

**DOTTORATO DI RICERCA IN  
GEOFISICA**

Ciclo XXVI

**Settore Concorsuale di afferenza: 04/A4**

**Settore Scientifico disciplinare: GEO/10**

**SEISMIC TOMOGRAPHIC FULL-WAVEFORM INVERSION  
FOR THE VRANCEA SINKING LITHOSPHERE STRUCTURE  
USING THE ADJOINT METHOD**

**Presentata da: Julie Baron**

**Coordinatore Dottorato**

Prof. Michele Dragoni

**Relatore**

Prof. Andrea Morelli



# TABLE OF CONTENTS

TABLE OF CONTENTS.....	i
LIST OF FIGURES.....	iii
LIST OF TABLES.....	v
INTRODUCTION.....	1
<b>I. THE VRANCEA REGION.....</b>	<b>5</b>
<b>I.I. TECTONIC AND SEISMIC SETTINGS .....</b>	<b>7</b>
I.I.1. EVOLUTION OF THE CARPATHIANS.....	7
I.I.2. SEISMICITY .....	10
<b>I.II. GEODYNAMICAL MODELS AND TOMOGRAPHIC STUDIES</b>	<b>13</b>
I.II.1. GEODYNAMICAL MODELS.....	13
I.II.2. TOMOGRAPHIC STUDIES.....	15
<b>II. FULL WAVEFORM INVERSION THEORY .....</b>	<b>19</b>
<b>II.I. ABOUT SEISMIC TOMOGRAPHY .....</b>	<b>21</b>
II.I.1. FROM RAY BASED APPROACH TO 3D TOMOGRAPHY.	21
II.I.2. THE INVERSE PROBLEM.....	23
II.I.3. SOLVING THE FORWARD PROBLEM .....	25
II.I.4. MISFIT CRITERIA.....	29
<b>II.II. THE ADJOINT METHOD.....</b>	<b>31</b>
II.II.1. SOLVING MINIMIZATION ALGORITHMS.....	31
II.II.2. COMPUTING THE GRADIENT WITH THE ADJOINT METHOD .....	34
II.II.3. APPLICATION TO STRUCTURE INVERSION .....	36
<b>II.III. SOURCE INVERSION: A KEY POINT FOR TOMOGRAPHIC     STUDIES .....</b>	<b>39</b>
II.III.1. SOURCE LOCATION.....	39
II.III.2. SOURCE MECHANISM .....	39
<b>III. AN INSIGHT INTO THE VARIABLE PROJECTION METHOD.....</b>	<b>43</b>
<b>III.I. INTRODUCTION .....</b>	<b>45</b>
III.I.1. MOTIVATIONS.....	45
III.I.2. THEORY .....	46
III.I.3. SOLVING A JOINT SOURCE SIGNAL AND LOCATION INVERSION PROBLEM WITH THE VPM .....	47
<b>III.II. COMPUTE SYNTHETIC WAVEFORMS IN AN     HOMOGENEOUS ISOTROPIC ACOUSTIC MEDIUM.....</b>	<b>49</b>
III.II.1. GREEN'S FUNCTION .....	49
III.II.2. WAVELET SIGNAL .....	51
<b>III.III. THE SOURCE SIGNAL PROBLEM .....</b>	<b>53</b>
<b>III.IV. THE SOURCE LOCATION PROBLEM.....</b>	<b>55</b>
III.IV.1. MISFITS EXAMPLES.....	55
III.IV.2. OPTIMIZATION ALGORITHMS .....	56
<b>III.V. APPLICATION.....</b>	<b>63</b>

<b>FULL WAVEFORM INVERSION WITH THE ADJOINT METHOD APPLICATION TO THE VRANCEA REGION USING LOCAL DATA FROM THE CALIXTO99 EXPERIMENT .....</b>	<b>67</b>
<b>IV.I. SETTING UP THE VRANCEA DATASET .....</b>	<b>69</b>
IV.I.1. DATA FROM THE 1999 CALIXTO EXPERIMENT .....	69
IV.I.2. VRANCEA 3D REFERENCE TOMOGRAPHIC MODEL ...	71
IV.I.3. SOURCE LOCATION AND MOMENT TENSOR.....	75
IV.I.4. PROCESSING THE DATABASE .....	78
<b>IV.II. SEISMIC WAVEFORM MODELING.....</b>	<b>83</b>
IV.II.1. WAVEFORM MODELING WITH SPECFEM.....	83
IV.II.2. FORWARD MODELING ON CINECA BLUE GENE/Q ...	85
<b>IV.III. STRUCTURE INVERSION .....</b>	<b>91</b>
IV.III.1. FROM MISFIT TO ADJOINT SOURCE .....	91
IV.III.2. FROM EVENT KERNELS TO FIRST UPDATE .....	94
IV.III.3. FROM INITIAL MODEL TO 3 <sup>RD</sup> UPDATE .....	97
 <b>CONCLUSIONS .....</b>	 <b>107</b>
<b>APPENDIX A - TABLES.....</b>	<b>111</b>
<b>APPENDIX B - VELOCITY MODELS.....</b>	<b>121</b>
<b>APPENDIX C - EVENT RELOCATION.....</b>	<b>125</b>
<b>APPENDIX D - SPECFEM3D .....</b>	<b>129</b>
<b>BIBLIOGRAPHY .....</b>	<b>139</b>
<b>ACKNOWLEDGMENTS .....</b>	<b>149</b>

# LIST OF FIGURES

Figure 1 - The Vrancea region in a European geo-topographic context. ....	7
Figure 2 - Sum up of the Carpathian tectonic evolution .....	8
Figure 3 - Vrancea, Bucaramanga and Hindu Kush nests.....	10
Figure 4 - Seismicity distribution in Romania .....	11
Figure 5 - Fault plane solutions from the Global CMT Harvard catalogue .....	12
Figure 6 - Geodynamic models for the Vrancea high velocity deepening volum.....	13
Figure 7 - Examples of seismic tomography studies.....	16
Figure 8 - NW-SE Vp profile across Vrancea .....	17
Figure 9 - NW-SE Vp/Vs ratio model cross section.....	18
Figure 10 - Sum up of the spectral element method.....	28
Figure 11 - Common misfit functions .....	30
Figure 12 - Construction of a cross-correlation travel-time sensitivity kernel .....	38
Figure 13 - Parameters describing the source mechanism: strike, dip and rake angles ..	39
Figure 14 - Morlet and Ricker wavelet .....	51
Figure 15 - Morlet and Ricker seismic signals .....	51
Figure 16 - Output from Wiener filtering .....	53
Figure 17 - Travel-time misfit function distribution. ....	55
Figure 18 - Waveform misfit function distribution. ....	56
Figure 19 - 2D representation of the waveform misfit distribution.....	58
Figure 20 - 3D representation of the waveform misfit distribution.....	58
Figure 21 - Location iteration 1 .....	61
Figure 22 - Seismic signals differences evolution 1 .....	61
Figure 23 - Location iteration 2 .....	62
Figure 24 - Seismic signals differences evolution 2 .....	62
Figure 25 - Waveform misfit distribution with two unknowns .....	63
Figure 26 - Optimization algorithm using the variable projection method 1.....	65
Figure 27 - Seismograms evolution through the VPM optimization algorithm. ....	65
Figure 28 - Optimization algorithm using the variable projection 2. ....	66
Figure 29 - Source signal evolution through the VPM algorithm. ....	66
Figure 30 - Distribution of the CALIXTO99 database (events and stations).....	69
Figure 31 - Distribution magnitude-depth for the CALIXTO99 input database .....	70
Figure 32 - 3D view of the initial tomographic model used for the inversion .....	72
Figure 33 - Slices view of the tomographic model from Tondi et al. (2009).....	73
Figure 34 - Slices views of the initial tomographic model used for inversion. ....	74
Figure 35 - Moment tensor informed input database .....	75
Figure 36 - Sample of the FPFIT fault plane solutions. ....	76
Figure 37 - Fault plane mechanism solutions.....	77
Figure 38 - Noise filtering .....	78
Figure 39 - Directly usable database .....	79
Figure 40 - Spectrograms .....	80
Figure 41 - Sample from the 990807 event at station A12.....	80
Figure 42 - Final event database distribution used for the full waveform inversion .....	81
Figure 43 - SPECFEM3D workflow.....	84
Figure 44 - Example of an irregular mesh .....	85
Figure 45 - Magnitude/Depth distribution of the 19 events dataset .....	86
Figure 46 - Minimum period resolution through the mesh .....	86
Figure 47 - Sample of forward simulations for the 990622 event.....	87
Figure 48 - Time shift normal distributions in the dataset.....	88
Figure 49 - Adjustments done for depth and reference time .....	89
Figure 50 - Time shift normal distributions in the dataset after adjustments .....	89
Figure 51 - Depth comparison with respect to the CALIXTO99 archive.....	90
Figure 52 - Reference time comparison with respect to the CALIXTO99 archive .....	90
Figure 53 - Selection windows example.....	91
Figure 54 - Misfit values computation.....	92

Figure 55 - Cross-correlation travel-time misfit distribution for P-S windows .....	93
Figure 56 - Computational costs involved by the smoothing operation.....	95
Figure 57 - Smoothing effect on a 5events misfit kernel in 3D.....	96
Figure 58 - Details of smoothing effect on a 5events misfit kernels on 2D planes .....	96
Figure 59 - Travel-time cross-correlation misfit measurements for 3 trial steps length...	98
Figure 60 - Travel-time cross-correlation distributions for 3 trial steps length .....	98
Figure 61 - Travel-time cross-correlation misfit measurements at first 3 iterations .....	99
Figure 62 - Misfit measurements among model updates. ....	99
Figure 63 - Cross-correlation measurements from M0 to M2 .....	100
Figure 64 - Cross-correlation signals for first 2 model updates at 3 stations .....	101
Figure 65 - Vp model relative variations at different depths .....	103
Figure 66 - Vs model relative variations at different depths .....	104
Figure 67 - Reference 1D velocity model. ....	122
Figure 68 - Geophysical discontinuities in the reference 3D model.....	122
Figure 69 - Vp and Vs first 2 updates velocity models at slice depths. ....	123
Figure 70 - Reference time shifts between NLLOC 1 and CALIXTO99 archive values..	127
Figure 71 - Reference time shifts between NLLOC 2 and CALIXTO99 archive values..	127
Figure 72 - External tomographic model file format for SPECFEM3D.....	130
Figure 73 - From irregular to regular spaced tomographic model.....	131
Figure 74 - Simplified call trees of the three SPECFEM executables .....	132
Figure 75 - Tasks management with core and virtual CPU.....	134
Figure 76 - SPECFEM3D efficiency on CINECA Sp6 architecture. ....	135
Figure 77 - Mesh set up for different resolution tests on Sp6. ....	135
Figure 78 - SPECFEM3D efficiency on CINECA BG/Q architecture. ....	137

# LIST OF TABLES

Table 1 - Instrumental characteristics of the CALIXTO Broadband seismic stations. ...	112
Table 2 - Source parameters from the CALIXTO99 archive .....	113
Table 3 - Source parameters from NLLOC.....	115
Table 4 - Source parameters from NLLOC after the corrections .....	117
Table 5 - Availability of the dataless-files from GFZ Potsdam .....	119
Table 6 - Reference time differences between the values given by the Amaru et al. (2008) catalogue, the CALIXTO99 experiment and the Tondi et al. (2009) study.....	126
Table 7 - Sample of possible combinations for building a mesh on FERMI-BG/Q.....	137

# INTRODUCTION

The Vrancea region, at the sharp bend of the southeast Carpathians, represents one of the most seismoactive zone in Europe. It is the place of an unusually intense seismicity with the presence of a cluster of intermediate-depth (60-180km) seismicity in a narrow, nearly vertical and NE-SW striking volume, contrasting with the low magnitude shallow seismicity spread across the Romanian territory. The strongest events with  $M_w > 7$  that occurred in the last century (1908, 1940, 1977 and 1986) caused high loss and trembled cities up to several hundred kilometers away from the epicenter. According to the Romanian historical catalogue (Radu 1979, 1991) such large events occur about three times per century. This relatively high potential of seismic hazard, makes the region one of the most studied areas in Europe. The strong intra-continental seismicity occurring at intermediate depth raised questions about the origin and the nature of such earthquakes and required a better understanding of the dynamics of this area. First geodynamical models were proposed after the 1940 earthquake with  $M_w = 7.7$  at a depth of approximately 160km. The central interrogation for seismic risk mitigation is the determination whether this singular seismogenic volume is geodynamically coupled to the crust. Such coupling would result in the occurrence of strong earthquakes in the crust and even in the aseismic zone above the volume between 40km and 70km, potentially causing heavy material and human costs. Nowadays, this potential coupling as well as the tectonic evolution of the Vrancea region still remains a bone of contention inside the scientific community.

To improve the knowledge of the geodynamics and to enlighten the relations between shallow and deep Earth's structure beneath Vrancea, several complementary studies were carried out in the last decades. They were realized in a large range of research fields, from geology to geodesy (e.g. using GPS data) and geophysics, including seismology. Progress in the applied methods and the combination of their results is a key point to propose a valid and robust geodynamical model, and to improve seismic risk mitigation for the Vrancea zone and the neighbour cities such as the capital Bucharest sheltering about two millions inhabitants. So far two main geodynamical models have been proposed to explain the dynamic of Vrancea: one in favour of a subduction-related process (Mc Kenzie, 1972; Fuchs et al., 1979; Oncescu, 1984; Linzer et al., 1998; Wortel and Spakman, 2000; Sperner et al., 2001; Heidbach et al., 2007), and, more recent, another supporting a delamination dynamic (Chalot and Prat, 2000; Knapp et al., 2005; Lorinczy and Houseman, 2009; Enciu and Knapp, 2009).

Seismic tomography aids unveiling the Earth subsurface properties from the information contained in waveform registrations of ground motion due to a seismic source, recorded by receivers at the surface. This is achieved by minimising a misfit function of the differences between candidate Earth's model waveforms and the real Earth's waveforms. This inverse method is applicable to different scales from global Earth's model to regional and local study.

Thus, several seismic tomography studies have been applied to the Vrancea area using teleseismic or local data from temporary experiments. Though they have revealed a cold nearly vertical lithospheric volume under Vrancea, collocated

## *Introduction*

with the seismic nest, they did not provide a unanimous final interpretation on the nature and origin of such geometry.

Nowadays, high-resolution seismic tomography could help to reveal a detailed subcrustal structure below Vrancea as well as the properties of the “down-going” seismogenic volume. Recent developments in this field as well as in parallel computing allow the collection of more information out of seismic data and the production of such fine models. The full waveform inversion (FWI) method together with the finite frequency theory, allow the nearly complete recovery of seismic waveform information content for tomographic inversion.

The aim of this study is to implement a full waveform inversion of the Vrancea region using local data collected during a temporary six months experiment undertaken in 1999 (CALIXTO99: Carpathian Arc Lithosphere X-Tomography). The study is based on the finite frequency theory with the simulation of synthetic seismograms using the SPECFEM3D Cartesian software, developed by Komatitsch et al. (2002a, b). The inversion follows a gradient-based algorithm for the iterative model update. In particular, I use the adjoint method implemented in SPECFEM3D to compute the model gradient for a given misfit. These recent methods offer a more accurate physical approximation of seismic waves dynamic, but also imply heavier computational costs. Main part of the inversion processing dealing with waveform simulation, gradient computation, smoothing of the gradient and model updating are done in parallel on a cluster at the Consorzio Interuniversitario CINECA.

In case of Vrancea, the seismic sources distribution with strong events at intermediate depth and smaller ones at shallow depth, limit the selection of the dataset to the strongest events at intermediate depth with acceptable signal-to-noise ratio. Consequently, the source-receiver geometry of the selected dataset is different from common tomographic cases. All seismic sources are clustered in the same location, resulting in sparse ray path coverage, quite similar rays and few crossing ray paths between different source-stations pairs. Implementing a local tomographic study only, these problems cannot be avoided. One would need to include teleseismic events, however that is not the scope of the current study. I thus show to which extent and at which costs this limited dataset can be exploited in the implementation of a full waveform inversion of the Vrancea zone, and if it allows to retrieve information on the velocity distribution at a better level than with a classical ray tomography.

This thesis is organised into four mainly independent parts. The first part gives an overview of the knowledge on the Vrancea region and the different geophysical studies, which have been carried out, as well as the geodynamical models proposed for this region. In the second part, I present an outline of the full waveform inversion and seismic tomography theory. This aims to expose the principles of inversion methods and the characteristics of the full waveform inversion, from the complete numerical solution of the waves equation to the finite frequency theory and the adjoint method. The third part gives an introduction to the variable projection method (VPM) that could be seen as an interesting tool to deal with simultaneous source and structure inversion, a key point in seismic tomographic studies. This part presents the work conducted during a three months internship at the Schlumberger Gould Research centre in Cambridge, UK. The fourth part presents the application of the FWI to the Vrancea case, using a discrete-quality local dataset.

## *Introduction*

In this part, I discuss in details, the preparation step of the tomographic study that consists in data collection, processing and selection, as well as the different phases of the inversion method, from seismic waveform modeling to model update. This part gives a good illustration on how this full waveform inversion method is conditioned by the initial set-up.



# I

## THE VRANCEA REGION

<b>I.I. TECTONIC AND SEISMIC SETTINGS .....</b>	<b>7</b>
I.I.1. EVOLUTION OF THE CARPATHIANS.....	7
I.I.2. SEISMICITY .....	10
<b>I.II. GEODYNAMICAL MODELS AND TOMOGRAPHIC STUDIES</b>	<b>13</b>
I.II.1. GEODYNAMICAL MODELS.....	13
I.II.2. TOMOGRAPHIC STUDIES.....	15



## I.I. TECTONIC AND SEISMIC SETTINGS

The Vrancea region, in Romania, is a complex tectonic zone, part of the highly arched South-East Carpathians orogeny (Figure 1).

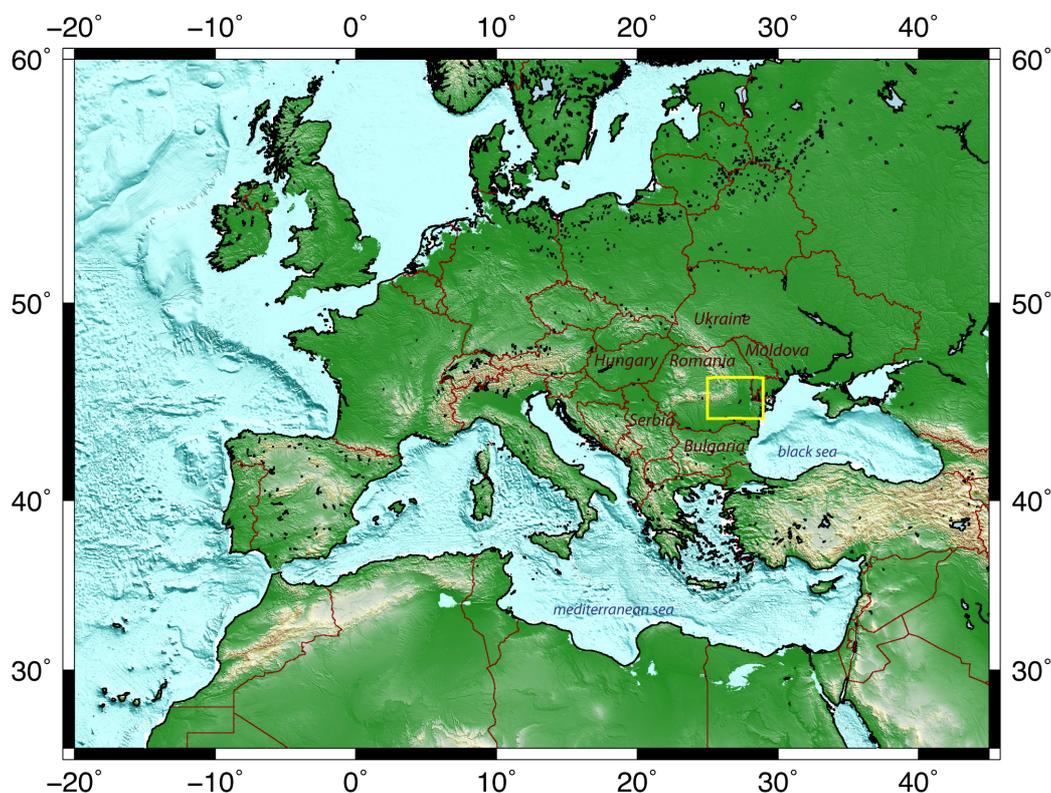


Figure 1 - The Vrancea region in a European geo-topographic context.

### I.I.1. EVOLUTION OF THE CARPATHIANS

Numerous authors have described the geology of the Carpathians: among others, Burchfield (1980), Foldary (1988), Linzer et al. (1998), Schmidt et al. (2008), Ismael Zadeh et al. (2012). The Carpathian belt, at the southeast part of the Alpine chain, finds its tectonic origins in the formation of the Alps and the closure of the oceanic lithosphere of the Tethys driven by the motion of the African and European plates in the late Jurassic.

The convergence between the African and Eurasian plates gives rise to a particularly complicated stress configuration on the Romanian territory because of the interaction between several subplates or intra-tectonic units enclosed between them. These accommodate the relative displacements driven by the main convergence motion as in a puzzle game. The literature does not bring a clear view on their delimitation and deformation stage (neither on their names) as their evolution is still a matter of debate. Otherwise, crossing different sources, I will attempt to give a synthesis of the Carpathians tectonic evolution based on the motion of these intra-tectonic units (Figure 2).

#### Jurassic times

In the Carpathian area, four main intra-tectonic units might have formed at Jurassic time, as proposed by Burchfield (1980): the Apulian (later named ALCAPA for Alps-Carpathians-Pannonian), Rhodopian (Dacia blocks in Hungarian literature), Moesian and North-Dobrogean fragments.

#### Cretaceous times

In late Jurassic/early Cretaceous, an extension motion between Tisca and Moesian platforms led to the aperture of two oceans in the Carpathian area: the Cealhau-Severin (between Dacia and European/Moesian foreland) and East Vadar (between Tisca and Dacia blocks) oceans. This motion inverted later in a convergence motion of the plates, leading to the first continental collision between the Rhodopian and Moesian units during Albian time and the gradual closure of the east Vadar ocean.

In late Cretaceous, the Apulian fragment rotated counter-clockwise moving northward, and the Rhodopian fragment rotated clockwise as it moved around the Moesian fragment. At this time, continental as well as oceanic collision occurred, with the collision of the Apulian with the Rhodopian platform. By the late Cretaceous, there was no more convergence in the South Carpathians.

#### Paleogene times

There is a lack of information to recover the entire Paleogene deformation in the Carpathians. The main motion consisted in the closure of the Severin ocean by subduction and the final closure of the Tethys in late Oligocene.

#### Neogene times

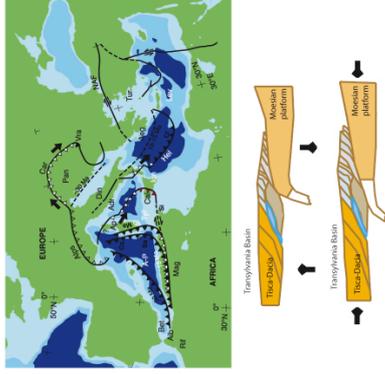
The last thrusting events occurred in the Miocene. Tisca-Dacia and ALCAPA plates moved into the Carpathian embayment forming the Miocene foreland fold-and-thrust-belt. In the late-early Miocene the Carpathian embayment subduction started. It stopped in the northern part in Middle Miocene, with continental collision between Tisca-Dacia and Moesian units and rotated from NW to SE before starting a gradual roll back. This retreat of the subducted oceanic lithosphere is associated with crustal thinning and upwelling of the asthenosphere underneath the Pannonian basin. The late Miocene defines the final emplacement of ALCAPA, Tisca-Dacia mega units and Alpine-Carpathian orogenic. The subsidence in the foreland, started in middle Miocene, is still active with a rate about 1mm/year.

#### From Pliocene to Nowadays

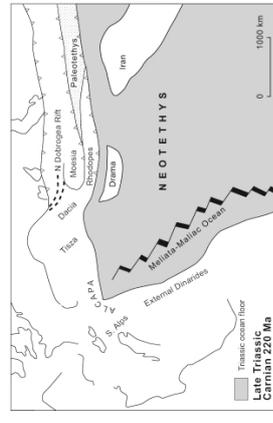
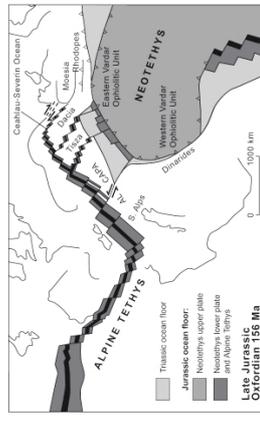
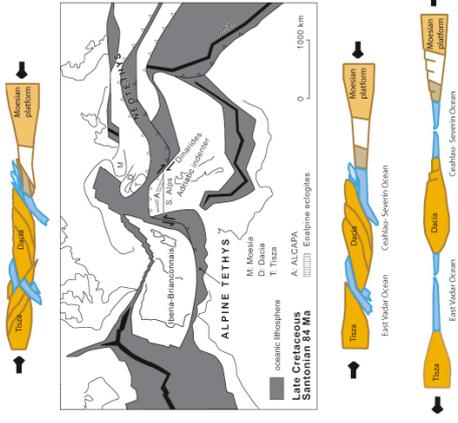
Plate convergence in the northern and eastern Carpathians seems to be currently inactive, while in the SE Carpathians active processes are still taking place, revealed by a particular seismicity in Vrancea. The Vrancea region is a particularly complex tectonic region of continental convergence characterized by at least three tectonic units in contact: the East European plate, Intra-Alpine and Moesian subplates. The evolution of the Vrancea lithosphere, which marks the youngest part of the subduction/collision along the Carpathians, is still under debates. Several models have been proposed into explain the geodynamic evolution of this subducted lithosphere and its present-day expression as a near vertical high velocity body beneath Vrancea.

