



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2020

Converted wave tomography: Developing a new inversion method for 3-D crustal shear wave velocities, with application to the Central Alps

Colavitti Leonardo

Colavitti Leonardo, 2020, Converted wave tomography: Developing a new inversion method for 3-D crustal shear wave velocities, with application to the Central Alps

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_FE83A10B9F0B0

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté des géosciences
et de l'environnement

**CONVERTED WAVE TOMOGRAPHY:
DEVELOPING A NEW INVERSION METHOD FOR
3-D CRUSTAL SHEAR WAVE VELOCITIES,
WITH APPLICATION TO THE CENTRAL ALPS**

Thesis for the degree of Doctor of Philosophy

presented by

LEONARDO COLAVITTI

Defence on 26 June 2020 in the presence of the jury composed of

Mr. György HETÉNYI	University of Lausanne	Thesis supervisor
Mrs. Christel TIBERI	University of Montpellier	External expert
Mr. Tobias DIEHL	Swiss Seismological Service	External expert
Mr. James IRVING	University of Lausanne	Internal expert
Mr. Jean-Luc EPARD	University of Lausanne	President

Alla mia famiglia/To my family

Abstract

Passive seismological investigations typically image the Earth's crust with direct P-waves or ambient noise correlation yielding S-wave information. While the first method requires local earthquakes to achieve high resolution, in the second method the depth penetration strongly depends on the recording network's aperture.

In this thesis I develop a new inversion method and implement the related software in which teleseismic P-to-S converted waves (receiver functions) are exploited to construct a fully 3-D structural and shear-wave velocity model of the crust. This method does not require local earthquakes, nor a large aperture seismic network, but a dense array of 3-component sensors with a station spacing similar to the expected crustal thickness. This new technique is first applied to the Central Alps, a tectonically complex area where imaging in 3-D is of pivotal interest.

The new method is composed of the following main elements. (1) An accurate ray propagator, which respects Snell's law in 3-D at any interface geometry, and allows P-to-S converted ray-paths to reach the recording station at <0.1 km accuracy. (2) A new model parameterization, with horizontally fixed but vertically flexible-position nodes, and layer-wise two velocity points defined to accommodate mapping both sharp discontinuities and gradients across layers. (3) A stochastic inversion procedure, combining simulated annealing and a pattern search algorithm, to find discontinuity depths and velocities across the crust by fitting grouped converted waves with synthetics. This inversion is performed locally for each point and its neighbours; it covers the entire study area step-wise with an overlap and at least two iterations.

The application to the Central Alps uses 20 years of high-quality data from permanent broad-band stations and from the temporary AlpArray Seismic Network. The initial model includes a Moho depth map and a 3-D P wave velocity model derived from past investigations. The 3-D inversion results at 25 km horizontal resolution provide a series of maps and cross-sections. The crustal thickness generally reflects well the roots of the Alpine orogen, and its jump between the European and Adriatic plates, including the Ivrea Geophysical Body. The lower crustal thickness is less well resolved, yet appears fairly constant. Average crustal V_p/V_s

ratios are relatively higher beneath the orogen. A low- V_p/V_s area in the European foreland correlates with lower crustal earthquakes, which we interpret as mechanical differences in rock properties, most likely inherited.

Our results are generally similar to those found by 3-D ambient noise tomography in the area. The new method inherently performs better at localizing discontinuities, and less well at imaging bulk anomalies. Thanks to sub-vertically propagating rays, our method maps the full crustal structure across the entire area of a seismic network. Future developments can incorporate joint inversions with gravity or other seismological tomography methods.

Résumé

La sismologie passive image la croûte terrestre typiquement par des ondes P directes ou des ondes S basées sur la corrélation du bruit ambiant. Si la première méthode requiert des séismes locaux pour une imagerie haute résolution, la deuxième méthode est limitée dans sa pénétration en profondeur par l'ouverture du réseau enregistrant des signaux.

Dans cette thèse je développe une nouvelle méthode d'inversion et j'implémente le logiciel correspondant pour exploiter des ondes converties P-en-S (fonctions récepteurs) pour construire un modèle 3-D structural et de vitesse d'onde S de la croûte. Cette méthode requiert ni séisme local, ni une grande ouverture du réseau, mais un déploiement dense de capteurs 3-composantes à un espacement comparable à l'épaisseur attendue de la croûte. La première application de cette nouvelle technique se focalise sur les Alpes Centrales, une région tectonique complexe où l'imagerie 3-D est un but important.

La nouvelle méthode se compose des éléments principaux suivants. (1) Un propagateur de rai exact, qui respecte la loi de Snell en 3-D à une géométrie d'interface quelconque, et permet aux rais convertis P-en-S d'arriver à <0.1 km de la station. (2) Un nouveau paramétrage de modèle, avec des nœuds horizontalement fixes mais verticalement flexibles, et deux points de définition des vitesses par couche pour permettre à la fois l'imagerie des discontinuités et celle des gradients dans les couches. (3) Une procédure d'inversion stochastique, combinant recuit simulé et un recherche de motifs, pour trouver la profondeur des discontinuités et des vitesses à travers la croûte en ajustant des synthétiques à des groupes d'ondes converties. Cette inversion est appliquée localement à chaque point et ses voisins, la procédure couvre toute la zone d'étude pas-à-pas avec un recouvrement et au moins deux itérations.

L'application aux Alpes Centrales utilise des données de haute qualité enregistrées sur plus de 20 ans par des stations large-bandes permanentes et par le du réseau temporaire du projet AlpArray. Le modèle initial inclut une carte de profondeur du Moho et un modèle 3-D en vitesse d'onde P d'études précédentes. Le résultat de l'inversion 3-D, à une résolution horizontale de 25 km, inclut une série de cartes et de profils. L'épaisseur de la croûte reflète bien la racine de l'orogène alpin, et le saut entre les plaques européenne et adriatique, y compris le corps d'Ivrée. L'épaisseur de la croûte inférieure est moins bien résolue mais

paraît relativement constante. Le rapport V_p/V_s moyenne de la croûte est relativement plus élevé sous la chaîne. Une zone de V_p/V_s faible dans l'avant-pays européen coïncide avec des séismes dans la croûte inférieure, ce que nous interprétons comme une différence dans les propriétés mécaniques des roches, probablement héritée.

Nos résultats sont généralement similaires à ceux trouvés par tomographie 3-D du bruit ambiant dans la région. La nouvelle méthode est plus performante à localiser des discontinuités, et moins bien pour l'imagerie des anomalies volumétriques. Grâce aux rais sub-verticaux, notre méthode image la structure de toute la croûte sous l'ensemble du réseau sismologique. Des développements futurs peuvent inclure des inversions conjointes avec la gravimétrie ou d'autres types de tomographies sismologiques.

Riassunto

In genere le indagini sismologiche di tipo passivo permettono di ottenere una immagine della crosta terrestre attraverso l'arrivo delle onde P dirette oppure mediante la correlazione del rumore ambientale da onde S. Mentre il primo metodo richiede l'utilizzo di terremoti locali per garantire un'alta risoluzione, nel secondo metodo la profondità di indagine dipende fortemente da quella che è l'apertura della rete sismica.

In questa tesi ho sviluppato un nuovo metodo di inversione ed il software associato dove le onde convertite da P ad S (funzione ricevitore) registrate da telesismi vengono utilizzate per la costruzione di un modello tridimensionale della crosta e delle onde di taglio. Questo metodo non richiede né l'utilizzo di terremoti locali, né una grande apertura di rete sismica, ma un denso sistema di sensori sismici a 3 componenti con uno spazio simile a quello dello spessore crostale atteso. Questa tecnica innovativa è stata applicata in primo luogo alle Alpi Centrali, un'area complessa dal punto di vista tettonico dove l'immagine 3-D del sottosuolo è di fondamentale interesse.

Il nuovo metodo si compone dei seguenti elementi principali. (1) Un propagatore di raggi accurato, che rispetta la legge di Snell in 3-D per ciascuna geometria di interfaccia, e permette ai percorsi dei raggi di onde convertite da P ad S di raggiungere la stazione corrispondente con una precisione < 0.1 km. (2) Una nuova parametrizzazione del modello, con una posizione dei nodi fissata orizzontalmente ma variabile verticalmente, e 2 valori di velocità definite per ciascun strato mappando sia discontinuità nette sia gradienti di velocità attraverso gli strati. (3) Una procedura di inversione di tipo stocastico, che coniuga un processo di ricottura simulata ed un algoritmo di ricerca per affinare la soluzione, allo scopo di scoprire le profondità delle discontinuità e le velocità della crosta attraverso una corrispondenza tra i dati osservati delle onde convertite con quelli sintetici. L'inversione viene eseguita per ciascun nodo del modello e per i suoi vicini; in questo modo l'area di studio viene via via coperta con una sovrapposizione di almeno due iterazioni.

L'applicazione alle Alpi Centrali utilizza 20 anni di dati sismici ad alta qualità ottenuta sia da stazioni sismiche permanenti a banda larga sia da stazioni temporanee relative al progetto AlpArray. Il modello iniziale comprende una mappa della Moho ed un modello di velocità 3-D

di onde P ottenuto da precedenti studi. I risultati della nostra inversione 3-D con risoluzione orizzontale di 25 km forniscono una serie di mappe e di profili verticali. In generale lo spessore crostale riflette bene quelle che sono le radici della catena alpina, ed i relativi balzi in profondità della placca Europea e di quella Adriatica, compreso il corpo di Ivrea. La struttura della crosta inferiore non è così ben definita, anche se sembra essere abbastanza costante. Il valore V_p/V_s mediato sull'intera crosta risulta essere relativamente più alto al di sotto dell'arco alpino. Una zona con valori di V_p/V_s bassi nell'avampaese europeo coincide con i terremoti avvenuti nella crosta profonda: ciò può essere interpretato come una differenza del comportamento meccanico nelle proprietà delle rocce, molto probabilmente eredita da processi preesistenti.

I nostri risultati sono in generale simili a quelli trovati da un altro studio 3-D di tomografia sismica da rumore ambientale condotto nell'area in esame. Per sua stessa natura, il nuovo metodo è più sensibile all'individuazione delle discontinuità, mentre risolve meno bene le anomalie di velocità. Grazie ai raggi pressoché sub-verticali, la nostra tecnica ci consente di mappare l'intera struttura della crosta lungo tutta l'area interessata dalla rete sismica. Sviluppi futuri potranno incorporare metodi di inversione combinando studi di gravità o altre tecniche di tomografia sismica.

Acknowledgements

I wanted to thank first the supervisor of this thesis Prof. György Hetényi for all the help, patience and the collaboration that has given to me during these 4 years. György always supported me from the first working day to the last page of this manuscript. In addition to being a big researcher, he is also a great person and I will certainly never forget him.

I am very grateful to the Swiss Seismological Service, Istituto Nazionale di Geofisica e Vulcanologia in Italy, University of Genova, Civil Defense of Bolzano, Mediterranean Network, Réseau sismologique et géodésique français, Zentralanstalt für Meteorologie und Geodynamik in Austria, University of Munich, German Research Centre for Geosciences and AlpArray project for collecting and sharing seismological data.

For the inversion part (Chapter 5), I say thanks to Prof. Niklas Linde (Institute of Earth Sciences in Lausanne, Switzerland) for his teachings regarding the uncertainty assessment in the geosciences, to Dr. Thomas Bodin (Laboratoire de Géologie de Lyon, France) and Prof. Malcom Sambridge (Australian National University) for nice discussions and several email exchanges. Another thanks also goes to Prof. Martin Knapmeyer (German Aerospace Center) for making available a 1-D ray tracing tool that I could implement (Chapter 3), Dr. Matteo Spada (Paul Scherrer Institute, Switzerland) and Dr. T. Diehl (Swiss Seimological Service, Zürich) for providing the initial models of Moho and P-velocity model of the Alps (Chapter 4). I don't want to forget scientific discussions at the Institute of Earth Sciences in Lausanne with Prof. Jean-Luc Epard (Chapter 1) regarding Alpine geology and Prof. Stefan Schmalholz for results interpretation (Chapter 7). Another thanks to Dr. Denis Lombardi and to Dr. Yang Lu (University of Southern California), who kindly shared their data for a comparison and to AlpArray Receiver Function Working Group for the fruitful advice during this research.

A special thanks goes to all my past and current colleagues of the OROG3NY Group, Dr. Kristel Chanard, Matteo Scarponi, Celso Alvizuri, Shiba Subedi and Dr. Konstantinos Michailos, who actively contributed both scientifically and from a human point of view to all beautiful moments during these years.

For financial support, I would like to thank the Swiss National Science Foundation for

funding this project (grant n. PP00P2_157627) and the Société Académique Vaudoise.

Moreover, I say thank you to the wonderful people I met in Lausanne, Giulia, Francesco, Daniele and Matteo and to long-time friends Luca, Marco, Andrea, Stefano, Ilik, Alessandro, Davide and Massimo for the unforgettable moments we have shared together.

Last but not least, a big thank you to my family for the priceless moral supports and continuous encouragements and a thought to the people in Italy and worldwide in this particular period we are living: let's keep the distance today to hug each other tomorrow!

*“Philosophers are people who know less and less about more and more,
until they know nothing about everything.*

*Scientists are people who know more and more about less and less,
until they know everything about nothing”.*

Konrad Lorenz

Contents

1	Geological framework of the Alps	18
1.1	Geological evolution	19
1.2	Tectonic framework and geological division	21
1.2.1	The Austroalpine thrust units	22
1.2.2	The Penninic zone	23
1.2.3	The Helvetic-Dauphinois zone	23
1.2.4	The Southern Alps	23
1.3	Geophysical studies in the Alps	24
1.3.1	Active geophysical investigations	24
1.3.2	Moho depth map	26
1.3.3	Receiver function studies	28
1.3.4	Local earthquake tomography	30
1.3.5	Ambient noise tomography	32
1.3.6	Gravity studies	35
2	Receiver function method and dataset	38
2.1	Fundamentals of receiver functions	39
2.1.1	RFs as a convolutional model	40
2.1.2	Spectral deconvolution (frequency domain)	41
2.1.3	Iterative deconvolution (time domain)	41
2.1.4	Example of a seismic record and a RF computation	42
2.2	Database construction	45
2.3	Quality Control	47
2.3.1	QC on ZNE components	47
2.3.2	QC on RF	49
2.4	H- κ Method	51
2.4.1	Basis of the Zhu-Kanamori technique	51
2.4.2	Preliminary results in the Central Alps	52

<i>CONTENTS</i>	12
2.4.3 Comparison with Lombardi et al., 2008	55
3 3-D ray tracing in heterogeneous media	59
3.1 Shooting schemes	60
3.1.1 Shooting method	60
3.1.2 Bending method	61
3.1.3 Grid based schemes	62
3.1.3.1 Fast Marching Method	63
3.1.3.2 Shortest path ray tracing	64
3.2 Ray Tracing Implementation in this work	66
3.2.1 P shooting in a global velocity model	67
3.2.1.1 TTBOX accuracy	68
3.2.2 Local dip of Moho	69
3.2.3 P-to-S piercing points detection	72
3.2.3.1 Quality Control on the piercing points	73
3.2.4 S shooting in a local velocity model	74
3.2.5 Shooting correction	75
3.2.6 Performances and shooting results	79
4 Model Parameterization	84
4.1 Model parameterization in classical tomography	85
4.2 Model parameterization for 3-D converted wave	86
4.2.1 Ray tracing accuracy using new model parameterization	87
4.3 New model parameterization applied to the Alps	88
4.3.1 Setup of the initial model geometry	89
4.3.1.1 Conrad discontinuity in the Alps	89
4.3.2 Setup of the initial velocity model	92
4.4 Extraction of 3-D velocity model	95
4.5 Computation of synthetic RF	96
4.5.1 The Gaussian convolution	98
4.6 Receiver function sensitivity on inversion parameters	100
4.6.1 Thicknesses	100
4.6.2 Velocities	102
5 Inversion strategy and setup	105
5.1 Inversion: optimization method	106
5.1.1 Deterministic global optimization	109

5.1.2	Stochastic optimization	112
5.1.2.1	Simulated annealing	113
5.1.2.2	Genetic algorithms	113
5.1.2.3	Neighbourhood Algorithm	114
5.1.3	Deterministic vs stochastic approaches	115
5.1.4	Example of nonlinear optimization for an analytical function	116
5.1.4.1	Optimization using Gradient Solver	117
5.1.4.2	MultiStart Optimization	118
5.1.4.3	Optimization using Genetic Algorithms	119
5.1.4.4	Optimization using Simulated Annealing	120
5.1.4.5	Optimization using Pattern Search	120
5.1.4.6	Choice of the Solver	121
5.2	Setup for 3-D converted wave tomography	122
5.2.1	Spatial separation of the inverse problem	122
5.2.2	Inversion at a node	124
5.2.2.1	Repeatability test on synthetic model	125
5.2.2.2	Adding Pattern Search	127
5.2.2.3	Pattern Search and its polling description	129
5.2.2.4	Number of iterations	132
5.2.3	Spatial grouping of rays	133
5.2.3.1	Bundle division	134
5.2.3.2	Threshold on number of rays per bundle	136
5.2.4	3-D Synthetic inversion recovery test	138
5.2.4.1	Waveform matches	138
6	Inversion results and representation	142
6.1	Overview of inversions and final model selection	143
6.1.1	Choice of the number of variables to invert	143
6.1.2	Inversion parameter and runs	145
6.1.3	Inversion rounds and misfit trends	147
6.1.4	Waveform fits	148
6.1.5	Choice of the final model	150
6.2	Inversion results representation	152
6.2.1	Inversion output	152
6.2.1.1	Interface depth representation	152
6.2.1.2	Velocities' representation	153

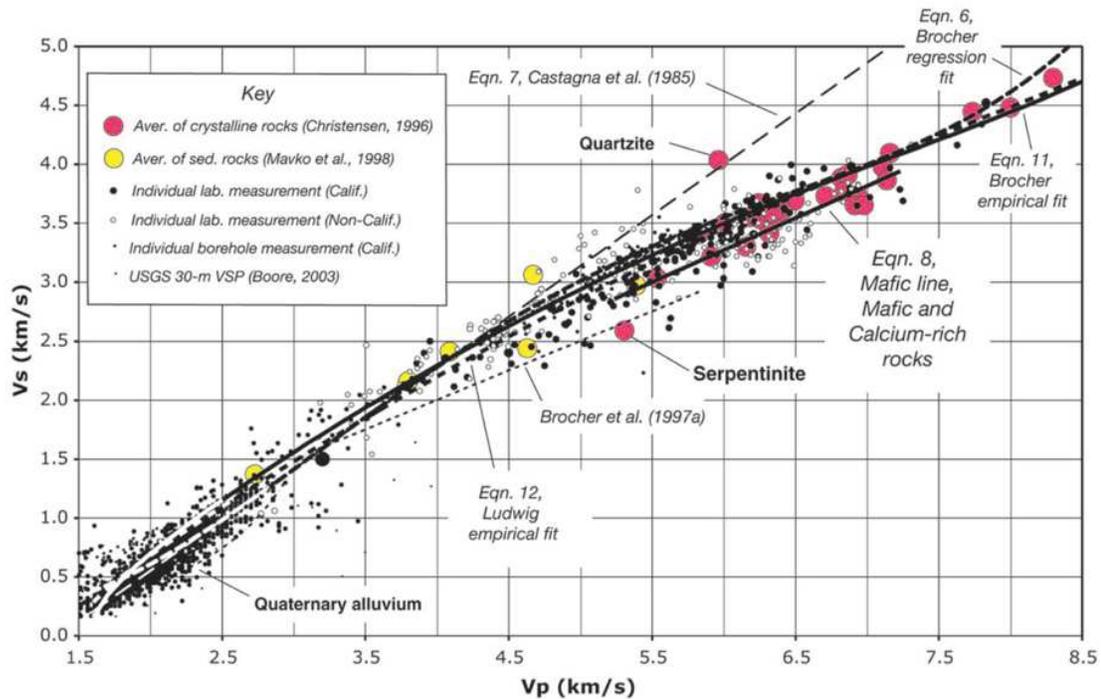
6.2.1.3	Interpolation of unresolved nodes	157
6.2.1.4	Results quality control by node	159
6.2.1.5	Velocity profile representation	161
7	3-D interpretation of the Central Alps	163
7.1	3-D P-to-S converted wave tomography results	164
7.1.1	Crustal interface structure	164
7.1.1.1	Conrad discontinuity depth	164
7.1.1.2	Moho discontinuity depth	165
7.1.2	Velocity structure	168
7.1.2.1	Absolute shear-wave velocities	168
7.1.2.2	Average crustal V_p/V_s	170
7.1.2.3	Velocity jump across the Conrad discontinuity and gradients inside the crustal layers	171
7.2	3-D V_s comparison with Ambient Noise Tomography and geological interpretations	173
7.2.1	Map of shown profiles	173
7.2.2	Cross-section A-A': ECORS-CROP	175
7.2.3	Cross-section B-B': Jura mountains-Po plain	177
7.2.4	Cross-section C-C': NFP-20 West	179
7.2.5	Cross-section D-D': Vosges-West Po basin	181
7.2.6	Cross-section E-E': Basel-Chiasso	183
7.2.7	Cross-section F-F': European GeoTraverse	185
7.2.8	Cross-section G-G': TRANSALP	187
7.3	3-D to 1-D crustal Moho and V_p/V_s comparison	190
7.4	Lower crustal seismicity and geodynamic interpretation	191
7.4.1	Lower crustal earthquakes in the northern Alpine foreland	191
7.4.2	Geodynamic considerations	193
8	Conclusions and future perspectives	197
8.1	Results summary	198
8.2	Uncertainties analysis and the method's applicability	201
8.3	Method limitations and future perspectives	203

Motivation

The goal of this work is to develop a new seismological tool to construct a fully 3-D shear-wave velocity model of the Earth's crust. The approach I implement is an inversion method based on seismic waves from remote natural earthquakes that convert from compressional (P) to shear (S) waves at the base of the crust. In this introductory part, I focus on the importance of a consistent 3-D model of the Earth's interior, on the role of S-waves in geophysics, and why this new technique is first applied on the Central Alps.

Three-dimensional models are very important and significant in all branches of Earth Sciences because geological structures can be very complicated, especially in subduction zones or in orogenic contexts. While it is needless to say that such structures are 3-D in reality, geoscientists are often constrained to build 1-D or 2-D models due to the lack of sufficient information regarding the Earth interior exploration, or decide do to so to simplify the understanding of a system. Yet, scientists have been building 3-D geological models of the Earth for over a century: the first models were conceived as tools to analyze and communicate the geometric shapes of subsurface geological features (Barringer, 1892; Cadman, 1927). At the same time, physical models were still being designed at the dawn of the digital era (Anstey, 1976). Once digital 3-D geological modelling took over from its physical predecessors, their usefulness expanded from the challenging task of 3-D geological visualization (Tipper, 1976), providing inputs for geophysical inversion schemes as well (Cordell and Henderson, 1968). Therefore, to feed such models and inversions with field data, whether they are collected in 1-D, 2-D, 3-D or 4-D manners, the need to step away from simplified representations of the Earth's interior and to build realistic 3-D spatial models increased. This was followed by improvements in 3-D modelling software and a reduction in their complexity and cost, so that this technology became widely accessible to the Earth science community. Following from the field of exploration geophysics, 3-D models are also more and more common in academic research, and, when sufficient data are available, 3-D imaging is the best approach to pursue.

Focusing on **shear (S) waves** is the other significant point of great importance in this work. Shear-waves are often put in the second line after P waves as they are most difficult to detect or pick on a seismogram or trace, they do not travel in liquids, are more difficult to generate in active surveys, and may even split due to material birefringence in anisotropic media which further complicates their use. Nevertheless, because S waves depend only on one of the Lamé parameters (the shear modulus μ) and P-waves depend on both Lamé constant (on μ and also on λ), the shear-waves carry complementary information with respect to P-wave alone and therefore give us a better idea of lithologies in a study area. Shear-waves exhibit different sensitivity to fluid content, partial melt and anisotropy, which can be connected to various geological processes occurring in sedimentary basins and in the lithosphere. The correlation between seismic waves' velocity and density have commonly been measured in the laboratory on real rock samples, which allows to infer the subsurface geometry of geological units and the bulk physical and chemical properties of the Earth's crust (Christensen and Mooney, 1995; Christensen, 1996; Brocher, 2005). Both structures and properties play a role in fundamental and applied research, ranging from geodynamics to earthquake hazard assessment (site effects) and geothermal exploration. The links between P wave velocities, S wave velocities and densities continue to be investigated, as they reveal interesting and useful correlations with lithology, such as the example shown on the figure below between V_p and V_s (Brocher, 2005). Shear-waves have already been used in exploration geophysics (Tatham and Stewart, 1993) and also in academic research (see section 1.3.4 and 1.3.5), nevertheless I consider they are underused compared to their potential and that therefore, further efforts to exploit S-waves should be done.



V_s as a function of V_p for common lithologies (from Brocher, 2005). Solid circles represent samples from Northern and Central California; unfilled circles represent other samples. Largest sizes represent averages of individual laboratory measurements, intermediate sizes show individual laboratory measurements, small sizes correspond to individual borehole measurements, and the smallest sizes correspond to USGS 30-m V_s Ps.

The application of my newly developed approach in the **Central Alps** is both a challenge and comes in a suitable context. The challenge is that the Alpine domain constitutes a very complex zone from a geological and structural point of view, and the imaging of dipping structures at depth is not straightforward. Numerous active and passive geophysical studies have been conducted in the past to explore the Alps. The currently most common techniques to achieve 3-D models are local earthquake tomography (LET), which requires local sources that are well distributed as well as a robust S-wave picker, and the ambient noise tomography (ANT), which requires a large aperture seismic network. Further details and examples are described in Chapter 1. While the P-wave velocity model from LET will serve as an important base for this study, here I aim at developing a 3-D S-wave velocity model across the entire depth range of the crust beneath a seismic array, which is the suitable context of this thesis. Beyond 20 years of high-quality data from permanent seismic broad-band stations, the AlpArray Seismic Network (Hetényi et al., 2018b) provides the needed coverage over the entire study area to implement and test the new inversion approach. This method differs from ANT and LET in that the exploited ray paths are closer to vertical and allow mapping the full crustal thickness beneath the array, without using local earthquakes, but simply a sufficiently dense array of 3-component seismometers.

Chapter 1

Geological framework of the Alps

The geology of the Alps has been studied for over 200 years, during which a large amount of knowledge and information became available including geological maps, cross-sections and geophysical investigations. All these studies provide a unique natural laboratory to improve our understanding of orogenesis and its relationship to present and past crust dynamics.

In this chapter I summarize the evolution of the Alps, focusing in particular on the genesis of the orogen, then on the structural framework and finally describing some of the existing geophysical studies. Regarding the geophysical investigations, I emphasize those conducted with passive methods, including receiver function studies and tomographic investigations since the main purpose of this work is to develop a new method to build a 3-D shear-wave velocity model of the crust with converted wave tomography.

1.1 Geological evolution

The Alps represent a collisional belt generated by the Cretaceous to present convergence of the Adriatic continental plate and a subducting plate including the Mesozoic Ocean and the European passive continental margin. The Alpine orogeny began in the eastern Austroalpine domain and involved the entire Tethys Ocean¹, gradually progressing from internal to external domains (Festa et al., 2019).

In the period between Late Triassic and Early Jurassic (200-175 Ma), the Tethys Ocean started to form between North America and Eurasia from one side and Africa and South America from the other side (Figure 1.1.1). The oceanic crust formed in this process the Piemont-Liguria Ocean, which is a part of the Tethys Ocean. At that time, a promontory of continental crust of the African plate called Adria plate was involved in subdividing the Tethys and in early Alps formation.

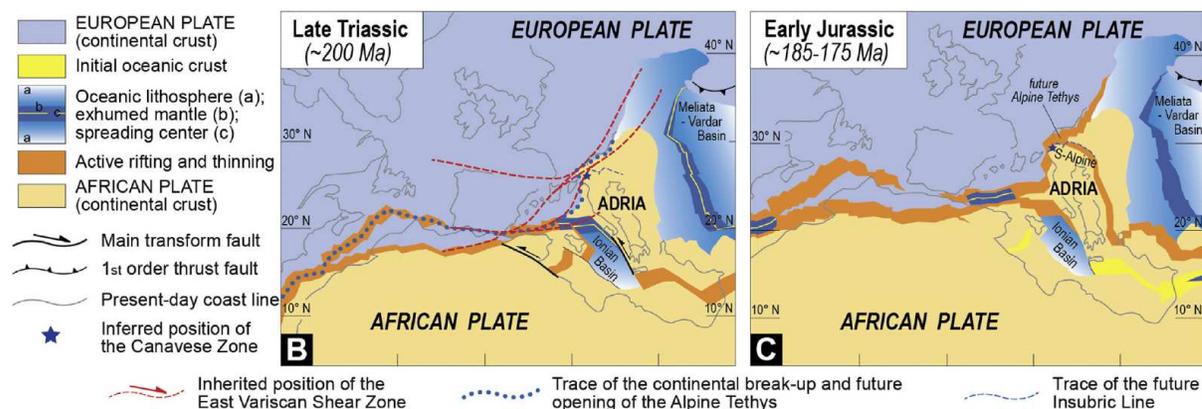


Figure 1.1.1: (B) Onset of the continental break-up (dotted blue lines) at Late Triassic, showing its close relation with the strike-slip faults inherited from the Early Permian tectonics (dashed red lines). (C) Location of the active rifting of the future Alpine Tethys and continental crust thinning at Early Jurassic (images modified from Stampfli et al., 2002; Schettino and Turco, 2011).

When the Atlantic Ocean formed between Africa and South America (about 140 Ma), Africa started moving northeast. During the Cretaceous, a first continental collision took place as the northern part of the Adriatic plate collided with Europe in a phase that is known as *Eo-Alpine phase*. The Alpine orogen formed during a long process of convergence and collision between the European and Adria continental domains during Mesozoic and Cenozoic times (Le Pichon et al., 1988; Dewey et al., 1989; Escher and Beaumont, 1997) and is represented in Figure 1.1.2.

¹The Tethys Ocean was an ocean originated during the Mesozoic Era between the ancient continents of Gondwana and Laurasia, before the opening of the Atlantic Ocean, during the Cretaceous Period.

When the Piemont-Liguria Ocean had completely subducted beneath the Adriatic plate in the Paleocene, the Briançonnais microcontinent, arrived at the subduction zone. This material later became the Penninic *nappes*², but a large part of the Briançonnais terrains subducted into the mantle and were lost. Meanwhile, the upper crust of the Adriatic plate was thrust over the European crust and this is known as the main collisional phase in the Alps formation.

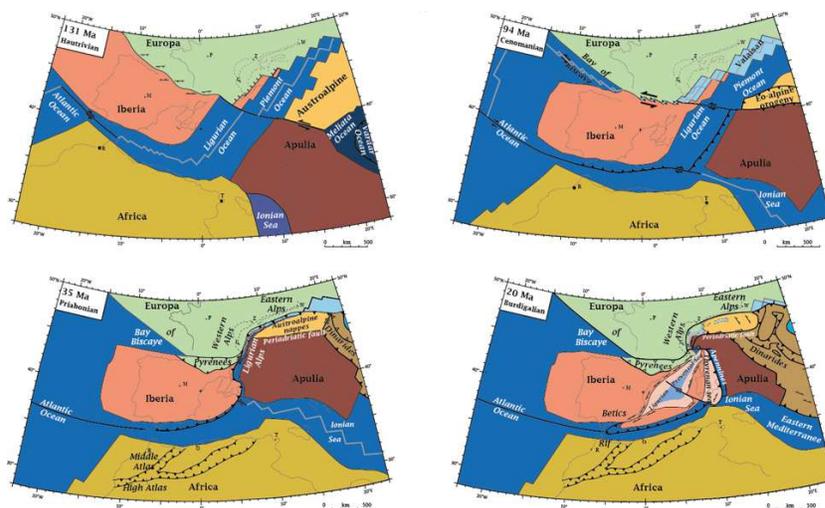


Figure 1.1.2: Paleogeographic reconstruction of the Western Mediterranean zone at 131 Ma, 94 Ma, 35 Ma and 20 Ma (from Bousquet et al., 2012b).

When the subducting slab broke off, the crust started moving up and led to the uplift of the thickened continental crust that evolved to extension. For the Alps, this extension could only occur in a West-East direction because the Adriatic plate was still converging from the South and a huge thrust zone evolved that would later become the *Periadriatic Line*³. With the exception of Austroalpine material, this thrust evolved at the boundary of the Adriatic and European plates while the central zones of the Alps rose and were subsequently eroded, with the formation of the High Tauern window (for details see Stampfli and Borel, 2004).

²In geology, a *nappe* is a large body of rock that has been moved more than 2 or 5 km above a thrust fault from its original position. These formations are characteristic in compressional tectonic settings like continental collision zones or on the overriding plate in subduction zones.

³The Adriatic plate is thought to still move independently of the Eurasian Plate in NNE direction with a small component of counter-clockwise rotation (Devoti et al., 2002). The fault zone that separates the Adriatic and Eurasian plate is the Periadriatic Line, that runs through the Alps.

1.2 Tectonic framework and geological division

From a geographic point of view, the Alps span from the Gulf of Genoa to Vienna, through the French-Italian Western Alpine arc and the East-West-oriented Central and Eastern Alps. At the South of Genoa, the Alpine range disappears, because it collapsed and was fragmented during the Late Neogene opening of the Tyrrenian basin, even if some segments of the Alpine belt are preserved in Corsica and Calabria (Dal Piaz et al., 2003). For more details about regional investigations for each country, see Michard et al. (1996) and Dal Piaz (1999) for the French-Italian Alps, Trumpy (1980) for Switzerland, Plöching (1995) and Neubauer (2000) for Austria, Castellarin et al. (1992) for Southern Alps.

From a geological point of view, the Alps are often divided into Eastern, Central and Western, even if these subdivisions are rather arbitrary (Figure 1.2.1). The division between the Eastern and Central Alps is approximately the line between St. Margrethen, Chur and Sondrio, while the division between the Central and Western Alps remains still unclear (Pfiffner, 2014).

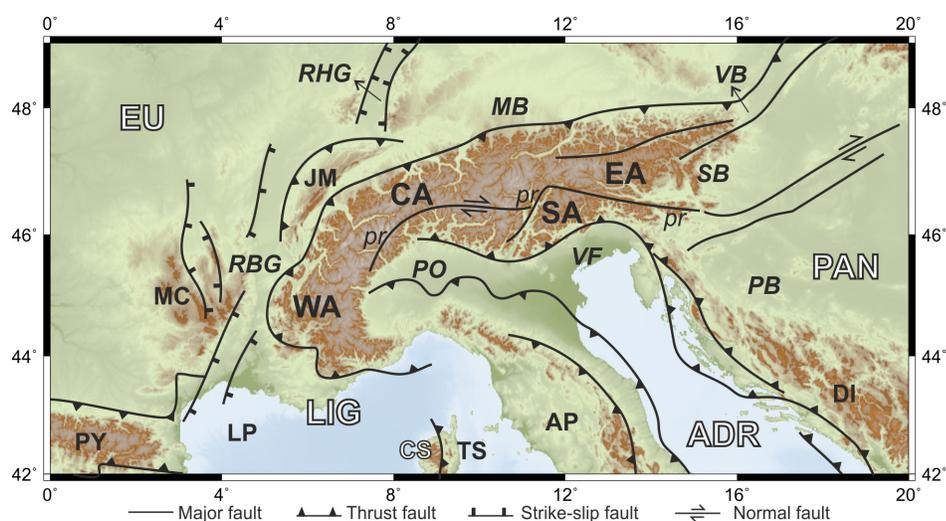


Figure 1.2.1: Simplified tectonic map of the Alps and surrounding areas showing the major tectonic features and faults, modified from Schmid et al. (2004) and Bousquet et al. (2012a). Geographic subdivision of the Alps: WA – Western Alps, CA – Central Alps, EA – Eastern Alps, SA – Southern Alps. Tectonic plates/blocks: EU – Eurasia, ADR – Adria, LIG – Liguria, PAN – Pannonia, CS – Corso-Sardinia. Basins: LP – Liguro-Provençal, TS – Tyrrenian, PO – Po, VF – Venetian-Friuli, MB – Molasse, PB – Pannonian, SB – Styrian, VB – Vienna. Graben: RBG – Rhône-Bresse, RHG – Rhine. Mountains and mountain belts: AP – Apennines, PY – Pyrenees, DI – Dinarides, JM – Jura, MC – Massif Central. Label *pr* identifies the Periadriatic Line.

In Figure 1.2.2 I show a tectonic map of the Central Alps, which is the focus area of our investigation with seismological tools to map the 3-D structure of the crust.

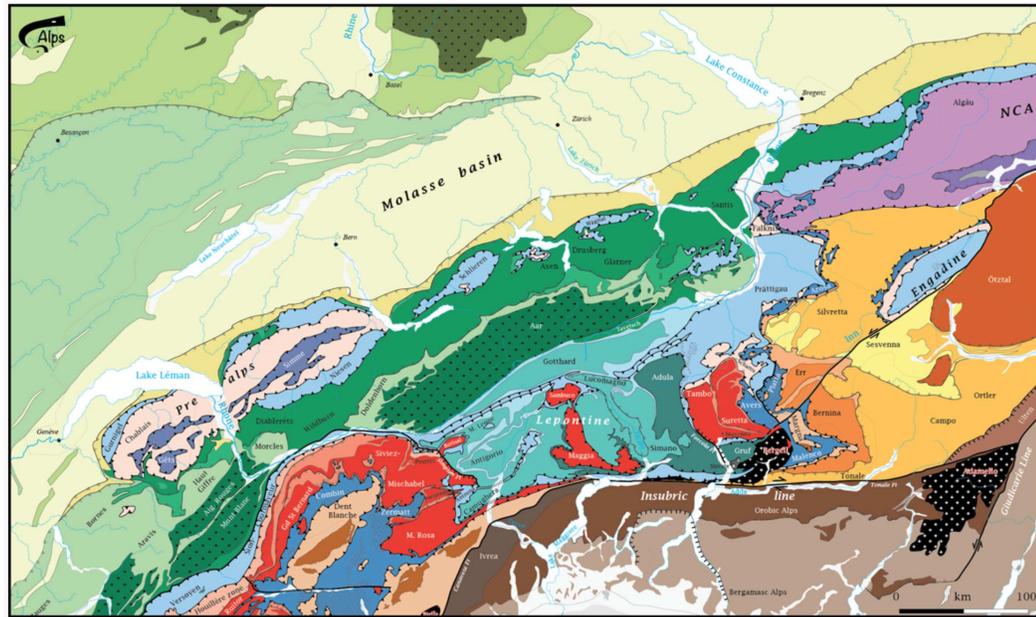


Figure 1.2.2: Tectonic map of the Central Alps. For the color legend of different units, see Bousquet et al. (2012b).

The main strike-slip Fault in the Alps is the Periadriatic Line that runs through the Alps from East to West. This represents the boundary between materials from the former European and Adriatic plates; at the South of this line we can find folded and thrust units of the Southern Alps. North of the Periadriatic Line, rocks from three main paleogeographic “domains” are found: the Helvetic or Dauphinois, the Penninic and the Austroalpine domains.

This subdivision is made according to rocks origin’s: the Helvetic Zone contains material from the European plate, the Austroalpine Zone material from the Adriatic plate, the Penninic Zone material from the domains that existed in between the two plates.

1.2.1 The Austroalpine thrust units

The Austroalpine is subdivided into two sectors, based on contrasting distribution, structural position and main deformation age: western and eastern. The western Austroalpine consists of the Sesia-Lanzo zone where the units are partly interleaved with the ophiolitic Piedmont zone, which as mentioned above, represents the major remnant of the Mesozoic Ocean. The eastern Austroalpine is a thick pile of cover and basement nappes which goes from the Swiss-Austrian border to the Pannonian Basin.

1.2.2 The Penninic zone

The Penninic is the classical name used to group the continental and oceanic nappes issued from the distal European continental margin and the Mesozoic Ocean, belonging to the subducting lower plate. The Penninic zone contains high-grade metamorphic rocks of different paleogeographic origins; these lithologies were deposited as sediments on the crust that existed between the European and Apulian plates before the Alps were formed.

1.2.3 The Helvetic-Dauphinois zone

The Helvetic and Dauphinois zone in the French part consists of crystalline duplexes⁴, sedimentary cover units, and nappes. This domain was strongly deformed from the Late Oligocene onwards, when the orogeny propagated onto the proximal European margin. In this specific zone, rift faults were largely reactivated and inverted. Basement and cover units were accreted in front of the exhumed Austroalpine-Penninic collisional wedge.

1.2.4 The Southern Alps

The Southern Alps are the typical example of a deformed passive continental margin in a mountain range (Bertotti et al., 1993). From the Neogene, the South Alpine thrust-and-fold belt developed and progressively propagated towards the Adriatic foreland, mainly reactivating Mesozoic extensional faults (Castellarin et al., 1992). The crystalline basement includes various kinds of Variscan metamorphic rocks derived from sedimentary and igneous protoliths, later intruded by igneous bodies of Permian age, among them there the famous Ivrea gabbro batholith, which was emplaced at the base of a gneissic crust.

⁴In structural geology, a *duplex* is a system of thrusts that branch off from a single fault below and merge with a thrust fault above. Duplexes form stacks of thrust-bounded rock bodies, which are bounded by roof and floor thrusts.

1.3 Geophysical studies in the Alps

The final goal of my work is to develop a new method where teleseismic P-to-S converted waves are used to map and construct a fully 3-D shear-wave velocity model of the crust. The method differs from ambient noise tomography and local earthquake tomography in its ray-paths being closer to vertical. In this section, I summarize the most important investigations performed in the Alps in recent years, considering both active and passive seismic studies (RFs, tomographic investigations), Moho map information and gravity studies, taking into account that some of these contributions will be used as the input of the initial model.

1.3.1 Active geophysical investigations

Most of our knowledge of the Alpine crust comes from active studies, through several controlled-source seismic campaigns (*CSS*). For a general overview of these active investigations, I refer to the work of Prodehl and Mooney (2012). The earliest seismic observations were performed by Reich (1952) and a first campaign was done in the zone of the Western Alps by Coron et al. (1963), followed in the seventies by the ALP75 (Miller et al., 1976).

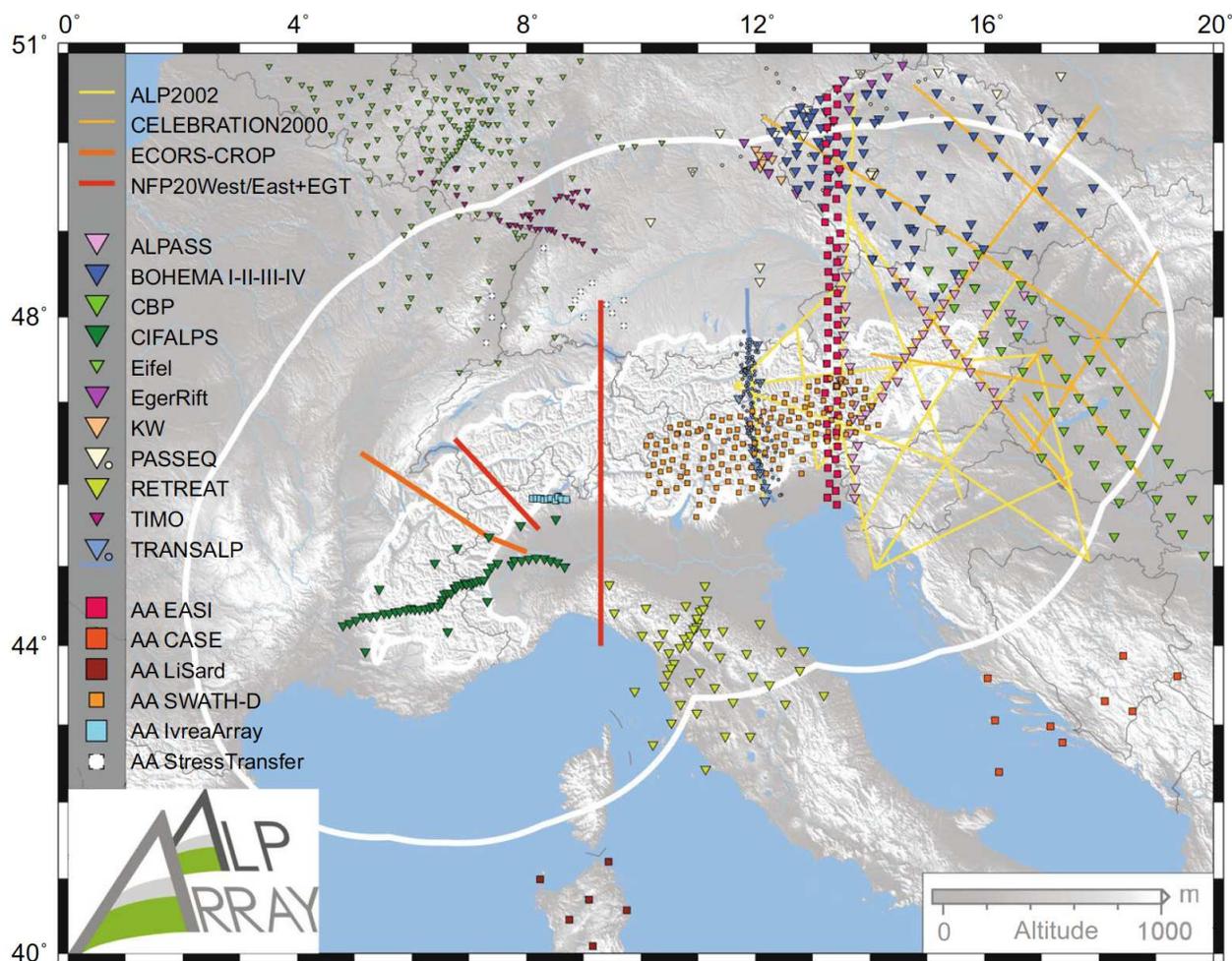


Figure 1.3.1: Location of some temporary seismicological projects in the Alpine area (from Hetényi et al., 2018a). Active seismic campaigns are shown as solid lines. Past passive seismicological networks are shown as symbols (broadband: large triangles, short period: small circles, mixed: small triangles). AlpArray complementary experiments are shown as squares.

Thanks to international collaboration, a lot of seismic studies were performed the last thirty years like the the *French-Italian Étude Continentale et Oceanique par Reflexion et Refraction Simique* or *Progetto Strategico Crosta Profonda ECORS-CROP* (Nicolas et al., 1990; Thouvenot et al., 1990), the *European Geotraverse EGT* (Blundell et al., 1992), the *Swiss National Research Project 20 Pfiffner et al. (1997)*, the *ALP2002* (Brückl and Bleibinhaus, 2003) and the *East-Alpine Reflection Seismic Traverse TRANSALP* (Kummerow et al., 2004). Some of these investigations are listed in Figure 1.3.1.

1.3.2 Moho depth map

The Mohorovičić discontinuity, named after the pioneering Croatian seismologist Andrija Mohorovičić and commonly referred as the Moho, represent the boundary between the Earth's crust and the mantle. It is defined by the distinct change in velocity of seismological waves as they pass through changing densities of rock (Rudnick and Gao, 2003)⁵.

The depth to the crust-mantle boundary in the Alpine region has been estimated by different seismic methods, which leads to a large set of published Moho map models in the Alpine region (e.g. Waldhauser et al., 1998; Stehly et al., 2009; Di Stefano et al., 2011; Spada et al., 2013; Molinari et al., 2015). Figure 1.3.2 shows a comprehensive compilation definition of the Moho depth with information about the uncertainties from the Alpine-Mediterranean region from the work of Spada et al. (2013), who proposed to combine in a quantitative way Controlled Source Seismology (CSS), Local Earthquake Tomography (LET) and Receiver Functions (RF) data.

⁵The Mohorovičić discontinuity was first identified in 1909 by Andria Mohorovičić, when he observed that seismograms from shallow-focus earthquakes had 2 sets of P and S waves, one that followed a direct path near the Earth's surface and the other refracted by a high-velocity medium (McLeish, 1992).

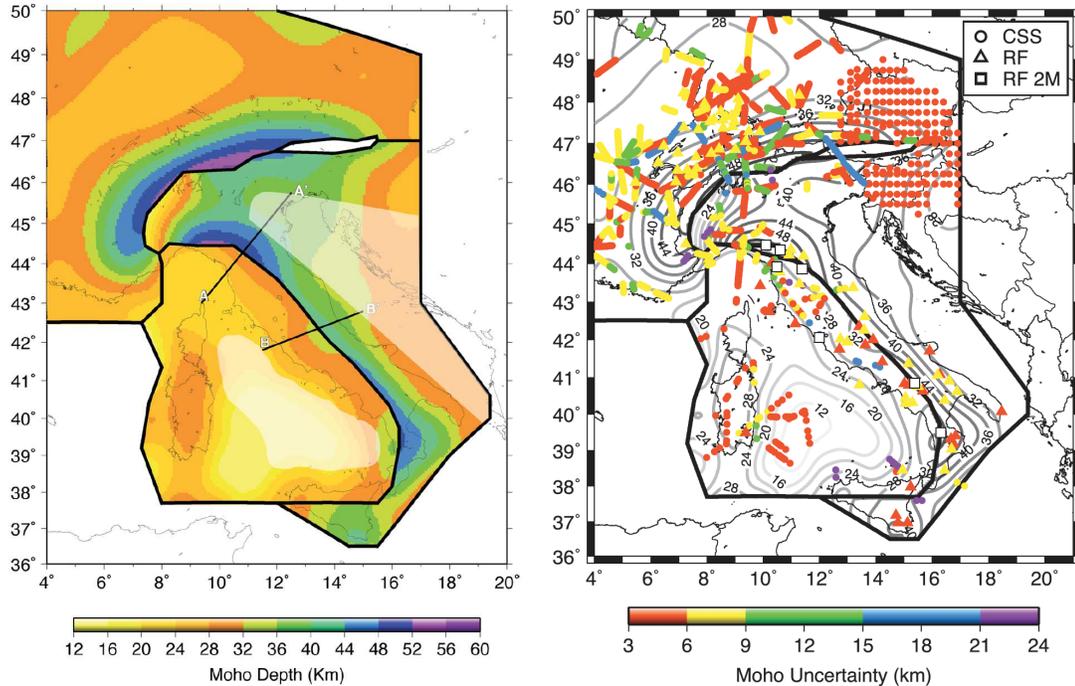


Figure 1.3.2: On the left, map describing the Moho topography for the European, Adriatic and Ligurian plates. Transparent areas denote regions of reduced reliability of interpolation results due to the presence of no reliable Moho data. On the right, CSS and RF Moho information for the three plates including their relative uncertainties used for the interpolation of the Moho map. Circles refer to CSS, triangles refer to high-quality RF and squares refer to RF where 2 Mohos are identified (figure from Spada et al., 2013).

The European Moho shows a typical thickness of a continental crust in the northern Alpine foreland (around 32 km) and a dipping Moho toward South-East in the Central Alps and toward South in the Eastern Alps. The deepest values with more than 55 km are observed in the central Alps along the southern plate boundaries, while in the Eastern Alps, CSS and RF information show a smoothly shallowing European Moho toward east from about 50 km beneath the Tauern window. Along the Western and Central Alps, the European Moho descends beneath the Ivrea body and the Adriatic Moho. The Ligurian plate is characterized by different kind of lithospheres in different stages of evolution from North to South and from West to East in the Ligurian Sea.

In the region of the Adriatic Sea, as in the European Moho, we find a typical depth of the continental crust for the Adriatic Moho (around 32 km). In the continental region, the Adriatic Moho is dipping towards the plate boundary with Liguria. Greater depths (~ 48 km) at the plate boundary characterize the northern Apennines and the Calabrian Arc, while in the centre part of the Apennines we observe a maximum of about 40 km. In the north, the Adriatic Moho smoothly dips from a typical young continental depth (~ 32 km) along the

axis of the Po plain to 40 km and more in the southern Alps.

On the West side, along the Europe-Adria plate boundary, the Adriatic Moho marks the top of Ivrea body, that in some places is located at depths less than 12 km.

The Moho uncertainties in the Spada et al. (2013) model is defined taking into account Fresnel volume and sampling geometries, following a new quality weighting scheme based on uncertainty of RF data, specifically designed to match the criteria in Waldhauser et al. (1998) for the CSS information. In order to provide an approximate quantitative description of the uncertainties of RF results, they divide, based on quality criteria, the data sets into four different classes from 0 (top) to 3 (not to be used), and two special (4 and 2M) for additional information.

The top *Class 0* includes a clear identification of the Ps (Moho) phase, a very high back-azimuth coverage (greater than 180°) and an uncertainty of ± 3 km, *class 1* with an azimuthal coverage between 150° and 180° with a less clear identification and an uncertainty of ± 6 km, *class 2* with a back-azimuth coverage between 90 and 150° with an uncertainty of ± 10 km and *class 3* includes the rejected stations with poorly identifiable Moho and back-azimuth coverage $\leq 90^\circ$. The two special classes include all stations with Moho signals that show the presence of specify additional interfaces.

The resulting Moho map denotes the simplest smoothest topography with the least number of separate interfaces that fits all available CSS and RF data within their respective individual uncertainty limits.

1.3.3 Receiver function studies

The receiver function (RF) technique is a way to image the inner structure of the Earth and its internal discontinuities by using the information from teleseismic earthquakes recorded at a three-component seismograph. A teleseismic P-wave will generate P-to-S conversions at boundaries, such as the Moho (crust-mantle boundary), beneath the seismograph; the difference in travel time between the generated S-wave and P-wave contains informations about the depth to the boundary and regarding the P- and S- wave velocities. If further reverberations are included more detailed structure can be resolved (Vinnik, 1977). For details about the Receiver Function method, see Chapter 2.

Recently in the Alpine context, investigations with RFs in Eastern Alps and Pannonian Basin took place (Bianchi and Bokermann, 2014; Kalmár et al., 2018). In the Western Alps the most important study performed so far with the RF method is the work of Lombardi et al. (2008) based on permanent stations. The goal was to establish a well-constrained Moho depth and crustal Poisson's ratios based on the $H - \kappa$ technique (Zhu and Kanamori, 2000). In this work, the receiver functions image a gently southeast dipping European Moho

beneath the northern Alpine foreland (depth ≤ 40 km). The dip increases from $5\text{-}10^\circ$ to about 15° towards the suture zone where the Moho depth reaches a maximum value around 55 km. South of the Periadriatic Line, the Adriatic Moho reaches a depth around 40-45 km possibly dipping N-NE. Main results in terms of Moho depth in the Western part of the Alps are shown in the Figure 1.3.3.

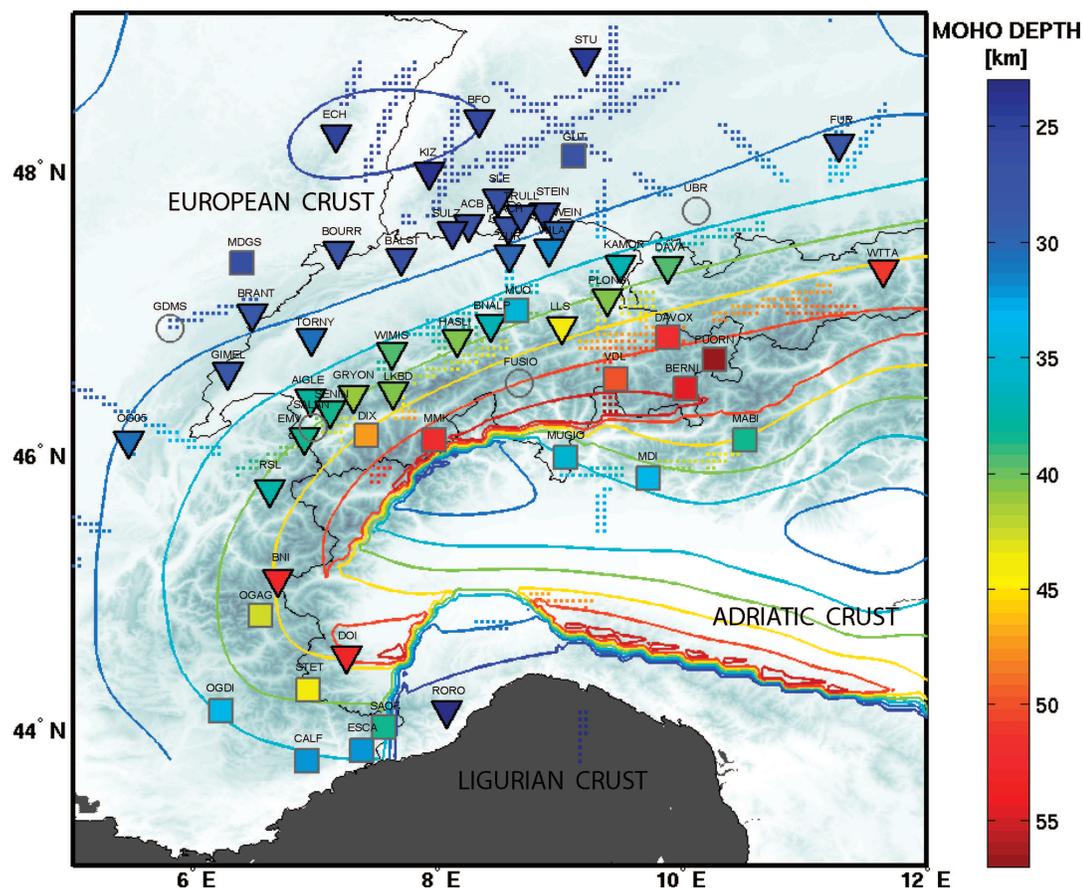


Figure 1.3.3: Moho depth variation in the Western-Central Alps. Reverse triangles represent Moho depth from $H - \kappa$ technique, squares depth from P_s times and $\frac{V_p}{V_s} = 1.75$. Stations without reliably identifiable P_s phase are shown as open circles: the symbols overlay Moho depth isolines and high-quality seismic profiles (dotted lines) with associated uncertainties smaller a 5 km from a CSS based model (from Lombardi et al., 2008).

Regarding Poisson's ratios, units in the Alpine foreland show a normal value near 0.25, the Helvetic nappes and Southern Alps have a lower average ratio of 0.22, which the authors attribute to upper-crustal thickening, while low ratios are in the southern Western Alps, compared to sites in the north, imply a discontinuity of the deep crustal structure of the Western Alps. A more detailed comparison between Lombardi's work and our results using the $H - \kappa$ approach will be discussed in Chapter 2.4.3.

1.3.4 Local earthquake tomography

Seismic tomography⁶ is a technique for imaging the Earth’s interior in 2-D or 3-D with seismic waves produced by earthquakes or explosions (Iyer and Hirahara, 1993; Rawlinson et al., 2010; Nolet, 2012).

In a technique similar to Computerized Tomography (*CT*) scans, where doctors use tomography scans to image the organs and bones, seismologists use seismic waves to create images of the Earth’s interior. The energy is typically released from the focus of an earthquake as seismic waves that spread outwards in all directions and travel through the Earth’s interior until it reaches the surface, where is recorded on seismograms. By recording the seismograms of many earthquakes, seismologists are able to create high-resolution images of Earth’s interior using the method of seismic tomography. Using the travel distance of the wave from the epicenter of the earthquake to the seismometer at the recording station, seismologists can invert the average velocity distribution of the seismic waves. The speed of the seismic waves depends on the type of rock materials, the temperature of the rocks (hot, partially molten rocks or cold rocks) and the “stability” of the area, tectonically active or intra-plate zone (Ammon et al., 2010). Many rock physics parameters contribute to changes in seismic velocities of rocks in addition to mineralogy, porosity, and *in situ* stress conditions such as pore fluid properties, which in turn depend on temperature and pressure (Birch, 1960; Spencer Jr and Nur, 1976; Toksoz et al., 1976; Christensen and Yuen, 1985; Wang et al., 1989; Dvorkin et al., 1999). High-frequency modeling of ultrasonic laboratory data (Mavko and Vanorio, 2010) suggests that frequency-dependent pore-fluid effects can impact observed effective stress coefficients for seismic velocities.

Unlike teleseismic tomography (Aki et al., 1977) that uses waves from distant earthquakes that deflect upwards to a local seismic array, Local Earthquake Tomography (LET) depends on earthquake sources within the volume to investigate. To do this, a possibly homogeneous distribution of stations is very important, as otherwise the model is not sampled in a homogeneous way and artifacts can be introduced into the solution (Lockett and Baptie, 2015). Figure 1.3.4 points out the importance of having a good coverage in the tomographic approach.

⁶The term tomography is derived from the Greek word *tomos*, which means “slice” or “section” and *graphie* meaning “drawing”.

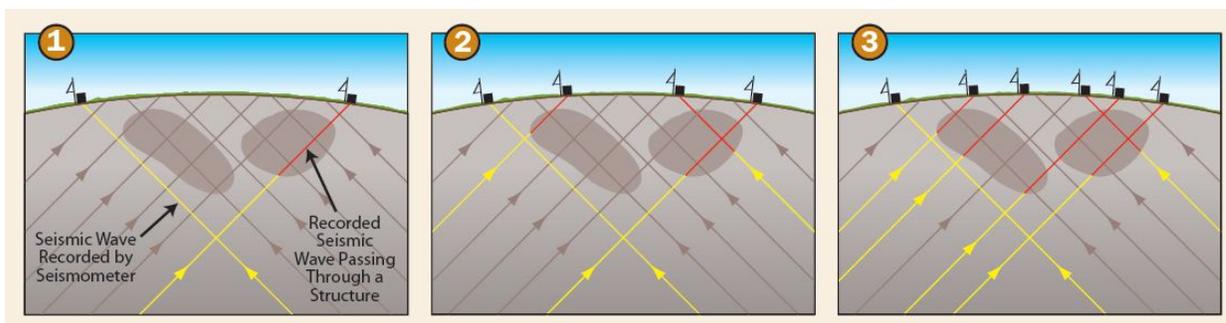


Figure 1.3.4: Importance of good coverage in the tomographic approach: by increasing the number of the recording stations (panels from 1 to 3), one gets much more data which will lead to image the shape and the size of the structures reducing the uncertainty in the tomographic results (image from <http://www.iris.edu/USArray>).

In panel 1) of Figure 1.3.4 we have two seismometers on the surface and only one recorded wave passes through part of one of the structures with different structures and scientists examining the recorded seismograms would infer that only one structure is present instead of two. In 2), when more seismometers are added, seismologists can detect two structures from the recorded seismographs, but they cannot determine the size and shape of each. In 3) with six seismometers catch enough recorded waves to start determining the borders of the structure.

During last 20 years, the structure of the Alpine crust was investigated using LET with several studies, e.g. Paul et al. (2001), Diehl et al. (2009) and Solarino et al. (2018). Here I focus in particular on the work of Diehl et al. (2009) which represents the best existing model of 3-D P-wave velocity structure of our study area in terms of resolution and high quality-data, as shown in Figure 1.3.5. In good agreement with previous 2-D CSS investigations, the Moho interface and the Ivrea body are well resolved, moreover a detailed analysis of focal depths in the 3-D P-wave model shows that all investigated earthquakes in the northern foreland have occurred within the European crust.

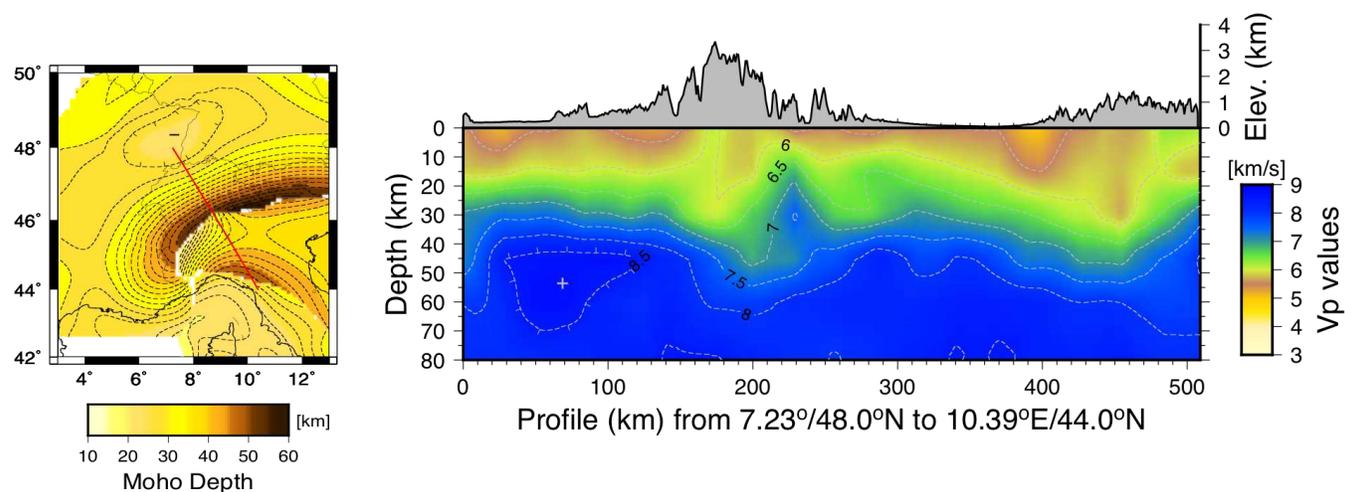


Figure 1.3.5: On the left, Moho depth map from Spada et al. (2013), on the right a cross-section of the 3-D P-velocity model from Diehl et al. (2009). The red line on the map represents the section Basel-Chiasso, where the P-velocity model is extracted. High velocity structure at shallow depth around 225 km along X axis represents the Ivrea Body. Vertical exaggeration for the Vp cross section is 2:1.

For this work, I decide to take the 3-D P-wave velocity model from Diehl et al. (2009) as a P-wave local velocity starting model. The choice of this model as the initial one is that it features both large aerial coverage over the Alps and the best resolution available in that domain: 25*25*15 km in the two horizontal and the vertical directions, and it has successfully imaged several known zones of the Alpine orogen.

1.3.5 Ambient noise tomography

The Ambient Noise Tomography (*ANT*) method provides a powerful tool for sampling the Earth’s shear-wave velocity structure V_s (Campillo and Paul, 2003; Shapiro et al., 2005; Yang et al., 2007). In *ANT* low amplitude seismic “noise” is used, which is present at all seismic stations, generated from a wide ranges of sources like waves and wind along shorelines and local man-made activities.

Correlation of noise between pairs of stations at tens of km spacing, stacked over many months, provide Green’s function of wave propagation between the stations, largely surface waves (see Figure 1.3.6). It has been proved (Weaver, 2005) that these signals are more robust at 5-20 s periods, where fundamental-mode Rayleigh waves sample the crust and the upper mantle, allowing 3-D imaging at these depths. *ANT* investigations are particularly useful in seismic areas, e.g. continental interiors, where *LET* is not feasible. However, depth penetration remains limited by the seismic network aperture. and wave propagation is mostly horizontal.

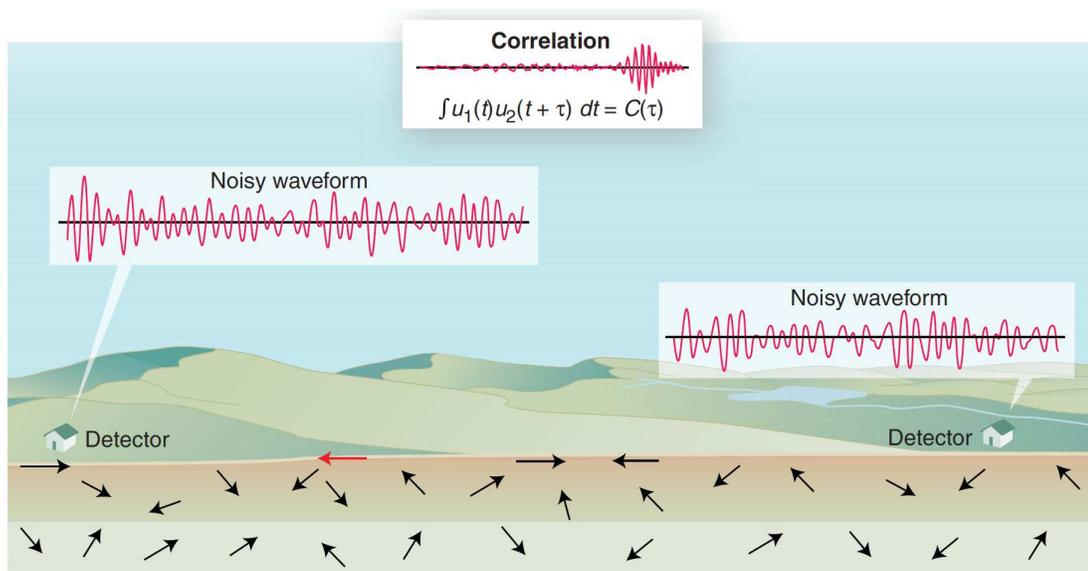


Figure 1.3.6: When a diffuse wave field is generated by distant sources, stations (detectors) report random signals. It can happen that a ray (for example, the one shown in red) passes through both detectors. As a result, the signals are weakly correlated (from Weaver, 2005).

Investigations by ambient noise tomography are particularly useful in aseismic areas, e.g. continental interiors. In the Alpine context, several studies were performed in the last 10 years (Stehly et al., 2009; Li et al., 2010; Verbeke et al., 2012; Molinari et al., 2015; Guidarelli et al., 2017) and one recent work with a large dataset and a high-resolution 3-D shear-wave velocity model of the crust (Figure 1.3.7) is published by Lu et al. (2018).

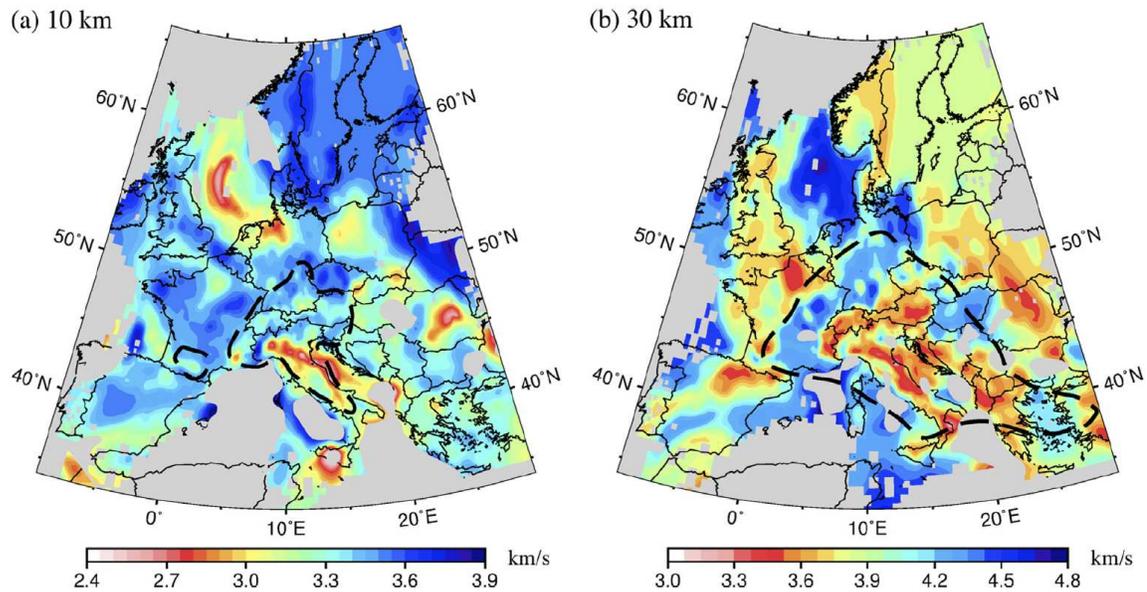


Figure 1.3.7: Depth slices with more than 10 crossing ray paths at 8 s period for the final V_s model at 10 km (a), and 30 km (b) for the model performed by Lu et al., 2018. The resolution of the area is 0.3° at 8 s. I plot only cells crossed by more than 10 paths. The black dashed lines enclose the well-resolved area.

Focusing on the interpretation of the Central Alps, we can observe that in the upper crust (10 km), the areas of lowest velocities (2.5–2.9 km/s) correspond to thick sedimentary basins such as the Po plain and the Adriatic basin, while the Moesian platform is not well resolved. At 30 km depth, variations in crustal thickness dominates the picture with lower velocities around 3.5 km/s in the crustal root of the mountain belts like the Alps, and high velocities (more than 4.1 km/s) in the areas sampling the mantle in the European plate foreland.

As we will see in the last chapter, the final 3-D crustal model based on P-to-S converted waves with a ray-paths sampling closer to vertical, is compared along profiles from the 3-D ambient noise tomography by Lu et al. (2018) to have a consistent comparison between 3-D structures obtained by different methods.

1.3.6 Gravity studies

Gravity studies are used to detect and identify subsurface bodies and anomalies within the Earth. These investigations are based on the density differences of rock bodies and their effect on the gravity acceleration. The intensity of the force of gravity due to a buried, anomalous mass is superimposed on the larger force of gravity due to the total mass of the Earth. Therefore, two components of gravity forces are measured at the Earth's surface: first, a general and relatively uniform component due to the total Earth and second, a component of much smaller size that varies due to lateral density change, which is called *gravity anomaly* (Telford et al., 1990). Usually, the observed raw gravity data is corrected for Earth rotation, latitude, tidal effects and gravity meter fluctuations to obtain the local gravity value. The *free air* gravity anomaly is further corrected for the gravity effect caused by the elevation difference between the station and sea level and is a common correction for oceanic gravity interpretation. The *Bouguer gravity anomaly* is further corrected for the mass that may exist between sea level. A simple-Bouguer anomaly has undergone a simplified removal of topographic effects, which suffices in relatively flat areas; a complete-Bouguer anomaly contains a terrain correction that uses a more complete representation of the local topography, which is necessary for accurate gravity anomaly values in mountainous area.

Over the last 45 years, both the Eastern (Granser et al., 1989; Ebbing et al., 2001) and the Western Alps (Kahle et al., 1976; Kissling, 1980; Kissling et al., 1983; Schreiber et al., 2010) have been studied by several gravity investigations. One of the first investigation in the Alps is from Rybach et al., 1980 where the so called "Swiss Geotraverse Basel-Chiasso" (same profile of the Figure 1.3.5) has been analyzed in detail (see Figure 1.3.8).

