Beckers, J., Neele, F., 1991. The effect of small-scale structure on normal mode frequencies and global inversions. *J. Geophys. Res.* 96, 501–515.
- Spetzler, J., Trampert, J., Snieder, R., 2001. Are We Exceeding the Limits of the Great Circle Approximation in Global Surface Wave Tomography? *Geophysical Research Letters* 28, 2341–2344. URL: <https://doi.org/10.1029/2000GL012691>.
- Tape, C., Liu, Q., Maggi, A., Tromp, J., 2009. Adjoint tomography of the southern California crust. *Science* 325, 988–992. URL: <https://science.sciencemag.org/content/325/5943/988>.
- Tarantola, A., 1984. Inversion of Seismic Reflection Data in the Acoustic Approximation. *Geophysics* 49, 1259–1266. URL: <https://doi.org/10.1190/1.1441754>.
- Trabant, C., Hutko, A.R., Bahavar, M., Karstens, R., Ahern, T., Aster, R., 2012. Data products at the IRIS DMC: Stepping stones for research and

- other applications. *Seismological Research Letters* 83, 846–854. URL: <https://doi.org/10.1785/0220120032>.
- Tromp, J., Komatitsch, D., Hjörleifsdóttir, V., Liu, Q., Zhu, H., Peter, D., Bozdağ, E., McRitchie, D., Friberg, P., Trabant, C., Hutko, A., 2010. Near Real-Time Simulations of Global CMT Earthquakes. *Geophysical Journal International* 183, 381–389. URL: <https://doi.org/10.1111/j.1365-246X.2010.04734.x>.
- Tromp, J., Tape, C., Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International* 160, 195–216. URL: <https://doi.org/10.1111/j.1365-246X.2004.02453.x>.
- Unidata, 2016. Network Common Data Form (netCDF). URL: <http://doi.org/10.5065/D6H70CW6>.
- Wang, K., Yang, Y., Basini, P., Tong, P., Tape, C., Liu, Q., 2018. Refined crustal and uppermost mantle structure of southern California by ambient noise adjoint tomography. *Geophysical Journal International* 215, 844–863. URL: <https://doi.org/10.1093/gji/ggy312>.
- Wang, Z., Dahlen, F.A., 1995. Validity of Surface-Wave Ray Theory on a Laterally Heterogeneous Earth. *Geophysical Journal International* 123, 757–773. URL: <https://doi.org/10.1111/j.1365-246X.1995.tb06888.x>.
- Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W., Tian, D., 2019. The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems* 20, 5556–5564. URL: <https://doi.org/10.1029/2019GC008515>.
- Woodhouse, J.H., Girnius, T.P., 1982. Surface Waves and Free Oscillations in a Regionalized Earth Model. *Geophysical Journal International* 68, 653–673. URL: <https://doi.org/10.1111/j.1365-246X.1982.tb04921.x>.
- Yamagishi, Y., Nakanishi, A., Miura, S., Kodaira, S., Sakaguchi, H., 2018. Development of a database and visualization system integrating various models of seismic velocity structure and subducting plate geometry around Japan. *Progress in Earth and Planetary Science* 5, 1–9. URL: <https://doi.org/10.1186/s40645-018-0207-4>.
- Zhou, H.W., 1996. A high resolution P wave model of the top 1200 km of the mantle. *J. Geophys. Res.* 101, 27791–27810.
- Zhu, H., Bozdağ, E., Peter, D., Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography. *Nature Geoscience* 5, 493–498. URL: <https://doi.org/10.1038/ngeo1501>.
- Zhu, H., Bozdağ, E., Tromp, J., 2015. Seismic structure of the European upper mantle based on adjoint tomography. *Geophysical Journal International* 201, 18–52. URL: <https://doi.org/10.1093/gji/ggu492>.

A. Appendix A

The convention adopted for the transverse isotropy is:

$$T_i = \ln \left(\frac{V_{SV}}{V_{SH}} \right) - \ln \left(\frac{V_{SV\,mean}}{V_{SH\,mean}} \right), \quad (14)$$

where $V_{SV\,mean}$ and $V_{SH\,mean}$ are the one-dimensional mean models for V_{SV} and V_{SH} , respectively. By mean model, we refer to the arithmetic mean of the three-dimensional model per depth.