**Figure 2 (next page) - Sum up of the Carpathian tectonic evolution. Figures modified from GSA, Merten et al. (2010), Schmidt et al. (2008) and Csontos et al. (2004).**

CENOZOIC		PERIOD	EPOCH	TECTONIC DESCRIPTION
AGE (Ma)	MAGNETIC POLARITY CHANGES	NEOGENE	HOLOCENE	renewed contraction evolution of eastern carpathian: slab break-off? delamination?
			QUATERNARY	
5	1 2 3 4 5	PLIOCENE	MIOCENE	Alpine-Carpathian orogenic final emplacement of Alcapa, Tisza-Dacia Mega Units last thrusting events start volcanic activity in Carpathians, start subsidence
				10
15	6 7 8 9 10 11 12	MIOCENE	MIOCENE	subduction from NW to SE, rollback stop subduction in northern part: continental collision
				20
25	13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	PALEOGENE	OLIGOCENE	Carpathian embayment subduction blocks rotations at 18Ma and 16 Ma Tisza-Dacia plate motion into Carpathian embayment: Miocene thrust belt
				30
35	28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	PALEOGENE	EOCENE	final closure of Thetys
				40
45	41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	PALEOGENE	EOCENE	closure of Severin ocean by subduction
				50
55	51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	PALEOGENE	PALEOCENE	Lake of information to recover the entire Paleogene deformations
				60



MESOZOIC		PERIOD	EPOCH	TECTONIC DESCRIPTION
AGE (Ma)	MAGNETIC POLARITY CHANGES	CRETACEOUS	LATE	convergence steps in South Carpathian rotation collision between Apulian and Rhodopian platforms Continental and oceanic subduction Apulian motion northward
90	31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	CRETACEOUS	EARLY	1st continental collision between Rhodopian and Moesian platforms. gradual closure of East Vadar Ocean (remnant of thetys)
				110
130	60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	JURASSIC	LATE	formation of Cealhau. and Severin ocean nappes.
				140
160	80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	JURASSIC	MIDDLE	Subduction starting of oceanic lithosphere of the Alpine Thetys.
				170
190	90 91 92 93 94 95 96 97 98 99 100	JURASSIC	EARLY	
				200
220	100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300	TRIASSIC	LATE	
				210
240	210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300	TRIASSIC	MIDDLE	
				230
250	240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300	TRIASSIC	EARLY	
				240



## I.1.2. SEISMICITY

Observation of global seismicity has provided a strong support for the tectonic plate theory founded on sea-floor spreading and lithosphere recycling at plate boundaries. Although seismic source physical process has been quite understood for shallow earthquakes that represent 75% of this global seismicity, the physical mechanism of the remaining 25% earthquakes deeper than 50km is still puzzling. From these, intermediate-depth earthquakes – in the range from 50 to 300km depth- occur most often in association with convergent plate boundaries within subducting lithosphere in the Wadati-Benioff zone. However, in few places around the world, an intra-continental unusual intermediate-depth seismic activity has been observed. These regions of concentrated high stationary activity relative to the surroundings are defined as nests (Zarifi et al., 2003; Prieto et al., 2012; Koulakov et al., 2010). In particular, nests related to subduction zone process count three main known examples: The Bucaramanga nest in Columbia, the Hindu Kush in Afghanistan and the Vrancea nest in Romania (Figure 3).

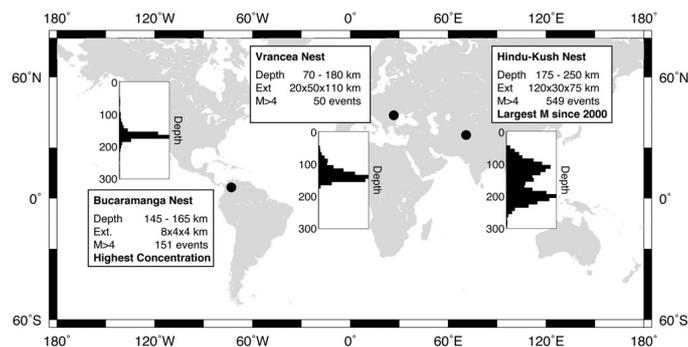


Figure 3 - Location and characteristics of the Vrancea, Bucaramanga and Hindu Kush nests. Based on the ISC 2000 to 2010 earthquakes catalogue. Prieto et al (2012).

The Vrancea region, located in the eastern part of the Carpathians arc bend in Romania, represents one of the most seismically active zones in Europe. The seismicity of the Romanian territory combines a shallow activity of moderate events following the regions of the orogenic belt (Bala et al., 2003) with an intermediate seismicity of strong events concentrated above the Vrancea area. This confined intermediate depth seismicity contrasts with the lack of activity elsewhere along the Carpathians.

After the 1977 earthquake of magnitude Mw 7,5, a telemetered seismic network of 17 short-period stations was installed in Romania. It has been modernized in 1990 through the installation of a complete seismological system for acquisition and data processing (earthquake localization and fast magnitude determination) (Ionescu and Neagoe, 2008).

The Romanian seismic network has started to be continuously developed from 2002, with the installation of new real-time seismic stations. It is currently composed of 82 broadband and short-period digital seismic stations.

Seismic data are analysed by the NIEP (National Institute for Earth Physics) and local earthquakes characteristics (location and fault plane mechanism) are archived on the ROMPLUS catalogue.

The seismicity distribution above Vrancea (Figure 4) suggests a nearly vertical small seismic volume. Looking more in details, there is a seismic gap around 70km depth, and the intermediate depth seismicity is concentrated from 80km in a volume of about 90\*20\*20km (Raykova et al., 2006; Hurukawa et al., 2008), with the main part of these events occurring around 130km depth. The particularity of the dimension and geometry of the Vrancea seismic volume makes difficult any interpretation of the dynamic of this zone.

Fault mechanisms of the Vrancea seismicity present a large range of solutions as shown in Figure 5 and in the study of Sandu et al. (2008), delineating a complex stress process. At crustal range, strike-slip mechanisms seem to have predominance above the whole range of other solutions (Ardeleanu et al., 2005, Bala et al., 2003). Instead, the fault plane geometry in the Vrancea subcrustal volume seems to reflect a predominant compressive stress regime.

In their study in 2003, Bala et al. have defined three distinct seismic zones inside the Vrancea intermediate depth seismic volume: one from 40 to 100km, the second from 100 to 140km and the third below 140km depth. The first domain is characterised by a reverse faulting on steeply dipping faults in a NE-SW direction. In the second domain, the nodal planes are slightly rotated towards a N-S direction and follow a strike slip movement. The deepest part has a less clear distribution of fault mechanisms, with predominance for dip-slip faulting.

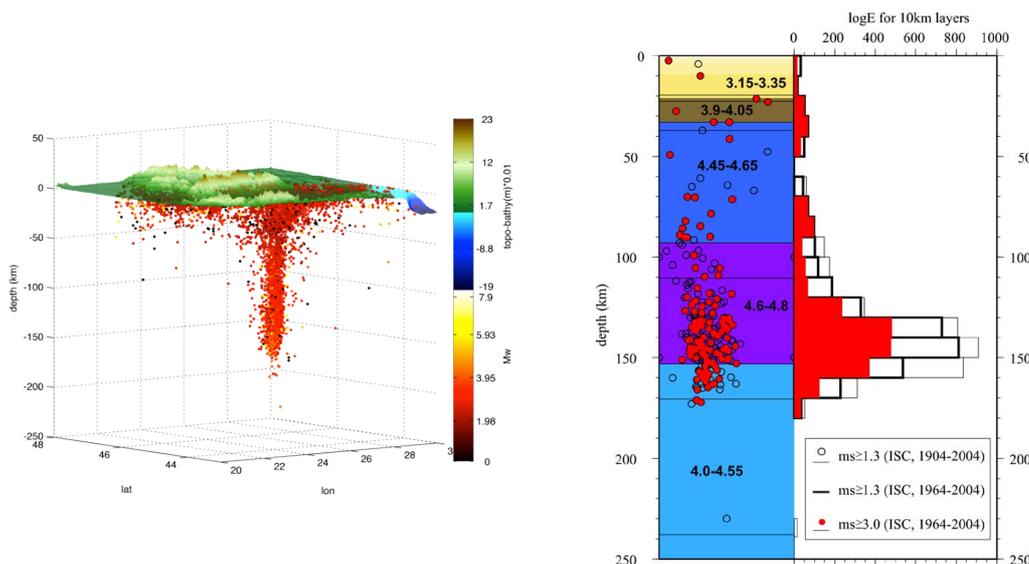


Figure 4 - Seismicity distribution in Romania. On the left, is given the 3D distribution based on the ROMPLUS catalogue with an evidence of the near vertical seismic prone volume. On the right, the figure from Raykova et al. (2006) shows the seismic-energy depth distribution of Vrancea. On can note the two main types of seismicity: a shallow and an intermediate depth one, with a seismicity decrease around 70km depth.

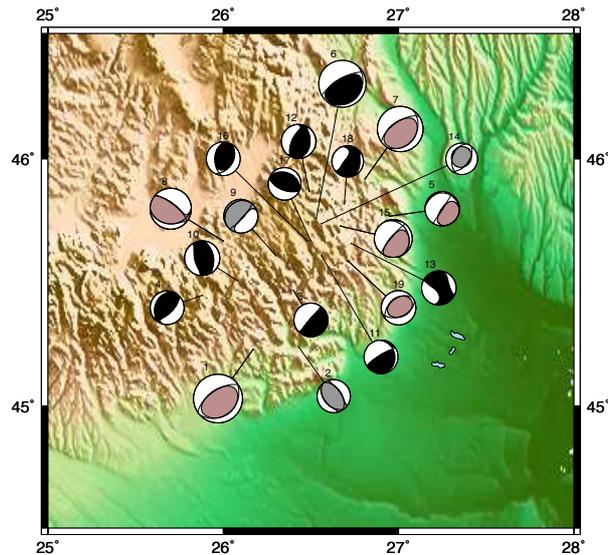


Figure 5 - Fault plane solutions from the Global CMT Harvard catalogue. Events are going from 1977 to 2009. Events with depth between 70 and 100km are indicated by the sand colour, deeper events are in black and event above 150km depth are in grey. No clear evidence for the stress field is seen, indicating a complex process.

Regarding not only the geodynamic side, understanding the nature of tectonic processes occurring in populated areas is also important for seismic risk mitigation. From a historical point of view, strong earthquakes with a magnitude greater than 7 occur in this area at a frequency about 3 per century. In the past, the events of 1908, 1940, 1977 and 1986 have caused large damages over the region and particularly in Bucharest, which is now populated by about 2millions inhabitants. This makes the Vrancea region, a favourite area for various complementary geological and geophysical studies. Despite of the abundance of studies in the region, the nature of this intermediate-depth seismicity cluster is still a bone of contention.

## I.II. GEODYNAMICAL MODELS AND TOMOGRAPHIC STUDIES

### I.II.1. GEODYNAMICAL MODELS

Seismic tomography studies have revealed a nearly vertical high velocity body beneath Vrancea, extending until 350km and whose NE part is the place of an intermediate depth seismogenic volume (Enciu et al., 2009; Martin et al., 2006; Wortel and Spakman, 2000). This high velocity body is interpreted as a cold and dense lithospheric material into the upper mantle.

The tectonic evolution of Vrancea is still under debates (Knapp et al., 2012; Sperner et al. 2001; Martin et al., 2006, Dando et al., 2011). The principal dissension lies on the type of this descending material and the degree of coupling between this material and the crust (Figure 6).

The first proposition stands for a remnant oceanic slab from a past subduction pulled beneath Vrancea by subduction retreat. In this first group, three hypotheses are competing: the attached slab (Mc Kenzie, 1972), the slab breakoff (Linzer et al., 1998; Girbacea and Frisch, 1998 and Wenze et al., 1998) and the weak coupled slab that defines a slab detachment in its final stage (Wortel and Spakman, 2000; Sperner et al., 2001; Muller et al., 2010).

The second proposition sets that this body is made of continental material resulting from a delamination process (Chalot-Prat and Girbacea, 2000; Knapp et al., 2005; Tondi et al., 2009; Fillerup et al., 2010, Ducea, 2011; Ren et al., 2012). Mantle delamination would have consisted of a thickening of the Carpathians lithosphere during continental collision, followed by its sinking in the mantle due to instability and the upwelling of the asthenosphere.

A coupling of both processes stands also as an explanation of the evolution of the lithospheric material (Carbunar and Radulian, 2011), as the alternative model of an unstable triple junction point between the East-European, Moesian and Intra-Alpine micro-plates moving at different velocity rates (Besutiu et al., 2001).

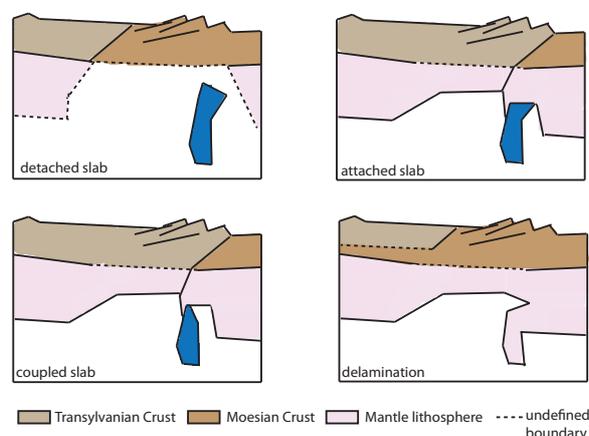


Figure 6 - Simplified sketch picture of the different geodynamic models proposed for the Vrancea high velocity deepening volume (redrawn after Knapp et al., 2005).

First geodynamical model related the Vrancea intermediate depth seismicity to the sinking of the relic of an oceanic slab, remnant of the closure of an oceanic crust floored basin (Mc Kenzie, 1972). Later, the presence of a seismic gap (to be interpreted as a low seismicity with respect to surroundings) between 40 and 70km depth, conducted Fuchs et al. (1979) to the hypothesis of a detached slab geodynamic model. The following tomographic studies revealed the shape of the seismic anomalies beneath the area. The latter, interpreted with regard to the seismicity gap, have supported the model of slab detachment (Onescu, 1984; Linzer et al., 1998; Girbacea and Frisch, 1998). In this dynamic, a subduction would have started westward of the Vrancea seismogenic zone in pre-Miocene times, and would have undergone a process of detachment and retreat toward southeast to position at the actual intermediate depth seismicity zone of Vrancea. Between the attached slab hypothesis and the detached one, the idea of a progressive decoupling of the slab along the Carpathians came out (e.g Sperner et al., 2001). Arguments that support this intermediate model defend that a slab break-off dynamic would be unable to explain the high strain rate inside the slab and that the spatial relationship between the Vrancea seismicity and the active surface deformation in the foreland basin suggests, instead, the existence of a mechanical coupling between the crust and the upper mantle (Enciu et al., 2009). For the last 20 years, these various subduction models have prevailed in most of the geophysical studies interpretation of the Carpathian-Pannonian system tectonic evolution (Fillerup et al. 2010).

However this subduction hypothesis seems to present some weaknesses. Indeed, such a process would imply some suture evidences from continental plates collision, whereas such a boundary has not been established yet (Knapp et al., 2005). To add some doubts on the subduction premise is the pronounced late Miocene/Pliocene subsidence in the Focsani basin of the foreland and the active surface deformation and crustal seismicity of the region. Moreover, if the subduction interpretation is supported by the presence of a Neogene volcanism within the hinterland, the relationship between this volcanism and the subduction process remains questionable (Koulakov et al., 2010; Fillerup et al., 2010).

The continental lithospheric delamination model instead, is distinguishable by its distance from oceanic subduction dynamic and appeared as another alternative to lithosphere recycling. This process is characterised by a thickening of the lithosphere due to the upwelling of the asthenosphere. In the Vrancea region, it would have result from the closure of an intra-continental basin during the Miocene. Such a dynamic induces gravity instability due to higher density in the lithosphere than in the asthenosphere and might be able to generate intermediate depth seismicity such as in the Vrancea case (Chalot-Prat and Girbacea, 2000; Knapp et al., 2005; Tondi et al., 2009; Koulakov et al. 2010).

The debate upon all these oceanic/continental lithosphere hypotheses comes from the uncertainty upon the evolution of the oceanic lithosphere created by the Cealhan-Severin ocean that opened during The Late-Early Cretaceous. Up to now, there is no clear evidence whether this oceanic lithosphere was entirely subducted by the late Cretaceous or whether a part of the oceanic material (post-Cretaceous) is contained in the Carpathian embayment (Ismael Zadeh et al., 2012).

Appart from qualitative interpretations of the Vrancea geodynamics which aim at describing the processes that explain the observed geophysical properties, quantitative interpretations, instead, seek to make the bridge between geophysical concept modelling and specific observed features.

To explain the distribution of the Vrancea seismogenic zone, Ismael Zadeh et al. (2000) carried out a study of viscous flow phase transition and dehydration on the stress field of the slab. From the shape similarity between their predicted depth distribution of the stress in the slab with the average seismic energy released by the earthquakes in a whole year, the process of rock dehydration appeared to them as a realistic candidate for earthquake triggering in the Vrancea slab. Successively in 2005, Ismael Zadeh et al. computed in details the stress repartition along the slab using a 2D thermo-mechanical model of the post Miocene slab subject to gravity. The maximum predicted stress occurred between 80 and 200km, whereas a minimum was found in the depth range of 40 to 80km. These areas correspond respectively to the intermediate-depth seismicity of Vrancea and to its seismic gap. Moreover, their model predicted crustal uplifts in agreement with the East Carpathian orogen, as well as a subsidence area coinciding with the Moesian and East European platforms. These results demonstrated that the Vrancea descending lithosphere could be a consequence of downwelling flows generated by slab subduction. These flows would have first teared down a part of the continental crust and would have successively had particular effects on the surface topography, influencing the Carpathian's orogen and its basins evolution.

In 2009, Manea and Manea deduced from their thermal model of the region that a "soft" attached slab was able to explain the particular Vrancea seismic distribution and rejected the hypothesis of detached slab as this model would not induce enough stress in this area for earthquake generation.

The origin and nature of the Vrancea lithospheric structure is still puzzling. There is a need to study and combine high-resolution data from the different fields of Earth's science such as geophysics, geochemistry, geodesy and seismology to better understand the geodynamical process behind the Vrancea region.

## **I.II.2. TOMOGRAPHIC STUDIES**

In 1984, Oncescu inverted teleseismic events recorded by the Romanian earthquake network. His tomographic images revealed a low-velocity structure between 40 and 80km depth and a high velocity structure between 80 and 160km depth coinciding with the Vrancea intermediate depth seismicity.

This localized fast anomaly has successively been observed in several seismic tomographic studies, which have confirmed the location of the Vrancea seismic nest inside the high velocity body (Figure 7).

Most of the times, these studies have used teleseismic events registered on the Romanian permanent network or during temporary experiment (Wortel and Spakman, 2000; Martin et al., 2006; Ren et al., 2012; Zhu et al., 2012), or at bigger scale on European seismic networks (e.g. Piromallo et al., 2003). Whereas first studies were using only the P waves arrival in the inversion, later studies

have added S waves arrivals and have started to work with local and regional events (Fan et al., 1998; Tondi et al., 2009; Koulakov et al., 2010).

These seismic tomography studies often depict the high seismic velocity anomaly in significantly different ranges. Nevertheless, despite uncertainty lying under such inversions, the location of the velocity anomaly beneath the Vrancea region is robust (Figure 7). The differences from one model to another are explained directly by the characteristics of the inversion, the approximations done in the forward problem and the data processing implemented (Romanowicz, 2003).

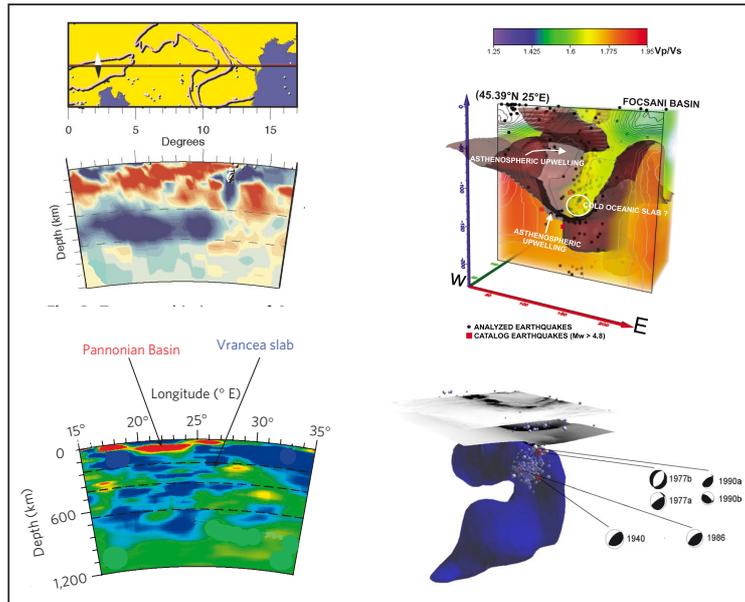


Figure 7 - Examples of seismic tomography studies. From left to right and up to down: Wortel and Spakman (2000), Tondi et al. (2009), Zhu et al. (2012) and Martin et al. (2006).

Most recent regional tomographic studies have been carried out by Ren et al. (2012, 2013) based on two important temporary projects done in the Carpathian-Pannonian basin: the Carpathian basin project from 2005 to 2007 involving 54 stations, and the South Carpathian project from 2009 to 2011 involving 56 stations.

In 2012, they retrieved some details of the upper mantle structure beneath the Carpathian-Pannonian basin from P wave teleseismic traveltimes residuals, using a 3D finite frequency kernel tomography (Figure 8). Their study was based on 1180 teleseismic earthquakes of  $M_w > 5.5$  that have been collected, between 2006 end 2011, either by the temporary stations of the two Carpathian's projects, either by the Romanian permanent seismic network, for a total of nearly 200 seismic stations. This dense coverage of the Carpathian-Pannonian basin allowed them to find interesting features about its upper mantle structure. In the Vrancea area they observed, in agreement with previous tomographic studies, a fast P velocity anomaly trending to the southern edge of the Carpathians. They defined its extension as a  $300 \times 200$  km area going from 100 to 400 km depth. Although this anomaly extends to SW of Vrancea, their tomographic images did not give any evidence for a remnant of subducted slab beneath the adjacent Eastern Carpathian. Instead the NW-NE Carpathian region, north of Vrancea, is dominated by slow anomalies.

The fast anomaly follows a nearly vertical shape from 80 to 400km depth and is bounded by two slow regions deeping to NW and SE. This particular distribution of the anomalies does not support the hypothesis of an attached slab underneath Vrancea. If the NW slow anomaly can be explained by an asthenospheric upwelling due to a delamination process, this is not the case for the SE slow anomaly. The authors thus support instead the gravitational instability model of Lorinczi and Houseman (2009) that can explain such a distribution. Another interesting outcome from their study comes from the mantle transition zone imaging that presents a fast P wave anomaly beneath the entire Carpathian area, without any connection with the upper mantle high velocity anomaly. This is interpreted as an evidence that the Vrancea structure post-dates the formation of this deeper high velocity anomaly as well as the extension of the Pannonian basin.

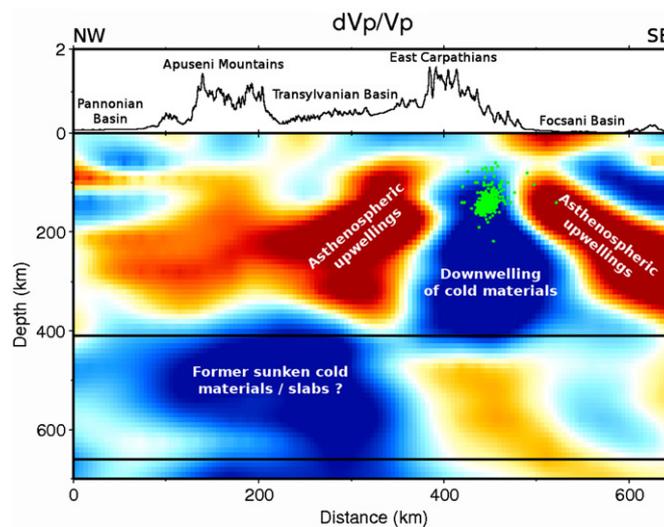


Figure 8 - NW-SE profile across the Vrancea Zone through the P-wave tomographic model of Ren et al. (2012). Green dots show the projected locations of intermediate-depth earthquakes in the Vrancea region.