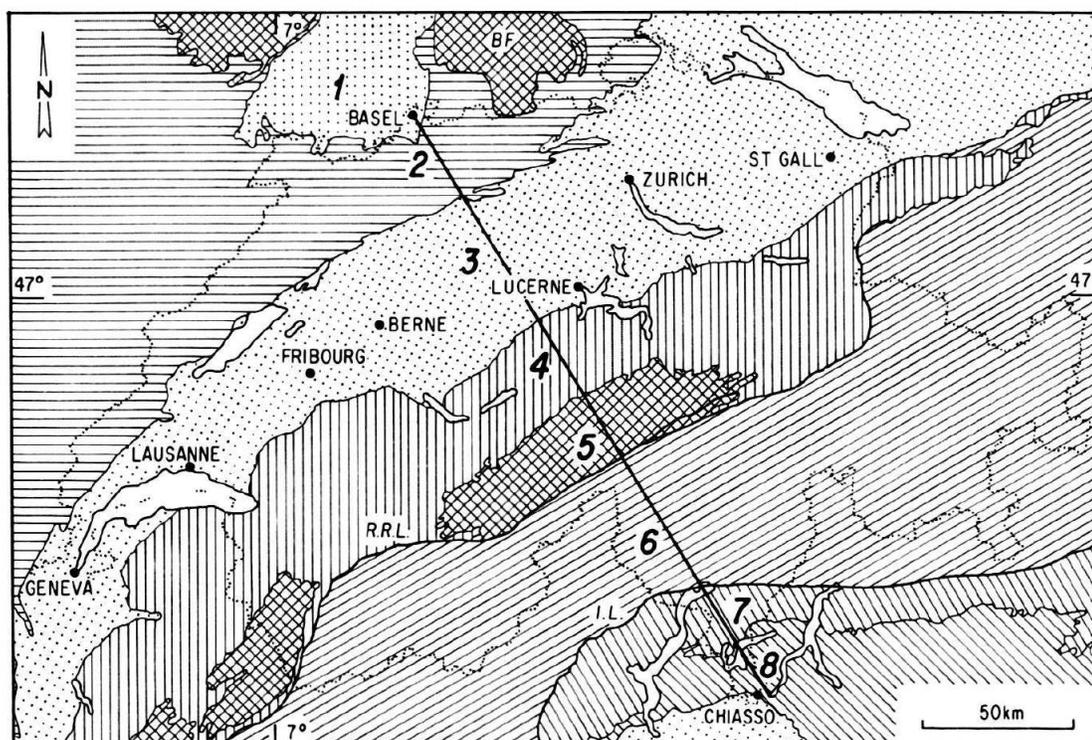


Figure 1.3.8: Swiss Geotraverse Basel-Chiasso. 1- Rhinegraben, 2- Jura mountains and undeformed foreland cover, 3- Molasse basin and Subalpine Molasse, 4- Helvetic nappes and remnants of Penninic cover rocks, 5- External massifs, 6- Subpenninic, Penninic and Austroalpine units, 7- South-Alpine basement with Ivrea zone, 8- South-Alpine cover (from Rybach et al., 1980).

This section across Central Switzerland crosses most of the main geologic units perpendicularly to their general strike. At first glance the Bouguer map of the Central Alps shows two main features: the distinct gravity low of the Alps due to the thickened orogenic root and the well-known positive anomaly due to the dense and shallow-lying rocks of the Ivrea-Verbano Zone (IVZ). The sediments of the Molasse basin have a significant effect to be taken into account in the study of the deeper crust (Kissling, 1980). Other tectonic units, for example the Quaternary sediments, produce anomalies of only a few mGals in amplitude or have a short spatial wavelength (Kissling et al., 1983). Based on the available geophysical and geological information, Kissling (1980) computed the negative gravity anomaly of the sediments of the Molasse basin and the positive gravity anomaly of the Ivrea-Verbano zone using a 3-D model (Figure 1.3.9).

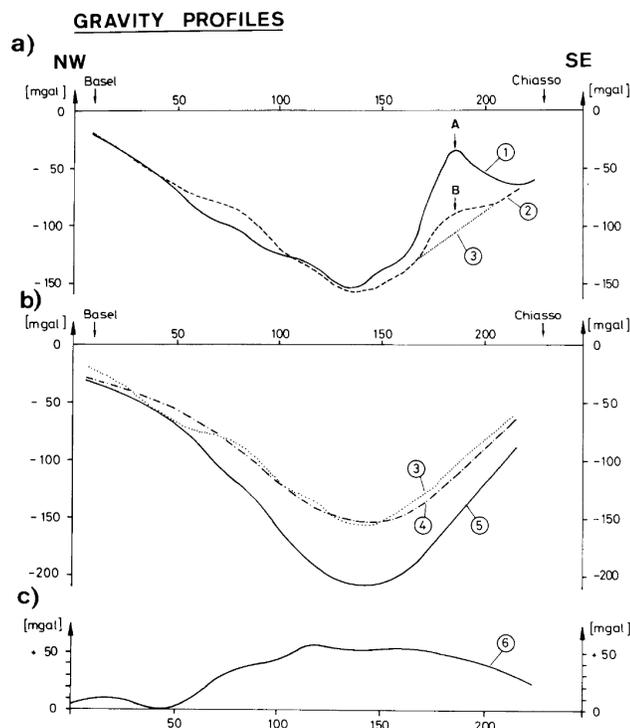


Figure 1.3.9: Gravity anomalies (curves 1 to 6) along the profile Basel-Chiasso (see Figure 1.3.8). Curve 1: Bouguer anomaly. A: effect of the Ivrea body. Curve 2: Gravity anomaly after elimination of the effects of the Ivrea body and of the Molasse sediments. Curve 3: Curve 2 without the local anomaly of Ticino (B). Curve 4: Bouguer anomaly after Kahle and Werner (1980), in which no mantle effects are taken into account. The average density contrast crust/mantle in this model is 0.37 g/cm^3 . Curve 5: the new synthetic curve with a density contrast of 0.50 g/cm^3 . Curve 6: the difference between curves 4 and 5. Image from Kissling et al. (1983).

Curve 2 of Figure 1.3.9 shows the corrected gravity profile along the transverse from Basel-Chiasso; the bulge on the southern end of the gravity profile is due to a local positive anomaly in southern Switzerland (Canton Ticino). This gravity anomaly along the Swiss Geotraverse has been interpreted in terms of a multi-layered crust in direct relationship to the established seismic models along the same profile.

In general, the most profitable results are obtained from the integrative interpretation of the gravity and seismic data (Robertsson et al., 2011). Even the reliability of the gravity modeling depends on additional data used in the interpretation, therefore it is necessary to bring together data from various geophysical techniques, which might be used in construction of integrative interpretation models. In light of this work, gravity data obtained from potential methods can contribute to the integration and the interpretation of the 3-D crustal model obtained with the proposed technique of the converted wave tomography.

Chapter 2

Receiver function method and dataset

In this chapter I describe briefly the fundamentals of the receiver function (RF) technique, then I introduce the dataset used in this work with the qualitative control. At the end, in order to determine the first order discontinuities, I compare the results I obtained for the crustal thicknesses and the V_p/V_s ratio using the $H - \kappa$ grid search (Zhu and Kanamori, 2000) with a previous work that used the same method for the Central Alps (Lombardi et al., 2008).

2.1 Fundamentals of receiver functions

The receiver function (RF) method has been broadly used in the last 30 years to retrieve information about discontinuities in the crust and in the upper mantle beneath three-component seismic stations, using natural earthquakes as sources. By using the elimination of the common characteristics from the seismic source, we are able to obtain the impulse response of the Earth just beneath the station. This technique was originally developed for single station analysis (Langston (1977b); Langston (1977a); Langston (1979)) and then was improved for different aspects, from the deconvolution to the inversion and applied on station networks. Some of these important developments on the technique are given in the studies of Owens and Crosson (1988), Ammon et al. (1990), Cassidy (1992), Sambridge (1999a), Frederiksen and Bostock (2000), Park and Levin (2000) and Zhu and Kanamori (2000). The main goal of this method is to locate sharp interfaces on the receiver side of the propagating waves.

Figure 2.1.1 shows the direct waves propagating through the Earth and arriving at a seismic station are clearly observable, in the meantime the impinging waves that undergo a phase conversion at a discontinuity are also visible, when the conversion occurs below the receiver. This conversion can be P-to-S or S-to-P.

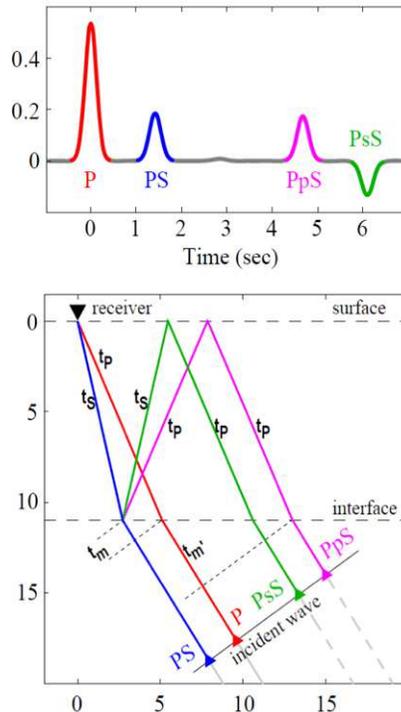


Figure 2.1.1: Above, example of synthetic receiver function showing the direct P-wave and the converted phases with a positive impedance contrast increasing with depth at an interface below the station. Below, corresponding path of the different phases from an incident wave (figure from Hetényi, 2007).

RFs are observable if a P-wave hits the interface at an angle close to the vertical ($\sim 20^\circ - 40^\circ$), which limits the range of useful epicentral distances to about $30^\circ - 90^\circ$. In practice, this means that earthquakes have to reach a certain magnitude to be detected, generally $M > 5.0 - 5.5$.

Beyond a P-to-S conversion, reverberations in the upper layer, called *multiples*, are also produced. These wave phases arrive later compared to the P-to-S conversion; the role of the multiples are very important because they provide additional information on the velocity structure above the interface.

Several RF investigations using advanced migration and different inversion strategy (Rondenay et al., 2005; Abe et al., 2011) were performed to image the Moho in different orogenic contexts with respect to the Alps (e.g. Himalaya, see Singer et al., 2017; Subedi et al., 2018).

2.1.1 RFs as a convolutional model

From a mathematical point of view, the recorded wavefield at a station on the Earth's surface can be described through a convolutional model. One can compose a seismogram as a combination of the source time function of the earthquake $S(t)$, the Earth's structure impulse response $E(t)$ and the instrument's impulse response $I(t)$. For a three-component seismogram, this can be written for the vertical, radial and tangential components, as:

$$Z(t) = S(t) * E_Z(t) * I(t), \quad (2.1.1)$$

$$R(t) = S(t) * E_R(t) * I(t), \quad (2.1.2)$$

$$T(t) = S(t) * E_T(t) * I(t). \quad (2.1.3)$$

where $*$ represents convolution, $R(t)$ and $T(t)$ are the radial and transverse components of the seismogram which are obtained by rotating the original Z-N-E component data into a coordinate system pointing towards the earthquake, and indexes Z, R and T refer to the vertical, radial and transverse component of the Earth structure E . The goal of the converted-wave seismology is to retrieve $E_R(t)$ and $E_T(t)$ from the original data, which are respectively called the radial and transverse (or tangential) receiver functions.

A general approximation to obtain RFs is that the observed vertical component $Z(t)$ represents well-enough the convolution of the earthquake's source time function with the instrument's response (Burdick and Helmberger, 1978):

$$S(t) * I(t) \simeq Z(t) \quad (2.1.4)$$

or, in other words, that $E_Z(t) \simeq 1$. Therefore, by deconvolving the radial and tangential components from the vertical component seismogram, one obtains the wanted radial and tangential receiver functions $E_R(t)$ and $E_T(t)$.

In RF techniques, two deconvolution methods are commonly used: the frequency-domain spectral deconvolution, and the time-domain iterative deconvolution.

2.1.2 Spectral deconvolution (frequency domain)

Convolution and deconvolution are performed in the frequency domain as multiplication and division, using Fourier-transforms. In the case of receiver functions:

$$E_R(f) = \frac{R(f)}{I(f) * S(f)} \simeq \frac{R(f)}{Z(f)}, \quad (2.1.5)$$

$$E_T(f) = \frac{T(f)}{I(f) * S(f)} \simeq \frac{T(f)}{Z(f)}, \quad (2.1.6)$$

Due to the limited bandwidth of the signals, the equation is not numerically stable (Langston, 1979). Clayton and Wiggins (1976) proposed to fill the troughs of the above spectral division to a minimum water level, which is usually 10^{-4} . Filtering is achieved by a Gaussian filter in the frequency domain applied to the result of the spectral division. The advantage of the method is that it is relatively quick, however it may produce sidelobes or long-period artifacts.

2.1.3 Iterative deconvolution (time domain)

Ligorria and Ammon (1999) presented an alternative deconvolution technique, following an iterative scheme. This approach determines by cross-correlation of the vertical component and the original horizontal component signal the location of the maximum correlation, and considers this location as a spike on the receiver function. The convolved effect of this spike is subtracted from the original horizontal component, and the procedure starts again. In this way, the difference between predicted and observed horizontal seismogram is minimized in the least square sense. The iteration is stopped either at a preset number, or at a given value of fit. At the end of the iterations the series of spikes represent the best transfer function between the vertical and horizontal signals, thus the receiver function.

The iterative deconvolution has the advantage of finding the “simplest” receiver function with high-frequency content but without side-lobe artifacts or long-period instability, and is therefore better adapted to map sharp interfaces, such as the ones in the lithosphere. However, the computation time is somewhat longer. Filtering is achieved on the original raw data. The ultimately obtained spike series must be convolved by a time-domain Gaussian to obtain RFs with the same frequency content as the filtered data.

2.1.4 Example of a seismic record and a RF computation

In Figure 2.1.2 I show an example of a seismic record and the corresponding RF computation at a station in Switzerland for the Chiapas earthquake, in Mexico (2017-09-08, 04:49:19 UTC, Mw 8.2).

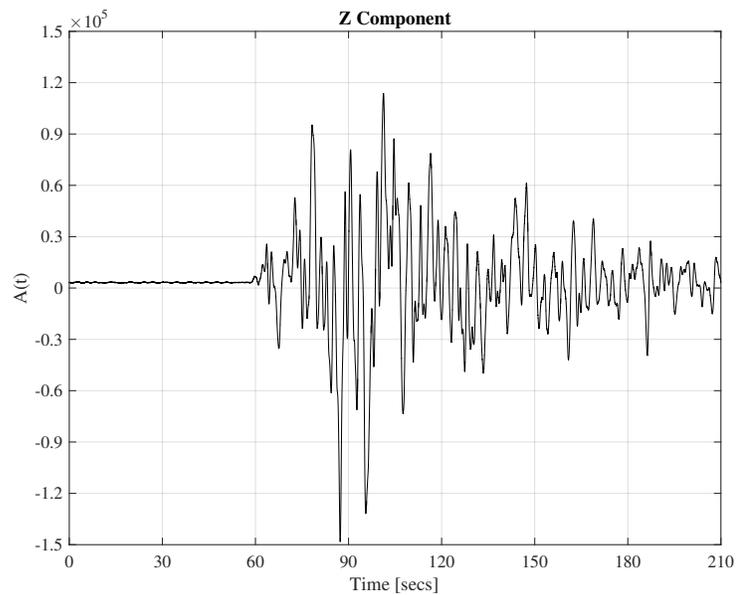


Figure 2.1.2: Vertical component recorded for the Chiapas earthquake, Mexico (2017-09-08, 04:49:19 UTC) at the station ZUR.

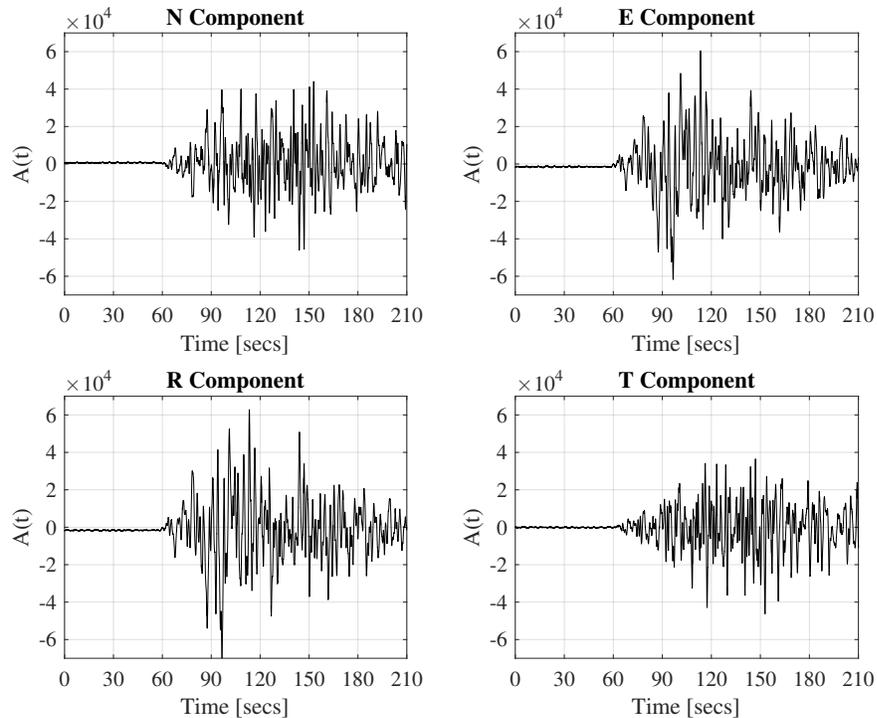


Figure 2.1.3: Above: original components, below: radial and transverse component after the rotation for the sample event.

For each event I saved 1 minute before the P-arrival and 2.5 minutes after. Radial and tangential components are oriented parallel and perpendicular to the direction of the incident wave, and are obtained through rotation of the original North N and East E component records. In the Figure 2.1.3 the original (N and E) and the rotational components (R and T) are shown for the same event of Figure 2.1.2.

For my work, I decided to use the iterative time domain deconvolution computed using the algorithm by Ligorria and Ammon (1999) using 100 iterations, a gaussian filter of order 2 to convolve the traces for a frequency band between 0.125 Hz and 0.5 Hz. This deconvolution method is used in several crustal investigations, e.g. Beck and Zandt (2002), Zor et al. (2003), Abt et al. (2010). The result for the sample event is shown in Figure 2.1.4.

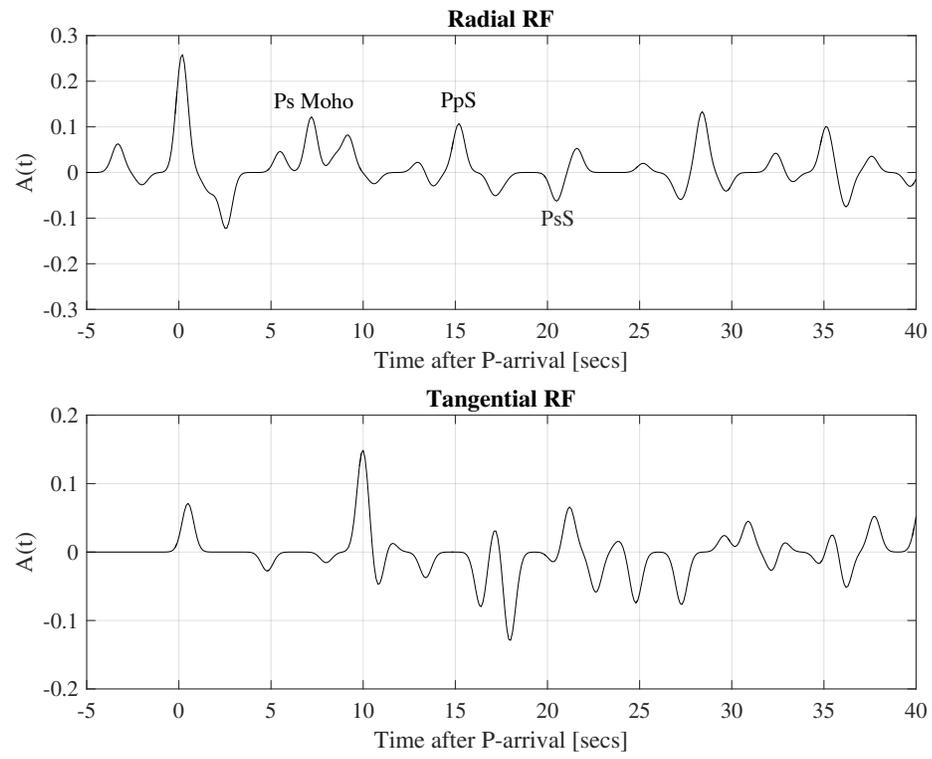


Figure 2.1.4: Radial and Tangential RF computed for the Chiapas event at station ZUR.

2.2 Database construction

I collect data to compute RFs in the zone of the Central Alps (which is a very complex area from a geological point of view, see Chapter 1) in a rectangular box from 5° to 12° E and 45° to 48.5° N where a dense network of stations is available. I consider both permanent broadband stations from Switzerland and for the surrounding countries and the temporary stations of the AlpArray project’s seismic network (Hetényi et al., 2018b); permanent stations are provided by nine agencies (Figure 2.2.1).

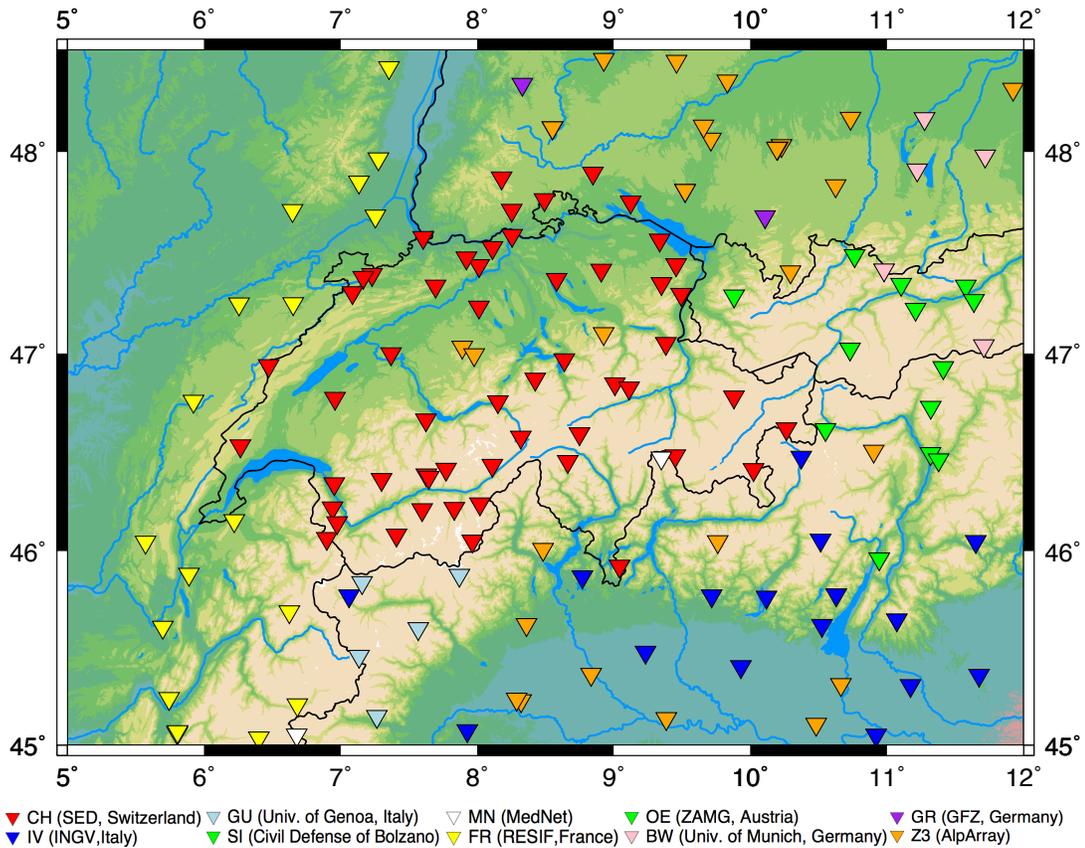


Figure 2.2.1: Stations used in the studied area. Two-letter abbreviation in the legend are seismic network codes of stations.

Temporary stations are part of the AlpArray project (Hetényi et al., 2018b), which represents a European consortium to advance our understanding of orogenesis. The AlpArray Seismic Network has been deployed with contributions from 36 institutions from 11 countries to achieve high-resolution geophysical images of structures from the surface down to the base of the mantle transition zone.

For our study, I consider all the teleseismic events from the website of U.S. Geological

Survey with $M > 5.2$ occurred from May 1995 (start of the earliest selected station) to June 2018, that has epicentral distance between 25° and 95° with respect to the reference point located in Lausanne, Switzerland. The distribution of the earthquakes considered for this work are represented in Figure 2.2.2.

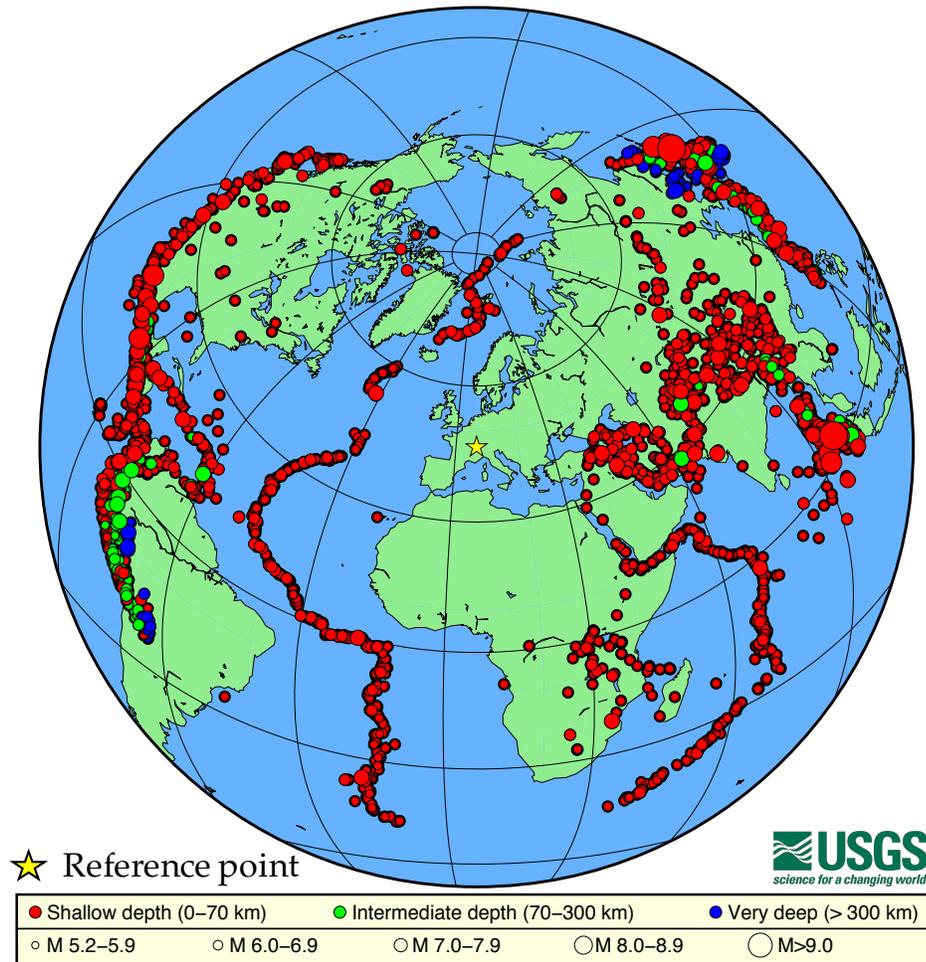


Figure 2.2.2: Distribution of earthquakes considered for this study. With yellow star is shown the reference point located in Lausanne, Switzerland.

The seismic sources originate from a broad area and back-azimuth range, including California in the West and Japan in the East; only the southernmost back-azimuth has few usable earthquakes. As we can see, from our location we can cover a broad range of events in terms of back-azimuth, including the earthquakes from California in the West until the teleseismic events in Japan in the East.

To collect data, we used *arlink_fetch* and *FDSNWS* (the international Federation of Digital Seismograph Networks Web Services) method, which allows us to request and obtain

data from several servers. In total, I have collected data from 6 450 events at 150 stations, yielding a total of 287 414 three-component records, in miniSEED format.

2.3 Quality Control

Considering the amount of data to analyze, it is impossible to perform a visual investigation of the entire dataset at each trace. For this reason, a semi-automatic approach proposed by Hetényi (2007) is used to select higher quality seismic records, both on Z-N-E components and in RFs.

2.3.1 QC on ZNE components

I applied first a quality control on the three-component of Z-N-E data. All traces are cut 30 seconds before and 150 seconds after the direct P-wave arrival and filtered according the frequency band that will be used for the computation of the RFs, which is between 0.125 Hz and 0.5 Hz.

For each single trace, I compute the following values:

- *rmsall*, root mean square (*rms*) of the whole trace;
- *rmsbg*, root mean square of the background signal, between 30 and 5 seconds before the P-arrival;
- *maxbg*, maximum of the background;
- *maxpk*, maximum of the peak, between -5 and 20 seconds with respect to the P-arrival.

The quality of a trace is measured using four control parameters, $C1$ to $C4$. A trace is of good quality if it respects the following criteria:

- It looks similar to other traces acquired at other stations for the same event (represented by the median):

$$median * C1 \geq rmsall \geq median / C2 \quad (2.3.1)$$

- It has a high signal to noise ratio (*SNR*) in amplitude:

$$maxpk \geq maxbg * C3 \quad (2.3.2)$$

- It has a high SNR compared to the background noise:

$$maxpk \geq rmsbg * \sqrt{2} * C4 \tag{2.3.3}$$

When the value of $C1$ and $C2$ increases, the range of acceptable traces increases. When $C3$ and $C4$ decrease, there are more traces included in the selection. Examples are shown in Figure 2.3.1 and 2.3.2.

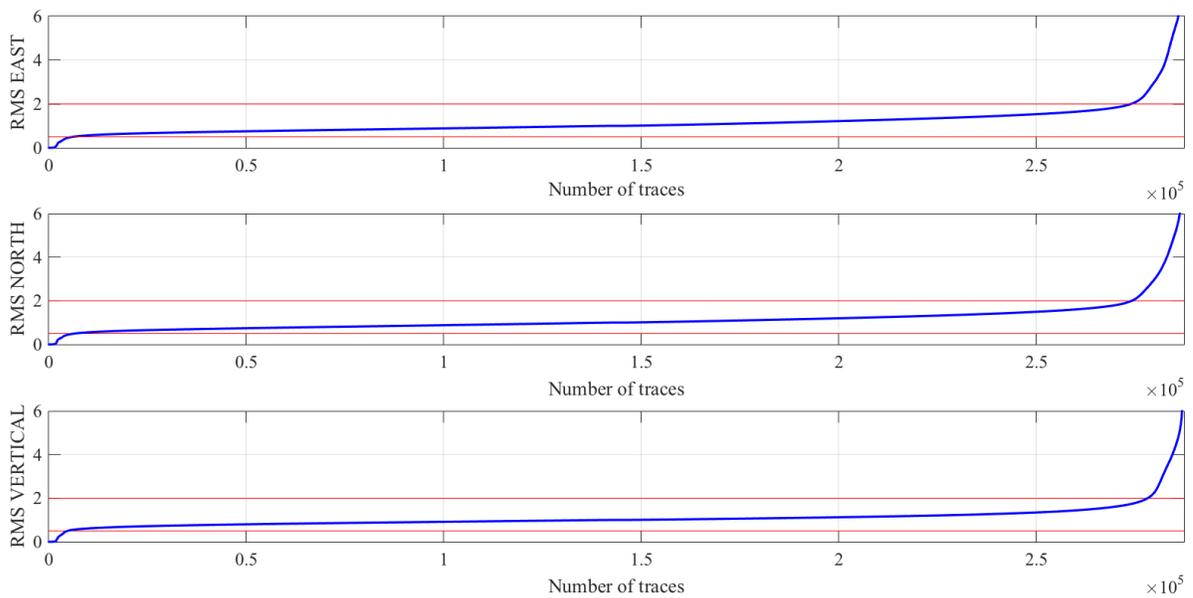


Figure 2.3.1: Quality control on East, North and Z component. Blue line represents the sorted values of RMS for each component, while the band defined by the red lines show the range where the ratio $\frac{rms_{small}}{median}$ is between $C1$ and $\frac{1}{C2}$.

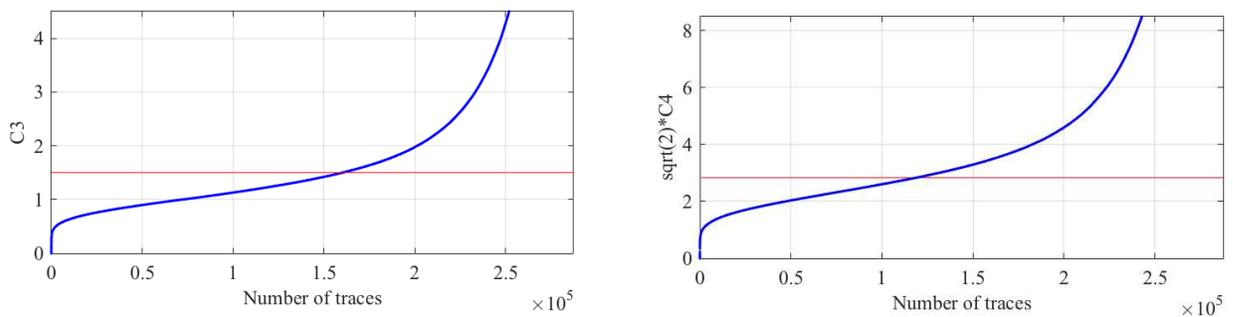


Figure 2.3.2: On the left, the blue line represents the ratio $\frac{maxpk}{maxbg}$ while $C3$ is in red; on the right the ratio $\frac{maxbg}{maxpk}$ against $\sqrt{2} * C4$.

2.3.2 QC on RF

I applied a second set of criteria that allows to select RFs to avoid low quality data, which can be produced for distant, low magnitude earthquakes or noisy recordings. A preliminary selection of RFs was performed based on the characteristics of the P-wave arrival. The position of the maximum absolute amplitude sample is chosen t_p along with its amplitude A_p and these should satisfy:

$$|t_p| \leq 2sec \quad (2.3.4)$$

$$0.05 \leq A_p \leq 0.9 \quad (2.3.5)$$

These criteria eliminate traces where the P-wave does not emerge from the background or where there are timing problems.

The effectiveness of the quality control is shown in Table 2.1. With increasing restrictivity of the first quality control (on the Z-N-E components), the second (on the RFs) eliminates less and less data and confirms the trend of the first quality control.

List name	C1	C2	C3	C4	QC1	(%)	QC2	QC2 / QC1 (%)
Whole Dataset					287 414	100		
Weak	4	4	1.1	1.5	186 693	65.0	45 777	24.5
Medium	2	2	2	3.5	70 801	24.6	33 292	47.0
Strong	2	2	10	15	13 055	4.5	10 046	77.0
This work	2	2	1.5	2	116 804	40.6	28 494	24.4

Table 2.1: Statistics on the efficiency of the quality control procedures.

For our dataset, I applied different quality control procedures in order to have a good selection of data. After tuning the parameters, I have 28 494 traces, which represent about the 10% of our initial dataset.

Figure 2.3.3 shows the flowchart of all the phases of the pre-processing from data download to the final quality control.

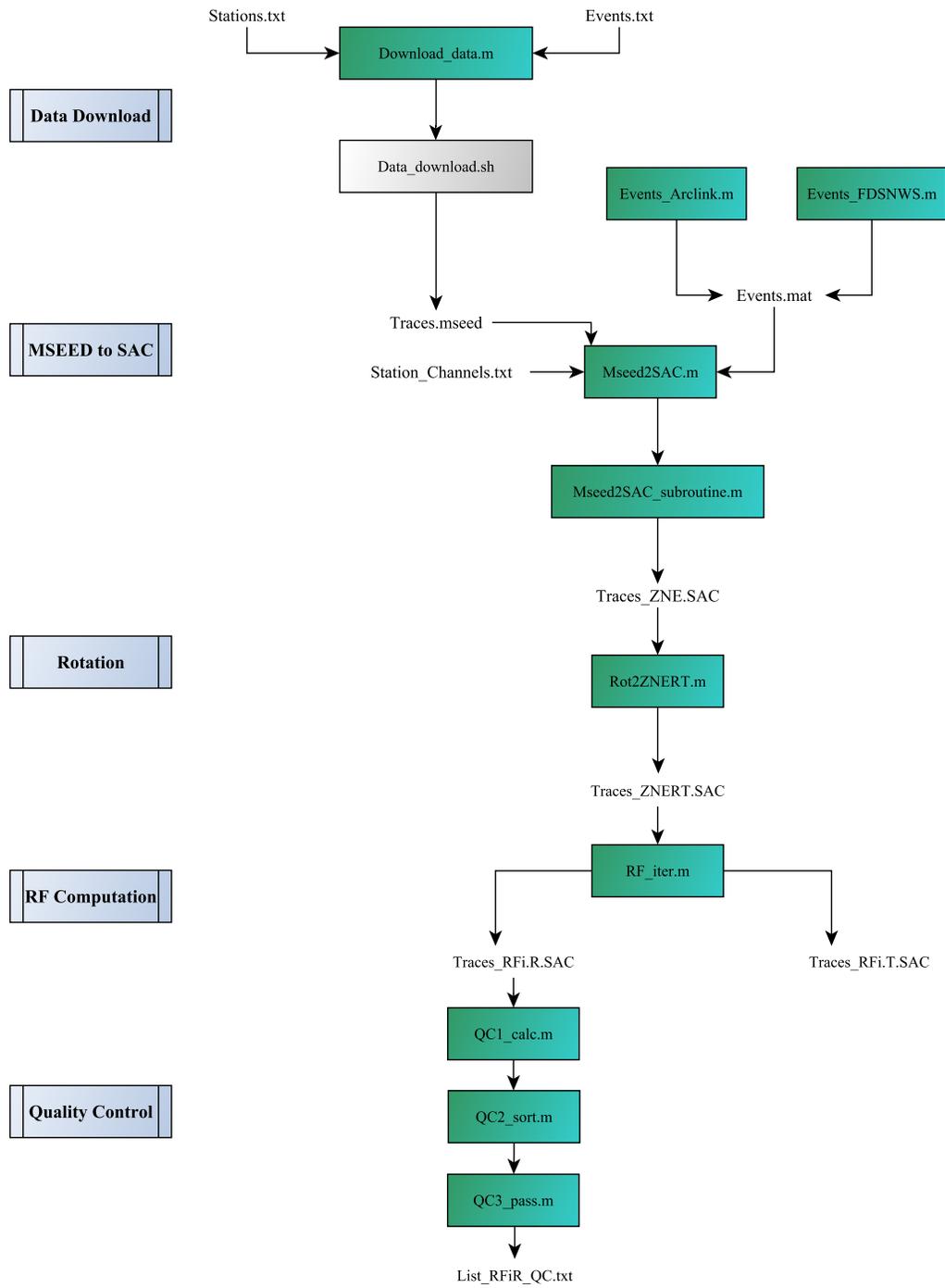


Figure 2.3.3: Flowchart of data preprocessing used in this work.

2.4 H- κ Method

One of the primary goals of this study is to determine the first-order discontinuities and define the main properties of the Earth's lithosphere. For this reason, in order to have a first idea of crustal thicknesses and average crustal V_p/V_s , I applied the simple stacking algorithm proposed by Zhu and Kanamori (2000) in our area. This technique has been extensively used in several seismic investigations (e.g. Nair et al., 2006; Lombardi et al., 2008; Piana Agostinetti and Amato, 2009; Licciardi et al., 2014).

2.4.1 Basis of the Zhu-Kanamori technique

Through a grid search in the H - V_p/V_s domain, I look for the values of H and V_p/V_s that maximize the stacking function:

$$S(H, V_p/V_s) = \sum_{j=1}^N w_1 r_j(t_1) + w_2 r_j(t_2) - w_3 r_j(t_3), \quad (2.4.1)$$

where N is the total number of RFs for that station, r_j is the j^{th} observed radial RF at times t_1, t_2, t_3 , the predicted arrival times for phases Ps, PpPs and PpSs + PsPs, respectively, calculated using the current values of $(H, V_p/V_s)$, while w_1, w_2 and w_3 are the weighting factors associated with each phase.

In the RF stacking algorithm proposed by Zhu and Kanamori (2000), the existing trade-off between crustal thickness and velocity is strongly reduced by including in the analysis additional multiple phases PpSs + PsPs. In this way, more robust estimates of crustal thickness and V_p/V_s -ratio can be obtained.

Although $H - \kappa$ method is broadly used and its implementation is straightforward to process automatically a large number of RFs and the possibility to estimate uncertainties associated with the results, it has several drawbacks:

- one has to assign *a priori* value for V_p ;
- the weights w_1, w_2, w_3 , associated with the amplitudes Ps, PpPs and PpSs + PsPs, are often chosen *ad-hoc*;
- it works well in regions with nearly flat discontinuities and sharp velocity jumps.

I employ this method in our study area to compare with the results of Lombardi et al. (2008) using the same approach.

2.4.2 Preliminary results in the Central Alps

I applied the $H - \kappa$ method to the stations (Figure 2.4.1) that are in the reference study performed by Lombardi et al. (2008).

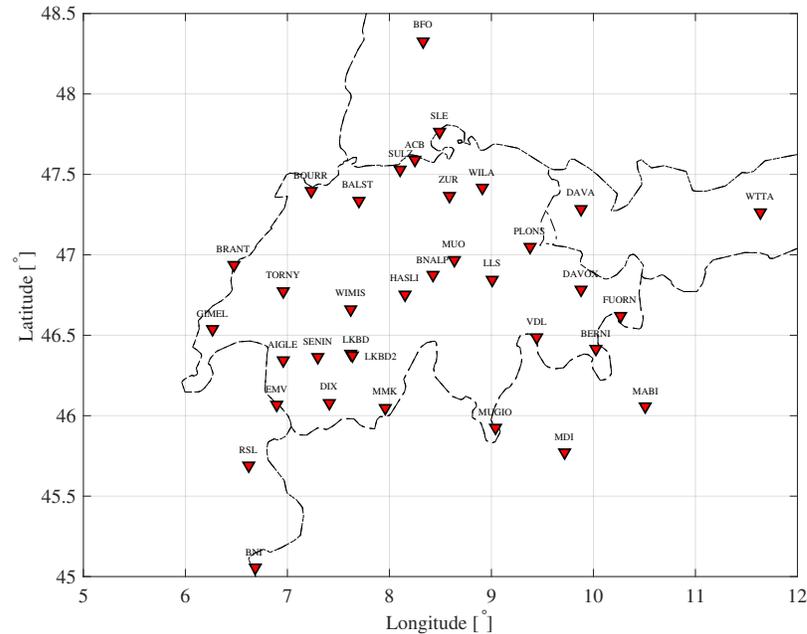


Figure 2.4.1: Stations in which I applied the $H - \kappa$ technique. Country borders are plotted with dashed lines.

To be consistent with this previous study, I compute RFs in a frequency band between 1 and 10 Hz and I adopt the same quality control procedure I described before. An example of stacked RFs is shown in the Figure 2.4.2 for the station HASLI (Hasliberg, Switzerland), which is in the center of our study zone.

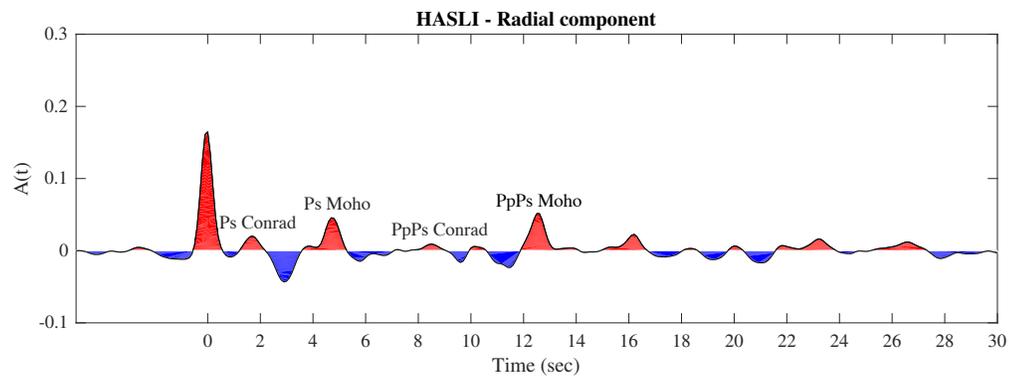


Figure 2.4.2: Radial component of RFs stack for the station HASLI (Hasliberg, Switzerland).

From the radial component, we can observe that the P_S *Moho* is 5 sec, while PpSs multiple phases appear between 10 and 14 sec. Figure 2.4.3 shows the back-azimuthal variation of RFs for the same station.

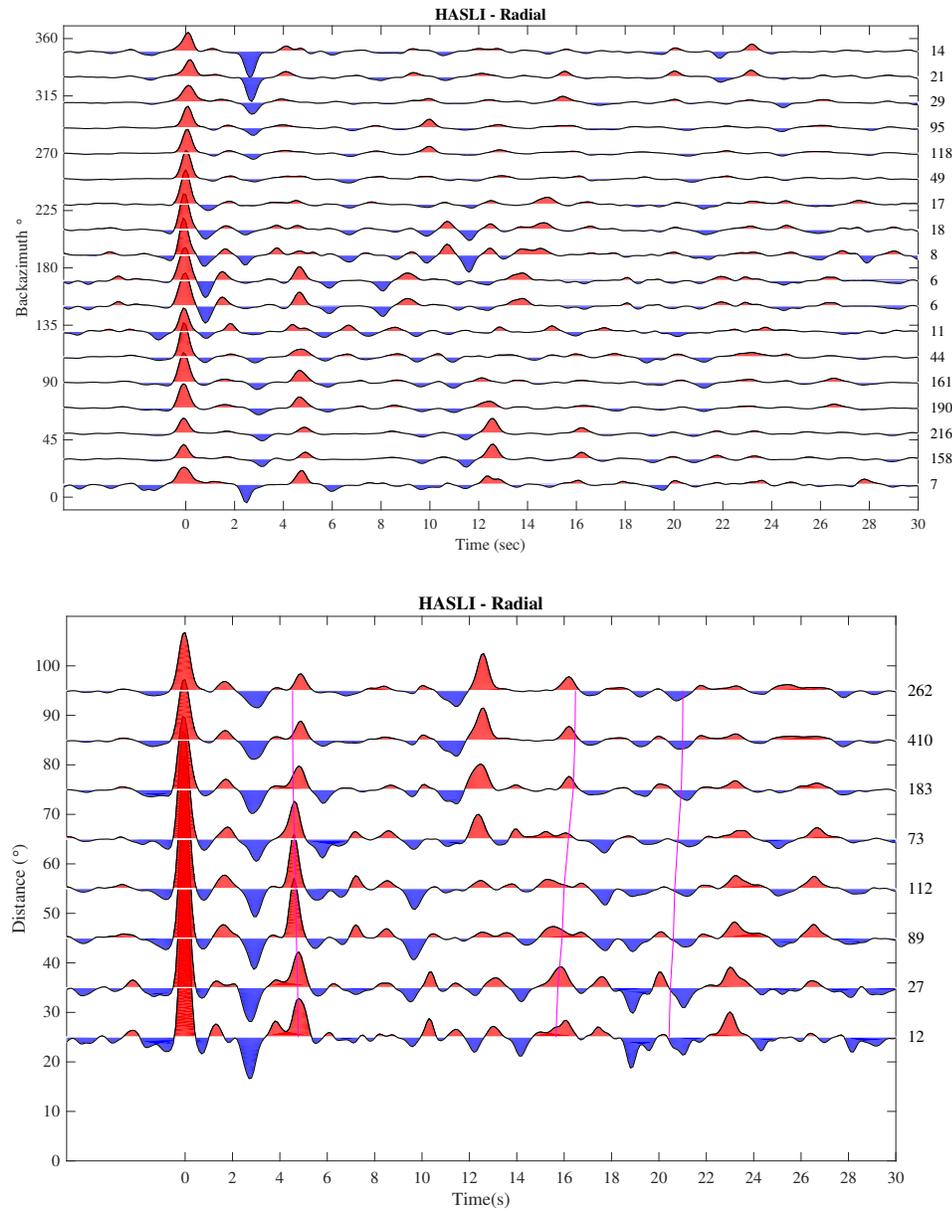


Figure 2.4.3: Above: examples of receiver function at the station HASLI plotted as a function of backazimuth where the RFs are stacked in bins of 20° . Below: RFs plotted as a function of epicentral distance. Red is used for drawing positive amplitude (increase of velocity with depth), blue for negative, numbers on the right are the number of RFs in a certain range of backazimuth or epicentral distance.

To compute the $H - \kappa$ result below each station, I selected only the events with magnitude $M > 6.0$, epicentral distance between 30° and 90° , assuming a crustal velocity of $V_p = 6.1$ km/s and I fixed the same weights for Ps and the multiples at 0.3 (same scheme as in Lombardi et al., 2008). The $H - \kappa$ stacking results for station MUO is shown in Figure 2.4.4. The identification of the maximum peak is not always univocal and often also has a considerable uncertainties, especially in the recognition of the V_p/V_s ratio.

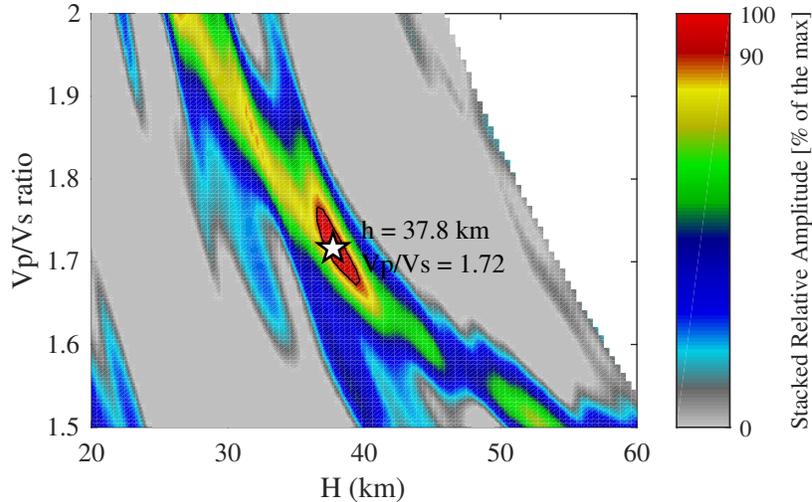


Figure 2.4.4: $H - \kappa$ stacking results for station MUO. Stacking function for a fixed crustal $V_p = 6.1$ km/s. Color scale represents the amplitude of the stacking function expressed as a percentage of the absolute value. White star represent the most likely peak in the $H - \kappa$ grid. Uncertainty is defined at 90% of the stacking function.

In this study, we have not considered the effects of seismic anisotropy nor the dipping interfaces on the back-azimuthal variation of RFs, although several studies both theoretically (Crampin, 1977; Levin and Park, 1997; Levin and Park, 1998) and by field investigations (Peng and Humphreys, 1997; Savage, 1998; Eckhardt and Rabbel, 2011; Ozacar and Zandt, 2004; Lucente et al., 2005) have shown the importance of these parameters in the RF analysis. In fact, in many cases the inversion process shows a trade-off between variable parameters, for instance, between layer dip and tilt of the anisotropic symmetry axis or between layer thickness and the type and amount of anisotropy (Baker et al., 1996; Savage et al., 2007; Li et al., 2007).

However, dipping interfaces are included in the new tool and are part of the model parameterization and ray tracing procedures are outlined in the following chapters.

2.4.3 Comparison with Lombardi et al., 2008

I applied the Zhu-Kanamori technique for the stations showed in Figure 2.4.1 and then I compared the results with the ones obtained by Lombardi et al. (2008). Results for Moho thickness are shown in Figure 2.4.5.

In both studies a general trend is recognizable, with some exceptions. In fact, I found high Moho depth values (≥ 35 km) in the zone inside the orogenic arc, and lower crustal thickness outside this area. This reflects the principle of isostasy and the presence of a crustal roots beneath collisional orogens (Fischer, 2002). Results for the average crustal V_p/V_s ratio are represented in Figure 2.4.6.

In the study zone, V_p/V_s values have a wide variability and with this technique we are not able to distinguish clearly the different geological domains (European basement, Mesozoic cover, Molasse, European margin and Austroalpine units). Main differences (Table 2.2) we found at each station are explained by the fact that in the work of Lombardi et al. (2008) is taken into account the correction of the local Moho dip.

To step forward from 1-D approaches and to take advantage of the new, dense station coverage with more data, I attach the 3-D problem for the construction of a velocity model that takes into account the lateral variations of geometries and velocities in the Earth's lithosphere.

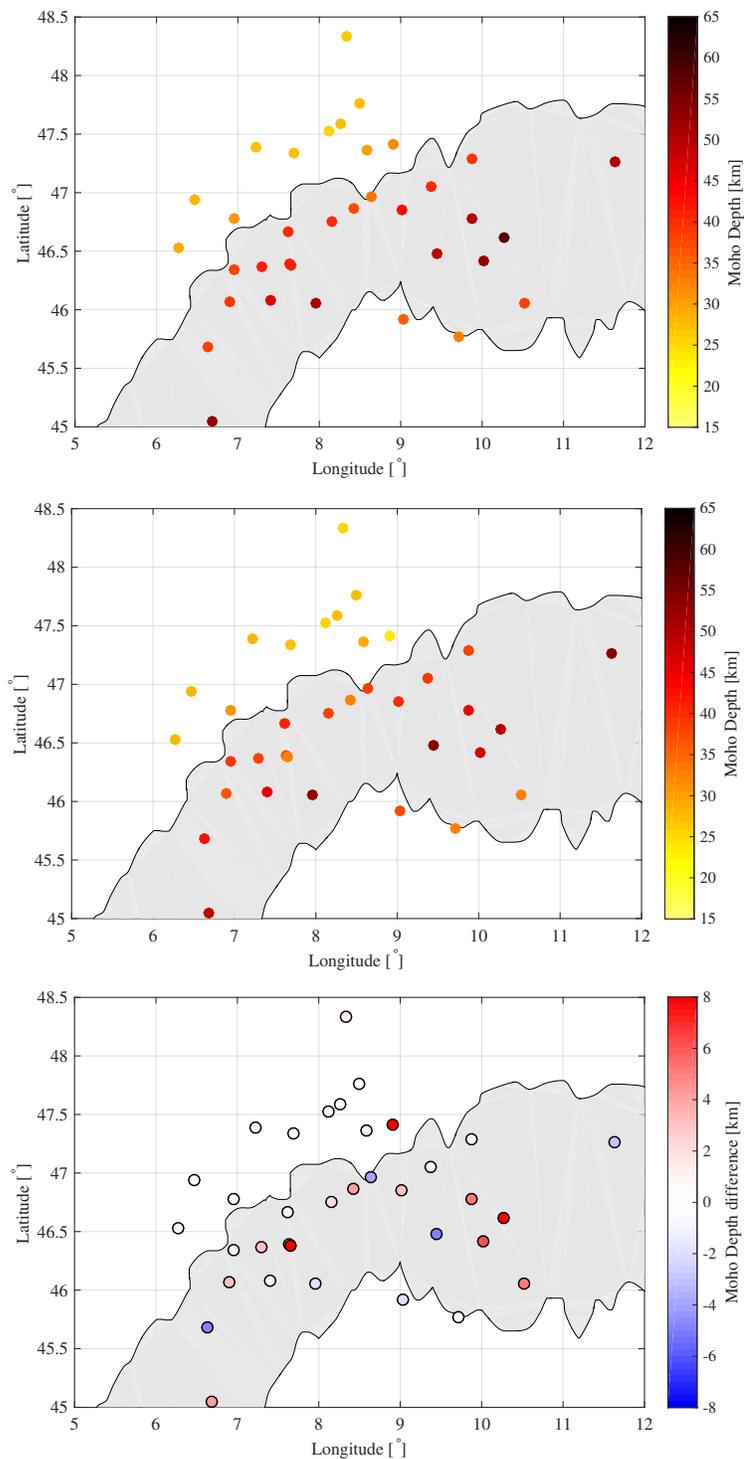


Figure 2.4.5: Comparison between Moho depth values obtained by $H - \kappa$ grid search from the work of Lombardi et al., 2008 (top), my work (center) and difference (bottom). Filled circles represent the Moho thickness below each station; in grey is shown the smoothed 800 m altitude line of the Alps. See Table 2.2 for details and uncertainties.

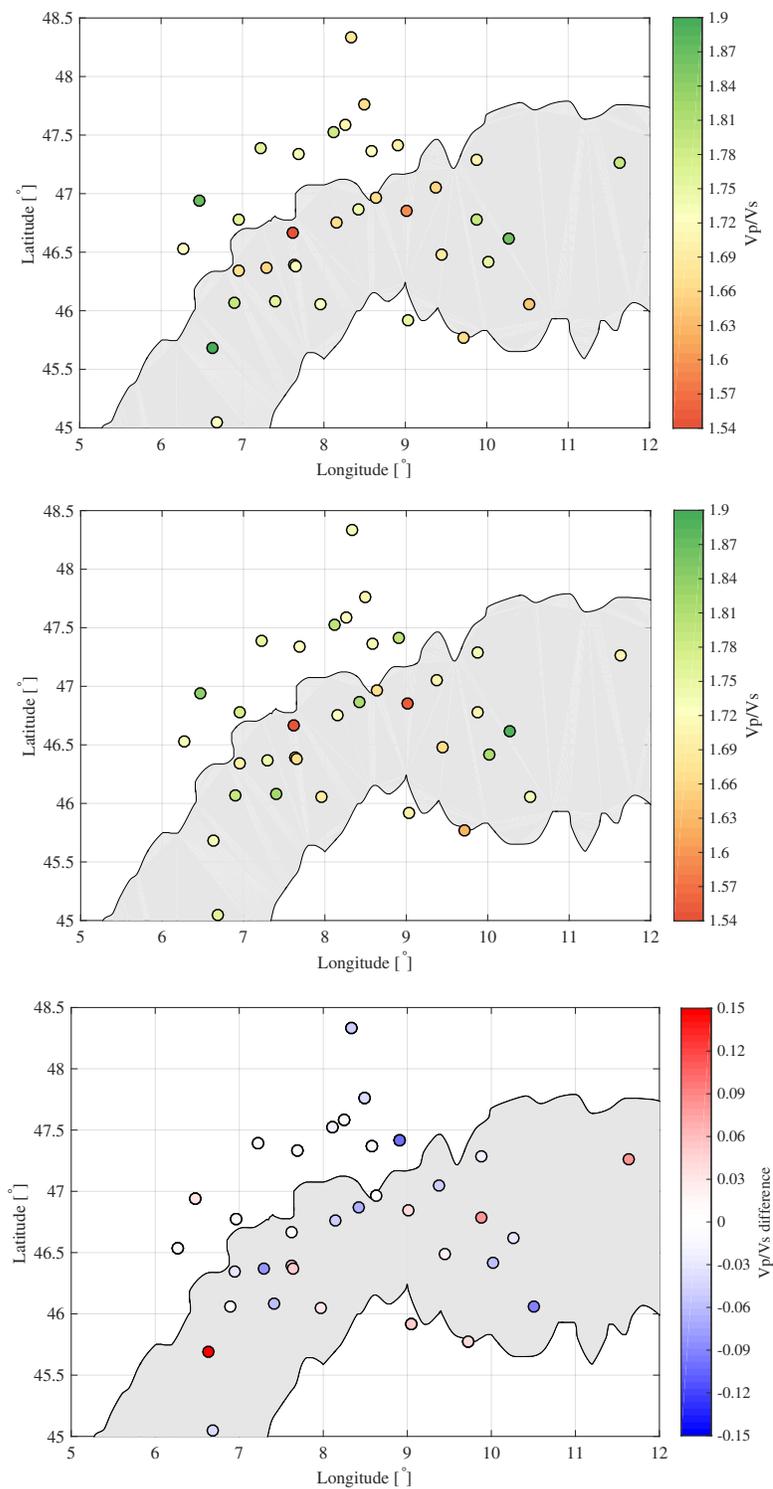


Figure 2.4.6: Comparison between V_p/V_s -ratio of the entire crust obtained by grid search from the work of Lombardi et al., 2008 (top), my work (center) and difference (bottom). Filled circles represent the V_p/V_s -ratio below each station; in grey is shown the smoothed 800 m altitude line of the Alps. See Table 2.2 for details and uncertainties.

Station	Numb. RFs	Numb. RFs	Moho (km)	Moho (km)	Vp/Vs	Vp/Vs
ACB	32	375	27±4	27±4	1.71±0.05	1.72±0.05
AIGLE	90	1433	38±5	39 ⁺⁵ ₋₃	1.67±0.09	1.70 ^{+0.05} _{-0.07}
BALST	91	1635	27±4	27±3	1.73±0.05	1.72±0.05
BERNI	90	2068	53±7	47±4	1.75±0.06	1.81±0.09
BFO	112	653	26±4	25±2	1.68±0.05	1.73±0.06
BNALP	83	1292	37±5	33±3	1.74±0.07	1.81 ^{+0.03} _{-0.09}
BNI	66	628	53±5	49±2	1.72±0.05	1.76±0.05
BOURR	72	1380	27±4	28±2	1.76±0.05	1.75±0.05
BRANT	32	980	28±4	28±3	1.87±0.06	1.84 ^{+0.05} _{-0.07}
DAVA	102	180	39±4	38±3	1.71±0.06	1.73±0.05
DAVOX	65	1554	50±7	45±5	1.79±0.06	1.71±0.08
DIX	79	1483	46±6	46±2	1.76±0.10	1.82±0.05
EMV	89	914	39±4	36±2	1.79±0.06	1.79±0.04
FUORN	103	1599	58±8	50 ⁺⁷ ₋₃	1.86±0.09	1.89±0.09
GIMEL	59	950	29±4	28±3	1.72±0.07	1.73±0.10
HASLI	105	1726	40±4	38±3	1.67±0.05	1.72±0.06
LKBD	22	320	41±4	37±5	1.71±0.06	1.66±0.06
LKBD2	29	182	41±5	33±6	1.72±0.09	1.67±0.06
LLS	132	1643	43±4	40±5	1.59±0.05	1.55±0.05
MABI	23	568	38±5	33 ⁺⁶ ₋₃	1.64±0.05	1.73 ^{+0.06} _{-0.10}
MDI	29	398	33±4	33±2	1.67±0.07	1.63±0.04
MMK	85	1640	51±7	53±2	1.72±0.08	1.69±0.05
MUGIO	56	1576	35±5	37±2	1.75±0.06	1.70±0.06
MUO	97	1883	34±5	38±5	1.67±0.06	1.67±0.05
PLONS	107	1739	40±4	39±2	1.66±0.05	1.71±0.05
RSL	32	351	38±5	43±3	1.90±0.07	1.73±0.07
SENIN	64	1163	41±4	38±2	1.66±0.06	1.74±0.04
SLE	73	1487	27±4	27±4	1.67±0.05	1.71±0.05
SULZ	82	1724	25±4	25±3	1.78±0.06	1.80±0.07
TORNY	58	1309	31±4	31±3	1.76±0.05	1.77±0.05
VDL	113	1829	49±6	54±5	1.69±0.06	1.67±0.10
WILA	63	1532	32±4	24±8	1.70±0.05	1.80±0.10
WIMIS	51	1238	40±4	41±5	1.54±0.05	1.54±0.05
WTTA	67	203	51±5	54±3	1.79±0.07	1.71 ^{+0.10} _{-0.02}
ZUR	74	1526	30±4	29±5	1.72±0.04	1.73±0.05

Table 2.2: Summary table of comparison between our results and those of Lombardi et al. (2008) using a grid search method. Results obtained in this work are the ones in the 3rd, 5th and 7th column, in red. Uncertainty is defined at 90% of the relative stacked function.

Chapter 3

3-D ray tracing in heterogeneous media

Forward modeling in geophysics is a fundamental operation that governs the equations relating the model and the observation. Here I describe the ray tracing procedure I implement in the Alpine context.

At the beginning of the chapter I provide an overview of a few approaches used in seismology for ray tracing, from the shooting method to grid based schemes. Then I illustrate step-by-step the local 3-D implementation I apply based on an existing global tool in 1-D considering a local Moho model (Spada et al., 2013) and a local 3-D P-velocity model (Diehl et al., 2009) for the Central Alps. In the construction of an accurate tool for the propagation of a P-to-S converted wave, I focus on the P-wave shooting, the interface discretization which respects Snell's law in 3-D and takes into account the Moho local dip, the S-wave shooting and the correction needed to reach accurate results. In the last part, I illustrate through some images the goodness and reliability of the shooting method.

3.1 Shooting schemes

Although 1-D Earth models are an acceptable approximation in some applications, lateral heterogeneity is significant in many regions of the Earth, like the Alpine area.

In the past few decades, the growing need for fast and accurate prediction of high frequency wave properties in 2-D and 3-D media has spawned a prolific number of grid and ray based solvers. One of the most common method has been ray tracing (Julian, 1977; Virieux and Farra, 1991; Cerveny, 2005), in which the trajectory of paths corresponding to wavefront normals are computed between two points. This approach is often highly accurate and efficient and naturally lends itself to the prediction of various seismic wave properties.

We can group ray tracing methods in two broad categories: shooting and bending.

3.1.1 Shooting method

In the *shooting methods*, the ray path is computed applying Snell's law at the presence of any interface. Shooting methods of ray tracing are widely used in seismology due to their conceptual simplicity and for high accuracy and efficiency.

The two point problem of finding a source-receiver path becomes an inverse problem in which the unknown is the initial direction of the ray and the function to be minimized is a measure of the distance between the ray end point and the receiver (Rawlinson et al., 2008). The main challenge that this class of methods faces is the non-linearity of the inverse problem, which increases proportional by the complexity of the medium.

The appropriate form of the equation required to solve the initial value problem depends on the choice of parameterization used to represent velocity variations. For example, in a medium described by constant velocity blocks, the ray path is described by a piecewise set of straight line segments; all that is required to solve the initial value problem is a repeated application of Snell's law at cell boundaries.

Shooting methods of ray tracing usually solve the boundary value problem by probing the medium with initial value ray paths and then exploiting information from the computed paths to better target the receiver. Figure 3.1.1 represents the basic concept in 2-D.

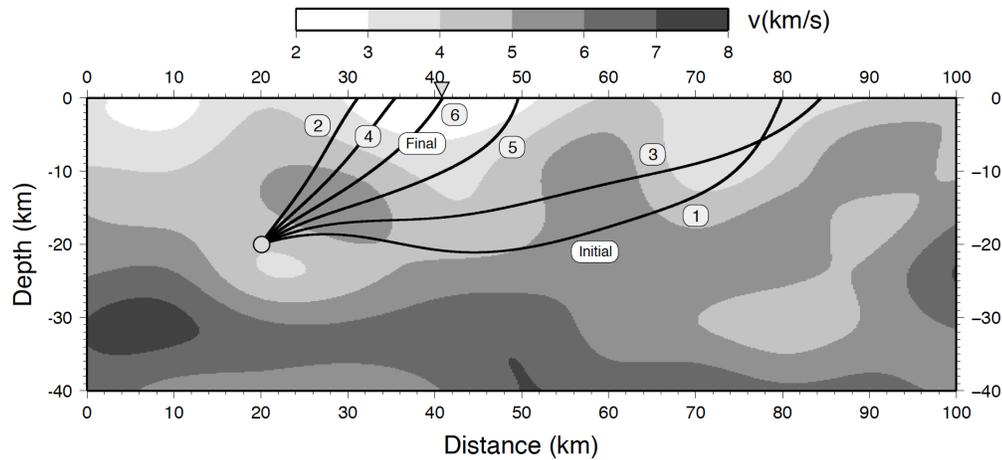


Figure 3.1.1: Principle of the shooting methods. The initial trajectory is updated until it converges at the receiver (from Rawlinson et al., 2008).

3.1.2 Bending method

The principle of the *bending methods* of ray tracing is to adjust iteratively the geometry of an initial arbitrary path that joins source and receiver until it becomes a true ray path, i.e., it satisfies Fermat's principle of minimum time, as shown in Figure 3.1.2.

A common approach to implement the bending method is to derive a boundary value formulation of the kinematic ray tracing equations which can be solved iteratively (Rawlinson et al., 2008).

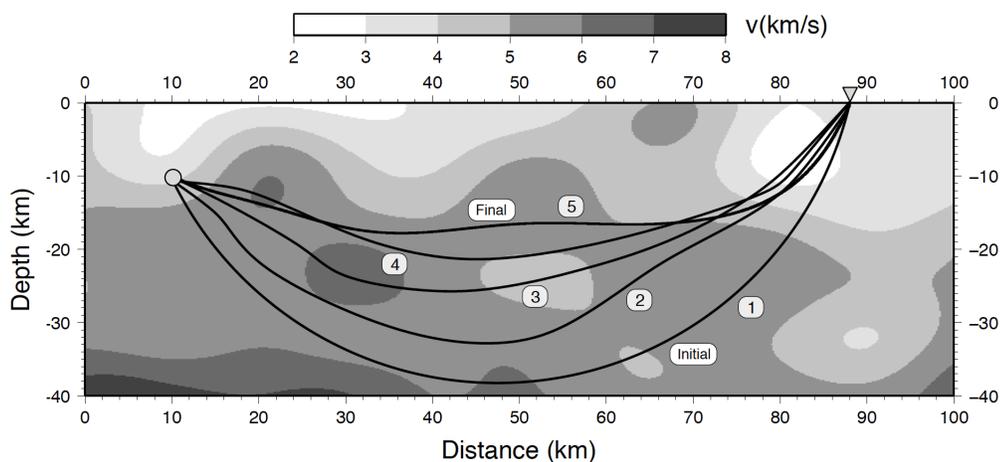


Figure 3.1.2: Principle of the bending method. In this case, an initial two point path is perturbed until it satisfies Fermat's principle (from Rawlinson et al., 2008).

A slightly different approach is represented by *pseudo-bending* method which avoid direct solution of the ray equations (Rawlinson et al., 2008). One of the most common pseudo-bending schemes is based upon the ray path being represented by a set of linearly interpolated points. Given some initial arbitrary path, the aim of the pseudo-bending method is to adjust sequentially the location of each point so that the path better satisfies the ray equations.

Um and Thurber (1987) apply the update scheme simultaneously from both end points of the ray path to the central point. This process is repeated until a convergence criterion is met. An initial path can be approximated in various ways, but one simple option is to begin with a three point ray joining source and receiver. Once the central point has been perturbed, two new points are introduced that bisect each line segment.

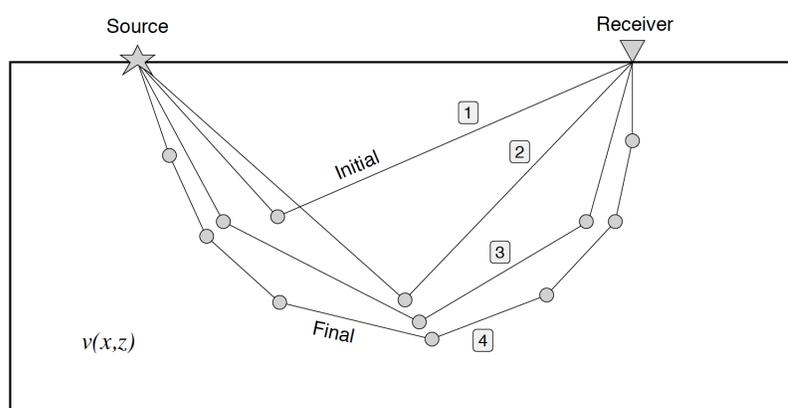


Figure 3.1.3: Principle of the pseudo-bending method (from Rawlinson et al., 2008).

In the schematic example of Figure 3.1.3, an initial ray is defined by three points. The center point is perturbed to satisfy Fermat's principle of stationary time. The number of path segments is then doubled and the process repeated until a suitably accurate path is obtained.

3.1.3 Grid based schemes

An alternative strategy to trace rays between source and receiver is to compute the traveltimes of the wavefront at all points of a grid which spans the medium (Rawlinson et al., 2008). Compared to conventional shooting methods, grid based traveltimes schemes have a number of clear advantages:

1. Most are capable of computing traveltimes to all points of a medium, and will locate diffractions in ray shadow zones;

2. The non-linearity of both ray shooting and bending means that they may fail to converge to a true two-point path, whereas most grid based schemes are highly stable and will find the correct solution even in strongly heterogeneous media;
3. Grid based schemes can be very efficient in computing traveltimes and path information to the level of accuracy required by practical problems. Ray tracing schemes can be inefficient if solution non-linearity is significant;
4. Most grid-based schemes consistently find first-arrivals in continuous media

Despite these advantages, grid based schemes have a number of limitations that should be considered prior to application. These include:

1. Accuracy is a function of grid spacing in 3-D halving the spacing of a grid will increase computation time by at least a factor of 8. Thus, computation time may become unacceptable if highly accurate traveltimes are required;
2. Most practical schemes compute first-arrivals only - thus, features such as wavefront triplications cannot be predicted;
3. Quantities other than traveltimes (such as amplitude) are difficult to compute accurately without first extracting path geometry and applying ray based techniques.

3.1.3.1 Fast Marching Method

Despite these advantages, grid based schemes have several limitations. These include that the accuracy is a function of grid spacing: in 3-D, halving the spacing of a grid will increase computation time by at least a factor of 8. Thus, computation time may become unacceptable if highly accurate traveltimes are required.

One of the more recently developed grid based schemes is the so called *Fast Marching Method* (Sethian, 1999; Popovici and Sethian, 2002). The term Fast Marching Method, usually abbreviated with FMM, was originally developed in the field of computational mathematics for solving various types of interface evolution problems.

In seismology, FMM has been used in the migration of coincident reflection profiles (Popovici and Sethian, 2002) and teleseismic tomography (Rawlinson et al., 2006a; Rawlinson et al., 2006b). A common feature of first-arrival traveltimes fields is that they are not spatially differentiable at every point. The FMM stencil describes how to calculate new traveltimes using known traveltimes from neighbouring grid points, but in order to populate a grid of nodes with traveltimes values, the order in which nodes are updated must be consistent with the direction of flow of information. FMM achieves this by systematically constructing

traveltimes in a downwind fashion from known values upwind by employing a *narrow band* approach. This is illustrated in Figure 3.1.4.

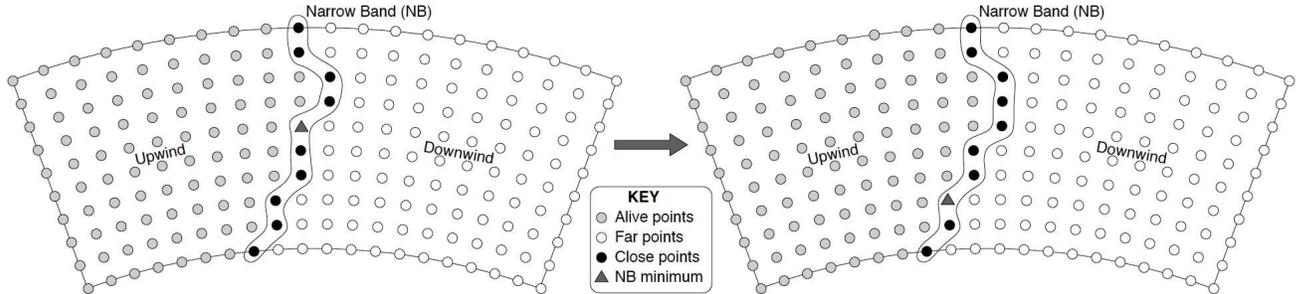


Figure 3.1.4: Narrow band evolution scheme used by the FMM for the ordered update of grid points (left). The narrow band advances from the close point with minimum traveltime (right). Figure from Rawlinson et al., 2008.