B. Appendix B

Following Tarantola, we use the logarithmic of the model parameters (velocities, density, etc.): $\ln(M)$. It follows that a perturbation is:

$$d\ln M = \ln(M) - \ln(M_{mean}) = \ln \left(\frac{M}{M_{mean}} \right), \quad (15)$$

where M is the value of the three-dimensional model, and M_{mean} is the one-dimensional mean model, i.e. the mean of all values of velocity, density, etc. at a given depth.

The above result, however, is not balanced as the mean of the perturbations is close, but not exactly zero. To overcome that problem, we use instead:

$$d\ln M = \ln \left(\frac{M}{M_{mean}} \right) - \left[\ln \left(\frac{M}{M_{mean}} \right) \right]_{mean}. \quad (16)$$

But, with some manipulation of the right-hand side:

$$\begin{aligned}
 d \ln M &= \ln(M) - \ln(M_{mean}) - [\ln(M) - \ln(M_{mean})]_{mean} \\
 &= \ln(M) - \ln(M_{mean}) + \ln(M_{mean}) - \frac{1}{N} \sum_{i=0}^N \ln(M_i) = \ln(M) - \frac{1}{N} \ln\left(\prod_{i=0}^N M_i\right) \\
 &= \ln(M) - \ln\left(\prod_{i=0}^N M_i\right)^{\frac{1}{N}} \\
 &= \ln(M) - \ln(M_{gmean}) \\
 &= \ln\left(\frac{M}{M_{gmean}}\right),
 \end{aligned} \tag{17}$$

where M_{gmean} is a one-dimensional geometric-mean model.

Chapter 10

Conclusion

We initiated the thesis with a review of spectral-element seismic wave simulations, geophysical inversion, and adjoint state method in seismology. Then, we proceed in Chapter 6 with a simplified explanation of the finite-frequency theory. Despite its limitations, our formulation correctly predicted the most relevant features of sensitivity kernels, including the fact that travel-time kernels are hollow. The step-by-step derivation and relatively easy equations should be understandable by an undergraduate student with a reasonable knowledge of classical physics and calculus. That work provided new insight into the banana-doughnut paradox, helping Earth scientists understand the finite-frequency theory, increasingly used in seismology and seismic exploration.

In Chapter 7, we revisited some previous tomographic studies of South America, before presenting our own model in Chapter 8. In that study, we constructed an adjoint waveform model, SAAM23, based on 3D spectral-element simulations of wave propagation for the South American continent using data from 112 earthquakes recorded by 1311 seismic stations. The model results from 23 conjugate gradient iterations where we assimilated three-component data, including both body and surface waves in the inversion. During the construction of the model, we used the exponentiated-phase misfit proposed by Yuan et al. (2020), a variant of the instantaneous phase misfit (Bozdağ et al., 2011), to better take the scattered waves into account while effectively mitigating the cycle skip problem of phase measurements.

At long wavelengths, SAAM23 is compatible with previous seismic-tomographic studies. Well resolved is the lithosphere of cratonic nuclei in the stable platform (Amazonian and São Francisco) as well as that of the covered Paranapanema and Parnaíba cratonic blocks. The Nazca Slab is imaged into the lower mantle and is shown to be continuous in the 300-500 km depth range down-dip from the Peruvian flat-slab segment. The slab crosses the mantle transition zone and dips into the lower mantle beneath the northern portion of South America. In the central and southern part

of South America, the slab somewhat flattens near the 650 km discontinuity before straightening and continuing into the lower mantle.

Lithospheric thicknesses estimated from published S -wave receiver functions agree well with our Lithosphere-Asthenosphere boundary (LAB) estimated using the most significant negative velocity gradient with depth in SAAM23. Our LAB depths agree remarkably well with the temperature model of Finger et al. (2021). We additionally found a strong correlation between the average age of the geochronological provinces in the Amazonian Craton and the positive lithospheric β_v anomalies of our model. Despite different formation ages, the fact that these formation ages are all Precambrian rules out that these differences in lithospheric structure are the result of simple cooling histories. Instead, we suggest that different degrees of iron depletion account for some of the inter-provincial differences in β_v and that westwards thinning of the lithosphere might be related to vigorous asthenospheric flow and its interaction with spatio-temporally changing mantle wedge behavior on the western margin of the continent. Like previous tomographic studies, we found no high-velocity anomalies beneath the Río de la Plata Craton. Our continental-scale, high-resolution, deeply-extending, full-waveform based, new tomographic model, SAAM23, provides a comprehensive and robust basis for further studies of the evolution of the South American Plate.