In my tomographic inversion of the Vrancea lithospheric structure, the reference model is taken from the Tondi et al. (2009) tomographic study.

Using jointly seismic data from temporary experiments together with gravimetric data, the authors have built a detailed 3D tomographic model of the Vrancea lithosphere until 230km depth. Their inversion was based on the Sequential Integrated Inversion method (Tondi et al., 2000) that follows a maximum likelihood regularization scheme. Their initial reference model was based on the Popa et al. (2001) 1D velocity model of Vrancea<sup>1</sup>. Their seismic database was made of P-S first arrivals from 156 local events records collected at 26 short period stations during the CALIXTO99 experiment as well as from the seismic records from the refraction experiments VRANCEA99 (12 record sections) and VRANCEA 2001 (11 record sections). Their gravity database, instead, was based on local Bouguer anomalies data from the Institute of Geodynamics of the Romanian Academy.

<sup>1</sup> Details about the 1D reference model can be found in Figure 67 in Appendix B.

This joint inversion gave an interesting snapshot of the main crustal and upper mantle features of the Vrancea lithosphere (Figure 9). At crustal level, until 35km depth, the study has evidenced a low velocity/density SE-NW arc shaped anomaly. In the upper mantle, above 45km depth, they observed a sign change of these anomalies and an increase of the Moho depth towards SE. In particular, they identified a high  $V_p/V_s$  ratio of the area between 40 and 80km depth, collocated with the Vrancea seismic gap and associated it with a delamination process. Finally, as in several other studies, their model is characterised by a low  $V_p/V_s$  anomaly oriented N-S in the upper part (from 45 to 70km) and NW-SE in the deeper part, from 70 to 180km. They found also a coupling between the subducted slab and the unsubducted Moesian lithosphere limited to the SW part of the model.

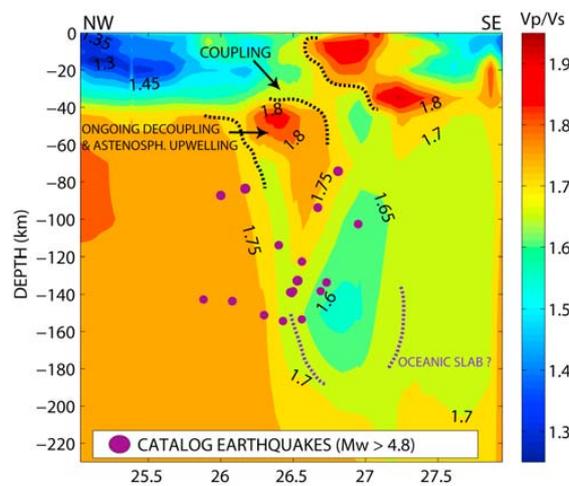


Figure 9 - NW-SE  $V_p/V_s$  ratio model cross section along VRANCEA2001 profil (from Tondi et al., 2009). The low  $V_p/V_s$  anomalies between 45 and 180km is interpreted as a downgoing slab. Instead, the high anomaly between 40 and 70km is interpreted as partial melting of the lithosphere due to a delamination process.

# II

## FULL WAVEFORM INVERSION THEORY

<b>II.I. ABOUT SEISMIC TOMOGRAPHY .....</b>	<b>21</b>
II.I.1. FROM RAY BASED APPROACH TO 3D FULL WAVEFORM TOMOGRAPHY .....	21
II.I.2. THE INVERSE PROBLEM.....	23
II.I.3. SOLVING THE FORWARD PROBLEM .....	25
II.I.3.1. The elastic wave equation.....	25
II.I.3.2. The spectral element method .....	26
II.I.4. MISFIT CRITERIA.....	29
<b>II.II. THE ADJOINT METHOD.....</b>	<b>31</b>
II.II.1. SOLVING MINIMIZATION ALGORITHMS.....	31
II.II.1.1 Introduction .....	31
II.II.1.2. Gradient based minimization algorithms .....	31
II.II.2. COMPUTING THE GRADIENT WITH THE ADJOINT METHOD .....	34
II.II.3. APPLICATION TO STRUCTURE INVERSION .....	36
<b>II.III. SOURCE INVERSION: A KEY POINT FOR TOMOGRAPHIC STUDIES .....</b>	<b>39</b>
II.III.1. SOURCE LOCATION.....	39
II.III.2. SOURCE MECHANISM .....	39



## II.I. ABOUT SEISMIC TOMOGRAPHY

---

Seismic tomography is a non-linear inverse problem based on the equation of waves motion. Its goal is to retrieve non-measurable physical Earth's properties, as seismic waves velocities, from seismograms recorded at seismic station on the surface.

### II.I.1. FROM RAY BASED APPROACH TO 3D TOMOGRAPHY

Detailed overviews of the state of the art in seismic tomography can be found on the publications of these different authors: Rawlinson et al. (2003, 2010), Tian et al. (2007), Trampert et al. (2005) and Liu and Gu (2012), who trace its evolution from its origins in 1970 to the recent trends developed in the last decade thanks to an increasing computing power.

The seismic tomography method comes from medical imaging that uses x-rays to map internal density distribution of the human body in terms of slice pictures (the word tomography means "slice picture" in the etymologic Greek). Seismic tomography thus intends to use the seismic waves generated from earthquakes or shots in the deep Earth and recorded at seismic stations to map the Earth structure.

The origin of seismic tomography is usually traced back to the late 70's, when Aki proposed the first 3D velocity model of the crust beneath California using ray-based tomography. Ten years later, Dziewonski opened the door to global tomography with the publication of the first global 3D model of the Earth's mantle (Dziewonski, 1984).

Following these pioneering efforts, a world of new applications and developments grew up together with the availability of new resources, from seismic network to data exchange facilities and computational resources. This evolution's thread is characterised by the constant desire to extract as much information as possible from as many data as possible through modelling and equation solvers that accurately depict Earth's properties, passing from 1D to 3D reference model in complex media and from first arrival seismic tomography to full waveform inversion.

Due to its simplicity, ray based tomography has been the most popular technique to carry out tomographic studies. It is based on the information given in the source-receiver travel times of seismic waves to find velocity perturbations.

Travel-time of a ray in a continuous velocity medium  $v(x)$  along the ray path  $L$  can be defined with the following non linear equation:

$$T = \int_{L(v)} \frac{1}{v(x)} dl, \text{ where } dl \text{ is a small segment of the ray path} \quad (1)$$

In an isotropic elastic medium the propagation of seismic waves is governed by the Eikonal equation:

$$\nabla_x T = \frac{1}{(v(x))^2} \quad (2)$$

This description relies on the high frequency assumption, considering weak and large-scale lateral heterogeneities: its validity applies when rays cross velocity variations of size larger than the wavelength. In this high frequency approximation, the seismic waves arrival times depend only upon the wave speed along the geometrical ray path between source and receiver and correspond to the highest frequency observable wave.

Limitations of the ray-based theory started to pop up with the notion of sensitivity kernel. Taking example on optical rays properties, it has been shown that seismic rays have Fresnel zones, defined as banana-like areas in which seismic rays can be influenced by a perturbation of the model. This introduced the concept of seismic signals sensitivity to perturbation off the classical ray-path (wave scattering) and their frequency dependence property.

In a wish to circumvent these limitations, finite-frequency theory started to be developed and implemented, using the ray theory in conjunction with the Born approximation.<sup>2</sup>

In the 2000', the studies of Marquering (1999) and Dahlen (2000), demonstrating the travel-time and amplitude frequency-dependent sensitivity, advocated the abandonment of the classical ray tomography in favour of the new-born banana-doughnut theory. This name comes from the particularity that for direct arrivals in global Earth model the sensitivity kernels looks like a banana in the vertical source receiver plane, and like a doughnut in the cross section perpendicular to this plane. These new results and conclusions have been first welcomed with skepticism by the scientific community. They indeed brought a paradox with the ray theory. Indeed, whereas in the ray theory the geometric ray path is taken as the only region where the travel-time is sensitive to Earth's structure, the results of Marquering showed the contrary. That is body wave travel-time are sensitive to heterogeneity anywhere else than along the ray path. However, this paradox – or misunderstanding- was only due to the deficiency of the ray theory to recognize the frequency dependence of the seismic waves: the sensitivity kernel of very high frequency waves is a very slender hollow, on the other hand low frequency waves are sensitive to wave speed heterogeneity farther off the ray, giving birth to their banana-doughnuts shape.

The last decade has thus been marked by the development of finite frequency techniques in 3D models and 3D sensitivity kernels computation at global and regional scale using even more realistic Earth models thanks to the growth of

---

<sup>2</sup> In Shearer (2005): “For sufficiently weak velocity and density perturbations, the single scattering Born approximation assumes that primary waves are unchanged by their passage through the scattering region. The total energy in the seismic wave-field therefore increases by the amount contained in the scattered waves and energy conservation is not obeyed. Thus this approximation is only valid when the scattered waves are much weaker than the primary waves, which is the case in the Earth when the velocity and density perturbations are relatively small (quantifying exactly how small depends upon the frequency of the waves and the source-to-receiver distance).”

In linearized inversion methods, the Born approximation is used to establish a linear relation between model perturbation and data residuals, defining the observed seismic signals as a perturbed synthetic signal.

In non linear methods, this approximation is used in the computation of the Fréchet derivatives by making the link between gradient and model perturbations, under the condition to have a starting model close enough to the true model.

powerful numerical techniques, such as the spectral element method (Komatitsch et al., 1999, 2002a,b). The finite-difference scheme, based on numerical solutions of the wave equation, including all types of waves in the synthetic seismograms, has opened the door to full waveform inversion. The latter, initiated in the early 1980's (Bamberger et al., 1982; Tarantola, 1984a,b) in seismic exploration allows to use portions or entire seismograms instead of single picks measurements to measure the misfit and minimise the objective function. Together with this new theory, several methods to deal with the inversion procedure have been developed, with the aim to adapt the tomographic schema to the data involved in the study. This has consisted in the definition of new misfit criteria to evaluate the reference model with the real data, from cross-correlation travel-time to amplitude misfit (Tromp et al., 2005) and recently, phase and envelope misfit (Fichtner et al., 2008). Owing to the high computational costs involved in the numerical solution of the wave equation, the minimization problem has to be solved iteratively through optimization algorithms, with local linear approximations. Most common algorithms include the steepest descent and the conjugate gradient algorithm (Tape et al., 2007), which both use the value of the gradient of the misfit to update the model.

Recently, the adjoint method has been used in tomographic inversion in the context of finite frequency theory, to compute efficiently the gradient of the misfit. Liu and Tromp (2006) demonstrated that the adjoint method encompasses finite frequency tomography and time reversal imaging by drawing a connection between seismic tomography, adjoint method, time reversal mirror and finite frequency banana-doughnut kernels. Using this method, the Fréchet derivatives of travel-time or amplitude misfit functions needed to build the gradient of the misfit, consist of a weighted sums of 3D banana-doughnut kernels. The adjoint method has then been expanded to other misfit criteria such as amplitude or envelope and phase misfit in several studies published in the last years (Fichtner et al., 2006a,b, 2009, 2011; Liu et al., 2006, 2008; Siemenski et al., 2007a,b; Tape et al., 2009, 2010; Tromp et al., 2008).

## II.1.2. THE INVERSE PROBLEM

When the physics of a system is known, we can use mathematics to solve the relevant equations and compute observables of the system. This is called a Forward problem. For instance, using elasticity and gravitation theory and knowing the distribution of mass and elastic parameters in the Earth, we can compute its moment of inertia, frequency of free oscillations etc.

Solving an Inverse problem is thinking the opposite way: we use a set of measurable data to retrieve the system parameters.

In geophysics, it is common to deal with a lack of information about parameters characterizing a system, as they are not directly measurable. These non-observable parameters constitute the model. Instead, information from instrumental observations refers to the data. It is thus quite common to deal with inverse problem in this field. One of the first inverse problems been posed in geophysics was on the way to determine the location and origin time of earthquakes from arrival times of seismic waves recorded by seismographs at the surface.

Solving a forward problem means predicting the observable parameters  $\mathbf{d}$  that correspond to a given model  $\mathbf{m}$  using the theoretical relation that describes the system. This theoretical prediction can be written as:

$$\mathbf{d} = g(\mathbf{m}) \quad (3)$$

Where the operator  $g$  expresses a mathematical relation of the physical system between the vector of data  $\mathbf{d}$  and the vector of model parameters  $\mathbf{m}$ .

In the following we look at the simple case of a linear inversion problem to introduce successively the non-linear one.

In the particular case of a linear relation, where  $g$  can be separated to represent a linear mapping of the model parameter into the data, this prediction can be written in the matrix form:

$$\mathbf{d} = \mathbf{G}\mathbf{m} \quad (4)$$

The inverse problem consists of finding the model parameterization  $\mathbf{m}$  that characterized the real state of the investigated problem from a set of observed data  $\mathbf{d}$ . In the best linear case, one would write it as:

$$\mathbf{m} = \mathbf{G}^{-1}\mathbf{d} \quad (5)$$

However, most of the time  $\mathbf{G}$  is not invertible. Indeed, most inversion problems are ill-posed: either underdetermined (the number of parameters is larger than that of equations) or overdetermined (the number of equations is larger than that of parameters), leading to a non invertible matrix system.

In such cases, the most general linear mapping from data to estimated model is then written in terms of the generalized inverse  $G^g$ :

$$\mathbf{m} = G^g \mathbf{d} \quad (6)$$

A common way to estimate a model is to look at the difference between the model and the data, measured with a given norm, and to minimize it. This cost function can be written as:

$$\chi = \|\mathbf{d} - \mathbf{G}\mathbf{m}\|_L \quad (7)$$

where  $\|\cdot\|_L$  stands for a given norm.

Taking for example, the least squares difference, based upon the  $L_2$  norm, the generalized inverse  $G^g$  is given by  $G^g = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T$  and the cost function is finally minimized by the following model estimate:

$$\mathbf{m} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d} \quad (8)$$

In non-linear inverse problems, the operator  $g$  cannot be separated to establish a mapping from the model parameters into the data. The point is that we cannot find explicitly the relation making the link from the data to the parameters of the model.

Thus grid-search methods or optimization methods are used to solve the non-linear inverse problem. These methods aims at minimizing a misfit function that describes how close the predicted data from a trial model fit the real data (similar to the definition of the linear case cost function). Whereas the grid search method

consists in looking over all the possible model parameterizations to identify the minimum misfit, optimization algorithms consist in starting from an initial model -sufficiently close to the real state of the system to ensure convergence- and correct it iteratively from misfit measurements until finding the best possible misfit. This minimization passes through the computation of the gradient of the misfit at each step of the model updating.

The solution of the inverse problem is thus very similar to the linear case minimization of a cost function, but this time the aim is to minimize iteratively the following function:

$$\chi = \|\mathbf{d} - g(\mathbf{m})\|$$

In seismic tomography, the forward problem is the solution of the seismic wave equation, the relation predicting the observable data from the model. Its inversion problem is based on finding the best model parameterization that minimizes a quantified difference between the observed seismic data and synthetic data computed solving the forward problem within trial models. The tomographic inversion is an ill-posed problem with some parts of the Earth's interior over-determined and other under-determined. Indeed it is characterised by an uneven distribution of source and receivers resulting in an uneven sampling of the medium by elastic waves (Trampert, 1998).

As this problem is ill-conditioned (small errors in the data causes large variations in the model) a regularization of the model is needed to avoid convergence problems (Kissling et al, 2001).

## II.1.3. SOLVING THE FORWARD PROBLEM

The inverse problem describing seismic tomography relies on the update of a physical Earth model through the minimization of a given misfit describing the differences between observed seismic data and corresponding synthetics waveforms. These later are obtained by solving the equation of motion relating the displacement field to external forces and to physical Earth properties such as the density and the elastic parameters.

### II.1.3.1. The elastic wave equation

The elastic wave equation in a continuous medium can be written as a function of the displacement vector  $\mathbf{u}$ , the Cauchy strain tensor  $\boldsymbol{\sigma}$  and a seismic source excitation vector  $\mathbf{f}$ .

$$\rho \ddot{\mathbf{u}} - \nabla \cdot \boldsymbol{\sigma} = \mathbf{f}, \text{ where } \rho \text{ is the density distribution} \quad (9)$$

Using the Hook's law and the definition of the strain tensor  $\mathbf{e}$ ,

$$\text{Hook's law: } \sigma_{ij} = c_{ijkl} e_{kl}, \text{ where } \mathbf{c} \text{ is the stiffness tensor} \quad (10)$$

$$e_{kl} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \text{ strain tensor}$$

this equation can be rewritten only in terms of displacement  $\mathbf{u}$ . On a single component this gives:

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial u_l}{\partial x_k} \right) = f_i \quad (11)$$

In 1D Earth models, one can use semi-analytical techniques to calculate the wave field generated by a point source. Instead, the 3D elastic wave equation cannot be solved in this way in realistic Earth's models and one has to rely on numerical solutions based on an approximate solution of the equation.

### II.I.3.2. The spectral element method

In the last three decades, together with the computational power growth, several numerical methods have been developed for the solution of the 3D elastic wave equation in complex geological models. These include finite-difference schemes (Igel et al., 2002), pseudo spectral (Fornberg et al., 1987; Carcione, 1990), finite elements (Lysmer et al., 1972) or discontinuous Galerkin methods (Dumbser et al., 2006). For seismic wave propagation on continental and global scales, the spectral element method has proven to be a good compromise between accuracy and computational costs (Komatitsch and Tromp, 2002a,b; Quin et al., 2009). Developed in computational dynamics by Patera in 1984, it has been introduced later to 3D elastodynamics (Seriani, 1998; Komatitsch et al. 1999, 2002a,b and Chaljub et al., 2000).

The Spectral Element Method is based upon the weak formulation of the wave equation and its solution through spatial discretization and a Gauss Lobatto Legendre integration quadrature scheme. The entire procedure can be divided in four main computational steps whose equations are detailed in Figure 10.

- 1) The weak formulation of the equation of motion is obtained by taking the dot product of the wave equation with a test vector, integrating it by parts and finally imposing boundary conditions. This formulation, used also in finite element and Discontinuous Galerkin methods, benefits from the advantage that free surface boundary conditions are automatically satisfied in comparison with finite difference method.
- 2) To solve this new equation a discretization of the domain is needed: integrals are subdivided in terms of smaller integrals over the volume and surface element. This consists in the decomposition of the computational domain  $G$  into  $n_e$  non-overlapping elements  $G_e$ , which are mapped to a reference cube  $\Lambda = [-1, 1]^3$  via an invertible transformation:

$$\mathcal{F}_e : [-1, 1]^3 \rightarrow G_e \quad \mathbf{x}(\boldsymbol{\xi}) = \mathcal{F}_e(\boldsymbol{\xi}) \quad \boldsymbol{\xi}(\mathbf{x}) = \mathcal{F}_e^{-1}(\mathbf{x})$$

- 3) The displacement field is represented onto a grid point in each reference element through an interpolation scheme using a polynomial approximation in terms of Lagrange polynomials of degree 4 to 10. Each integral of the weak equation of motion is interpolated in terms of triple product of Lagrange polynomials.
- 4) Finally a Gauss Lobatto Legendre quadrature is undertaken to evaluate these integrals at elemental level.

The equation can finally be written in terms of a mass matrix, a stiffness matrix and a source term and surface matrix. The diagonality of the mass matrix allows solving easily the system.

**3D elastic wave equation**

$$\rho \ddot{\mathbf{u}} - \nabla \cdot \boldsymbol{\sigma} = \mathbf{f}$$

with boundary and initial conditions

$$\mathbf{n} \cdot \boldsymbol{\sigma} = \mathbf{0} \quad \mathbf{u}(t = 0) = \mathbf{0}$$

on free surface the traction is null  $\frac{\partial \mathbf{u}}{\partial t}(t = 0) = \mathbf{0}$

$\mathbf{n}$  denotes the unit outward normal on the free surface

**weak formulation**

$$\int_G \rho \mathbf{w} \cdot \ddot{\mathbf{u}} d^3 \mathbf{x} + \int_G \nabla \mathbf{w} \cdot \boldsymbol{\sigma} d^3 \mathbf{x} = \int_G \mathbf{w} \cdot \mathbf{f} d^3 \mathbf{x}$$

with implicit free surface boundary conditions and initial conditions

$$\int_G \rho \mathbf{w} \cdot \mathbf{u}(t = 0) d^3 \mathbf{x} = \mathbf{0}$$

$$\int_G \rho \mathbf{w} \cdot \dot{\mathbf{u}}(t = 0) d^3 \mathbf{x} = \mathbf{0}$$

**discretization**

decompose  $G$  into  $n_e$  disjoint subdomains  $G_e$  called "elements"  
 map each element on the reference cube  $\mathcal{F}_e : [-1, 1]^3 \rightarrow G_e \quad \mathbf{x}(\boldsymbol{\xi}) = \mathcal{F}_e(\boldsymbol{\xi})$   
 weak equation solved for each element individually

$$\int_{-1}^1 \rho \mathbf{w} \cdot \ddot{\mathbf{u}} J(\boldsymbol{\xi}) d^3 \boldsymbol{\xi} + \int_{-1}^1 \nabla \mathbf{w} \cdot \boldsymbol{\sigma} J(\boldsymbol{\xi}) d^3 \boldsymbol{\xi} = \int_{-1}^1 \mathbf{w} \cdot \mathbf{f} J(\boldsymbol{\xi}) d^3 \boldsymbol{\xi}$$

where  $J$  is the Jacobian of  $\mathcal{F}_e$

**interpolation**

the displacement field  $\mathbf{u}$  is approximated with Lagrange polynomials of degree  $N$

$$\mathbf{u}_p(\mathbf{x}, t)_{G_e} = \sum_{i,j,k=1}^{N+1} u_p^{ijk}(t) \psi_{ijk}(\mathbf{x}) \quad \psi_{jki}^e[\mathbf{x}(\boldsymbol{\xi})] = l_i^{(N)}(\xi_1) l_j^{(N)}(\xi_2) l_k^{(N)}(\xi_3),$$

$\boldsymbol{\xi} \in [-1, 1]$

**integration**

numerical integration with Gauss Lobatto Legendre quadrature

matrix system to solve  $\mathbf{M} \cdot \dot{\mathbf{u}} + \mathbf{K} \cdot \mathbf{u} = \mathbf{f}$

with  $\mathbf{M}$  diagonal mass matrix and  $\mathbf{K}$  stiffness matrix

Figure 10 - Sum up of the spectral element method.

## II.I.4. MISFIT CRITERIA

The choice of the misfit is related to the type of seismic tomography applied. This choice is also closely linked to the type of data exploitable for the inversion and their quality.

Common misfit functions are based on cross-correlation travel-time measurements, amplitude variations (Tromp et al., 2005; Tape et al., 2007; Tian et al., 2007; Fichtner et al., 2008) or waveform differences (Tromp et al., 2008). Their mathematical definitions can be found in Figure 11.

Cross-correlation travel-time and amplitude measurements provide robust estimates of time and amplitude differences in selected seismic phases of seismogram. If made at different frequency ranges, they can characterise all information contained in the seismic signal. Nevertheless, their applicability is limited to scenarios where seismic waveforms have similar shape and single phases are clearly separable. As they depend on event magnitude it may be also necessary to account for source time function in the computation of these misfits.

The waveform difference misfit, defined by differences between observed and synthetic seismograms and applicable to the whole seismogram, is often used when the quality of the data does not allow isolating particular phases.

Mixing phase and amplitude information into a single observable, the waveform misfit carries out the non-linearity of the seismic waves amplitude to Earth's structure in the vicinity of the sources and receivers. Being quite a sensitive method as its value is mostly influenced by large amplitude waveforms, allowing outliers to become predominant, attention has to be paid during its implementation. This misfit is not well indicated in the frame of gradient-based minimisation algorithm as it presents multiple local minima (see III.VI, Variable Projection Method, for some illustrations).

Recently a new kind of misfit functions has been used in tomographic inversion, which separates data in terms of phase and amplitude (envelope) information (Fichtner et al., 2009; Bozdogan et al., 2011). Whereas amplitude is affected by several factors like source properties and local structure at source and receiver, the phase is directly related to seismic wave velocity and thus it seems better to look at these parameters separately. Its application is however reduced to high quality data as the sensitivity to noise is high and to similar waveforms to avoid discontinuity of the phase.

The choice of the misfit function is highly dependent on the data quality and frequency ranges of investigation for the tomographic inversion. Besides in the frame of full waveform tomography using the adjoint method, it affects the construction of the sensitivity kernel and adjoint sources, as for each kind of misfit function one has to define a given adjoint problem. Examples of adjoint sensitivity kernels based on cross-correlation travel-time measurements, amplitude and waveform misfit can be found in Liu and Tromp (2006, 2008). Details for kernel computation based on phase and envelope misfit can be found in Fichtner et al. (2009).