Alive points have their values correctly computed, *close* points lie within the narrow band and have trial values with *alive* points only, and *far* points have no values calculated. The narrow band is evolved by identifying the close point with minimum traveltime, tagging it as alive and then updating any adjacent close or far point, the latter being re-tagged as close. Using this approach, the shape of the narrow band approximates the shape of the first-arrival wavefront, and the idea is to propagate the band through the grid until all points become alive. Further FMM examples and detailed analysis of computational efficiency can be found in Rawlinson and Sambridge (2004a) and Rawlinson and Sambridge (2004b).

3.1.3.2 Shortest path ray tracing

Shortest path ray tracing or *SPR* (Nakanishi and Yamaguchi, 1986; Moser, 1991; Cheng and House, 1996) is another common method for determining first-arrival traveltimes of a velocity field. Rather than solving a differential equation, a network is formed by connecting neighbouring nodes with traveltime path segments.

The shortest time path between two points corresponds to a true ray path. Shortest path networks are commonly defined in terms of either a cell or a grid centered framework. A particular way of creating a network is to use a regular grid of velocity nodes and to form linear connections between adjacent velocity nodes (Moser, 1991), as shown in Figure 3.1.5.

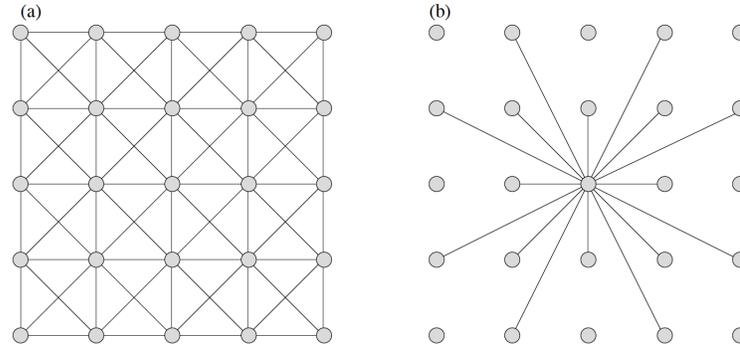


Figure 3.1.5: Shortest path network built on a grid of velocity modes (Moser, 1991). **a)** Use of a stencil with at most 8 connections per node; **b)** 16-node stencil, which allows smaller path deviations. Figure from Rawlinson et al., 2008.

The advantage of a grid formalism for the shortest path network is that varying velocity fields can be more accurately represented and interfaces are more easily inserted. Once a network structure and method of traveltime determination between node pairs has been chosen, the next step is to use a shortest path algorithm to compute the full traveltime field and associated ray paths.

The basic approach is conceptually simple, with nodes divided into two groups: a set of P nodes with known traveltimes, and a set of Q nodes with unknown traveltimes Moser (1991). Initially, Q contains M elements and P is empty, with the traveltimes of nodes in Q set to an arbitrary large value. The scheme is initiated from a source node by adding it to P , and then computing traveltimes to all neighbouring nodes as defined by the forward star (Figure 3.1.6).

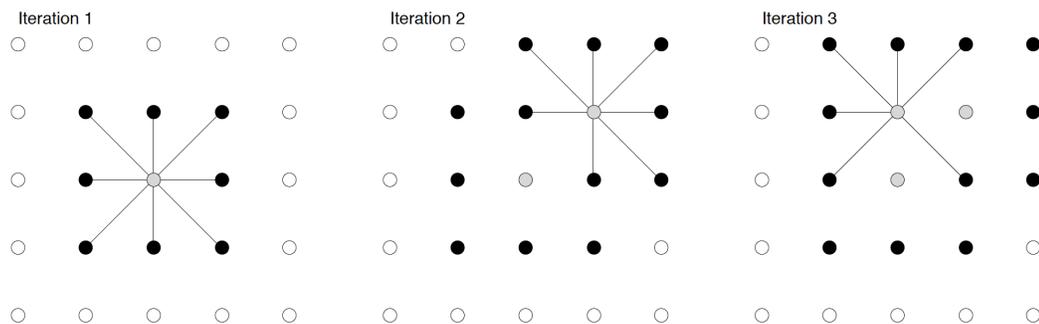


Figure 3.1.6: Three iterations of a simple shortest path scheme using an initial star with 8 connections Grey dots have known traveltimes, black dots have trial traveltimes and white dots are yet to have traveltime computed. Figure from Rawlinson et al. (2008).

These constitute trial traveltime values, and the method proceeds by choosing the trial

node with minimum traveltime, adding it to P and then computing trial times to all neighbouring nodes in Q. If the node in Q already has a trial value from a previous iteration, then the one with minimum traveltime is selected. The complete traveltime field is found in exactly M iterations, and ray paths are obtained by simply recording the update order of nodes.

In this work, I decide to use a shooting method of ray tracing because is relatively simple to implement and it shows a high accuracy.

3.2 Ray Tracing Implementation in this work

P-to-S ray tracing in 3-D represents a very delicate operation as it is necessary to take into account the lateral variations of the velocity structures and the local dip of the discontinuities in order to build an accurate tool of ray propagation for converted waves. The scheme I follow is summarized in the flowchart of the Figure 3.2.1.

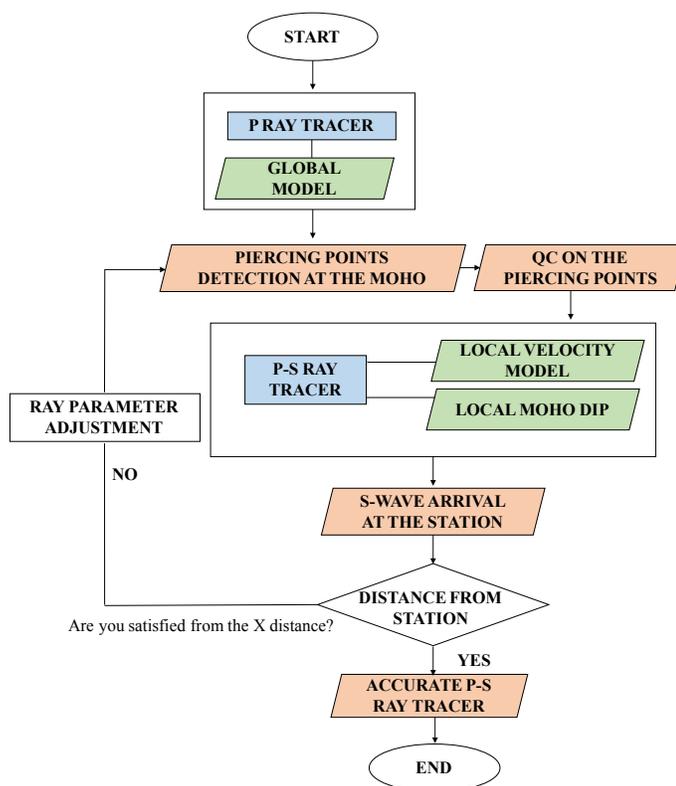


Figure 3.2.1: Flowchart used for the ray tracing procedure I implement in this work. Each block corresponds to a script I develop in my tool. Start and end points are represented by an oval, parallelograms represent an input (green) or an output (red), rectangles represent a process, a diamond indicates a decision.

First, I do a P-ray propagation from a source (in this study represented by an earthquake) to a station in 1-D spherical global velocity model. For this propagation I use the MatLab software *TTBOX* (Knapmeyer, 2004; Knapmeyer, 2005) and as input model the global velocity model *IASP91* (Kennett and Engdahl, 1991). After this operation, I go back along that ray considering the actual Moho dip (for the Alps, reference Moho model is given by Spada et al., 2013): in this I we obtain a map of the piercing point at the local Moho. Then, in order to simulate the P-to-S conversion, I go forward in the propagation to a station from that piercing point with a S-wave considering a local velocity model (for the Alps, reference velocity is given by Diehl et al., 2009). Finally I adjust the ray parameter to make the S-waves arrive at the station. All these steps are explained in detail in the sections below.

3.2.1 P shooting in a global velocity model

For propagation of the P-wave in a global velocity model, I decide to use a shooting scheme, implementing an existing software, called *TTBOX* (Knapmeyer, 2004; Knapmeyer, 2005), which is a useful and very efficient tool for the computation of teleseismic travel time in 1-D models. In the original version, this software presents several constraints, since ray paths and travel times are computed for a 1-D spherical isotropic velocity model. Under these assumptions, the propagation of seismic rays can be computed applying Snell's law, which describes ray geometry by seismic velocity and the ray parameter.

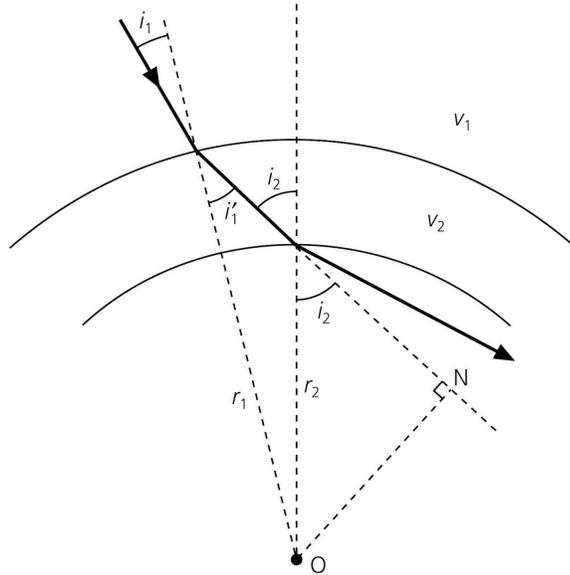


Figure 3.2.2: Geometry of Snell's law for spherical Earth (from Stein and Wysession, 2003).

Considering Figure 3.2.2, if v_1 and v_2 represent the velocities above and below r_1 and i_1 , i'_1 , and i_2 are the angles shown, by Snell's law:

$$\frac{r_1 \sin i_1}{v_1} = \frac{r_2 \sin i'_1}{v_2} \quad (3.2.1)$$

Because both equal the length ON, we can write:

$$\frac{r_1 \sin i_1}{v_1} = \frac{r_2 \sin i_2}{v_2} \quad (3.2.2)$$

Therefore, the ray parameter p for a spherical Earth is:

$$p = \frac{r_n \sin i_n}{v_n} \quad (3.2.3)$$

where

r is the radial distance from the center of the Earth,

v is the velocity of that point,

n is the number of the n^{th} discontinuity,

i is the incidence angle between the ray path and the radius vector.

3.2.1.1 TTBOX accuracy

TTBOX uses the flat Earth transformations to evaluate both epicentral distance and travel time as functions of ray parameter. This allows to replace a spherical symmetric model for the crust by a Cartesian one and avoid the complications of polar coordinates. The accuracy of the solution therefore depends on the quality of the *depth sampling* (Knapmeyer, 2005): in fact, velocity models may be defined by a list of depth samples or piecewise continuously by layer polynomials that are discretized into a list of samples before computing paths and times.

TTBOX supports two kinds of depth sampling: equidistant in spherical Earth and equidistant in flat Earth (Cartesian) coordinates; rays are computed using a *shooting method*.

The ray parameter needed to arrive at a given epicentral distance, and is determined by a three-stage optimization process.

1. The epicentral distance as a function of ray parameter is approximately computed by shooting lots of rays;

2. This approximation of the ray-parameter function Δp is checked for approximate solutions;
3. The solutions found at the step 2 are refined by a *binary search* algorithm, that finds the position of a target value within a sorted array.

A ray parameter is returned which allows hitting the target distance within $\pm 0.001^\circ$.

To speed up the time spent for running the P ray tracing, I compute a table which give us the ray parameter p knowing the depth and the epicentral distance (see Figure 3.2.3).

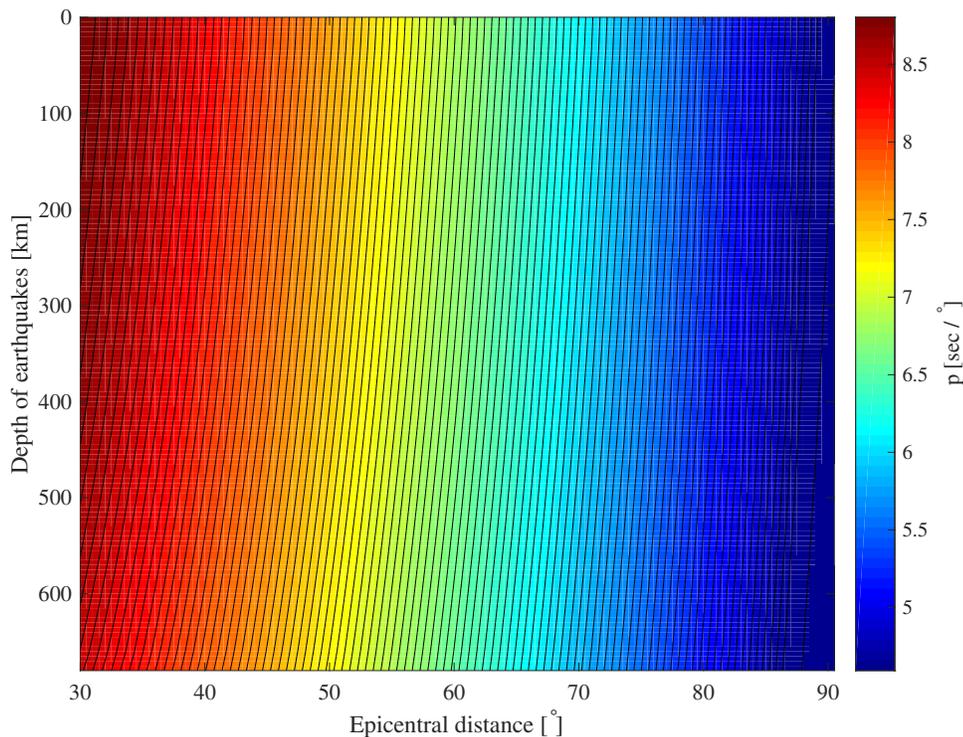


Figure 3.2.3: Pre-computed table generated with *TTBOX* of ray parameter p in function of depth and epicentral distance.

3.2.2 Local dip of Moho

Since the 1-D model is too simplistic as it is thought in the original version of the software, in order to move in a 3-D context, I develop my own ray tracing using the first P-to-S conversion in a local model. First, I need initial information about the Moho map in the Alps, which represents the main interface where the wave conversion takes place. For the application to the Central Alps, I decide to use the Moho map from the work of Spada et al. (2013). We choose this model that uses a combination of CSS and LET (Waldhauser et al., 1998;

Wagner et al., 2012) because it presents a new quality classification based on uncertainty of RF data. We preferred the Spada et al. (2013) model with respect to the one we computed in the second chapter derived by Zhu and Kanamori (2000) because the $H - \kappa$ has several limitations since it assumes a simple one-layer crustal model with a well-defined horizontal Moho interface and a fixed P-wave velocity values across the entire crust.

I convert the initial map in longitude and latitude (Figure 3.2.4) into a Cartesian system centered at 10° E, 46° N and in km towards E and N, respectively (Figure 3.2.4).

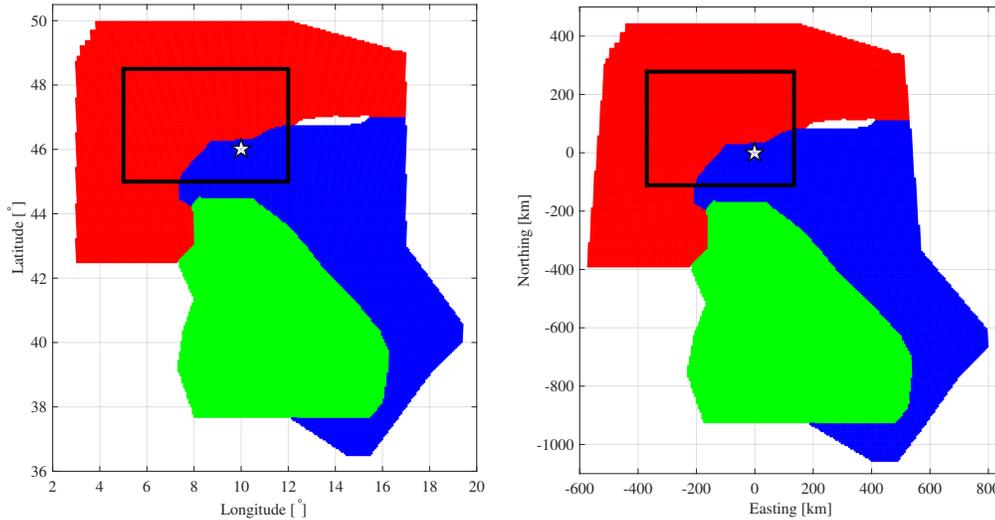


Figure 3.2.4: Spherical (left) and Cartesian (right) Moho map of Europe (red), Adria (blue) and Liguria (green) plates. The black rectangle is my area of investigation. The origin of the system (10° , 46°) is indicated with a white star. Plate division and Moho depth values taken from Spada et al. (2013).

In our area (5 - 12° E, 45 - 48.5° N), I take each element of the Moho map to define Delaunay triangles, in order to calculate the local Moho dip and its orientation (Figure 3.2.5).

We can observe that dip values vary from small values (0 - 5°) up to angles of about 20° at the Adria-Europe plate jump. Moreover, in the European plate, intermediate values of the dip (around 10 - 12°) seem to follow the shape of the Alpine arc. These figures illustrate the importance of taking into account the significant changes of dip within relatively small distances in the ray tracing phase. To speed up the calculation, for each station I consider a cylinder (Figure 3.2.6) and I reconstruct the local dip of each element of our initial Moho model.

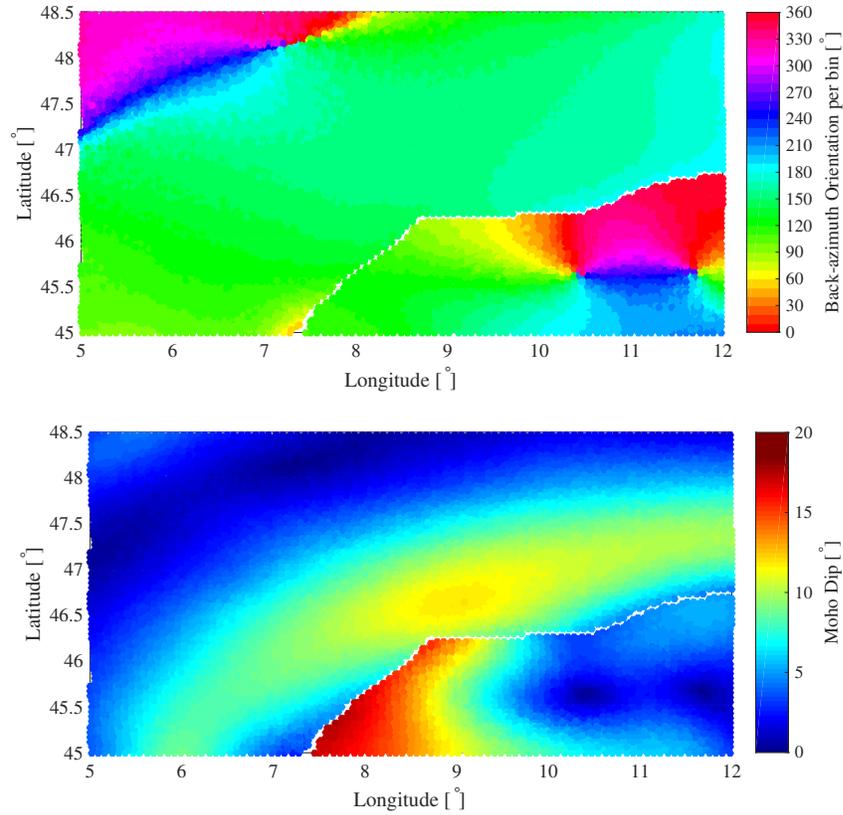


Figure 3.2.5: Moho dip orientation (top) and local dip (bottom) in the study area. The Europe-Adria plate boundary is clearly visible (white stripe).

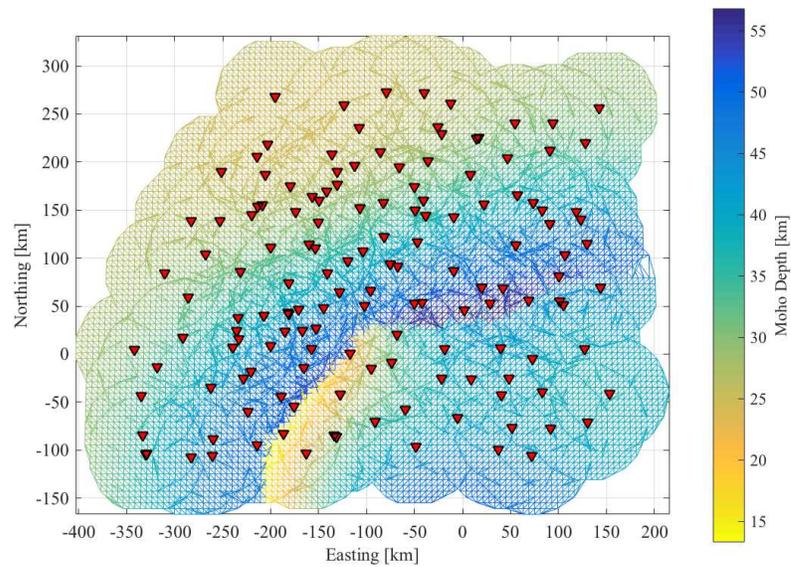


Figure 3.2.6: Map of the cylinders of Moho considered for the study area.

3.2.3 P-to-S piercing points detection

Since in the original version of the software, converted waves are not supported, I decide to compute the piercing points between the P-wave traveltimes and the discretized surface which represents the local Moho from Spada et al. (2013). Figure 3.2.7 (plan view in Figure 3.2.8) shows the global P rays arriving at the station ZUR (Zürich) which lies completely on European plate.

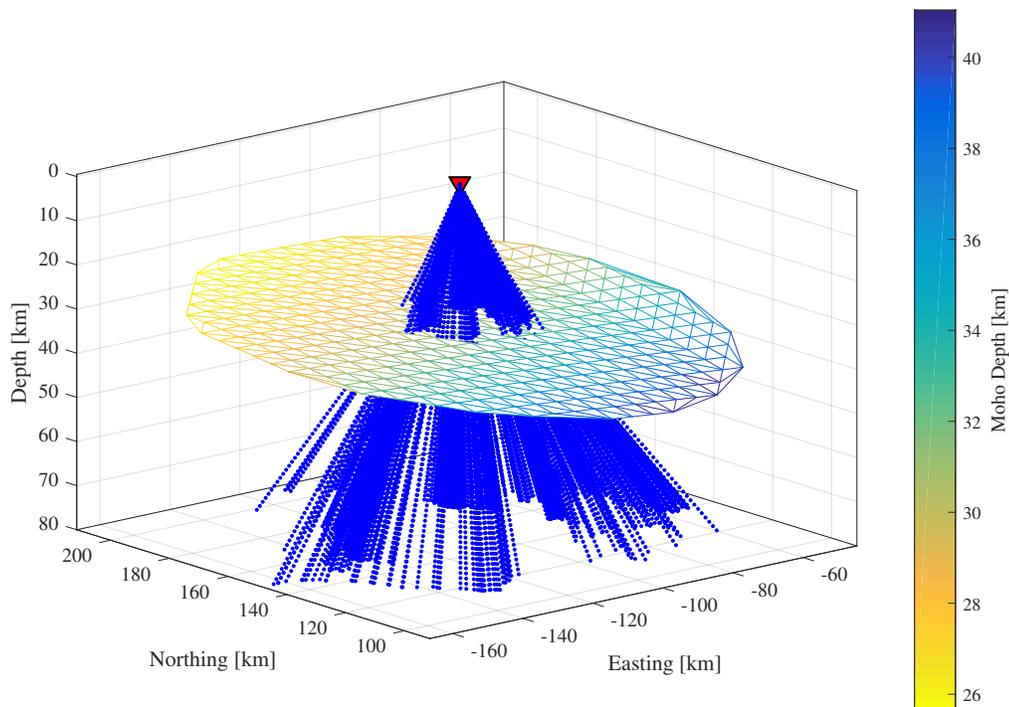


Figure 3.2.7: P-wave shooting considering a global velocity model for the station ZUR. The surface discretized in triangles represents the local Moho model from Spada et al. (2013).

To find the coordinates of the piercing points at the Moho in Cartesian space, I used the Möller-Trumbore ray-triangle intersection algorithm (Möller and Trumbore, 2005), which is a fast method for computing the intersection of a ray (our P-wave raypaths) and a triangle (the smallest element in which the Moho surface can be divided) in 3-D without needing pre-computation of the plane equation of the plane containing the triangle.

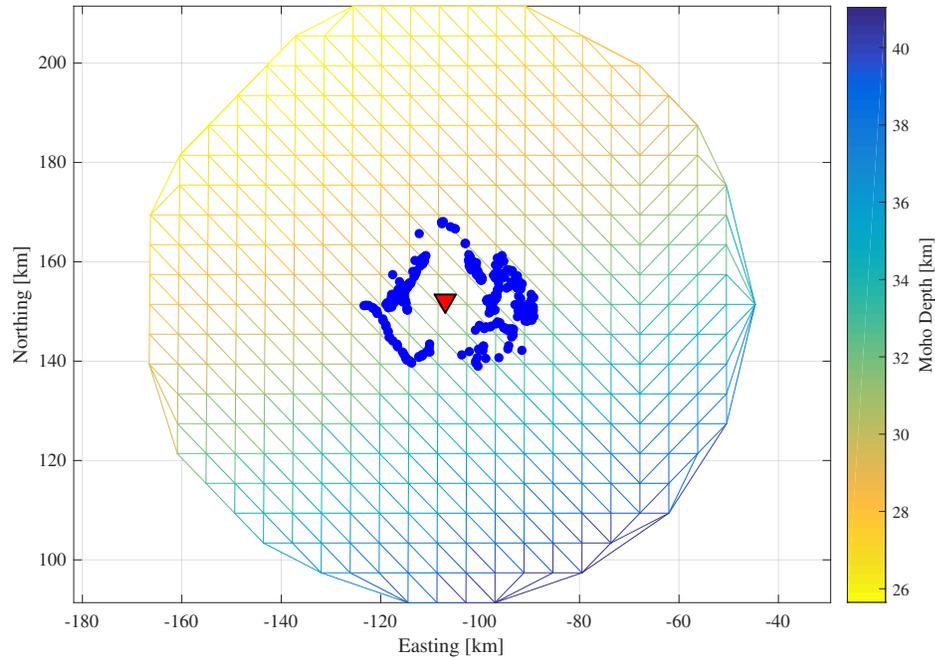


Figure 3.2.8: Piercing points at the local Moho after first shooting of P-to-S conversion for the station ZUR.

A limitation of our ray tracing tool is that while it takes into account the local dip of any discontinuities (Moho, Conrad) to properly simulate the shooting of P-to-S conversion, it does not follow the full ray-path of multiples. That would require a more complex ray propagator, which is feasible, and to consider that the amplitude of multiples need to be spread to three conversions or reflections, which is a separate, large, problem.

Given the piercing points, the next fundamental step is the construction of an S-wave propagator from the Moho interface to the surface.

3.2.3.1 Quality Control on the piercing points

As already mentioned above, the Central Alps represent a very complex area in terms of Moho geometry. For this reason, in areas at a plate boundary, as in the case of the station ABSI, some rays cross the Moho surface between the Adria and the Europe plates (Figure 3.2.9), in a “wall” that is physically not a Moho discontinuity.

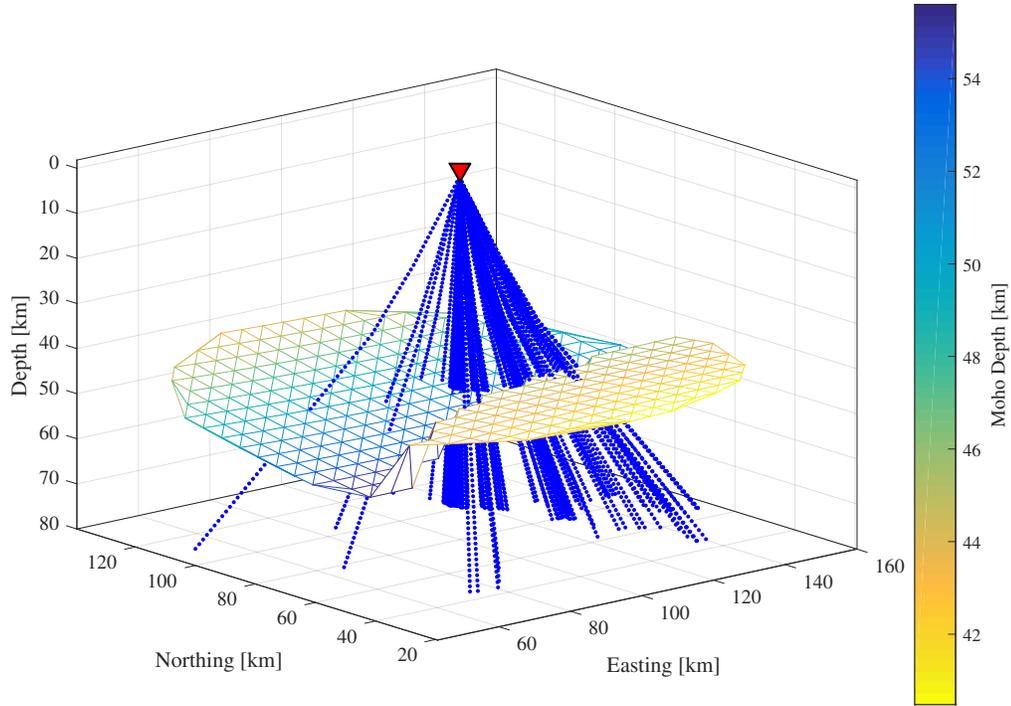


Figure 3.2.9: P-wave shooting considering a global velocity model for the station ABSI. The surface discretized in triangles represents the local Moho model from Spada et al. (2013). Note the plate jump that creates a “wall”, which does not represent a Moho discontinuity.

In such cases the intersection algorithm between the raypaths and the Moho surface, yields piercing points that don’t make sense since the surface inside the plate margin is only an effect of the point cloud interpolation. Piercing points that lie in such areas are eliminated, both in the first and in the corrected (second) shooting based on the local Moho dip compared to the maximum dip in our study area for any plate. This control on the piercing points reduced the dataset by about 15% (from 28 494 to 24 148 traces).

3.2.4 S shooting in a local velocity model

As the initial local shear-wave velocity structure, I took the 3-D P-wave velocity model from Diehl et al. (2009) and divided it by a classical value of $V_p/V_s = 1.73$. Then, given the velocities around the raypath, I used an interpolation with inverse squared distance weighting of the 4 enclosing neighbour points where we know the velocity to obtain the local V_s at the ray. In this way I am able to propagate an S-wave from the bottom to the top of the crust to reach the corresponding station accurately. Starting from the depth of the piercing point at the Moho and moving towards the surface, I take into account the nearest upper neighbour

plane in depth of Diehl et al. (2009) and then we consider staggered grids at the midpoint of the tomographic model. With this approximation, I expect to know velocities at the grid mid-points with respect to the tomographic work. For example, if we know velocities at 45, 30, 15, 0 km at depth from tomographic investigation, in our S-propagation I will have velocity values at 37.5, 22.5 and 7.5 km at depth (Figure 3.2.10). For the shallowest part, between surface and station elevation, I interpolate velocities in the layer between 0 and 5 km above the surface, which are also part of the tomographic grid Diehl et al. (2009).

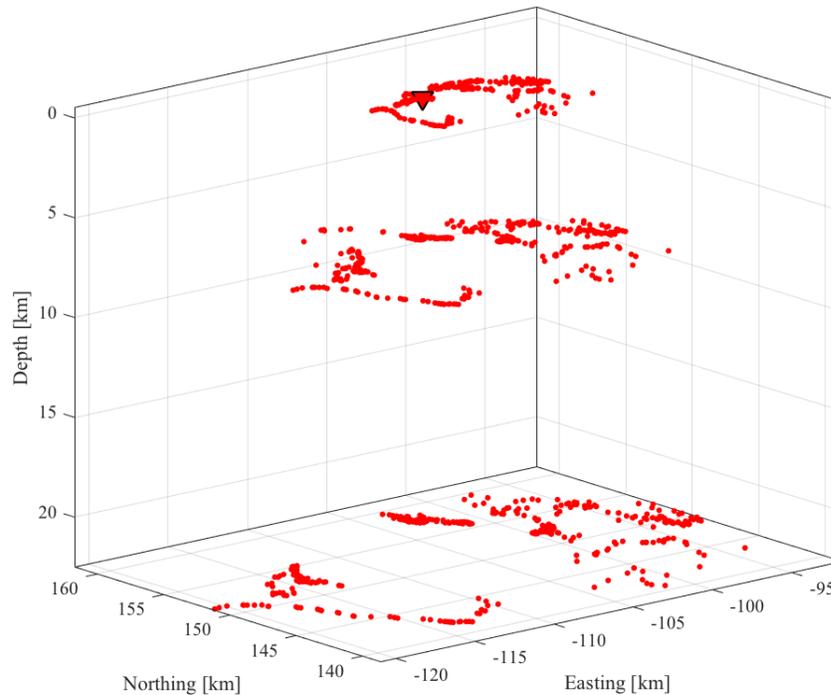


Figure 3.2.10: Example of piercing point positions after S-wave propagation considering the 3-D velocity model from Diehl et al. (2009) and my own ray tracing implementation.

3.2.5 Shooting correction

In order to produce the P-to-S conversion, I test a way to see how much I have to decrease the ray parameter p (or increase the epicentral distance Δ), so that the P-S converted wave hits the same point at the P-wave. In other words, our goal is to minimize the distance $X = dP - dS$, as shown in Figure 3.2.11.

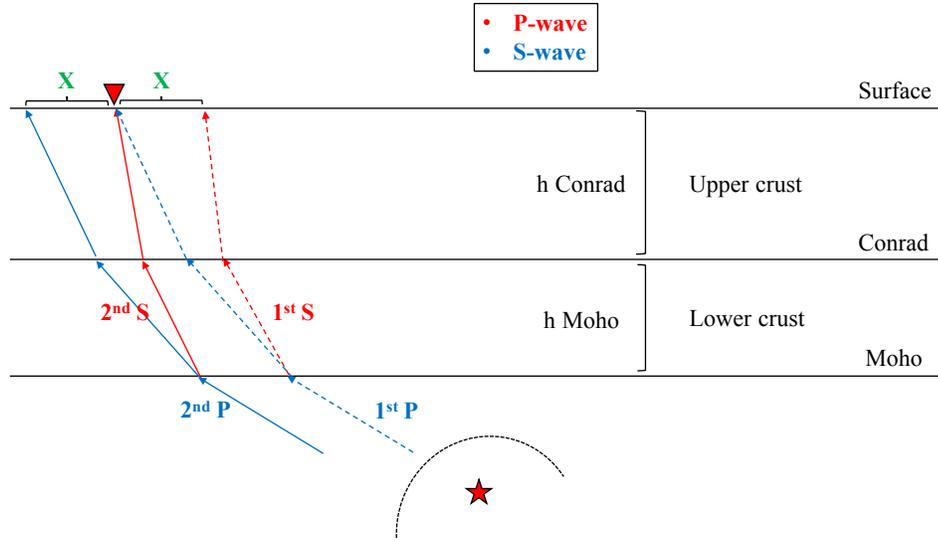


Figure 3.2.11: Graphical representation of the distance calculation between P direct wave and P - S converted wave. The earthquake focus (further away) is represented with a red star, while the station is a reverse triangle.

In other words, for the final shooting, the ray parameter is corrected proportionally to the value of the distance X , simulating that the station is not anymore at the original position but in a new position considering the addition of X , in the opposite direction from the arrival of the first shooting. In the Figure 3.2.12 we can observe how the X distance changes with respect to the ray parameter p and the epicentral distance Δ .

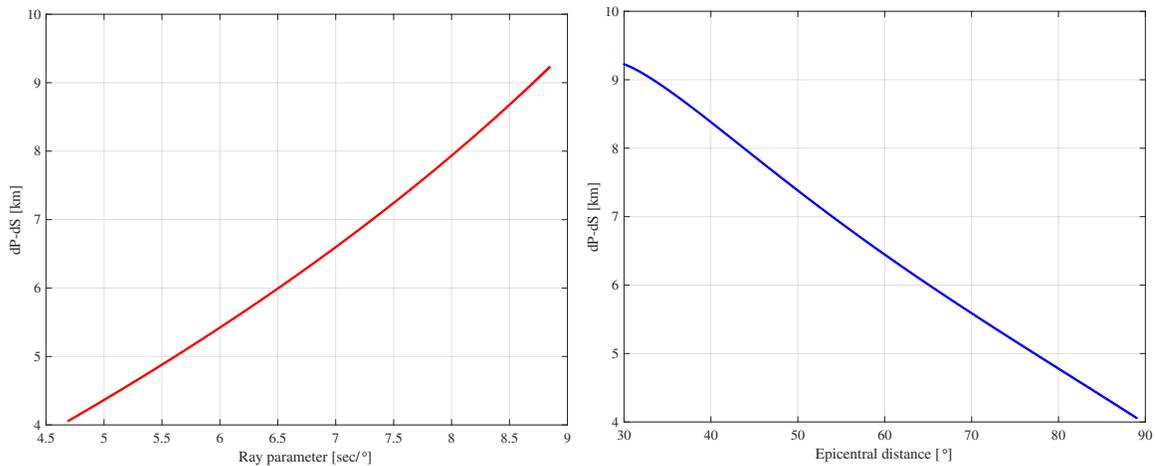


Figure 3.2.12: Distance between P - S converted wave and P -wave arriving at the surface as a function of the ray parameter and the epicentral distance for a source at 30 km depth using the global 1-D velocity model IASP91.

As we can see, the X distance is minimum for small values of p (minimum X around at

4 km for a value of $p = 4.692 \text{ sec}/^\circ$) and for large epicentral distances ($\Delta \approx 89^\circ$). As we observed in Figure 3.2.13 after the first shooting, we have to adjust the ray parameter to make the crustal S-waves arrive at the station.

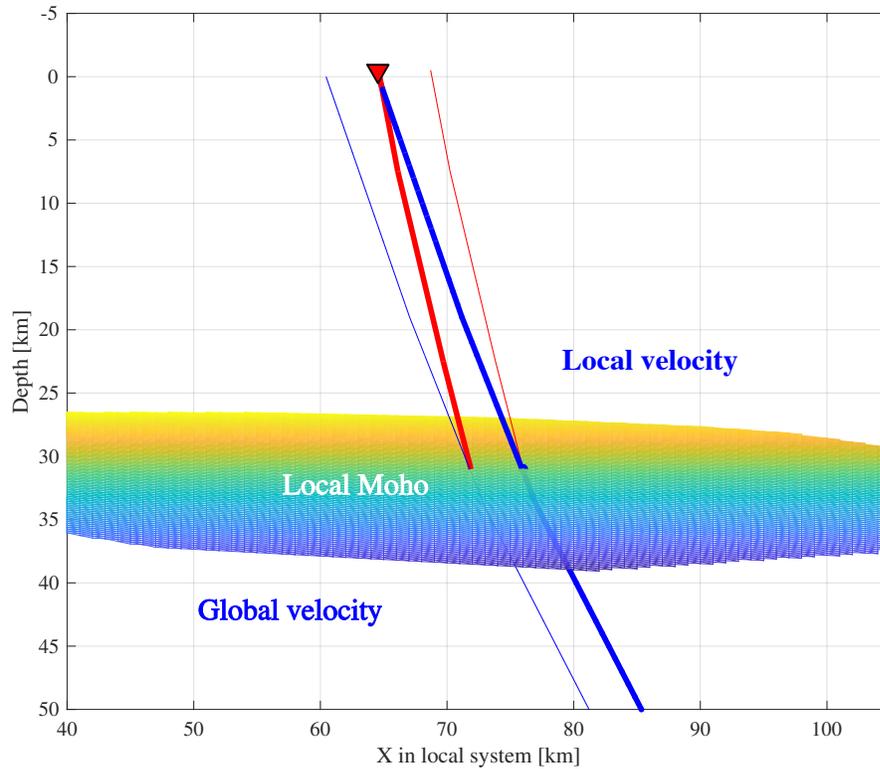


Figure 3.2.13: Example of correct P-to-S wave shooting. Blue thick line represents first shooting of global P-wave, red thin line first shooting of P-to-S conversion. Blue thin line is the second shooting of global P-wave, and red thick line is the second shooting of P-to-S wave, reaching the station.

After this correction, I shoot again a P-wave in the global model using the original tool of Knapmeyer, 2004 (where I find also a slightly different set of piercing points computed previously) and then I use the S-wave shooting procedure I develop and I described before considering the local velocity model from Diehl et al. (2009) including local Moho dip.

Figure 3.2.14 shows the improvement in the ray tracing of P-to-S wave between first and second shooting as the ray arrives to the surface.

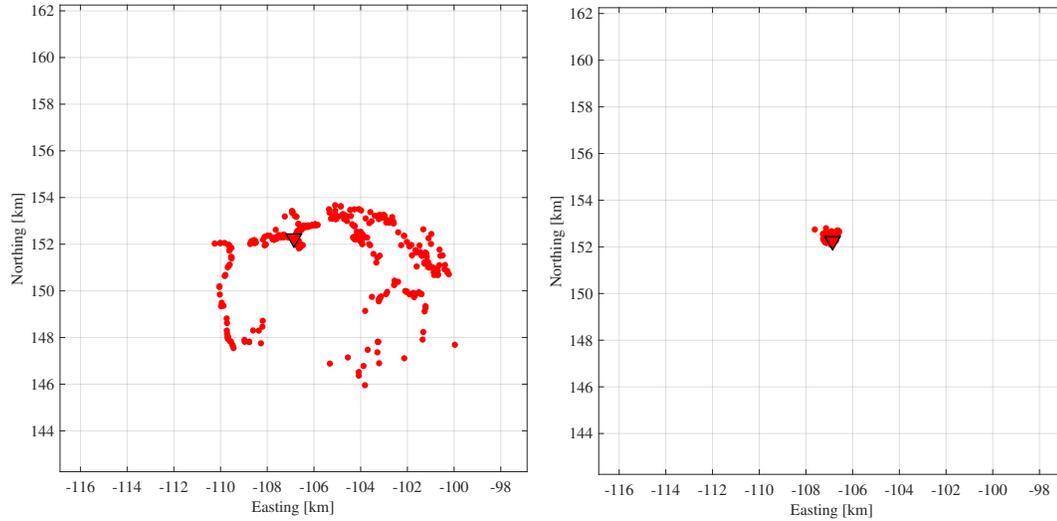


Figure 3.2.14: *P-to-S-wave arrival at a station: first shooting (left), second shooting (right) after correction, at the station ZUR.*

We can observe that the distance of the ray arrival from the station in Figure 3.2.13 decreases from mean values around 3 km to distances around 80 m for the examined case at station ZUR.

The graph in Figure 3.2.15 shows the improvement of this distance from the station between the first and the second shooting using local Moho model from Spada et al. (2013) and local velocity model from Diehl et al., 2009.

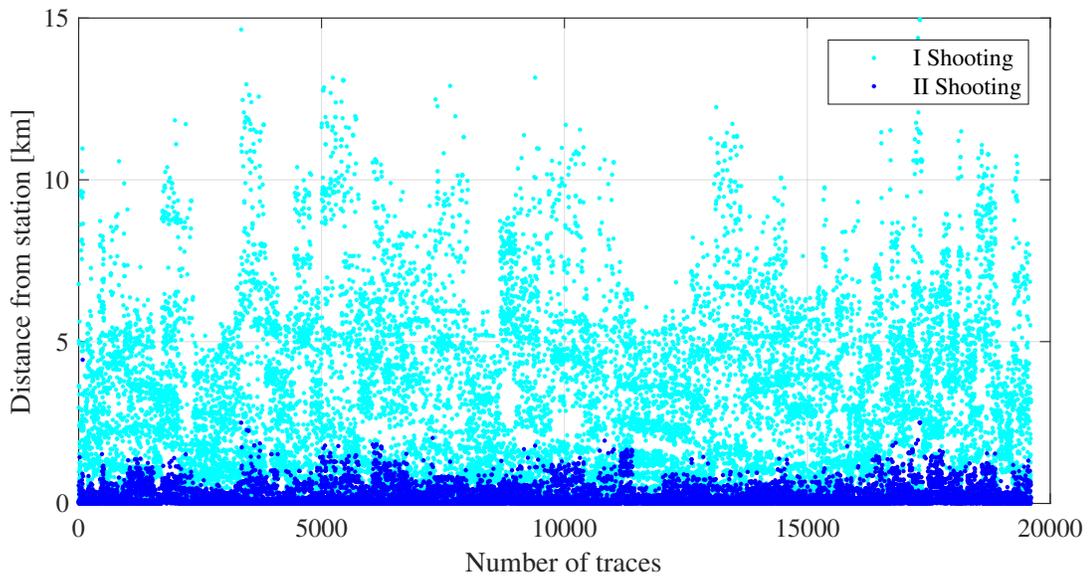


Figure 3.2.15: Comparison between first and second shooting using local Moho model from Spada et al., 2013 and local velocity model from Diehl et al., 2009.

Considering all the selected traces, the first shooting between the arriving S-wave and the station has a mean of 3 km (with a median of 2.4 km); after the adjustment of the ray parameter, the mean of this distance decreases to 180 m (median of 60 m), which represents a very good accuracy for our shooting tool. Meanwhile, the computation time remained very reasonable (13 minutes for the computation of the whole dataset, 0.04 sec for 1 trace).

3.2.6 Performances and shooting results

In this section I describe the accuracy of the ray tracing procedure I implemented. Figure 3.2.16 represents the shooting's distribution in our area.

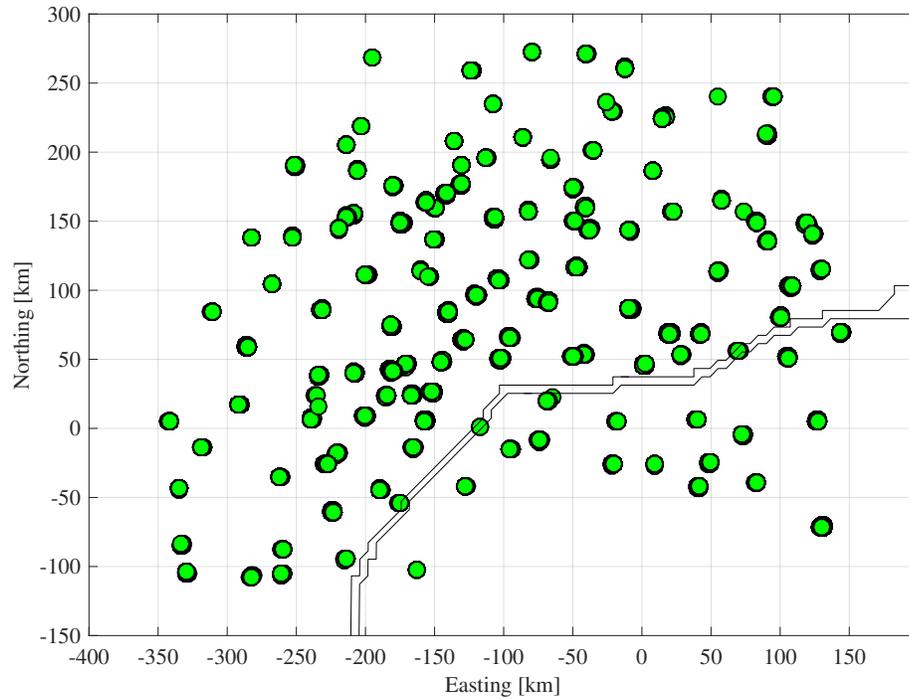


Figure 3.2.16: Shooting's distribution in our study area. Green dots represent S -arrival at the station, black double line plate boundary between European and Adriatic plate.

The position of the dots have a perfect match with respect to the position of the stations in the study area.

In Figures 3.2.17, 3.2.18 and 3.2.19 I represent the location of the rays at depth planes used for ray tracing in the 3-D space, respectively at 7.5, 22.5 and 37.5 km.

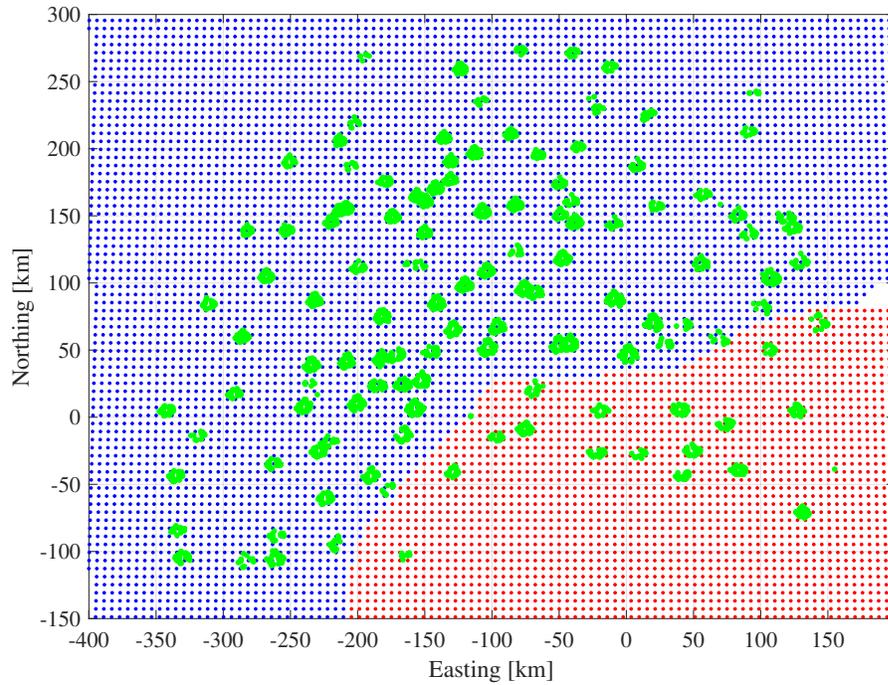


Figure 3.2.17: *Depth slice at 7.5 km. Green circles represent the position of the S-ray, blue points belong to European Moho, red points to Adriatic Moho.*

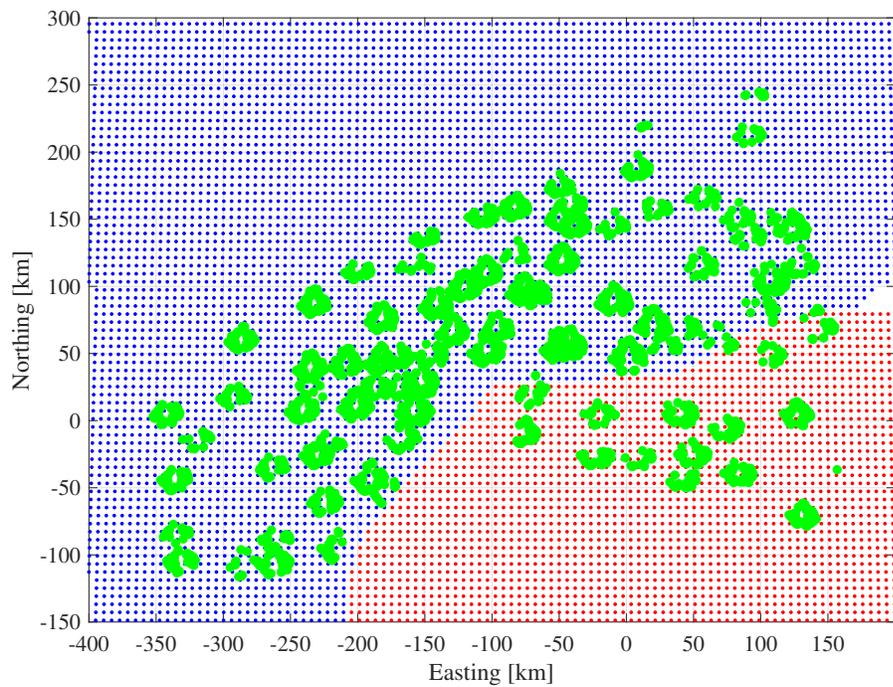


Figure 3.2.18: *Depth slice at 22.5 km. Green circles represent the position of the S-rays, blue points belong to European Moho, red points to Adriatic Moho.*

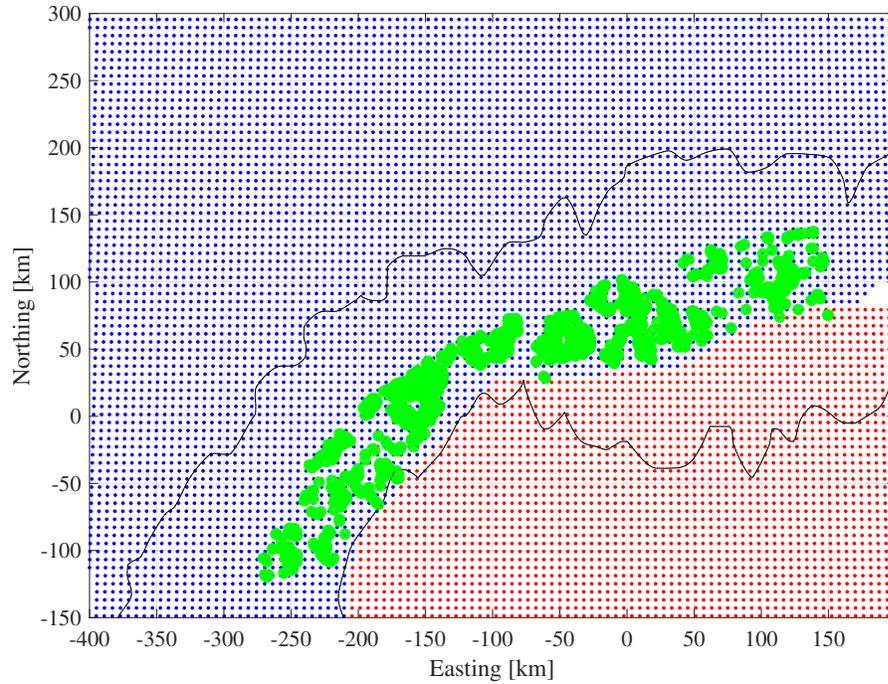


Figure 3.2.19: Depth slice at 37.5 km. Green circles represent the position of the S-rays, blue points belong to European Moho, red points to Adriatic Moho. Black line shows the Alpine arc's smoothed 800 m altitude line.

In particular, Figure 3.2.19 shows that the deepest piercing points of the S-waves are located inside the Alpine arc, where the crust in the assumed model is thicker. A good indication regarding the limit of the seismic resolution is shown in Figure 3.2.20, where ray coverage at the local Moho depth is shown with the corresponding Fresnel zone. This considers the volume within which the waves are sensitive to velocity variations. The dimension of the Fresnel zone can be calculated by simple geometry (Sheriff, 1996).

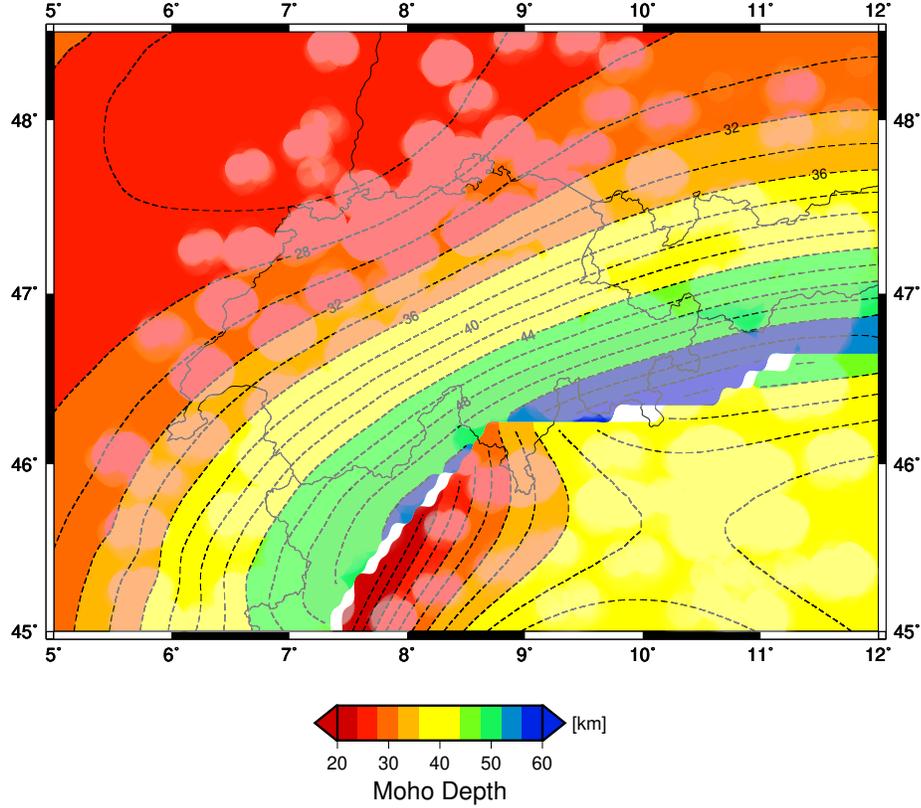


Figure 3.2.20: Moho depth map according to the model of Spada et al. (2013) and ray coverage of my final dataset. White shaded circles represent the position of piercing points for the final crustal S-wave. Dimension of the radius of each circle is proportional to the Fresnel zone.

The Fresnel zone radius depends on the wavelength, which is itself a function of frequency and velocity of the wave. Applying the calculation of the Fresnel zone for the S-wave in the Central Alps case, the frequency I refer to is the maximum frequency at which I computed the RFs $f_{max} = 0.5$ Hz and for the S-velocity I use typical velocities derived from the IASP91 model at the depths of the piercing points at the initial Moho model from Spada et al. (2013).

For example, if the depth of a piercing point at the Moho is $z = 30$ km, the corresponding S-wave velocity $V_{s_{iasp91}}$ is 3.75 km/s and the Fresnel zone F is:

$$F = \sqrt{\frac{z * V_{s_{iasp91}}}{2 * f_{max}}} \quad (3.2.4)$$

$F \cong 10.6$ km, which represents a typical Fresnel zone at the Moho depth. Considering all the depths of the piercing points for the Figure 3.2.20 the range of variation of the Fresnel radius goes between 9.3 and 15 km. Overall, the coverage of the crustal volume by P-to-S waves in the Central Alps is very good, and shows that a densely spaced network is suitable for mapping Vs using receiver function method.

Chapter 4

Model Parameterization

In seismology the model parameterization is a simplified representation of a seismic structure by a set of model parameters. A new important contribution of this study is that I define in depth a flexible model with separate velocities above and below each discontinuity. This flexibility allows to accommodate a velocity gradient within each layer and investigate a velocity jumps across discontinuities.

In the second part, I describe the use of a velocity gradient across a layer in the computation of individual ray paths and how I extract the corresponding 1-D velocities model from the 3-D to compute synthetic receiver functions. The new model parameterization is applied to the context of the Central Alps.

4.1 Model parameterization in classical tomography

In this section, I discuss the model parameterization in tomography, which is the way of modeling and representing the subsurface structure of the Earth (Tong et al., 2019). The factors that affect the choice of a particular model parameterization include forward solvers, inverse algorithms, *a priori* knowledge of Earth's structure and resolution capability of available data set (Kissling et al., 2001). Different functions have been used to express the 3-D structure of the Earth's interior (Rawlinson et al., 2010; Zhao, 2015), where blocks and grids are the most basic forms of parameterization. In relation to the accuracy of the forward model, the initial velocity and its updates are usually discretized on small blocks or fine grids and this approach (with some variants) have become very popular in local earthquake tomography as well as in teleseismic tomography (Zhao et al., 1992; Thurber and Eberhart-Phillips, 1999; Zhao, 2009). The typical parameterization used in classical tomography for a 3-D domain is shown in Figure 4.1.1.

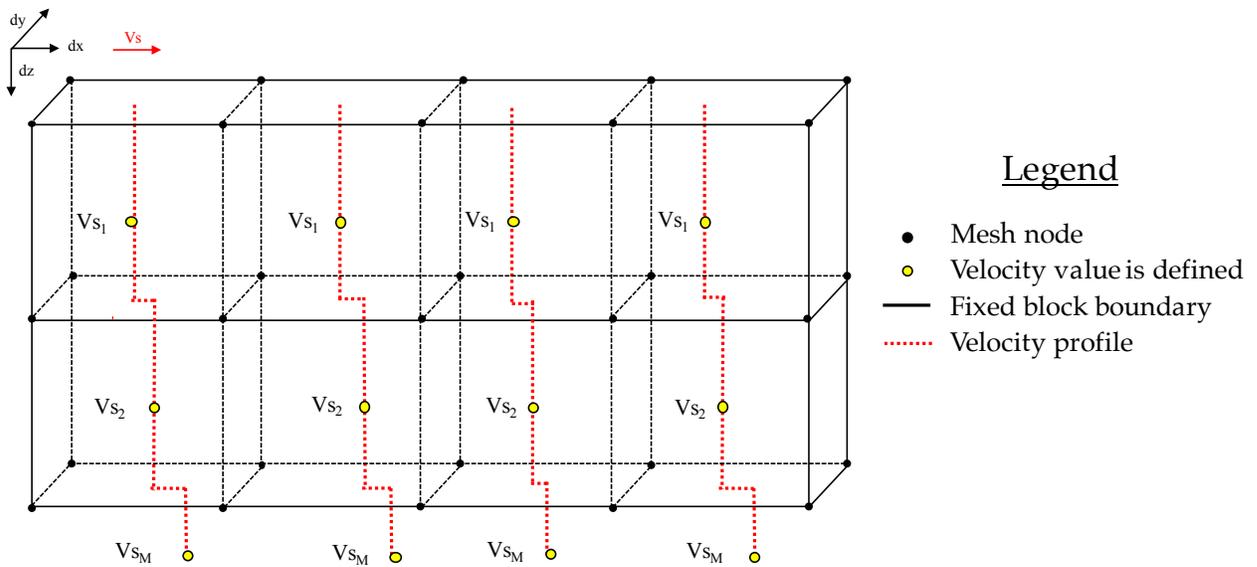


Figure 4.1.1: 3-D Sketch of typical model parameterization used in classical tomography. Block boundaries are fixed in space. Velocity values represent a volume.

In this approach velocity is defined inside each block and has usually a constant value. Even though interfaces in reality may vary in depth, the block geometries are usually fixed. Such grids are typically linearly interpolated to obtain velocities in other points of the 3-D space. The main drawback of this type of parameterization is that sharp changes in velocity and interfaces (e.g. Conrad or Moho) depth are poorly resolved. Therefore, it may be too simplistic to represent the geological complexity of a given study area.

4.2 Model parameterization for 3-D converted wave

In order to surmount these drawbacks and to aim at images which better reflect the complexity of the Earth's interior, I define a new model parameterization, which is shown in detail in Figure 4.2.1.

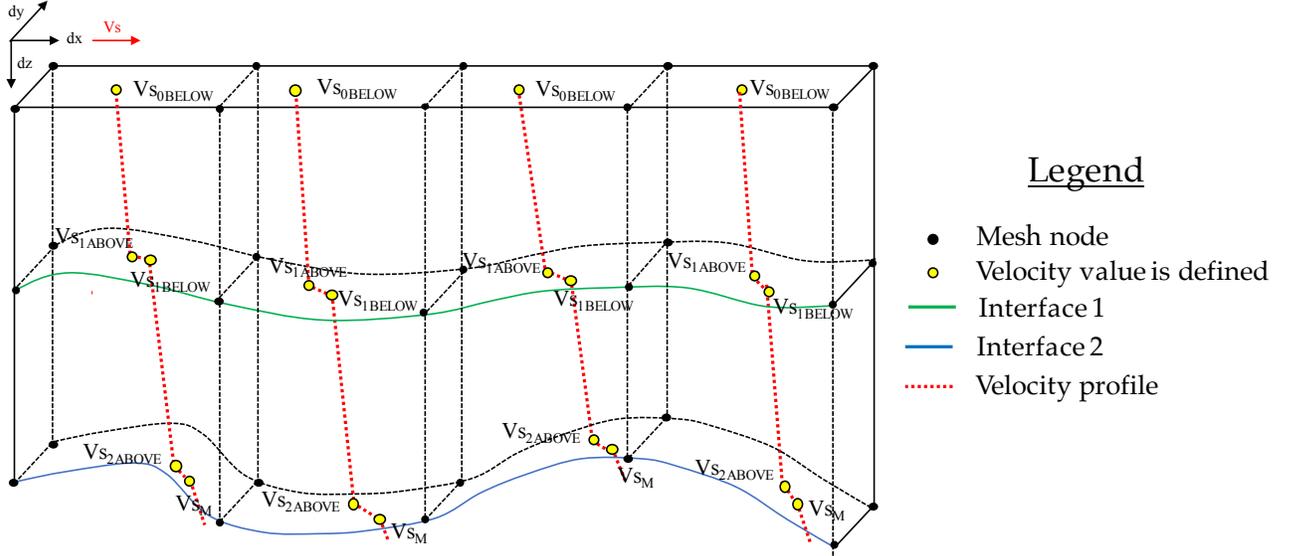


Figure 4.2.1: 3-D model parameterization used in this work for RF tomography. Layer boundaries are flexible in depth. Velocity profiles belong to a grid node (a point in map view).

Taking advantage of one of the features of the RF technique, which is more sensitive to the geometry of discontinuities with respect to tomography, I parameterize a multi-layer crustal model with a flexible nodes along the vertical direction. Moreover, I define a velocity above and below each interface and this allows us to accommodate fixed nodes in horizontal direction but with both velocity jumps between adjacent layers and velocity gradients across layers.

Expressing this in terms of variables:

- NX, NY, NZ , are the number of the nodes in the x, y, z directions;
- CX, CY, CZ , are the coordinates along the x, y, z directions.

At each interface:

- VP_{above} , P-wave velocity above the interface;
- VP_{below} , P-wave velocity below the interface;

- $V_{S_{above}}$, S-wave velocity above the interface;
- $V_{S_{below}}$, S-wave velocity below the interface;

For the implementation in the Alps context, I define V_s from Diehl's V_p using a constant initial $V_p/V_s = 1.73$. Considering that the velocity value above the surface is not used for RF calculation, and below the Moho is kept fixed at the mantle value of the IASP91 model, I reduced the number of independent variables at each node to 4.

4.2.1 Ray tracing accuracy using new model parameterization

Once the new model parameterization has been defined, I adapt the ray shooting tool that I developed for this work, considering the lateral variations of the velocity structures and the local dip of the discontinuities in the Alpine region. Figure 4.2.2 shows the improvement from the first to the second corrected shooting considering the model parameterization described above, in terms of distance of S-wave arrival at the surface from the station.

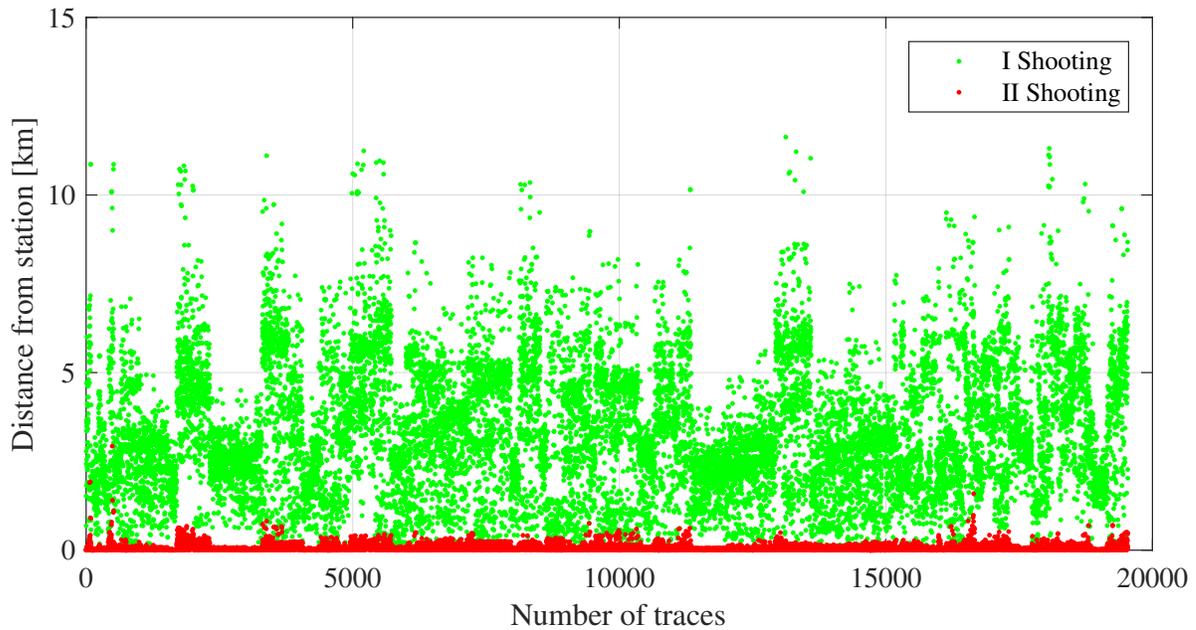


Figure 4.2.2: Distance of S-wave piercing point at the surface from the station. Comparison between first and second shooting using the new model parameterization.

Considering all the selected traces, the first shooting between the arriving S-wave and the station has a mean of 3.3 km (with a median of 3.0 km), and after the adjustment of the ray parameter the mean distance drastically decreases to 60 m (median 33 m), which

confirms the goodness of the shooting implementation and its accuracy already shown in the previous chapter. Figure 4.2.3 shows the comparison between the final corrected shooting considering the model parameterization used in classical tomography and the new model parameterization developed in this work.

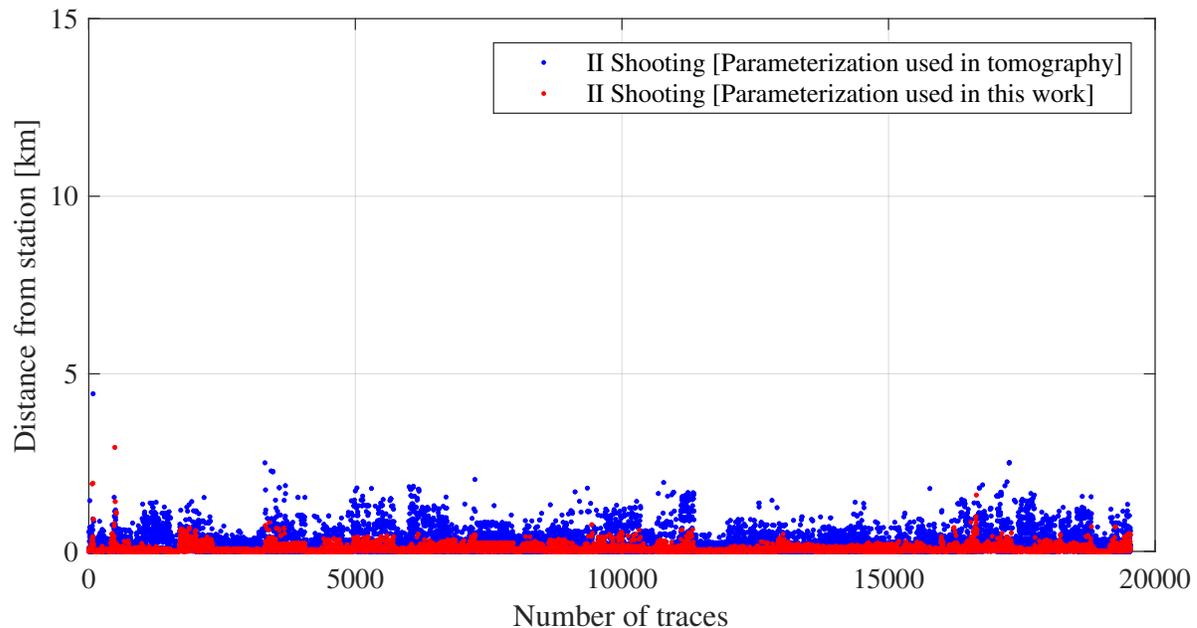


Figure 4.2.3: Comparison between the final shooting of the ray tracer I develop with the model parameterization used in tomography (blue points) and new model parameterization defined in this work (red points).

In particular, the graph shows that the distance from the station becomes even smaller if I use the model parameterization defined in this work (mean of distances 180 m) as input model with respect to the model parameterization defined in classical tomography (mean of distances 60 m). The improvement is due to the higher number of defined velocity nodes, that reflects a better approximations of reality.

4.3 New model parameterization applied to the Alps

In this section I define the initial setup of the model geometry and velocities in the Central Alpine context, using the new model parameterization formulation.

4.3.1 Setup of the initial model geometry

It is known that the crust is not homogeneous but has a vertically layered structure. In 1925 the seismologist Victor Conrad separated arrivals from a Tauern (Eastern Alps) earthquake of 1923 into P_g and S_g waves in an upper crustal layer and faster P^* and S^* waves that travelled with velocities 6.29 km/s and 3.57 km/s, respectively, in a deeper layer.

Because the P^* and S^* velocities are significantly slower than corresponding upper mantle velocities, Conrad deduced that they were head waves from a lower crustal layer. The interface separating the continental crust into an upper crustal layer and a lower crustal layer is called the Conrad discontinuity, in honour of Victor Conrad. Influenced by early petrological model of crustal composition and by comparison with seismic velocities in known materials, seismologists referred to the upper and lower crustal layers as the granitic layer and the basaltic layer, respectively.

This petrological separation is now known to be simplistic as more gradual changes within the crust occur as a result of a regional tectonic and metamorphic history. In contrast to the Moho, which is a primary discontinuity and is usually sharp, the Conrad discontinuity may be poorly defined or absent in some areas.

4.3.1.1 Conrad discontinuity in the Alps

In the Alpine domain, a continuous layer between the granitic layer and the Moho was not consistently found by seismic observations. Most geophysicists take for granted the existence of a so-called “intermediate” layer (Hart, 1969), that can often be subdivided into 2 parts.

In some areas, the intermediate layer seems to begin with high velocities. The range in which a velocity of 6.3 km/s is found with a positive velocity gradient can be defined as a top of the intermediate layer. In and outside Alps, the Conrad discontinuities lies between 15 and 20 km (Hart, 1969). Exceptions are known, especially the gravity highs at the southern inner margin of the Alps (Ivrea Zone, Garda Lake). The inner structure of the intermediate layer seems to vary remarkably. The greatest thickness of the intermediate layer has been found in the Central Alps in Switzerland and in Western Austria Closs (1965).