The thesis ends with Chapter 9 in which we presented SphGLLTools, a toolbox designed to facilitate easy visualization, processing, and analysis of tomographic models defined on spectral-element meshes sampled at every GLL point by expanding them onto spherical harmonics while keeping the top 80 km as a block model. We also lead the reader through a comprehensible yet intuitive explanation of the theory and concepts used by the routines. SphGLLTools allows for easy and practical visualization of spectral-element models using either direct interpolation or the flexible expansion using spherical harmonics while taking advantage of GMT6 (Wessel et al., 2019) to create high-quality images. The additional routines in the package that operates on spectral-element meshes in parallel for interpolation, remeshing, and smoothing are useful during the workflow that encompasses creating a new adjoint model and future studies using different resolutions computational resources.

Bibliography

- Affonso, G., Rocha, M., Costa, I., Assumpção, M., Fuck, R., Albuquerque, D., Portner, D., Rodríguez, E., and Beck, S. (2021). Lithospheric Architecture of the Parapanema Block and Adjacent Nuclei Using Multiple-Frequency P-Wave Seismic Tomography. *Journal of Geophysical Research: Solid Earth*, 126(4):e2020JB021183.
- Aki, K., Christoffersson, A., and Husebye, E. S. (1977). Determination of the three-dimensional seismic structure of the lithosphere. *Journal of Geophysical Research*, 82(2):277–296.
- Aki, K. and Lee, W. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model. *Journal of Geophysical research*, 81(23):4381–4399.
- Bartorelli, A., Teixeira, W., and de Brito Neves, B. B. (2020). *Geocronologia e Evolução Tectônica do Continente Sul-Americano: a contribuição de Umberto Giuseppe Cordani*, volume 1. Solaris Edições Culturais.
- Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., and Pugmire, D. (2016). Global adjoint tomography: first-generation model. *Geophysical Journal International*, 207(3):1739–1766.
- Bozdağ, E., Trampert, J., and Tromp, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, 185(2):845–870.
- Cahill, T. and Isacks, B. L. (1992). Seismicity and shape of the subducted Nazca plate. *Journal of Geophysical Research: Solid Earth*, 97(B12):17503–17529.
- Casio Computer, C. L. (2015). Nodes and Weights of Gauss-Lobatto Calculator.
- Celli, N., Lebedev, S., Schaeffer, A., Ravenna, M., and Gaina, C. (2020). The upper mantle beneath the South Atlantic Ocean, South America and Africa from waveform tomography with massive data sets. *Geophysical Journal International*, 221(1):178–204.

- Ciardelli, C. (2015). Modelagem de Ondas Sísmicas por Diferenças Finitas e Elementos Espectrais: Futura Aplicação em Tomografia de Forma de Onda. *Undergraduate thesis, University of São Paulo, unpublished.*
- Cordani, U. G. (1984). Estudo preliminar de integração do Pré-Cambriano com os eventos tectônicos das bacias sedimentares brasileiras. *Bol. Ciênc. Técn. Petról.*, 15:1–70.
- Dahlen, F., Hung, S.-H., and Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes-I. Theory. *Geophysical Journal International*, 141(1):157–174.
- Demanet, L. (2015). 18.325 topics in applied mathematics: Waves and imaging. In *Mathematics–MIT Course No. 18.325*. Cambridge MA. MIT OpenCourseWare.
- Durand, S., Debayle, E., Ricard, Y., Zanolli, C., and Lambotte, S. (2017). Confirmation of a change in the global shear velocity pattern at around 1000 km depth. *Geophysical Journal International*, 211(3):1628–1639.
- Dziewonski, A. M. and Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the earth and planetary interiors*, 25(4):297–356.
- Dziewonski, A. M., Hager, B. H., and O’Connell, R. J. (1977). Large-scale heterogeneities in the lower mantle. *Journal of Geophysical Research*, 82(2):239–255.
- El-Baghdady, G. I., El-Azab, M., and El-Beshbeshy, W. (2015). Legendre-Gauss-Lobatto Pseudo-spectral Method for One-Dimensional Advection-Diffusion Equation. *Eur. J. Oper. Res*, 2(1):29–35.
- Epperson, J. F. (1987). On the Runge example. *The American Mathematical Monthly*, 94(4):329–341.
- Feng, M., Van der Lee, S., and Assumpção, M. (2007). Upper mantle structure of South America from joint inversion of waveforms and fundamental mode group velocities of Rayleigh waves. *Journal of Geophysical Research: Solid Earth*, 112(B4).
- Fichtner, A. (2010). *Full seismic waveform modelling and inversion*. Springer Science & Business Media.