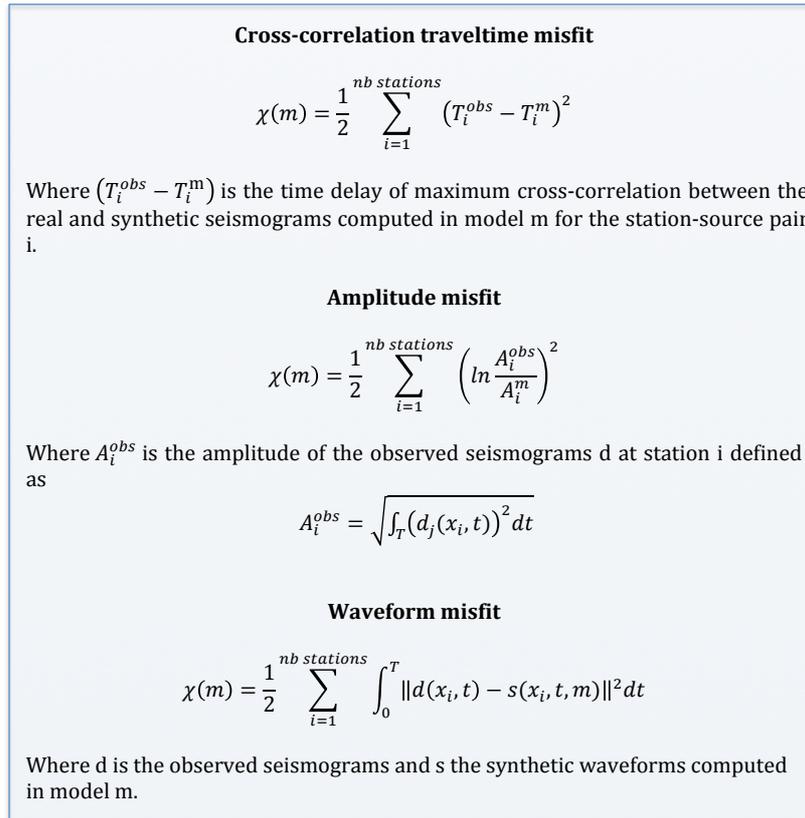


Figure 11 - Common misfit functions used for seismic tomography full waveform inversion.

## II.II. THE ADJOINT METHOD

---

### II.II.1. SOLVING MINIMIZATION ALGORITHMS

#### II.II.1.1 Introduction

Two main categories of optimization algorithms are used to solve minimization problems. The distinction is based on the use of the gradient of the objective function.

In non-gradient based methods, such as grid search method, only objective function evaluations are used to find the optimum point. If such method may allow finding the global minimum, they require a large number of computations as they intend to test as many solutions as possible.

Instead, gradient based approaches involve the computation of continuous first derivatives of the objective function and possibly higher derivatives. Compared with non-gradient based method they generally require less computation to converge to an optimum point, even if in this case the convergence to the global minimum is not guaranteed, but can go instead to a local one.

The simplest gradient-based methods are based solely on the first derivatives of the objective function (the gradient) and usually require more iterations than methods based on also second order derivatives (Hessian).

In case of large-scale inversion problem such as tomographic inversion, involving numerous model parameters variables, non-gradient based methods are not efficient and Gradient based algorithms are used instead to find a minimum of the misfit. In such a case, the search starts from a given starting point (initial model) and use the misfit and its derivatives values at this point into test a following point (model update) that gives a lower misfit value. A stopping criterion is then fixed into close the search when the misfit value is close enough to zero.

These optimization algorithms involve the computation of the gradient of the misfit, that is the perturbation of the misfit function due to a perturbation of the models parameters, and sometimes, higher order derivatives depending on the method used. Computing the partial derivatives of the misfit regarding each model parameters can become a tedious task. To deal with it, an alternative method has been developed that allows getting access to the gradient of the misfit function without computing each derivative, it is the adjoint method.

#### II.II.1.2. Gradient based minimization algorithms

A seismic tomographic inversion problem is based on the minimization of a physically meaningful quantified difference between observed and synthetic waveforms computed for a given Earth model. This difference is expressed in terms of a misfit function that depends non-linearly on the model parameters.

The goal of the inversion procedure is to find iteratively the best model parameter values that lead to a minimum of the misfit function. That is, we start from a given reference model and we successively compute model updates that converge to a minimum of the misfit function.

Different optimization algorithms can be used to solve the inverse problem iteratively. Their efficiency and usability depend on the functions involved in the optimization scheme that is the order of the derivatives used. Algorithms, such as Newton's one, involving the Hessian matrix of second order derivatives, is faster than an only gradient based one, such as the conjugate gradient method, require more computations and can be hardly implemented. In the following, I will focus on two methods involving only the computation of the gradient of the misfit function: the steepest descent and the conjugate gradient algorithm.

### II.II.1.2.1. Linear search method: steepest descent

The main principle of a linear search method is based on choosing a direction from the current state of the objective function along which to search for the next iteration, regarding a given step length to advance the optimization.

Given a misfit function  $\chi$  for a model  $m$  of  $n$  parameters, the optimization problem can be written as :

$$\min_m \chi(m) \quad (12)$$

The misfit function can be written in terms of Taylor expansion using a search direction  $p$  and a step length  $l$  into make appearing the first and second derivatives functions.

First order partial derivatives are defined in the gradient vector  $g = \nabla\chi = [g_1 \ g_i \ g_n]$  of the misfit function  $\chi$  where  $i$  points to a given model parameter to invert.

$$g_i = \frac{\partial\chi}{\partial m_i} \quad (13)$$

Second order partial derivatives are defined in the  $n \times n$  Hessian matrix:

$$H = \nabla^2\chi = \begin{bmatrix} H_{11} & H_{1i} & H_{1n} \\ H_{i1} & H_{ii} & H_{in} \\ H_{n1} & H_{ni} & H_{nn} \end{bmatrix} \quad (14)$$

$$H_{in} = \frac{\partial^2\chi}{\partial m_i \partial m_n}$$

The Taylor expansion of the misfit function ( 15 ) is then written in terms of the gradient and the Hessian matrix.

$$\chi(m + lp) \approx \chi(m) + lp^T \nabla\chi + \frac{1}{2} l^2 p^T \nabla^2\chi p + \dots \quad (15)$$

Where the maximum descent is found at

$$p = - \frac{\nabla\chi}{\|\nabla\chi\|}$$

Thus, the steepest descent optimization algorithm, defined by Cauchy in 1847, is characterised by setting the search direction  $p$  to the negative of the gradient direction of the misfit function, at every iteration. The model update can be written as:

$$m_{n+1} = m_n + \delta m, \text{ where } \delta m = lp = -lg \quad (16)$$

The optimal step length  $l$  for each iteration is so that

$$\tilde{\chi}(l) = \min_{l>0} \chi(m + lp) \quad (17)$$

It is usually found by line search over a small number of trial step length  $l^{\text{test}}$  starting for example from

$$l_n^{\text{test}} = \frac{-2\chi(m_n)}{g_n} \quad (18)$$

Then, by computing the test model  $m_n^{\text{test}} = m_n + l_n^{\text{test}} p_n$  and calculating the corresponding misfit and gradient measurements  $\chi(m_n^{\text{test}})$  and  $g_n^{\text{test}}$ . After interpolating  $\chi(m_n + lp_n)$  by quadratic or cubic polynomial, one can find the analytical minimum of this polynomial and take it as an approximation of the optimal step length  $l$  (Tape et al., 2007).

The choice of the trial step length is done most efficiently with the help of intuition with the scope to get physically plausible models.

### II.II.1.2.2. The conjugate gradient method

The first nonlinear conjugate gradient method was proposed by Fletcher and Reeves in 1964. It is built within an iterative descent scheme, identical to the steepest descent with some modifications. The conjugate gradient method uses the recurrence relation:

$$m_{n+1} = m_n + l_n p_n \quad (19)$$

Where the positive step length  $l_n$  is found by line search, and the search directions  $p_n$  are generated from the rule  $p_n = \beta_{n+1} p_n - g_{n+1}$ , with  $p_0 = -g_0$  as defined in the steepest descent method. If  $\|p_{n+1}\| < \epsilon$ , where  $\epsilon$  is a suitably small number, then  $m_{n+1}$  is the best model, otherwise the iteration cycle has to continue.

The  $\beta_n$  term is a conjugate gradient update parameter, whose formulation changes according to the method employed. For example, Fletcher and Reeves propose this kind of formulation:

$$\beta_{n+1} = \frac{\nabla\chi(m_{n+1}) \cdot \nabla\chi(m_{n+1})}{\nabla\chi(m_n) \cdot \nabla\chi(m_n)} \quad (20)$$

As seen in this part, the main step in these optimization algorithms is the computation of the gradient of the misfit function, defined as the vector of first order derivatives of the misfit function, also named Fréchet derivatives. This computation becomes a heavy task for a complex problem involving a lot of variables. Fortunately, we can use substitutive methods that allow dealing with such computation such as the adjoint method.

## II.II.2. COMPUTING THE GRADIENT WITH THE ADJOINT METHOD

Adjoint state methods have been first introduced in inverse problem theory in the early 1970's by Chavent to compute the gradient of a functional depending on state variables that were solutions of a forward equation.

In seismic tomography, the computation of the gradient of the misfit that quantifies the discrepancies between observed seismic waveforms and waveforms computed from a given physical Earth model, requires the computation of the derivatives of the wave field with respect to model parameters. Dealing with computationally expensive solution of the seismic wave equation and several unknown model parameters, this can become a laborious task. This is where the adjoint method comes into play, enabling an efficient computation of the gradient of the misfit.

Adjoint method is built on the properties given by the adjoint operator of a given operator representative of a relation between different variables.

In the mathematical words used by Tarantola in 1984 to introduce this notion, we can write the following state: let  $\mathbb{M}$  and  $\mathbb{D}$  be two Hilbert spaces with scalar products  $\langle \cdot, \cdot \rangle$  and let  $G$  be an operator from  $\mathbb{M}$  into  $\mathbb{D}$ . The adjoint of  $G$ , denoted  $G^*$ , is the operator from  $\mathbb{D}$  into  $\mathbb{M}$  defined by:

$$\langle Gm, d \rangle = \langle m, G^*d \rangle \text{ for all } m \text{ and } d \text{ in the Hilbert spaces } \mathbb{M} \text{ and } \mathbb{D} \quad (21)$$

Let now take a physical observable  $u$ , that depends on the Earth parameters, the space and the time. It is linked to external sources and model parameters by the wave equation through the operator  $L$  so that:

$$L(u, m) = f \quad (22)$$

We are looking for the gradient of the misfit function  $\chi(m)$  with respect to model parameters  $\nabla_m \chi \delta m$ . Computing this gradient directly – for example by finite difference- would involve as many explicit computations of the misfit function as they are parameters  $m$ , which can become practically impossible.

From the chain rule, the search for  $\nabla_m \chi \delta m$  can be expressed as:

$$\nabla_m \chi \delta m = \nabla_u \chi \delta u \quad (23)$$

Here, the computation of the derivatives of the wave field  $\delta u$  with respect to the model parameters is an hard task. However, this is the point where the adjoint method comes to play its role.

To introduce the principle of the adjoint method, we use the augmented functional method (Plessix, 2006), also called associated Lagrangian (Tromp et al., 2005; Liu and Tromp, 2006). We define a function  $A$  that is the sum of the wave equation and the misfit equation, and look at its derivative using (23).

$$A = L(u, m) + \chi(m) \quad (24)$$

$$\nabla A = \nabla L(u, m) + \nabla_u \chi \delta u \quad (25)$$

The derivative of the wave equation  $\nabla L(u, m)$  is developed and written in terms of a scalar product, introducing a test function  $\varphi$ .

$$\nabla L(u, m) = \nabla_u L(u, m)\delta u + \nabla_m L(u, m)\delta m = 0 \quad (26)$$

$$\langle \nabla_u L(u, m)\delta u, \varphi \rangle + \langle \nabla_m L(u, m)\delta m, \varphi \rangle = 0 \quad (27)$$

At this point, the property of the adjoint operator is used to isolate  $\delta u$  from the rest of the variables:

$$\langle \nabla_u L(u, m)\delta u, \varphi \rangle = \langle \delta u, \nabla_u L^*(u, m)\varphi \rangle \quad (28)$$

$$\langle \nabla_m L(u, m)\delta m, \varphi \rangle = \langle \delta m, \nabla_m L^*(u, m)\varphi \rangle \quad (29)$$

A new derivative of the wave equation ( 26 ) is obtained:

$$\nabla L(u, m) = \langle \delta u, \nabla_u L^*(u, m)\varphi \rangle + \langle \delta m, \nabla_m L^*(u, m)\varphi \rangle = 0 \quad (30)$$

Going back to the derivatives of the augmented functional  $\nabla A$ , we now have:

$$\nabla A = \nabla L(u, m) + \nabla_u \chi \delta u = \nabla_m \chi \delta m \quad (31)$$

$$\nabla_m \chi \delta m = \langle \delta u, \nabla_u L^*(u, m)\varphi \rangle + \langle \delta m, \nabla_m L^*(u, m)\varphi \rangle + \langle \delta u, \nabla_u \chi \rangle \quad (32)$$

$$\nabla_m \chi \delta m = \langle \delta u, \nabla_u L^*(u, m)\varphi + \nabla_u \chi \rangle + \langle \delta m, \nabla_m L^*(u, m)\varphi \rangle \quad (33)$$

To avoid the computation of  $\delta u$ , we finally have to set:

$$\langle \nabla_u L^*(u, m)\varphi + \nabla_u \chi \rangle = 0 \quad (34)$$

This is the adjoint equation, whose solution is the adjoint wave field  $\varphi$ . The whole adjoint problem is formed by subsidiary conditions, required for the existence of the adjoint operator ,and this system of equations:

$$\langle \nabla_u L^*(u, m)\varphi + \nabla_u \chi \rangle = 0 \quad (35)$$

$$\nabla_m \chi \delta m = \langle \delta m, \nabla_m L^*(u, m)\varphi \rangle$$

Thus, the computation of the gradient comes from the solution of the adjoint problem, and can be done for any  $\delta m$  without the explicit knowledge of  $\delta u$ .

The adjoint problem can be solved for all types of linear wave equation operator  $L$ . The least squares objective function is one of them, and is widely used in geophysical application allowing to implement the adjoint method in seismic tomography problems.

Solving the adjoint problem for the seismic wave equation using a least squares misfit is equivalent to the solution of the seismic wave equation under a particular set of boundary and temporal conditions. The point is that, in the adjoint problem, the temporal conditions are at the terminal state whereas they are at the initial one in the direct problem. Besides, the source term of the adjoint equation is defined by the negative of the residual. The whole adjoint problem solution can be thus interpreted as the propagation of the residual backward in time: we define a new system of equations where the source terms

are changed into residuals between synthetic and data and the wave equation is solved in reverse time.

The adjoint method starts being quite used in tomographic studies. Several detailed descriptions and applications can be found in the works of Plessix (2006), Fichtner et al. (2009, 2011), Tape et al. (2007, 2009, 2010) and Tromp et al. (2008, 2010).

### II.II.3. APPLICATION TO STRUCTURE INVERSION

Structure inversion aims to compute Earth's seismic velocity and densities using information contained in seismic records.

Taking a travel time misfit function in a frame of a full waveform inversion, I will show in this part how the adjoint method comes into play in the computation of the gradient of the misfit, and its link with the sensitivity kernels at the base of the finite-frequency theory.

The travel-time misfit and its variations are expressed as ( 36 ) and ( 37 ), with the sign convention such that a negative travel-time indicates a delay in the synthetic arrivals  $T_i^m$  relative to the observed arrivals  $T_i^{obs}$ .

$$\chi(m) = \frac{1}{2} \sum_{i=1}^n (T_i^{obs} - T_i^m)^2 \quad (36)$$

$$\delta\chi(m) = - \sum_{i=1}^n (T_i^{obs} - T_i^m) \delta T_i \quad (37)$$

where  $n$  is the number of stations

Under the finite frequency assumption, the travel-time anomaly  $\delta T_i$  can be expressed in terms of a wave speed perturbation  $\delta lnc$  and a finite frequency sensitivity kernel  $K_i$ , also called banana-doughnut kernel:

$$\delta T_i = \int_V K_i \delta lnc dx^3 \quad (38)$$

The variations of the misfit can thus be written in terms of a sum of banana-doughnuts kernels (or misfit kernel  $\mathcal{K}$ ) and the speed perturbation:

$$\delta\chi(m) = - \int_V \sum_{i=1}^n (T_i^{obs} - T_i^m) K_i \delta lnc dx^3 \quad (39)$$

$$\delta\chi(m) = \int_V \mathcal{K} \delta lnc dx^3 \quad \text{with} \quad \mathcal{K} = - \sum_{i=1}^n (T_i^{obs} - T_i^m) K_i$$

$\mathcal{K}$  the misfit kernel is defined as a weighted sum of banana-doughnut kernel  $K_i$ , where the weights are the travel-time residuals.

The model is discretized and expanded in basis functions  $B_i$  to solve the inverse problem ( 40 ). In spectral element method, the basis functions are Lagrange polynomials.

$$\delta lnc = \sum_{k=1}^N \delta m_k B_k(x), \text{ where } N \text{ is the number of model parameters} \quad (40)$$

The variations of the misfit can be thus written in terms of model perturbations:

$$\delta \chi(m) = \sum_{i=1}^N \int_V \mathcal{K} \delta m_k B_k(x) dx^3 \quad (41)$$

The gradient of the misfit  $g = \frac{\partial \chi}{\partial m}$  can thus be expressed in terms of misfit kernel:

$$g = \sum_{k=1}^N g_k \quad g_k = \frac{\partial \chi}{\partial m_k} = \int_V \mathcal{K} B_k(x) dx^3 \quad (42)$$

Following this scheme, we see that the gradient of the misfit can be obtained from the computation of the misfit kernel ( 42 ).

At this step, the adjoint method allows finding an analytical expression of the misfit kernel in terms of traveltime adjoint field. One is exempt to compute the misfit kernel from the weighted sum of individual banana–doughnut kernels for each source–station pair. Instead, the misfit kernel is constructed from the interaction between a forward and an adjoint wave field, which simultaneously propagate backward seismic waves from adjoint sources containing the residual measurements (Figure 12). By this way, building the misfit kernel requires only two simulations per earthquake (Tromp et al., 2005):

- 1) a forward simulation which solves the equation of motion for each source-station pairs in the model
- 2) an adjoint simulation, which solves simultaneously the equation of motion for each source-adjoint source pairs in the model and computes the interaction field between forward and adjoint wave fields. Adjoint sources are constructed by time reversing the synthetic velocity recorded at the station and multiplying by residual. Examples of adjoint source construction for different misfit function are indicated in the articles of Tromp et al. (2005) or Fichtner et al. (2009).

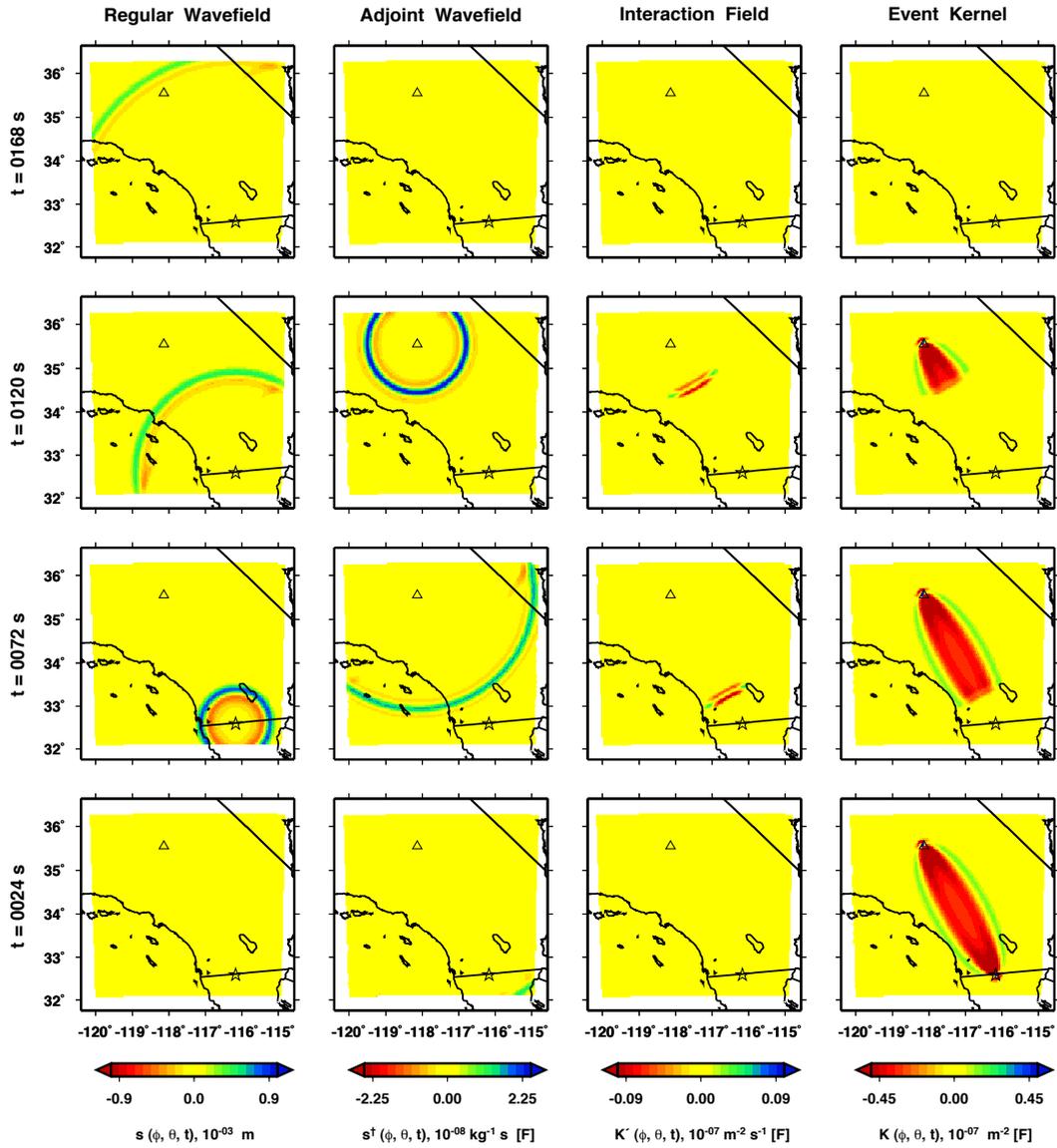


Figure 12 - Construction of a cross-correlation travel-time sensitivity kernel from Tape et al. (2007). Sources and receivers are represented by stars and triangles respectively. Each row represents a different time step. The kernel is constructed via the interaction between the forward wave-field (first column) and the adjoint one (second column). The interaction field (third column) is the instantaneous product of these two wave-fields, which is integrated to form the sensitivity kernel (fourth column).

## II.III. SOURCE INVERSION: A KEY POINT FOR TOMOGRAPHIC STUDIES

---

The source of an earthquake can be fully defined by the location of its hypocentre, its source time function and its fault plane solution -or moment tensor- that describes the physical process in play. Seismograms, containing information about the source and seismic waves path, are used to estimate these parameters. A source inversion would consist in solving two inversion problems: a non-linear problem to find the location of an earthquake, and a linear one to estimate its physical mechanism.

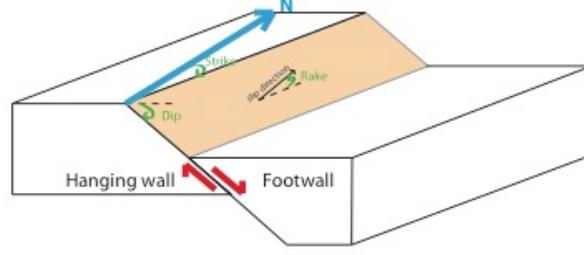
### II.III.1. SOURCE LOCATION

Source location is a key point in seismology and has become even more important with the growth of Earthquake Early Warning Systems that require quick and reliable earthquake location. Such requirements now are possible with the growth of computational power and the development of new methods.

Before wide computer availability, earthquake location was done manually using the “circle and chord method” or the Wadati diagram. Now, the estimate of hypocentre location and origin time can be divided in two main groups: grid search methods and iterative methods. In grid search method (Lomax, 2000) one seeks to minimize the least square of the travel-time residuals between observed arrival times and synthetic arrival times computed in a given Earth model over all possible location and origin time. Instead, iterative methods are based on linearizing the inversion problem starting from an initial solution for the unknown parameters.

### II.III.2. SOURCE MECHANISM

The description of the physics of seismic sources is a major research interest in seismology. Earthquakes mechanisms are described by their focal mechanisms - or fault plane solutions- under the assumption that the rupture occurs along a planar fault surface. The focal mechanism is defined by three angles describing the fault orientation (Figure 13). The strike  $\phi$  ( $0^\circ$  to  $360^\circ$  clockwise from north) is the angle of the fault with respect to North, the dip  $\delta$  ( $0^\circ$  to  $90^\circ$  against the horizontal) is the pending angle of the fault plane and the rake  $\lambda$  ( $-180^\circ$  to  $+180^\circ$  against the horizontal) is the direction of slip on the fault. The fault plane solution, together with the information about the static seismic moment  $M_0$ , can be represented by the seismic moment tensor  $M_{ij}$ . The latter describes the stress glut due to the dislocation taking place on the fault. Thus, methods for the estimate of the physics of seismic source differ between fault plane solution search and moment tensor estimation.