Figure 4.3.1 shows the interpreted ECORS-CROP transect and the variability in crustal thicknesses not only between European and Adriatic plate, but also of the upper and lower crust within the same plate.

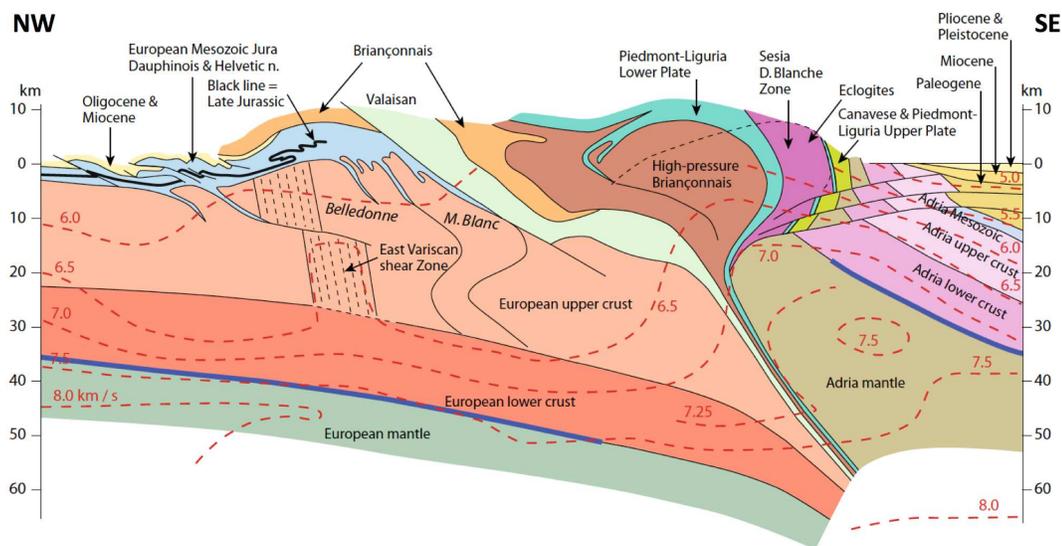


Figure 4.3.1: ECORS-CROP transect by Schmid et al. (2017). The section is superimposed with and partly based on vertical cross sections across the P-wave velocity model of Diehl et al. (2009) indicated by red dashed iso-velocity contours. The Moho (blue line) is taken from a combination of controlled source seismology, local earthquake tomography and receiver function analysis (Spada et al., 2013).

The geological interpretation of the P-wave velocity model of Diehl et al. (2009) allows new details concerning the outlines of the Adriatic mantle wedge and integrates its outline into the structure of the Alps. The interpretation of these findings reveal that the upper plate of the Eastern Alps is defined by the Austroalpine nappe system, while the upper plate of the Western Alps is formed by non-metamorphic klippen of the Piedmont-Ligurian Oceanic crust and lithosphere (Molli et al., 2010).

In general, a subdivision of upper and lower crust by a well-defined boundary is often, but not always seen on interpreted cross-sections. However, as Escher et al. (1997) point it out, the mid-crustal structure is clearly less constrained than the near-surface geology and the Moho (Figure 4.3.2).

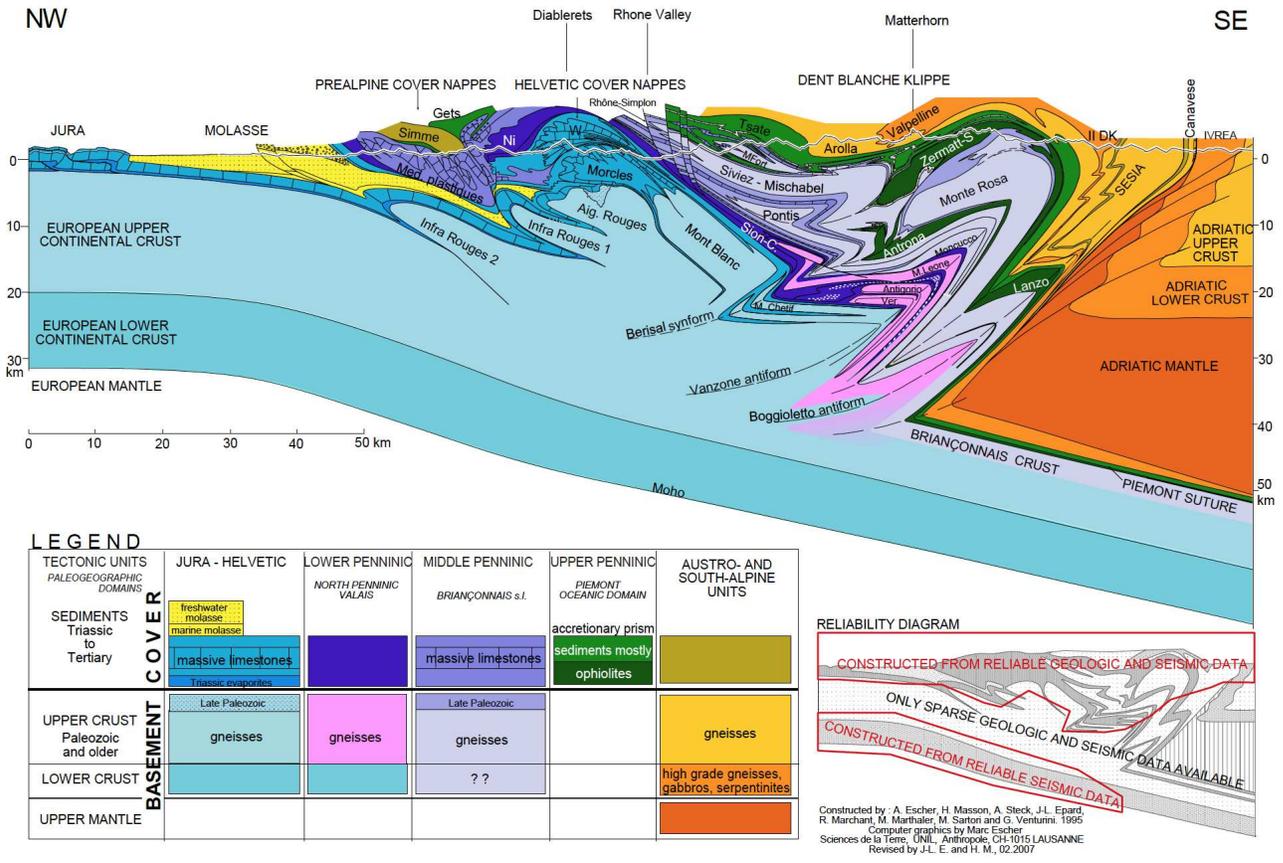


Figure 4.3.2: Schematic geological profile through the Western Swiss-Italian Alps from the Mont Tendre (Jura) in the northwest to the Val Sesia in the southeast. It is based partly on the results of seismic data. The overall interpreted structure shows that the main response to compressional stress was the subduction of the European (and Briançonnais) lithosphere. It also suggests that the deformation and stacking of nappes took place within the upper part of the downgoing continental crust, while the lower crust subducted in a passive way together with the lithospheric mantle (modified after Escher et al., 1997).

Based on these geological models and geophysical *a priori* information, I define the initial geometry of the Conrad discontinuity in my model with respect to the Moho depth model of Spada et al. (2013), and consider a lower crustal thickness of 12 km everywhere (Figure 4.3.3).

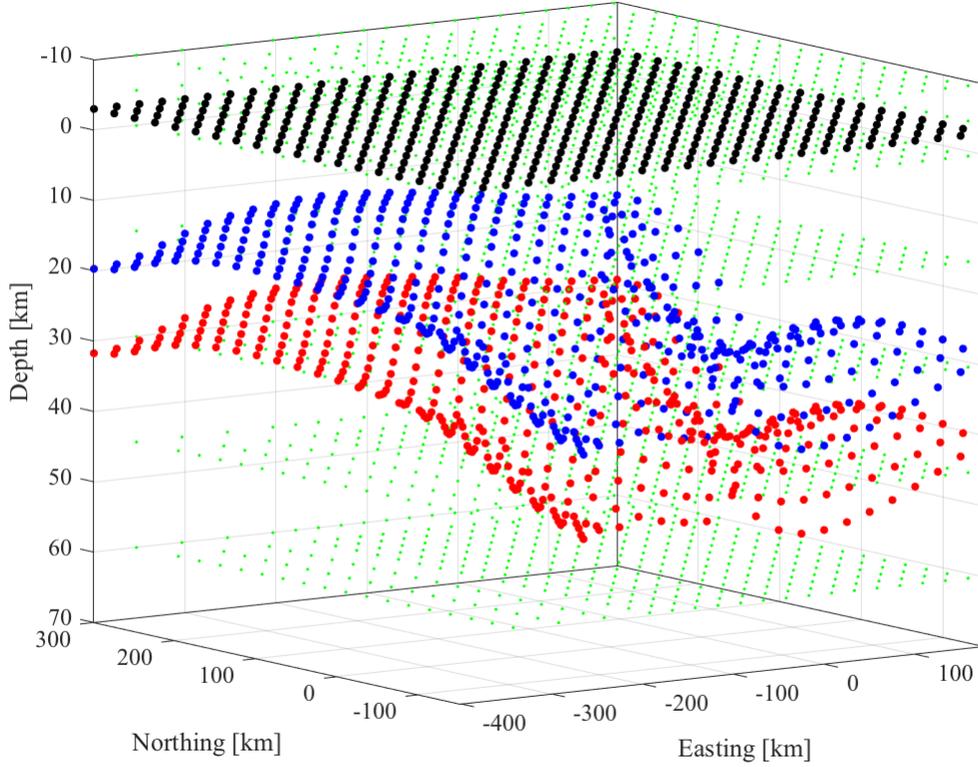


Figure 4.3.3: Model geometry of our study area in 3-D. Small green nodes represent velocity according to the local earthquake tomography by Diehl et al. (2009). Bigger dots in black, blue and red represent respectively surface, Conrad and Moho interface. Vertical axis is exaggerated.

During the inversion, the Conrad geometry will be able to evolve independently. For the definition of the surface, the elevation of the highest station (at 3 000 m above sea level) is chosen to define velocities, but rays only propagated to the actual elevation of the corresponding station.

4.3.2 Setup of the initial velocity model

To construct the initial model for the Central Alps, I build a 2-layer model of the crust and consider 3 interfaces: the surface, an intra-crustal boundary (possible Conrad discontinuity) and the Moho discontinuity. Regarding Moho model, I decide to use the same geometry of Spada et al. (2013) as initial model. For the velocity model parameterization, I decide to introduce a P-velocity jump $\Delta V_{P_{T91}}$ derived from the IASP91 model (Kennett and Engdahl, 1991), equal to 0.7 km/s. In Figure 4.3.4 I show the P-wave velocity values at the depth slice of 15-km given by the model of local earthquake tomography by Diehl et al. (2009).

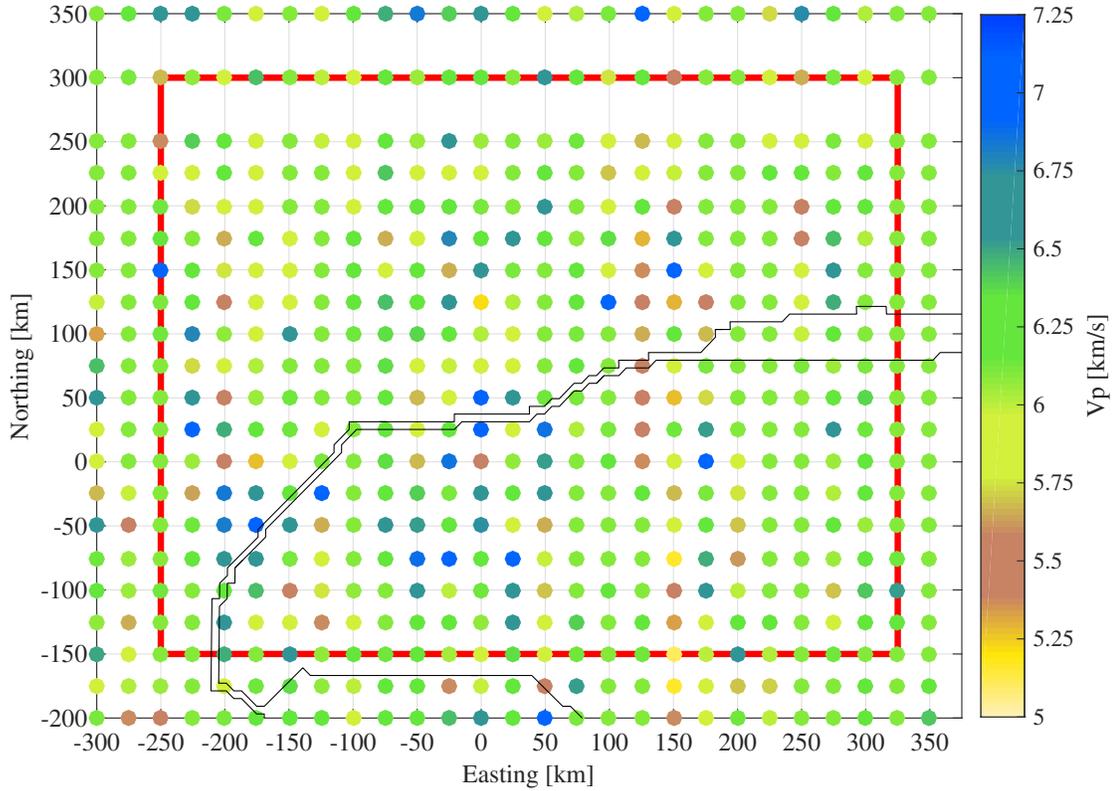


Figure 4.3.4: P-wave velocity at 15-km depth from the local earthquake tomography model of Diehl et al. (2009). The red box is our study area; thin double line indicates the plate boundary.

I resample the starting model defining velocities every 25×25 km, and then apply the new model parameterization, which considers a velocity value above and below each discontinuity. In the case of 2 layers with variable velocities and 3 interfaces bounding them, the number of independent P-wave velocity variables is 4:

- $VP_{BELOW-SURFACE}$, P-velocity below the surface;
- $VP_{ABOVE-CONRAD}$, P-velocity above the Conrad discontinuity;
- $VP_{BELOW-CONRAD}$, P-velocity below the Conrad discontinuity;
- $VP_{ABOVE-MOHO}$, P-velocity above the Moho discontinuity.

Figure 4.3.5 and Figure 4.3.6 represent respectively the P-wave velocity derived from the local earthquake tomography model by Diehl et al. (2009) and considering our parameterization at the depth slice of 15-km (one of the plane where we know the velocity from the tomography model by Diehl et al. (2009)).

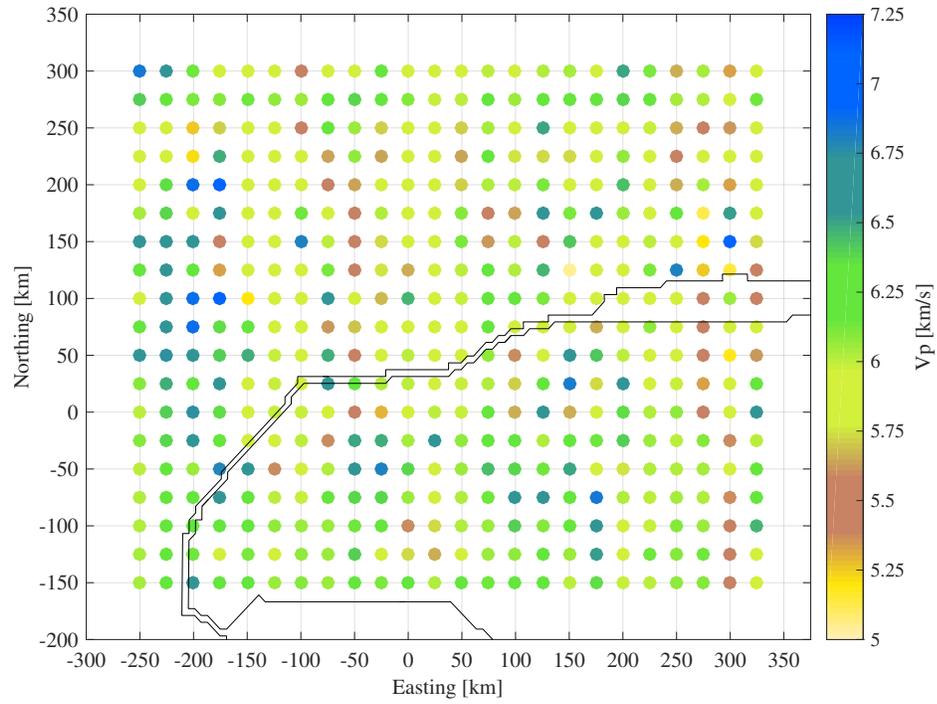


Figure 4.3.5: P-wave velocity model defined above 15 km depth according to our model parameterization.

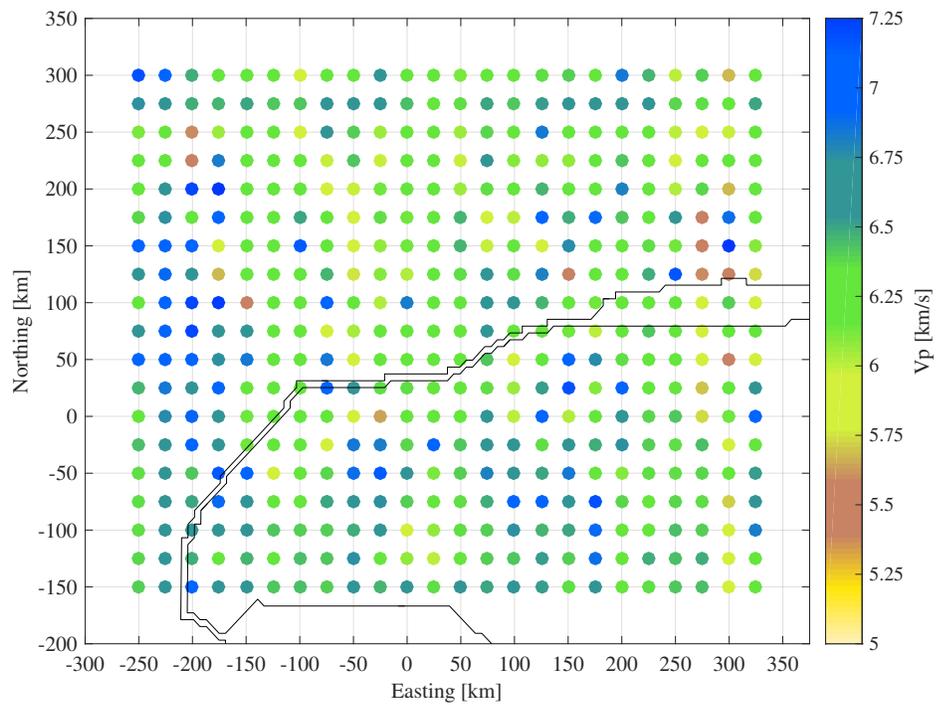


Figure 4.3.6: P-wave velocity model defined below 15 km depth according to our model parameterization.

As expected, both of the figures show a little perturbation with respect to the starting model given by tomography presented in Figure 4.3.4, with velocity values increasing from above to below by definition of our parameterization.

4.4 Extraction of 3-D velocity model

For each observed trace, synthetic RFs are computed using 1-D velocity models which are extracted from the 3-D velocity model along the raypath using the ray tracing procedure described in Chapter 3. Figure 4.4.1 shows an example of extraction of the S-velocity along a raypath for the station ZUR and its profile.

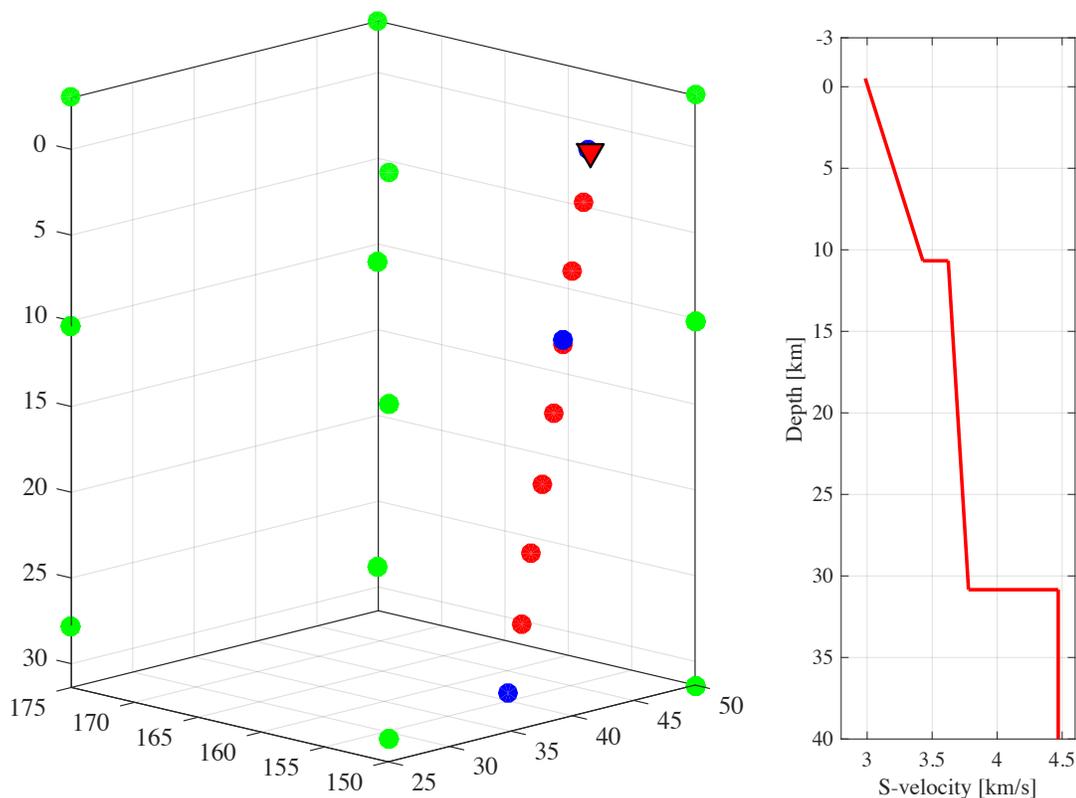


Figure 4.4.1: On the left, example of extraction of the velocity along a ray-path for the station ZUR. Green points shown the position where velocities value are known from the tomography of Diehl et al. (2009); red points define the step from the Moho to the surface; blue points (at Moho, Conrad and surface, overlapped by the station) were new velocity model is defined considering a weighted mean at the enclosing mesh nodes. On the right, example of the V_s profile according to our parameterization for the same trace.

In the 3-D approach, the local crustal velocity is obtained using a weighted mean based on the inverse of the square of the distance from the 8 enclosing neighbour nodes, 4 below

and 4 above with respect to the point, where we know the velocity from our input model. The local S-wave velocity is initially set by dividing the P-wave velocity model by a constant V_p/V_s of 1.73. This approach is compatible with the S-wave propagator that I developed in Chapter 3. The ray and the 1-D velocity model are sampled at regular intervals in depth, within a layer, and then at interfaces. In the example shown in Figure 4.4.1 I use 4 km depth sampling.

With this scheme I retrieve a 1-D velocity for each trace, and then I generate the corresponding synthetic RF. This waveform is then compared with the corresponding observed RF. An example of this comparison is shown in Figure 4.4.2.

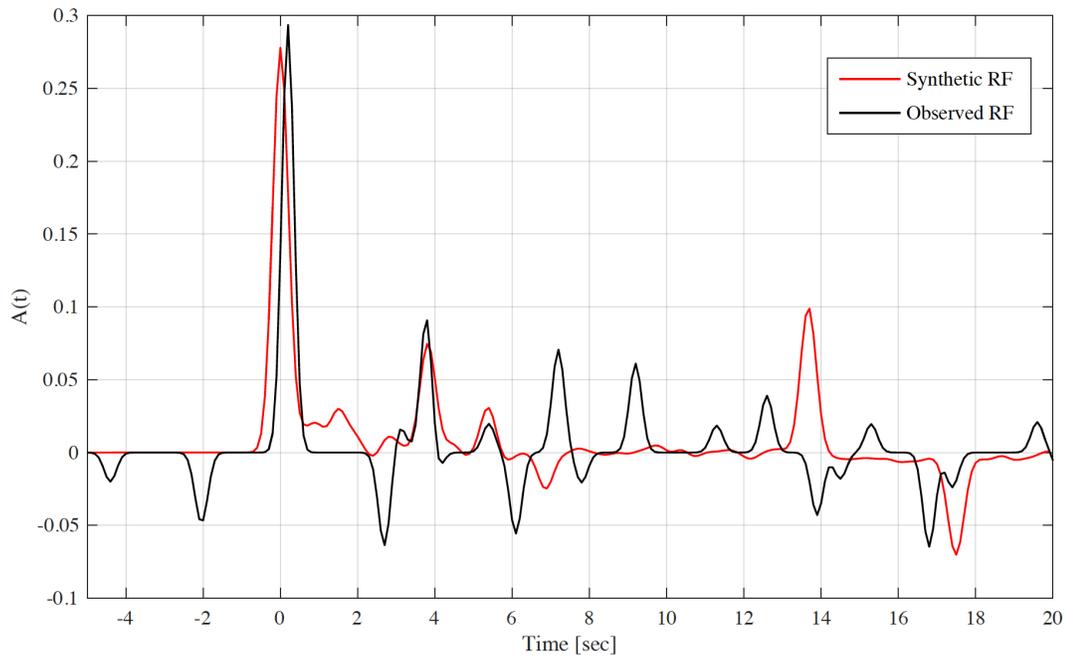


Figure 4.4.2: Comparison between observed (black) and synthetic (red) RF for the Chiapas earthquake, Mexico (2017-09-08, 04:49:19 UTC) at the station ZUR.

The difference between the observed and the synthetic RF will be optimized through inversion, by choosing an objective function to minimize. Full details will be discussed in Chapter 5.

4.5 Computation of synthetic RF

For the synthetic RF computation, the model parameters in each layer are the thickness, the S-wave velocity at the upper boundary, the S-wave velocity at the lower boundary, and

the velocity ratio between P and S waves (V_p/V_s). The S-wave velocity in each layer is constructed by linearly connecting the values at the upper and lower boundaries, to give a sequence of constant velocity gradient segments separated by velocity discontinuities as the full crustal velocity model. For RF calculation, I use the code of Shibutani et al. (1996) based on frequency domain deconvolution.

Table 4.1 shows respectively the 1-D input velocity structure for an example of velocity model used in this work.

Layer	H (km)	Vs1 (km/s)	Vs2 (km/s)	Vp/Vs	Qp	Qs
Upper Crust	20	3.20	3.40	1.73	1450	600
Lower Crust	12	3.60	3.70	1.73	1450	600
Mantle	Half-space	4.47	4.47	1.80	1450	600

Table 4.1: Input velocity model for receiver function calculation considering this section.

As we already mentioned in the first part of this chapter, the model parameterization that uses different velocities within each layer, has the possibility to accommodate velocity gradients between nearby layers. Since in this study the primary focus is not the signal's amplitude, I keep the default value for the seismic attenuation Q_p and Q_s at 1450 and 600 respectively. Figure 4.5.1 shows the synthetic RF generated by the models with and without a velocity gradient.

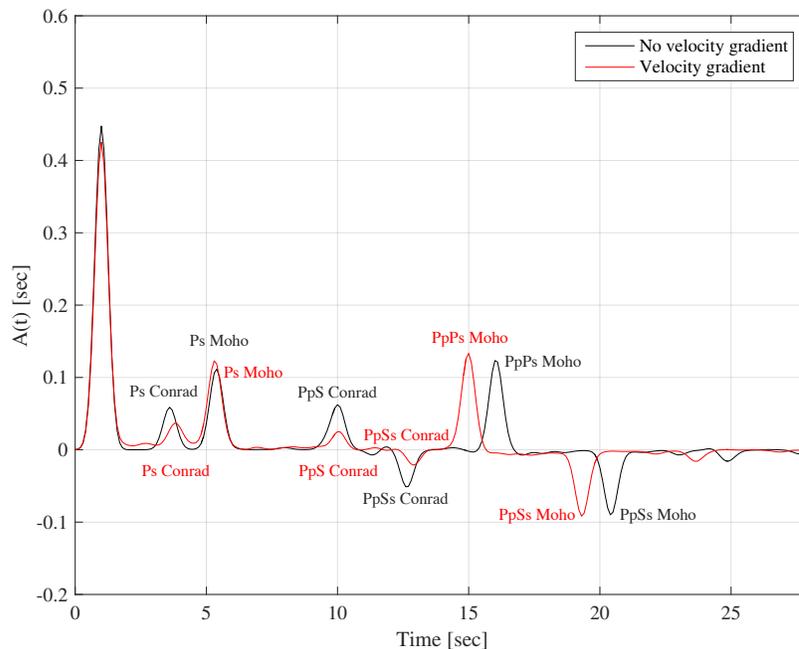


Figure 4.5.1: Synthetic RF generated by a model without velocity gradient (IASP91) in black and with a velocity gradient (input in Table 4.1) in red. P-wave arrival is at 1 sec.

The influence of the each parameter (crustal thicknesses, V_p/V_s ratio and velocity gradient) on the RF waveform, affecting the arrival times of the Ps phase and the multiples, will be discussed in section 4.5.

4.5.1 The Gaussian convolution

In order to obtain receiver function waveforms from the series of spikes from the iterative deconvolution, a convolution with a Gaussian is necessary. The width of the Gaussian depends on the maximum frequency content of the seismograms before deconvolution. Ammon (1997) has determined parameter for the Gaussian function to produce curves of the required width, which I report in Table 4.2.

agauss	f (Hz)	Pulse Width (s)
10	4.8	0.50
5	2.4	0.75
2.5	1.2	1.00
1.25	0.6	1.50
1	0.5	1.67
0.625	0.3	2.10
0.5	0.24	2.36
0.4	0.2	2.64
0.2	0.1	3.73

Table 4.2: Value of the Gauss filter (first column). The frequency at which the value $G(f)$ corresponds to the 10% of the passing signal and the pulse width is the time domain (Ammon, 1997).

In practice, the following equation relates the input Gauss parameter $agauss$ and the highest frequency content f_{max} .

$$agauss = (2 * \sqrt{2 * \ln 2})^2 * f_{max} \quad (4.5.1)$$

where \ln is the natural logarithm and f_{max} is the maximum frequency to which the RFs are computed. In our case, f_{max} being 0.5 Hz, the value of the $agauss$ parameter is 2.7726.

4.6 Receiver function sensitivity on inversion parameters

It is useful to know what are the effects of perturbations of model parameters (thicknesses and velocities) on the receiver function waveforms. To visualize these, I compute a series of synthetic tests, considering the 1-D *IASP91* velocity model (Kennett and Engdahl, 1991) as initial structure (Figure 4.6.1).

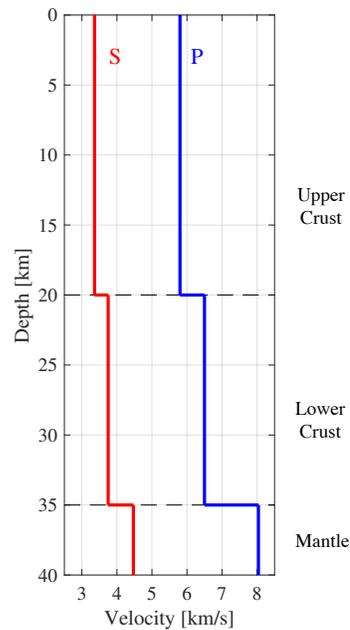


Figure 4.6.1: Crustal velocities in the *IASP91* model (Kennett and Engdahl, 1991): *P*-waves in blue, *S*-waves in red.

I perturb the initial velocity model by varying one parameter at a time, to clearly see the effects on the receiver functions.

4.6.1 Thicknesses

Figure 4.6.2 represent the situation the Conrad depth is varied, from the initial 20 km value (thick line in the plot).

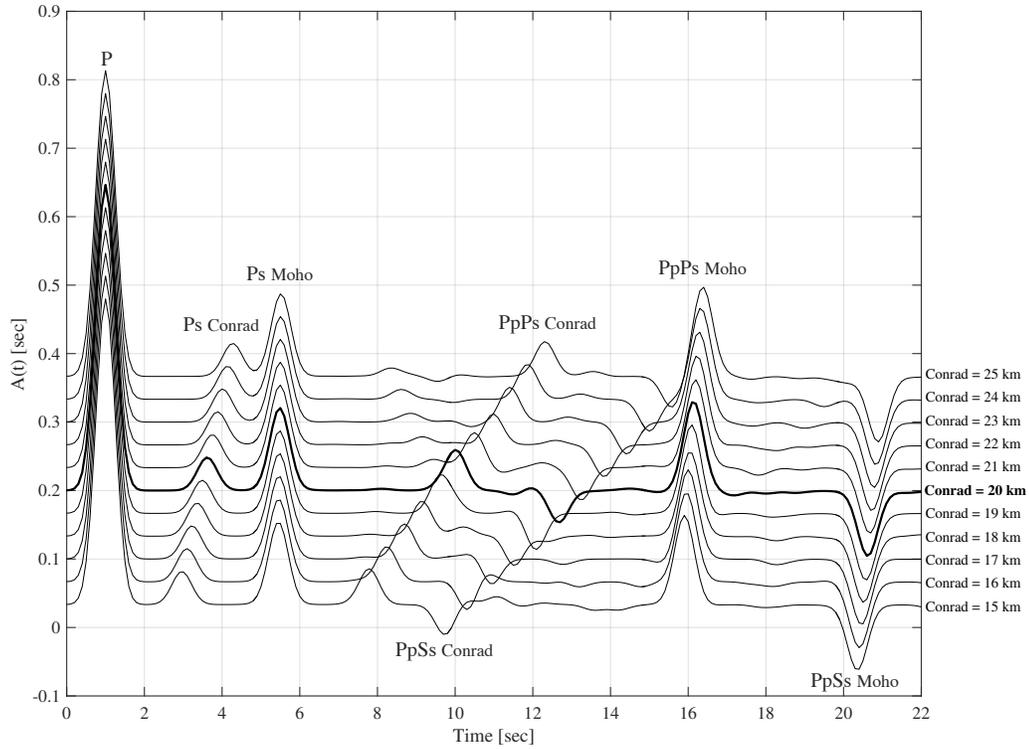


Figure 4.6.2: Synthetic test varying the depth of the Conrad discontinuity. Thickest line shows the initial value with the Conrad at 20 km depth. Direct P-wave arrival is at 1 sec.

What is observed is that the arrival time for the Ps conversion from the Conrad and the corresponding multiples PpPs and PpSs increase as the Conrad depth increases.

The same pattern is observed when the Conrad depth is kept fix but the Moho depth is varied (Figure 4.6.3).

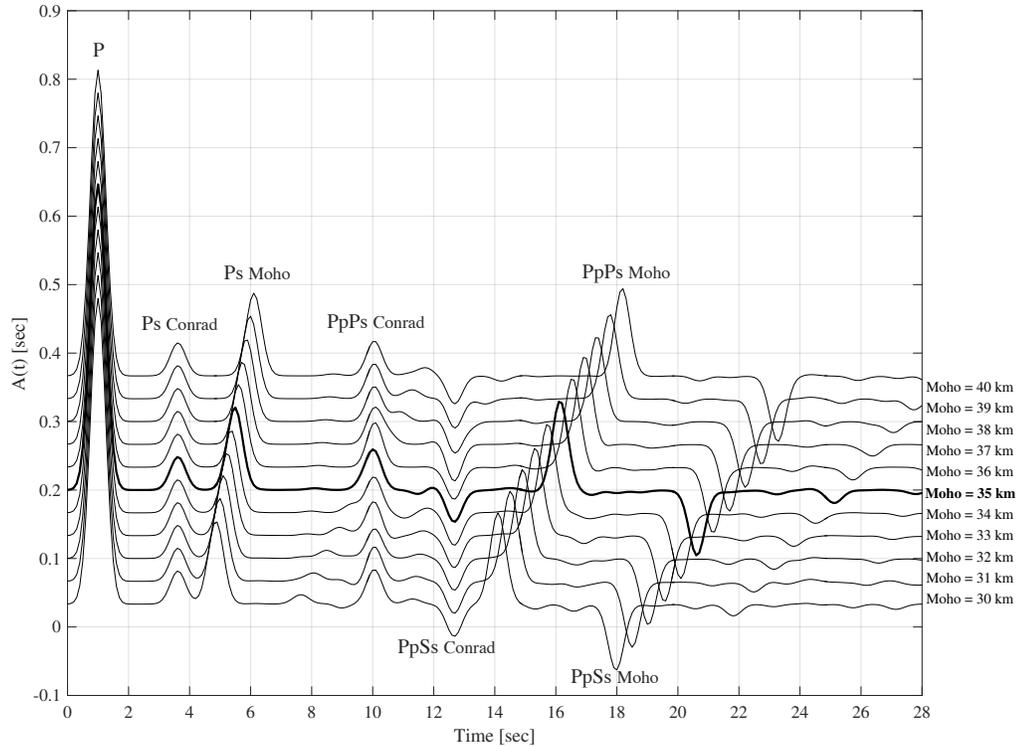


Figure 4.6.3: Synthetic test varying the depth of the Moho discontinuity. Thickest line shows the initial value with the Moho at 35 km depth. Direct P-wave arrival is at 1 sec.

In both cases shown above, there is also a variation in the amplitude of the peaks, but is negligible compared to the variation observed in arrival time.

4.6.2 Velocities

I performed similar tests varying one velocity value at a time. Figure 4.6.4 shows the synthetic test when the upper crustal V_p/V_s is varied, from the initial value of 1.73 (default value taken from *IASP91* model). The general effect is that all the wave conversions are delayed in time with increasing V_p/V_s and slightly increase their peak amplitude.

Similar results are obtained when perturbing the V_p/V_s in the lower crust (Figure 4.6.5), now using a value of 1.78 as reference (Zandt and Ammon, 1995). The main difference observed is that the converted phases for the intra-crustal layer remain at the same arrival times, and only the Moho conversions are delayed. The amplitude change is still small, its trend is the same for the Moho conversion, but opposite for the Conrad, as the velocity variation across it is decreased.

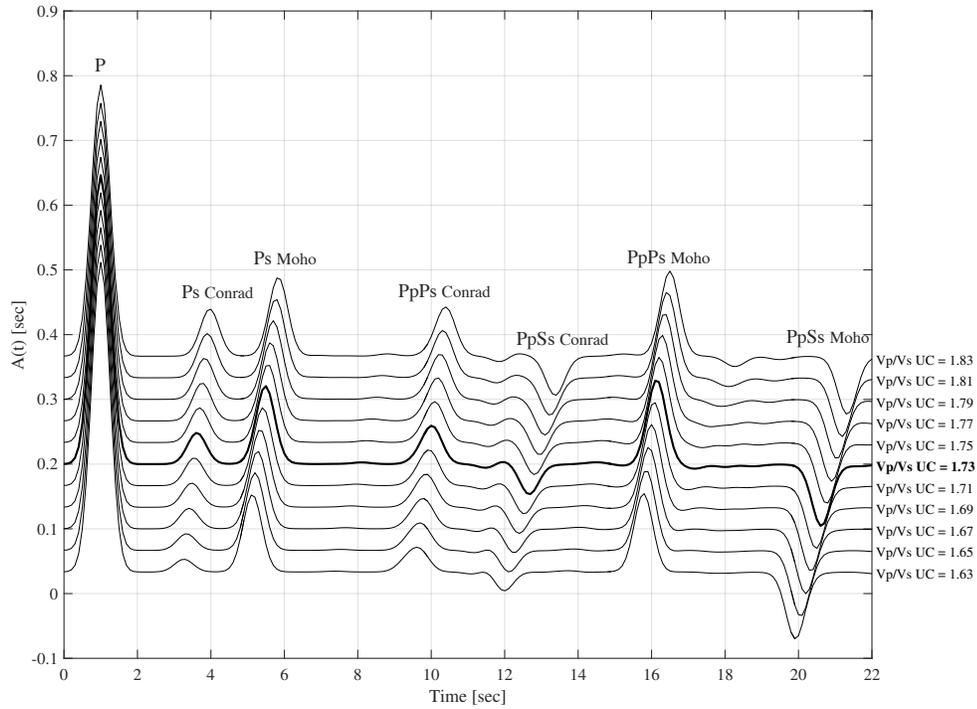


Figure 4.6.4: Synthetic test varying the V_p/V_s in the upper crust. Thickest line shows the initial value.

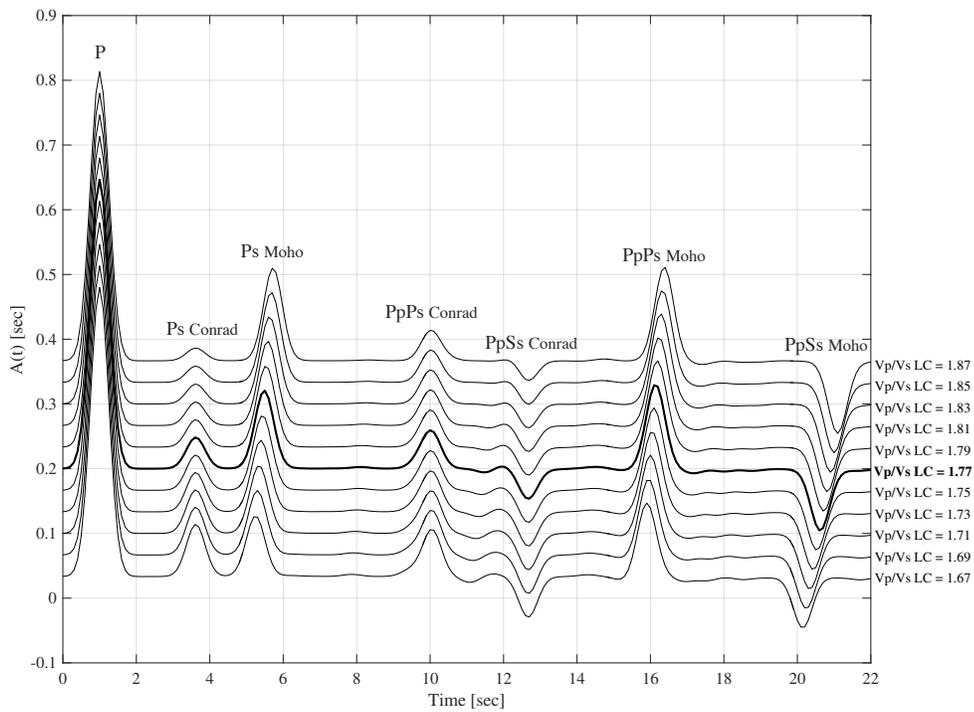


Figure 4.6.5: Synthetic test varying the V_p/V_s in the lower crust. Thickest line shows the initial value.

The new model parameterization with different velocity values above and below each discontinuity allows to investigate the velocity jump between layers. Moreover, it constitutes a simplification in the *a priori* V_p model replacing the 2 absolute values by the V_p jump (1 parameter), at the cost of fixing the mid-point of the jump. When we look at the variation of the velocity jump across the Conrad, the change in the amplitude of the waveforms is further emphasized with respect to the other parameters (Figure 4.6.6).

The change in amplitude is further emphasized when the velocity jump across the Conrad is varied (Figure 4.6.6).

In the case where $\Delta V_p \text{Conrad} = 0$, I do not observe any Conrad conversion, in this case the velocity profile is continuous from the top of the upper crust to the top of the lower crust. When negative velocity jumps are tested, a low velocity zone is created and this results in a negative amplitude Ps Conrad conversion and also flipped polarity multiples.

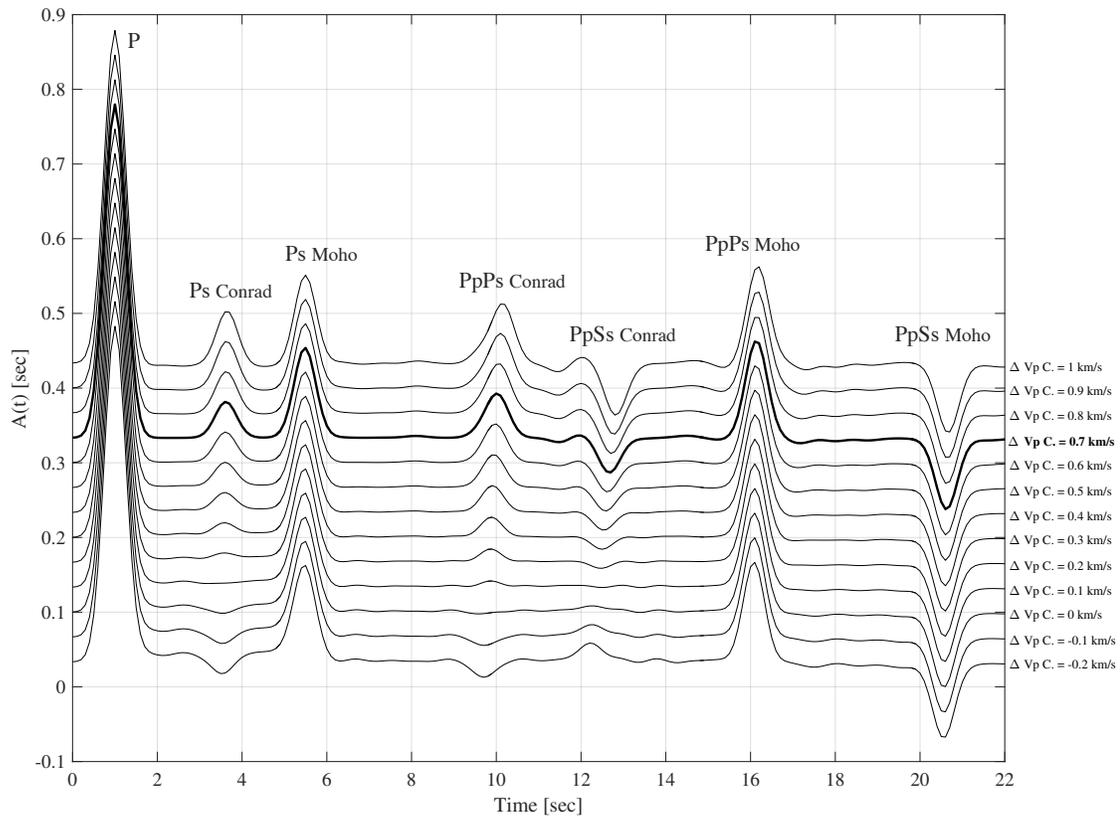


Figure 4.6.6: Synthetic test varying the velocity jump between upper and lower crust.

Chapter 5

Inversion strategy and setup

In this chapter I describe briefly the literature on global optimization methods, from deterministic techniques to the stochastic approaches with a particular focus on Monte Carlo methods that are currently popular in geophysics. To justify the choice of the method adopted for this study and before applying the inversion strategy to the seismic data, I show the efficiency of some inversion techniques considering an analytical function with several local minima.

For the application to the Central Alps the 3-D inversion of crustal structure by converted waves proceeds iteratively, by visiting every node on the map following a traveling salesman path (TSP). At each node, receiver function rays in the surrounding volume are considered for inversion, and are bundled into sub-blocks and ranges of back-azimuth. The velocity model at the given node is inverted using the technique of simulated annealing (SA), followed by a pattern search algorithm to avoid falling in a local minimum.

The main parameters in the inversion process (e.g. the number of iterations to use) have been tuned by simulating observations with *a priori* known models, which also allows to verify the accuracy and the applicability of the approach.

5.1 Inversion: optimization method

Inversion of geophysical data involves searching for optimal sets of Earth model parameters that can be used to compute synthetic data that match with observations (Sen and Stoffa, 2013). The data misfit (or fitness) is generally measured by a suitably defined objective function. Therefore, the essential elements of a model-based inversion algorithm include *data*, *model*, *forward problem*, *objective function* and *optimization method*.

We will restrict ourselves to discrete data d and model m defined by the following vectors:

$$d = [d_1 d_2 d_3 \dots d_N]^T \quad (5.1.1)$$

$$m = [m_1 m_2 m_3 \dots m_M]^T \quad (5.1.2)$$

where, in general, $N \neq M$ that means that the number of data parameters is different with respect to the model parameters.

Physical laws allow us to make predictions: given a complete description of a physical system, we can predict the results of measurements. This is the forward problem described in Chapter 3 that can be represented by the following equation:

$$d^{syn} = G(m) \quad (5.1.3)$$

where G is a nonlinear forward modeling operator that works on the model vector to generate synthetic data vector d^{syn} .

The *inverse problem* uses the actual results of measurements to infer the parameters that characterize the system. The traditional interaction between forward and inverse problems is schematized in Figure 5.1.1.

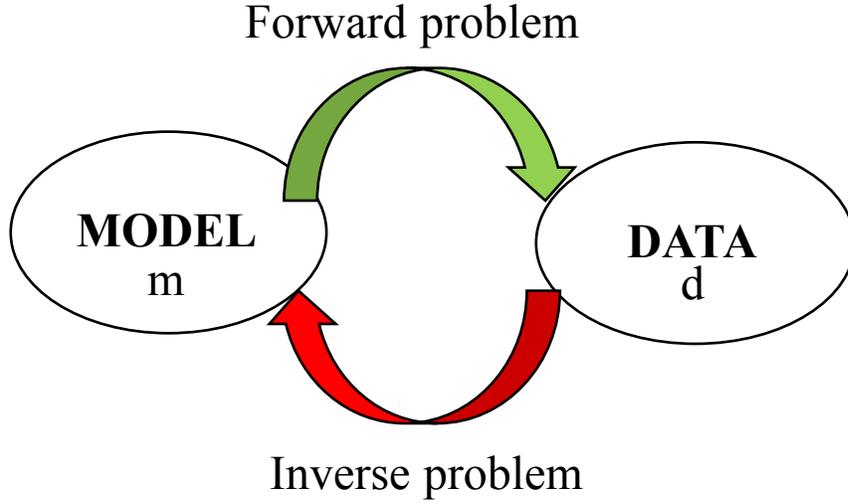


Figure 5.1.1: Interaction between forward and inverse problems. In an inverse problem, the characteristics of an unknown system are estimated based on its observed output data.

The main equation for the inverse problem is:

$$m = G^{-1}(d) \quad (5.1.4)$$

where G^{-1} represents the inverse problem operator that is able to estimate the model m given the data d .

The next step is to define an objective function that measures a misfit between the observed and synthetic data using a suitably defined norm. The data residual is given by:

$$\Delta d = d^{obs} - d^{syn} \quad (5.1.5)$$

where d^{obs} is the observed data vector and the data misfit norm can be written as:

$$\| \Delta d \|_p = \left[\sum_{i=1}^N | \Delta d_i |^p \right]^{\frac{1}{p}} \quad (5.1.6)$$

where p is an integer representing the order of the norm. The most commonly used norm is an L_2 norm for which $p = 2$ resulting in a minimization of the mean square error.

In a general inversion approach, synthetic data are generated for an assumed model and compared with the observed data. If the match between observed and synthetic data is acceptable, the model is accepted as the solution, otherwise, the model is changed and the synthetics are recomputed and again compared with the observations. This iterative forward modeling procedure is repeated until an acceptable match is obtained between data and synthetics. Thus in this approach, inversion is viewed as an optimization process in which a model is sought that best explains the observations (Figure 5.1.2).

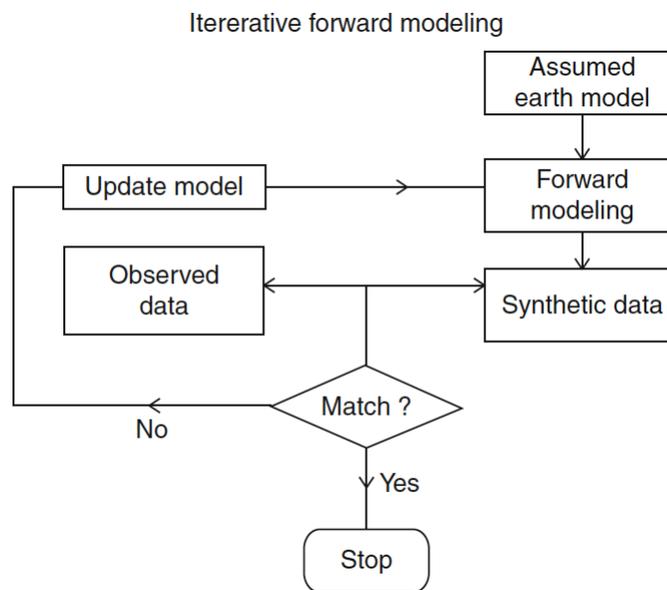


Figure 5.1.2: Principle of model-based inversion by optimization.

Typical examples in geophysics can be given in the gravity or seismological domain. In the first case, the density distribution of the rocks is the source of the gravity field; in the second scenario, the goal is to estimate the location and type of seismic source and a structure from the observed seismic field. The inverse problem whose goal is to estimate Earth model parameters from observations is nontrivial due to the fact that data are almost always inadequate, inconsistent and insufficient.

In the solution of any inverse problem, there are three important questions arising:

1. Does the solution exist?
2. Is this solution unique?
3. Is it stable?

The question of the existence of a solution is related to the mathematical formulation of the inverse problem. From the physical point of view, there should be some solution because the application is the real geological structure inside the Earth. However, from a mathematical point of view, there may be no suitable numerical model to fit our observed data on a given set of models.

The concept regarding the uniqueness of the solution can be illustrated by the following formula. Suppose we have 2 different models, m_1 and m_2 , and two different sources, s_1 and s_2 , that produce the same data d_0 :

$$A(m_1, s_1) = d_0, A(m_2, s_2) = d_0 \quad (5.1.7)$$

In this case, it is impossible to distinguish the two models, m_1 , m_2 , from the given data. The problem of the non uniqueness of the solution in the inverse problem was shown for the first time in Backus and Gilbert (1967), that represents the foundations of geophysical inverse theory.

The last question of solution stability is also a key issue in inversion theory since geophysical data are always contaminated by some noise δd : the problem is whether the difference in the response of different modes is greater than noise level.

According to Hadamard (1902), if all three questions mentioned above have a positive answer, then a mathematical problem can be expressed correctly. Therefore, the mathematical problem is considered to be well-posed, if its solution exists, is unique and is stable; however, unfortunately, most natural science problems are ill-posed. Tikhonov and Arsenin (1977) develop the basis of the theory of ill-posed problem solutions, introducing a regularization method to approximate an ill-posed problem by a number of well-posed problems.

Applications of inverse theory in the Earth Sciences are discussed in many books and reviews including Aki and Richards (1980), Parker (1994), Sen and Stoffa (2013), Tarantola (2005), Moorkamp et al. (2016) and Menke (2018). Here I give a brief overview based on the considerations for 3-D converted wave tomography.

5.1.1 Deterministic global optimization

The goal of an optimization algorithm is to search for the minimum of an objective function¹. Deterministic global optimization is a branch of numerical optimization which focuses on finding the global solutions of an optimization problem (Neumaier, 2004). Deterministic algorithms can handle millions of data and model parameters using advanced simulation tools; these algorithms, based on the concept that future behavior can be predicted precisely from the past behavior of a set of data, provide one final model, where the error is a function of the chosen regularization operator. Most of the local optimization methods are deterministic in nature and the success is largely dictated by the choice of the starting model. Only if the starting model is close to the globally optimal model can we expect to reach the global minimum using local optimization methods.

One of the most common deterministic techniques to find a local minimum of a function is the **gradient (or steepest) descent** (Robbins and Monro, 1951; Gill and Murray, 1981),

¹The objective function can have multiple local minima; the minimum of all the local minima is called a global minimum. In most applications, once an objective function has been defined, the goal is to find the global minimum.

where some steps are taken proportional to the negative of the gradient of the function to minimize at the current point. A more sophisticated evolution of this method is the **conjugate gradient method** (Liu and Storey, 1991), that is useful for the optimization of both linear and non-linear systems. The approach is generally used as an iterative algorithm, applicable to sparse systems that are too large to be managed by a direct search implementation. Figure 5.1.3 shows the evolution path of the gradient sampling algorithm on a typical non-smooth objective function, as compared to the behavior of the classical steepest descent algorithm (Vanbiervliet et al., 2008).

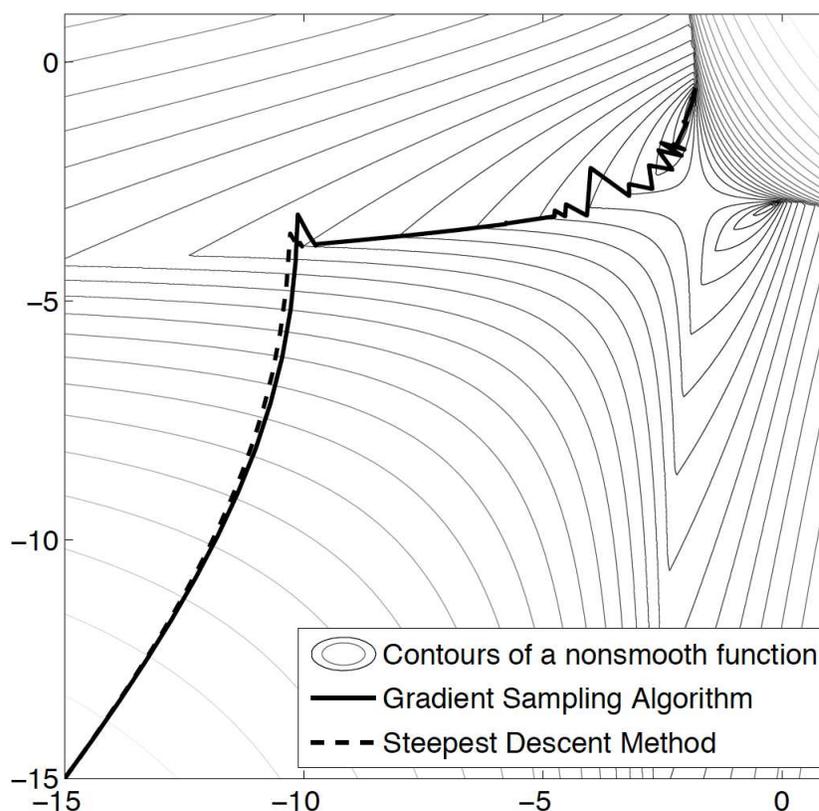


Figure 5.1.3: Comparison of the behavior of the classical steepest descent method (dashed line) and the gradient sampling algorithm (full line) on a typical non-smooth 2-D function (figure from Vanbiervliet et al., 2008).

It is seen that the steepest descent, showed with a dashed line, indeed strands upon reaching a ridge of non-smooth points, whereas the solid line representing the gradient sampling algorithm, is able to proceed towards the non smooth local minimum by following the steepest descent direction.

Another gradient method of optimization includes the **Newton method** (Fischer, 1992)

that corresponds to obtaining at the current point m_n the paraboloid that is tangent to the function $S(m)$ and that has the same local curvature, and jumping to the point where this tangent paraboloid reaches its minimum (Figure 5.1.4).

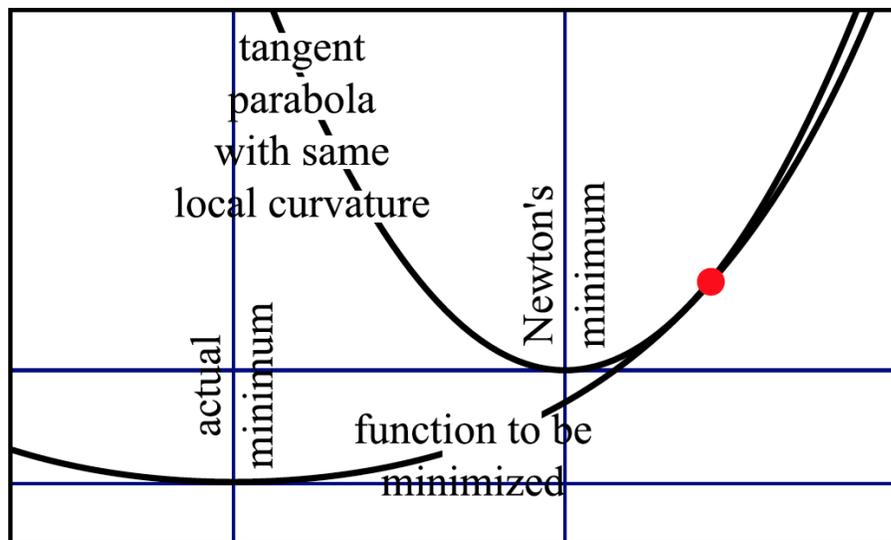


Figure 5.1.4: The Newton method provides the minimum of the parabola that is tangent to the 1-D function to be minimized (from Tarantola, 2005). Red dot indicates the current point.

The main advantage of the gradient-based method is that, when it works, it can be very efficient. The disadvantage is that the local properties of the function to be optimized may be of little interest if the function is complex (Tarantola, 2005).

Deterministic algorithms also include **Tabu search** (Glover, 1986) that uses a local or neighborhood search procedure to iteratively move from one potential solution x to an improved solution x' in the neighborhood of x , until some stopping criterion has been satisfied. At each step worsening moves can be accepted if no improving move is available; moreover, prohibitions are introduced to discourage the search from coming back to previously visited solutions. The implementations of tabu search use memory structures that describe the visited solutions. If a potential solution has been previously visited within a certain short-term period or if it has violated a rule, it is marked as *tabu* (forbidden), so that algorithm does not consider that possibility repeatedly.

Finally, there is another class of algorithms, called **uniform cost search** which operates in a brute-force way. These algorithms represent the best solution for a search problem which does not involve the use of heuristics². In particular, grid search methods represent

²A heuristic method is an approach to find a solution to a problem. The name comes from the ancient Greek word “eurisko”, meaning to “find”, “search” or “discover”.

a manually specified subset of the hyperparameter space of a learning algorithm (Claesen and De Moor, 2015). Since these methods evaluate the objective function at every point in the parameter space, they require an high computational cost. One example of a grid search method is the Zhu and Kanamori (2000) technique, as already mentioned in Chapter 2. For such methods, a solution is possible for a reduced parameter space, but not in the case of more complex problem like the one regarding the Central Alps.

5.1.2 Stochastic optimization

Stochastic optimization methods include randomness, such as generating and using random. In these approaches, the global minimum of a multimodal objective function is measured by a probability density function (*PDF*). The global maximum of this function corresponds to the best data fit model, and locating it would be unfeasible with techniques based on gradient methods or matrix inversion, unless one started near the solution.

Monte Carlo (MC) methods are a broad class of computational algorithms that rely on repeated random sampling to obtain results. Most of the Monte Carlo approaches are commonly associated with a Bayesian probability, and perform a random search of a parameter space, which is a multidimensional region describing the set of all possible values that Earth models can take: each point in the parameter space represents an Earth model (Sambridge and Mosegaard, 2002). For general probability distributions, Monte Carlo³ methods represent a good solution to perform an extensive exploration of the model space. Monte Carlo sampling aims to solve 2 problems of a probability distribution in a large-dimensional space:

1. Locating the regions of significant probability.
2. Sampling the whole of the regions densely enough.

Discovering the location of the regions is the most difficult problem, and due to the great emptiness of large-dimensional spaces, mathematics alone cannot solve it; it is the particular geometry of the problem that may help (Tarantola, 2005). Once one has been able to come close to one of these regions, some techniques as *Gibbs sampler* or *Metropolis algorithm*⁴ are able to perform a *random walk*, which consists in a perturbative sequence of random changes to a point in a multidimensional space.

³These methods were called Monte Carlo by the team at Los Alamos that was at the origin of the Metropolis sampling algorithm.

⁴Gibbs sampler (Geman and Geman, 1987) is a Markov chain Monte Carlo (MCMC) algorithm for obtaining a sequence of observations which are approximately from a specified multivariate probability distribution, when direct sampling is difficult. The Metropolis (or Metropolis-Hastings) algorithm was developed by Metropolis and Ulam (1949), Metropolis et al. (1953) and Hastings (1970). It is a Markov chain Monte Carlo (MCMC) method and has no memory, in the sense that each step depends only on the previous step.

5.1.2.1 Simulated annealing

Simulated annealing (SA) approach was originally developed to simulate the statistical mechanics of systems in equilibrium but very quickly gained attention as a general purpose global optimization method. Physically, annealing consists of heating a solid until thermal stresses are released, then cooling it very slowly in the ambient temperature. Ideally, the substance is heated until it melts, and then cooled very slowly until a perfect crystal is formed.

Simulated annealing (Kirkpatrick et al., 1983; Geman and Geman, 1987) is a numerical method based on Metropolis sampling algorithm of Metropolis et al. (1953), using an analogy between the process of physical annealing and the mathematical problem of obtaining the global minimum of a function that may have local minima. The power of simulated annealing is that it can be used in cases where the model-data relationship is highly nonlinear and produces multimodal data misfit functions.

The SA technique begins with an initial model m_0 , with associated error, or energy $E(m_0)$. It draws a new model m_{new} from a flat distribution of models within the predefined limits. The associated energy $E(m_{new})$ is then computed, and compared against $E(m_0)$. If the energy of the new state is less than the initial state, the new state is considered to be good. In this case, the new model is accepted and replaces the initial model unconditionally. However, if the energy of the new state is larger than the initial state, m_{new} is accepted with the probability of $e^{-(E(m_{new})-E(m_0))/T}$, where T is a control parameter called annealing temperature that controls if the “bad” model should be carried over to the new model (see section 5.2.2). The same process is repeated for a large number of times, with the annealing temperature gradually decreasing according to a predefined scheme.

5.1.2.2 Genetic algorithms

Genetic algorithms (GA) were first used by geophysicists in the early 1990s and fall into the class of Monte Carlo techniques, since they also use random numbers to control components of the search. Like simulated annealing, the metaphor underlying genetic algorithms is a natural optimization process, in this case biological evolution Goldberg (1989). In contrast to the basic form of simulated annealing, which keeps one set of parameters that are continually updated, GA works on an ensemble of sets of parameters, with less emphasis placed on any particular member (Sambridge and Mosegaard, 2002).

5.1.2.3 Neighbourhood Algorithm

Neighbourhood algorithm (NA, Sambridge, 1999a; Sambridge, 1999b) represents a Monte Carlo direct search technique, developed specifically for sampling in geophysical inverse problems. The approach makes use of concepts from the field of computational geometry and bears little resemblance to genetic algorithms and simulated annealing. The main idea with this method is to generate a set of samples, at each generation, whose sampling density function is built from all previously selected model using the neighbourhood approximation.

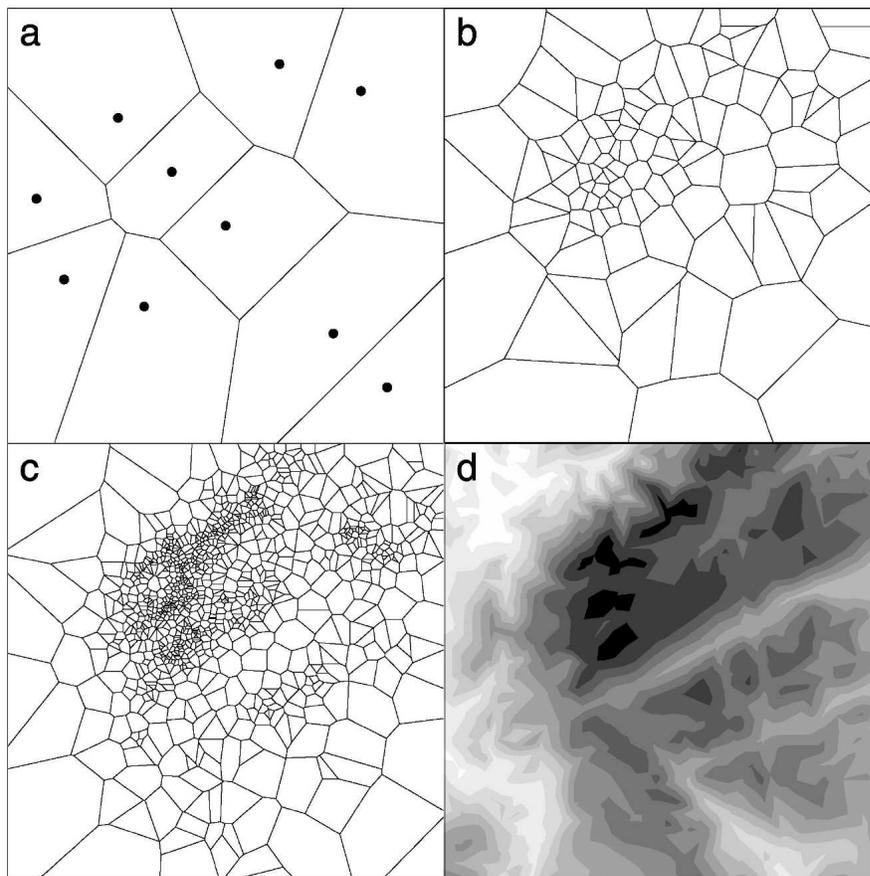


Figure 5.1.5: Voronoi cells drawn around the sampling produced by a neighbourhood algorithm example (Sambridge, 1999a) for a simple 2-D problem. a) 10 quasi-uniform random points and their Voronoi cells. b) The Voronoi cells about the first 100 samples generated by a Gibbs sampler using the neighbourhood approximation. c) Similar to b), but for 1000 samples. d) Contours of the test objective function. Note the higher concentration of Voronoi cells in the darker regions where the fit is high (figure from Sambridge, 1999a).

This is a partition of parameter space into Voronoi cells about each of the previous models; in this way the information in the previous samples drives the search for new models. At

regular intervals the approximation is updated, and sampling can concentrate in multiple regions. The number of vertices of Voronoi cells grows exponentially as the dimension of the parameter space increases, because of the curse of dimensionality (Sambridge, 1999a). Figure 5.1.5 shows an example of the neighbouring sampling algorithm.

At each iteration a choice must be made on how to sample from the current approximate fitness function. Like a genetic algorithm, the neighbourhood approach updates a population of models at each iteration but does so using the Voronoi cell concept to identify “promising regions” of the parameter space.

5.1.3 Deterministic vs stochastic approaches

One important question to address is whether a Monte Carlo technique (like SA, GA, NA) or a linearized approach (based on matrix inversion) is more appropriate for a particular problem. The answer depends on the nature of the relationship between data and model, the number of unknowns the available computational resources (Sambridge and Mosegaard, 2002). As the data-model relationship becomes more complex, the misfit function will also increase in complexity and Monte Carlo techniques will be more advantageous for two reasons. The first is that they avoid local minima better with respect to the deterministic approaches, the second is that Monte Carlo techniques are usually more reliable in appraising the solution, because they avoid derivatives, and they estimate uncertainty by means of model variance and resolution matrix. A very useful framework for comparing different search algorithms is to consider the trade-off between exploration and exploitation (Sambridge and Mosegaard, 2002).

The term *exploration* indicates the improvement in the objective function looking in different regions and different amount of coverage of the parameter space. *Exploitation* is perpendicular to this: we make decisions of where to sample next by only using the previous sampling and sometimes just the current best fit model. Figure 5.1.6 shows an attempt to classify various techniques according to these definitions.

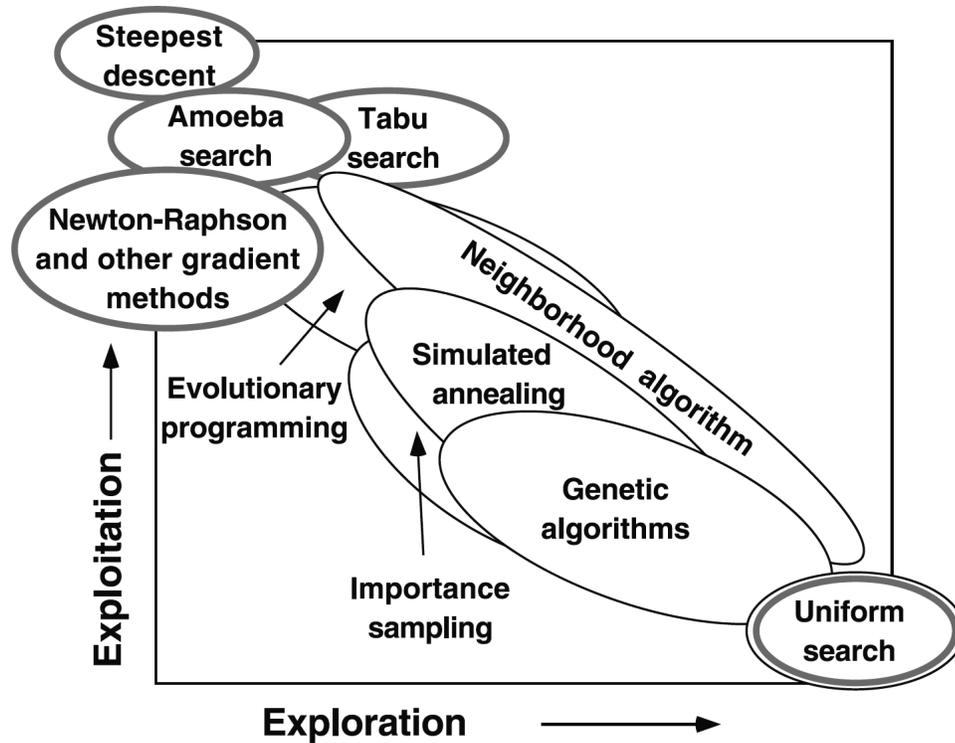


Figure 5.1.6: A schematic representation of various search/optimization algorithms in terms of the degrees to which they explore the parameter space and exploit information. Shaded borders indicate a deterministic method (from Sambridge and Mosegaard, 2002).

Regarding optimization, the rule of thumb is that the more explorative an algorithm is, the less likely it will fall into local minima, but the less efficient it will be at converging on a solution. Conversely, the exploitative algorithms will be more efficient at convergence but prone to entrapment, and hence the final result will depend on the starting point. Examples of methods that lie at the extremes would be a uniform search, which is completely explorative, and a steepest descent algorithm, which is completely exploitative.

5.1.4 Example of nonlinear optimization for an analytical function

Before applying the inversion strategy to the seismic data, in this section I show the efficiency of a few inversion techniques on an analytical function that has more than one minima. In the MatLab software, the command *peaks* produces a highly nonlinear function in 2-D that has many peaks and sinks. This function has the following analytical form:

$$f(x, y) = 3 * (1 - x)^2 * e^{-x^2 - (y+1)^2} - 10 * (x/5 - x^3 - y^5) * e^{-x^2 - y^2} - 1/3 * e^{-(x+1)^2 - y^2}$$

5.1.4.1 Optimization using Gradient Solver

The function above represents a challenging surface to minimize using a gradient based solver. This represents a very challenging surface to minimize using a gradient based solver.