- Finger, N.-P., Kaban, M., Tesauro, M., Haeger, C., Mooney, W., and Thomas, M. (2021). A Thermo-Compositional Model of the Cratonic Lithosphere of South America. *Geochemistry, Geophysics, Geosystems*, 22(4):e2020GC009307.
- French, S. and Romanowicz, B. (2014). Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. *Geophysical Journal International*, 199(3):1303–1327.
- Gibowicz, S. J. (2009). Seismicity induced by mining: Recent research. *Advances in geophysics*, 51:1–53.
- Giles, M. B. and Pierce, N. A. (2000). An introduction to the adjoint approach to design. *Flow, turbulence and combustion*, 65(3):393–415.
- Grand, S. P. (2002). Mantle shear-wave tomography and the fate of subducted slabs. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 360(1800):2475–2491.
- Hasegawa, A. and Sacks, I. S. (1981). Subduction of the Nazca plate beneath Peru as determined from seismic observations. *Journal of Geophysical Research: Solid Earth*, 86(B6):4971–4980.
- Heintz, M., Debayle, E., and Vauchez, A. (2005). Upper mantle structure of the South American continent and neighboring oceans from surface wave tomography. *Tectonophysics*, 406(1-2):115–139.
- Herak, M. (2010). Andrija Mohorovičić: 2010-centennial of discovery of the Moho. *Studia geophysica et geodaetica*, 56(1):293–298.
- Herrin, E., Tucker, W., Taggart, J., Gordon, D. W., and Lobdell, J. L. (1968). Estimation of surface focus P travel times. *Bulletin of the Seismological Society of America*, 58(4):1273–1291.
- Houser, C., Masters, G., Shearer, P., and Laske, G. (2008). Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms. *Geophysical Journal International*, 174(1):195–212.

- Igel, H. (2017). *Computational seismology: a practical introduction*. Oxford University Press.
- James, D. E. (1999). Cenozoic formation of the Central Andes: a geophysical perspective. *Geology and ore deposits of the Central Andes*.
- Jeffreys, H. and E., B. K. (1940). Seismological Tables. *Nature*.
- Jordan, T. H. (1988). Structure and formation of the continental tectosphere. *Journal of Petrology*, (1):11–37.
- Kennett, B. and Engdahl, E. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2):429–465.
- Kennett, B. L., Engdahl, E., and Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1):108–124.
- Koelemeijer, P., Ritsema, J., Deuss, A., and Van Heijst, H.-J. (2016). SP12RTS: a degree-12 model of shear-and compressional-wave velocity for Earth’s mantle. *Geophysical Journal International*, 204(2):1024–1039.
- Komatitsch, D. and Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical journal international*, 139(3):806–822.
- Komatitsch, D. and Tromp, J. (2002a). Spectral-element simulations of global seismic wave propagation-I. Validation. *Geophysical Journal International*, 149(2):390–412.
- Komatitsch, D. and Tromp, J. (2002b). Spectral-Element Simulations of Global Seismic Wave Propagation—II. Three-Dimensional Models, Oceans, Rotation and Self-Gravitation. *Geophysical Journal International*, 150(1):303–318.
- Lehmann, I. (1936). P’. *Publications du Bureau Central Scientifiques*, 14:87–115.
- Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., and Pugmire, D. (2020). Global adjoint tomography-model GLAD-M25. *Geophysical Journal International*, 223(1):1–21.