**Figure 13 - Parameters describing the source mechanism: strike, dip and rake angles defined on the hanging wall of the two moving blocks. Example of a normal fault characterized by a positive rake defined as the angle from the strike direction to the slip direction.**

A usual way to deal with the fault plane solution consists in using the polarity of the first seismic arrivals and projects them on the focal sphere (FPFIT, Reasenber and Oppenheimer, 1985). In modern seismology the fault plane solution is also found using grid search algorithms over the three angles, by minimizing the misfit between observed waveforms and synthetic seismograms computed with a source model defined by a given combination of (strike, dip, rake) angles (Grid3D, Liu 2004).

From the fault plane solution, one can define the moment tensor solution using the relation between the moment tensor components and the three focal mechanism angles.

A common approach for moment tensor inversion is the approximation of seismic sources by an equivalent forces model that corresponds to the linear wave equation neglecting non-linear effects in the near source region (Aki and Richards, 1980). These equivalent forces are defined as producing displacements at the Earth's surface that are identical to those from the actual forces in play.

Starting from the representation theorem for seismic sources from Aki Richards (1980), the displacement generated by a point source and recorded at stations on surface can be expressed in terms of a linear combination of time dependent moment tensor convolved with the derivative of the Green function with respect to spatial coordinates. The representation theorem for seismic sources relates the displacement  $u_n$  due to a distribution of equivalent body force densities  $f_k$  with the components  $\mathcal{G}_{nk}$  of its Green function.

$$u_n(x, t) = \int_{-\infty}^{\infty} \int_V \mathcal{G}_{nk}(x, t; r, \bar{t}) f_k(r, \bar{t}) dV(r) d\bar{t} \tag{43}$$

where V is the source volume

Expanding the Green's function into a Taylor series around the centroid  $r = \xi$ , and introducing the time dependent force moment tensor, ( 43 ) can be rewritten in terms of a convolution:

$$u_n(x, t) = \sum_{m=1}^{\infty} \frac{1}{m!} \mathcal{G}_{nk,j_1..m}(x, t; \xi, \bar{t}) * M_{kj_1..m}(\xi, \bar{t}) \tag{44}$$

Assuming that all components of the time dependent seismic moment tensor have the same time dependence  $s(\bar{t})$ , this can be simplified in:

$$u_n(x, t) = M_{kj} [\mathcal{G}_{nk,j} * s(\bar{t})], \quad (45)$$

where  $M_{kj}$  are the component of the seismic moment tensor

It results that the displacement is a linear function of the moment tensor elements and the terms in brackets.

The moment tensor  $M_{kj}$  can be decomposed into six elementary moment tensor  $M_n$  (Kikushi and Kanamori, 1991). These elementary tensors consist of five double couple sources and an explosive source.

$$M_{kj} = \sum_{n=1}^6 a_n M_n \quad (46)$$

The best estimation for the coefficients  $a_n$  can be obtained by minimizing the misfit between observed seismic waveforms and synthetic signals computed in a given Earth model and written as a linear combination of elementary tensors.

$$\chi = \sum_{j=1}^n \int \left[ d_j(t) - \sum_{n=1}^6 a_n \mathcal{G}_{jn,k}(x, t; \xi, \bar{t}) \right]^2 dt \quad (47)$$

In the last years, several methods for source inversion have been developed taking advantage of growing computer resources enabling the computation of synthetic seismograms in even more realistic Earth models.

Given that the centroid of the stress gives a better position for the equivalent point source than the hypocentre, which describes only the rupture initialisation, it has become used to describe seismic sources parameters at this point. The CMT algorithm developed by Dziewonski (1983) form the basis of numerous catalogues (CMT global project, Ekstrom et al., 2012). Recently a CMT full waveform inversion (Liu, 2004, Kim et al., 2011) poses and solves the source inversion problem as a unique problem, solving for location and physical parameter in a single full waveform inversion frame.

The source location procedure is influenced by the velocity model assumption used to compute travel-times. The accuracy of the model used impacts on the estimate of the source mechanics and further in the tomographic inversion. It is thus usual to do a source inversion inside the proper reference model and data coverage before starting the structure inversion or at least correct them at a given update of the velocity model. Valentine and Woodhouse (2010) suggested a unified approach to source and structure inversion to ensure self-consistency during the tomographic inversion. This is a notion I have approached in the following partIV through the Variable Projection Method (Golub and Pereyra, 1973) and its possible application to seismic tomography. Although this solution has not been implemented in my full-waveform inversion of the Vrancea region, this part has nonetheless a strong pedagogical value. Indeed, it gives a detailed introduction to the Variable Projection Method with its application in a simplified environment and illustrates clearly some key notions raised in this part.



# III

## AN INSIGHT INTO THE VARIABLE PROJECTION METHOD

<b>III.I. INTRODUCTION</b> .....	<b>45</b>
III.I.1. MOTIVATIONS.....	45
III.I.2. THEORY.....	46
III.I.3. SOLVING A JOINT SOURCE SIGNAL AND LOCATION INVERSION PROBLEM WITH THE VPM.....	47
<b>III.II. COMPUTE SYNTHETIC WAVEFORMS IN AN HOMOGENEOUS ISOTROPIC ACOUSTIC MEDIUM</b> .....	<b>49</b>
III.II.1. GREEN'S FUNCTION.....	49
III.II.2. WAVELET SIGNAL.....	51
<b>III.III. THE SOURCE SIGNAL PROBLEM</b> .....	<b>53</b>
<b>III.IV. THE SOURCE LOCATION PROBLEM</b> .....	<b>55</b>
III.IV.1. MISFITS EXAMPLES.....	55
III.IV.1.1. Travel-time misfit.....	55
III.IV.1.2. Waveform misfit.....	56
III.IV.2. OPTIMIZATION ALGORITHMS.....	56
III.IV.2.1. Gradient computation.....	57
III.IV.2.2. Iteration.....	59
<b>III.V. APPLICATION</b> .....	<b>63</b>



## III.I. INTRODUCTION

---

### III.I.1. MOTIVATIONS

Seismic tomography inversion, aimed at finding Earth's seismic velocity distributions, strongly depends on source parameters. Indeed, this method based on waveforms contents carries out in the inversion problem the sensitivity of seismic waveforms to both structure and source patterns. Consequently, accurate determinations of event locations and focal mechanisms are required for any tomographic inversion.

Source parameters are usually obtained prior to the tomographic inversion, from a simplified/global Earth model, such as the solution distributed by the global CMT (formerly Harvard) project. However, the inaccuracy of the tomographic model used for source parameterization can come out in tomographic inversion results. It is thus on common use to treat the source parameterization alongside with the velocity inversion to insure consistency of the tomographic study.

Sequential source and structure inversion consisting of a step-by-step update of the source and structure parameter, each time fixing one of the two is commonly used (Hara 2000). However, this type of inversion leads to a slow rate of convergence and can retain bias from the initial model (Valentine and Woodhouse, 2010). Combining source and structure inversion into a single inverse problem seems to be a better approach. This has previously been investigated by several authors in related problems for structure and earthquakes location simultaneous inversions (Spencer 1980, Harvey 1998, Tape 2007, Morelli and Dziewonski 1991). Recently, an implementation based on inverse problem linearization and regularization has been proposed by Valentine and Woodhouse (2010) to retrieve complete source parameters (Moment tensor and location) together with velocity structure.

The separation of variables approach, also known as the variable projection method (VPM), suggested by Golub and Pereyra (1973) can be another alternative strategy. This method is used in non linear least-squares problems, when it is possible to identify a subset of parameters from which the dependence is linear (Golub and Pereyra 2003, Van Leewen 2009). The idea is to define the linear parameters as a function of the non-linear one, so that the inversion problem becomes unique involving only the resolution of the non-linear parameters. Applied to a tomographic inversion, one could think about inverting for both source parameters and velocity structure, defining the moment tensor parameters (linear problem) as functions of the Earth parameters and source location (non linear problem).

In this part, I present the work I have done during my three months internship at Schlumberger Cambridge in England on the variable projection method. This is an introduction to the method with its application to a 1D acoustic case for the particular simplified problem of finding both seismic source signal and source location. All this study has been carried out with MATLAB coding and synthetic examples. The different steps of the work are detailed, from the computation of

synthetic waveforms to the implementation of the VPM, showing how both inversion problems are solved individually and then together.

### III.I.2. THEORY

The principle of an inversion method lies on the quest of the best model parameterization that characterized as close as possible the real state of the investigated problem. In other words, the inversion process is based on the minimization of a quantified difference or misfit between real data and synthetic data computed with a guessed model describing the problem.

Least squares misfit has many applications in signal modeling and parameter estimation problems. Whereas efficient methods have been developed for linear problem resolution, methods for non-linear problem are generally more binding. Existing approaches mainly fall in two main categories: minimization techniques and linearization techniques such as Gauss-Newton methods.

Otherwise, in some cases the least squares misfit presents both linear and non linear components. In 1973, Golub and Pereyra have defined the new class of separable nonleast square problem together with the method to solve it, called the variable projection method.

In this type of problem, linear and non-linear components are solved “separately” by introducing the solution of the linear problem inside the non linear problem, reducing the complexity of the problem.

Given a set of observations  $d_{i,j=1..m}$ , a separable non linear least squares problem can be defined as :

$$r(\mathbf{a}, \boldsymbol{\alpha}) = \sum_{i=1}^m \left\{ d_i - \sum_{j=1}^n a_j \varphi_j(\boldsymbol{\alpha}, t_i) \right\}^2 \quad (48)$$

where  $r$  is the least square function to minimize,  $a_{j,j=1..n}$  is the linear parameter to be determined together with the  $k$  dimensional vector of non linear parameters  $\boldsymbol{\alpha}$  defined by the non linear function  $\varphi_j$  and  $t_i$  are independent variables associated with observation  $d_i$ .

Written with matrix notation, this leads to :

$$r(\mathbf{a}, \boldsymbol{\alpha}) = \|\mathbf{d} - \boldsymbol{\phi}(\boldsymbol{\alpha})\mathbf{a}\|^2 \quad (49)$$

Where the columns of  $\boldsymbol{\phi}(\boldsymbol{\alpha})$  are the non linear function  $\varphi_j$  evaluated at all  $t_i$  values.

If I knew the non linear parameter, then the linear parameter  $\mathbf{a}$  would be easily recoverable solving the linear least squares problem:

$$\mathbf{a} = \boldsymbol{\phi}(\boldsymbol{\alpha})^+ \mathbf{d} \quad (50)$$

where  $\boldsymbol{\phi}(\boldsymbol{\alpha})^+$  is the Moore-Penrose generalised inverse of  $\boldsymbol{\phi}(\boldsymbol{\alpha})$ .

The goal of the VPM is thus to transform the misfit functional in a modified functional depending only on the non linear parameters together with an inherent subsidiary solution for the linear parameters, defined as a function of

the non linear parameter. One can consider the following variable projection functional:

$$r_2(\boldsymbol{\alpha}) = \|\mathbf{d} - \boldsymbol{\phi}(\boldsymbol{\alpha})\boldsymbol{\phi}(\boldsymbol{\alpha})^+ \mathbf{d}\|^2 \quad (51)$$

Once a minimum  $\hat{\boldsymbol{\alpha}}$  is found for  $r_2$ , then  $\mathbf{a}$  is found through (50).

### III.I.3. SOLVING A JOINT SOURCE SIGNAL AND LOCATION INVERSION PROBLEM WITH THE VPM

The context is placed in the case where I am looking for an estimation of the seismic source signal and its location using a least squares misfit. Whereas the first problem is linear, the second one is non linear. I seek to minimize the difference between the observed data  $\mathbf{d}$  and synthetic data  $\mathbf{s}$  defined as:

$$\mathbf{s} = \mathbf{y} * g(m) \quad (52)$$

That is  $\mathbf{s}$  is the convolution product between a wavelet signal and the Green's function, depending on the model parameters. In our case, the Green function contains information about the source location. The misfit is written in its simplest form as:

$$\chi(\mathbf{y}, m) = \frac{1}{2} \|\mathbf{y} * g(m) - \mathbf{d}\|^2 \quad (53)$$

or in matrix notation:  $\chi(\mathbf{y}, m) = \frac{1}{2} \|\mathbf{G}\mathbf{y} - \mathbf{d}\|^2$

Where  $\mathbf{G}$  is the operator convolution with the Green's function  $g(m)$ . Its dependency to model parameter  $m$  is hidden for a better reading.

One can note that I am using a waveform misfit here. Indeed, only this kind of misfit allows implementing easily the VPM. This misfit allows defining the linear problem with respect to the source signal inversion problem. Other misfits based on secondary observable such as travel-time misfit do not allow this linearity and thus cannot be used for the VPM.

The VPM involved minimizing

$$\chi(\mathbf{m}) = \frac{1}{2} \|\mathbf{G}(\mathbf{G}^t \mathbf{G})^{-1} \mathbf{G}^t \mathbf{d} - \mathbf{d}\|^2 \quad (54)$$

with respect to the non linear parameter  $m$ , subject to the constraint that  $\mathbf{y}$  minimizes

$$\mathbf{y} = (\mathbf{G}^t \mathbf{G})^{-1} \mathbf{G}^t \mathbf{d} \quad (55)$$

This non-linear inversion problem, involving only one parameter estimation, can then be solved with traditional method, for any choice of  $m$ .



## III.II. COMPUTE SYNTHETIC WAVEFORMS IN AN HOMOGENEOUS ISOTROPIC ACOUSTIC MEDIUM

---

I use the following Fourier convention:

$$\begin{cases} \hat{h}(\omega) = \int_{-\infty}^{\infty} h(t) e^{i\omega t} dt \\ h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{h}(\omega) e^{-i\omega t} d\omega \end{cases} \quad \text{with } \omega \text{ the angular frequency } \omega = 2\pi f$$

The compressional wave velocity  $c$  is fixed at 7km/s.

A seismic signal can be interpreted as an initial wavelet signal representing the seismic source convolved with its response to the Earth's medium defined as the Green's function.

$$u_k(t) = s(t) * g_k(t)$$

### III.II.1. GREEN'S FUNCTION

The Green's function is the response of the Earth's medium to a unit impulsive force in space and time. In the time domain, the 2D Green's function  $g^{2D}$  in a homogeneous acoustic medium satisfies the following system:

$$\begin{cases} \left( \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) g^{2D}(x, t; x', t') = -\delta(x - x') \delta(t - t') \\ g^{2D}(x, t; x', t') = 0 \text{ for } t < t' \end{cases} \quad (56)$$

With analytical solution:

$$g^{2D}(x, t; x', t') = \frac{1}{2\pi} \frac{H(\tau - \frac{r}{c})}{\sqrt{(\tau^2 - (\frac{r}{c})^2)}}$$

Where:

and  $H$  is the Heaviside function

$$\begin{cases} \tau = t - t' \\ r = \sqrt{(x - x')^2 - (z - z')^2} \end{cases} \quad \begin{cases} H(k) = 0 \text{ for } k < 0 \\ H(k) = 1 \text{ for } k \geq 0 \end{cases}$$

These results can be transposed to the frequency domain by Fourier transform, where the Greens' function is denoted by a capital letter:

$$G^{2D}(x, \omega; x') = \int_{-\infty}^{\infty} g^{2D}(x, t; x', t') e^{i\omega t} dt \quad (57)$$

$$G^{2D}(x, \omega; x') = \int_{-\infty}^{\infty} \frac{1}{2\pi} \frac{H(\tau - \frac{r}{c})}{\sqrt{(\tau^2 - (\frac{r}{c})^2)}} e^{i\omega\tau} d\tau$$

The Heaviside function is related to the Hankel function through the following equation (Chapman 2004):

$$\int_{-\infty}^{\infty} \frac{1}{2\pi} i\mathcal{H}_0^1(a\omega) e^{-i\omega t} d\omega = \frac{2H(t-a)}{\pi\sqrt{t^2-a^2}} \quad (58)$$

$\mathcal{H}_0^1$  is the Hankel function of the first kind of order 0 :  $\mathcal{H}_n^1(x) \sim \sqrt{\frac{2}{\pi x}} e^{i(x - n\frac{\pi}{2} - \frac{\pi}{4})}$

Introducing the function  $h(t)$  and its Fourier transform  $\hat{h}(\omega)$ , ( 58 ) can be rewritten as:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{h}(\omega) e^{-i\omega t} d\omega = h(t) = \frac{2H(t-a)}{\pi\sqrt{t^2-a^2}} \quad (59)$$

$$\hat{h}(\omega) = i\mathcal{H}_0^1(a\omega) = \int_{-\infty}^{\infty} \frac{2H(t-a)}{\pi\sqrt{t^2-a^2}} e^{i\omega t} dt$$

By analogy with ( 58 ) we finally find a simple formulation of the 2D Green's function in a homogeneous acoustic medium in terms of the Hankel function:

$$\int_{-\infty}^{\infty} \frac{1}{2\pi} \frac{H(\tau - \frac{r}{c})}{\sqrt{(\tau^2 - (\frac{r}{c})^2)}} e^{i\omega\tau} d\tau = \frac{i}{4} \mathcal{H}_0^1\left(\frac{\omega}{c} r\right) \quad (60)$$

$$G^{2D}(x, \omega; x') = \frac{i}{4} \mathcal{H}_0^1\left(\frac{\omega}{c} r\right) \quad (61)$$

To avoid the high discrepancy values introduced by the Green's function, I use a smoothing operator in form of a boxcar function convolved with the Green's function. The boxcar function is defined in the time domain as:

$$box(t) = \begin{cases} 1 & -\frac{a}{2} < t \leq \frac{a}{2} \\ 0 & elsewhere \end{cases} \quad (62)$$

In the frequency domain the Fourier transform of the boxcar is a function cardinal sine:

$$\widehat{box}(\omega) = asinc\left(\omega \frac{a}{2}\right) \quad (63)$$

Having set the mathematical description of the Green's function, I have now to define the seismic source signal that is described by a wavelet signal.

### III.II.2. WAVELET SIGNAL

The seismic source signal can be either described by a Ricker wavelet commonly used in seismic prospection

$$r(t) = (1 - 2\pi^2 f_0^2 t^2) e^{-\pi^2 f_0^2 t^2} \tag{64 a}$$

or a Morlet wavelet  $m(t)$  (Figure 14).

$$m(t) = \frac{1}{\sqrt{\pi}} e^{i\omega_0 t} e^{-t^2} \tag{65 b}$$

The displacement seismic signal is finally the convolution of the smoothed Green function with the wavelet signal (Figure 14 and Figure 15).

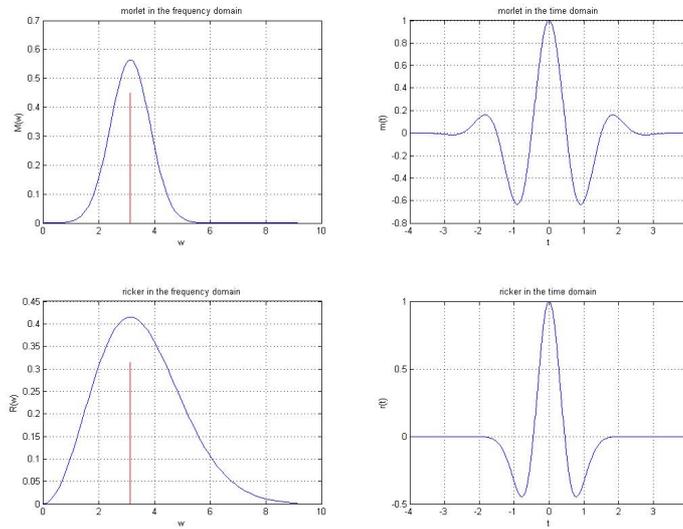


Figure 14 - Morlet and Ricker wavelet for peak frequency  $F_0 = 0,5\text{Hz}$ .

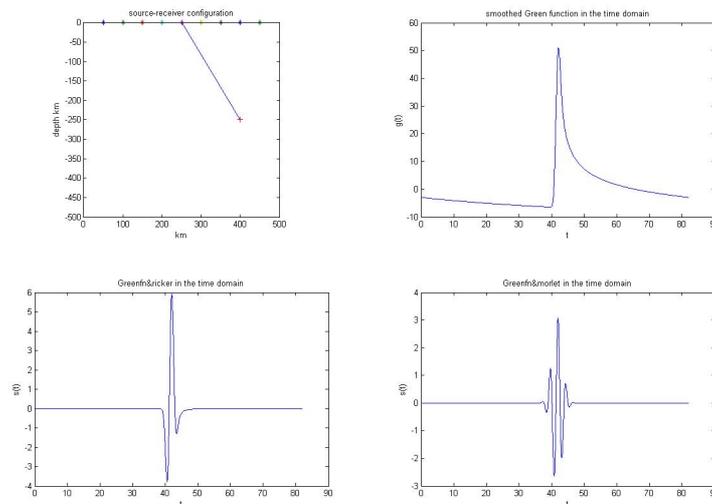


Figure 15 - Source-receiver configuration, smoothed Green function, Morlet and Ricker seismic signals for peak frequency  $F_0 = 0,5\text{Hz}$ .



### III.III. THE SOURCE SIGNAL PROBLEM

The seismic source signal inversion is a linear problem, whose solution can be written in terms of the pseudo inverse of the convolution operator with the Green function  $\mathbf{G}$ :

$$\mathbf{y} = (\mathbf{G}^t \mathbf{G})^{-1} \mathbf{G}^t \mathbf{d} \quad (66)$$

The source signal  $\mathbf{y}$  is thus found from the autocorrelation matrix of the Green's function  $\mathbf{G}^t \mathbf{G}$  and the correlation matrix of the Green's function with observed data  $\mathbf{G}^t \mathbf{d}$ , as the transpose of the convolution operator is the correlation. This operation is equivalent to Wiener filtering method, which consists in finding a filter that transforms an input wavelet in the desired output shape in a least-square sense.

In our implementation I use a priori information about the source signal: I already have an idea of its shape. I assume that it is a Ricker wavelet. However, I do not have information on the dominant frequency of this signal. The Wiener filter operation consists in finding the deconvolution filter that minimizes the misfit between an a priori input signal and the data. Depending on the dominant frequency of the input wavelet the Wiener filter gives different results (Figure 16). The best result on the waveform misfit is obtained when overestimating the dominant frequency of the wavelet signal.

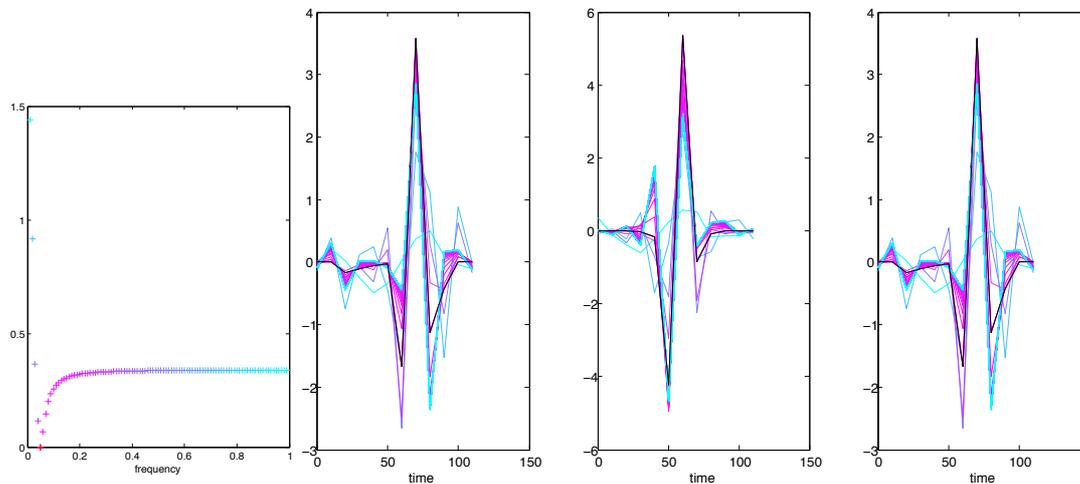


Figure 16 - Output from Wiener filtering for a three stations line geometry. The Wiener filter has been computed for all the stations to recover the source signal. The color legend is relative to the accuracy of the output wiener signal, the warmer is the color the closer is the signal to real data. The left figure shows the variation of the misfit with respect to the predominant frequency used.



## III.IV. THE SOURCE LOCATION PROBLEM

In simple non-linear inverse problems, the best model parameterization can be found with a grid search over all the possible sets of parameters. The combination giving the minimum of the misfit forms the best model estimate.

I illustrate this concept with two different misfits: the travel time and the waveform misfit function. Even if the travel-time misfit cannot be used in the frame of the VPM, it is interesting to see these functions shapes to illustrate the notions of local and global minima and the difference between the two misfits.

### III.IV.1. MISFITS EXAMPLES

#### III.IV.1.1. Travel-time misfit

The travel-time misfit function is defined as the least-square difference over the station-source pair of the cross-correlation time delay between the data and the synthetics computed within a given model  $m$ :

$$\chi_{tt} = \frac{1}{2} \sum_{i=1}^n (T_i^{obs} - T_i^m)^2 \quad (67)$$

where  $n$  stands for the number of receivers and  $(T_i^{obs} - T_i^m)$  is the time delay of maximum cross-correlation between the real and synthetic seismograms.