Let's first consider a starting point at $x_0 = [0, 1.5]$. Using the gradient method I find the right solution as reported in Figure 5.1.7.

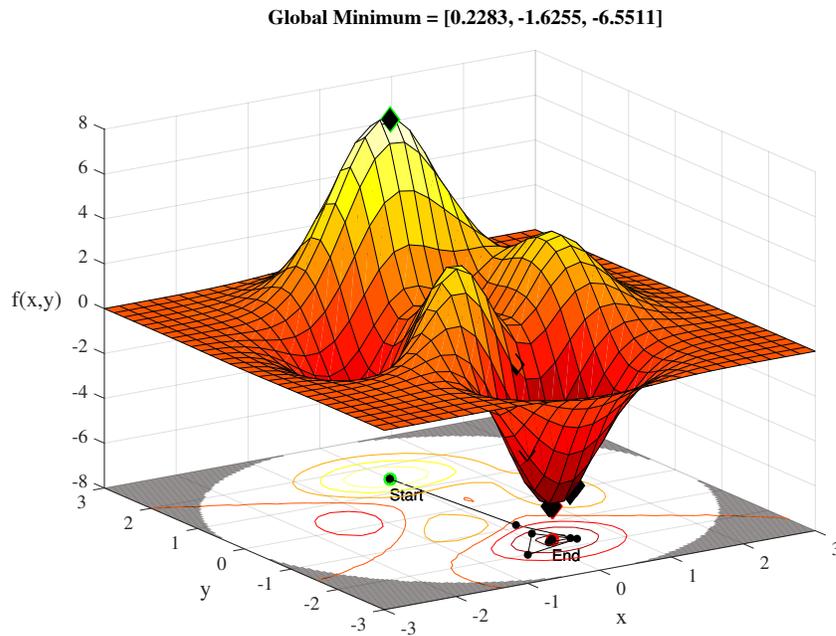


Figure 5.1.7: Minimization of the “peaks” function using a gradient method. In the x - y plane, the green point is the starting point $[0, 1.5]$, while the red point shows the global minimum.

The top plot is a surface plot of the *peaks* function, while the lower plot is a contour plot of the upper plot, with the constrained boundary in grey. If I move the starting point in another region, for example assuming $x_0 = [-3, 3]$, the solution finds a local minimum, as represented in Figure 5.1.8.

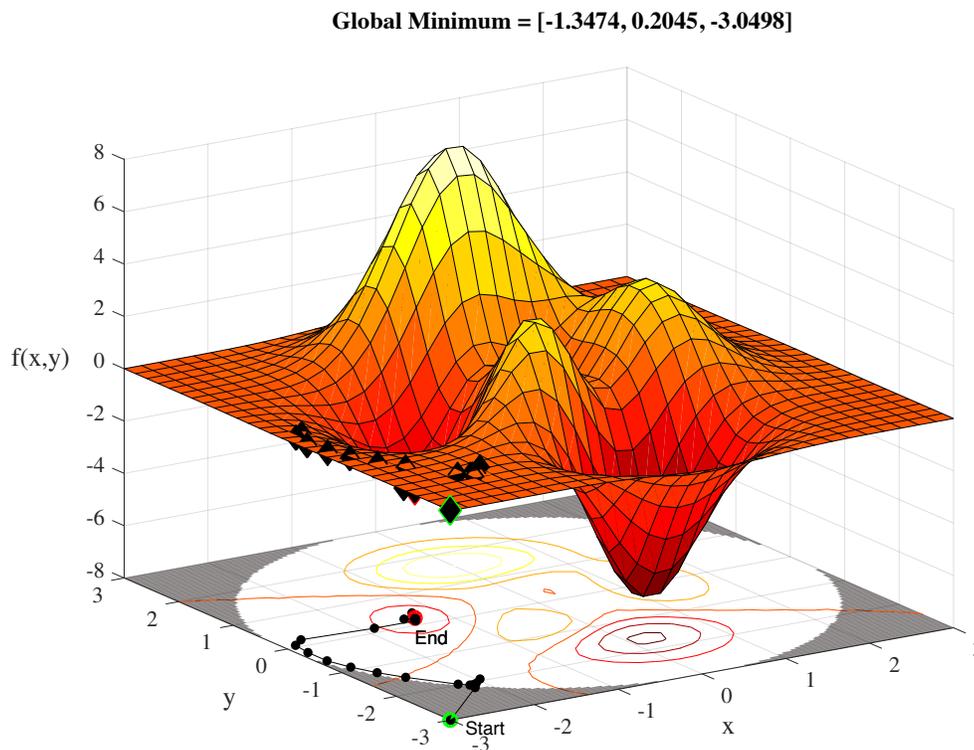


Figure 5.1.8: Minimization of the “peaks” function using gradient method. In the x - y plane, green point is the starting point $[-3, 3]$, while red point shows global minimum.

This is because the gradient solver does not guarantee a global solution for nonlinear optimization problems like this one.

5.1.4.2 MultiStart Optimization

Multistart optimization is a typical approach when we suspect that the problem is nonlinear and we attempt to get around this issue by using multiple starting points: for example, by generating 4 random starting points. In this case the global minimum was found at least once: for this reason, multistart can be an effective method (Figure 5.1.9).

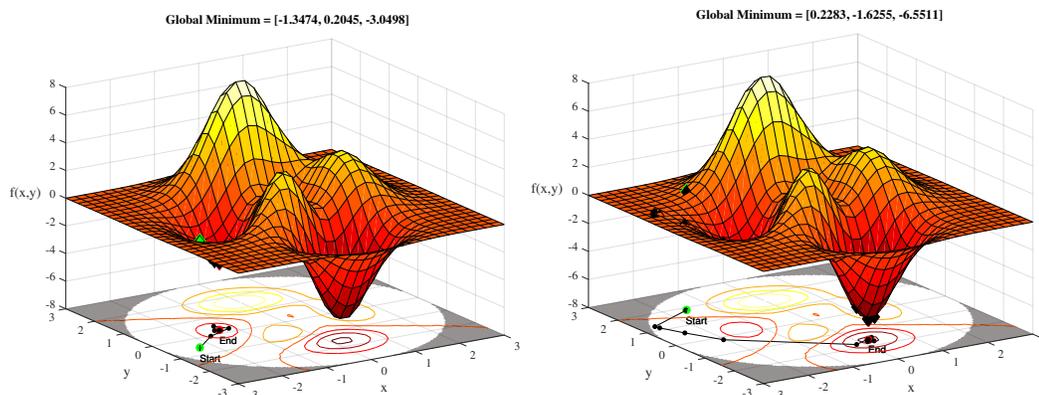


Figure 5.1.9: Minimization of the “peaks” function using MultiStart optimization. On the right, the successful case where the global minimum was found.

5.1.4.3 Optimization using Genetic Algorithms

One of the advantages of the solvers in the Genetic Algorithm is that they are useful for highly non linear problems, but they have the drawback that they are expensive in terms of computational cost and are limited to solving problems of relatively smaller size.

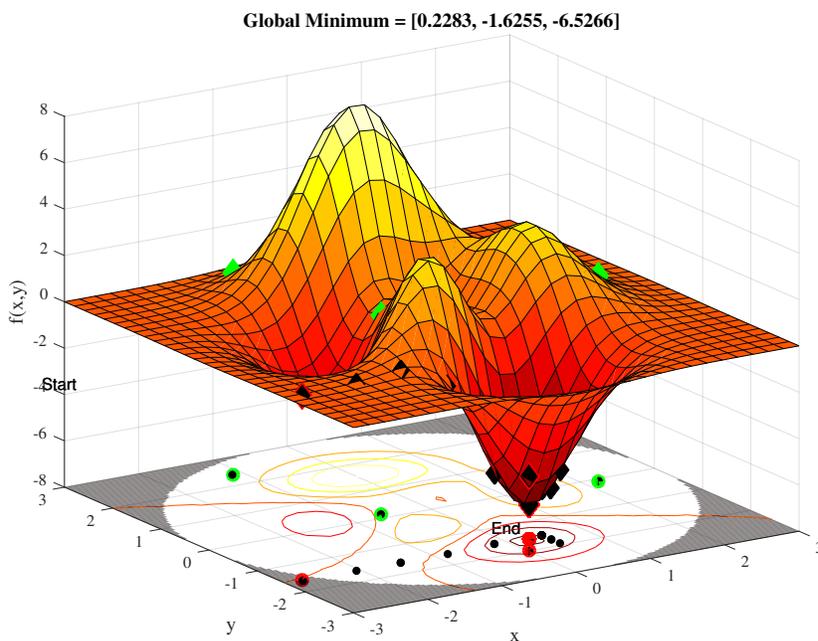


Figure 5.1.10: Minimization of the “peaks” function using the genetic algorithm method. In the x - y plane, green points are the starting points, while 'End' shows the global minimum.

Figure 5.1.10 shows how the GA progresses towards a solution. It uses a stochastic search,

based on an initial population of points (green), and then selects a subpopulation for creating a new population to evaluate. At the end of this process, the global minimum is found.

5.1.4.4 Optimization using Simulated Annealing

The Simulated Annealing solver samples randomly the domain and reduces gradually the search radius around a found minimum; periodically, the solver will reset and start to search in a wider radius. This is the mechanism that allows this solver to jump out of a local minimum and to ultimately find the global minimum (see Figure 5.1.11).

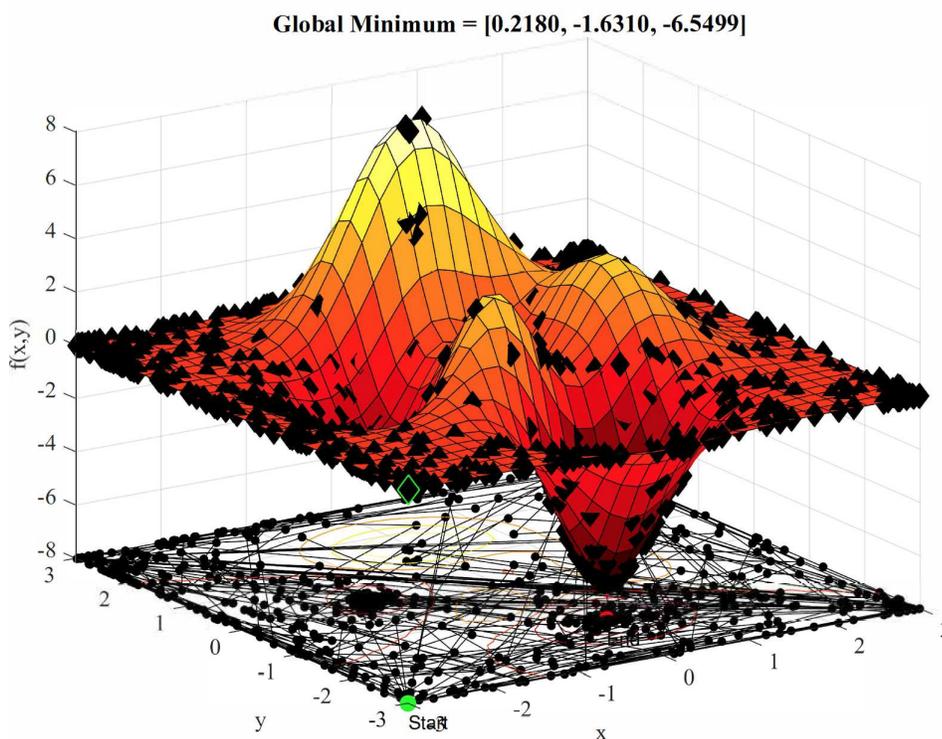


Figure 5.1.11: Minimization of the “peaks” function using the simulated annealing approach. The starting point is in green and the global minimum is shown in red. Black dots represent the tested points.

5.1.4.5 Optimization using Pattern Search

Pattern search (or direct search) is a family of numerical optimization methods that does not require a gradient (Hooke and Jeeves, 1962). Pattern search solver starts from an initial point provided for the input and contracts the search radius as it explores the domain for the maximum value; this algorithm is very efficient on a highly rough surface like this current example (Figure 5.1.12), on which gradient based solvers do not work well.

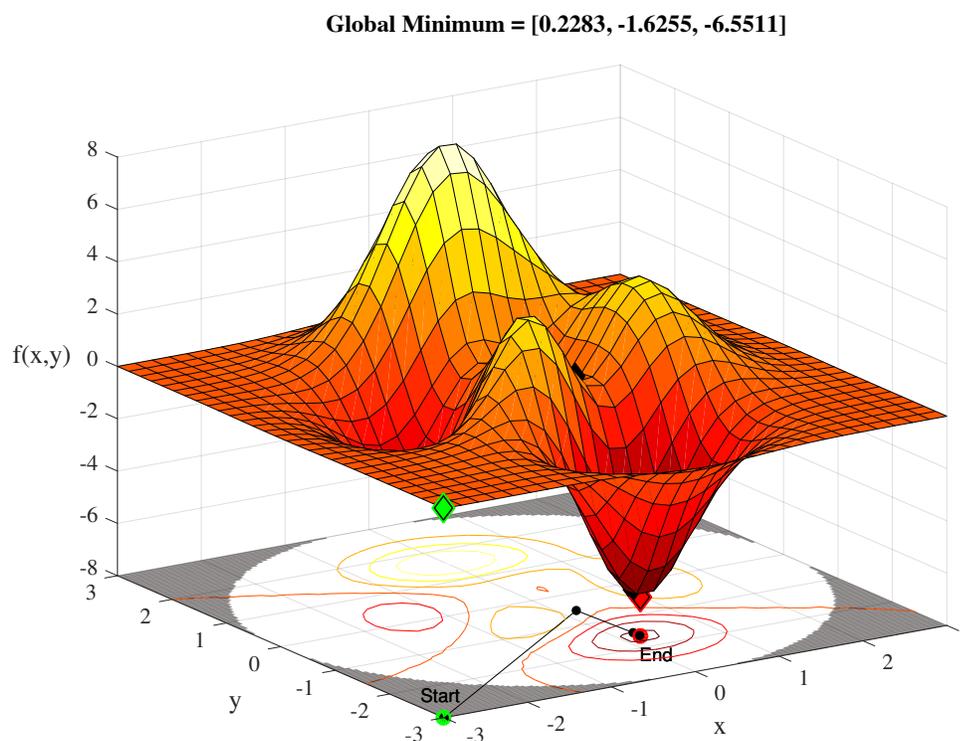


Figure 5.1.12: Minimization of the “peaks” function using the pattern search method. In the x - y plane, the green point is the starting point $[-3, 3]$ and the red point shows the global minimum.

5.1.4.6 Choice of the Solver

Table 5.1 shows the performance and cost overview of various solvers on an analytical function example. It is clear that gradient based solver is the best approach, at least for non-linear functions. However, as shown in section 5.1.4.1, the right solution is not always found. Note that even if run with a random starting point, it is still faster than other methods and often results in fewer function calls.

Among the gradient-free methods, we can see that the order of preference in terms of function calls is PS, SA, GA. And if we considered time to solution, it would be SA, PS and then GA.

Solver	Function calls	Time (s)
Gradient (lucky guess)	47	0.174108
Gradient (random start)	189	0.843512
Genetic Algorithm	5220	17.886585
Simulated Annealing	2406	1.226381
Pattern Search	453	3.991359

Table 5.1: Results from different optimization solvers on an analytical functions.

A good recommendation is that if you think you have a nonlinear problem (like the receiver function minimization), you can't use a gradient based solver. From the results of this analytical example, we find out that simulating annealing is a good compromise between PS and GA. In this analytical example, simulated annealing is a good compromise between PS and GA, taking into account both solution accuracy and computational cost.

5.2 Setup for 3-D converted wave tomography

In this section I describe the practical implementation of the inversion of converted rays to retrieve 3-D crustal structure. The main elements, discussed by section, are the following:

- *Spatial separation of the inverse problem*: since the matrix linking the model to the data is sparse, i.e. only the nearest receiver functions affect the structure and velocity of a given node, I proceed to a node-wise inversion visiting every node in a consecutive order. This is performed using a Traveling Salesman itinerary, and multiple rounds are performed to let the inversion result stabilize.
- *Inversion at each node*: the corresponding structure and velocity parameters are inverted, to minimize the objective function between observed and synthetic data. Synthetic tests are performed to assess the performance of the inversion and to set its parameters.
- *Spatial grouping of rays*: during the inversion, receiver functions following similar ray-paths are grouped into a bundle.
- *Synthetic model recovery test*: before applying this approach to real data in the Alpine area, I test the method and its limitations simulating observations with a known synthetic model, to verify that the approach is able to reconstruct it.

5.2.1 Spatial separation of the inverse problem

The inversion method I propose takes advantage of the fact that the matrix G linking model to the data is sparse; this is because only the nearest receiver functions affect the model properties at a given node. Therefore, as the matrix is *sparse*, so most of its elements are zero, I can run inversion locally at each node, instead of a large inversion for the entire model domain. This requires visiting every node going neighbour by neighbours, and to do at least 2 rounds to stabilize the results.

I divide the model space into a 25x25 km mesh in map-view (Figure 5.2.1), and those nodes to which at least one RF raypaths belongs. At this mesh size, any RF path across the crust remains within an area bounded by 2-by-2 mesh elements.

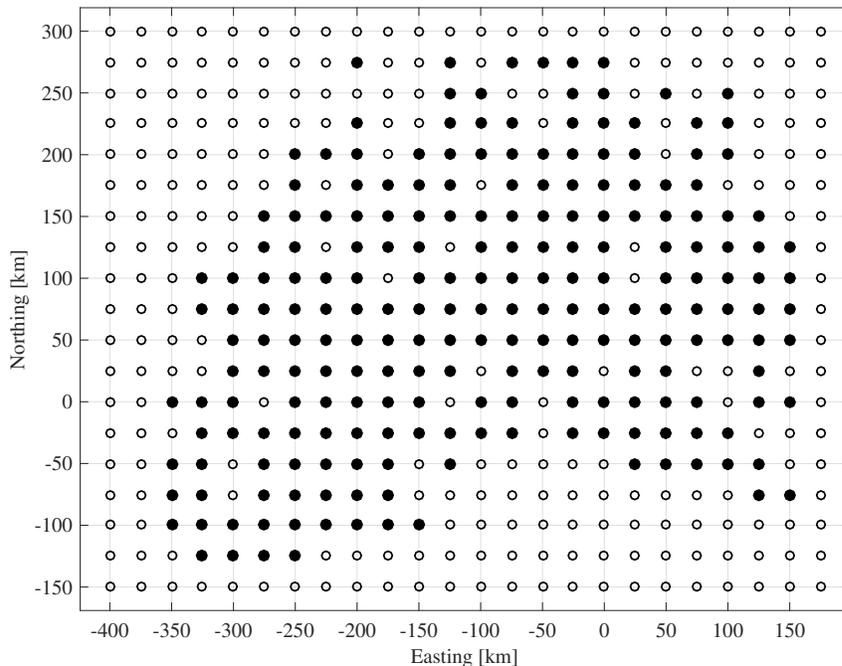


Figure 5.2.1: Grid node map at Moho depth in the study area for the real dataset. Black nodes indicate at least 1 ray, white nodes have no data available. Origin is at $10^\circ E$, $46^\circ N$.

In order to visit every node only once and to visit all nodes by the end of the tour, I adopt the traveling salesman problem (TSP). This is a non-deterministic polynomial-time hardness problem in combinatorial optimization (*Biggs, 1986*) and is important in operations of research and theoretical computer science. TSP solves for the most efficient trajectory possible given a set of points and distances that must all be visited⁵. In our case, the problem can be applied to the most efficient route to travel between various nodes going neighbour to neighbours. Figure 5.2.2 represents a possible path applying the TSP algorithm to the grid of the study area. In this case, I first run the inversion at a selected initial node with available data, and then I move up and right in clockwise sense until I reach the center of the grid.

⁵The origins of the traveling salesman problem are unclear. A handbook for traveling salesman from 1832 mentions the problem and includes example of tours of cities through Germany and Switzerland, but contains no mathematical treatment (Voigt, 1831).

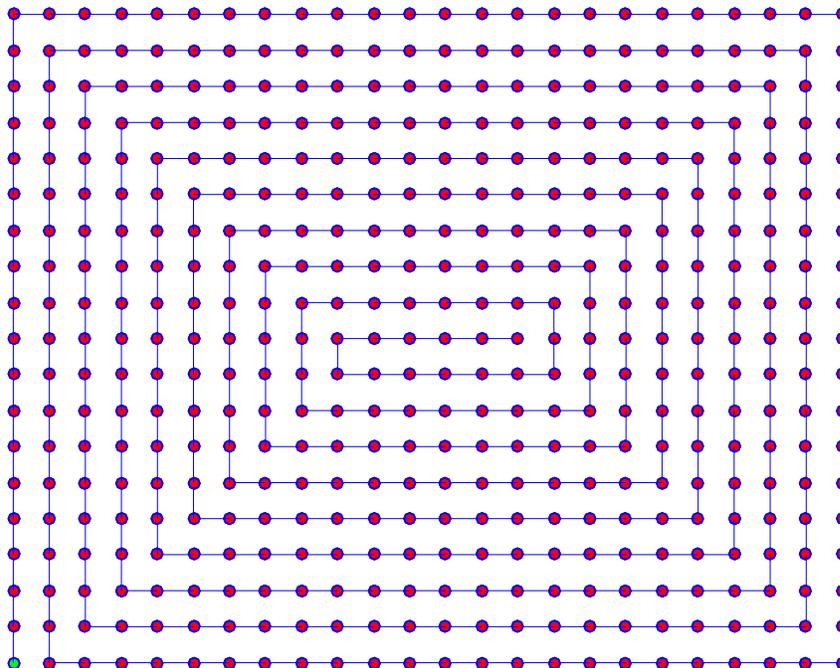


Figure 5.2.2: Possible path applying the TSP algorithm to our grid. In this case, the path chosen is a kind of spiral which starts at the bottom left corner (green point) and converges to the center of the grid.

For our purposes, I applied the TSP following different paths at each round of the inversion, so that the final inversion result is not influenced by the order of the nodes visited during the same inversion tour.

5.2.2 Inversion at a node

As mentioned in section 5.1, I adopt the simulated annealing technique where, at each iteration, a new point is generated randomly. The distance of the new point from the current is based on a probability distribution proportional to the temperature. The simulated annealing algorithm accepts all new points that have a lower value of the objective function as currently, but also, with a certain probability, points that raise the objective function value.

In the MatLab implementation of the simulating annealing, the option *temperature* specifies how the temperature will be lowered at each iteration over the course of the algorithm.

- *Initial Temperature* represents the start of the algorithm. The initial temperature can be a vector with the same length of unknowns. I set this parameter to T_0 (standard value), as I do not have *a priori* constraints.
- *Temperature Function* is the function used to update the temperature schedule. Let T_0

be the initial temperature and k denote the annealing parameter (which is the same as the iteration number until reannealing), the possible options are:

- $@temperatureexp = T_0 * 0.95^k$
- $@temperaturefast = T_0/k$
- $@temperatureboltz = T_0/\ln(k)$

These options are shown in Figure 5.2.3.

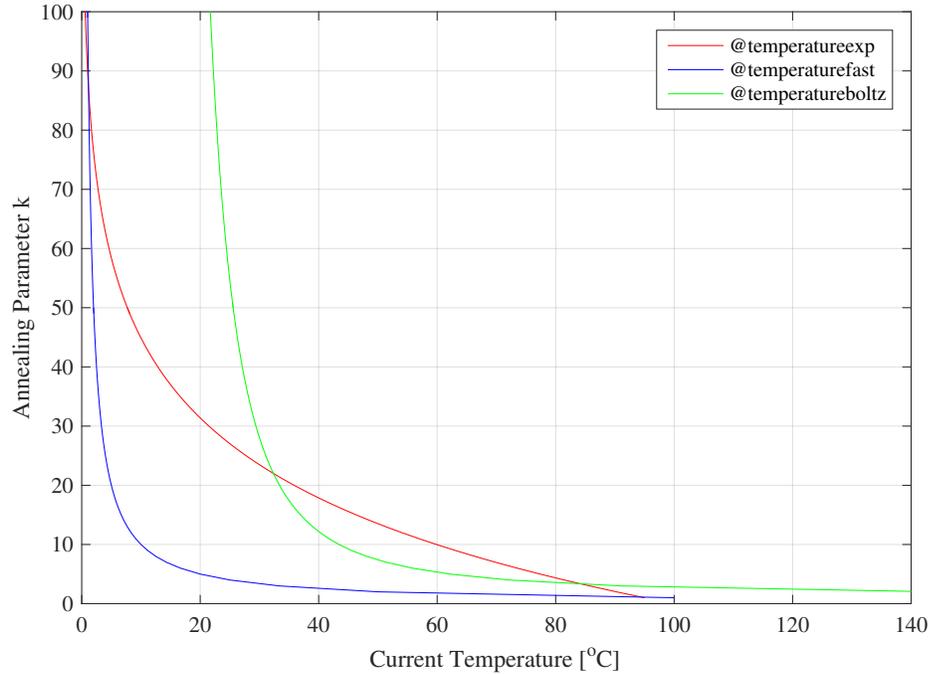


Figure 5.2.3: Graph of the temperature in the SA algorithm that shows the behavior of the current temperature as a function of the annealing parameter k .

For the application in this study, since I do not have a priori information, I choose the temperature function $@temperaturefast$, which is the default value in MatLab.

5.2.2.1 Repeatability test on synthetic model

Before application to the observed data (Chapter 6), I first tune the parameters for the inversion in order to understand the limits and the repeatability of the outcomes. I consider a synthetic 1-D velocity (with depths taken from *IASP91* but different velocity values) to produce a synthetic RF with 5 parameters.

$$x_0 = [35, 20, 1.73, 1.78, 0.35]$$

where the elements of the vector are respectively the Moho depth [km], Conrad depth [km], V_p/V_s of the upper crust, V_p/V_s of the lower crust and $\Delta V_p \text{ Conrad}$ [km/s].

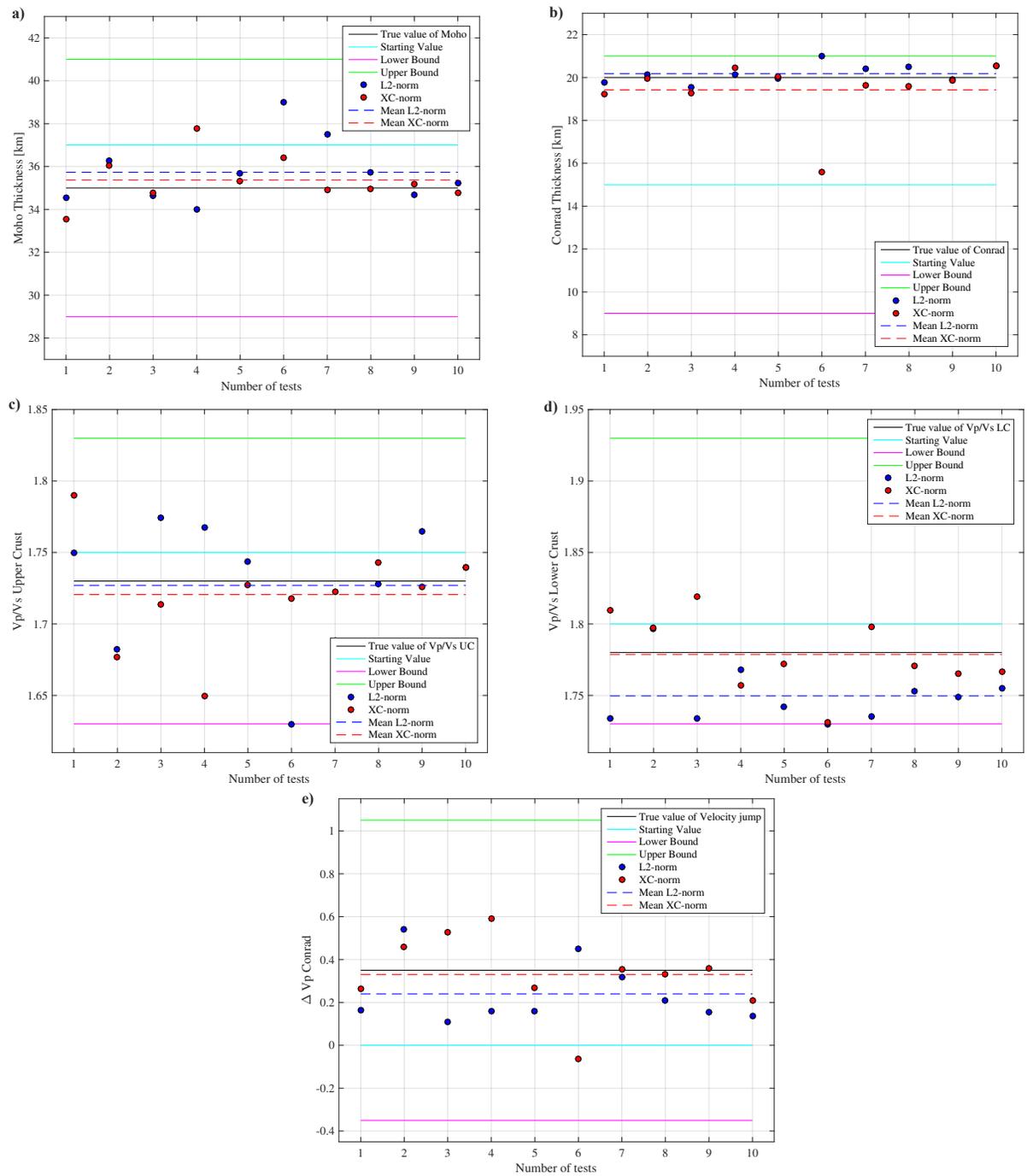


Figure 5.2.4: Comparison between L2-norm and cross-correlation. a) Moho; b) Conrad; c) V_p/V_s ratio at the upper crust; d) V_p/V_s ratio at the lower crust; e) $\Delta V_p \text{ Conrad}$.

This represents the solution to reach if the data are not affected by noise or algorithm

problems. In order to test the repeatability of the results, I performed several tests using the vector $x_i = [37, 15, 1.75, 1.80, 0]$ as a starting point, and considering as lower bound the values $l_b = [29, 9, 1.63, 1.73, -0.35]$ and, as upper bound the vector $u_b = [41, 21, 1.83, 1.93, 1.05]$. The target of the objective function is to minimize a misfit between theory and observations (Igel, 2017), and to measure waveform fitting I test 2 norms: the L2-norm and the cross-correlation norm. Figure 5.2.4 show the recovered values of the 5 variables in x_0 obtained using the two norms in 10 synthetic tests.

For the discontinuity depths (Figure 5.2.4 a and b), the average values for the Moho are respectively 35.73 km for L2-norm and 35.37 for cross-correlation (real value is 35 km), while the average value for the Conrad is 20.18 km and 19.42 (real value is 20 km). During the 10 tests, only one value deviates significantly from the true value (test n.6, which has a particular effect on the Conrad thickness when using the cross-correlation norm).

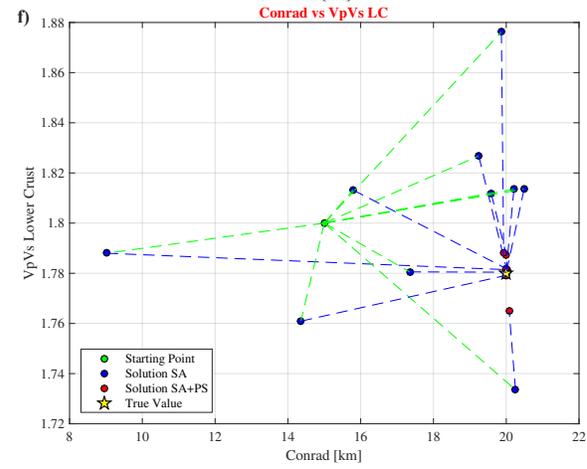
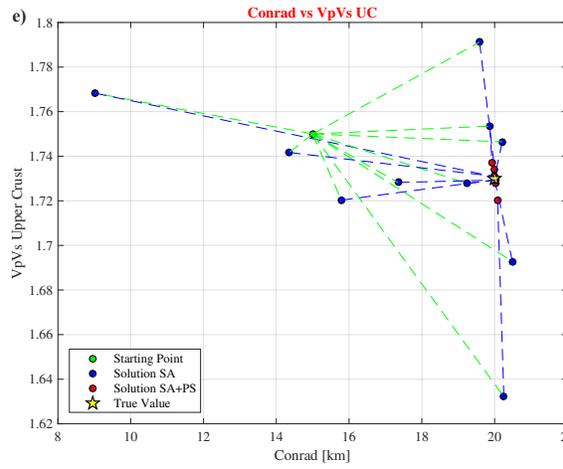
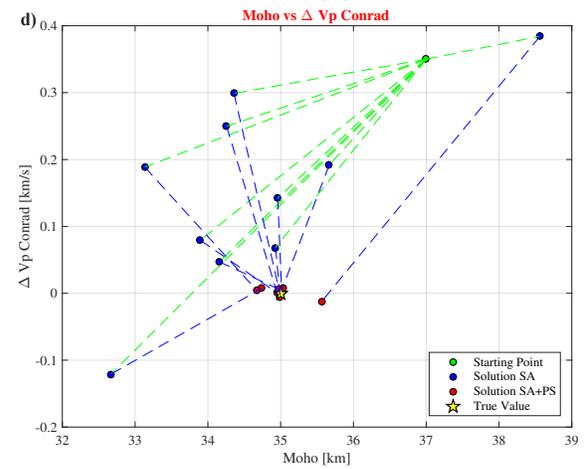
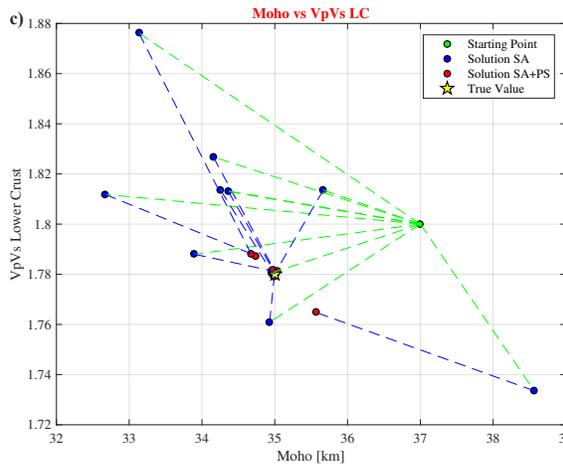
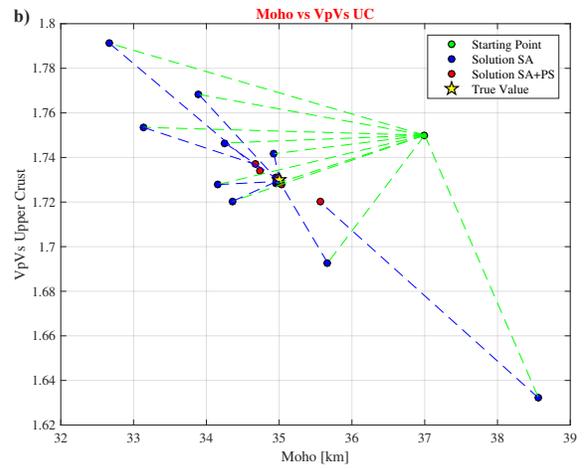
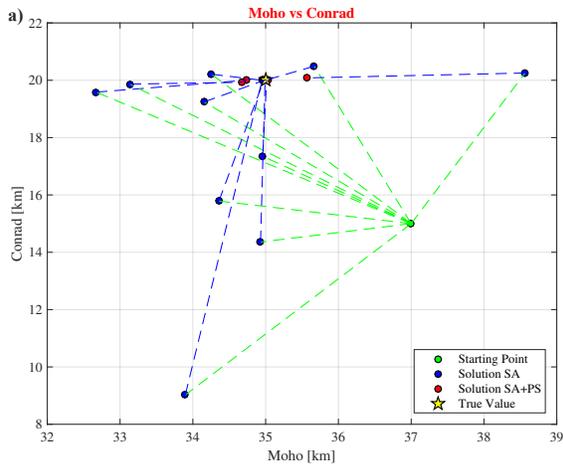
Regarding the velocity values (Figs. 5.2.4 c, d and e), the spectrum of variations seems to be wider with respect to the spectrum of thickness variations, and the recovery of the V_p/V_s and $\Delta V_p \text{Conrad}$ values is not really satisfactory. Several other tests have been carried out considering no velocity jump at the Conrad for the synthetic model whether the technique is capable of verify that the obtained results show similar patterns as above.

In light of these results, I decide to expand the inversion algorithm to a potentially better approach, and, for that, to use the L2-norm as objective function.

5.2.2.2 Adding Pattern Search

As mentioned in section 5.1, using the *patternsearch* solver after the Simulating Annealing algorithm, makes the solution more stable. Pattern search methods (Powell, 1973; Lewis and Torczon, 1999) are derivative-free methods where the current iterate is updated by sampling the objective function at a finite number of points along a suitable set of search directions, and the aim of the sampling is to find a decrease of the function value. More details of the pattern search algorithm adopted in this work are described in the section 5.2.2.3.

To demonstrate the progress of the solution quality when applying this solver as well, I performed several tests with a synthetic model, as in the previous sub-section, with $x_0 = [35, 20, 1.73, 1.78, 0]$. The results and the trade-off of the parameters are shown on Figure 5.2.5.



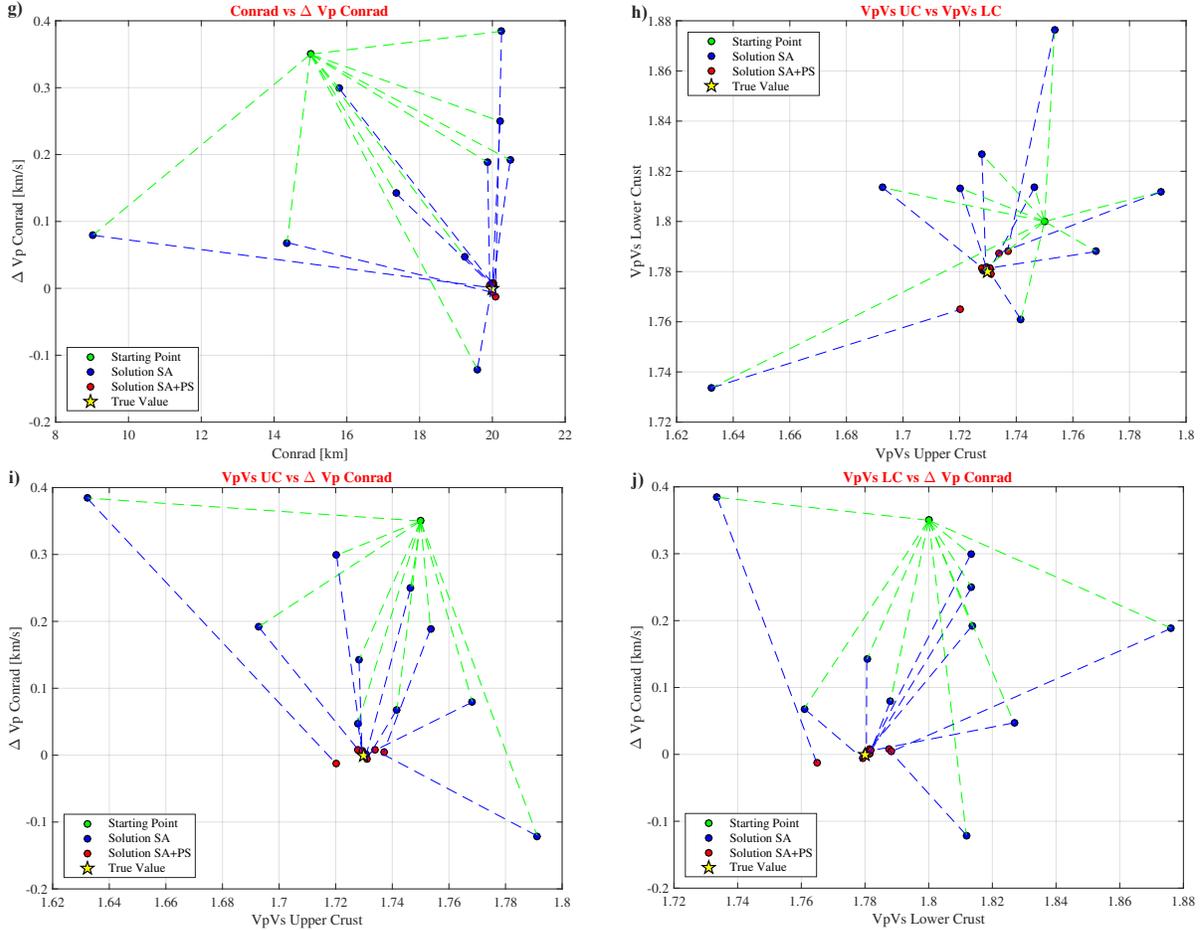


Figure 5.2.5: Solution of ten 1-D synthetic recovery tests (from a to j) for 5 parameters, using simulating annealing alone (blue) and simulating annealing combined with pattern search algorithm (red). See axes labels for shown parameters.

The results on these 10 tests clearly demonstrate that the pattern search algorithm efficiently improves the RF inversion. Moho and Conrad depth are recovered well within 0.5 km, and Vp/Vs within 0.01. The solution computed with SA+PS solver is always closer to the real value I want to replicate than SA alone. For this reason, I decide to use a pattern search solver after the simulated annealing computation in the implemented new algorithm.

5.2.2.3 Pattern Search and its polling description

The *pattern search* algorithms starts from an initial point in the parameter space, and polls its neighbourhood in every direction of the space by mapping the misfit values compared to the initial one. In two dimensions, this resembles a poll along the cardinal directions (East, North, West, South), and the similar principle is implemented in higher dimensions.

Starting from the initial point, the misfit value is evaluated at D_{mesh} distance (default value: 1) in all directions in the order specified above. As soon as a point with lower misfit is found, the polling is *successful*: this point becomes the new current point, and D_{mesh} is multiplied by the expansion factor (default value: 2) for broader search. If, at a given point, none of the neighbours has lower misfit, the polling is unsuccessful: in this case D_{mesh} is multiplied by the contraction factor (default value: 0.5) for coarser search from the same point. The pattern search continues until a given mesh size is reached ($1 \cdot 10^{-6}$ in the *Mesh Options* of MatLab).

The strategy in the directional search can be chosen to be *complete*, meaning that all neighbours are polled before the next point is chosen. This avoids the bias of ordering the directions, at the cost of higher computation time. Synthetic recovery tests performed on a 1-D velocity model demonstrate that full polling of all neighbours produces a better result (Figure 5.2.6), and that the additional computational cost is acceptable.

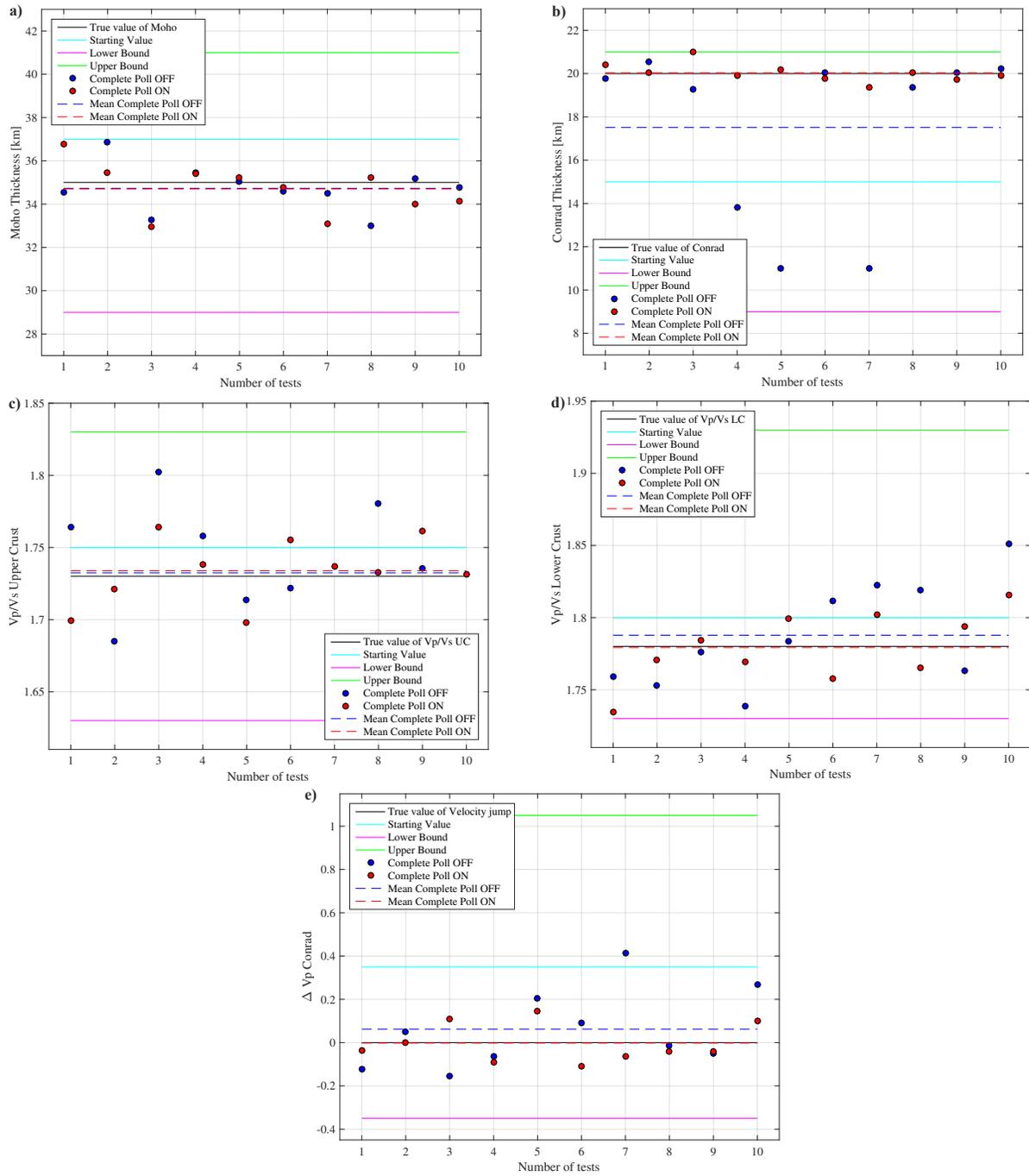


Figure 5.2.6: Comparison between incomplete (blue) and complete (red) poll. **a)** Moho; **b)** Conrad; **c)** Vp/Vs ratio at the upper crust; **d)** Vp/Vs ratio at the lower crust; **e)** ΔV_p Conrad.

A common characteristic we observed from this test is that is preferable to use the option *Complete Poll ON* to reach the results accurately.

5.2.2.4 Number of iterations

Another important decision to take before working with real data is the number of iterations to use during the inversion procedure.

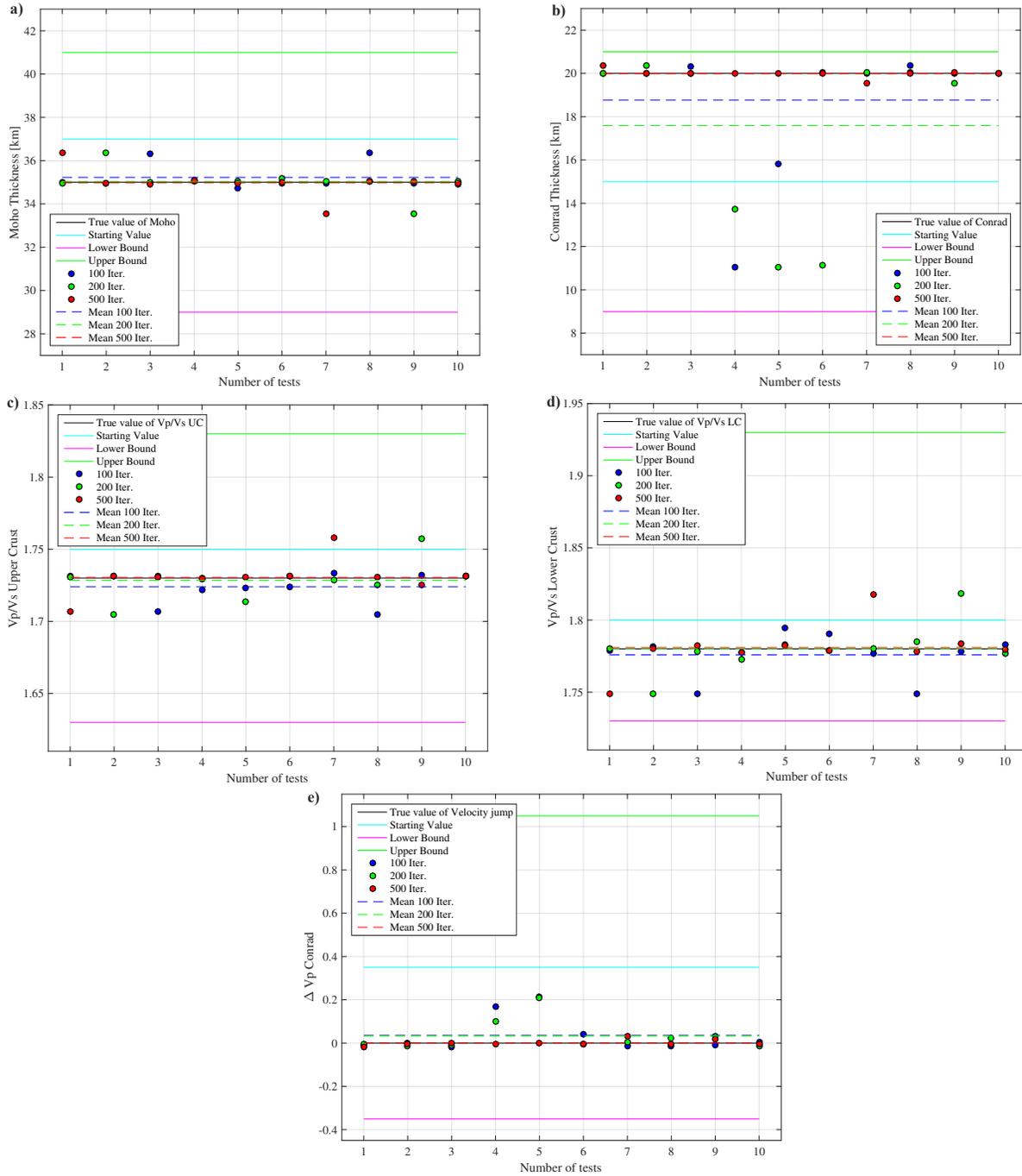


Figure 5.2.7: **a)** Moho; **b)** Conrad; **c)** Vp/Vs ratio at the upper crust; **d)** Vp/Vs ratio at the lower crust; **e)** ΔV_p Conrad considering 100 (blue), 200 (green) and 500 (red) iterations.

Also for this case, I run ten 1-D synthetic RF recovery tests using a known velocity model. The results are shown in Figure 5.2.7. What is seen from the tests is that with 5 parameters to invert 500 iterations generally (but not always) recover the input values. Therefore, 500 is the minimum iteration number to carry out. Clearly, this number must be increased when working on real data, which are always more complex than synthetic examples.

5.2.3 Spatial grouping of rays

The proposed inversion method is adapted to treat as many RFs as in the dataset. However, when rays geometries are very similar, it is a reasonable step to group them together into a stack “or bundle”. Figure 5.2.8 shows the ray coverage map of the real dataset at the Moho depth in the Alpine area, with 25x25 km cells along the X-Y directions (24x19 cells). The ray coverage reflects the station configuration in the Central Alpine zone with high number of rays in the center of the area (up to 765 rays per cell) and a empty cells in the south-eastern part of the model.

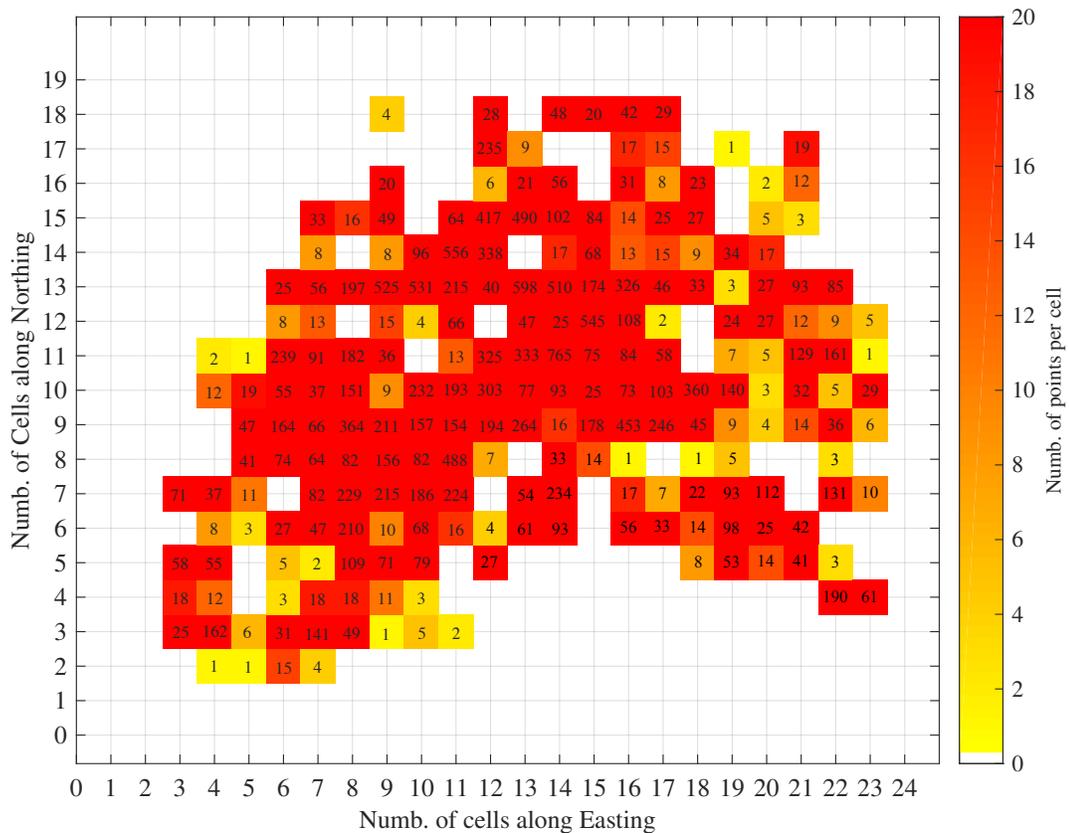


Figure 5.2.8: Number of rays crossing each cell in the study area at the Moho level. Red cells are crossed by at least 20 points.

In order to reduce the computation time, I stack together traces that have similar characteristics, which also reduces noise and improves overall data quality. This process is very common in geophysics and is often applied to the receiver functions method (e.g. Kumar et al., 2010; Sippl et al., 2017).

5.2.3.1 Bundle division

Around each node, what is new in our proposal is to stack RFs by similar raypaths, I divide the 25x25 km area into 25 sub-blocks. The choice of the dimension of the sub-blocks (5x5 km) was made because it is compatible to the Fresnel zone. I also group data into 6 sectors, taking back-azimuth bins every 60° (Figure 5.2.9), which can be changed but is a good compromise between achieved results and computation time.

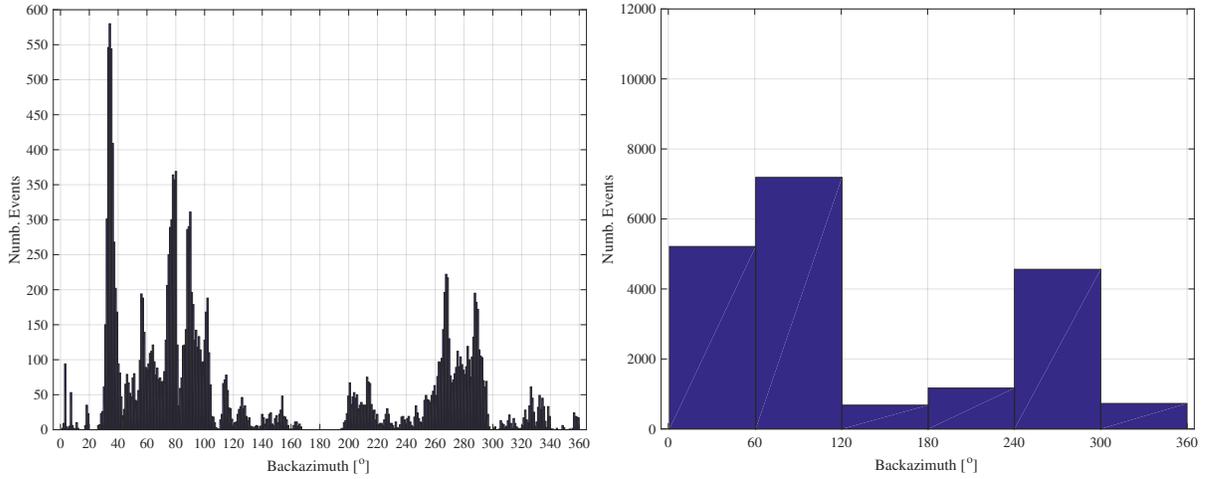


Figure 5.2.9: Back-azimuth distributions for all events available in the whole dataset (left) and histogram of the events grouped every 60° (right).

The discretization into back-azimuth is also reported in Table 5.2. As observed, the vast majority of the rays (>60%) is between 0 and 60°.

Backazimuth [°]	Number of Events (% tot)
0 - 60	5 163 (26.43%)
60 - 120	7 234 (37.03%)
120 - 180	683 (3.50%)
180 - 240	1 164 (5.96%)
240 - 300	4 565 (23.27%)
300 - 360	728 (3.73%)

Table 5.2: Number of events grouped into Back-azimuth families.

This step of grouping traces into bundles can be made because rays that come from the same back-azimuth and within the same sub-block sample the same portion of the Earth crustal structure. To demonstrate the stacking operation, I show a node with numerous rays.

Figure 5.2.10 **a)** represents the piercing points for the node in the center of the diagram, where the colormap indicates the back-azimuth of the RFs, which reflects the different spatial distributions of the teleseismic earthquakes. Initially, 531 rays cross this block.

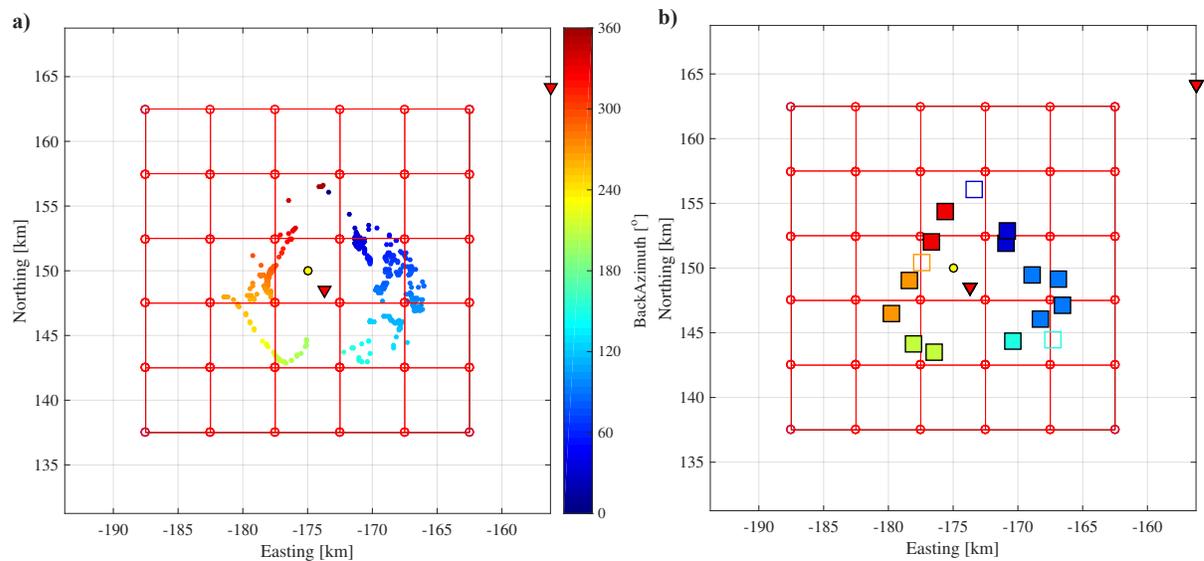


Figure 5.2.10: **a)** Example of piercing points at the Moho interface for the node $(i, j) = (10, 13)$. Stations are shown with the reverse red triangles, the node at the center of the block is represented with filled yellow circle, red squares represent the sub-blocks and the color of the piercing points indicates the back-azimuth orientation. **b)** Filled colorful squares indicates the new piercing points after the bundling. Empty squares indicate that the bundling is not done since because there is only 1 trace which is kept.

Given 25 sub-blocks and 6 sectors, the slowest scenario is 150 piercing points at any node. In the example considered here the effective number of bundles is 13 (5.2.10 **b)**).

In the best case scenario for ray coverage, and the worst for the efficiency of bundling, for each node at each discontinuity I can have 150 (25*6, number of sub-blocks*number of back-azimuth sectors) piercing points. This represents a considerable reduction on the number of piercing points, especially in the zones with a very good data coverage.

3-D spatial information of the ray-path as well as the ray parameter is averaged for each bundle. The bundling operation reduced the dataset to invert from 19 602 RFs to 1 795 bundles, which represents a 10-fold decrease.

5.2.3.2 Threshold on number of rays per bundle

In order to guarantee result robustness and decrease the computation time, I decide to introduce a threshold on the number of rays for each bundle (Figure 5.2.11).

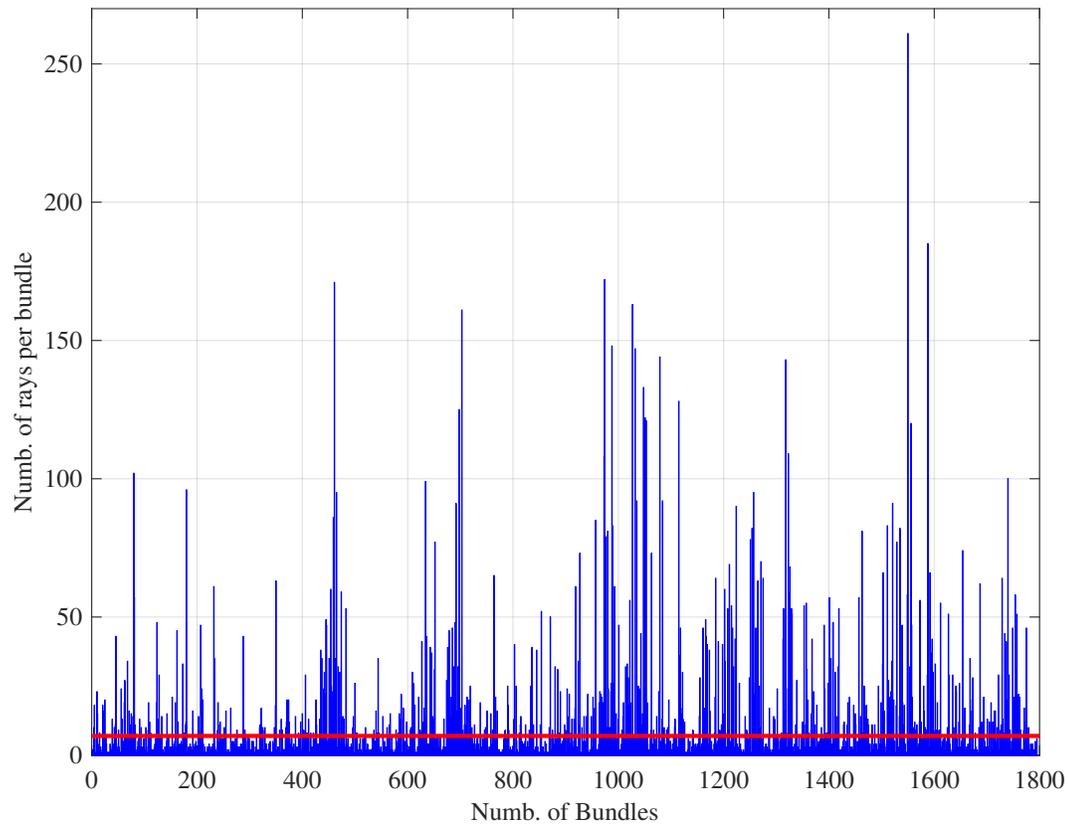


Figure 5.2.11: Number of rays per bundle. Red line shows the minimum threshold I set (7 rays per bundle).

Setting the threshold to 7 rays per bundle (red line on Figure 5.2.11), I obtain 613 bundles for a total number of 16 226 rays, which constitutes more than 80% of the number of rays I have at the beginning. Figure 5.2.12 shows us what is the real ray coverage on the study area, after the bundling.

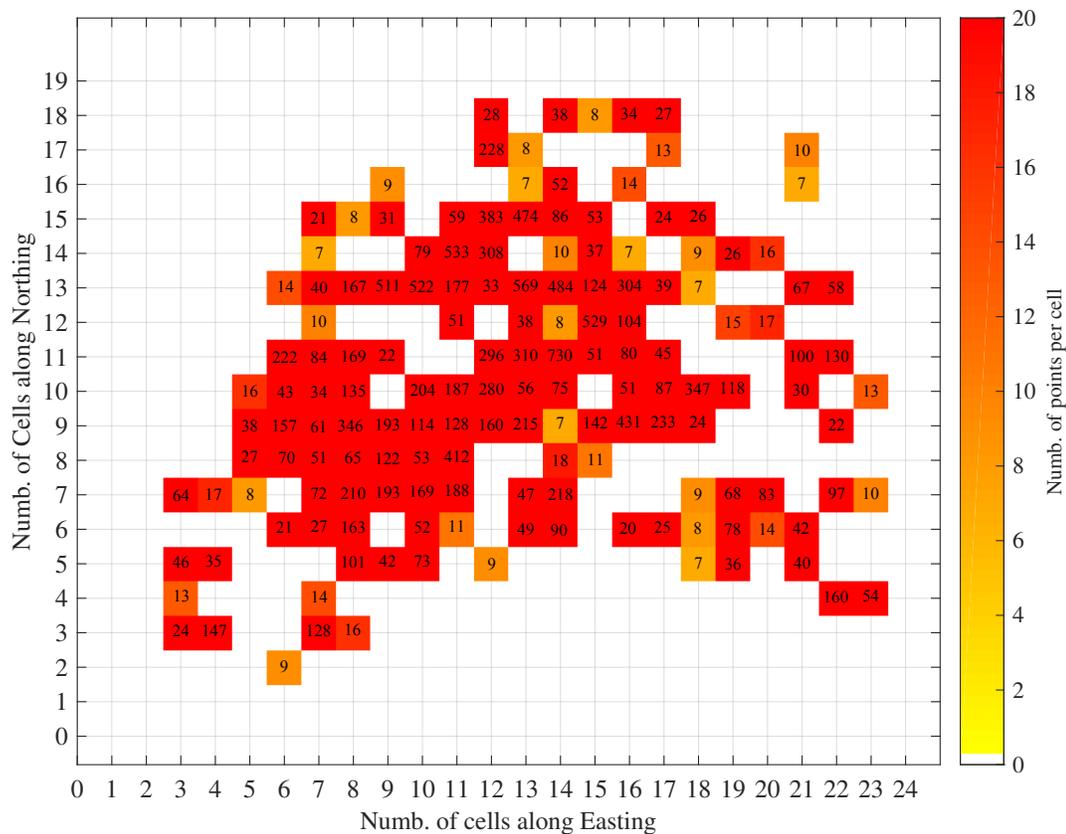


Figure 5.2.12: Number of rays crossing each cell in the study area at the Moho level after bundling. Red cells are crossed by at least 20 points.

Compared with Figure 5.2.8, the ray coverage is less dense, especially in the areas that are close to the plate boundary between Europe and Adria.

5.2.4 3-D Synthetic inversion recovery test

To test the sensitivity and the performance of the inversion approach proposed here, I carry out a 3-D synthetic recovery test using a known velocity model, and synthetic RF data generated from that model. The 3-D model is extended from a 1-D model, with $x_0 = [35, 20, 1.73, 1.78, 0]$ values corresponding to Moho depth [km], Conrad depth [km], upper crustal V_p/V_s , lower crustal V_p/V_s , V_p jump across the Conrad [km/s], with V_p in each layer fixed and constant. The synthetic data includes as many traces as the real dataset, and is bundled the same way.

The initial point of the inversion is at $x_i = [37, 15, 1.75, 1.80, 0]$. The inversion is run with 500 iterations, using the SA+PS inversion at each node, and TSP across the area. The final result obtained at a randomly node is $x_s = [35.0137, 19.4902, 1.7321, 1.77809, 0.04296]$. The recovery of the discontinuity depth, as in the 1-D case, is excellent. The recovery of the V_p/V_s ratios and of the velocity jump across the Conrad is also good, although the relative errors are larger than for the discontinuity depths. Figure 5.2.13 shows the waveform fit between the synthetic data and the best fit at the end of the inversion, and demonstrates that the newly implemented approach technically works.

5.2.4.1 Waveform matches

In Figure 5.2.13 is shown the waveforms misfit between observed and calculated RF for a single bundle after 500 iterations of simulated annealing and pattern search solver.

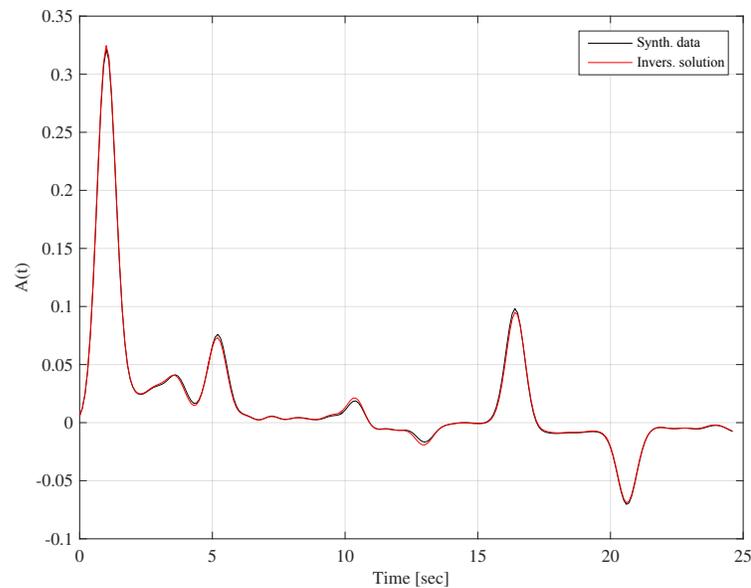
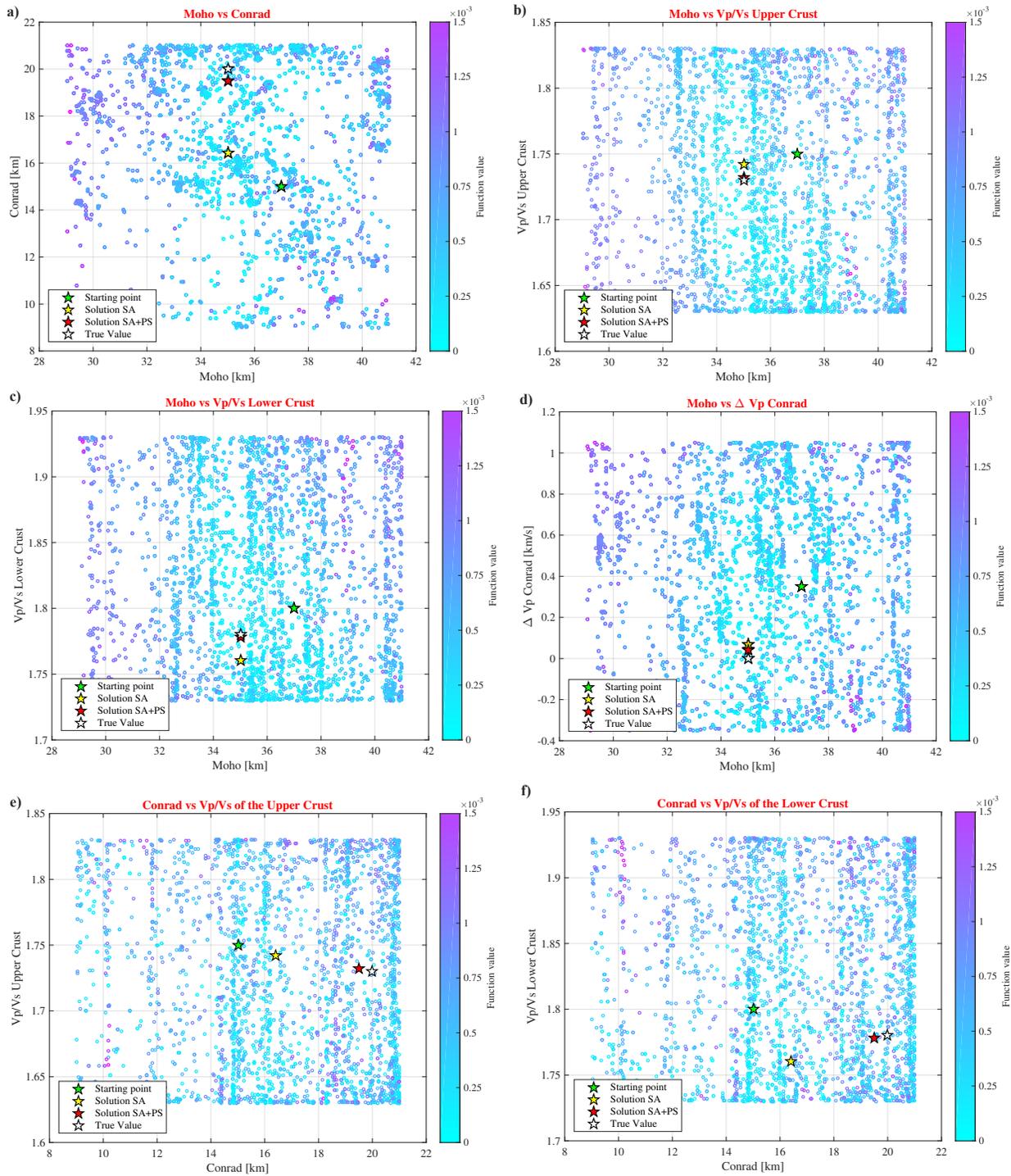


Figure 5.2.13: Waveform match between a synthetic RF at a node (black) and the best RF-fit (red).

As observed in the Figure 5.2.13, the waveform misfit between RF corresponding to constant iso-velocity model per layer and the output model from the inversion is minimal and a sign of the efficiency of the method used.

To visualize the exploration of the parameter space during the SA+PS inversion approach, Figure shows 5.2.14 the tested pairs of parameters and the value of the misfit (L2-norm). Each step of the SA phase is represented, together with the final point after PS (details of the PS steps are unfortunately not available as output from the MatLab function). The V_p/V_s values change about 0.02 in the PS phase. Overall, the randomness of the SA phase is well visible on the figures, as the parameters space is fairly well sampled by points.