- Lu, C., Grand, S. P., Lai, H., and Garnero, E. J. (2019). TX2019slab: A new P and S tomography model incorporating subducting slabs. *Journal of Geophysical Research: Solid Earth*, 124(11):11549–11567.
- Luo, Y., Modrak, R., and Tromp, J. (2013). Strategies in adjoint tomography. In *Handbook of Geomathematics: Second Edition*, pages 1943–2001. Springer Berlin Heidelberg.
- Mantovani, M. and Brito Neves, B. d. (2005). The Paranapanema lithospheric block: Its importance for Proterozoic (Rodinia, Gondwana) supercontinent theories. *Gondwana Research*, 8(3):303–315.
- Menke, W. (2012). *Geophysical data analysis: Discrete inverse theory*. Academic press.
- Moczo, P. (1998). *Introduction to Modeling Seismic Wave Propagation by the Finite-Difference Methods*. Disaster Prevention Research Institute, Kyoto University.
- Mohorovičić, A. (1909). *Das Beben vom 8. X*, volume 9. Jahrbuch met. obs. Zagreb.
- Mokashi, P. (2014). What is the difference between strong form and weak form?
- Montelli, R., Nolet, G., Dahlen, F., and Masters, G. (2006). A catalogue of deep mantle plumes: New results from finite-frequency tomography. *Geochemistry, Geophysics, Geosystems*, 7(11).
- Müller, G. (2007). *Theory of elastic waves*. Geoforschungszentrum Potsdam.
- Oldham, R. D. (1906). The constitution of the interior of the Earth, as revealed by earthquakes. *Quarterly Journal of the Geological Society*, 62(1-4):456–475.
- Oyhantçabal, P., Siegesmund, S., and Wemmer, K. (2011). The Río de la Plata Craton: a review of units, boundaries, ages and isotopic signature. *International Journal of Earth Sciences*, 100(2-3):201–220.
- Patera, A. T. (1984). A spectral element method for fluid dynamics: laminar flow in a channel expansion. *Journal of computational Physics*, 54(3):468–488.

- Plesa, A.-C., Bozdog, E., Rivoldini, A., Knapmeyer, M., McLennan, S., Padovan, S., Tosi, N., Breuer, D., Peter, D., Staehler, S., et al. (2020). Seismic Velocities Distribution in a 3D Mantle: Implications for InSight Measurements.
- Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- Press, W. H., William, H., Teukolsky, S. A., Saul, A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press.
- Rapela, C. W., Fanning, C. M., Casquet, C., Pankhurst, R. J., Spalletti, L., Poiré, D., and Baldo, E. G. (2011). The Rio de la Plata craton and the adjoining Pan-African/brasiliano terranes: their origins and incorporation into south-west Gondwana. *Gondwana Research*, 20(4):673–690.
- Ritsema, J., Deuss, a. A., Van Heijst, H., and Woodhouse, J. (2011). S40rts: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltimes and normal-mode splitting function measurements. *Geophysical Journal International*, 184(3):1223–1236.
- Ritsema, J., van Heijst, H. J., and Woodhouse, J. H. (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science*, 286(5446):1925–1928.
- Rocha, M., Assumpção, M., Affonso, G., Azevedo, P., and Bianchi, M. (2019). Teleseismic P wave tomography beneath the Pantanal, Paraná, and Chaco-Paraná Basins, SE South America: Delimiting lithospheric blocks of the SW Gondwana assemblage. *Journal of Geophysical Research: Solid Earth*, 124(7):7120–7137.
- Schaeffer, A. and Lebedev, S. (2013). Global shear speed structure of the upper mantle and transition zone. *Geophysical Journal International*, 194(1):417–449.
- Schimmel, M., Assumpção, M., and VanDecar, J. (2003). Seismic velocity anomalies beneath SE Brazil from P and S wave travel time inversions. *Journal of Geophysical Research: Solid Earth*, 108(B4).