A source is located at the point  $(x,z)$  (250km,250km) of the grid. As this inversion involves only two parameters, the best location estimate within a grid search over all  $(x,z)$  is feasible. This gives Figure 17 using a 3 stations line array spaced by 250km (that is:  $x = 0 \ 250 \ 500$ km) and a 25km spaced grid involving 110 computations. The minimum of the misfit function gives the  $(x,z)$  combination that best fits the real source location.

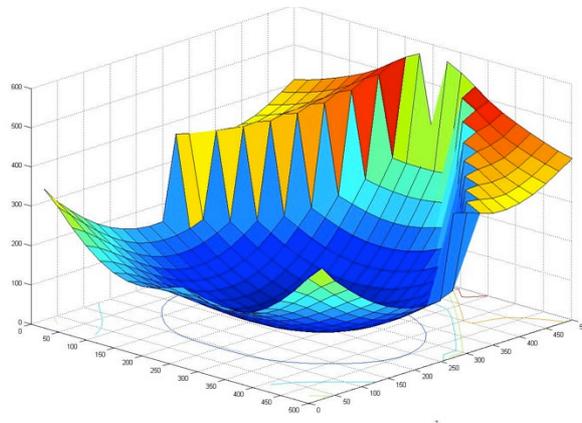


Figure 17 - Travel-time misfit function distribution from a grid search on all points of the 25km spaced grid model for a source located at (250,250) and 3 stations located at  $x(0 \ 250 \ 500)$ . The minimum of the misfit gives the best model parameterization.

### III.IV.1.2 Waveform misfit

The waveform misfit function is defined as the weighted least-square difference between the synthetic and real seismograms waveforms:

$$\chi_{wf} = \frac{1}{2} \sum_{i=1}^n \int_0^T w_i \|s(x_i, t, m) - d(x_i, t)\|^2 dt \quad (68)$$

where  $n$  stands for the number of receivers.

The weight  $w_i$  is defined as the inverse of squared data registered at a given receiver  $i$ .

Taking the same example as in the previous section, a grid search over all the possible model parameterizations give the waveform misfit distribution of Figure 18.

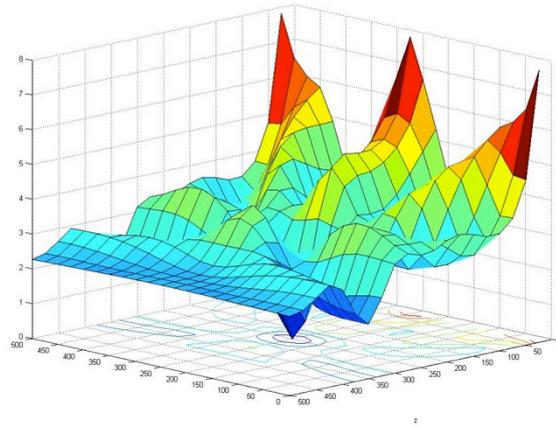


Figure 18 - Waveform misfit function distribution from a grid search on all points of the 25km spaced grid model for a source located at (250,250) and 3 stations located at  $x(0\ 250\ 500)$ . The minimum of the misfit gives the best model parameterization.

The global minimum of the misfit function is clearly seen, but I note also the presence of several local minima (in comparison with the travel-time misfit function that has only a global minimum), which in case of optimization algorithms search could lead to a wrong estimation of the best model parametrization.

### III.IV.2. OPTIMIZATION ALGORITHMS

In case of heavy inversion problems, where the model parameters to find are numerous involving a lot of computations, the grid search algorithm is not efficient anymore and optimization algorithms are used to find the minimum of the misfit. In such case, instead of computing the misfit for all points of the grid search and look at the minimum, the search starts from a given starting point (initial model) of the grid and uses the misfit and gradient misfit values at this point to find a following point of the grid (model update). A stopping criterion is subsequently fixed to close the search when the misfit is close enough to zero.

These optimization algorithms involve the computation of the gradient of the misfit, that is the perturbation of the misfit function due to a perturbation of the models parameters. This involves computing the partial derivatives of the misfit regarding each model parameters, which in some case can become a heavy task. To deal with it, an alternative method has been developed, that allows getting access to the gradient of the misfit function without computing each derivative, it is the adjoint method we have detailed in a preceding part of the thesis.

In the following, I will present the way an inversion is done with a gradient-based optimization algorithm avoiding the computation of the misfit at all points of the grid search.

### III.IV.2.1. Gradient computation

A first step in the gradient based optimization algorithm is the computation of the gradient of the misfit function., which indicates the direction of the model update. For a model  $m(m_1, \dots, m_i, \dots, m_n)$  of  $n$  parameters  $m_i$ , the gradient  $g$  of the misfit function  $\chi$  is defined as the vector  $g = [g_1 \ g_i \ g_n]$ , where  $i$  points to a given model parameter to invert:

$$g_i = \frac{\partial \chi}{\partial m_i} \quad (69)$$

In case of my 2D example, I seek to find the two parameters  $(x, z)$  of the model  $m$  that gives the best source location. The gradient can be written as:

$$g = \left( \frac{\partial \chi}{\partial x}, \frac{\partial \chi}{\partial z} \right) \quad (70)$$

From ( 68 ), the variation of the waveform misfit function  $\delta \chi_{wf}$  in the time domain is given by:

$$\delta \chi_{wf} = \sum_{i=1}^n \int_0^T w_i [s(x_i, t, m) - d(x_i, t)] \delta s(x_i, t, m) dt \quad (71)$$

where  $\delta s(x_i, t, m)$  is the perturbation in the displacement field due to a model perturbation  $\delta m$ .

I need to find the variations of the misfit function due to a particular model parameter that are its partial derivatives  $\frac{\partial \chi}{\partial m_i}$ . This implies to find the partial derivatives of  $s(x_i, t, m)$ .

$$\frac{\partial \chi_{wf}}{\partial m_j} = \sum_{i=1}^n \int_0^T w_i [s(x_i, t, m) - d(x_i, t)] \frac{\partial s(x_i, t, m)}{\partial m_j} dt \quad (72)$$

In our case, the synthetic displacement field  $s(x_i, t, m) = s(x_i, t, (x, z))$  is defined from the convolution of the Green's function with the wavelet signal. As the wavelet signal is independent of the model parameters  $(x, z)$ , the derivatives of  $s$  reduce to the derivatives of the Green's function convolved with the wavelet signal. In the frequency domain, this becomes the product of the derivatives of the Hankel function with the wavelet signal:

$$G^{2D}(x, \omega; x') = \frac{i}{4} \mathcal{H}_0^1\left(\frac{\omega}{c} r\right) \tag{73}$$

$$\frac{\partial s}{\partial x'} = \frac{\partial G}{\partial r} \frac{\partial r}{\partial x'} \qquad \frac{\partial s}{\partial z'} = \frac{\partial G}{\partial r} \frac{\partial r}{\partial z'}$$

A useful property of the derivative of the Hankel function of the first kind of order zero is that it is proportional to the Hankel function of the first kind of first order:

$$\frac{\partial \mathcal{H}_0^1\left(\frac{\omega}{c} r\right)}{\partial r} = -\frac{\omega}{c} \mathcal{H}_1^1\left(\frac{\omega}{c} r\right) \tag{74}$$

Using  $\frac{\partial r}{\partial x'} = \frac{1}{2\sqrt{r}}(-2x + 2x')$  and equivalent for the parameter z, I obtain the partial derivatives of the displacement field with respect to the location parameters and finally the partial derivatives of the misfit of the gradient.

The gradient gives the direction to go for a better model parametrization. Where the misfit value is high over a region, the gradient indicates to look outside and presents some ‘divergence’ pattern. On the contrary, when in a minimum misfit area, the gradient indicates to pursue the search inside the area with a kind of ‘convergent’ behaviour (Figure 19 and Figure 20).

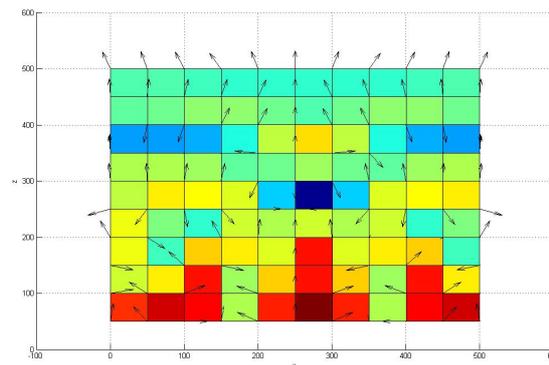


Figure 19 - 2D representation of the waveform misfit distribution with the gradient direction given at all points of the grid.

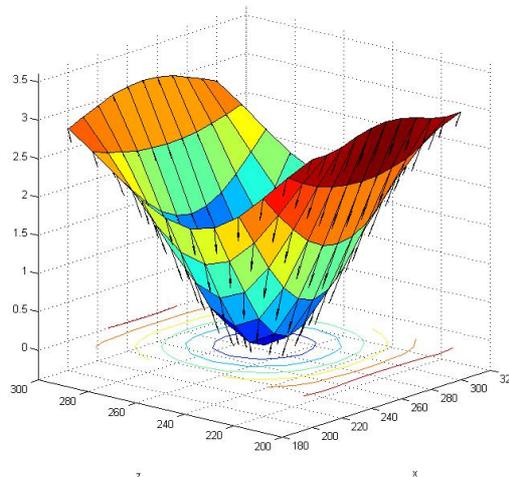


Figure 20 - 3D representation of the waveform misfit distribution with the gradient direction given for the area around the global minimum.

### III.IV.2.2. Iteration

The optimization algorithm implemented in MATLAB uses the Trust-Region-Reflective(TRR) scheme<sup>3</sup>. For the minimization of a given function  $f(x)$ ,  $f$  is approximated by its quadratic Taylor expansion  $\varphi_i$  in the neighborhood  $N$  of the current point  $x_i$  to find the following iteration point giving a smaller value of  $f$ .  $N$  is the trust region containing the point  $x_{i+1}$ . The step  $stp_i$  between two iteration points is approximately found by solving:

$$\varphi_i(stp) = g^T stp + \frac{1}{2} stp^T H stp \quad (75)$$

Where  $g$  and  $H$  are the gradient and Hessian of  $f(x_i)$  respectively. When the variation of  $f(x)$  comes below a chosen tolerance, the algorithm ends.

The method of least-squares fitting with TRR is implemented by a library function in MATLAB (`lsqnonlin`). This function is such that it finds the coefficients  $x$  that solve:

$$\min \sum_i f_i(x)^2 \text{ where } f \text{ is the vector } [f_1, f_2, \dots, f_n]. \quad (76)$$

For our particular purpose of waveform misfit least-squares minimization, I seek to minimize

$$\chi_{wf} = \frac{1}{2} \sum_{i=1}^n \int_0^T w_i \|s(x_i, t, m) - d(x_i, t)\|^2 dt \quad (77)$$

Thus  $f$  in (76) can be written as a matrix  $F$  such as:

$$F = \sqrt{\frac{W}{2}} (S - D) \quad (78)$$

Where  $S$ ,  $D$ ,  $W$  represent the synthetic and data vectors and the weight matrix, in particular:

$$S = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_i \\ \vdots \\ S_n \end{pmatrix} \quad \text{where } S_i = \begin{pmatrix} s_i^1 \\ s_i^2 \\ \vdots \\ s_i^l \end{pmatrix} \text{ is the synthetic signal of } l \text{ time steps at station } i$$

$$D = \begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_i \\ \vdots \\ D_n \end{pmatrix} \quad \text{where } D_i = \begin{pmatrix} d_i^1 \\ d_i^2 \\ \vdots \\ d_i^l \end{pmatrix} \text{ is the observed signal of } l \text{ time steps at station } i$$

<sup>3</sup> <http://www.mathworks.it/it/help/optim/ug/least-squares-model-fitting-algorithms.html#broz0i4>



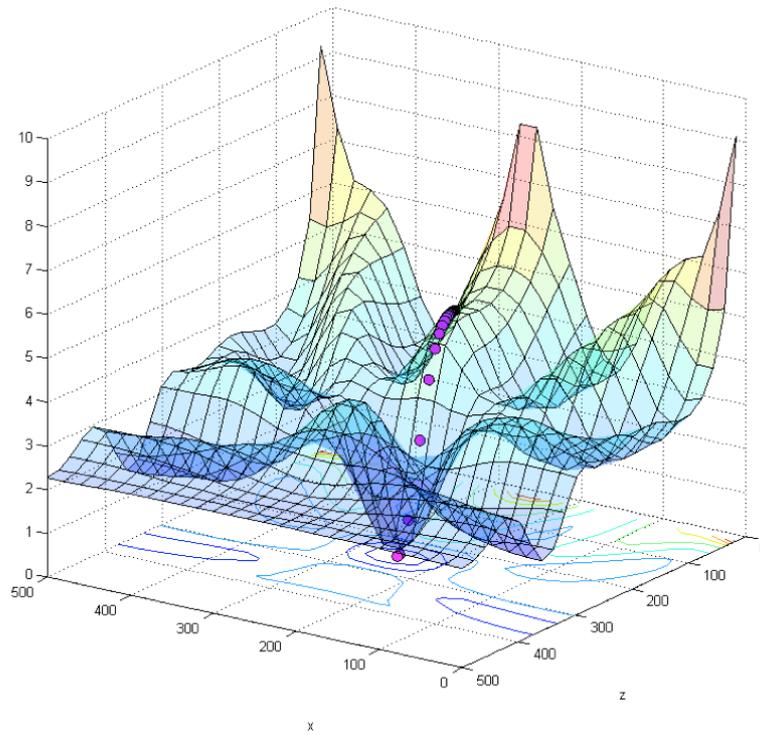


Figure 21 - Location iteration through the TRR algorithm. Iterations (filled circles) start from cold color at point (250, 150)km to hot color . In this case the global minimum in (250,250)km is found after 33 iterations.

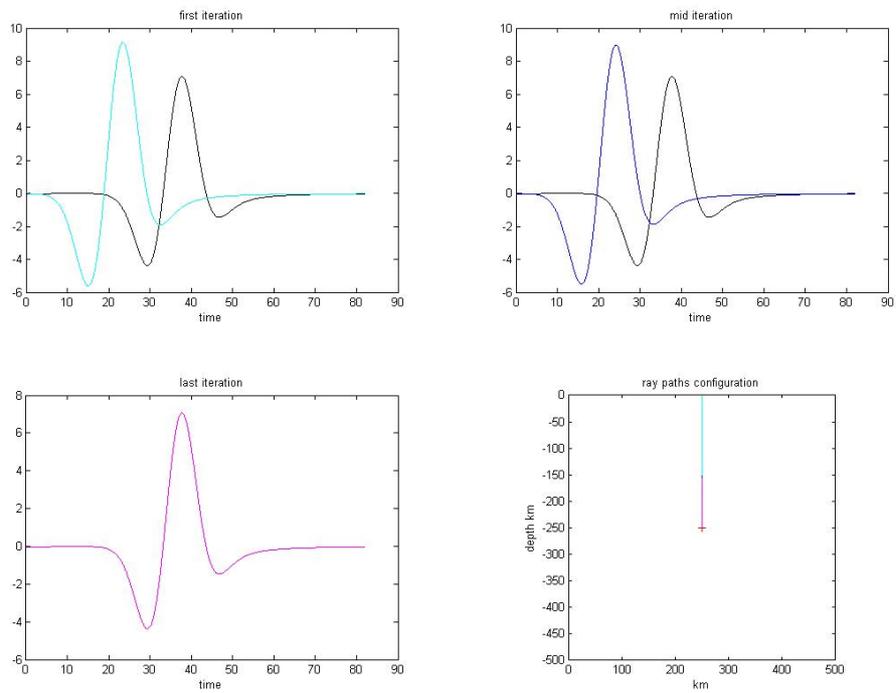


Figure 22 - Seismic signals differences evolution from first iteration at point (250,150)km to last one at a particular receiver. The real source location is indicated by a red cross.

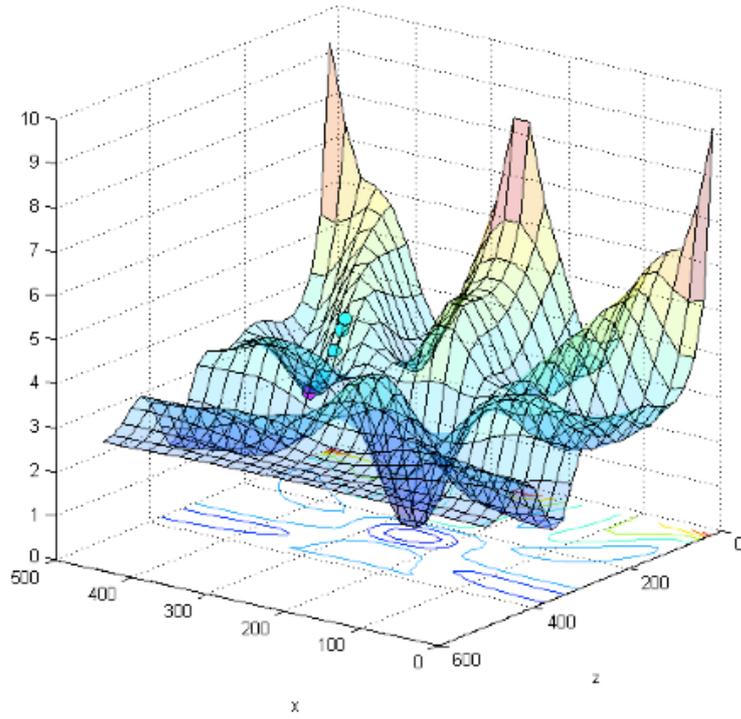


Figure 23 - Location iteration through the TRR algorithm. Iterations (filled circles) start from cold color at point (400, 150) to hot color. In this case the search is trapped in a local minimum and the search ends after 55 iterations.

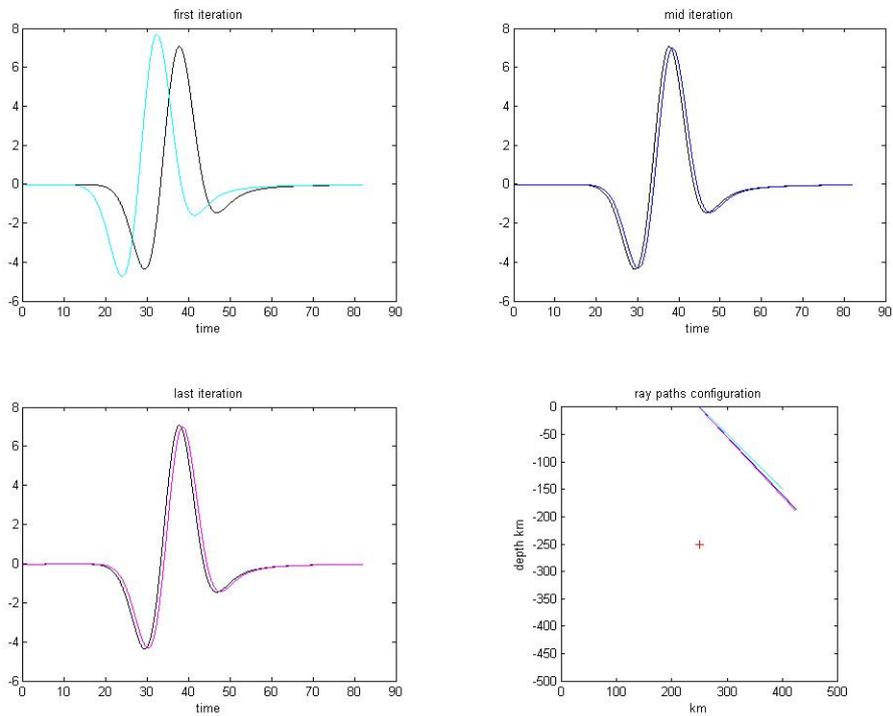


Figure 24 - Seismic signals differences evolution from first iteration at point (400,150)km to last one at a particular receiver.

### III.V. APPLICATION OF THE VARIABLE PROJECTION METHOD

Applying the Variable Projection Method, the waveform misfit shape comes out quite different from the previous case with a known source signal. In this case, I obtain a less focused misfit minimum over the location parameters (Figure 25).

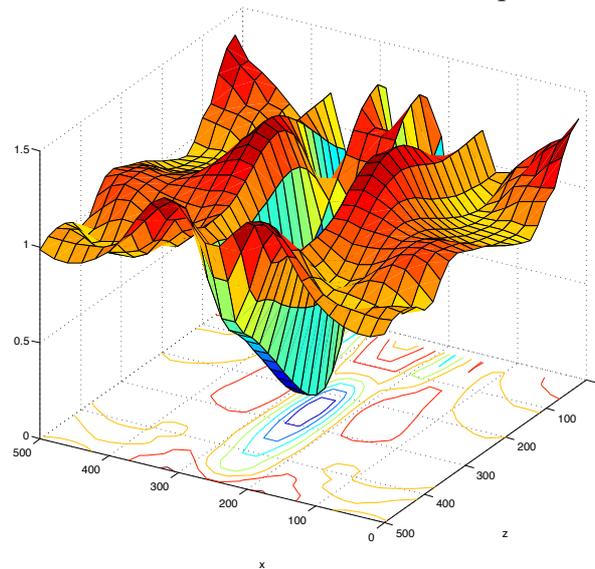


Figure 25 - Waveform misfit distribution with two unknowns: the source signal and its position.

Looking at the optimization part, as before the success of the optimization algorithm to find the global minimum is strongly dependent on the starting point, or starting model (Figure 26 to Figure 29).

The Variable Projection Method appears to be an efficient tool to solve least squares minimization problems involving linear and non linear parameters. In case of a seismic tomography inversion, one is interested in solving for the geophysical parameters of the Earth structure as well as the source parameters, to carry out the inversion in a more accurate frame and obtain consistent models. Solving for the source parameters would involve the source location as well as the moment tensor estimation. A problem, which is alone a combination of linear and non linear problem and is commonly solved through different schemes. Recently, methods for solving for the whole source parameters have been developed, but still distinctly to the geophysical model inversion, following a sequential update of both models (source model vs velocity model) (Liu et al. 2006).

Valentine and Woodhouse (2010) have opened the door to joint inversion of both source and geophysical model through linearization solver. In their study, they solve for a least squares minimization of the waveform misfit, treating each seismograms as depending linearly on both Earth model and source vector. The VPM, absent in the tomographic field, does not imply any linearization approximation and thus should be seen as an interesting alternative to carry out this type of inversion. In its simplest implementation it implies to use the waveform misfit and thus might not be applicable for some studies where this

misfit is not usable due, for example, to noisy waveforms. Still, it can have a huge range of applications and could be developed for the use of other misfit functions.

In the frame of my full waveform inversion for the Vrancea region, the VPM could have been used to invert for source and velocity model. However, its application in such a frame, implying a more complicate solution, required a lot of work and I did not implement it here. Nevertheless, the example shown here has a strong pedagogical value and provides useful insights in the method as well as in iterative optimization techniques.

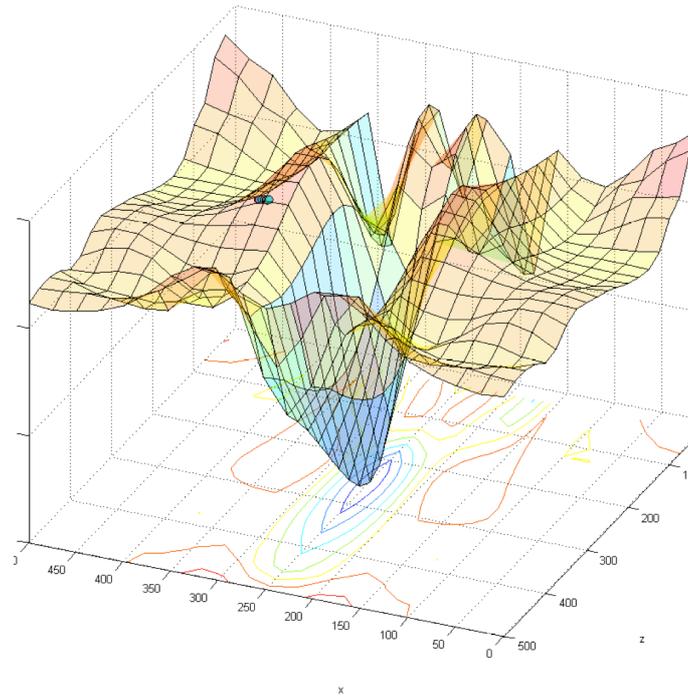


Figure 26 - Optimization algorithm using the variable projection method at test point 1.