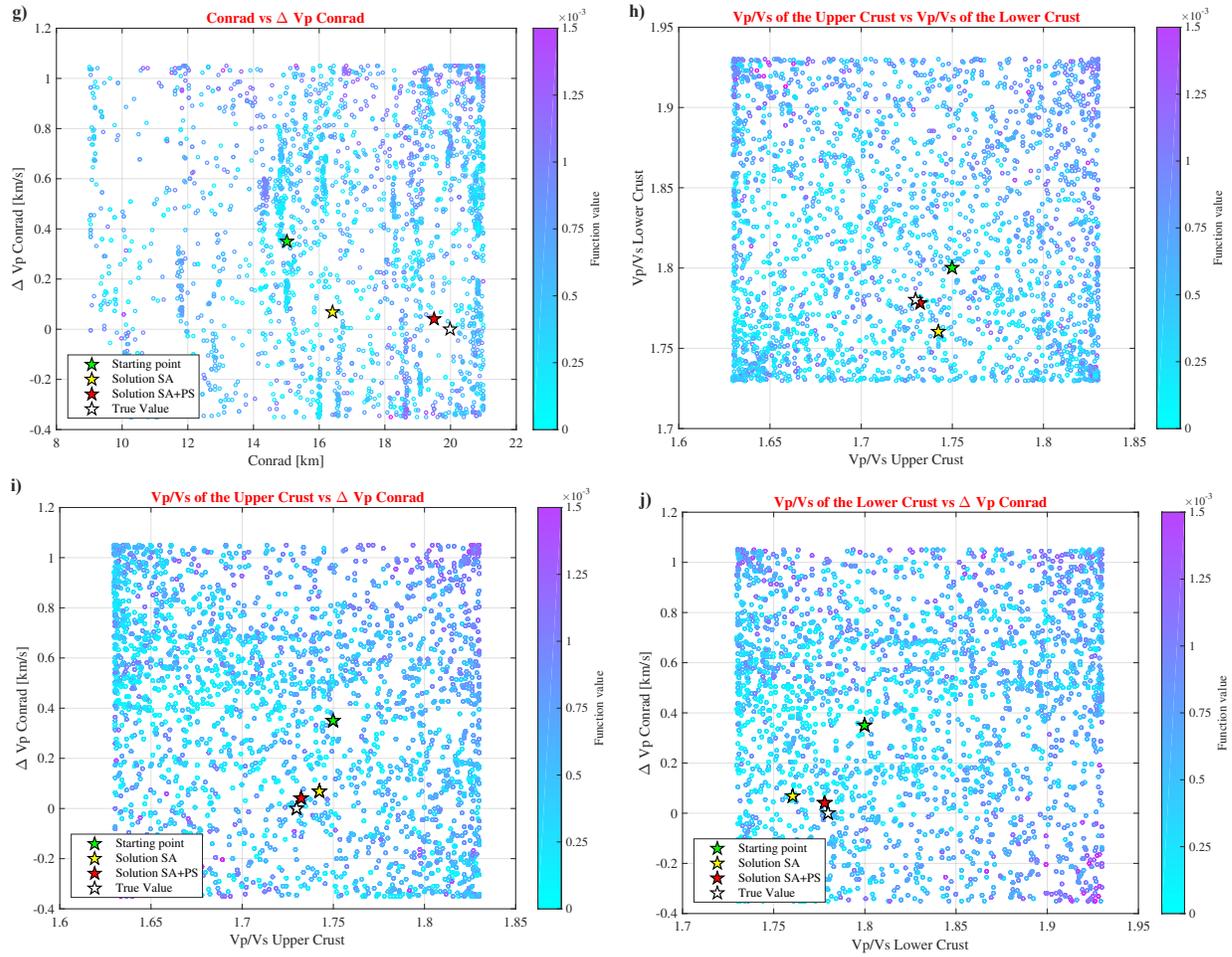


Figure 5.2.14: Exploration of parameter space (from a to d) for 5 parameters; see axes labels for shown parameters. Scatter points represent the value of the objective function to minimize before applying the pattern search solver.

I do not see clear correlations between the inverted parameters in the sampling. The objective function value is smaller near the solution, but the PS phase is certainly needed to converge to the good result. In conclusion, all the synthetic tests and sensitivity studies carried out in this chapter allowed me to define the minimum set-up of the new inversion approach that I apply to real data in the following chapter.

Chapter 6

Inversion results and representation

At the beginning of this chapter, I describe the inversions I carried out for converted wave tomography in the Central Alps. Part of the discussion is to consider the number of the possible variables to be inverted, which greatly influences the computation time. Based on these considerations, I choose a set of optimal and feasible parameters. After a careful analysis between the quality of the solution and the computation time, for the final inversion I proceed with 4 independent parameters, which are Moho depth, intra-crustal discontinuity (Conrad) depth, V_p/V_s ratio for the full crust and P-wave velocity jump at the Conrad. By performing a few rounds of the spatial traveling salesman path (described in Chapter 5) to invert structure at nodes, I observe that the overall misfit improves.

In the second part of the chapter, I show the various presentation styles, and describe the interpolation method used to fill in gaps at nodes not resolved by the inversion. I also establish a quality control based on the absolute misfit per node in order to present a more reliable model.

6.1 Overview of inversions and final model selection

The final model has been chosen after considering the number of variables to be inverted, and a set of inversions run to see the sensitivity of the results based in inversion parameters. These are described in the following sub-sections.

6.1.1 Choice of the number of variables to invert

In section 4.2 I introduced a new model parameterization suitable for converted wave tomography. Considering a model with 2-layers in the crust bounded by 3 interfaces (surface, Conrad and Moho), I can distinguish different scenarios according to the number of variables to invert (Figure 6.1.1), which go from more flexible and more time consuming to simpler and computationally more efficient implementations. Table 6.1 gives an overview of these possibilities, with a color code of fixed and unknown parameters.

N. ind. Vars.	10	9	6	5	4	3
Conrad depth	Z Conrad	Z Conrad	Z Conrad	Z Conrad	Z Conrad	Z Conrad
Moho depth	Z Moho	Z Moho	Z Moho	Z Moho	Z Moho	Z Moho
P velocity	Vp below surface	Vp below surface	Vp below surface	Vp below surface	Vp below surface	Vp below surface
	Vp above Conrad	Vp above Conrad	Vp above Conrad	Vp jump	Vp jump	Vp jump
	Vp below Conrad	Vp below Conrad	Vp below Conrad	across the Conrad	across the Conrad	across the Conrad
	Vp above Moho	Vp above Moho	Vp above Moho	Vp above Moho	Vp above Moho	Vp above Moho
S velocity	Vs below surface	Vs below surface	Vs below surface	Vp/Vs upp. crust	Vp/Vs crust	Vp/Vs crust
	Vs above Conrad	Vs above Conrad	Vs above Conrad			
	Vs below Conrad	Vs below Conrad	Vs below Conrad	Vp/Vs low. crust		
	Vs above Moho	Vs above Moho	Vs above Moho			

Table 6.1: Overview of inversion strategies with different number of parameters to invert. Green cells represent independent variables to invert, red cells the *a priori* fixed variables. See text for details.

The full scenario is represented by the inversion with 10 independent variables at each node: 2 variables for thicknesses (Moho depth and Conrad depth where the starting Moho is from Spada et al. (2013) and the initial Conrad is defined as the Moho - 12 km), 4 for Vp and 4 for Vs values (Fig. 6.1.1 a). The number of variables drops to 9 if I consider the depth of the Moho fixed using an *a priori* model (Fig. 6.1.1 b), for example the model proposed by Spada et al. (2013). For further reduction of the number of parameters, one can consider the Vp model to be known (for example from local earthquake tomography), which results in 6 independent variables (Fig. 6.1.1 c): 2 for thicknesses and 4 for Vs.

6.1.2 Inversion parameter and runs

I carried out several inversions using different parameters of the inversion itself (e.g. number of iterations) in order to choose a set of optimal parameters (Table 6.2).

Name	N	N. iter	Lower Bounds	Upper Bounds	Baz. Bin	N. rxb	N. stat.	Time (h)
Inv 1	5	500	-12, -6, 1.63, 1.70, -0.35	12, 6, 1.63, 1.70, 1.05	45°	10	150	8
Inv 2	5	1000	-12, -6, 1.63, 1.70, -0.35	12, 6, 1.83, 1.90, 1.05	45°	7	150	16
Inv 3	5	2000	12, -6, 1.63, 1.70, -0.35	12, 6, 1.83, 1.90, 1.05	45°	7	150	30
Inv 4	5	4000	-12*, -10*, 1.6, 1.7, -0.35	12, 10, 1.82, 1.90, 1.05	60°	7	150	35
Inv 5	4	500	-12, -6, 1.60, -0.35	12, 6, 1.90, 1.05	45°	7	150	9
Inv 6	4	1000	-12, -10, 1.60, -0.35	12, 10, 1.90, 1.05	60°	7	150	14
Inv 7	4	1000	-12*, -10*, 1.6, -0.35	12*, 10*, 1.9, 1.05	60°	7	150	12.5
Inv 8	4	1000	-12*, -10*, 1.6, -0.35	12*, 10*, 1.9, 1.05	60°	7	75	8
Inv 9	4	2000	-12*, -10*, 1.6, -0.35	12*, 10*, 1.9, 1.05	60°	7	150	23
Inv 10	4	4000	-12*, -10*, 1.6, -0.35	12*, 10*, 1.9, 1.05	60°	7	150	31
Inv 11	4	4000	-12*, -10*, 1.6, -0.35	12*, 10*, 1.9, 2.00	60°	7	150	33
Inv 12	4	4000	-12*, -10*, 1.6, 1.7	12*, 10*, 1.8, 1.9	60°	7	150	37.5
Inv 13	3	4000	-12*, -10*, 1.6	12*, 10*, 1.9	60°	7	150	30

Table 6.2: Inversions performed for the Central Alps domain. N is the number of variables to invert, N_{iter} is the number of iterations, Lower and Upper Bounds are respectively the lower and upper limits for the variables to invert. Moho depth and Conrad depth were relative constraints, and the * symbol means there was an absolute constraint on those values. Baz. bin indicates the angle of back-azimuthal grouping into bundles, N_{rxb} is the threshold of the minimum number of rays per bundle, N_{stat} the number of stations involved, Time is the effective computation time.

All the inversions were run considering the simulated annealing algorithm and a pattern search for every node visited by the traveling salesman path.

From Inv 1 to Inv 4 I consider the inversion with 5 parameters, from Inv 5 to Inv 12 I examine the inversion with 4 independent parameters, while Inv 13 has 3 parameters. I varied the number of iterations from a minimum number of 500 (established by the synthetic tests in Chapter 5) up to 4000, taking into account that the increase in the number of iterations is directly proportional to the increase of the computation time.

Regarding the range of variation for each of variables value, i.e. the lower and upper bounds, I have adopted 2 strategies:

- *Relative* ranges of variation (Inv 1, 2, 3, 5, 6) where the lower and upper bounds for Moho and Conrad depth refers to the current round of the inversion;

- *Absolute* ranges of variation (Inv 4, 7, 8, 9, 10, 11, 12, 13 reported with * in Table 6.2), where an additional *a priori* absolute range was imposed on Moho and Conrad depth. For the case applied to the Central Alpine domain, the Moho depth was allowed to vary from 15 to 65 km, while the Conrad depth from 10 to 55 km.

Other parameters I have tuned to apply a spatial smoothing are the back-azimuthal sector division (45° or 60°), the threshold on the number of rays per bundle (10 rays for Inv 1, 7 rays for the other inversions) and the number of stations (usually all, and for Inv 8 I visually selected the stations with very good RF waveform quality). The inversions have been launched using a MacBook Pro with 8 GB of RAM and with a 3.1 GHz Intel Core i7 processor. Figure 6.1.2 shows the computation time cost for the inversions listed in Table 6.2.

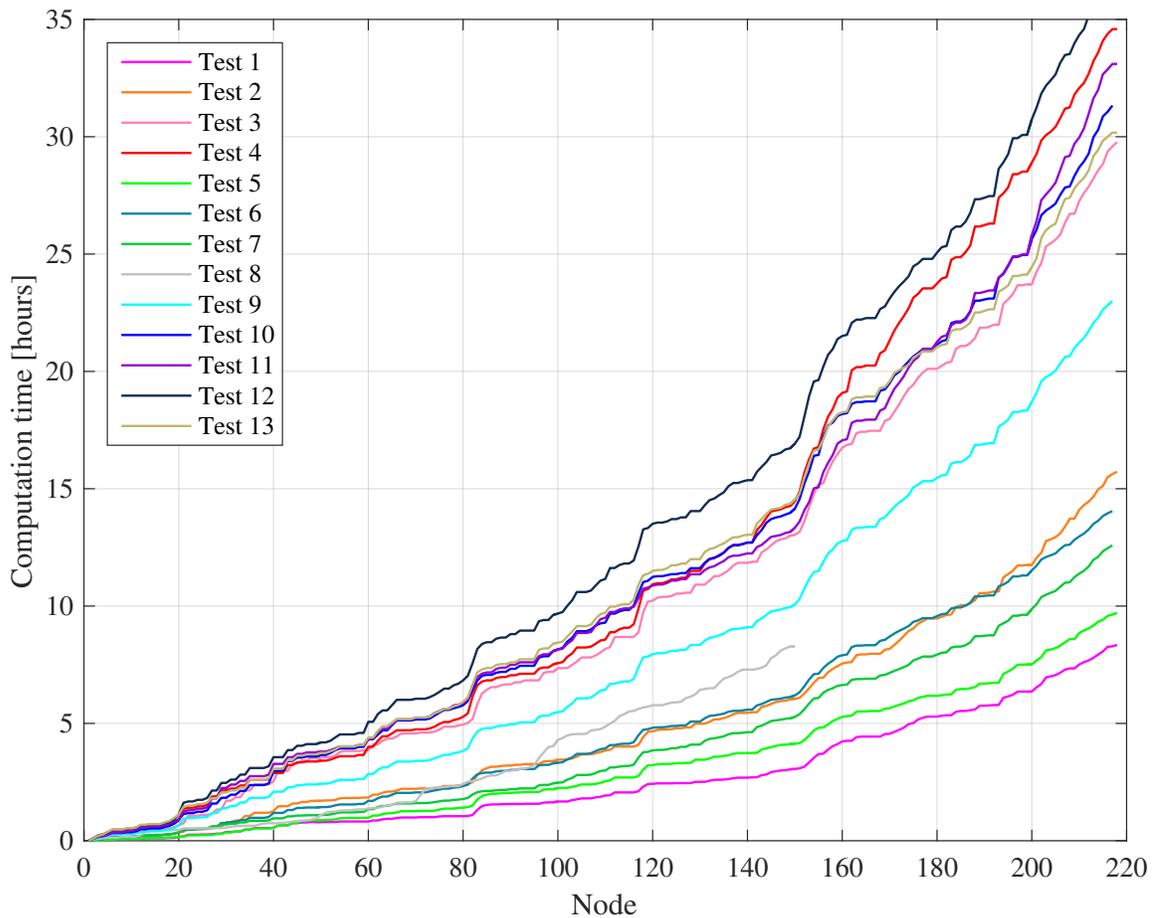


Figure 6.1.2: Computation time cost for the inversions listed in Table 6.2. Inversion 8 stops at node 150 because it considers a fewer number of stations than the other inversions performed.

6.1.3 Inversion rounds and misfit trends

As already mentioned, the inversion procedure is performed considering the technique of simulated annealing, followed by a pattern search algorithm (Chapter 5). The geometry and velocity structure is updated step by step at each node, but always (except at the end) including neighbouring, not-yet-inverted-for nodes in the velocity interpolation. At the end of a complete round of the inversion, the output velocity model from the inversion constitutes the starting model for the next round of the computation, during which I expect the solution to improve, because each node is visited for the second time, so in theory only minor adjustments in the solution are expected. I decided to vary the traveling salesman path scheme in each round, changing the order of the visited nodes. For each round of the inversion, the ray-paths of the bundles are recomputed using the current structure and velocity model, and the new (slightly changed) ray coverage is considered for constructing bundles. In the graph of Figure 6.1.3, I show the relative misfit trend per node for Inversion 10 during 3 consecutive rounds of the traveling salesman problem, obtained respectively with 4000 (31 h), 1500 (19 h) and 1500 (19 h) iterations using the simulating annealing and the pattern search algorithm for the inversion.

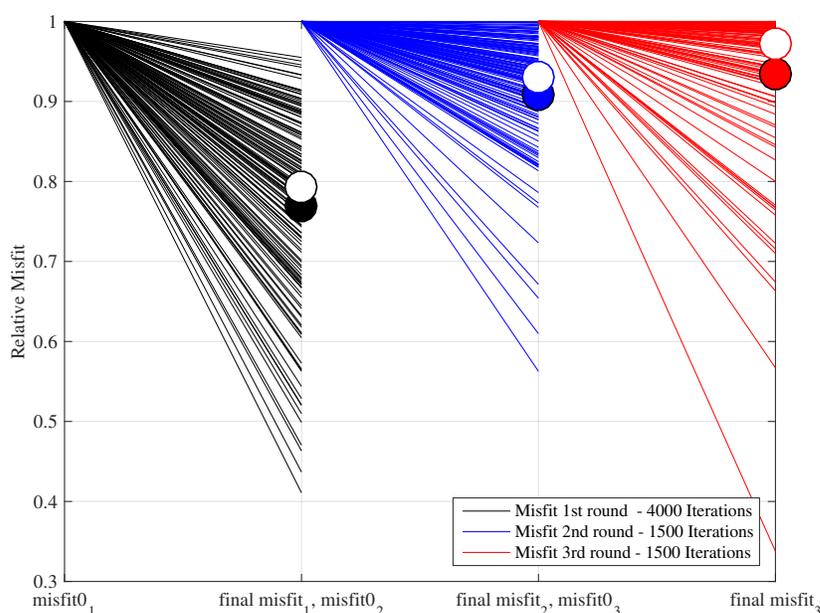


Figure 6.1.3: Relative misfit trends for inversion 10 (see Table 6.2) during inversion rounds. Y-axis shows the relative improvement (final misfit/initial misfit) at each node. Values close to 0 show a big improvement, values close to 1 a slight improvement in the relative misfit per node. Filled black, blue and red circles represent respectively the mean after 1st (0.7702), 2nd (0.9076) and 3rd (0.9349) round of the inversion. Empty black, blue and red circles represent the respective medians (0.7925, 0.9298 and 0.9715).

Figure 6.1.3 shows the misfit reduction $misfit_R$ at each node, defined as:

$$misfit_R = \frac{misfit_0 - misfit_n}{misfit_0} \quad (6.1.1)$$

where $misfit_0$ represents the initial misfit and $misfit_n$ the final misfit at the round n of the inversion. The reason for showing relative misfit reductions by rounds instead of absolute misfit values is that the bundles are recomputed in each round, and therefore a given node considers different set of receiver functions in each round. Values close to 1 indicate a slight variation on the relative misfit, while values far below 1 show big improvements of the relative misfit. Apart from a very small number of nodes, the general trend shows that a few rounds of traveling salesman paths clearly improve the overall misfit. In fact, values of mean and median after the different rounds of inversion suggest a convergence towards a stable solution.

6.1.4 Waveform fits

The observed receiver functions and bundles, and the respectively computed synthetic traces are minimized using the least-squares norm L2-norm. Figure 6.1.4 shows a few examples of the waveform match between observed (black) and synthetic (red) RF for a single bundle at 4 different stations. At first glance and as we expect, the observed RF data are rather complicated, partly from the complex structure, and also as they contain noise. To increase the signal-to-noise ratio (SNR), I did the grouping from single traces to bundles considering a certain range of back-azimuth sector, and also applied a minimum number of rays per bundle as a threshold (Chapter 5).

The misfit calculated at each node is the sum of misfits for each bundle, weighted by the number of rays. The misfit computed at each node $misfit_{node}$ is the mean of misfits for each included bundle, weighted by the number of rays:

$$misfit_{node} = \frac{\sum normL2(ORF, SRF) * Nr_{bundle}}{Nr_{node}} \quad (6.1.2)$$

where $normL2(ORF, SRF)$ represents L2-norm difference between observed and synthetic RFs for a bundle, Nr_{bundle} is the number of rays per bundle, and Nr_{node} is the number of rays per node that exceed the established threshold of the number of rays per bundle.

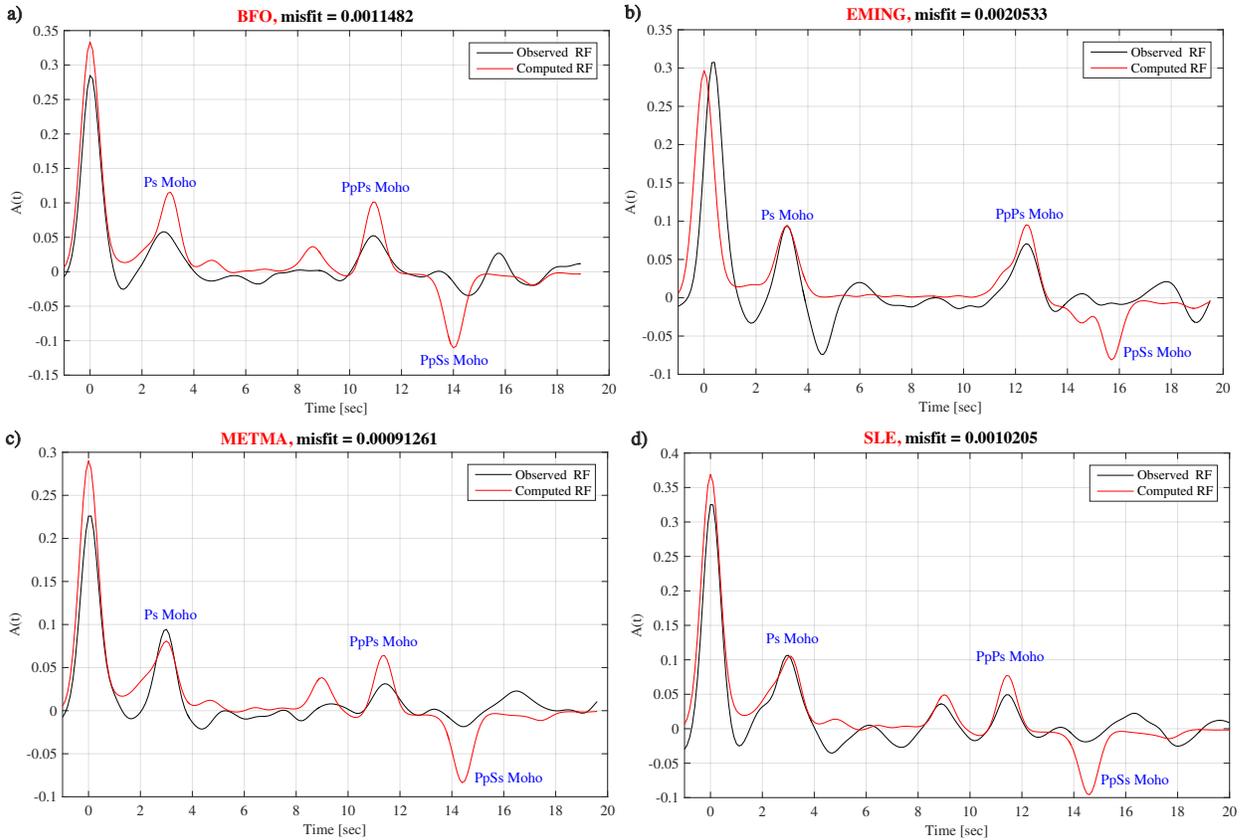


Figure 6.1.4: Waveform matches for a single bundle between the receiver functions observed (black) and synthetic (red) receiver functions by the inversion using simulated annealing and pattern search method (Inversion 10 in Table 6.2). **a)** BFO, fit with 16 rays in the bundle; **b)** EMING, fit with 19 rays in the bundle; **c)** METMA, fit with 28 rays in the bundle; **d)** SLE, fit with 94 rays in the bundle. Blue labels indicate the Ps conversion at the Moho and the subsequent multiples.

Figure 6.1.4 shows few examples of good fit between observed and synthetic RF waveforms. The match of the converted Ps and PpPs phases for the Moho (and sometimes additional peaks) is remarkable¹, while it is usually more difficult to identify the PpSs multiple and the conversions from the intra-crustal discontinuity. Observed amplitude variations may be due to the use of default values for the seismic attenuation Q_p and Q_s .

The waveforms of the synthetic data in Figure 6.1.4 (red curve) show that the Ps Conrad is not present but is combined with the Ps Moho peak, while the multiple PpPs and PpSs phases are separated in the lower and in the upper crust. In order to discriminate the first

¹Since the forward model of the waveform inversion uses the method of Shibutani et al. (1996), the dipping of the main interfaces is not accurately modeled and this might also explain why some Moho multiples are not fitted properly (e.g. PpSs phase in Figure 6.1.4).

Ps conversion between Conrad and Moho, it would therefore be appropriate to refer to RF computed at higher frequencies.

6.1.5 Choice of the final model

The ten inversions run described in the section 6.1 also represent my search of a sufficient set of parameters to get stable results, while results themselves should look similar when changing inversion parameters. In this search, it turned out that the number of iterations plays a key role, and it has to be sufficiently high to sample a large enough portion of the parameter space. The second most important choice is the level of value bounds, which expands the size of the search. Here, the absolute bounds were essential, to avoid divergence of parameters, especially Moho and Conrad depth. The minimum criteria on rays, stations and the size of the back-azimuth bins for bundling turned out to have relatively smaller role in determining the results, they primarily affected the computation time and spatial smoothness of the results.

Considering the tested inversion parameter sets (Table 6.2) and result, I ultimately choose and stopped at the inversion 10 (Inv 10), which uses 4 independent variables to invert: Moho depth, Conrad depth, V_p/V_s for the full crust and the P-wave velocity jump at the Conrad. This inversion run was performed with the technique of simulated annealing and pattern search using 4000 iterations for the first round of traveling salesman path, and 1500 iterations for the subsequent rounds. Regarding the range of variation of the parameters' values, in the first round the Moho depth can vary from -12 to +12 km with respect to the initial model of Spada et al. (2013), while the Conrad depth can change from -10 to +10 km compared to the starting value (chosen at 12 less than the initial Moho depth). From the second round onwards, I added an absolute constraint, as discussed above (Moho 15-65 km, Conrad 10-55 km). The average V_p/V_s for the full crust can vary from 1.6 to 1.9, and the P-wave velocity jump at the Conrad from -0.35 to 1.05 km/s.

For the spatial smoothing, in this inversion run I grouped rays into bundles every 60° , and a threshold of 7 rays per bundle was considered together with all stations available in the study area. I decide to consider this inversion as the final one because due to the estimated and "felt" complexity of the crust in the Alpine domain, it was a reasonable compromise, and it seemed to be extremely difficult to meaningfully separate the average crustal V_p/V_s ratio into a upper and a lower crustal value. The results of this inversion are presented in Chapter 7, in the ways described in section 6.2. The pattern of inversion results from the other runs performed here are similar to the selected model, but sometimes less robust or presenting more spatial gaps.

Since we observed that the P-velocity jump at the Conrad is very difficult to estimate,

we have decided to test a further simplified inversion, in which both the ΔVp_{CONRAD} and the ΔVs_{CONRAD} are set to 0, satisfying in this way the conditions with 3 variables to invert (Inv 13 in Table 6.1.2). The result of this test showed already the main features with the crustal thickness that generally reflects well the roots of the Alpine orogen and the average crustal Vp/Vs ratios relatively higher beneath the orogen.

Figure 6.1.5 shows the difference between Conrad, Moho and Vp/Vs for the full crust between the inversion with 4 parameters (Inv 10) and inversion with 3 parameters (Inv 13).

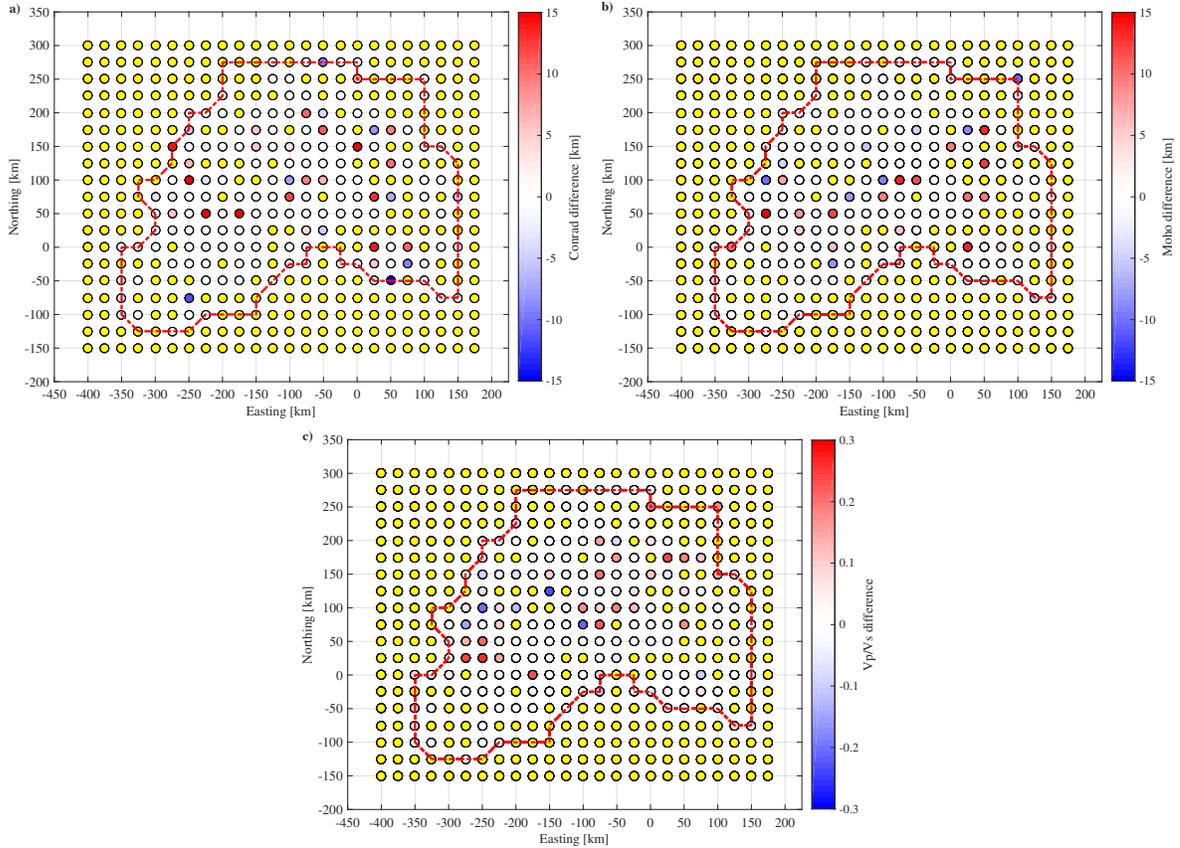


Figure 6.1.5: **a)** Conrad depth difference between inversion with 4 variables and inversion with 3 variables. **b)** Moho depth difference between inversion with 4 variables and inversion with 3 variables. **c)** Vp/Vs difference between inversion with 4 variables and inversion with 3 variables. Images from the 4-variable inversion are from Inv 10, which is considered the final result.

Most of the nodes have values close to zero except for some nodes within the Alpine domain with maximum variations at around 15 km for the values of Conrad and Moho and 4 nodes with a difference of 0.3 regarding the Vp/Vs ratio. For all 3 maps, in most cases, the difference is positive, meaning that the inversion with 4 variables mostly finds values with Conrad and Moho deeper and with higher Vp/Vs with respect to the inversion using 3 variables.

6.2 Inversion results representation

In this section I describe the model parameters obtained directly from the output of the inversion with 4 independent variables, and the representation of interfaces and velocities. Furthermore, I briefly report the interpolation scheme of the results at the nodes not resolved by the inversion, as well as at nodes of high misfit to obtain high quality results.

6.2.1 Inversion output

As a result of the inversion procedure, I produce several point-wise maps with information on the interfaces (Conrad and Moho depth) and velocities (average V_p/V_s for the entire crust and P-wave velocity jump at the Conrad).

6.2.1.1 Interface depth representation

Figure 6.2.1 shows the depth map of the Moho discontinuity obtained during the 1st and 2nd round of the inversion, for the finally selected inversion (Inv 10), and similar maps are produced for the Conrad discontinuity depth.

The investigated area (contoured by the red dashed line) contains in total 218 nodes. Considering the ray coverage of the study area, I obtain results at 154 nodes. From Figure 6.2.1 I can observe that the general pattern in terms of Moho depth geometry is similar between the 1st and the 2nd round of the inversion, and only minor changes occur in the position of the nodes solved by the inversion as a result of slightly varying ray coverage between the two rounds, which are highlighted in the Figure 6.2.1 concerning the difference between the two Moho depth maps. Another observations is that the nodes that have high values after the first round are often higher after the second round, reaching the upper allowed bounds.

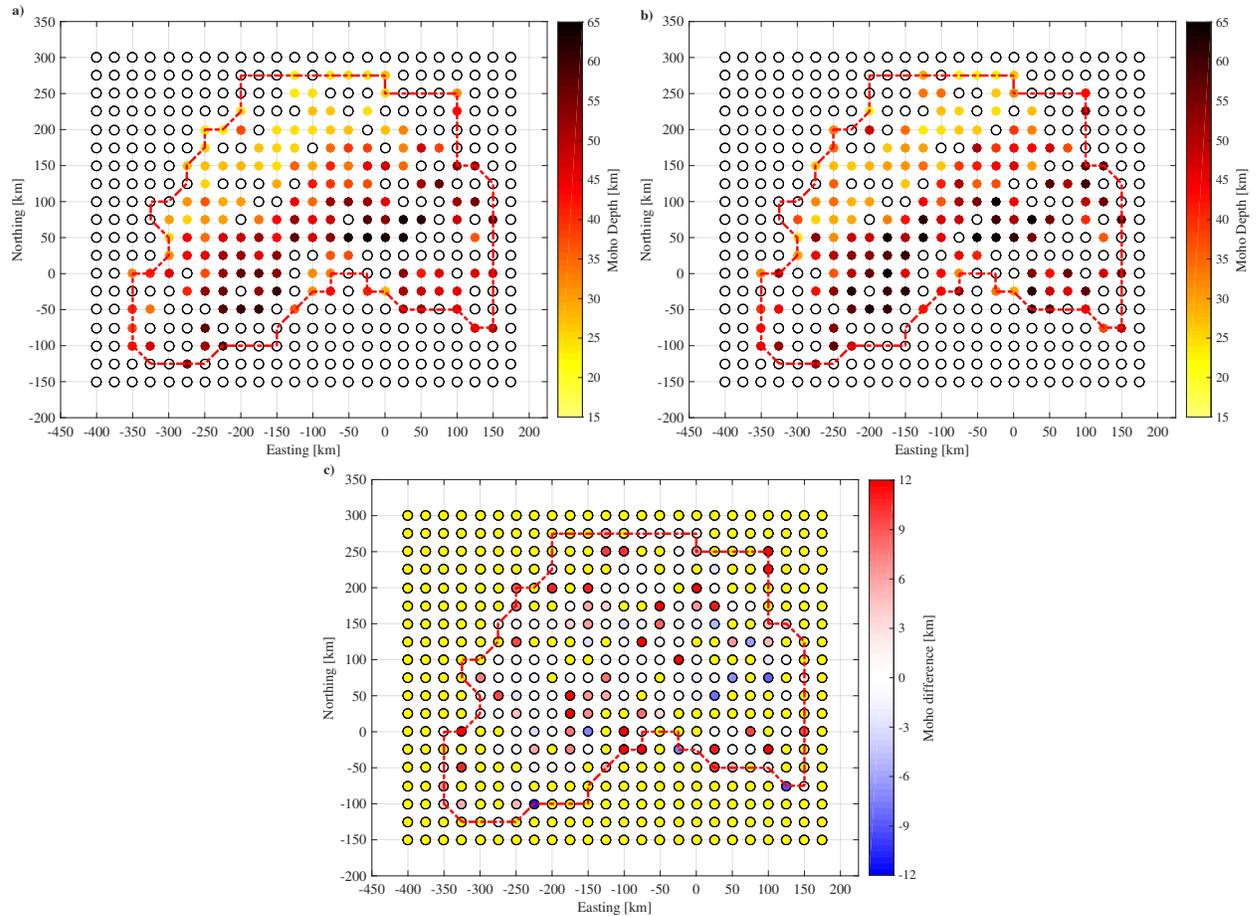


Figure 6.2.1: **a)** Moho depth map after the 1st round of inversion. **b)** Moho depth map after the 2nd round of inversion. **c)** Difference between 2nd and 1st round results. Red dashed line indicates the border of the study area, white dots indicate nodes not resolved by the inversion. The color scale in Figure **a)** and **b)** spans from the imposed lower to upper bound for Moho depth. Maps are from the inversion Inv 10, which is considered the final result.

6.2.1.2 Velocities' representation

Regarding the velocities' representation, the inversion yields maps for the average crustal V_p/V_s and for the P-wave velocity jump at the Conrad. As for the Moho depth map, also for the V_p/V_s map (Figure 6.2.2) there are some but not too many differences (Figure 6.2.2 **c)** between the second and the first round of the inversion, even if, compared to the result of the interfaces, a particular pattern on the map is less evident.

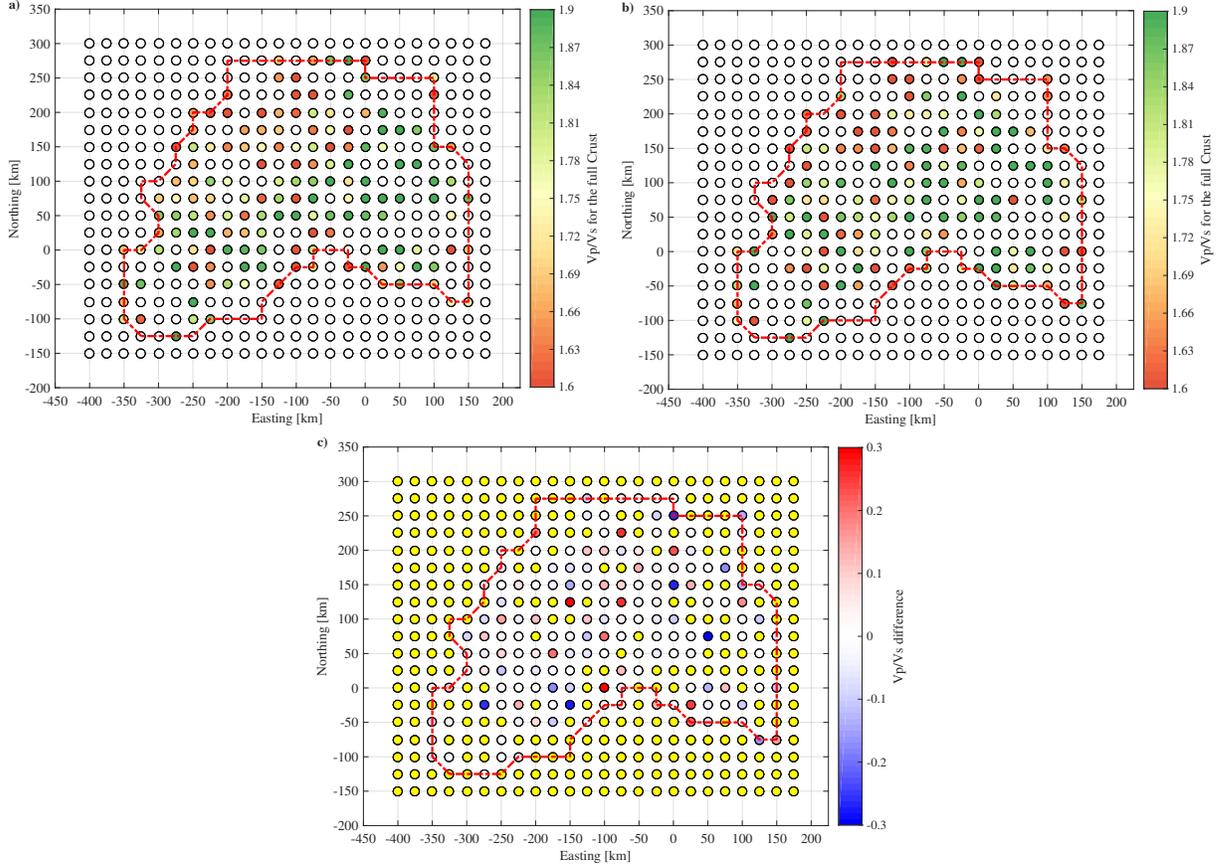


Figure 6.2.2: **a)** Average crustal V_p/V_s map after the 1st round of inversion. **b)** Average crustal V_p/V_s map after the 2nd round of inversion. **c)** Difference between 2nd and 1st round results. Red dashed line indicates the border of the study area, white dots indicate nodes not resolved by the inversion. The color scale in Figure **a)** and **b)** spans from the imposed lower to upper bound for V_p/V_s ratio. Maps are from the inversion Inv 10, which is considered the final result.

I also tested an inversion which considers different values of V_p/V_s for the upper and for the lower crust, but keeps the P-velocity jump across the Conrad at zero (Inv 12 in the Table 6.2). Results showed that often the V_p/V_s of the upper crust is larger than the V_p/V_s of the lower crust. This means that the RF data suggests that there is a V_s jump across the Conrad even if there is no V_p jump there. This is also reflected in a separate PpPs phase for the Conrad and the Moho on the RF waveform fits (see Figure 6.1.4), while the Ps phase of the Conrad is close to that of the Moho and may simply cause its broadening, not a separate peak.

Similar to the previous cases, I find little differences between the first and the second round of the inversion in terms of P-wave velocity jump at the Conrad with, in general,

values close to 1 km/s, which is the upper limit of the model (Figure 6.2.3). These values are suspicious not only because they are close to the maximum value of the search range, but also because it seems high compared to result obtained by active seismic refraction studies.

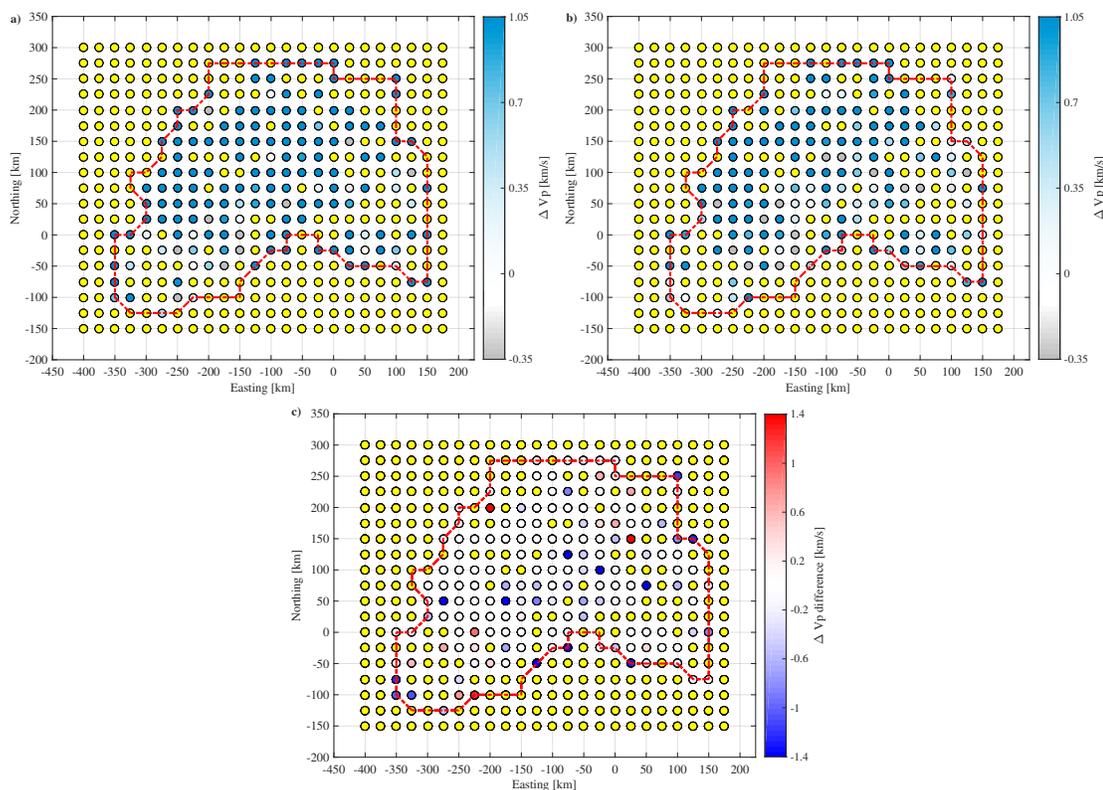


Figure 6.2.3: a) Map of P-wave velocity jump at the Conrad after the 1st round of the inversion. b) Map of P-wave velocity jump after the 2nd round. Red dashed line indicates the perimeter of the study area, yellow dots indicate nodes not resolved by the inversion. The color scale spans from the imposed lower to upper bound for ΔVp_{CONRAD} . Maps are from the inversion Inv 10, which is considered the final result.

In fact, seismic refraction data show that upper and middle crust has P-velocities between 6.0 and 6.2 km/s. Below the Penninic nappes the lower crust thickens remarkably, merging probably with the high-velocity zone of 6.6 km/s at a depth of about 21 km, which has been interpreted as the top of the indenting lower crust of the Adriatic promontory of the African plate (Stäuble et al., 1993; Ye et al., 1995). Moreover, a low average P-velocity from 6.2 km/s to 6.4 km/s in the lower crust below northern Switzerland was reported earlier by Deichmann and Rybach, 1989 based on the investigation of earthquakes with foci reaching down to the Moho. High velocities between 6.6 and 6.8 km/s were found below the Molasse Basin while low P-velocity value of 6.2 km/s for a highly reflective lower crust between the

more commonly found higher velocity of 6.5-6.6 km/s below the northern Molasse Basin. These data suggest that the V_p jump is at most 0.4-0.5 km/s at different geological domains and does not reach higher values.

In general the interpretation of the P-velocity jump at the Conrad is complex, and the result can contain some artifacts due to the parameterization, i.e. that the mid-point of the jump in V_p is fixed. To further investigate this point, I carried out an inversion test (Inv 11 in the Table 6.2) by setting the upper limit of the P-velocity jump at 2.0 km/s (Figure 6.2.4), which was not a realistic value, yet the results as in Figure 6.2.3, most nodes are close to the upper limit of the model and this makes too high amplitude Conrad PpPs multiples.

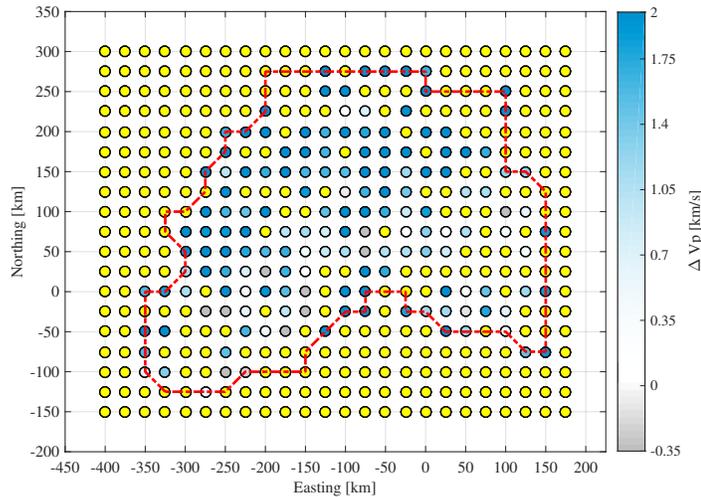


Figure 6.2.4: Map of P-wave velocity jump at the Conrad with upper limit at 2.0 km/s (Inv 11).

Looking at the waveforms, it is still possible to see some peaks for the Conrad PpPs multiples in the data, even if it is not possible to unambiguously discriminate whether this is signal or noise. The Ps Conrad phase is usually close to the Ps Moho and causes its broadening. This means that in case of a thin lower crust, the RF is not an efficient tool in resolving ΔV_p .

Based on these results, shear-wave velocity maps and cross-sections, which represents the ultimate goal of our converted-wave tomography method, can be reconstructed. First, the P-wave velocity above (or below) the Conrad is set, from the *a priori* selected mid-point and by subtracting (or adding) half of $\Delta V_{PCONRAD}$, then the four values of V_p are divided by the average crustal V_p/V_s ratio to obtain the four V_s values. This allows to plot V_s maps along interfaces (above or below), V_s depth slices (interpolated to a constant depth), and V_s profiles (projected to a cross-section, see examples below).

6.2.1.3 Interpolation of unresolved nodes

One inherent property of our model parameterization on a fixed mesh-size grid is that in areas of lacking ray coverage, there are unresolved nodes of structure and velocity. In order to fill these gaps, I use the 8 surrounding nodes, but limited to the same tectonic plate, and I compute an average weighted by the inverse of the square of the distance to these nodes. However, as very often not all of these 8 nodes are resolved by the inversion procedure, I distinguish 3 cases for this interpolation:

- a) if the number of neighbouring nodes resolved by the inversion are 0, then the value at the current node remains the value from the starting model (no relevant new information);
- b) if there is only 1 neighbouring node resolved by the inversion, the value at the current node is the mean between the its own initial value and the value of the resolved neighbour node;
- c) if the number of neighboring nodes resolved by the inversion and on the same tectonic plate is 2 or more (up to a maximum of 8), then I assign the weighted average as described above.

Figure 6.2.5 shows the Moho map interface before (above) and after (below) this interpolation scheme was applied.

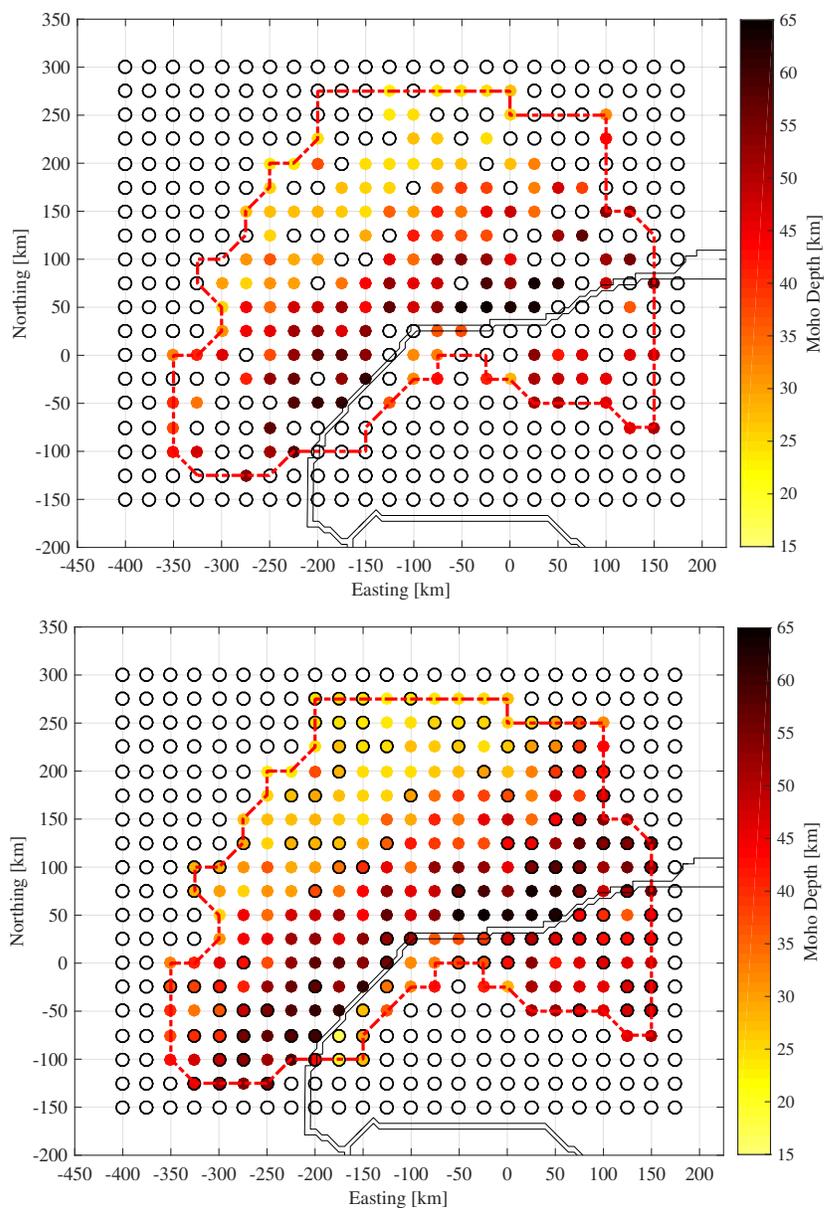


Figure 6.2.5: Moho depth map. Above, first round result of inversion 10, colored dots represent the inversion results at the resolved nodes. Below, same map after applying the spatial interpolation scheme (see text for details). The direct inversion outputs are simple colored dots, while the interpolated values are distinguished by a black edge. Thin double line indicates the plate boundary. Red dashed line indicates the perimeter of the study area. Maps are from the inversion Inv 10, which is considered the final result.

6.2.1.4 Results quality control by node

During the inversion run I observed that despite the 3-D scheme, some nodes remain with a high misfit value, either because of noisy stations, or because of lower ray coverage near the edge of the model. To obtain inversion model results based only on high quality data, and to obtain a smoother image of the model, I decide to introduce a maximum threshold based on the absolute values of the misfit to accept results as they are, and to replace values by interpolation at nodes above this threshold (Figure 6.2.6). The choice of the threshold has been made visually, looking at the distribution curve of the absolute misfit per node.

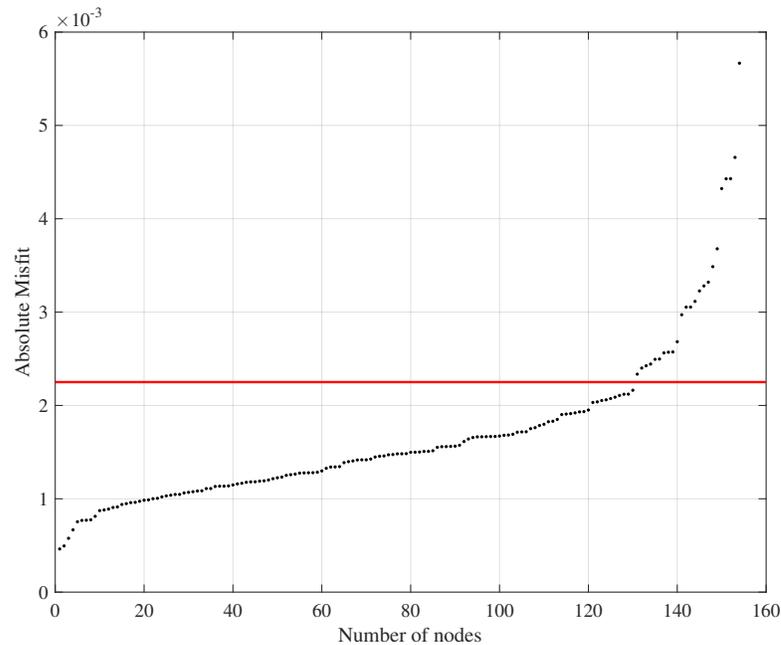


Figure 6.2.6: Value of the absolute misfit per node, sorted in increasing order, for inversion 10, round 1. The red line with misfit of 0.0025 is chosen to cut the curve at its steeper increase, and represents the threshold for which the quality control is carried out.

Figure 6.2.7 shows the misfit map of Inv 10 round 1, with the nodes that exceed the set is threshold colored in red (16% of the available nodes). The majority of the nodes are close to the edge of the study area or to the plate boundary.

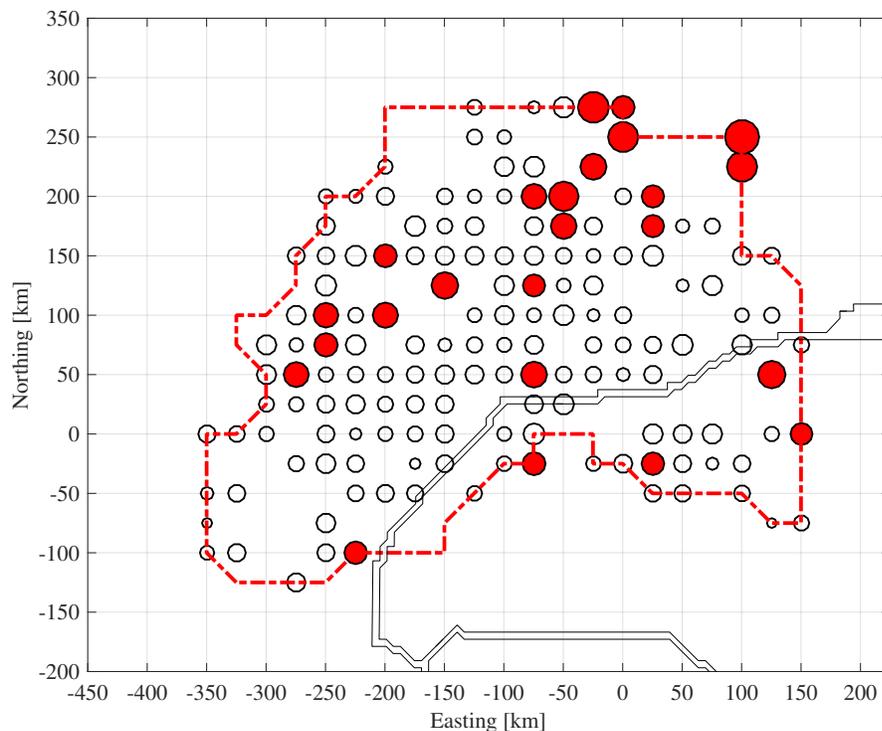


Figure 6.2.7: Point-wise map inversion quality at each node, from inversion 10 round 1, with circle size proportional to absolute misfit. The nodes that exceed the threshold defined by the red line in Figure 6.2.6 are filled in red. Thin double line indicates the plate boundary, red dashed line contours the study area.

Once the high misfit nodes are selected, their values are interpolated following the same scheme as the one performed for unresolved nodes (section 6.2.1.3). Figure 6.2.8 shows the final interpolated map for the study area, where the two sources of interpolated nodes are highlighted.

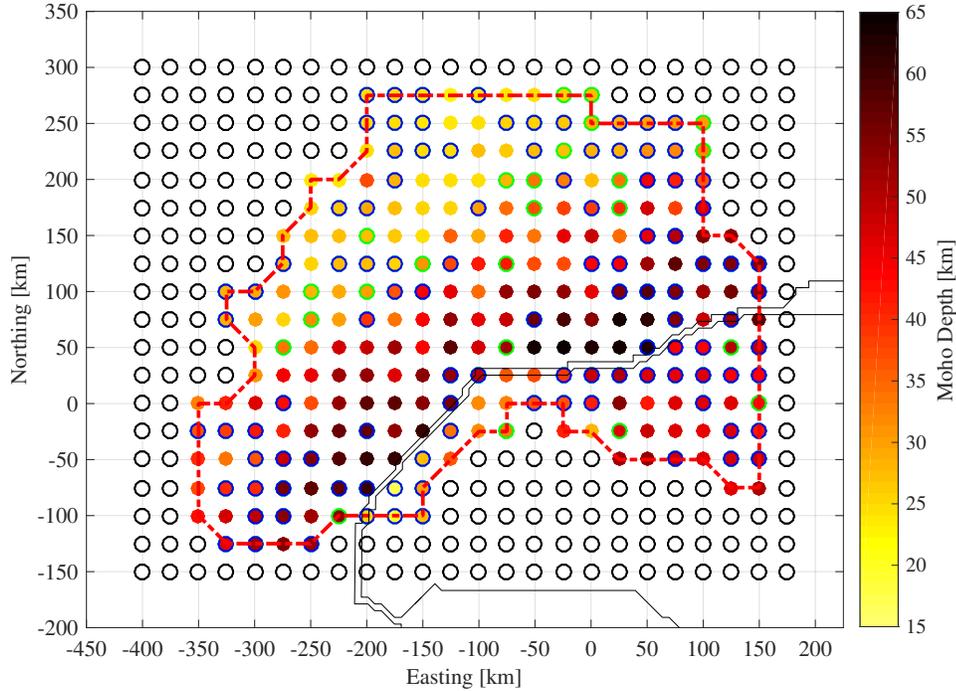


Figure 6.2.8: Moho depth map after interpolation of empty and poor quality nodes. Colored points without an edge are the direct results from the inversion, blue edge represents interpolated value at unresolved nodes, green edge represents interpolated value following misfit based quality control. Thin double line indicates the plate boundary, red dashed line contours the study area.

6.2.1.5 Velocity profile representation

Combining the inversion results in terms of velocities (V_p/V_s , $\Delta V_{pCONRAD}$) and the *a priori* set V_p values, I can reconstruct shear-wave velocities at any point in the 3-D volume. From this, V_s profiles along any cross-section can be plotted by projection, for interpretation and comparison with earlier results (Chapter 7). The main advantage of the model parameterization I defined is that a sharp velocity jump at the Conrad can be imaged. Therefore, interpolations of V_s along a profile is performed within each layer, i.e. within the upper crust using V_s below surface and V_s above Conrad values, and similarly in the lower crust, following the geometry of the Moho and Conrad discontinuities. Moreover, the interpolation of unresolved and low-quality nodes done as described above (sections 6.1.2.3 and 6.1.2.4) is already taken into account and can be directly represented on V_s profiles. An example of such a profile is shown in Figure 6.2.9.

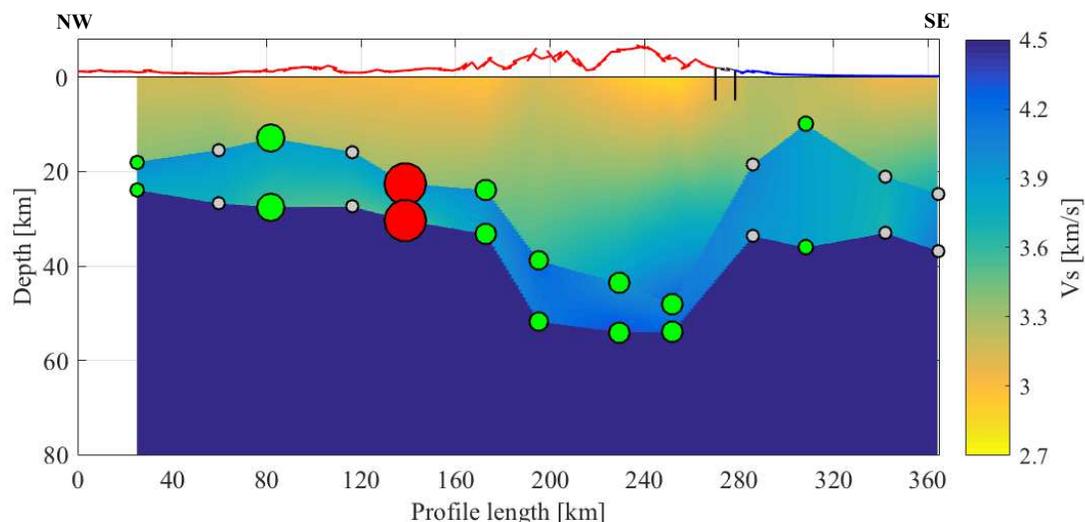


Figure 6.2.9: V_s profile along a cross-section between Vosges (NW, North-West) and West Po basin (SE, South-East). The velocity in the mantle is fixed (blue at $V_s=4.5$ km/s), velocities within the two crustal layers are interpolated. Dots show the nodes projected from the 3-D volumes considering an area of 90 km with respect to the transect. Grey points are those not resolved directly by the inversion, green and red points are those accepted or discarded after performing the quality control and the size of the circle is proportional to the absolute misfit. Red solid line represents the topography in European plate, blue solid line the topography in Adriatic plate; black solid vertical lines the plate boundary between the plates. Vertical exaggeration is 2:1.

In the case where a point is removed due to the large misfit (red dots in Figure 6.2.9), we proceed to the interpolation with the next node, while in the case of a node not directly resolved from the inversion (grey dots in Figure 6.2.9), the initial reference of Spada et al., 2013 is taken.

In Chapter 7 I will use these representation styles, color scales and interpolation steps to describe and interpret the results for the Central Alps.

Chapter 7

3-D interpretation of the Central Alps

The main goal of this chapter is to show the results I obtained with the newly developed 3-D P-to-S converted wave tomography method, and to compare them with other geological investigations and geophysical studies obtained with other investigations. In the first part, I present maps of the crustal structure with the information of intra-crustal boundary, Moho discontinuity and lower crustal thickness. Then, I move to the velocity structure which is represented through absolute V_s images for each interface (below surface, above Conrad, below Conrad, above Moho), the average crustal V_p/V_s ratio, the velocity jump across the Conrad discontinuity and the maps of the S-wave velocity within the upper crust and the lower crust.

In the central part of the chapter, I compare the results I obtained with the P-to-S converted wave tomography technique with corresponding geological transects, and another study obtained by ambient noise tomography method by Lu et al. (2018), through the presentation of a number of V_s cross-sections. The sections are selected in areas of interest for the comparison, giving preference to the more reliably resolved areas, those crossed by a higher amount of rays. The comparison is a starting point to show that the results obtained with the new method are reliable and the discussion can be improved referring to active-source experiments within the study area in more detail.

In the final part, I compare briefly thicknesses and velocities for the 3-D results with the 1-D outcomes using the $H - \kappa$ method. Finally, I make a qualitative comparison between the V_p/V_s map for the whole crust and the lower crustal seismicity distribution in the northern Alpine foreland, and discuss its geodynamic implications.

7.1 3-D P-to-S converted wave tomography results

In this section, I show maps obtained with the P-to-S converted wave tomography approach regarding the crustal structure (Moho and Conrad discontinuity depths, lower crust thickness) and the velocity structure through some images of the absolute shear-wave velocities for each interface (below surface, above Conrad, below Conrad, above Moho), the average crustal V_p/V_s ratio, the velocity jump across the Conrad discontinuity and the S-wave velocity gradients within the upper crust and the lower crust. All results shown are those obtained in Inversion 10, round 2.

7.1.1 Crustal interface structure

The crustal structure (with the Moho and Conrad interfaces) is probably the most reliable data that comes out of our model since it is a parameter that can be better constrained with respect to the V_p/V_s values.

7.1.1.1 Conrad discontinuity depth

Figure 7.1.1 shows the result of the inversion with the map of the intra-crustal discontinuity depth.

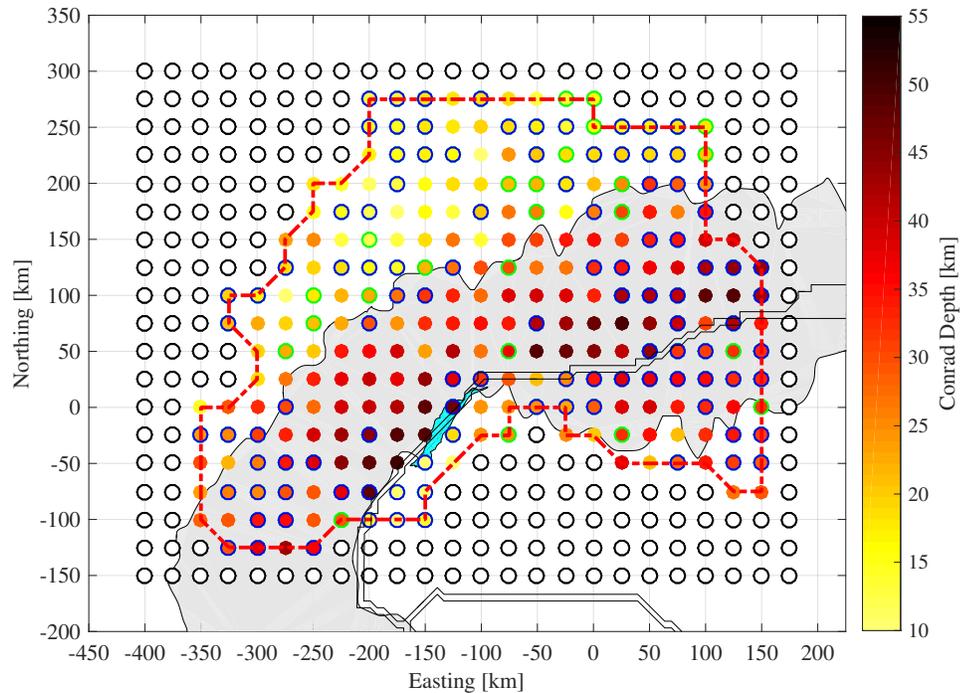


Figure 7.1.1: Final Conrad depth map. Colored points without an edge are the direct results from the inversion, blue edge represents interpolated value at unresolved nodes, green edge represents interpolated value following misfit based on quality control. Filled grey area shows the Alpine arc's smoothed 800 m altitude line while filled cyan area represents the geological Ivrea Verbano Zone (IVZ). Thin double line indicates the plate boundary, red dashed line contours the study area. The color scale spans from the imposed lower to upper bound for the Conrad depth variable during the inversion.

We observe that in the Alpine domain the depth of the upper crust is thicker than in the surrounding area, with values around 30-35 km. According to the result of my model, even thicker values for Conrad depths (> 45 km) are recorded in the Alpine arc within the European plate, following the Europe-Adria plate boundary. Outside the Alpine arc, the values of intra-crustal discontinuity are in the range between 15 and 25 km, with the only exception in the South-East at the edge of the investigation area (contoured by red dash line) at the Adriatic plate.

7.1.1.2 Moho discontinuity depth

The final Moho map (Figure 7.1.2) confirms the trend we have already seen on the previous map for the intra-crustal discontinuity.

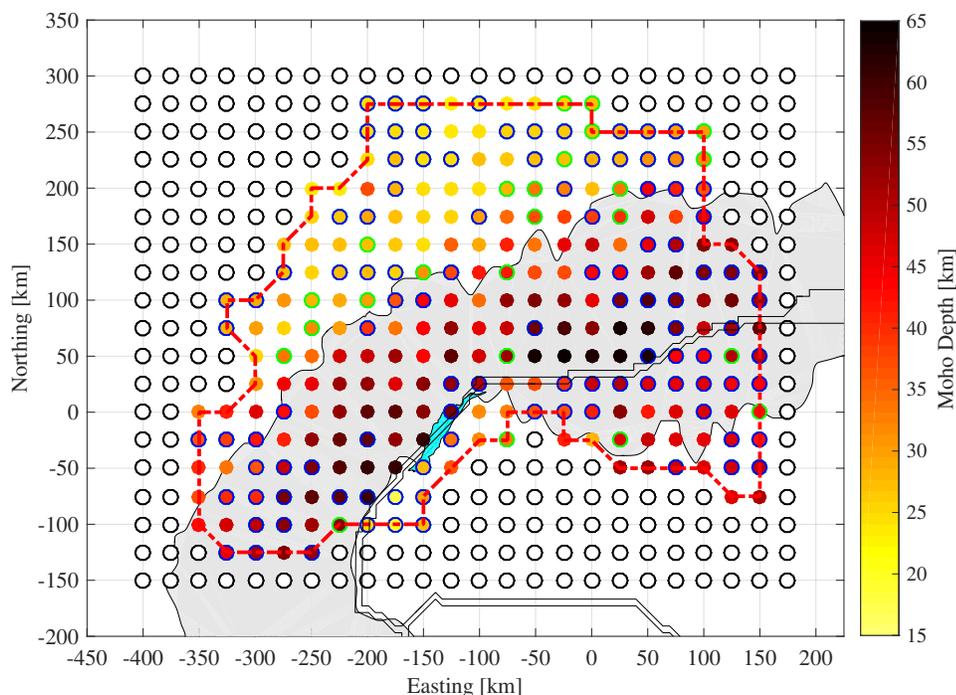


Figure 7.1.2: Final Moho depth map. Colored points without an edge are the direct results from the inversion, blue edge represents interpolated value at unresolved nodes, green edge represents interpolated value following misfit based on quality control. Filled grey area shows the Alpine arc's smoothed 800 m altitude line while filled cyan area represents the Ivrea Verbano Zone (IVZ). Thin double line indicates the plate boundary, red dashed line contours the study area. The color scale spans from the imposed lower to upper bound for the Moho variable during the inversion.

The main remarkable result is that the Moho depth reflects very well the roots of the Alpine orogen, with values deeper than 40 km within the Alpine arc. We find deeper values (at around 60 km) on the European plate following the plate boundary. As we observed for the Conrad map, even for the Moho outside the Alpine arc, the values are in the range between 20 and 25 km, apart from southeastern margin of the study area, where the values remain relatively high (40-45 km).

Figure 7.1.3 shows the difference between the final Moho result and initial Moho model, which is the one of Spada et al. (2013). A large part of the nodes present values close to zero (whitish colors in Figure 7.1.3), which means a very small difference between final Moho and the initial value. Moreover, many nodes have positive variations with respect to the model proposed by Spada et al. (2013), especially in the center and in the West of the study area inside the European plate.

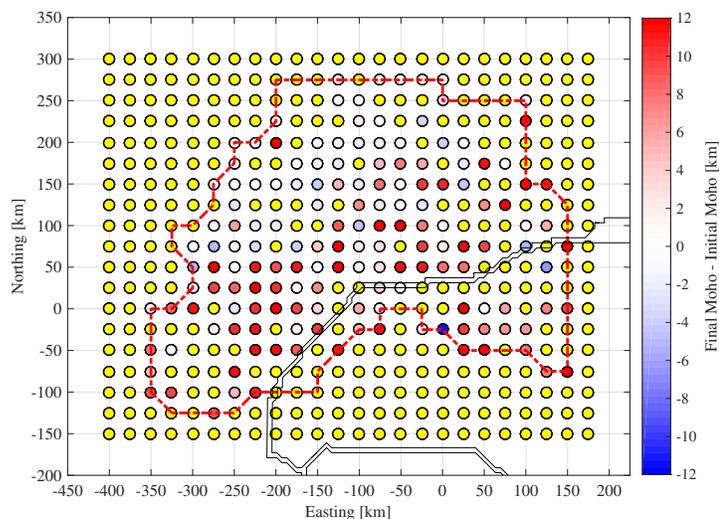


Figure 7.1.3: Difference between final Moho result and initial Moho model. Colored points from blue to red colors show respectively a negative or a positive variation compared to the initial model. Null variations are marked by white dots. Yellow points represent nodes not directly solved from the inversion procedure.

This variation in terms of Moho difference seems important, however not unrealistic, as the study of Spada et al. (2013), partly based on the RF method at individual stations, counts with a Moho uncertainty of ± 6 km for the stations involved in the investigated area. Figure 7.1.4 shows the map of lower crustal thicknesses, which is obtained as the difference between Moho and Conrad depth from the inversion results.