- Schubert, B. (2003). The spectral-element method for seismic wave propagation: theory, implementation and comparison to finite difference methods. *Master's thesis, Ludwig-Maximilians-Universität München, unpublished.*
- Schuster, G. (2017). Seismic inversion: Society of Exploration Geophysicists. *ISBN (Print)*, pages 978–1.
- Sengupta, M. K. and Toksöz, M. N. (1977). Three dimensional model of seismic velocity variation in the Earth's mantle. *Geophysical Research Letters*, 3(2):84–86.
- Simões Neto, F., Julià, J., and Schimmel, M. (2019). Upper-mantle structure of the Borborema Province, NE Brazil, from P-wave tomography: implications for rheology and volcanism. *Geophysical Journal International*, 216(1):231–250.
- Tape, C., Liu, Q., Maggi, A., and Tromp, J. (2009). Adjoint tomography of the southern California crust. *Science*, 325(5943):988–992.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.
- Tassinari, C. C. and Macambira, M. J. (1999). Geochronological provinces of the Amazonian Craton. *Episodes-Newsmagazine of the International Union of Geological Sciences*, 22(3):174–182.
- Van der Lee, S., James, D., and Silver, P. (2001). Upper mantle S velocity structure of central and western South America. *Journal of Geophysical Research: Solid Earth*, 106(B12):30821–30834.
- VanDecar, J. C., James, D. E., and Assumpção, M. (1995). Seismic evidence for a fossil mantle plume beneath South America and implications for plate driving forces. *Nature*, 378(6552):25–31.
- Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W., and Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, 20(11):5556–5564.

Yuan, Y. O., Bozdağ, E., Ciardelli, C., Gao, F., and Simons, F. J. (2020). The exponentiated phase measurement, and objective-function hybridization for adjoint waveform tomography. *Geophysical Journal International*, 221(2):1145–1164.

Appendix A

Book Cover

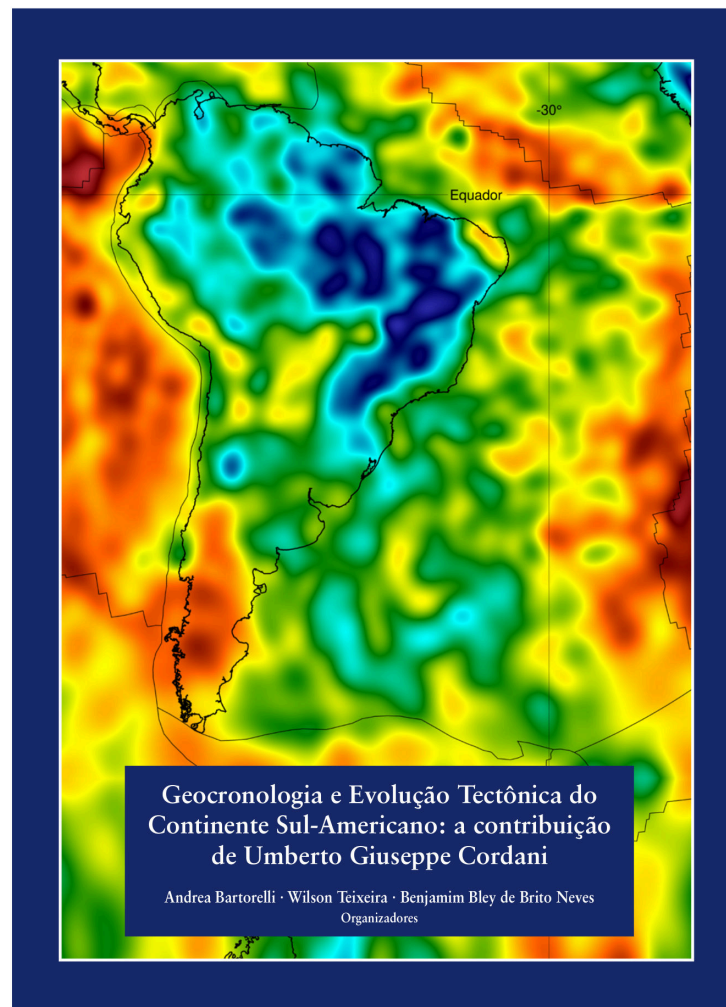


Figure A.1: Book cover for the book *Geocronologia e Evolução Tectônica do Continente Sul-Americano: a contribuição de Umberto Giuseppe Cordani*. (Bartorelli et al., 2020) showing a depth slice of the β_v anomaly for SAAM23 (Chapter 8) at 100 km depth. The bluish colors represent regions with seismic velocities above the average of 4.44 km/s (green), whereas the reddish colors stand for lower-than-average velocity regions. High velocities suggest a thicker lithosphere, such as in the Amazonian Cratons, São Francisco Craton, and other cratonic blocks.