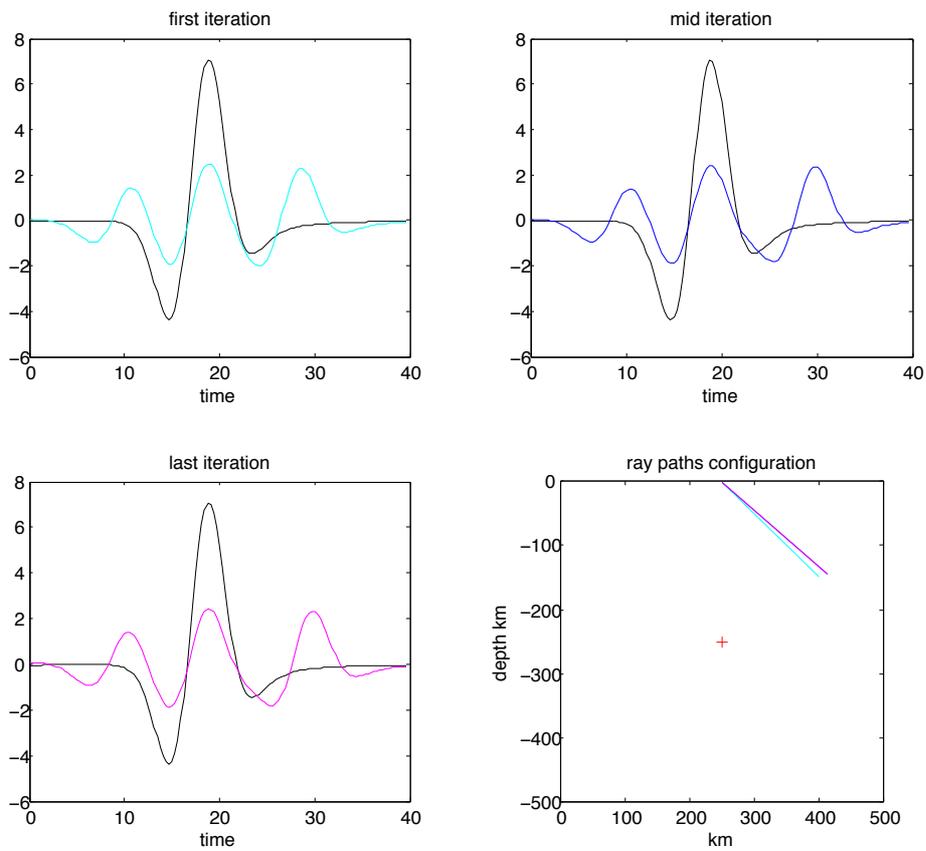


Figure 27 - Seismograms evolution through the VPM optimization algorithm.

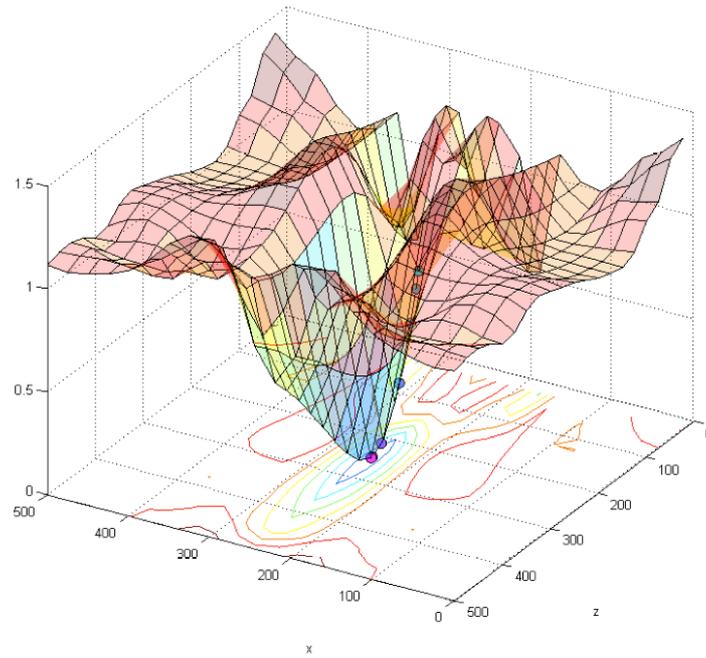


Figure 28 - Optimization algorithm using the variable projection method at test point 2.

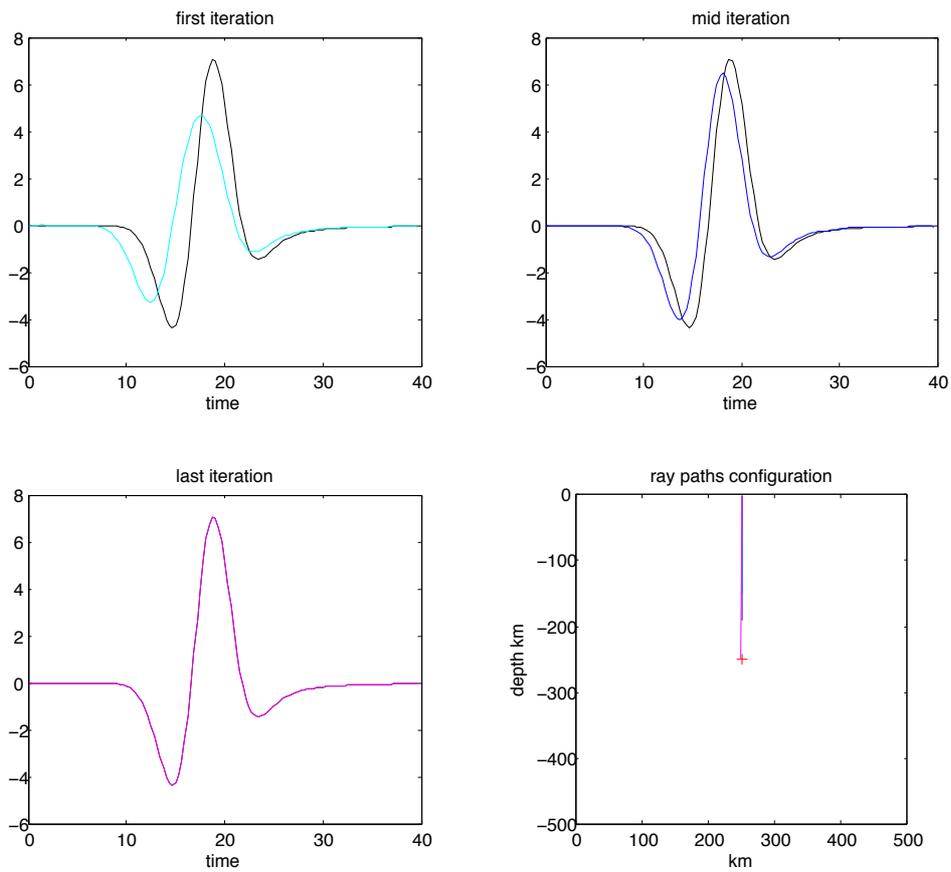


Figure 29 - Source signal evolution through the variable projection method algorithm.

# IV FULL WAVEFORM INVERSION WITH THE ADJOINT METHOD APPLICATION TO THE VRANCEA REGION USING LOCAL DATA FROM THE CALIXTO99 EXPERIMENT

<b>IV.I. SETTING UP THE VRANCEA DATASET .....</b>	<b>69</b>
IV.I.1. DATA FROM THE 1999 CALIXTO EXPERIMENT .....	69
IV.I.1.1. The CALIXTO 1999 experiment .....	69
IV.I.1.2. Input database study .....	70
IV.I.1.3. Data banking and delivery .....	70
IV.I.2. VRANCEA 3D REFERENCE TOMOGRAPHIC MODEL ...	71
IV.I.2.1. Input model .....	71
IV.I.2.2. Corrections on the model .....	71
IV.I.3. SOURCE LOCATION AND MOMENT TENSOR IN THE 3D MODEL .....	75
IV.I.4. PROCESSING THE DATABASE .....	78
IV.I.4.1. Noise processing .....	78
IV.I.4.2. Instrumental responses .....	78
IV.I.4.3. Frequency content .....	79
IV.I.4.4. Final database for inversion .....	81
<b>IV.II. SEISMIC WAVEFORM MODELING .....</b>	<b>83</b>
IV.II.1. WAVEFORM MODELING WITH SPECFEM .....	83
IV.II.1.1. Introduction .....	83
IV.II.1.2. Build the computational mesh with the internal mesher .....	84
IV.II.2. FORWARD MODELING ON CINECA BLUE GENE/Q ...	85
IV.II.2.1. The Vrancea CALIXTO set up .....	85
IV.II.2.2. First waveforms comparisons .....	87
<b>IV.III. STRUCTURE INVERSION .....</b>	<b>91</b>
IV.III.1. FROM MISFIT TO ADJOINT SOURCE .....	91
IV.III.1.1. Selecting the windows .....	91
IV.III.1.2. Compute the misfit and adjoint sources .....	92
IV.III.2. FROM EVENT KERNELS TO FIRST UPDATE .....	94
IV.III.2.1. Misfit kernels .....	94
IV.III.2.2. Update .....	97
IV.III.3. FROM INITIAL MODEL TO 3 <sup>RD</sup> UPDATE .....	97
IV.III.3.1. Results .....	97
III.3.2. Discussion .....	105



## IV.I. SETTING UP THE VRANCEA DATASET

### IV.I.1. DATA FROM THE 1999 CALIXTO EXPERIMENT

#### IV.I.1.1. The CALIXTO 1999 experiment

The six months international tomography experiment CALIXTO (Carpathian Arc Lithosphere X-Tomography) conducted in South-Eastern Romania between May and November 1999, involved 120 seismic stations, from which ninety short period and thirty broadband stations, and the registration of 173 local events (Figure 30).

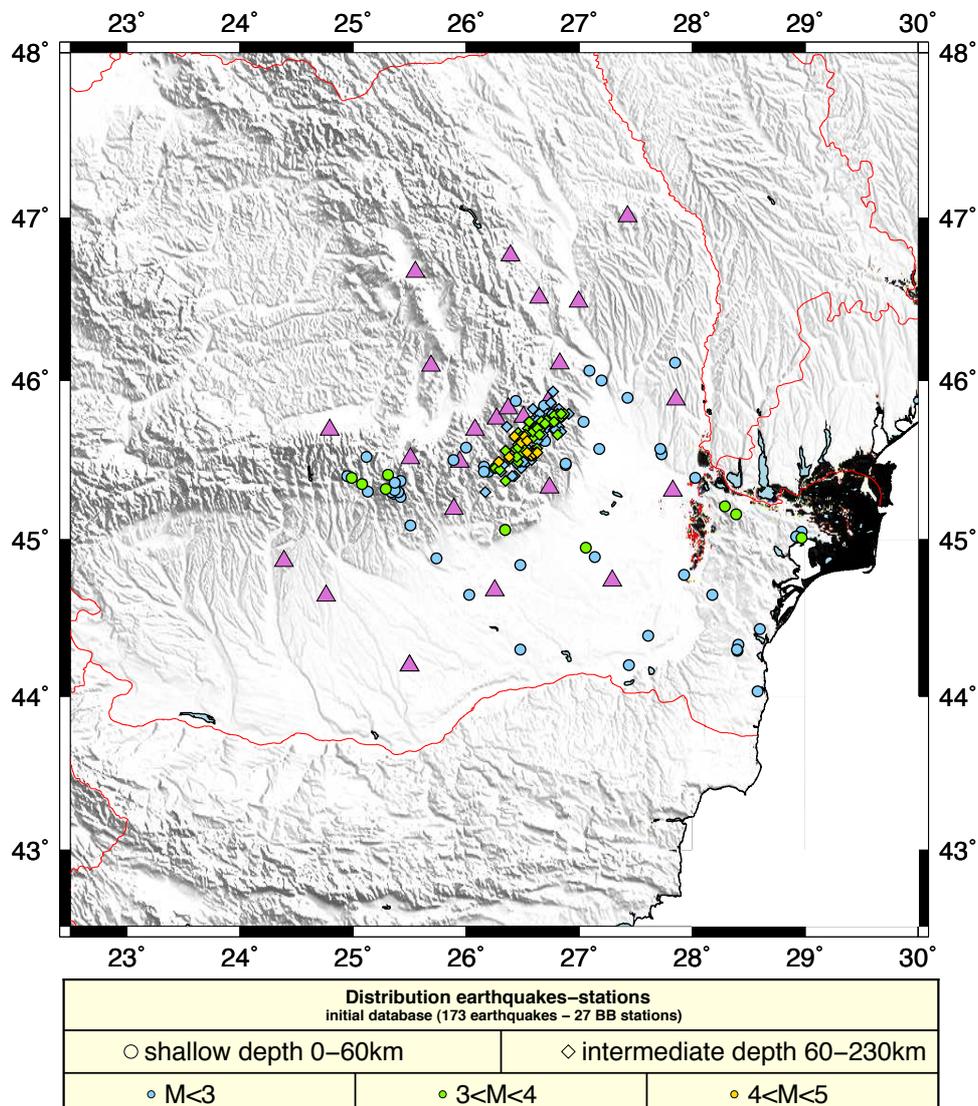


Figure 30 - Distribution of the CALIXTO99 broadband seismic stations together with the recorded local events.

Our database is built on the registrations of 27 three-components broadband seismic stations with diverse instrumentation configurations from which Guralp, Streickensen and Lennartz sensors (Table 1 in Annexes).

### IV.I.1.2. Input database study

The database contains 173 local events with magnitude  $M_w$  from 1.8 to 4.8, and a depth distribution from '0' km to 157km.

The primary earthquake information details, which are hypocentre location and magnitude, were taken from the CALIXTO99 archive documentation.

This database is a good representation of the regional seismicity, with small earthquakes at shallow depth from 0 to 60km and bigger events at intermediate depth from 80 to 160km, concentrated in the Vrancea zone (Figure 31). It even gives a snapshot of the seismic gap present in the depths around 70km (Raykova et al., 2006).

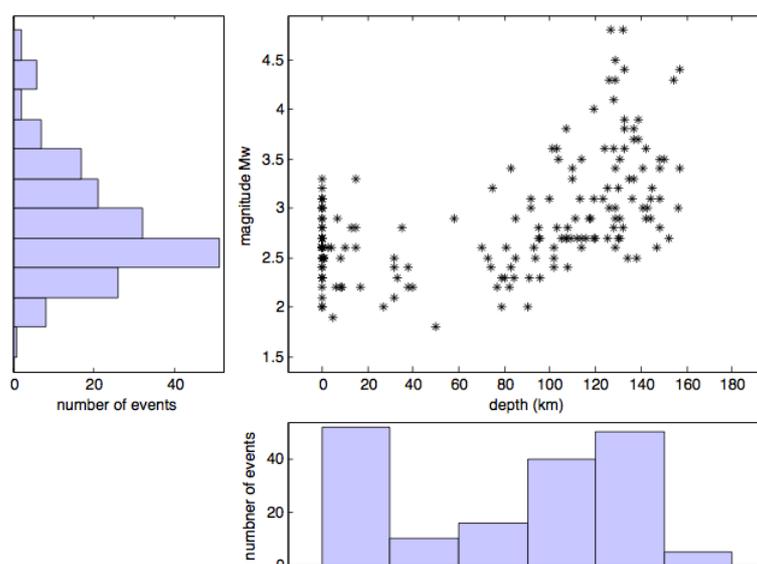


Figure 31 - Distribution magnitude-depth for the CALIXTO99 input database. Most earthquakes are in the magnitude range of  $2 < M_w < 3.5$  and two main hypocentre depths are underlined: from 0 to 20 km and from 100 to 150 km. The first group might be biased by the fact that many event are presented with a 0 km depth.

### IV.I.1.3. Data banking and delivery

There are different ways of archiving seismic data. The international standard format for the exchange of digital seismological data is the SEED (“Standard for the Exchange of Earthquake Data”) format. SEED files contain digital time series data and metadata information such as station and channel identifier, sample rate and instrument response<sup>4</sup>. The miniseed-dataless format consists in putting the waveforms data on a “miniseed” file, separately from the metadata “dataless” file. While SEED is a complex standard with a large set of data packets, called Blockette, miniSEED requires a limited amount of storage and

<sup>4</sup> The signal recorded at seismic stations is a convolution of the seismic source, the Earth’s structure and the recording system. The latter is characterised by its instrumental response that acts as a filter over the seismic signal in a given frequency range. To get a meaningful ground motion signal, one has thus to deconvolve the signal from the instrument response.

thus appears more convenient. Nevertheless, using miniseed-dataless data adds major difficulties in post-processing of temporary experiment records, especially when dataless files are physically separated from the miniseed files, lost or unavailable.

The CALIXTO99 records had not been processed and archived on any seismic portal at the beginning of the study. I got them at the Institut de Physique du Globe of Strasbourg where a copy of the database is conserved in form of compact disks in miniseed format without any dataless volume. These ones were obtained on request at the GeoForschungsZentrum (GFZ) Potsdam that was just starting the archiving of the database.

The first processing step has consisted in retrieving SAC traces from the miniseed format using convertor codes (as mseed2sac or the rdseed routine) that allow rebuilding the complete seismograms with stations information as well as the instrumental response files of the stations from the respective dataless files.

## **IV.I.2. VRANCEA 3D REFERENCE TOMOGRAPHIC MODEL**

### **IV.I.2.1. Input model**

My reference tomographic model is taken from the Tondi et al. (2009) tomographic study.

Using jointly seismic data from temporary experiments together with gravimetric data, the authors have built a detailed 3D  $V_p$ ,  $V_s$  and density tomographic model of the Vrancea area from 25-28 latitude degrees to 44,36-46,43 longitude degrees for a 230km depth.

This model gives an interesting snapshot of the main crustal and upper mantle features of the Vrancea lithosphere and is a good start for my tomographic inversion.

### **IV.I.2.2 Corrections on the model**

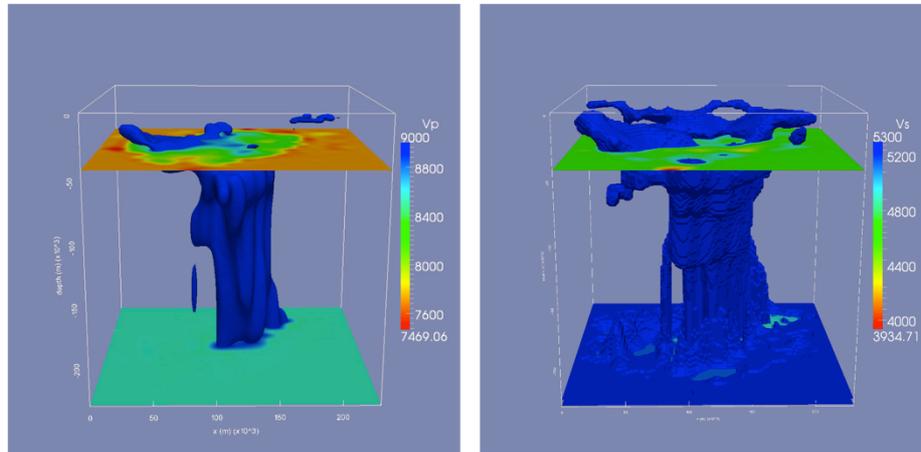
One of the weaknesses of a joint inversion is the possible introduction of some inaccuracy in the model. In the case of the Tondi et al. (2009) study, where seismic rays do not constrain a region of the model, the gravimetric data tend to fill this lack with all remaining residuals, leading to unexpected velocity values (Figure 33).

Consequently, to become suitable for our tomographic scope, this model required some corrections. In a first time, the 3D model has thus been smoothed and corrected such that the velocity from a 1D reference model<sup>5</sup> has been automatically assigned to the nodes non-resolved by the seismicity. In a second step, a manual correction was necessary to better identify the remaining nodes to correct. The correction was done on any slices depth of the initial 3D model, such

---

<sup>5</sup> Details about the 1D reference model can be found in Figure 67 in Appendix B

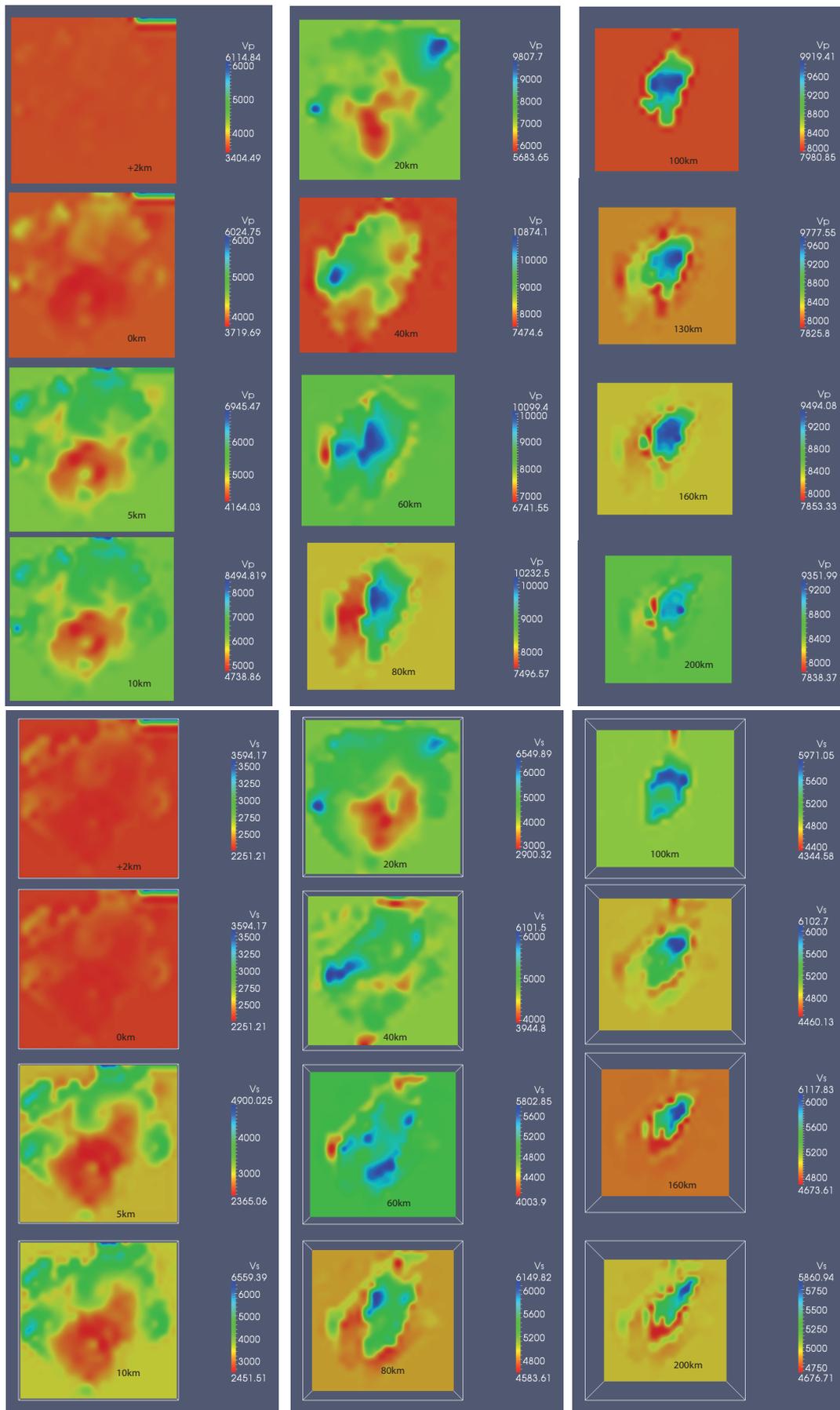
that a maximum threshold velocity value has been assigned to the unwanted nodes. In this way, I removed the high velocity spots from the model (Figure 32 and Figure 34).