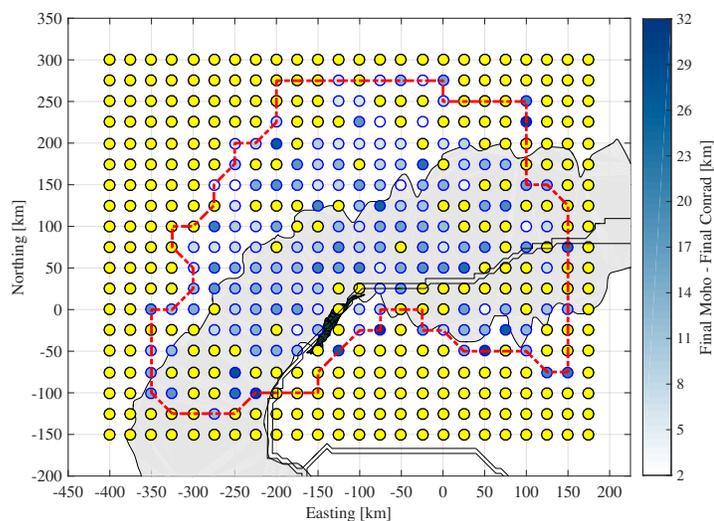


Figure 7.1.4: Thickness of the lower crust: relatively thin areas are represented by whitish colors, relatively thick areas are shown with bluish colors. Yellow points represent nodes not directly solved from the inversion procedure. Other display elements are as on previous figures.

Apart from some outliers (with values of lower crust thickness locally greater than 25 km, mostly at the edge of the study area), the majority of the nodes have values from 15 to 20 km in the Alpine arc, which drops to values around 10 km or less in the northern Alpine foreland (the initial value was set at 12 km).

7.1.2 Velocity structure

The other parameter obtained with the newly developed methodology is the 3-D velocity structure. Below I show maps of the absolute shear-wave velocities for each interface (below surface, above Conrad, below Conrad, above Moho), the average crustal V_p/V_s ratio, the velocity jump across the Conrad discontinuity, and the S-wave velocity gradient across the upper and the lower crust.

7.1.2.1 Absolute shear-wave velocities

As already discussed in Chapter 2, receiver functions are more sensitive to relative velocity variations rather than absolute shear-wave velocities, therefore the maps from the inversion results are not necessarily easy to interpret (Figure 7.1.5).

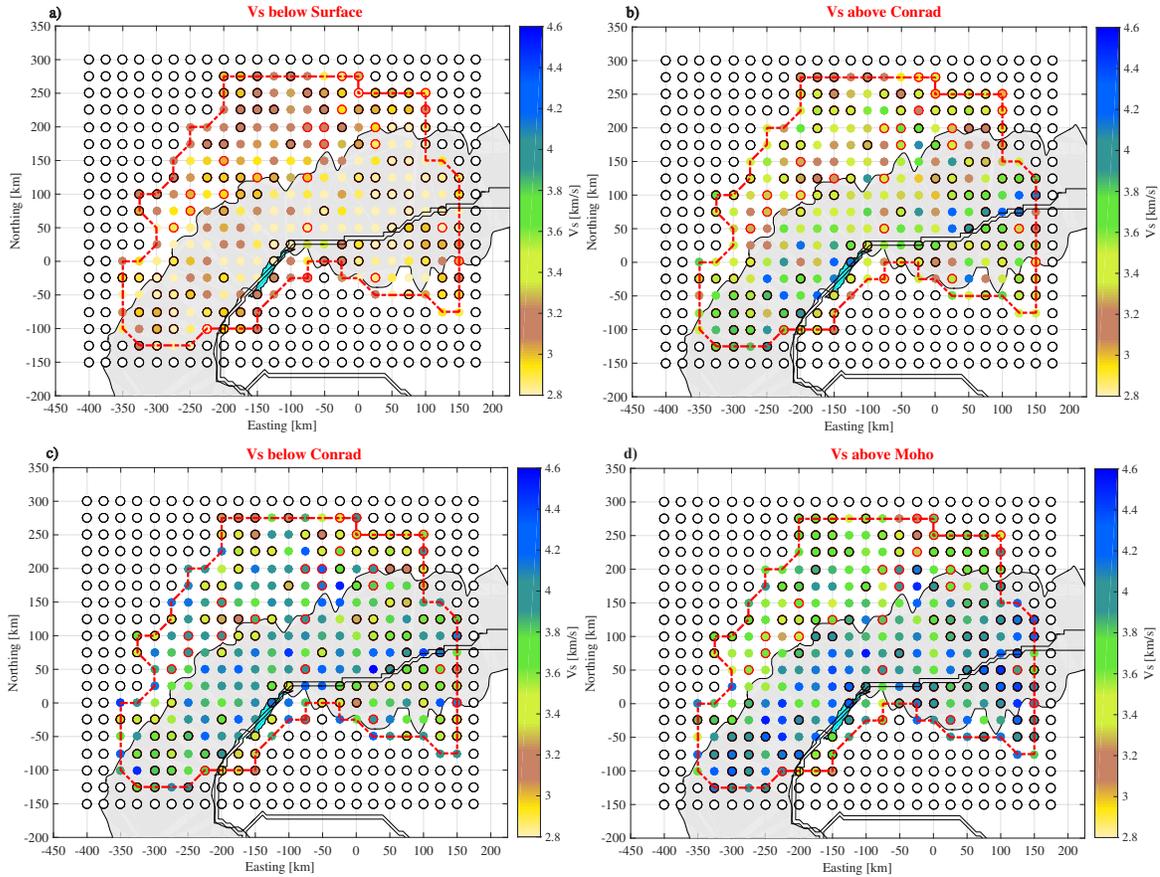


Figure 7.1.5: Final shear-wave velocity map at each interface. **a)** V_s below surface; **b)** V_s above Conrad; **c)** V_s below Conrad; **d)** V_s above Moho. Colored points without an edge are the direct results from the inversion, black edge represents interpolated value at unresolved nodes, red edge represents interpolated value following misfit based quality control. Other display elements are as on previous figures.

Regarding the V_s below the surface (Figure 7.1.5 **a**), we observe that usually in the Alpine arc we found lower values of S-waves velocity (2.8-3.0 km/s), while in the foreland we have higher values (3.0-3.2 km/s). Our model parameterization, currently using only two layers, does not allow for resolving the Molasse basin Labhart (2005); a parameterization with an additional layer would likely resolve sedimentary basin (Hetényi et al., 2006) but in our approach this would come with a high computational cost. Shear-wave velocity maps above and below the Conrad (Figure 7.1.5 **b** and **c**, respectively) are both more difficult to interpret, there is more local variability. The only feature that is clearly detected that high velocities (>4.2 km/s) are found in the European plate very close to the plate boundary Europe-Adria (Figure 7.1.5 **b**) and these values extend along the Alpine arc on the map of S velocity below the Conrad (Figure 7.1.5 **c**). This trend is confirmed and is much more evident on the map of V_s above the Moho discontinuity, where high velocities are found beneath the Alps (4.3-4.5

km/s), while lower velocities (3.6-3.8 km/s) are present in the Alpine foreland.

7.1.2.2 Average crustal V_p/V_s

Figure 7.1.6 shows the map with the average crustal V_p/V_s ratio for the entire crust.

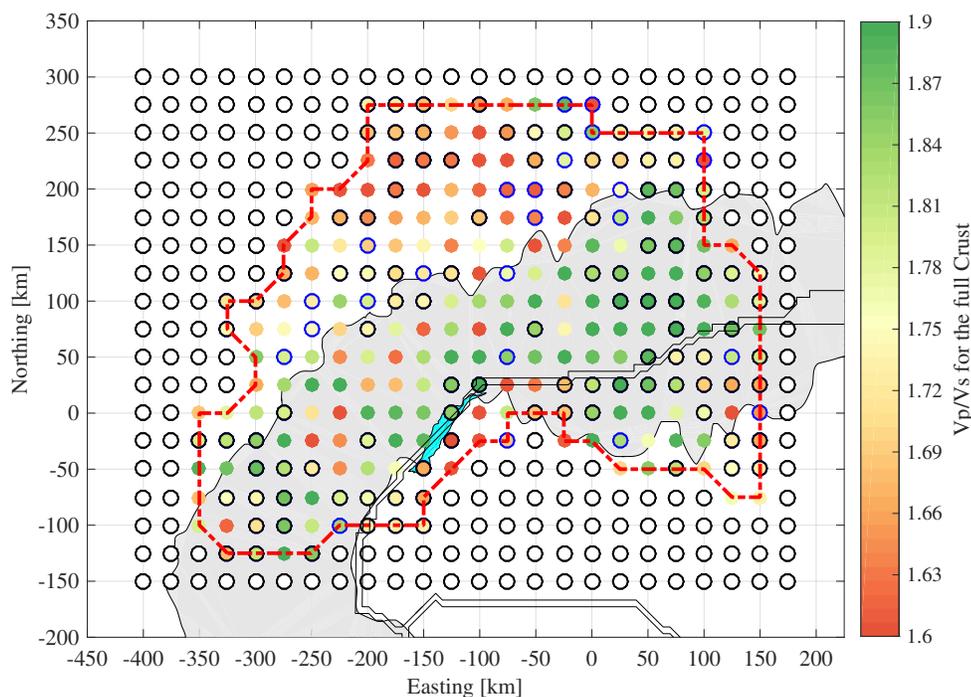


Figure 7.1.6: Final V_p/V_s map for the full crust. Colored points without an edge are the direct results from the inversion, black edge represents interpolated value at unresolved nodes, blue edge represents interpolated value following misfit based quality control. Other display elements are as on previous figures. The color scale spans from the imposed lower to upper bound for the V_p/V_s map during the inversion.

Although it is less evident with respect to the Moho map of Figure 7.1.2, also for the V_p/V_s ratio I observe a similar trend, with low V_p/V_s (1.60-1.70) in the foreland and higher values inside the Alpine arc (1.80-1.90), with some local variations¹. The petrophysical properties density ρ , compressional wave velocity and heat generation exhibit considerable variation with rock type (Rybach and Buntebarth, 1982). Our results broadly reflect of Lombardi et al. (2008) who used individual station RFs to estimate crustal V_p/V_s and correlated them with tectonic units known at the surface; however, between the regular V_p/V_s values in the foreland (Variscan basement, Mesozoic cover, Molasse basin) and higher V_p/V_s values in the Alps (Suture Zone), a lower V_p/V_s region in between, proposed to correlate with the

¹From Wadati inversion of local earthquake data in Switzerland, the average V_p/V_s ratio for the mid to upper crust is about 1.70-1.71 (Kummerow et al., 2012).

Helvetic nappes, is not resolved when using a 3-D approach. Regarding the reliability of the V_p/V_s map we can assert that the obtained result is robust since even the inversion with only 3 parameters (Conrad, Moho, V_p/V_s) pointed out a similar V_p/V_s map (see difference in Figure 6.1.5 c) with generally lower V_p/V_s in the foreland.

7.1.2.3 Velocity jump across the Conrad discontinuity and gradients inside the crustal layers

Figure 7.1.7 represents the map of the velocity jump across the intra-crustal discontinuity.

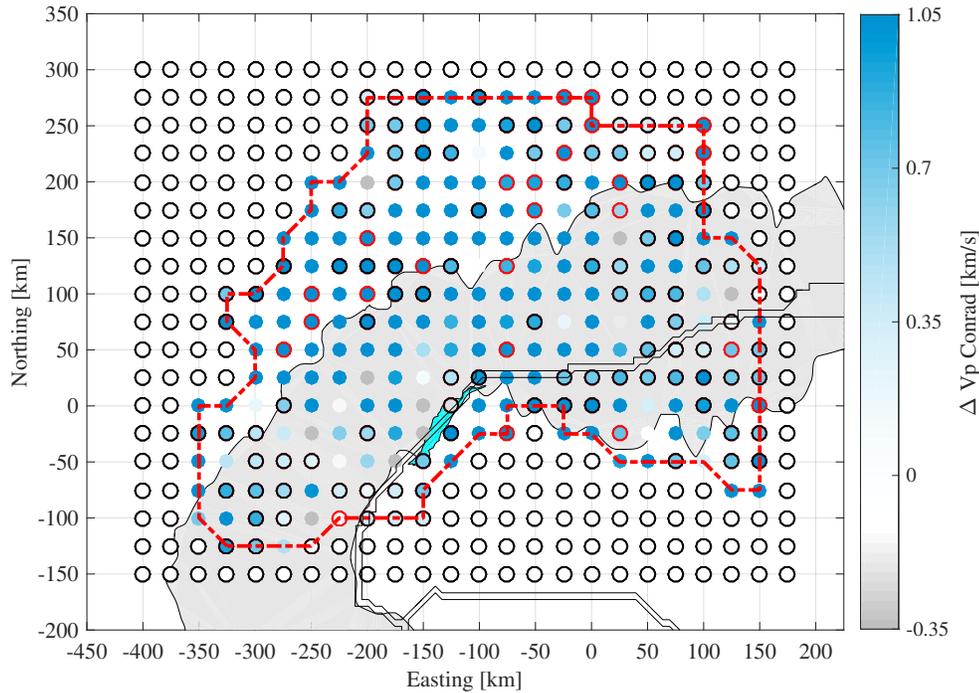


Figure 7.1.7: Final map of P-wave velocity jump across the Conrad. Colored points without an edge are the direct results from the inversion, black edge represents interpolated value at unresolved nodes, red edge represents interpolated value following misfit based quality control. Other display elements are as in previous figures. The color scale spans from the imposed lower to upper bound for $\Delta V_{PCONRAD}$ during the inversion.

As a reference, $\Delta V_{PCONRAD}$ is 0.7 km/s in the iasp91 velocity model. As we can observe, this result suggests high values (> 0.7 km/s) are present both inside and outside the Alpine arc, with locally small, null or negative values. This is not easy to interpret, while earlier RF studies have not found a clear Conrad discontinuity, but active seismic surveys did image a more reflective lower crust with respect to the upper crust. This result may suffer from the limitation of how $\Delta V_{PCONRAD}$ was implemented (symmetrically shifted V_s values above and

below the interface, with a fixed middle point). The result also underlines the complexity of the Alpine crustal structure.

To shed light on this question from another angle, I also compute the Vs gradient across the upper and the lower crust. The new model parameterization makes this calculation straightforward:

$$\gamma V_{SUC} = \frac{V_{S\text{ABOVE-CONRAD}} - V_{S\text{BELOW-SURFACE}}}{H_{\text{MOHO}}} \quad (7.1.1)$$

$$\gamma V_{SLC} = \frac{V_{S\text{ABOVE-MOHO}} - V_{S\text{BELOW-CONRAD}}}{H_{\text{MOHO}} - H_{\text{CONRAD}}} \quad (7.1.2)$$

for upper crust and lower crust, respectively.

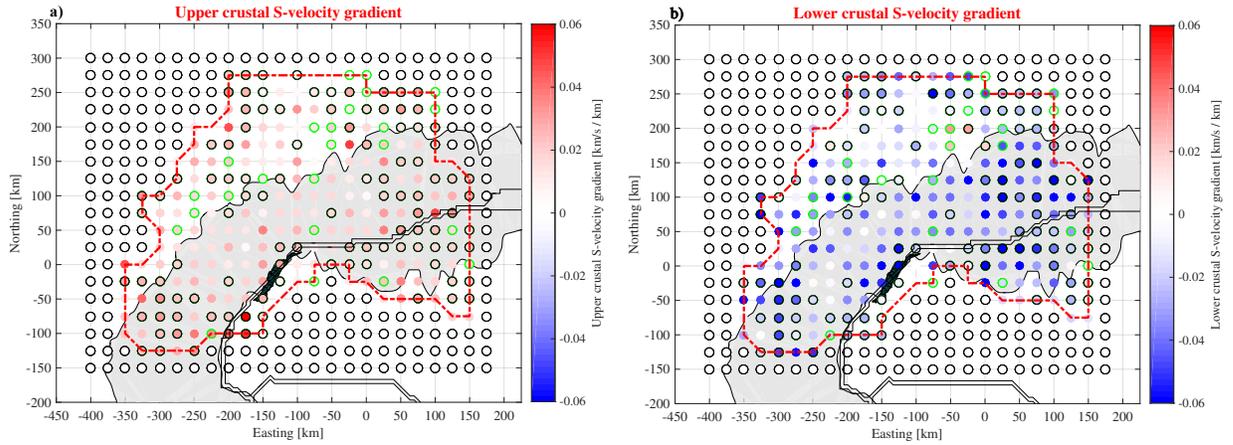


Figure 7.1.8: **a)** S-wave velocity gradient across the upper crust; **b)** S-wave velocity gradient across the lower crust. Colored points without an edge (in the scale blue-white-red) represent the direct result from the inversion, black edge represents interpolated value at unresolved nodes, green edge represents interpolated value following misfit based quality control. Other display elements are as in previous figures.

Figure 7.1.8 shows the shear-wave velocity gradients across the upper and lower crust. In the upper crust the majority area shows values between 0.02 and 0.04 km/s / km, which represents an indication that Vs increases with depth. In the lower crust, there main pattern is negative values, usually with values of -0.04 km/s / km or more, expressing shear-wave velocity decrease with depth because of large P-velocity jump at the Conrad. Only some areas of the thickened European crust in the Central Alps display no or positive gradients in the lower crust.

While the Vs increase in the upper crust and its gradient appear normal, a decrease in the lower crust is surprising, and could be an artifact. Considering that the velocity jump across

the Conrad was found to be large (Fig. 7.1.7), it is possible that lower crustal Vs gradients are subsequently distorted. A future investigation, inverting for more parameters per node and allowing larger flexibility of the velocity structure across the Conrad, may reveal the amplitude jump across the Conrad and Vs gradient in the lower crust.

7.2 3-D Vs comparison with Ambient Noise Tomography and geological interpretations

Here I show and discuss the 3-D results I get with the P-to-S converted wave tomography method with those obtained by ambient noise tomography method by Lu et al. (2018), through the representation of some Vs cross-sections in depth.

7.2.1 Map of shown profiles

Figure 7.2.1 shows the map of the sections performed in this work. The profiles were carried out in areas of interest for the comparison with other investigations where this was possible, looking at the areas with a good ray coverage.

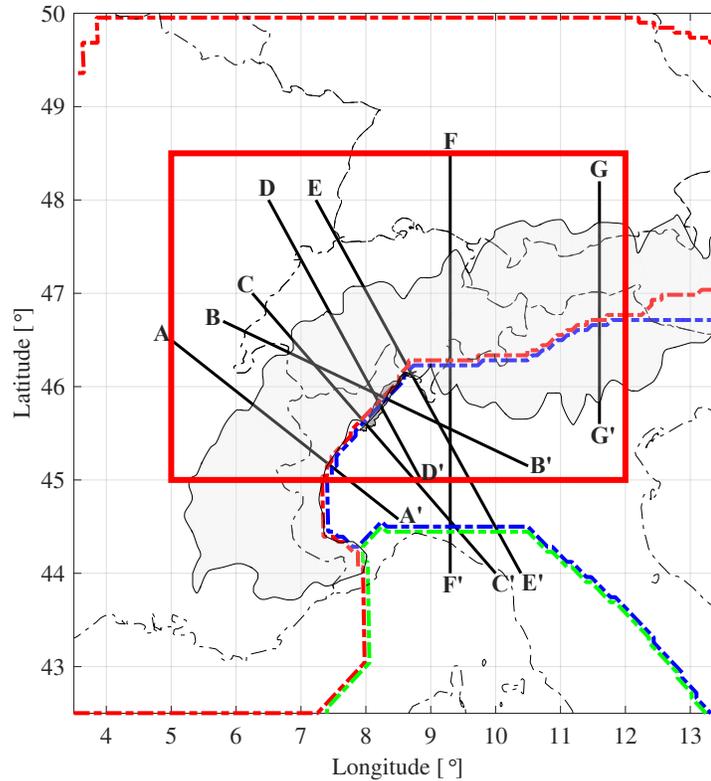


Figure 7.2.1: Map of the sections shown in this work. Red box is the study area, filled light grey area shows the Alpine arc's smoothed 800 m altitude line while filled dark grey area represents the Ivrea Geophysical Body (IGB). Dashed red, blue and green lines represent the boundary of Europe, Adria and Liguria plate, respectively. taken from the model of Spada et al. (2013). Thin black dashed lines represent the contours of the country borders. AA': ECORS-CROP profile, BB': Jura mountains-Po plain, CC': NFP-20 West, DD': Vosges-West Po basin, EE': Basel-Chiasso profile, FF': European GeoTraverse profile, GG': TRANSALP profile.

The coordinates of the sections are:

- A-A': ECORS-CROP: Lon 5.0°, Lat 46.5°; Lon 8.5°, Lat 44.58°
- B-B': Jura mountains-Po plain: Lon 5.8°, Lat 46.7°; Lon 10.5°, Lat 45.15°
- C-C': NFP-20 West: Lon 6.25°, Lat 47.0°; Lon 10.0°, Lat 44.0°
- D-D': Vosges-West Po basin: Lon 6.5°, Lat 48.0°; Lon 8.85°, Lat 45.0°
- E-E': Basel-Chiasso: Lon 7.23°, Lat 48.0°; Lon 10.39°, Lat 44.0°
- F-F': European GeoTraverse: Lon 9.30°, Lat 48.5°; Lon 9.30°, Lat 44°
- G-G': TRANSALP: Lon 11.6°, Lat 48.2°; Lon 11.6°, Lat 45.6°

7.2.2 Cross-section A-A': ECORS-CROP

A reference crustal seismic profile across the Alps was planned within a cooperation between the French ECORS (*Étude Continentale et Océanique par Reflexion et Refraction Sismique*) and the Italian CROP (*Progetto Strategico Crosta Profonda*) Projects. The seismic refraction and wide-angle reflection at high resolution preliminary survey was acquired in 1985 (Nicolas et al., 1990; Thouvenot et al., 1990; Finetti, 2005).

The Vs-cross section ECORS-CROP profile I obtained in this study is presented in Figure 7.2.2.

Result shows a Moho interface at around 30 km which increases to 55 km at the down-dip end of the European plate, with an intermediate value (40 km) beneath the European foreland, mapped at lower Vs at the surface. In the Adria domain, the Moho is as shallow as 25 km, representing the Ivrea Geophysical Body (IGB) at depth, and then it increases again southwards to 40 km below the Po plain, which also show slightly lower Vs at the surface. The horizontal location of the IGB aligns perfectly to the plate boundary, as proposed by earlier studies. The lower crustal thickness is fairly constant along this profile.

I compared my result with the one obtained by ambient noise tomography by Lu et al. (2018), which is shown in Figure 7.2.3.

Although in ambient noise tomography sharp interfaces are not clearly defined, Lu et al. (2018) chose to express them as probabilities of having an interface, based on the vertical gradient of the Vs model. In that sense, the two Vs models are rather similar in the European plate, with a southeastward dipping Moho, and a good agreement even on the Moho depths. My model locates the Moho discontinuity a little deeper (55-60 km) with respect to what is found in the model of Lu et al. (2018). There is also a good agreement in the Adriatic domain, with values of the Moho depth around 30 km and with milder dip. The IGB appears with sub-mantle but still high velocities in the ANT model.

Regarding shear-wave velocities, my model presents a velocity at around 4.0 km/s in the lower crustal layer, which is in agreement with what is found in the ambient noise tomography model. One evident difference is that in the Lu et al. (2018) model the low velocity zone (2.5-2.7 km/s) near the right end of the profile is more pronounced than in the my results. This is not a surprise, as the corresponding sediments of the Po plain affect more horizontally propagating noise than the vertically propagating teleseismic waves. Moreover, in our model the eastern part is not resolved by the RF data due to a poor ray coverage and the apparent resolution in the uppermost crust (e.g. Po basin) is the result from the *a priori* tomographic model.

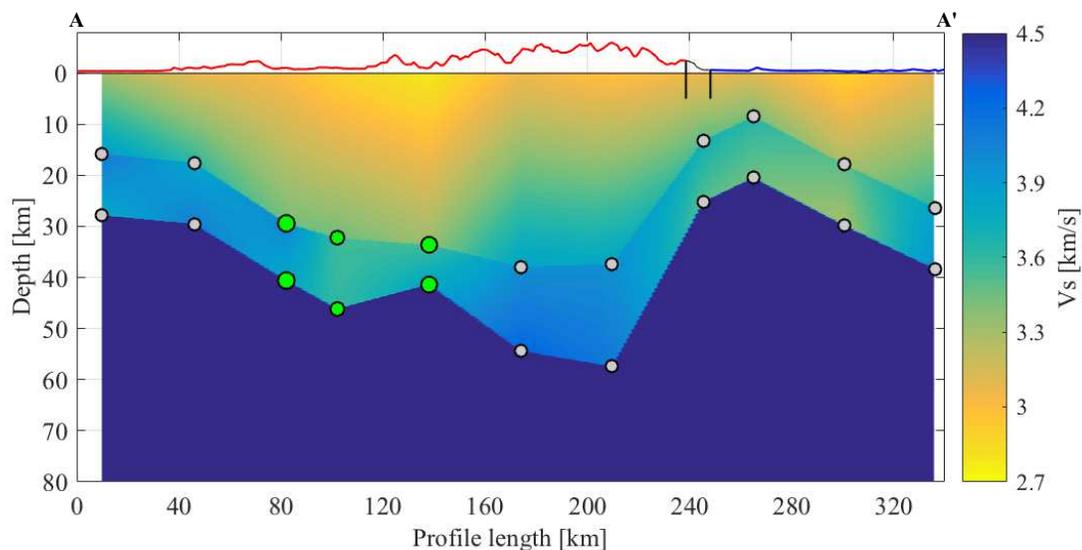


Figure 7.2.2: V_s cross-section ECORS-CROP profile obtained in this study. Background velocities represent the result of the inversion, where dots show the projection of the model nodes. Grey points are those not resolved directly by the inversion, while the green points are those accepted after performing the quality control and the size of the circle is proportional to the absolute misfit. Red solid line represents the topography in European plate, blue solid line the topography in Adriatic plate, black solid vertical line the plate boundary between the two plates. Vertical exaggeration is 2:1.

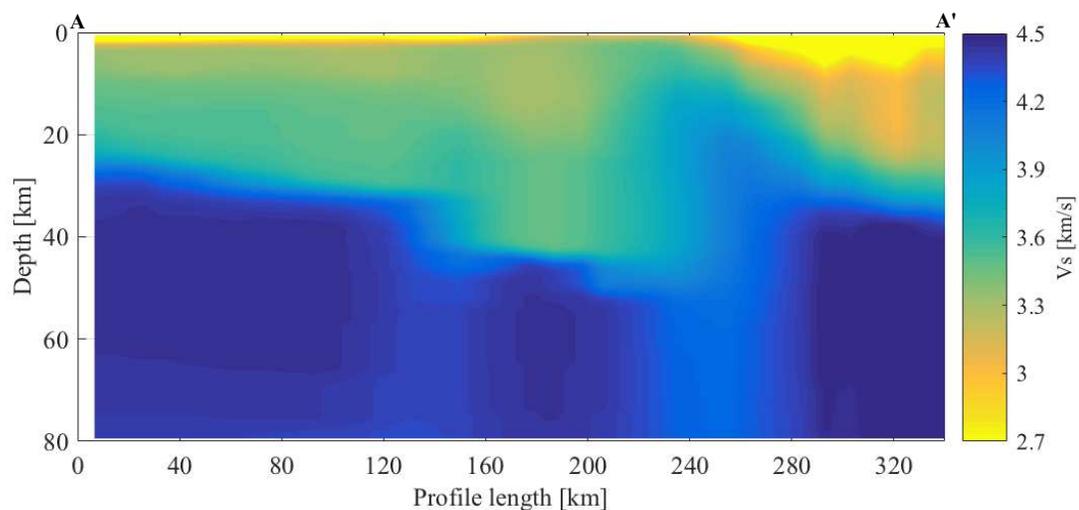


Figure 7.2.3: V_s cross-section along the ECORS-CROP profile obtained by ambient noise tomography by Lu et al. (2018). Vertical exaggeration is 2:1.

7.2.3 Cross-section B-B': Jura mountains-Po plain

Figure 7.2.4 shows the result for a section that cross in a perpendicular way the Ivrea Verbano Zone (IVZ) starting from the Massif du Jura, which is a sub-alpine mountain range located north of the Western Alps (Hölder and Lotze, 1964), crossing the Geneva Lake, the canton of Valais, into Po Plain, until the South-West of the city of Verona, in Italy.

In the cross-section we can observe a progressive deepening of the Moho (values from 25 to 50 km from West to East) in the European plate that reflects the principle of isostasy at the presence of a crustal roots beneath collisional orogens Fischer (2002). In the Adriatic domain we find a small Moho depth variation, with typical continental crustal values of 35-40 km below the Po plain. The supposed position of the Ivrea Geophysical Body is not resolved sharply, in lack of close-enough nodes to the profile, but the general difference of Moho depth across the Europe-Adria plate boundary, of about 20 km, is clear. The lower crust seems somewhat thicker in Europe than in Adria. These main characteristics are also found in the tomographic model of Lu et al. (2018) presented in Figure 7.2.5.

As in our model, we observe a progressive variation in depth of the crust-mantle interface in the European domain and a rather constant Moho at 40 km below the Po plain. As occurred for the section A-A', also for this profile our model is not able to detect lower velocities surroundings beneath the sediments, that instead are found with values around 2.7 km/s in the model of Lu et al. (2018).

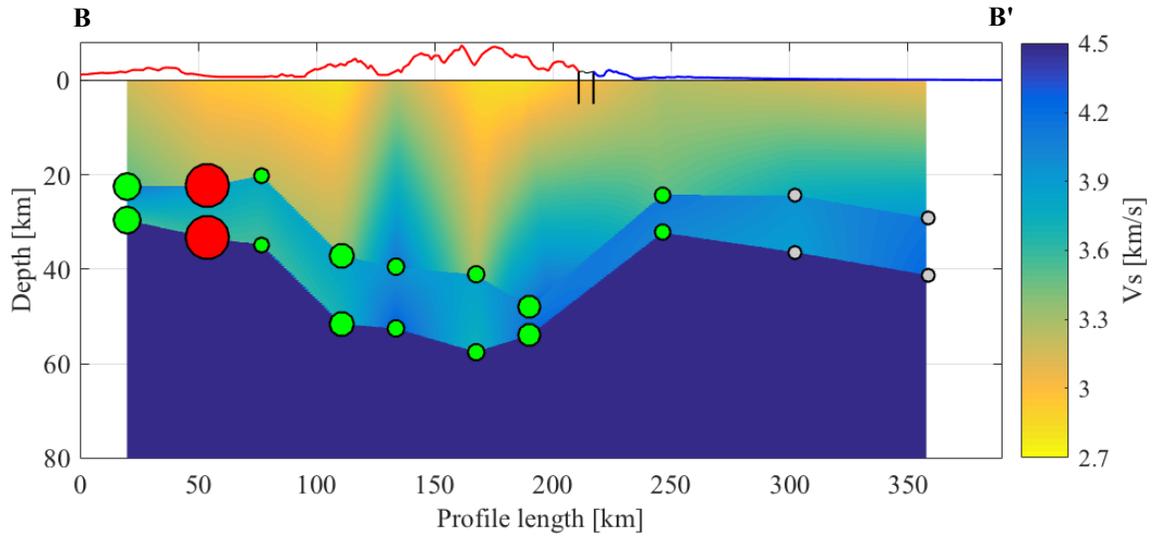


Figure 7.2.4: $B-B'$ V_s cross-section (Jura mountains-Po plain) obtained in this study. Background velocities represent the result of the inversion, where dots show the projection of the model nodes. Grey points are those not resolved directly by the inversion, green and red points are those accepted or discarded after performing the quality control, and the size of the circle is proportional to the absolute misfit. Red solid line represents the topography in European plate, blue solid line the topography in Adriatic plate; black solid vertical lines the plate boundary between the plates. Vertical exaggeration is 2:1.

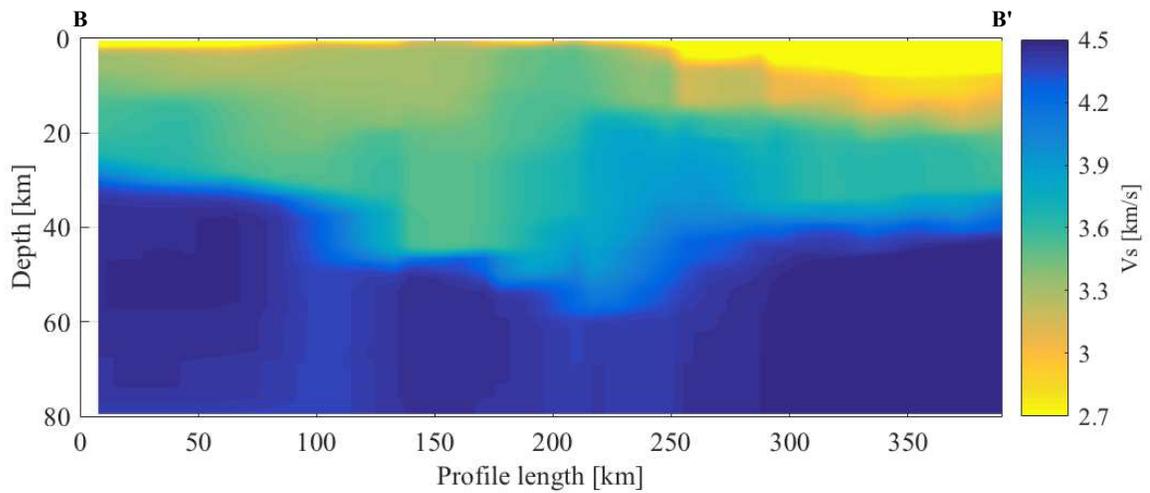


Figure 7.2.5: V_s cross-section $B-B'$ (Jura mountains-Po plain) obtained by the ambient noise tomography by Lu et al. (2018). Vertical exaggeration is 2:1.

7.2.4 Cross-section C-C': NFP-20 West

Figure 7.2.6 **a)** shows the result for the section NFP-20 West (Kissling et al. (2006); Schmid et al. (2017)). This transect follows the trace of the Swiss geophysical campaign NFP 20 “Deep Structure of Switzerland”, whose results were published in Pfiffner et al. (1997).

In our model, I observe a steep slope of the Moho that corresponds to the plate boundary: in the European domain the Moho is at nearly 60 km depth, and then a major jump towards the East sees the Moho appear at a depth of less than 30 km, in correspondence with the Ivrea Geophysical Body with typical mantle velocities ($V_s = 4.5$ km/s). In the Adriatic sector, the Moho depth increases again to the SE. In general the same features are imaged by the ambient noise tomographic model (Figure 7.2.6 **b)**, but in our results the plate boundary is better resolved using vertical rays at the plate boundary (at 200 km distance of the profile length), while it appears less clearly and shifted or smeared to the SE (at $x=240$ km distance) in the Lu et al. (2018) model.

Figure 7.2.6 **c)** and **d)** show 2 tectonic transects covering the same investigation area of the profile NFP 20-West. In both of these geological sections (7.2.6 **c)** is by Schmid et al., 2017, **d)** is by Escher et al., 1997) we can observe a good match both in slab shape and in the position of the division into European upper and lower crust compared with our V_s results. The Moho jump we observe in our model is located within the Sesia Unit, that is known to exhibit eclogite facies conditions in its most internal parts only Bousquet et al. (2012a), immediately adjacent to the Ivrea mantle rocks. However, the Ivrea Geophysical Body is thought to be even denser, and thus of higher velocity, and this location, according to the most 3-D analysis (Scarponi et al., 2020).

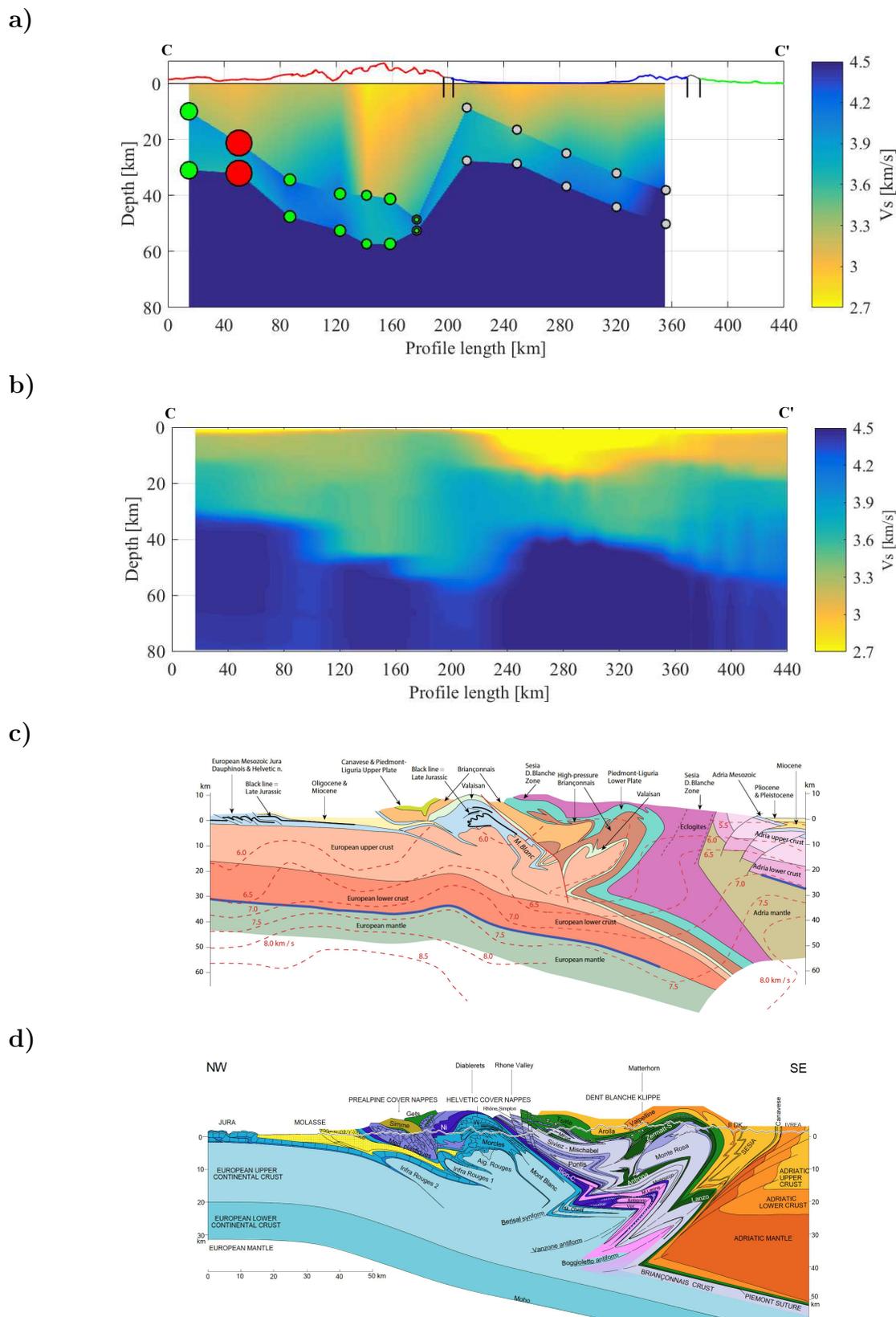


Figure 7.2.6: **a)** V_s cross-section NFP-20 West obtained in this study. **b)** V_s cross-section NFP-20 from ambient noise tomography by Lu et al. (2018). **c)** NFP 20-West geological-tectonic transect by Schmid et al. (2017). **d)** Schematic geological profile by Escher et al. (1997) through the Western Swiss-Italian Alps from the Mont Tendre (Jura) in the NW to the Val Sesia in the SE. See Figure 4.2.5 for legend details.

7.2.5 Cross-section D-D': Vosges-West Po basin

Figure 7.2.7 shows the result for a transect which goes from the Vosges, a range of mid-elevation mountains in eastern France bounding the Upper Rhine Graben, crosses Switzerland and ends in the western part of the Po basin, to the South of the city of Milan.

I choose this section since it is an area with a very good ray coverage and therefore we have consistent information with the uncertainty on the nodes examined. Regarding the European domain, we observe a flat Moho with values around 25 km in the Western part, and then a rather steep deepening to depths around 60 km in the Eastern part. There is a Moho step across the plate boundary, to the Ivrea Geophysical Body, after which the Moho smoothly goes to values at 40 km in the Adriatic plate. Aside from one point, the lower crustal thickness seems fairly constant, and the crustal root mirrors the Alpine topography. The observed characteristics in terms of velocities and interfaces' location are in perfect agreement with what was found in the tomographic study by ambient noise, as observed in Figure 7.2.8, with the same mapping differences of the Moho jump and the sedimentary basin velocities as along the previous cross-section.

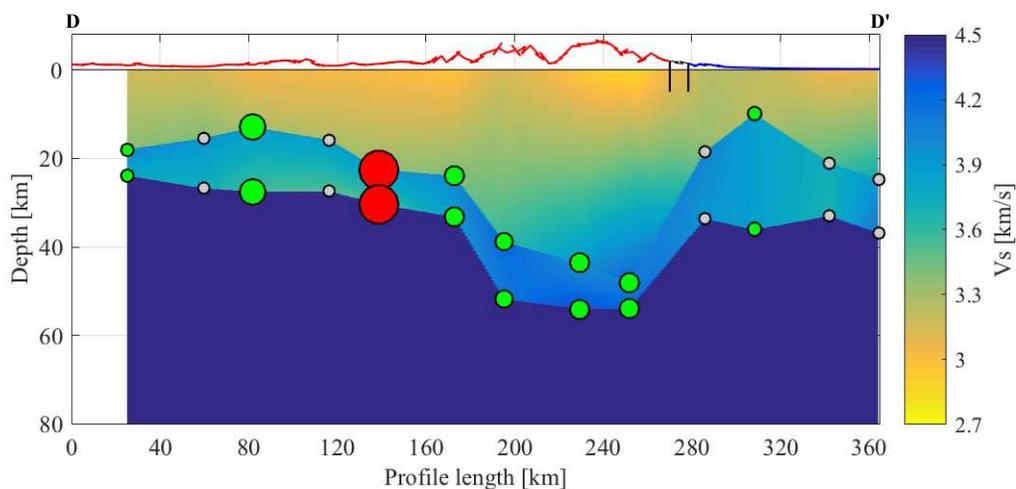


Figure 7.2.7: D-D' Vs cross-section obtained in this study. Background velocities represent the result of the inversion, where dots show the projection of the model nodes. Grey points are those not resolved directly by the inversion, green and red points are those accepted or discarded after performing the quality control and the size of the circle is proportional to the absolute misfit. Red solid line represents the topography in European plate, blue solid line the topography in Adriatic plate; black solid vertical lines the plate boundary between the plates. Vertical exaggeration is 2:1.

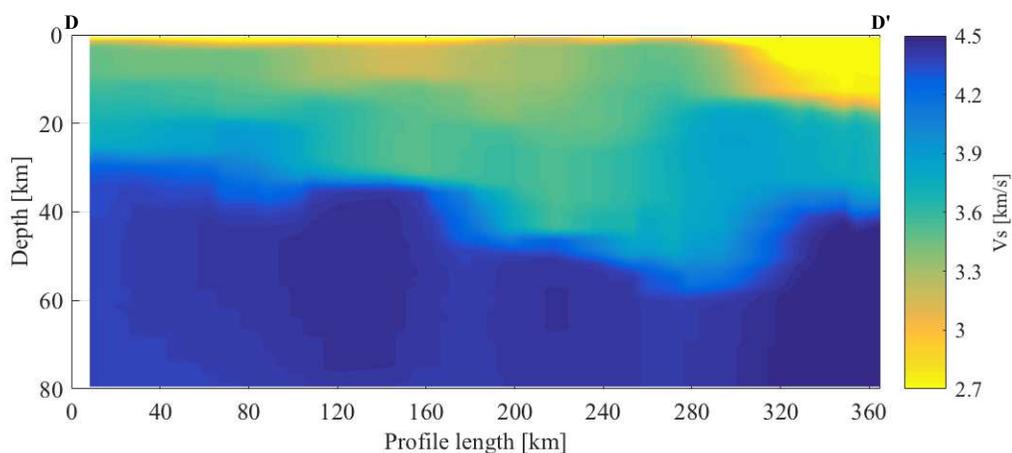


Figure 7.2.8: D-D' Vs cross section obtained by the ANT by Lu et al. (2018). Vertical exaggeration is 2:1.

7.2.6 Cross-section E-E': Basel-Chiasso

Figure 7.2.9 represents the Basel-Chiasso profile which crosses the Central Alps and Switzerland from North-Northwest to the South-Southeast, and expands a bit beyond.

Also in this case we see that the principle of isostasy is reflected: the rather shallow Moho depth values in the European foreland (around 30 km) increase 50 km-depth beneath the Alpine arc. This profile does not cross the Ivrea Verbano Zone, therefore the jump across the plate boundary is less spectacular. Nevertheless, as already seen in the previous profiles, the Moho depth of the Adria plate increase southeastwards, first to values typical for the continental crust, and then further deeper, as a sign of Adria subducting beneath the Liguria plate (not imaged by our dataset).

Figure 7.2.10 obtained by ambient noise tomography shows a very similar pattern with respect to the results I obtained both in terms of discontinuities and in terms of velocities. Again, the Po plain sediments are better imaged with ambient noise tomography.

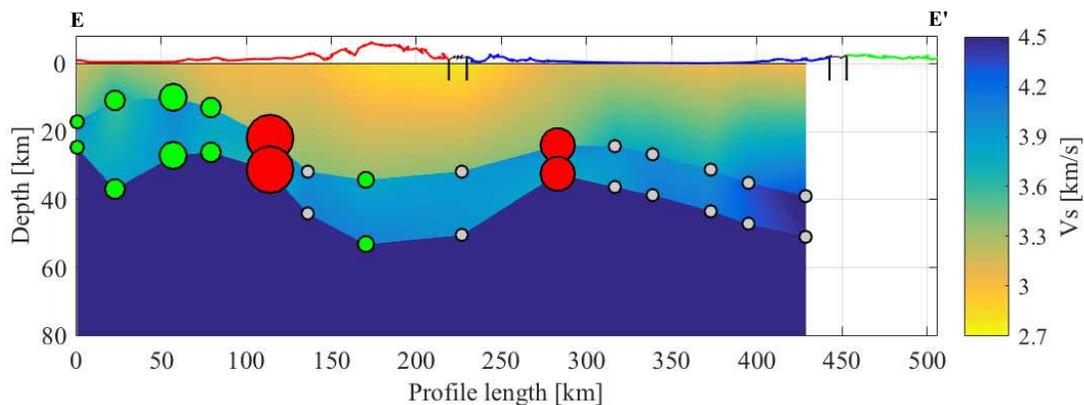


Figure 7.2.9: V_s cross-section along Basel-Chiasso profile obtained in this study. Background velocities represent the result of the inversion, where dots show the projection of the model nodes. Grey points are those not resolved directly by the inversion, green and red points are those accepted or discarded after performing the quality control and the size of the circle is proportional to the absolute misfit. Red solid line represents the topography in European plate, blue solid line the topography in Adriatic plate, green solid line the topography in Ligurian plate; black solid vertical lines the plate boundary between the plates. Vertical exaggeration is 2:1.

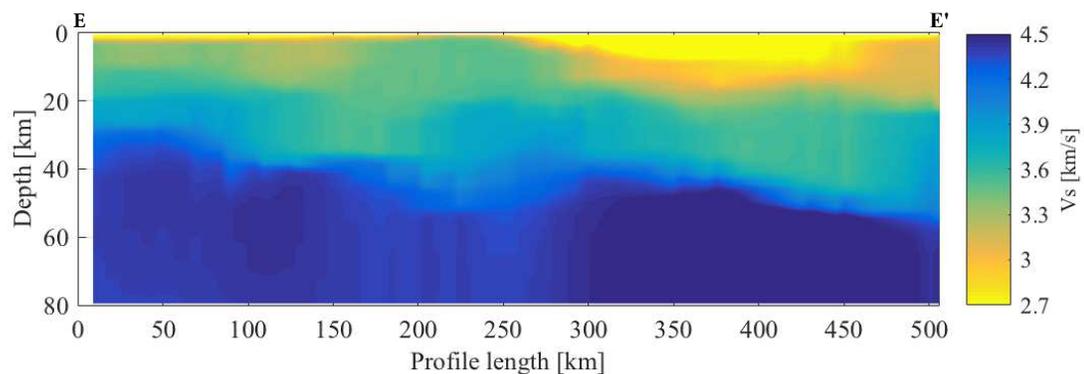


Figure 7.2.10: V_s cross-section along Basel-Chiasso line obtained by the ANT by Lu et al. (2018). Vertical exaggeration is 2:1.

Finally I report the V_p cross-section along the same profile obtained from the local P-wave earthquake tomographic investigation by Diehl et al. (2009), which also in this case shows a similar trend regarding fast or slow areas than the surrounding zones.

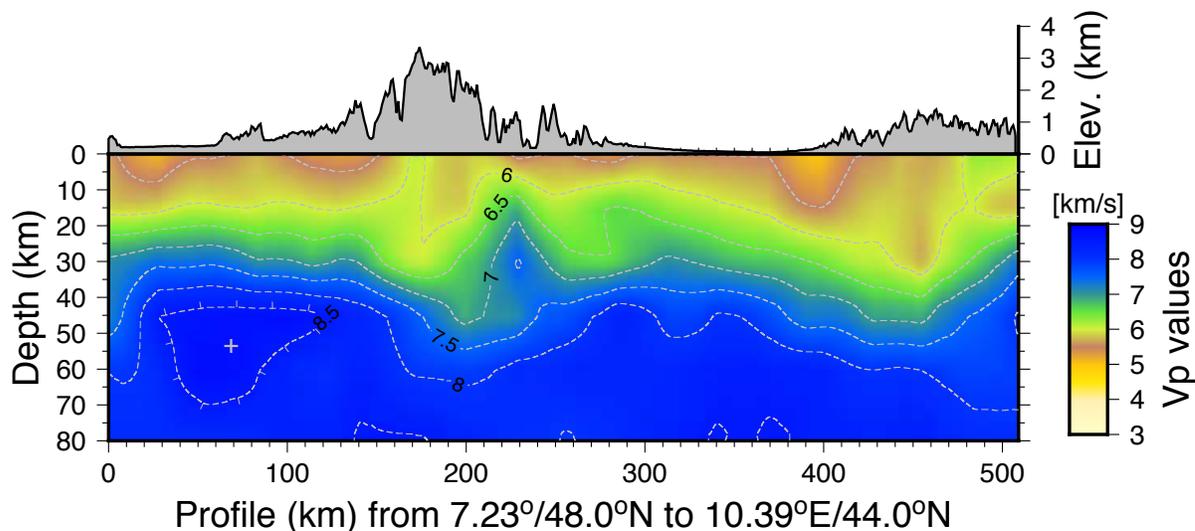


Figure 7.2.11: V_p cross-section along Basel-Chiasso line obtained by the local earthquake tomography by Diehl et al. (2009). Vertical exaggeration is 2:1.

7.2.7 Cross-section F-F': European GeoTraverse

Figure 7.2.12 shows the result from my study along a small part of the multidisciplinary European GeoTraverse transect (Galson and Mueller, 1986; Blundell et al., 1992), which crosses the Central Alps from North to the South along 9.3°E .

In the European domain the Moho discontinuity goes from values around 30 km in the northern part to values of 60 km in the middle of the Central Alps, again expressing a crustal root. There is uncertainty at the nodes in the European foreland (big red circles in Figure 7.2.12), and the supposedly slow sediments are not imaged. Across the Europe-Adria plate boundary, there is a clear step in the Moho. In the Adriatic sector, the crust-mantle interface is shallower and lies first shallower than 40 km in the Alps, with then a slight inclination southwards, entering beneath the Liguria plate (outside the profile).

The corresponding ambient noise tomography model (Figure 7.2.13) highlights the same features I detect, with the difference that it identifies the sediments with low V_s (2.7 km/s), and images a clearly smaller Moho step than my model across the Europe-Adria plate boundary at the center of the transect.

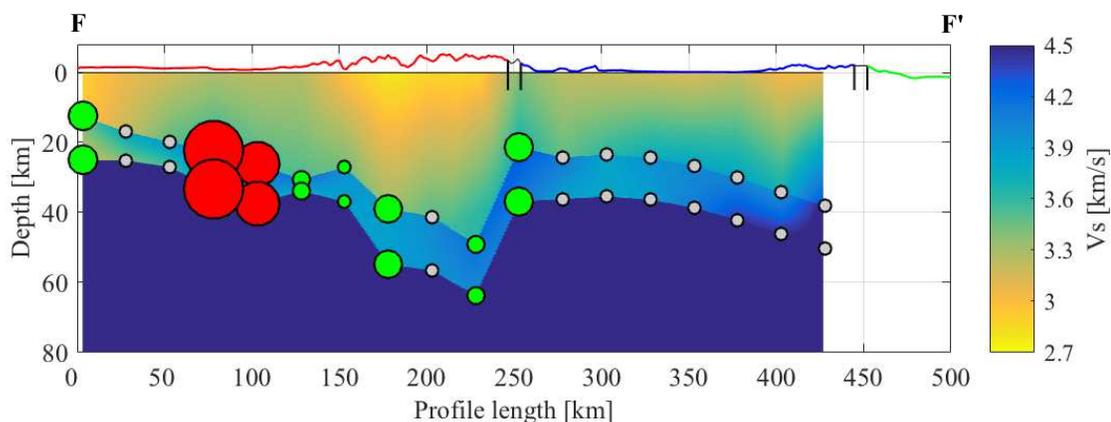


Figure 7.2.12: V_s cross-section along European GeoTraverse transect obtained in this study. Background velocities represent the result of the inversion, where dots show the projection of the model nodes. Grey points are those not resolved directly by the inversion, green and red points are those accepted or discarded after performing the quality control and the size of the circle is proportional to the absolute misfit. Red solid line represents the topography in European plate, blue solid line the topography in Adriatic plate, green solid line the topography in Ligurian plate; black solid vertical lines the plate boundary between the plates. Vertical exaggeration is 2:1.

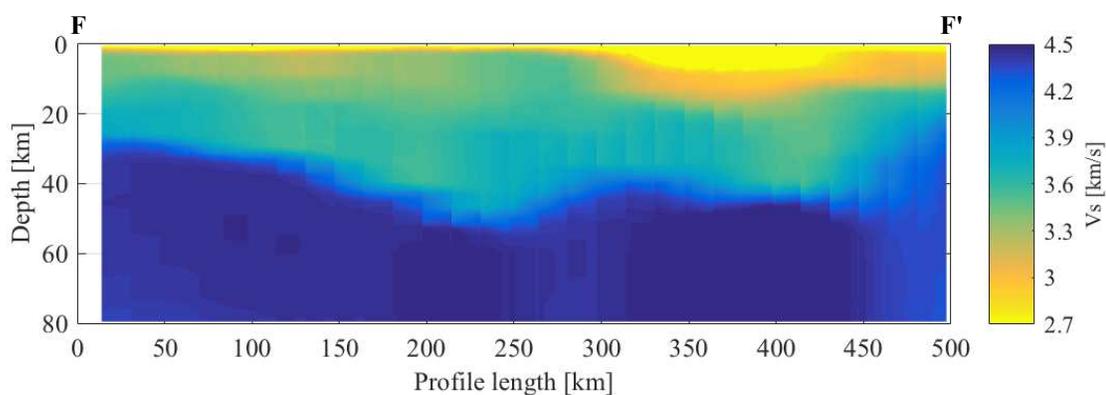


Figure 7.2.13: V_s cross-section along European GeoTraverse transect obtained by the ambient noise tomography by Lu et al. (2018). Vertical exaggeration is 2:1.

7.2.8 Cross-section G-G': TRANSALP

The TRANSALP project (Castellarin et al., 1992; Gebrande et al., 2001; Gebrande et al., 2002; Lippitsch et al., 2003; Lüschen et al., 2004; Kummerow et al., 2004) is an international and multidisciplinary research program for investigating orogenic processes driven by the collision of continental lithospheric plates. The target area is the Eastern Alps to show the collision of the Adriatic-African plate with the European plate. The combined passive and active seismic TRANSALP experiment produced an unprecedented high-resolution crustal image of the Eastern Alps between Munich and Venice (Kummerow et al., 2004).

Figure 7.2.14 shows the result obtained with our method of 3-D P-to-S converted wave tomography of the TRANSALP profile.

In the northern part of the profile, the Moho is at about 30 km depth and then it goes deeper until it reaches values up to 55 km within the European Alpine domain. On the other side of the plate boundary, in the Adriatic sector, the Moho resumes at a shallower depth, at around 50 km depth beneath the Alps, to then gently shallow towards to South (45 km). The results from ambient noise tomography display similar patterns (Figure 7.2.15), and also capture a relatively sharp crust-mantle velocity transition. However, ANT results do not clearly mark the crust-mantle discontinuity in the Adriatic plate, displaying a broader vertical gradient; this the most difficult part to interpret of the profile also according to previous investigations (Finetti, 2005).

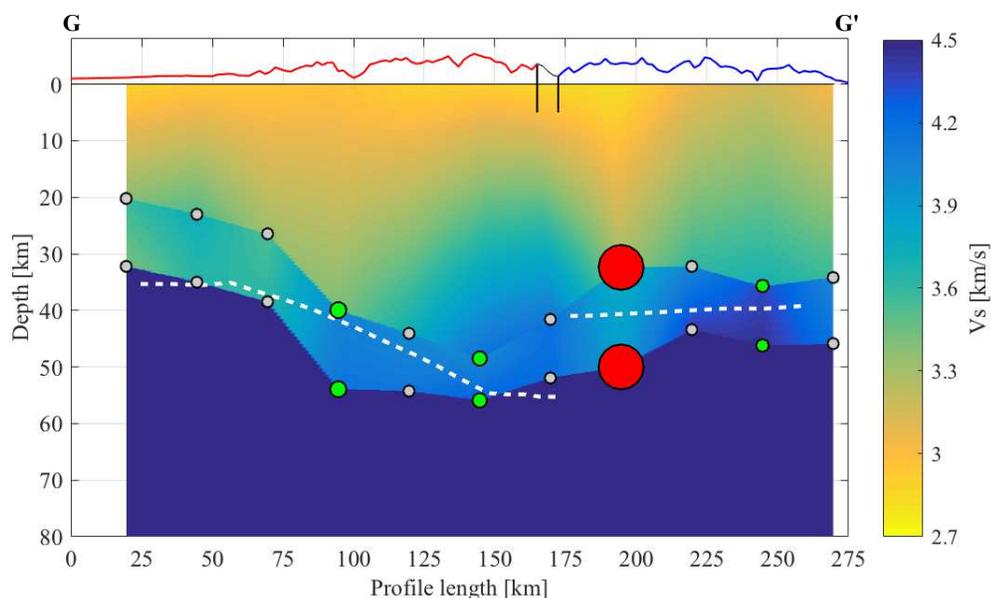


Figure 7.2.14: V_s cross-section along TRANSALP profile in this study. Background velocities represent the result of the inversion, where dots show the projection of the model nodes. Grey points are those not resolved directly by the inversion, green and red points are those accepted or discarded after performing the quality control and the size of the circle is proportional to the absolute misfit. White dashed line is the Moho from Kummerow et al. (2004). Red solid line represents the topography in European plate, blue solid line the topography in Adriatic plate; black solid vertical lines the plate boundary between the plates. Vertical exaggeration is 2:1.

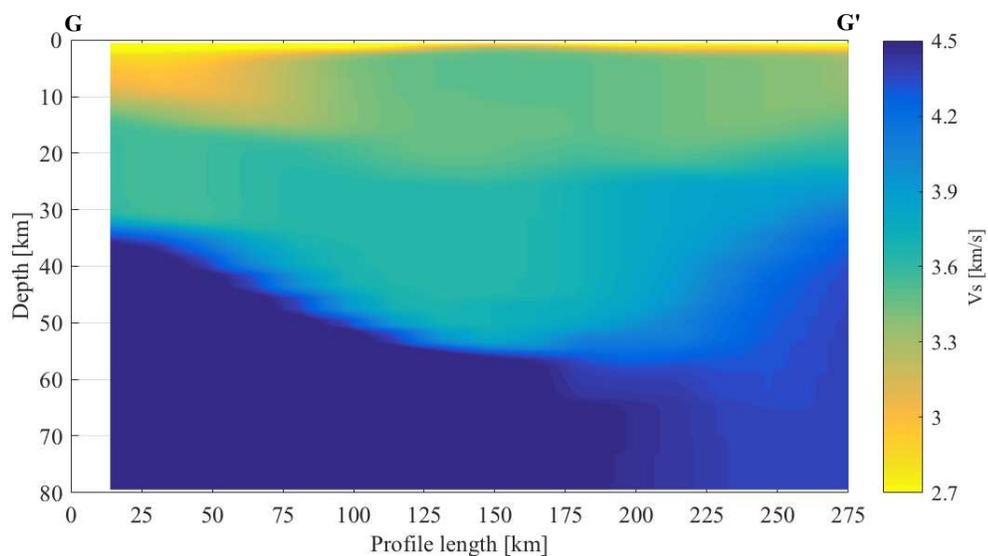


Figure 7.2.15: V_s cross-section along TRANSALP profile obtained by the ambient noise tomography by Lu et al. (2018). Vertical exaggeration is 2:1.

Our model finds an almost perfect match with the interpretation of the depth migrated short-period receiver functions TRANSALP line by Kummerow et al. (2004) shown in Figure 7.2.16. This work constitutes the first high-resolution investigation in this region, combining receiver function (RF) and active seismic techniques aiming at better constraining the deep architecture of the crust.

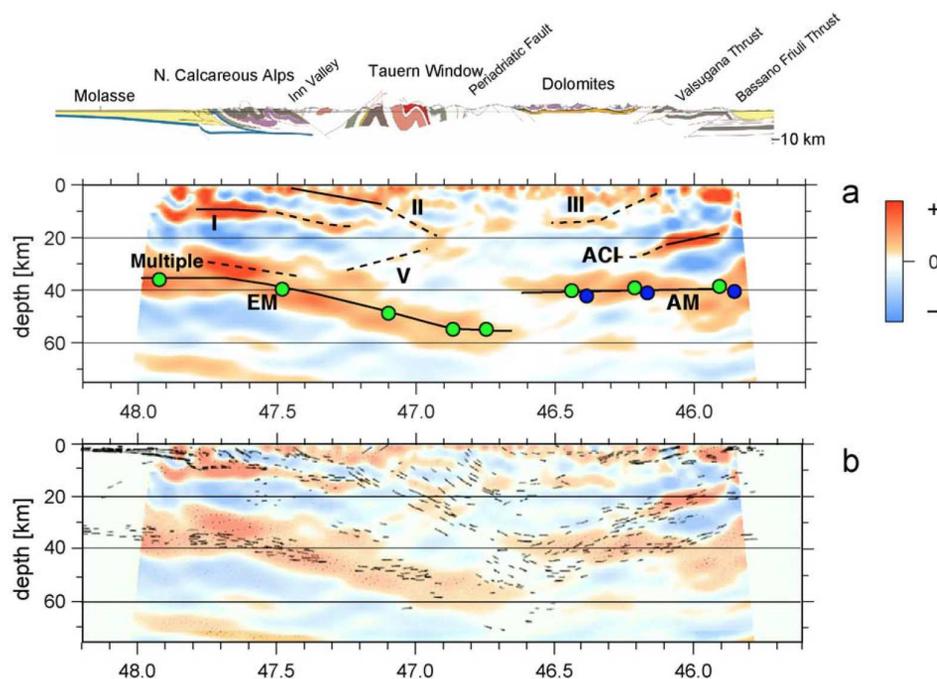


Figure 7.2.16: (a) Depth migrated short-period receiver functions along the TRANSALP line with simplified geological section according to the work of Kummerow et al. (2004). EM = European Moho, AM = Adriatic Moho, ACI = Adriatic Crust Interface, a dipping discontinuity in the middle and lower Adriatic crust. Green and blue circles select values for crustal thickness. Times are converted to depth values using a constant V_p/V_s ratio of 1.73 and $V_p = 6.0$ km/s (green circles) and $V_p = 6.3$ km/s (blue circles), respectively. (b) Superposition of receiver functions and seismic reflection line drawings.

In the original TRANSALP results, the European Moho dips gently southward from 35 km beneath the northern foreland to a maximum depth of 55 km beneath the central part of the Eastern Alps, whereas the Adriatic Moho is imaged primarily by receiver functions at a relative constant depth of about 40 km. The European Moho (EM) and the Adriatic Moho (AM) are easily identified in Figure 7.2.16. The Moho step is not only observed in the direct Moho conversions, but also in the near-vertical reflection data. The results obtained with our new approach agree well with the migrated receiver function data, while the flat Adriatic receiver function Moho seems to be somewhat different from the line drawings of the near-vertical reflection data.

7.3 3-D to 1-D crustal Moho and V_p/V_s comparison

Here I compare the results obtained with the 3-D P-to-S converted wave tomography method with the outcomes obtained through the 1-D $H-\kappa$ grid search technique (Zhu and Kanamori, 2000). For the latter, I take the results obtained by myself and by the investigation of Lombardi et al. (2008) who included dip corrections, as already presented in Chapter 2.

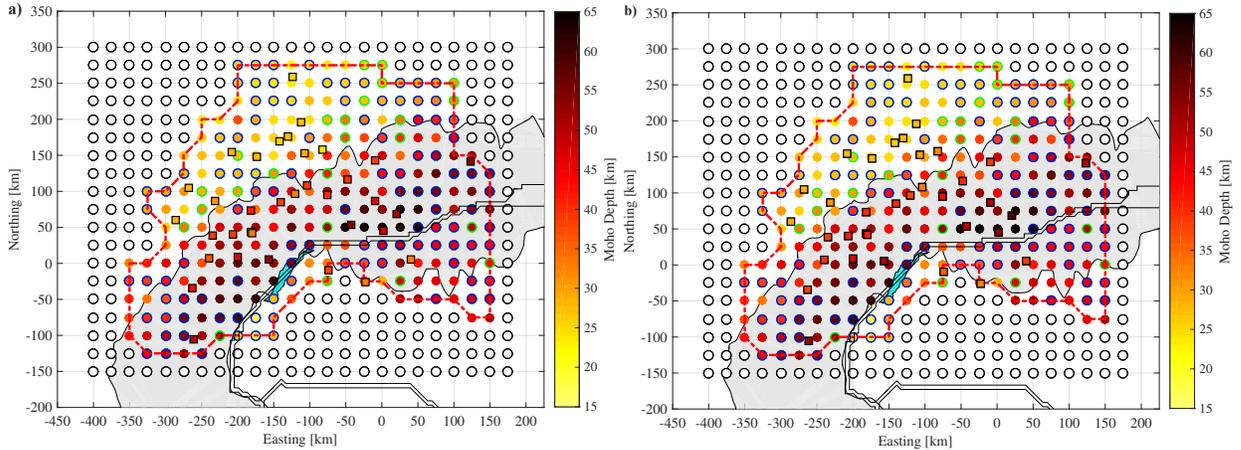


Figure 7.3.1: Final Moho depth map from the 3-D approach (circles) with superimposed 1-D results obtained in this work (a) and by Lombardi et al. (2008) and (b) using the $H-\kappa$ technique (squares). Colored points without an edge are the direct results from the inversion, blue edge represents interpolated value at unresolved nodes, green edge represents interpolated value following misfit based quality control. Other display items as on earlier figures.

Figure 7.3.1 shows the Moho depth comparison. There is a good match between the values obtained by the 3-D inversion and the ones obtained by the inversion of single station data (which has many intrinsic limitations as a method, e.g. one single value for crustal V_p), with thick crust within the Alpine area, reaching 60 km in the European domain along the Europe-Adria plate boundary and Moho depth values between 20 and 35 km in the Alpine foreland. However, locally, 5 km or larger differences exist at a few places, especially in the center of the Alpine arc (stations DAVOX, FUORN, VDL, WILA) where the Moho depth seems to be deeper compared to the surrounding area. What we observe is that the 1-D result of the work of Lombardi et al. (2008) gives in general deeper values with respect to the our 1-D Moho values, although the biggest differences do not seem to be connected to the local dip, which in the Alpine arc assumes values around 10-12°.

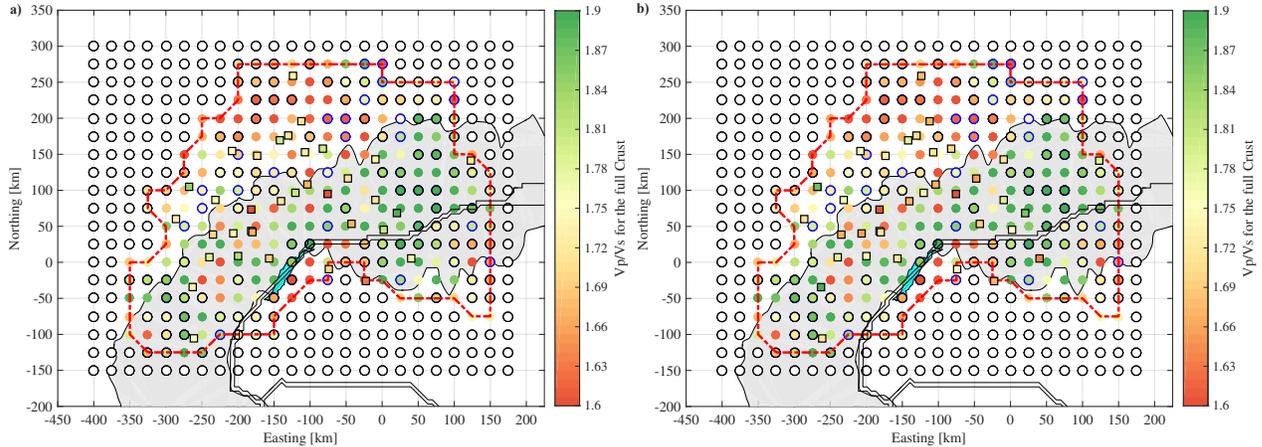


Figure 7.3.2: Final average crustal V_p/V_s map from the 3-D approach (circles) with superimposed 1-D results obtained in this work (a) and by Lombardi et al. (2008); (b) using the $H - \kappa$ technique (squares). Display as in Figure 7.3.1. The color scale spans from the imposed lower to upper bound for the V_p/V_s variable.

In several points of the map, differences of 0.1 or larger appear. For example, within the European foreland our 3-D model has values lower than 1.70 while both in our 1-D inversion and in Lombardi et al. (2008) work the V_p/V_s is higher, up to values equal to 1.78. Note that the single station V_p/V_s estimates' uncertainty on 57 stations is based between 0.04 and 0.18 (average: 0.07, median: 0.06), so part of the difference can be accounted by this effect. The remaining part of the difference is still an open question to discuss.

7.4 Lower crustal seismicity and geodynamic interpretation

Finally, in order to have an idea of the relation between seismic velocities, crustal rheology and rock composition, I make a qualitative comparison between the V_p/V_s map and the lower crustal seismicity distribution in the northern Alpine foreland.

7.4.1 Lower crustal earthquakes in the northern Alpine foreland

While the majority of the Alpine area shows relatively high V_p/V_s values, the European foreland has a contiguous area with consistently low V_p/V_s ratios. The same area is also known to host lower crustal earthquakes, as reported by Singer et al. (2014) based on Swiss and EMSC catalogue data and detailed reanalysis of focal depths and mechanisms. I compare my final 3-D velocity ratio results and the distribution of these earthquakes (Figure 7.4.1) to investigate whether there is a geodynamic interpretation possible at a larger scale.

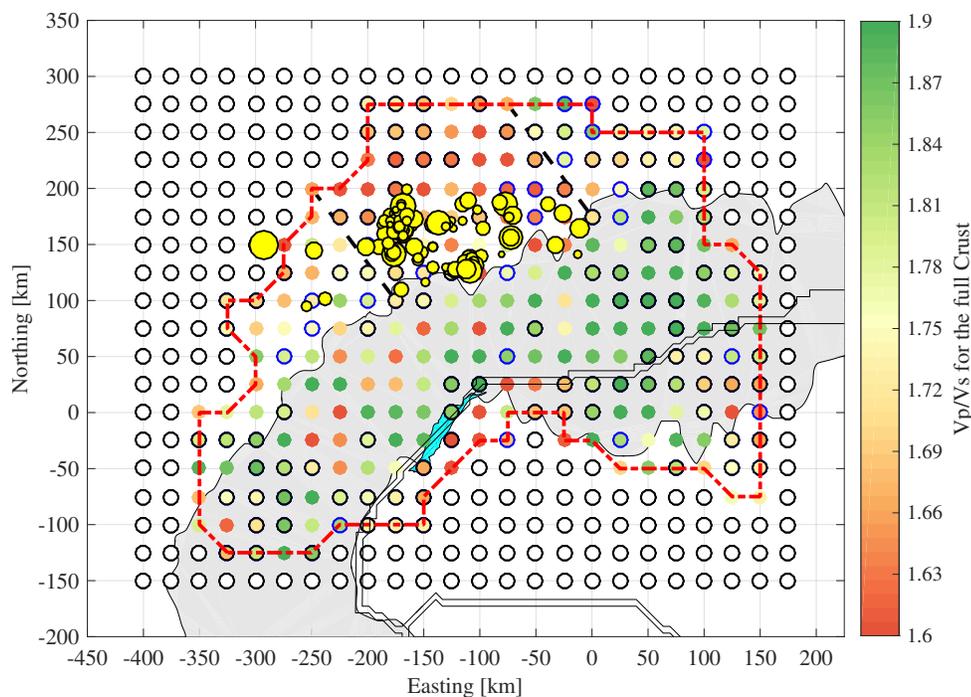


Figure 7.4.1: Average crustal V_p/V_s map after interpolation of empty and poor quality nodes (Figure 7.1.6) to which is superimposed with yellow dots the lower crustal seismicity in the Alpine region as reported in Singer et al. (2014). Size of the circle is proportional to the magnitude of the event. Two black dashed line segments represent represent the V_p/V_s sector when $V_p/V_s < 1.70$.

As we can observe on Figure 7.4.1, 95% (all except 3 in the foreland, and an event near Besançon) lie in a zone of the model with crustal V_p/V_s values lower than < 1.70 . They do not seem to align along well definable faults or fault zones. According to Singer et al. (2014) this focus of seismicity in Northern Switzerland correlates well with the lateral extent of the European slab beneath the Central Alps where it is still attached to the European lithosphere, based on tomographic arguments (Lippitsch et al., 2003).

A zoom of the area of interest is reported in Figure 7.4.2. The western bound of the area is roughly along the Bern - Biel - Porrentruy - Belfort line; this is further NE than the shallow Fribourg Lineament (Vouillamoz et al., 2017). The eastern bound starts East of Lindau, crosses NE of Ravensburg, and has a projected end between Albstadt and Tübingen. There is no primary geological boundary or marker at the surface, according to the map Geological Map of Switzerland, at the 1:500 000 scale, published by the Swiss Federal Office of Topography (Swisstopo). Most of this area is covered by the foreland basin sediments. On the other hand, the nature of the lower crust is not well known, and it can be that the segment with lower crustal earthquakes and relatively low V_p/V_s is inherited.

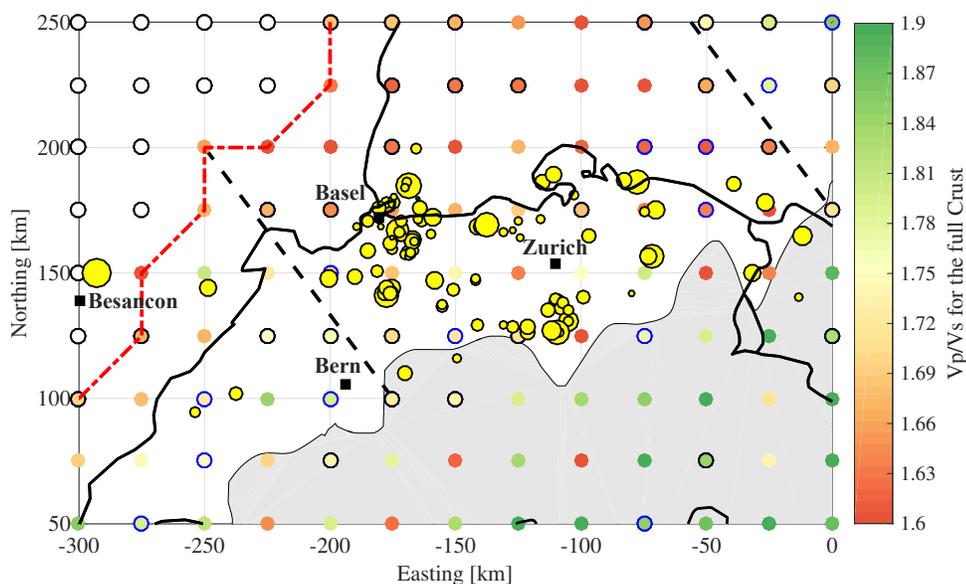


Figure 7.4.2: Zoom of the Figure 7.4.1, black squares show the location of the main cities in the study area, black solid lines are the borders of the country, dashed black contours the selected area with V_p/V_s value less than 1.70.

7.4.2 Geodynamic considerations

As a general trend one could expect that the crust is density sorted, and that at the lower crustal level we should find rocks with more mafic compositions, with a high Poisson's ratio and a relatively high V_p/V_s (>1.75), as reported in several works in literature (e.g. Barton, 1986; Zandt and Ammon, 1995; Christensen, 1996). Based on these considerations, the lower crust in the area of full-crustal $V_p/V_s < 1.70$ cannot be extremely mafic.

The reasons why there are lower crustal earthquakes in a zone of low V_p/V_s and not elsewhere can be possibly linked to several factors, such as the thermal state, regional tectonic strain rate, petrological composition and fluid concentration. All these parameters play a role in controlling the depth of transition between brittle and ductile deformation within the crust.

Earthquakes which are clear evidences of brittle deformation are very important for our understanding of the properties of the continental crust, and this is particularly true in a region of such structural complexity and of ongoing tectonic activity as the Alps and its foreland (Deichman and Baer, 1990). In view of the high temperatures and an assumed felsic composition, hence quartz or plagioclase rheology in the lower crust, as deduced from surface heat flow and seismic velocities, the deep seismicity below the northern Alpine foreland could be a consequence of the upward migration of fluids under near-lithostatic pore pressures

(Deichmann and Rybach, 1989).

It remains a problem of reconciling the existence of seismicity in a depth range where temperatures are expected to be high, and where V_p/V_s ratio is low. In general, the combined effect of increasing pressure and temperature will result in an increase of both velocities and Poisson's ratio (Kern, 1982). However, under the influence of a high temperature gradient, some rocks can exhibit the opposite trend, suggesting that the assumption of a dry crust is indeed unrealistic (Deichmann and Rybach, 1989). Moreover, the behavior of seismic velocities for a given rock in the presence of water is quite complex, because it depends critically on the degree of saturation on pore pressure. The role of water is also critical for the assessment of the crust's rheology: in the presence of water, both the frictional resistance to failure and the resistance to ductile deformation are decreased relative to a dry rock (Deichmann, 1992). One possible source of water in the lower crust could be dehydration reactions: there are several minerals, which undergo dehydration (or decarbonation) reactions at higher temperatures, thereby releasing fluids into existing or newly formed pore spaces (Fyfe, 2012). This mechanism has been suggested in conjunction with the transformation of serpentinite to olivine, as a possible explanation for deep subduction-zone earthquakes. However, serpentinite weakening is unlikely to be a factor controlling lower-crustal seismicity below northern Alpine foreland, first because of its geological history, and second because a significant amount of either serpentinite or olivine (Christensen, 1966) is inconsistent with the low V_p/V_s ratio obtained in this study and the low Poisson's ratios (0.23-0.24) computed in the study of Deichmann and Rybach (1989).

While fluids may explain the presence of lower crustal earthquakes, the spatial concentration of lower crustal earthquakes and of low crustal V_p/V_s ratios along the arc of the Alpine foreland still needs to be explained. The surface heat flow is not known to vary significantly along the Molasse foreland basin, as illustrated by geothermal project initiatives all along the basin from Geneva to Bavaria (Chelle-Michou et al., 2017). The strain rate, which is an important factor influencing rock deformation, is also unlikely to vary a lot in the slowly deforming Alpine domain, especially along the arc. What can easily vary spatially is the composition of the rocks in the lower crust. This would explain the consistently low V_p/V_s ratios, and also the earthquakes in two possible ways. First, if these rocks carry more water, following dehydration they may provide the fluids mentioned above. Second, and independently of the fluids, these rocks may have weaker rheologies than in the surrounding areas of the lower crust. This means that under the same strain and stress conditions, these mechanically weaker rocks (such as granites and gneisses) reach their strength limit and yield in form of brittle ruptures. At the same conditions, stronger rocks (such as granulites or dry diabase) would still support the mechanical stress and remain apparently aseismic. A

quantitative demonstration for our domain remains to be done, using rheological properties reported in the literature, such as by Rybacki and Dresen (2004) and Afonso and Ranalli (2004). Such compositional variations in the lower crust are most likely inherited from the geological history of these zones.

Inherited structures in general play a key role in controlling strain localization on a local scale, during early stages on rifting (Piqué and Laville, 1996; Tommasi and Vauchez, 2001) or also in orogeny as reactivated structures. Inheritance can be defined as the difference between a *real* and an *idealized* lithosphere (Manatschal et al., 2015). Figure 7.4.3 shows in a conceptualized manner the 3 types of inheritance in a section across a lithosphere at the end of an orogenic collapse but before the onset of rifting.

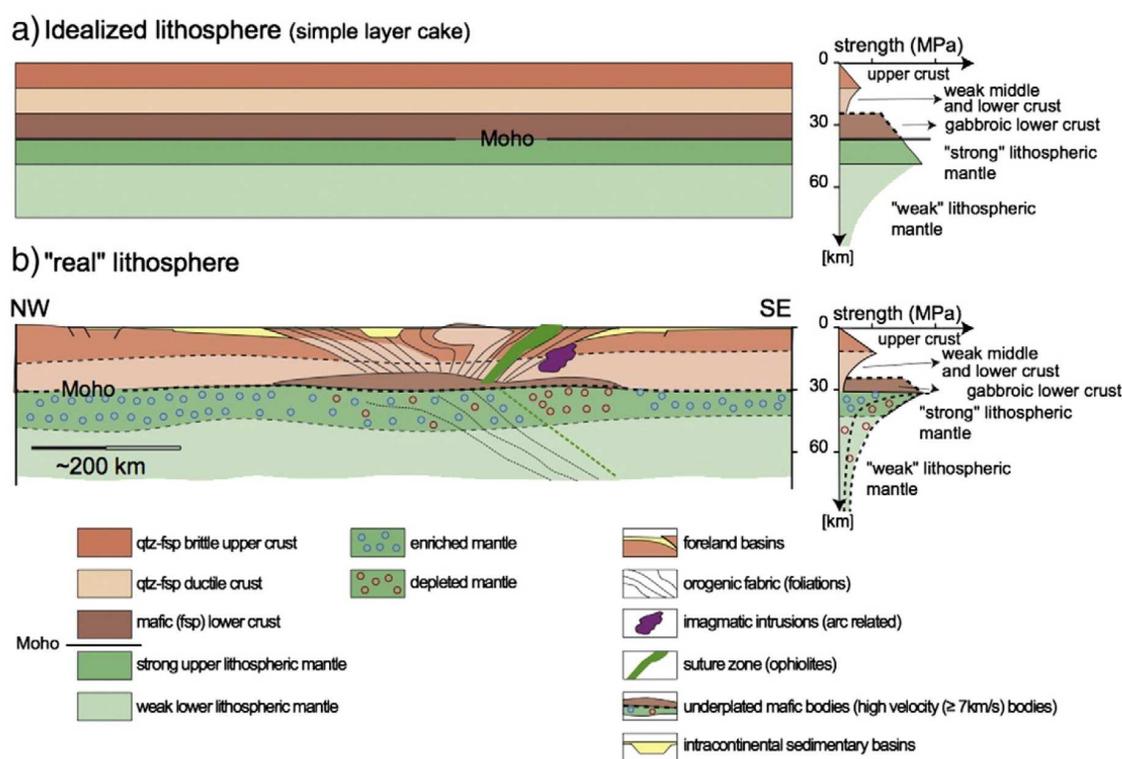


Figure 7.4.3: Two different type of lithospheric scale sections (from Manatschal et al. (2015)). (a) Idealized lithosphere made of a thermally equilibrated layer cake; (b) Real post-orogenic lithosphere with inherited structural and compositional complexity.

The idealized lithosphere is defined as a thermally equilibrated horizontally homogeneous layer-cake lithosphere, made of quartzo-feldspatic upper and middle crust and a mafic lower crust and a peridotitic mantle. In contrast to this idealized lithosphere, a real lithosphere may differ by its thermal state, compositional heterogeneities and structural complexity. Manatschal et al. (2015) distinguish between three types of inheritance, called thermal (a

function of the age of the lithosphere), structural (with mechanically “weak” zones) and compositional (related to rheology difference).

Based on these concepts, the geodynamic interpretation of the study zone like the Central Alps is a complex exercise in which some simplifications are generally considered. These areas are represented by the continental domains, which are the results of polyphase tectonic, magmatic, and metamorphic history since Variscan times and even earlier. Therefore, it cannot be described by a few physical parameters only, but it needs to take into account the structural and time-integrated inheritance of the lithosphere. This is increasingly difficult with increasing geological time and increasing depth. For better understanding the lower crust, which is unlikely to be sampled by continental drilling, joint inversion of geophysical data and multi-parameter numerical modeling of its evolution may allow to step forward.

Chapter 8

Conclusions and future perspectives

In this last chapter, I compare the newly obtained, final 3-D Moho depth model of the Central Alps with other previous investigations using different methods, and then I summarize the results presented and discussed in the previous chapter. I describe briefly the uncertainties and highlight the applicability of the method with its own limits. Finally, I suggest some ideas for future developments of the presented technique.

8.1 Results summary

The first application of the new method presented in this thesis allowed to quantitatively constrain the Alpine crustal structure. Figure 8.1.1 shows the geometry of the main interface, the final Moho depth model of the study area. In this 3-D image no smoothing or regularization technique is applied and this is the reason why in some points the curve has some “spiky” geometries.

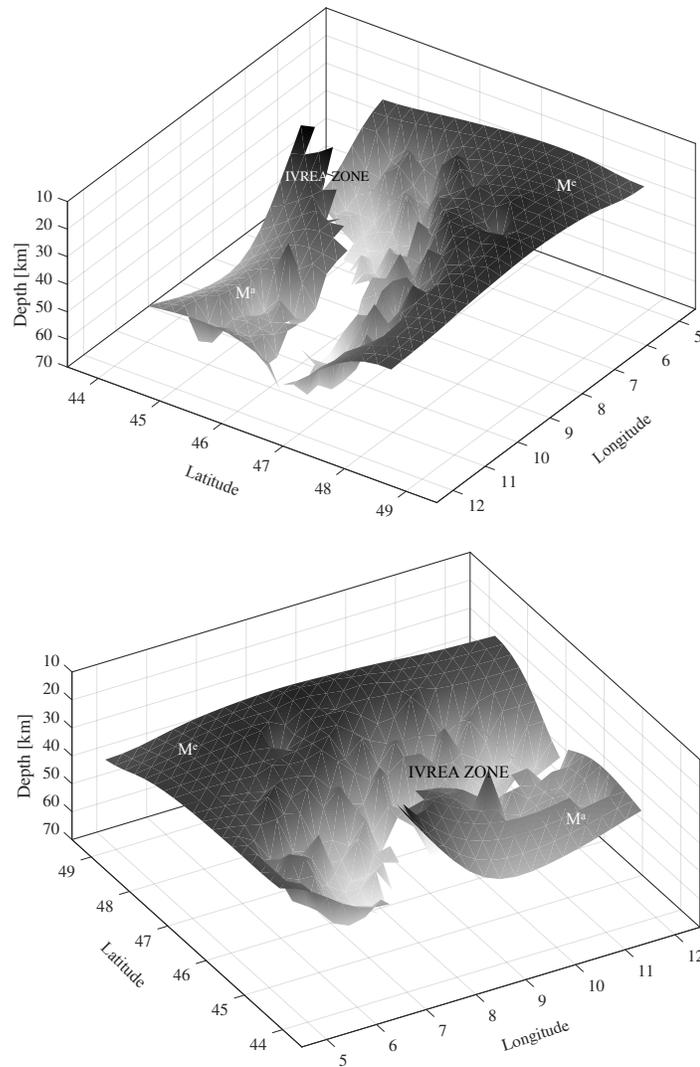


Figure 8.1.1: Perspective SW-oriented (top) and NE-oriented (bottom) view on the Moho depth model beneath the Central Alpine region obtained in this study. The European and Adriatic Moho are indicated with M^e and M^a .

In general, the obtained Alpine Moho topography reflects the present large-scale Alpine tectonic structure resulting from the collision of the African with the European Plate. The

two Moho interfaces at either side of the Insubric Line confirm a southward subduction of Europe under the shallower lying Adria in the Central Alps, with a decreasing depth difference from West to East in our study area. Our results highlight and confirm the unusually superficial crust-mantle interface beneath the Ivrea Verbano Zone, at around 15-km depth, as reported also by seismological and gravity investigations (Berckhemer et al., 1968; Kissling et al., 1983; Scarponi et al., 2020). In general Moho depths are in good agreement with previous studies for those regions with dense and reliable controlled-source data (e.g. Nadir, 1988; Valasek, 1992; Kissling, 1993; Hitz, 1995; Waldhauser et al., 1998; Spada et al., 2013).

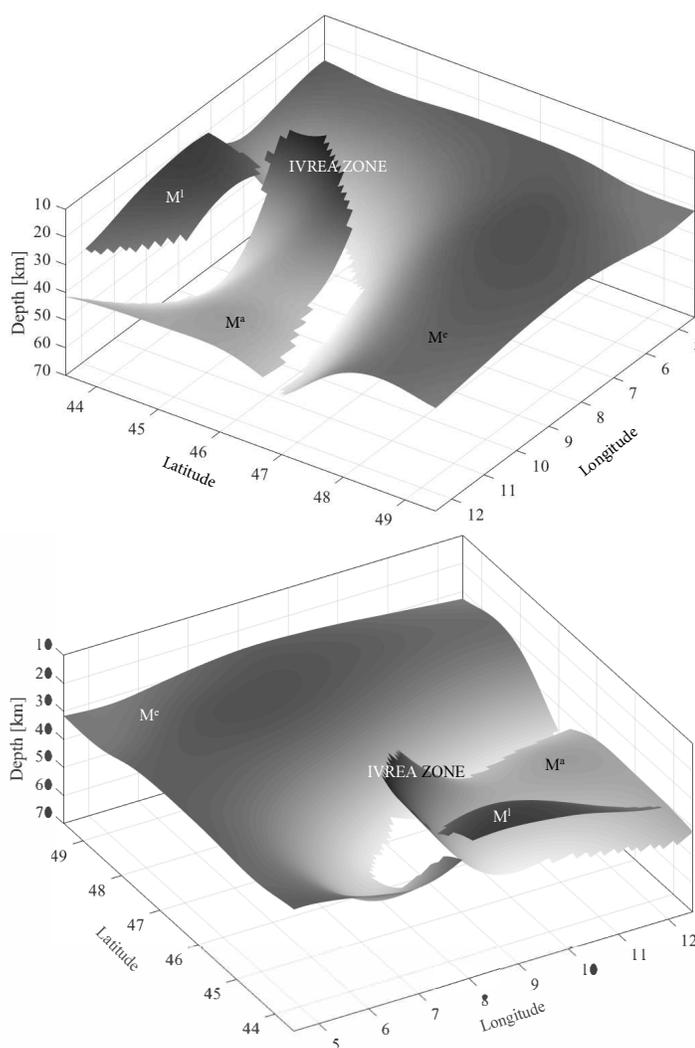


Figure 8.1.2: Perspective SW-oriented (top) and NE (bottom) on the Moho depth model below the Alpine region according to Spada et al. (2013), in a study combining receiver functions and controlled-source seismology data. European, Adriatic and Ligurian Moho are indicated with M^e , M^a and M^l .

When comparing the model of Spada et al., 2013 (Figure 8.1.2), which was the initial

model of our inversion, with our results, the Moho topography is rather similar at long wavelength, yet our results reveal some smaller wavelength variations. These could be realistic considering the complexity of Alpine tectonics, and that the Moho model of Figure 8.1.2 appears spatially more smoothed.

This is because in Spada et al. (2013) Moho, the interface is generated following the principle of simplicity; they search for the smoothest Moho interface that fits all available information within their uncertainties (Kissling, 1993) and then they define the simplicity of a seismic interface based on the degree of continuity and of surface roughness of this interface. To take into account for documented Moho offsets along plate boundaries, three main *a priori* surfaces are used to fit the data: interfaces described in Wagner et al. (2012) based on analysis of the LET (Diehl et al., 2009), CSS (Waldhauser et al., 1998) and teleseismic tomography information in the region (Lippitsch et al., 2003).

The real Moho geometry is most likely somewhere in between the locally irregular geometry proposed in this work (some of this irregularity coming from data coverage and *a priori* models) and the structure proposed by Spada et al. (2013) which appears very smooth on the regional level (this smoothness coming from interpolation between sparse seismic data). In this sense, an important constraint would be to include in our model some active seismic lines or dense RF profiles.

The new inversion method I developed yields a fully 3-D shear-wave velocity model of the crust based on receiver functions only. The resulting velocity profiles are comparable to those obtained by ambient noise tomography: while the new method inherently performs better in localizing sharp discontinuities, it is less well adapted to image bulk anomalies as tomography does. The maps of average crustal V_p/V_s ratio shows higher values in the Alps compared to the foreland. Within the European foreland, a zone of low V_p/V_s (<1.70) coincides with lower crustal earthquakes, which I interpret as inherited difference in rock properties.

The successful test of the new method on the Central Alps and using a wealth of high-quality data validated the applicability of the new method, which consists in the following main elements:

- An accurate ray propagator, which respects Snell's law in 3-D at any interface geometry, and allows P-to-S converted ray-paths to reach the recording station with an accuracy at the order of a hundred meters (Chapter 3).
- A new model parameterization of velocities: it is a fixed rectangular mesh in map view, but the interfaces have flexible depths; separate velocities are define above and below each discontinuity which allows to map both sharp discontinuities and gradient across layers (Chapter 4).

- A stochastic inversion procedure, combining simulated annealing and a pattern search algorithm, to find discontinuity depths and velocities across the crust by fitting grouped converted waves with synthetics (Chapter 5).

8.2 Uncertainties analysis and the method's applicability

The accuracy of the ray tracing implemented in this new tool is much higher than in most previous passive seismology studies. The reliability of the inversion is maximized by the quality and the number of rays passing through each cell. This is why I included a detailed quality control on the data, and introduced a threshold on the number of rays for each bundle to invert. In this regard, to save computation time keeping the robustness of the stochastic inversion method, it would be appropriate to launch the inversion with a significant random portion of the initial data (e.g. 50% of the dataset) and verify that the final results do not differ too much from those obtained with the entire dataset.

The uncertainty analysis of the results would ideally be based on the number of sampled models during inversion, and their assessment based on misfit and coverage of the parameter space. As I have shown in Chapter 5, the joint use of the simulated annealing and the pattern search algorithms in the MatLab implementation unfortunately precludes doing this. The reason is that while all sampled model parameter sets are available during the simulated annealing phase, no information can be retrieved during the pattern search phase and because the result from the last SA step to the final (SA+PS) step differ, using the SA steps alone would not be representative. Therefore, before a new technical implementation is performed, either in another software, or by rewriting the PS algorithm to have full control, I am not able to carry out a proper uncertainty analysis of the obtained results and this is probably the weakest point of the presented implementation. Another important limitation is that *a priori* information, studies with different resolution and significant uncertainties are mapped into the final S-wave velocity, which can create a null variation of the initial model or, worse, some artifacts, especially in areas where the ray coverage is not so high.

The newly developed tool is organized into different modules and is parameterized so that it can be adapted to be applied to other areas where a sufficient coverage of stations is available. The method does not require local earthquakes, nor a large aperture seismic network, but a dense array of 3-D component sensors with a spacing similar to the expected crustal thickness. In this study I showed a first application to the Central Alps; a natural extension would be to extend to the entire Alpine domain using AlpArray Seismic Network data (Figure 8.2.1).

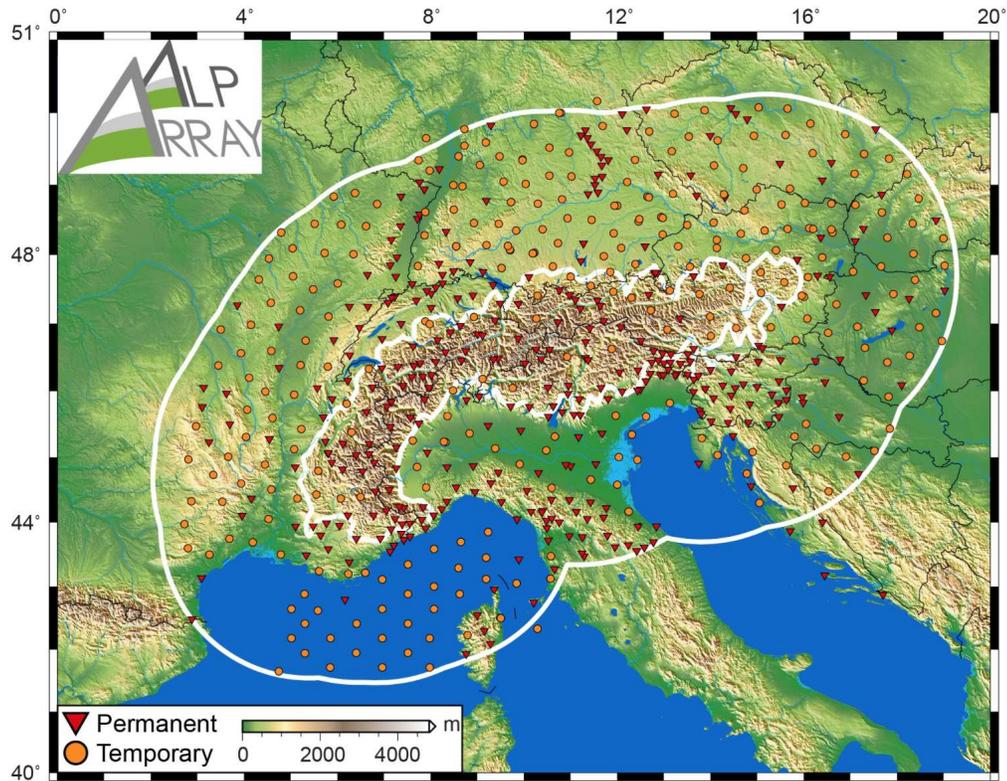


Figure 8.2.1: Distribution of seismic stations in the frame of the AlpArray project (from Hetényi et al., 2018b).

Regarding application to other geological contexts (Figure 8.2.2), the USArray network may be too sparse for Moho depth imaging, but the lithosphere-asthenosphere boundary could be mapped with S-to-P converted waves¹. The dense network of broadband seismographs makes Japan a suitable target to apply this new method, and other arrays exist (e.g. IberArray) and will be implemented in the near future (e.g. AdriaArray).

¹It is always preferable to constrain the initial models with available information related to the geological context of reference (e.g. active seismics, gravity studies).

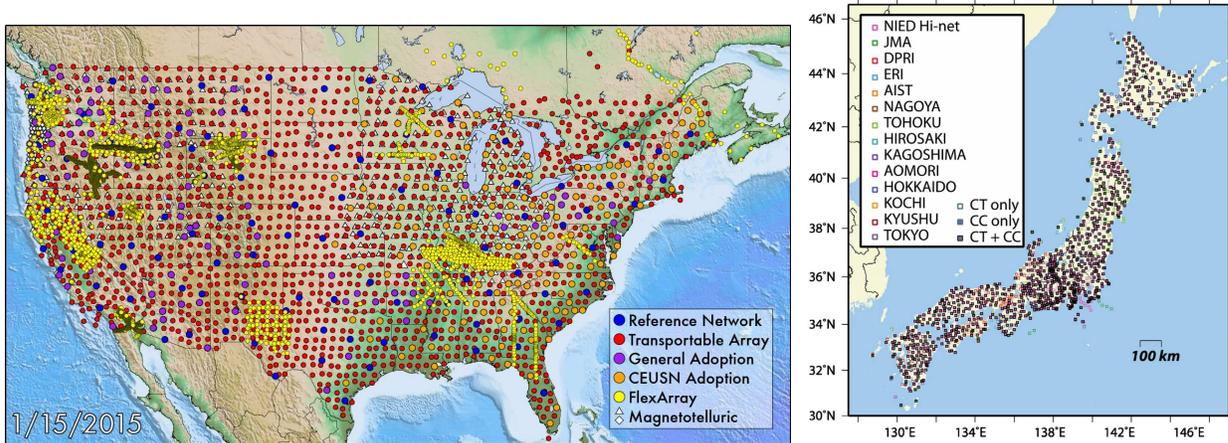


Figure 8.2.2: On the left status of USArray stations (image from Usarray.org), on the right distribution of seismic stations from Japan Meteorological Agency (from Yano et al., 2017).

8.3 Method limitations and future perspectives

The method proposed in this study presents different limitations, either by its nature or due to simplifying assumptions in the presented implementation to save computation time. While the latter can be addressed easily by more powerful computational power, the former could be improved by progressing with the methodology.

Regarding the ray tracing procedure, the accuracy of the developed tool considering the local dip is remarkable (accuracy of the order of hundred meters), however I focused on simulating the converted P-to-S wave raypath, and, inherently, not that of the multiples. This could be very useful for imaging, although not easy to implement as it already requires a high reliability of an initial 3-D model.

For the inversion, the main limiting assumption in the application to the Central Alps was the choice of the number of parameters to invert for. In this work, I showed the final results with 4 independent parameters: Conrad and Moho depth, V_p/V_s ratio for the entire crust, P-velocity jump at the Conrad. This can be easily improved for a more flexible parameterization of the model, and by performing the inversion with more than 4 parameters, up to 10, as discussed in section 6.1. Another assumption that could be considered as a limitation is the choice of the initial model, and whether that is a strong constraint or not. In our case, both the Moho depth and the P-wave velocity model were evolved inputs, but inversion could be launched in a much less constrained case. For both of these aspects, the corresponding increase in the computation time needs to be accounted for, either with time, or with parallel programming.

A similar limiting assumption adopted in this work was to assume a model structure consisting of only 2 layers, upper crust and lower crust, meaning 2 discontinuities. A more realistic choice could be to add a sedimentary layer, which would have helped to produce better results in the basins adjacent to the Alps (e.g., Po plain, Molasse Basin). If one considers mapping the entire lithosphere, allowing for further discontinuities would be useful. In the last two decades a class of algorithm using transdimensional Bayesian formulation has developed and appears in several problems related to RF inversion (Piana Agostinetti and Malinverno, 2010; Bodin et al., 2012; Sambridge et al., 2006; Young et al., 2013; Sambridge et al., 2013; Kolb and Lekić, 2014). With this kind of approach, the number of model parameters is a variable and a Markov chain Monte Carlo is used to provide a solution. Therefore, in a future development of our method, one might decide not to fix *a priori* the number of layers so the dimension of the model space becomes itself a variable. This certainly requires high computational power, especially that each layer has 2 velocities defined at each node.

Another choice that can effect the implementation of the method is the shape of the cells used in the model parameterization. In this work I used square meshes in map view. Other possible shapes filling the space can be triangular (Li et al., 2014) or hexagonal meshes. Furthermore, one could consider a variable cell size mesh, adapting to data density, as for example in recent tomography studies (Schaefer et al., 2011; Lu et al., 2018). While these choices should not drastically influence the results, they may allow a better recovery or representation of reality in a model. For mantle convection simulations, a new kind of overset grid, named Yin-Yang grid, for spherical geometry is proposed Kageyama and Sato, 2004: these grids, composed of two identical component, are combined in a complementary way to cover a spherical surface with partial overlap on their boundaries. This type of grids require the cost of parallel computers.

A clear improvement to our method would be to use and implement in 3-D the stochastic Neighbourhood Algorithm (Sambridge, 1999a; Sambridge, 1999b) for the inversion procedure. This approach was specifically built for sampling in geophysical inverse problems, and in this sense could likely be better than the combination of simulated annealing algorithm and pattern search used in this work. The computation time should be on the same order of magnitude (see Figure 5.1.5). The main advantage of implementing the Neighbourhood Algorithm would be to access the entire series of tested parameter sets, and therefore to quantitatively access the model results' uncertainties.

Finally, the main future target to better constrain the 3-D structure of the lithosphere is to improve the method proposed in this study by integrating other geophysical data and hence move towards joint inversion approaches, since a study based only on RF itself may

not be sufficiently reliable. The main goal is to constraint multiple physical properties of rocks, to obtain a geophysical model consistent with multiple datasets. Our method exploits sub-vertically propagating S waves in the crust to retrieve Vs. It could be used together with surface wave dispersion or ambient noise tomography for still inverting Vs but with waves of different geometry and sensitivity (Julia et al., 2000; Chang et al., 2004). To include other rock properties, such as Vp and density, jointly inverting gravity and seismic tomography data should be implemented (Tiberi et al., 2003; Basuyau and Tiberi, 2011). The advantage is that gravity and seismic data are complementary in terms of their best resolution areas and sensitivities. The relationship between rock velocities and rock density can be assumed based on empirical relations (Nafe and Drake, 1957; Birch, 1961), or can be part of the inversion problem itself.

As a main conclusion: in the application of this new method, as with any other method, some simplifying assumptions are usually taken at start, nevertheless the capacities should be explored for imaging in fully in 3-D, as this is the nature of our environment and therefore a key element in all branches of solid Earth research.

Bibliography

- Abe, Y., Ohkura, T., Hirahara, K., and Shibutani, T. (2011). Common-conversion-point stacking of receiver functions for estimating the geometry of dipping interfaces. *Geophysical Journal International*, 185(3):1305–1311.
- Abt, D. L., Fischer, K. M., French, S. W., Ford, H. A., Yuan, H., and Romanowicz, B. (2010). North american lithospheric discontinuity structure imaged by Ps and Sp receiver functions. *Journal of Geophysical Research: Solid Earth*, 115(B9).
- Afonso, J. C. and Ranalli, G. (2004). Crustal and mantle strengths in continental lithosphere: is the jelly sandwich model obsolete? *Tectonophysics*, 394(3-4):221–232.
- 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. R. and Richards, P. G. (1980). *Quantitative Seismology: Theory and Methods*. Freeman.
- Ammon, C. (1997). Isolating the receiver response Langston’s source equalization procedure.
- Ammon, C. J., Randall, G. E., and Zandt, G. (1990). On the nonuniqueness of receiver function inversions. *Journal of Geophysical Research: Solid Earth*, 95(B10):15303–15318.
- Ammon, C. J., Velasco, A. A., and Lay, T. (2010). An introduction to earthquakes. In *Seismic Waves and Earth Interior, Saint Louis University, EAS-A193 Class Note*.
- Anstey, N. A. (1976). Three-dimensional seismic display. US Patent 3,931,609.
- Backus, G. E. and Gilbert, J. (1967). Numerical applications of a formalism for geophysical inverse problems. *Geophysical Journal International*, 13(1-3):247–276.
- Baker, G. E., Minster, J. B., Zandt, G., and Gurrola, H. (1996). Constraints on Crustal Structure and Complex Moho Topography beneath Piñon Flat, California, from Teleseismic Receiver Functions. *Bulletin of the Seismological Society of America*, 86(6):1830–1844.

- Barringer, D. (1892). Apparatus for illustrating geological formations. US Patent 477,633.
- Barton, P. (1986). The relationship between seismic velocity and density in the continental crust - A useful constraint? *Geophysical Journal International*, 87(1):195–208.
- Basuyau, C. and Tiberi, C. (2011). Imaging lithospheric interfaces and 3D structures using receiver functions, gravity, and tomography in a common inversion scheme. *Computers & geosciences*, 37(9):1381–1390.
- Beck, S. L. and Zandt, G. (2002). The nature of orogenic crust in the central Andes. *Journal of Geophysical Research: Solid Earth*, 107(B10):ESE–7.
- Berckhemer, H. et al. (1968). Topographie des Ivrea-Körpers abgeleitet aus seismischen und gravimetrischen daten. *Schweiz. Mineral. Petrogr. Mitt.*, 48(1):235–246.
- Bertotti, G., Picotti, V., Bernoulli, D., and Castellarin, A. (1993). From rifting to drifting: tectonic evolution of the South-Alpine upper crust from the Triassic to the Early Cretaceous. *Sedimentary Geology*, 86(1-2):53–76.
- Bianchi, I. and Bokelmann, G. (2014). Seismic signature of the Alpine indentation, evidence from the Eastern alps. *Journal of geodynamics*, 82:69–77.
- Biggs, N. (1986). The traveling salesman problem - a guided tour of combinatorial optimization. *Bulletin of the London Mathematical Society*, 18(5):514–515.
- Birch, F. (1960). The velocity of compressional waves in rocks to 10 kilobars: 1. *Journal of Geophysical Research*, 65(4):1083–1102.
- Birch, F. (1961). The velocity of compressional waves in rocks to 10 kilobars: 2. *Journal of Geophysical Research*, 66(7):2199–2224.
- Blundell, D. J., Freeman, R., Mueller, S., and Button, S. (1992). *A continent revealed: The European Geotraverse, structure and dynamic evolution*. Cambridge University Press.
- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research: Solid Earth*, 117(B2).
- Bousquet, R., Oberhansli, R., Schmid, S., Berger, A., Wiederkeher, M., Robert, C., Moller, A., Rosenberg, C., Zeilinger, G., Molli, G., et al. (2012a). *Metamorphic framework of the Alps- Carte metamorphique des Alpes CCGM/CGMW*. CCGM/CGMW.

- Bousquet, R., Schmid, S., Zeilinger, G., Oberhänsli, R., Rosenberg, C., Molli, G., Robert, C., Wiederkehr, M., and Rossi, P. (2012b). Tectonic framework of the Alps, CCGM/CGMW.
- Brocher, T. M. (2005). Empirical relations between elastic wavespeeds and density in the Earth's crust. *Bulletin of the seismological Society of America*, 95(6):2081–2092.
- Brückl, E. and Bleibinhaus, F. (2003). The ALP2002 refraction experiment and its relation to TRANSALP. In *Transalp conference. Memorie di Scienze Geologiche*, volume 54, pages 239–242.
- Burdick, L. and Helmberger, D. V. (1978). The upper mantle P velocity structure of the Western United States. *Journal of Geophysical Research: Solid Earth*, 83(B4):1699–1712.
- Cadman, W. (1927). The Golden Lanes of Greenwood County, Kansas. *AAPG Bulletin*, 11(11):1151–1172.
- Campillo, M. and Paul, A. (2003). Long-range correlations in the diffuse seismic coda. *Science*, 299(5606):547–549.
- Cassidy, J. (1992). Numerical experiments in broadband receiver function analysis. *Bulletin of the Seismological Society of America*, 82(3):1453–1474.
- Castellarin, A., Cantelli, L., Fesce, A., Mercier, J., Picotti, V., Pini, G., Prosser, G., and Selli, L. (1992). Alpine compressional tectonics in the Southern Alps. relationships with the N-Apennines. In *Annales tectonicae*, volume 6, pages 62–94.
- Cerveny, V. (2005). *Seismic ray theory*. Cambridge university press.
- Chang, S.-J., Baag, C.-E., and Langston, C. A. (2004). Joint analysis of teleseismic receiver functions and surface wave dispersion using the genetic algorithm. *Bulletin of the Seismological Society of America*, 94(2):691–704.
- Chelle-Michou, C., Do Couto, D., Moscarriello, A., Renard, P., and Rusillon, E. (2017). Geothermal state of the deep Western Alpine Molasse basin, France-Switzerland. *Geothermics*, 67:48–65.
- Cheng, N. and House, L. (1996). Minimum travelttime calculation in 3-D graph theory. *Geophysics*, 61(6):1895–1898.
- Christensen, N. I. (1966). Elasticity of ultrabasic rocks. *Journal of Geophysical Research*, 71(24):5921–5931.

- Christensen, N. I. (1996). Poisson's ratio and crustal seismology. *Journal of Geophysical Research: Solid Earth*, 101(B2):3139–3156.
- Christensen, N. I. and Mooney, W. D. (1995). Seismic velocity structure and composition of the continental crust: A global view. *Journal of Geophysical Research: Solid Earth*, 100(B6):9761–9788.
- Christensen, U. R. and Yuen, D. A. (1985). Layered convection induced by phase transitions. *Journal of Geophysical Research: Solid Earth*, 90(B12):10291–10300.
- Claesen, M. and De Moor, B. (2015). Hyperparameter search in machine learning. *arXiv preprint arXiv:1502.02127*.
- Clayton, R. W. and Wiggins, R. A. (1976). Source shape estimation and deconvolution of teleseismic bodywaves. *Geophysical Journal International*, 47(1):151–177.
- Closs, H. (1965). Explosion seismic studies in the Alps and in the German Federal Republic. In *The Upper Mantle Project Symposium, New Delhi*, volume 1964, pages 94–102.
- Cordell, L. and Henderson, R. G. (1968). Iterative three-dimensional solution of gravity anomaly data using a digital computer. *Geophysics*, 33(4):596–601.
- Coron, S., Closs, H., and Labrouste, Y. (1963). Aperçu gravimétrique sur les Alpes Occidentales. *Annee Geophysique Internationale, CNRS, Série*, 13(2):31–37.
- Crampin, S. (1977). A review of the effects of anisotropic layering on the propagation of seismic waves. *Geophysical Journal International*, 49(1):9–27.
- Dal Piaz, G. (1999). The Austroalpine-Piedmont nappe stack and the puzzle of Alpine Tethys. *Memorie di Scienze Geologiche*, 51(1):155–176.
- Dal Piaz, G. V., Bistacchi, A., and Massironi, M. (2003). Geological outline of the Alps. *Episodes*, 26(3):175–180.
- Deichman, N. and Baer, M. (1990). Earthquake focal depths below the Alps and northern Alpine foreland of Switzerland. In *The European Geotraverse: Integrative studies. Results from the earth science study centre*. 5, pages 277–288.
- Deichmann, N. (1992). Structural and rheological implications of lower-crustal earthquakes below northern Switzerland. *Physics of the Earth and Planetary Interiors*, 69(3-4):270–280.

- Deichmann, N. and Rybach, L. (1989). Earthquakes and temperatures in the lower crust below the northern Alpine foreland of Switzerland. *Properties and processes of Earth's lower crust*, 51:197–213.
- Devoti, R., Ferraro, C., Gueguen, E., Lanotte, R., Luceri, V., Nardi, A., Pacione, R., Rutigliano, P., Sciarretta, C., and Vespe, F. (2002). Geodetic control on recent tectonic movements in the central Mediterranean area. *Tectonophysics*, 346(3-4):151–167.
- Dewey, J., Helman, M., Knott, S., Turco, E., and Hutton, D. (1989). Kinematics of the western Mediterranean. *Geological Society, London, Special Publications*, 45(1):265–283.
- Di Stefano, R., Bianchi, I., Ciaccio, M., Carrara, G., and Kissling, E. (2011). Three-dimensional Moho topography in Italy: New constraints from receiver functions and Controlled Source Seismology. *Geochemistry, Geophysics, Geosystems*, 12(9).
- Diehl, T., Husen, S., Kissling, E., and Deichmann, N. (2009). High-resolution 3-D P-wave model of the Alpine crust. *Geophysical Journal International*, 179(2):1133–1147.
- Dvorkin, J., Mavko, G., and Nur, A. (1999). Overpressure detection from compressional-and shear-wave data. *Geophysical Research Letters*, 26(22):3417–3420.
- Ebbing, J., Braitenberg, C., and Götze, H.-J. (2001). Forward and inverse modelling of gravity revealing insight into crustal structures of the Eastern Alps. *Tectonophysics*, 337(3-4):191–208.
- Eckhardt, C. and Rabbel, W. (2011). P-receiver functions of anisotropic continental crust: a hierarchic catalogue of crustal models and azimuthal waveform patterns. *Geophysical Journal International*, 187(1):439–479.
- Escher, A. and Beaumont, C. (1997). Formation, burial and exhumation of basement nappes at crustal scale: a geometric model based on the Western Swiss-Italian Alps. *Journal of Structural geology*, 19(7):955–974.
- Escher, A., Hunziker, J., Marthaler, M., Masson, H., Sartori, M., and Steck, A. (1997). Geological framework and structural evolution of the Western Swiss-Italian Alps. *Deep Structure of the Swiss Alps: Results of the National Research Program 20 (NRP 20)*, pages 205–222.
- Festa, A., Balestro, G., Borghi, A., De Caroli, S., and Succo, A. (2019). The role of structural inheritance in continental break-up and exhumation of Alpine Tethyan mantle (Canavese Zone, Western Alps). *Geoscience Frontiers*.

- Finetti, I. R. (2005). *CROP project: deep seismic exploration of the central Mediterranean and Italy*. Elsevier.
- Fischer, A. (1992). A special newton-type optimization method. *Optimization*, 24(3-4):269–284.
- Fischer, K. M. (2002). Waning buoyancy in the crustal roots of old mountains. *Nature*, 417(6892):933–936.
- Frederiksen, A. and Bostock, M. (2000). Modelling teleseismic waves in dipping anisotropic structures. *Geophysical Journal International*, 141(2):401–412.
- Fyfe, W. S. (2012). *Fluids in the Earth's crust: Their significance in metamorphic, tectonic and chemical transport process*, volume 1. Elsevier.
- Galson, D. and Mueller, S. (1986). An introduction to the European Geotraverse Project: first results and present plans. *Tectonophysics*, 126(1):1–30.
- Gebrande, H., Lüschen, E., Bopp, M., Bleibinhaus, F., Lammerer, B., Oncken, O., Stiller, M., Kummerow, J., Kind, R., et al. (2002). First deep seismic reflection images of the Eastern Alps reveal giant crustal wedges and transcrustal ramps. *Geophysical Research Letters*, 29(10):92–1.
- Gebrande, H., Lüschen, E., Lammerer, B., Oncken, O., Stiller, M., Neubauer, F., Millahn, K., Grassl, H., Bertelli, L., Angeleri, G., et al. (2001). European orogenic processes research transects the Eastern Alps. *Eos, Transactions American Geophysical Union*, 82(40):453–461.
- Geman, S. and Geman, D. (1987). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In *Readings in computer vision*, pages 564–584. Elsevier.
- Gill, M. and Murray, W. (1981). Wright. practical optimization.
- Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & operations research*, 13(5):533–549.
- Goldberg, D. (1989). Genetic algorithms in search, optimization, and Machine Learning. *NN Schraudolph and J*, 3(1).
- Granser, H., Meurers, B., and Steinhauser, P. (1989). Apparent density mapping and 3D gravity inversion in the Eastern Alps 1. *Geophysical Prospecting*, 37(3):279–292.

- Guidarelli, M., Aoudia, A., and Costa, G. (2017). 3-D structure of the crust and uppermost mantle at the junction between the Southeastern Alps and External Dinarides from Ambient Noise Tomography. *Geophysical Journal International*, 211(3):1509–1523.
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, pages 49–52.
- Hart, P. J. (1969). *The Earth's crust and upper mantle: structure, dynamic processes, and their relation to deep-seated geological phenomena*. Number 13. American Geophysical Union.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications.
- Hetényi, G. (2007). *Evolution of deformation of the Himalayan prism: from Imaging to Modelling*. PhD thesis, Université Paris Sud-Paris XI.
- Hetényi, G., Cattin, R., Vergne, J., and Nábělek, J. L. (2006). The effective elastic thickness of the India Plate from receiver function imaging, gravity anomalies and thermomechanical modelling. *Geophysical Journal International*, 167(3):1106–1118.
- Hetényi, G., Epard, J.-L., Colavitti, L., Hirzel, A. H., Kiss, D., Petri, B., Scarponi, M., Schmalholz, S. M., and Subedi, S. (2018a). Spatial relation of surface faults and crustal seismicity: a first comparison in the region of Switzerland. *Acta Geodaetica et Geophysica*, 53(3):439–461.
- Hetényi, G., Molinari, I., Clinton, J., Bokelmann, G., Bondár, I., Crawford, W. C., Dessa, J.-X., Doubre, C., Friederich, W., Fuchs, F., et al. (2018b). The AlpArray seismic network: a large-scale European experiment to image the Alpine Orogen. *Surveys in geophysics*, 39(5):1009–1033.
- Hitz, L. (1995). The 3D crustal structure of the Alps of eastern Switzerland and western Austria interpreted from a network of deep-seismic profiles. *Tectonophysics*, 248(1-2):71–96.
- Hölder, H. and Lotze, F. (1964). *Handbuch der Stratigraphischen Geologie*. Enke.
- Hooke, R. and Jeeves, T. (1962). *Jacm. 8, 212 (1961)*. *Google Scholar OG Ludwig, Ph. D. thesis, Carnegie Institute of Technology, Pittsburgh, Pa.*
- Igel, H. (2017). *Computational seismology: a practical introduction*. Oxford University Press.

- Iyer, H. and Hirahara, K. (1993). *Seismic tomography: Theory and practice*. Springer Science & Business Media.
- Julia, J., Ammon, C. J., Herrmann, R., and Correig, A. M. (2000). Joint inversion of receiver function and surface wave dispersion observations. *Geophysical Journal International*, 143(1):99–112.
- Julian, B. (1977). Three-dimensional seismic ray tracing. *Journal of Geophysics*, 43:95–113.
- Kageyama, A. and Sato, T. (2004). Yin-Yang grid: An overset grid in spherical geometry. *Geochemistry, Geophysics, Geosystems*, 5(9).
- Kahle, H.-G., Klingele, E., Egloff, R., et al. (1976). The variation of crustal thickness across the Swiss Alps based on gravity and explosion seismic data. *Pure and Applied geophysics*, 114(3):479–494.
- Kahle, H.-G. and Werner, D. (1980). A geophysical study of the Rhinegraben. II. Gravity anomalies and geothermal implications. *Geophysical Journal International*, 62(3):631–647.
- Kalmár, D., Hetényi, G., Süle, B., and Bondár, I. (2018). Receiver function analysis in the Eastern Alps–Pannonian Basin transition zone. In *EGU General Assembly Conference Abstracts*, volume 20, page 464.
- Kennett, B. and Engdahl, E. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2):429–465.
- Kern, H. (1982). Elastic-wave velocity in crustal and mantle rocks at high pressure and temperature: the role of the high-low quartz transition and of dehydration reactions. *Physics of the Earth and Planetary Interiors*, 29(1):12–23.
- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.
- Kissling, E. (1980). *Krustenaufbau und Isostasie in der Schweiz*. PhD thesis, ETH Zurich.
- Kissling, E. (1993). Deep structure of the Alps - what do we really know? *Physics of the Earth and Planetary Interiors*, 79(1-2):87–112.
- Kissling, E., Husen, S., and Haslinger, F. (2001). Model parametrization in seismic tomography: a choice of consequence for the solution quality. *Physics of the Earth and Planetary Interiors*, 123(2-4):89–101.

- Kissling, E., Mueller, S., and Werner, D. (1983). Gravity anomalies, seismic structure and geothermal history of the Central Alps. In *Annales de Geophysique*, volume 1, pages 37–46.
- Kissling, E., Schmid, S. M., Lippitsch, R., Ansorge, J., and Fügenschuh, B. (2006). Lithosphere structure and tectonic evolution of the Alpine arc: new evidence from high-resolution teleseismic tomography. *Geological Society, London, Memoirs*, 32(1):129–145.
- Knapmeyer, M. (2004). TTBox: A MatLab toolbox for the computation of 1D teleseismic travel times. *Seismological Research Letters*, 75(6):726–733.
- Knapmeyer, M. (2005). Numerical accuracy of travel-time software in comparison with analytic results. *Seismological Research Letters*, 76(1):74–81.
- Kolb, J. and Lekić, V. (2014). Receiver function deconvolution using transdimensional hierarchical Bayesian inference. *Geophysical Journal International*, 197(3):1719–1735.
- Kumar, P., Kind, R., and Yuan, X. (2010). Receiver function summation without deconvolution. *Geophysical Journal International*, 180(3):1223–1230.
- Kummerow, J., Kind, R., Oncken, O., Giese, P., Ryberg, T., Wylegalla, K., Scherbaum, F., Group, T. W., et al. (2004). A natural and controlled source seismic profile through the Eastern Alps: TRANSALP. *Earth and Planetary Science Letters*, 225(1-2):115–129.
- Kummerow, J., Reshetnikov, A., Häring, M., and Asanuma, H. (2012). 3-D V_p/V_s Ratio Distribution in the Geothermal Reservoir at Basel, Switzerland, from Microseismic Data. In *AGU Fall Meeting Abstracts*.
- Labhart, T. (2005). *Geologie der schweiz*, 210 pp.
- Langston, C. A. (1977a). Corvallis, Oregon, crustal and upper mantle receiver structure from teleseismic P and S waves. *Bulletin of the Seismological Society of America*, 67(3):713–724.
- Langston, C. A. (1977b). The effect of planar dipping structure on source and receiver responses for constant ray parameter. *Bulletin of the Seismological Society of America*, 67(4):1029–1050.
- Langston, C. A. (1979). Structure under Mount Rainier, Washington, inferred from teleseismic body waves. *Journal of Geophysical Research: Solid Earth*, 84(B9):4749–4762.
- Le Pichon, X., Bergerat, F., and Roulet, M.-J. (1988). Plate kinematics and tectonics leading to the Alpine belt formation; a new analysis. *Geological Society of America Special Papers*, 218:111–132.

- Levin, V. and Park, J. (1997). P-SH conversions in a flat-layered medium with anisotropy of arbitrary orientation. *Geophysical Journal International*, 131(2):253–266.
- Levin, V. and Park, J. (1998). P-sh conversions in layered media with hexagonally symmetric anisotropy: a cookbook. In *Geodynamics of Lithosphere & Earth's Mantle*, pages 669–697. Springer.
- Lewis, R. M. and Torczon, V. (1999). Pattern search algorithms for bound constrained minimization. *SIAM Journal on Optimization*, 9(4):1082–1099.
- Li, F., Xu, T., Zhang, M., Wu, Z., Wu, C., Zhang, Z., and Teng, J. (2014). Seismic traveltime inversion of 3D velocity model with triangulated interfaces. *Earthquake Science*, 27(2):127–136.
- Li, H., Bernardi, F., and Michelini, A. (2010). Surface wave dispersion measurements from ambient seismic noise analysis in Italy. *Geophysical Journal International*, 180(3):1242–1252.
- Li, J., Tian, B., Wang, W., Zhao, L., and Yao, Z. (2007). Lateral Variation in the Sedimentary Structure of West Bohai Bay Basin Inferred from P-Multiple Receiver Functions. *Bulletin of the Seismological Society of America*, 97(4):1355–1363.
- Licciardi, A., Agostinetti, N. P., Lebedev, S., Schaeffer, A., Readman, P., and Horan, C. (2014). Moho depth and V_p/V_s in Ireland from teleseismic receiver functions analysis. *Geophysical Journal International*, 199(1):561–579.
- Ligorria, J. P. and Ammon, C. J. (1999). Iterative deconvolution and receiver-function estimation. *Bulletin of the seismological Society of America*, 89(5):1395–1400.
- Lippitsch, R., Kissling, E., and Ansorge, J. (2003). Upper mantle structure beneath the Alpine orogen from high-resolution teleseismic tomography. *Journal of Geophysical Research: Solid Earth*, 108(B8).
- Liu, Y. and Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: Theory. *Journal of optimization theory and applications*, 69(1):129–137.
- Lombardi, D., Braunmiller, J., Kissling, E., and Giardini, D. (2008). Moho depth and Poisson's ratio in the Western-Central Alps from receiver functions. *Geophysical Journal International*, 173(1):249–264.

- Lu, Y., Stehly, L., Paul, A., and Group, A. W. (2018). High-resolution surface wave tomography of the European crust and uppermost mantle from ambient seismic noise. *Geophysical Journal International*, 214(2):1136–1150.
- Lucente, F. P., Piana Agostinetti, N., Moro, M., Selvaggi, G., and Di Bona, M. (2005). Possible fault plane in a seismic gap area of the southern Apennines (Italy) revealed by receiver function analysis. *Journal of Geophysical Research: Solid Earth*, 110(B4).
- Lockett, R. and Baptie, B. (2015). Local earthquake tomography of Scotland. *Geophysical Journal International*, 200(3):1538–1554.
- Lüschen, E., Lammerer, B., Gebrande, H., Millahn, K., Nicolich, R., Group, T. W., et al. (2004). Orogenic structure of the Eastern Alps, Europe, from TRANSALP deep seismic reflection profiling. *Tectonophysics*, 388(1-4):85–102.
- Manatschal, G., Lavier, L., and Chenin, P. (2015). The role of inheritance in structuring hyperextended rift systems: Some considerations based on observations and numerical modeling. *Gondwana Research*, 27(1):140–164.
- Mavko, G. and Vanorio, T. (2010). The influence of pore fluids and frequency on apparent effective stress behavior of seismic velocities. *Geophysics*, 75(1):N1–N7.
- McLeish, A. (1992). *Geological science*. Nelson Thornes.
- Menke, W. (2018). *Geophysical data analysis: Discrete inverse theory*. Academic press.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.
- Metropolis, N. and Ulam, S. (1949). The Monte Carlo method. *Journal of the American statistical association*, 44(247):335–341.
- Michard, A., Goffé, B., Chopin, C., and Henry, C. (1996). Did the Western Alps develop through an Oman-type stage? *Eclogae Geologicae Helvetiae*, 89(1):43–80.
- Miller, H. et al. (1976). A lithospheric seismic profile along the axis of the Alps. *Pure and Applied geophysics*, 114(6):1109–1130.
- Molinari, I., Verbeke, J., Boschi, L., Kissling, E., and Morelli, A. (2015). Italian and Alpine three-dimensional crustal structure imaged by ambient-noise surface-wave dispersion. *Geochemistry, Geophysics, Geosystems*, 16(12):4405–4421.

- Möller, T. and Trumbore, B. (2005). Fast, minimum storage ray/triangle intersection. In *ACM SIGGRAPH 2005 Courses*, page 7. ACM.
- Molli, G., Crispini, L., Malusà, M., Mosca, P., Piana, F., and Federico, L. (2010). Geology of the Western Alps-Northern Apennine junction area: a regional review. *Eds Marco Beltrando, Angelo Peccerillo, Massimo Mattei, Sandro Conticelli, and Carlo Doglioni Journal of the Virtual Explorer*, 36(3).
- Moorkamp, M., Lelièvre, P. G., Linde, N., and Khan, A. (2016). *Integrated imaging of the Earth: Theory and applications*, volume 218. John Wiley & Sons.
- Moser, T. (1991). Shortest path calculation of seismic rays. *Geophysics*, 56(1):59–67.
- Nadir, S. (1988). *Structure de la croûte continentale entre les Alpes occidentales et les Alpes Ligures et ondes S dans la croûte continentale à l'ouest du Bassin de Paris*. PhD thesis, Paris 7.
- Nafe, J. E. and Drake, C. L. (1957). Variation with depth in shallow and deep water marine sediments of porosity, density and the velocities of compressional and shear waves. *Geophysics*, 22(3):523–552.
- Nair, S. K., Gao, S. S., Liu, K. H., and Silver, P. G. (2006). Southern African crustal evolution and composition: Constraints from receiver function studies. *Journal of Geophysical Research: Solid Earth*, 111(B2).
- Nakanishi, I. and Yamaguchi, K. (1986). A numerical experiment on nonlinear image reconstruction from first-arrival times for two-dimensional island arc structure. *Journal of Physics of the Earth*, 34(2):195–201.
- Neubauer, F. (2000). *Aspects of geology in Austria: printed on behalf of the Austrian National Committee of Geology and financially supported by the Austrian Federal Ministry of Education, Science and Culture*. Österr. Geologische Ges.
- Neumaier, A. (2004). Complete search in continuous global optimization and constraint satisfaction. *Acta numerica*, 13:271–369.
- Nicolas, A., Hirn, A., Nicolich, R., and Polino, R. (1990). Lithospheric wedging in the Western Alps inferred from the ECORS-CROP traverse. *Geology*, 18(7):587–590.
- Nolet, G. (2012). *Seismic tomography: with applications in global seismology and exploration geophysics*, volume 5. Springer Science & Business Media.

- Owens, T. J. and Crosson, R. S. (1988). Shallow structure effects on broadband teleseismic P waveforms. *Bulletin of the Seismological Society of America*, 78(1):96–108.
- Ozacar, A. A. and Zandt, G. (2004). Crustal seismic anisotropy in central Tibet: Implications for deformational style and flow in the crust. *Geophysical Research Letters*, 31(23).
- Park, J. and Levin, V. (2000). Receiver functions from multiple-taper spectral correlation estimates. *Bulletin of the Seismological Society of America*, 90(6):1507–1520.
- Parker, R. L. (1994). *Geophysical inverse theory*, volume 1. Princeton university press.
- Paul, A., Cattaneo, M., Thouvenot, F., Spallarossa, D., Béthoux, N., and Fréchet, J. (2001). A three-dimensional crustal velocity model of the southwestern Alps from local earthquake tomography. *Journal of Geophysical Research: Solid Earth*, 106(B9):19367–19389.
- Peng, X. and Humphreys, E. D. (1997). Moho dip and crustal anisotropy in northwestern Nevada from teleseismic receiver functions. *Bulletin of the Seismological Society of America*, 87(3):745–754.
- Pfiffner, O. A. (2014). *Geology of the Alps*. John Wiley & Sons.
- Pfiffner, O.-A., Lehner, P., Heitzmann, P., Mueller, S., and Steck, A. (1997). *Deep structure of the Swiss Alps: results of NRP 20*. Birkhäuser.
- Piana Agostinetti, N. and Amato, A. (2009). Moho depth and Vp/Vs ratio in peninsular Italy from teleseismic receiver functions. *Journal of Geophysical Research: Solid Earth*, 114(B6).
- Piana Agostinetti, N. and Malinverno, A. (2010). Receiver function inversion by trans-dimensional Monte Carlo sampling. *Geophysical Journal International*, 181(2):858–872.
- Piqué, A. and Laville, E. (1996). The central Atlantic rifting: Reactivation of Palaeozoic structures? *Journal of Geodynamics*, 21(3):235–255.
- Plöschinger, B. (1995). Tectonics of the Northern Calcareous Alps: A review. *Mem Sci Geol Padova*, 47:73–86.
- Popovici, A. M. and Sethian, J. A. (2002). 3-D imaging using higher order fast marching traveltimes. *Geophysics*, 67(2):604–609.
- Powell, M. J. (1973). On search directions for minimization algorithms. *Mathematical programming*, 4(1):193–201.

- Prodehl, C. and Mooney, W. D. (2012). *Exploring the Earth's crust: history and results of controlled-source seismology*, volume 208. Geological Society of America.
- Rawlinson, N., Hauser, J., and Sambridge, M. (2008). Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Advances in Geophysics*, 49:203–273.
- Rawlinson, N., Kennett, B. L., and Heintz, M. (2006a). Insights into the structure of the upper mantle beneath the murray basin from 3D teleseismic tomography. *Australian Journal of Earth Sciences*, 53(4):595–604.
- Rawlinson, N., Pozgay, S., and Fishwick, S. (2010). Seismic tomography: a window into deep Earth. *Physics of the Earth and Planetary Interiors*, 178(3-4):101–135.
- Rawlinson, N., Reading, A. M., and Kennett, B. L. (2006b). Lithospheric structure of Tasmania from a novel form of teleseismic tomography. *Journal of Geophysical Research: Solid Earth*, 111(B2).
- Rawlinson, N. and Sambridge, M. (2004a). Multiple reflection and transmission phases in complex layered media using a multistage fast marching method. *Geophysics*, 69(5):1338–1350.
- Rawlinson, N. and Sambridge, M. (2004b). Wave front evolution in strongly heterogeneous layered media using the fast marching method. *Geophysical Journal International*, 156(3):631–647.
- Reich, H. (1952). Seismische Beobachtungen bei großen Steinbruchsprengungen und deren geolgoische Ergebnisse. *Zeitschrift der Deutschen Geologischen Gesellschaft*, pages 152–197.
- Robbins, H. and Monro, S. (1951). A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407.
- Robertsson, J., Blanch, J., and Gupta, H. (2011). Encyclopedia of solid earth geophysics.
- Rondenay, S., Bostock, M. G., and Fischer, K. M. (2005). Multichannel inversion of scattered teleseismic body waves: practical considerations and applicability. *Geophysical Monograph-American Geophysical Union*, 157:187.
- Rudnick, R. and Gao, S. (2003). Composition of the continental crust. *The crust*, 3:1–64.
- Rybach, L. and Buntebarth, G. (1982). Relationships between the petrophysical properties density, seismic velocity, heat generation, and mineralogical constitution. *Earth and Planetary Science Letters*, 57(2):367–376.

- Rybach, L., Mueller, S., Milnes, A., Ansorge, J., Bernoulli, D., and Frey, M. (1980). The Swiss Geotraverse Basel-Chiasso: A review. *Eclogae Geologicae Helvetiae*, 73(2):437–462.
- Rybacki, E. and Dresen, G. (2004). Deformation mechanism maps for feldspar rocks. *Tectonophysics*, 382(3-4):173–187.
- Sambridge, M. (1999a). Geophysical inversion with a neighbourhood algorithm. I. Searching a parameter space. *Geophysical journal international*, 138(2):479–494.
- Sambridge, M. (1999b). Geophysical inversion with a neighbourhood algorithm. II. Appraising the ensemble. *Geophysical Journal International*, 138(3):727–746.
- Sambridge, M., Bodin, T., Gallagher, K., and Tkalčić, H. (2013). Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1984):20110547.
- Sambridge, M., Gallagher, K., Jackson, A., and Rickwood, P. (2006). Trans-dimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, 167(2):528–542.
- Sambridge, M. and Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3):3–1.
- Savage, M. K. (1998). Lower crustal anisotropy or dipping boundaries? Effects on receiver functions and a case study in New Zealand. *Journal of Geophysical Research: Solid Earth*, 103(B7):15069–15087.
- Savage, M. K., Park, J., and Todd, H. (2007). Velocity and anisotropy structure at the Hikurangi subduction margin, New Zealand from receiver functions. *Geophysical Journal International*, 168(3):1034–1050.
- Scarponi, M., Hetényi, G., Berthet, T., Baron, L., Manzotti, P., Petri, B., Pistone, M., and Müntener, O. (2020). New gravity data and 3D density model constraints on the Ivrea Geophysical Body (Western Alps). *Geophysical Journal International*.
- Schaefer, J., Boschi, L., and Kissling, E. (2011). Adaptively parametrized surface wave tomography: methodology and a new model of the European upper mantle. *Geophysical Journal International*, 186(3):1431–1453.
- Schettino, A. and Turco, E. (2011). Tectonic history of the western Tethys since the Late Triassic. *Bulletin*, 123(1-2):89–105.

- Schmid, S. M., Fügenschuh, B., Kissling, E., and Schuster, R. (2004). Tectonic map and overall architecture of the Alpine orogen. *Eclogae Geologicae Helvetiae*, 97(1):93–117.
- Schmid, S. M., Kissling, E., Diehl, T., van Hinsbergen, D. J., and Molli, G. (2017). Ivrea mantle wedge, arc of the Western Alps, and kinematic evolution of the Alps–Apennines orogenic system. *Swiss Journal of Geosciences*, 110(2):581–612.
- Schreiber, D., Lardeaux, J.-M., Martelet, G., Courrioux, G., and Guillen, A. (2010). 3D modelling of Alpine mohos in southwestern Alps. *Geophysical Journal International*, 180(3):961–975.
- Sen, M. K. and Stoffa, P. L. (2013). *Global optimization methods in geophysical inversion*. Cambridge University Press.
- Sethian, J. A. (1999). *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, volume 3. Cambridge university press.
- Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307(5715):1615–1618.
- Sheriff, R. E. (1996). Understanding the Fresnel zone. *AAPG Explorer*, pages 18–19.
- Shibutani, T., Sambridge, M., and Kennett, B. (1996). Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia. *Geophysical Research Letters*, 23(14):1829–1832.
- Singer, J., Diehl, T., Husen, S., Kissling, E., and Duretz, T. (2014). Alpine lithosphere slab rollback causing lower crustal seismicity in northern foreland. *Earth and Planetary Science Letters*, 397:42–56.
- Singer, J., Kissling, E., Diehl, T., and Hetényi, G. (2017). The underthrusting Indian crust and its role in collision dynamics of the Eastern Himalaya in Bhutan: Insights from receiver function imaging. *Journal of Geophysical Research: Solid Earth*, 122(2):1152–1178.
- Sippl, C., Kumar, A., and Dettmer, J. (2017). A Cross-Correlation-Based Approach to Direct Seismogram Stacking for Receiver-Side Structural Inversion. *Bulletin of the Seismological Society of America*, 107(3):1545–1550.

- Solarino, S., Malusà, M. G., Eva, E., Guillot, S., Paul, A., Schwartz, S., Zhao, L., Aubert, C., Dumont, T., Pondrelli, S., et al. (2018). Mantle wedge exhumation beneath the Dora-Maira (U) HP dome unravelled by local earthquake tomography (Western Alps). *Lithos*, 296:623–636.
- Spada, M., Bianchi, I., Kissling, E., Agostinetti, N. P., and Wiemer, S. (2013). Combining controlled-source seismology and receiver function information to derive 3-D Moho topography for Italy. *Geophysical Journal International*, 194(2):1050–1068.
- Spencer Jr, J. W. and Nur, A. M. (1976). The effects of pressure, temperature, and pore water on velocities in westerly granite. *Journal of Geophysical Research*, 81(5):899–904.
- Stampfli, G., Borel, G. D., Marchant, R., and Mosar, J. (2002). Western Alps geological constraints on western Tethyan reconstructions. *Journal of the Virtual Explorer*, 8:77.
- Stampfli, G. M. and Borel, G. D. (2004). The TRANSMED transects in space and time: constraints on the paleotectonic evolution of the Mediterranean domain. In *The TRANSMED Atlas. The Mediterranean region from crust to mantle*, pages 53–80. Springer.
- Stäubli, M., Pfiffner, O.-A., and Smithson, S. (1993). Crustal structure and reflectivity of the Swiss Alps from three-dimensional seismic modeling: 1. Helvetic Nappes. *Tectonics*, 12(4):911–924.
- Stehly, L., Fry, B., Campillo, M., Shapiro, N., Guilbert, J., Boschi, L., and Giardini, D. (2009). Tomography of the Alpine region from observations of seismic ambient noise. *Geophysical Journal International*, 178(1):338–350.
- Stein, S. and Wysession, M. (2003). An introduction to Seismology. *Earthquakes, and Earth*.
- Subedi, S., Hetényi, G., Vergne, J., Bollinger, L., Lyon-Caen, H., Farra, V., Adhikari, L. B., and Gupta, R. M. (2018). Imaging the Moho and the Main Himalayan Thrust in Western Nepal with receiver functions. *Geophysical Research Letters*, 45(24):13–222.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*, volume 89. SIAM.
- Tatham, R. H. and Stewart, R. R. (1993). Present status and future directions of shear-wave seismology in exploration. Technical report, CREWES Research Report, 5.
- Telford, W. M., Telford, W., Geldart, L., Sheriff, R. E., and Sheriff, R. (1990). *Applied Geophysics*, volume 1. Cambridge university press.

- Thouvenot, F., Paul, A., Senechal, G., Hirn, A., and Nicolich, R. (1990). ECORS-CROP wide-angle reflection seismics: constraints on deep interfaces beneath the Alps. *Mémoires de la Société géologique de France*, 156:97–106.
- Thurber, C. and Eberhart-Phillips, D. (1999). Local earthquake tomography with flexible gridding. *Computers & Geosciences*, 25(7):809–818.
- Tiberi, C., Diament, M., Déverchère, J., Petit-Mariani, C., Mikhailov, V., Tikhotsky, S., and Achauer, U. (2003). Deep structure of the baikal rift zone revealed by joint inversion of gravity and seismology. *Journal of Geophysical Research: Solid Earth*, 108(B3).
- Tikhonov, A. N. and Arsenin, V. Y. (1977). Solutions of ill-posed problems. *New York*, pages 1–30.
- Tipper, J. C. (1976). The study of geological objects in three dimensions by the computerized reconstruction of serial sections. *The Journal of Geology*, 84(4):476–484.
- Toksoz, M. N., Cheng, C. H., and Timur, A. (1976). Velocities of seismic waves in porous rocks. *Geophysics*, 41(4):621–645.
- Tommasi, A. and Vauchez, A. (2001). Continental rifting parallel to ancient collisional belts: an effect of the mechanical anisotropy of the lithospheric mantle. *Earth and Planetary Science Letters*, 185(1-2):199–210.
- Tong, P., Yang, D., and Huang, X. (2019). Multiple-grid model parametrization for seismic tomography with application to the San Jacinto fault zone. *Geophysical Journal International*, 218(1):200–223.
- Trumpy, R. (1980). *Geology of Switzerland: a Guide-book: An Outline of the Geology of Switzerland*. Wepf and Company.
- Um, J. and Thurber, C. (1987). A fast algorithm for two-point seismic ray tracing. *Bulletin of the Seismological Society of America*, 77(3):972–986.
- Valasek, P. (1992). The tectonic structure of the Swiss Alpine crust interpreted from a 2D network of deep crustal seismic profiles and an evaluation of 3D effects. *PhD Thesis, Swiss Fed. Inst. Technology Zurich*.
- Vanbiervliet, J., Verheyden, K., Michiels, W., and Vandewalle, S. (2008). A nonsmooth optimisation approach for the stabilisation of time-delay systems. *ESAIM: Control, Optimisation and Calculus of Variations*, 14(3):478–493.

- Verbeke, J., Boschi, L., Stehly, L., Kissling, E., and Michelini, A. (2012). High-resolution Rayleigh-wave velocity maps of Central Europe from a dense ambient-noise data set. *Geophysical Journal International*, 188(3):1173–1187.
- Vinnik, L. (1977). Detection of waves converted from P to SV in the mantle. *Physics of the Earth and planetary interiors*, 15(1):39–45.
- Virieux, J. and Farra, V. (1991). Ray tracing in 3-D complex isotropic media: an analysis of the problem. *Geophysics*, 56(12):2057–2069.
- Voigt, B. F. (1831). Der handlungsreisende, wie er sein soll und was er zu thun hat, um aufträge zu erhalten und eines glücklichen erfolgs in seinen geschäften gewiss zu sein. *Commis-Voageur, Ilmenau*.
- Vouillamoz, N., Mosar, J., and Deichmann, N. (2017). Multi-scale imaging of a slow active fault zone: contribution for improved seismic hazard assessment in the Swiss Alpine foreland. *Swiss Journal of Geosciences*, 110(2):547–563.
- Wagner, M., Kissling, E., and Husen, S. (2012). Combining controlled-source seismology and local earthquake tomography to derive a 3-D crustal model of the western Alpine region. *Geophysical Journal International*, 191(2):789–802.
- Waldhauser, F., Kissling, E., Ansorge, J., and Mueller, S. (1998). Three dimensional interface modelling with two-dimensional seismic data: the Alpine crust-mantle boundary. *Geophysical Journal International*, 135(1):264–278.
- Wang, Z., Nur, A. M., et al. (1989). Effects of CO₂ flooding on wave velocities in rocks with hydrocarbons. *SPE Reservoir Engineering*, 4(04):429–436.
- Weaver, R. L. (2005). Information from seismic noise. *Science*, 307(5715):1568–1569.
- Yang, Y., Ritzwoller, M. H., Levshin, A. L., and Shapiro, N. M. (2007). Ambient noise Rayleigh wave tomography across Europe. *Geophysical Journal International*, 168(1):259–274.
- Yano, T. E., Takeda, T., Matsubara, M., and Shiomi, K. (2017). Japan unified high-resolution relocated catalog for earthquakes (JUICE): crustal seismicity beneath the Japanese Islands. *Tectonophysics*, 702:19–28.
- Ye, S., Ansorge, J., Kissling, E., et al. (1995). Crustal structure beneath the eastern Swiss Alps derived from seismic refraction data. *Tectonophysics*, 242(3-4):199–221.

- Young, M. K., Rawlinson, N., and Bodin, T. (2013). Transdimensional inversion of ambient seismic noise for 3D shear velocity structure of the Tasmanian crust - Transdimensional shear velocity inversion. *Geophysics*, 78(3):WB49–WB62.
- Zandt, G. and Ammon, C. J. (1995). Continental crust composition constrained by measurements of crustal Poisson's ratio. *Nature*, 374(6518):152.
- Zhao, D. (2009). Multiscale seismic tomography and mantle dynamics. *Gondwana Research*, 15(3-4):297–323.
- Zhao, D. (2015). *Multiscale seismic tomography*. Springer.
- Zhao, D., Hasegawa, A., and Horiuchi, S. (1992). Tomographic imaging of P and S wave velocity structure beneath northeastern Japan. *Journal of Geophysical Research: Solid Earth*, 97(B13):19909–19928.
- Zhu, L. and Kanamori, H. (2000). Moho depth variation in southern California from teleseismic receiver functions. *Journal of Geophysical Research: Solid Earth*, 105(B2):2969–2980.
- Zor, E., Sandvol, E., Gürbüz, C., Türkelli, N., Seber, D., and Barazangi, M. (2003). The crustal structure of the East Anatolian plateau (Turkey) from receiver functions. *Geophysical Research Letters*, 30(24).