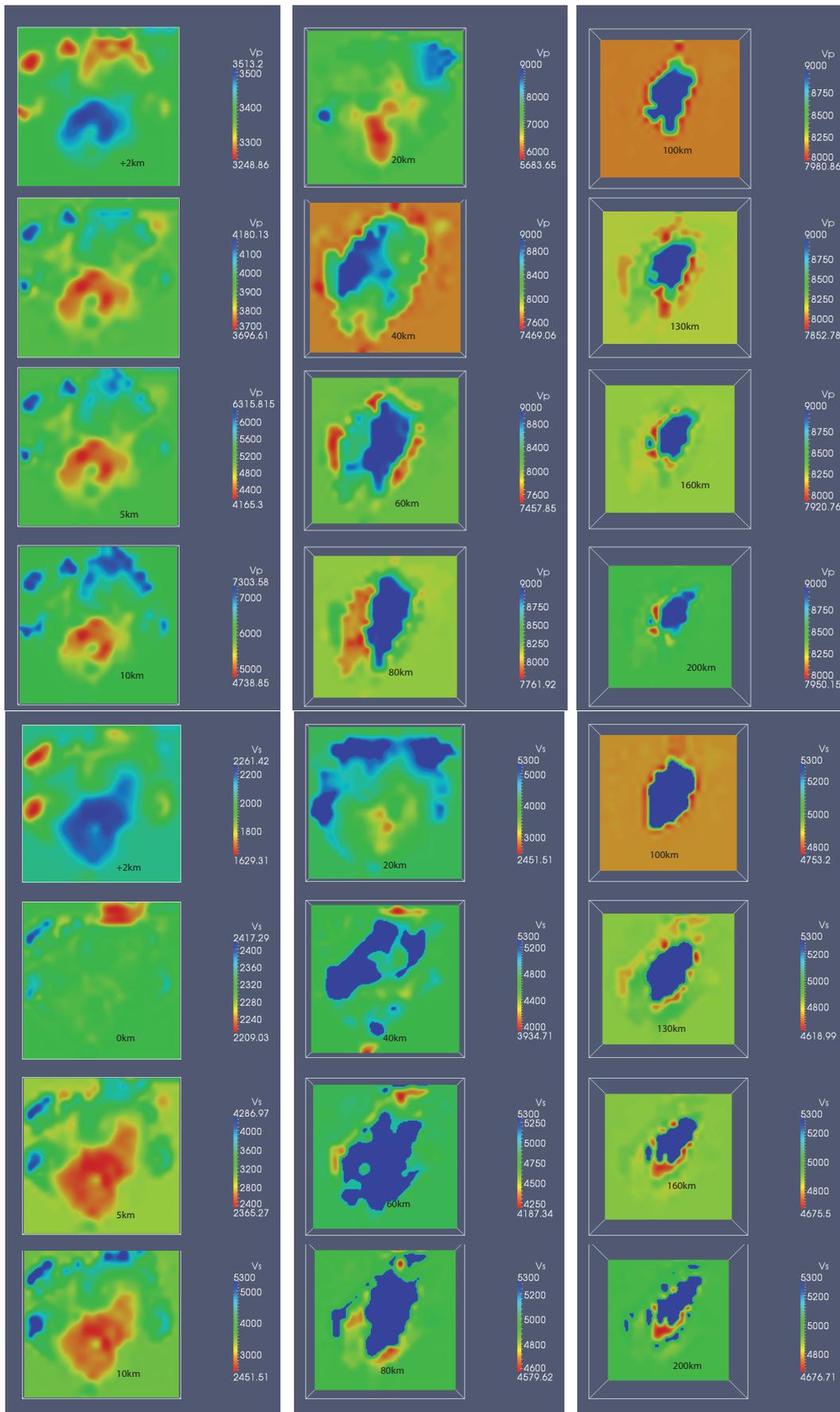


**Figure 32 - 3D view of the corrected initial tomographic model used for the inversion. Velocities are given in m/s.**

**Figure 33 - Slices view at different depths of the initial tomographic model from Tondi et al. (2009). Velocities are given in m/s. Some artefacts are found at shallow depth as well as unexpected velocity values between 40 and 80km depth with a P velocity reaching 10, nearly 11km/s.**

**Figure 34 - Slices views at different depths of the corrected initial tomographic model used for inversion. Velocities are given in m/s. Artefacts at shallow depths have been removed and a maximum velocity threshold has been imposed for the whole model (9km/s for Vp and 5,3Km/s for Vs).**





### IV.I.3. SOURCE LOCATION AND MOMENT TENSOR IN THE 3D MODEL

In their study, Tondi et al. (2009) have relocated the source of 61 events and computed their moment tensor parameters (strike, dip and rake) using the NLLOC<sup>6</sup> and FPFIT<sup>7</sup> algorithms respectively. As a preliminary source reference parameterization for the tomographic inversion, I sought to work with seismic sources parameters computed in my reference 3D model. This 61 events database has thus been used as a reference for our starting event database. This reduced by more than a half the number of events that could be “directly” investigable. Distribution of this event database in terms of magnitude and depth became a little bit less representative of the regional seismicity (Figure 35). In term of magnitude, I have lost the Gaussian-like shape of the event repartition (light violet in the figure), with a Gaussian centred on  $M_w=2,5$  and obtain a quite uniform distribution of events (dark violet in the figure). In terms of depth, I have lost a huge part of the 0-20km depth events, whereas the 80-160km depth events have been reduced by a half. However, the two main depth groups of events have been conserved. If I have lost the representation of the regional seismicity in terms of magnitude distribution, there is still a good representation of these in terms of depth.

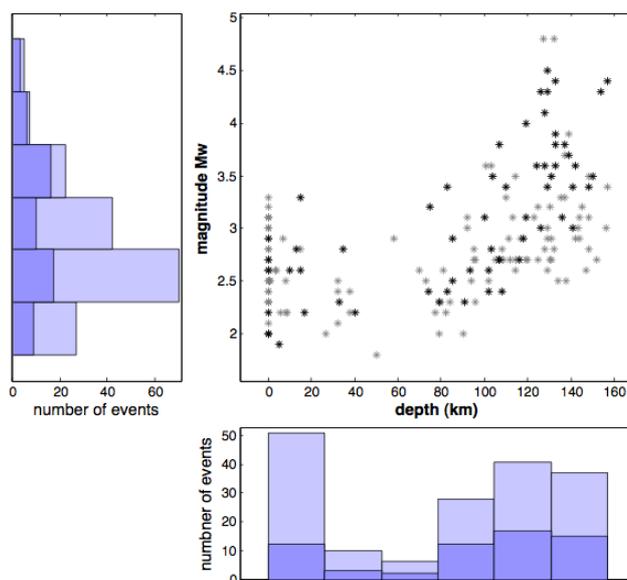


Figure 35 - Distribution magnitude-depth for the CALIXTO99 moment tensor informed input database (black dots and dark violet bars) in comparison with the initial database (grey dots and light violet bars)

<sup>6</sup> The Non Linear LOcation earthquake location algorithm (Lomax, 2000) uses synthetic pick arrival times of seismic waves to estimate the reference time and hypocentral coordinates of an event. It is a grid-search algorithm based on the probabilistic formulation presented by Tarantola (1987) for inverse problems. It relies on the definition of probability density function to get information about the source parameters.

<sup>7</sup> The Fault Plane FIT code (Reasenber and Oppenheimer, 1985) is based on first motion polarities to find the double couple fault plane solution of an event. For each double-couple source model obtained, FPFIT formally estimates the uncertainty in the model parameters (strike, dip, rake) and calculates a uniformly distributed set of solutions within the range of estimated uncertainty.

For consistency, the source parameterization should be updated for each model modification or update. In my case, this has been done in my corrected 3D reference model and after some corrections of the dataset before starting the inversion. I have used the same protocol as Tondi et al. (2009): firstly by relocating the events in my 3D reference model from corrected picking times<sup>8</sup> with NLLOC, secondly by looking for the fault plane parameters using the FPFIT algorithm together with the new event locations obtained.

As a result, it is important to note that the stations coverage and the deep location of the CALIXTO99 events has led to a large uncertainty over the fault plane parameters, with an ample panel of possible solutions (Figure 36).

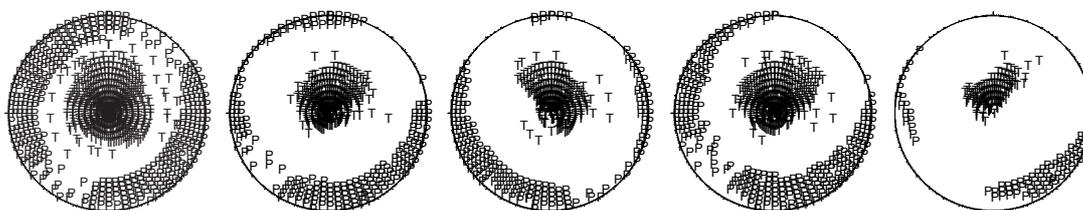
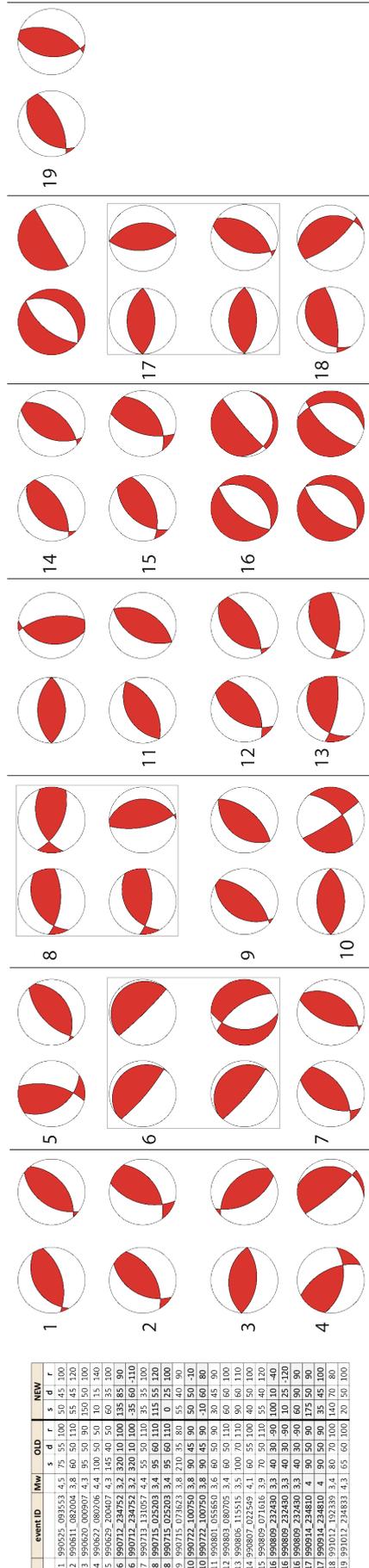


Figure 36 - Sample of the FPFIT fault plane solutions for the five first events of the database (see Table 5 in Annexes A). On the focal spheres are represented all the possible principle axis of the fault plane solutions (P and T points). The diversity in the solutions denotes an important uncertainty about the fault plane geometry.

The moment tensor parameterization obtained with the new source location has given nearly the same fault plane characteristics for the main part of the events (Figure 37). Moreover, regarding the uncertainty among the fault plane definition, it is not surprising to deal with different solutions, even with great differences. I noted as an example, that for some events, highlighted with grey boxes in the figure, there were two possible fault plane solutions, with sometimes very distinct geometries (see event 8 and 10 in Figure 37). Aware of the difficulty to define well-constrained focal mechanisms from the data, I did not give a strong faith to the solutions obtained.

Figure 37 - Fault plane mechanism solutions. Comparison between old and new FPFIT solutions. In each column are given: on the left side the old FPFIT beach balls, and on the right side the new ones. Details for each fault plane mechanism is given in the summary table on left side. Grey boxes stand for multiples solutions for a same event (event 6,8,10,16 and 17).

<sup>8</sup> Details about the picking time correction is given in Appendix C.



## IV.I.4. PROCESSING THE DATABASE

### IV.I.4.1. Noise processing

The broadband input database consists of 10.063 records. As some traces were erroneous or damaged with blockette reading problems, I had to use some the procedure to correct the miniseed volume and keep only non-erroneous files<sup>9</sup>. This resulted in about 30% of loss of the broadband records that includes the loss of entire stations records, where all data for a given station are unreadable. This is the case for eight broadband stations of the network that all have a Guralp-3ESP configuration.

This readable database contains 7.315 broadband traces. A noise selection criterion is done by manual check of the whole remaining database. This step is the costliest one in term of data loss. At the end of the selection only 16% of the initial SAC database remain. Considering only the 61 events, which have a moment tensor parameterization available, this leads to a number of 903 readable seismic traces (Figure 38).

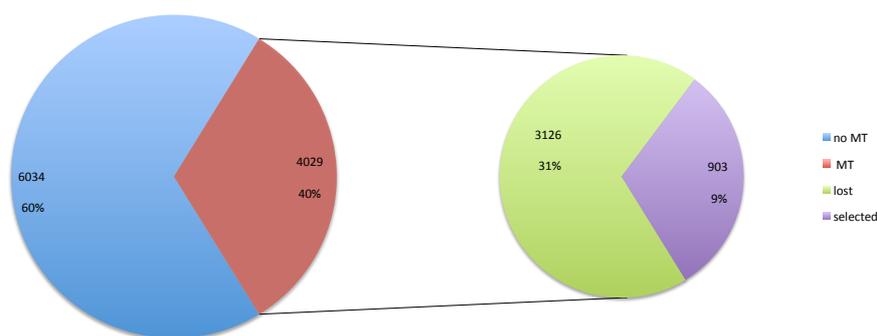


Figure 38 - Amount of remaining traces after the noise filtering among the moment tensor (MT) informed broadband database. The moment tensor informed events represent 40% of the input readable database. At the end of the noise selection, 9% of these events are conserved.

### IV.I.4.2. Instrumental responses

The final data pre-processing step is the instrumental correction procedure by deconvolution of the signal with the instrumental response. For this, dataless files of the miniseed data have to be available into create the response file used to do such a correction. Due to the lack of available dataless volumes, I have been able to handle this processing only for nineteen broadband stations, represented in the following Figure 39 and detailed in Table 5 in Appendix A.

<sup>9</sup> In my case, I have used the msfix tool.

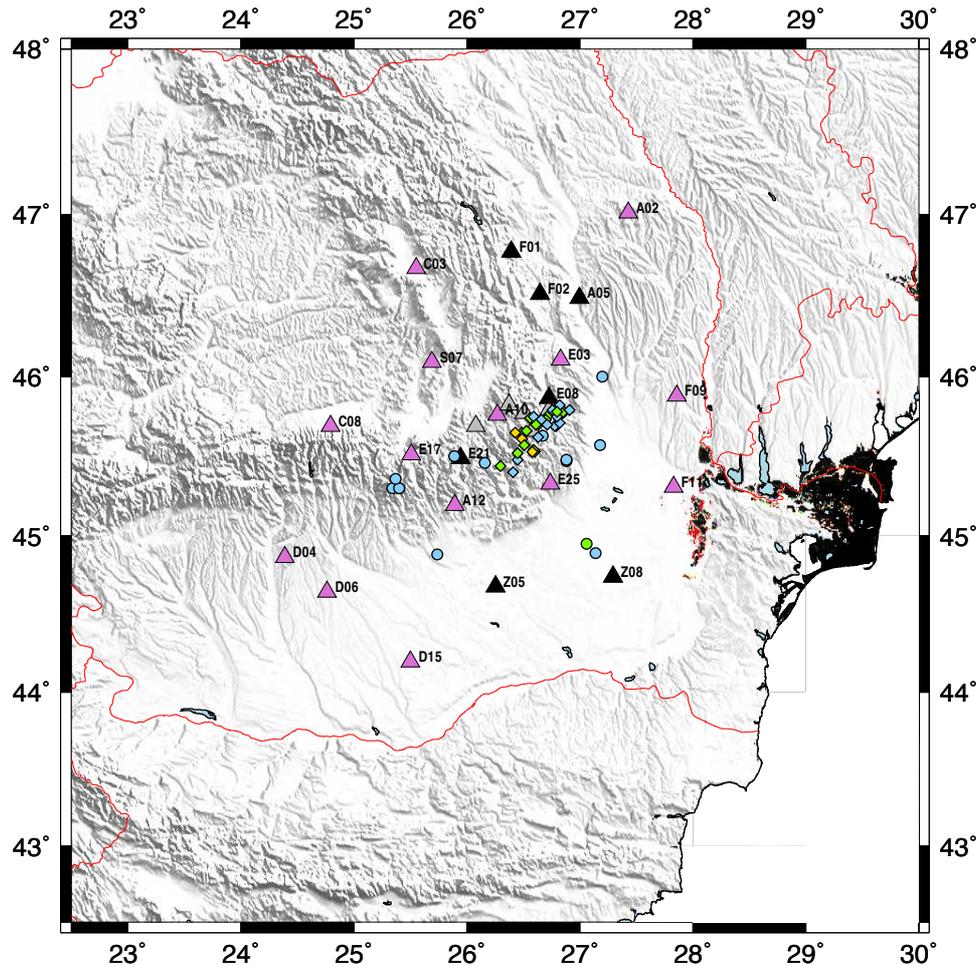


Figure 39 - Directly usable database - distribution of the moment tensor informed earthquakes and instrumental response dotted stations. Stations with a response file are presented in orchid colour, for case of unreadable registrations for a whole station the triangle is darkened. Grey stations are stations without an available response file. The 61 moment tensor informed events are presented with the same legend as in Figure 30: circles stand for shallow depths, diamonds for intermediate depths; blue colour for  $M < 3$ , green for  $3 < M < 4$  and yellow for  $4 < M < 5$ .

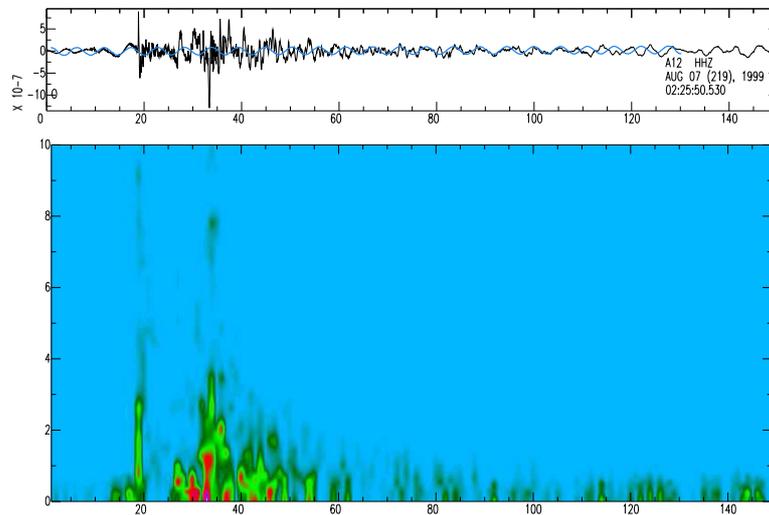
#### IV.I.4.3. Frequency content

The tomographic inversion implementation depends on the frequency range to investigate. This one governs the scale of the velocity anomalies illuminated by the sensitivity kernels<sup>10</sup>. Working at a low frequency range, one is able to individuate large scale heterogeneities but will remain blind of small ones that can be seen only at smaller frequency. This frequency range depends on one hand on the data quality and noise level of the seismic traces, on the other hand, on the solver capacity to compute accurately high frequency synthetic waveforms. Indeed, high frequency computations are limited by computational resources and power as, such simulations require ever smaller spectral elements and thus a large amount of computations.

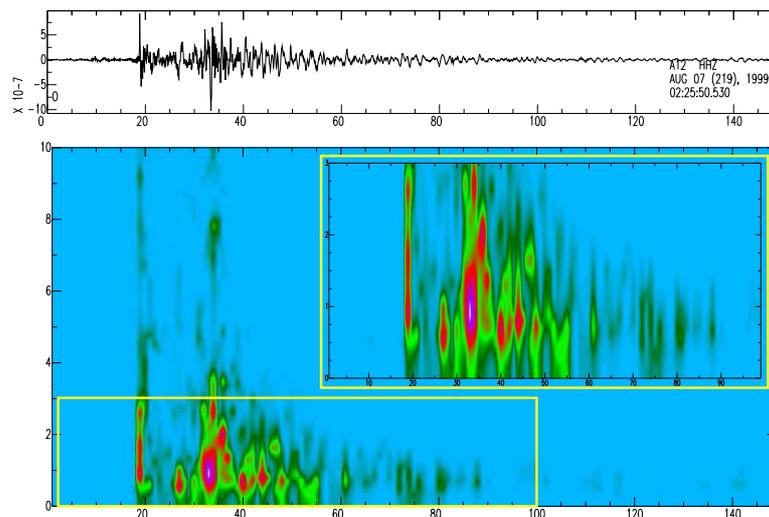
<sup>10</sup> Sensitivity kernels have been defined previously in II.I.2 pp. 22. They can be seen as fat seismic rays illuminating the path from seismic source to receiver.

Spectrograms of the seismic records, showing the frequency distribution along time, give crucial information on the data frequency content (Figure 40 and Figure 41). The frequency content of my event database is divided in a low frequency band around 0.2Hz and a high frequency band above 1Hz. Looking at the waveforms, I could note that the traces are disturbed by a five seconds period noise (blue sinusoid on Figure 40). This noise range corresponds to the low frequency content visible in the spectrograms (0,2Hz band) and can be removed applying a high-pass filter on the seismic record (Figure 41). It results that the main frequency content of the signal stays between 0,4 and 2Hz.

Up to now, computational resources allow reaching synthetic seismograms simulations until 1Hz accuracy. Thus, the final frequency range I could retrieve from the records stays between 0,4 and 0,8 Hz, which corresponds to periods from 1,25 to 2,5seconds. This range of frequency is considered as high frequency for full waveform inversion and brings heavier computations than a usual lower frequency full waveform inversion.



**Figure 40 - The spectrograms are a representation of the frequency distribution (y axe) along the time (x axe). Sample from the 990807 event at station A12, after deconvolution in a frequency band between 0,1 and 10 Hz. The signal stays in a frequency band between 0.2 and nearly 3Hz.**



**Figure 41 - Sample from the 990807 event at station A12, after deconvolution and high pass filtering at 0,4 Hz. The right side yellow box is a zoom of the left side box between 0 and 3Hz. The main part of the seismic signal stays between 0,4 and 2Hz.**

#### IV.I.4.4. Final database for inversion

To implement a full waveform inversion of the Vrancea region, I built a concise database taking heed of the quality of the seismic records and their stations coverage.

Owing to the particular seismicity of the Vrancea region, the largest events are all concentrated in the same area at deep depth (Figure 42). Strongest events benefit from cleaner signals, as the seismic energy is able to surround the micro-seismic noise. At shallower depth, events are sparser and of minor magnitude (below Mw 3) with a poor quality signal.

The final database is made of the strongest deep local events recorded during the CALIXTO99 experiment by at least 3 seismic stations, with a magnitude range from 3,4 to 4,5 (see orange lines in Table 4 in Appendix A).

This is a quite reduced amount of seismic events for a local tomographic inversion that do not give a large lateral coverage of the Vrancea lithosphere. Nevertheless this comes together with the particularity of the Vrancea nest and the CALIXTO99 station's network. This tomographic study aims to show in which extent the finite-frequency theory and the full waveform approach allow retrieving interesting information from the Vrancea seismogenic volume.

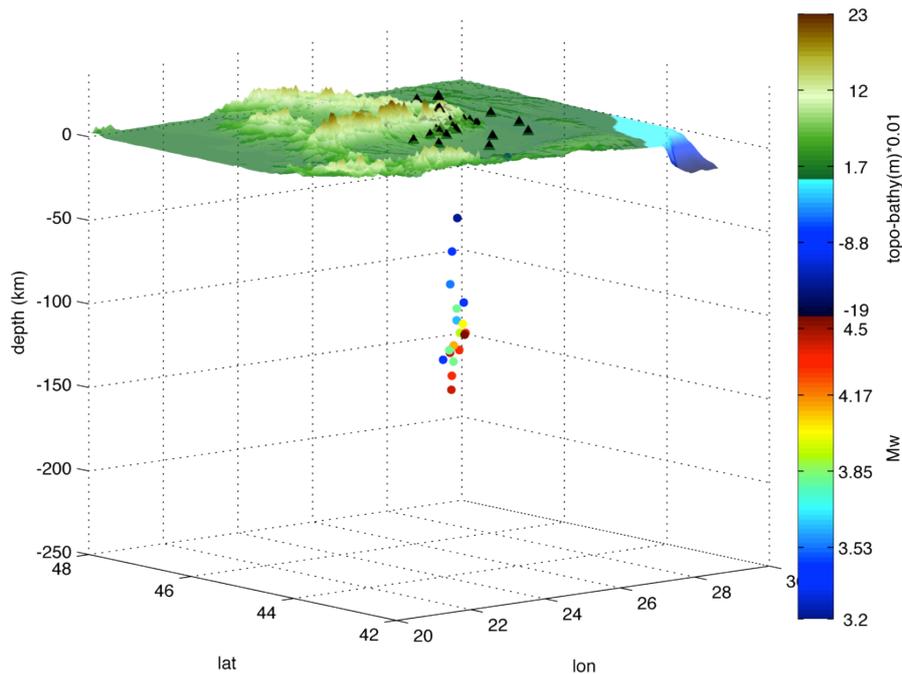


Figure 42 - Final event database distribution used for the full waveform inversion of the Vrancea region. The topography is reduced by a factor of 0.01. 19 events registered by at least 3 stations with a magnitude Mw between 3,4 and 4,5 have been selected.



## IV.II. SEISMIC WAVEFORM MODELING

---

Most of the time, seismic tomographic studies require an important amount of computation that an individual computer is not able to sustain. High Performance Computing (HPC) facilities, based on parallel computing, thus aim to host such kind of large scale computations. During these three years, I have used the facilities of the CINECA HPC centre in Bologna, Italy.

### IV.II.1. WAVEFORM MODELING WITH SPECFEM

#### IV.II.1.1 Introduction

SPECFEM tools gather a set of open-source parallel software packages freely available via the Computational Infrastructure for Geodynamics (CIG). The software simulates seismic wave propagation using the Spectral Element Method, which is being actively researched, extended and applied in the geophysical context (Komatitsch et al. 1999, 2002a,b, 2004, 2005, 2010; Luo et al. 2009; Peter et al. 2011; Tromp et al. 2008; Fichtner et al. 2009). The several variants of SPECFEM cover the range from 1D implementation to 2D and 3D simulations for global or regional scale tomographic inversion.

In particular, the SPECFEM3D package simulates forward and adjoint seismic wave propagation at local or regional scale on arbitrary unstructured hexahedral meshes.

A SPECFEM simulation consists in (Figure 43):

1) Mesher/Partitioner: defining the mesh and computational geometry <sup>11</sup>

Meshing consists in designing an equivalent Earth's model geometry in which to compute the seismic wave equation.

The partitioner allows dividing this skeleton between computational resources.

2) Databases: building the computational mesh

When the mesh geometry is defined, the computational grid can be built up. P-S velocity and density values are assigned at each point of the geometrical mesh so that it becomes a computational environment where to solve the equation of motion.

3) Solver: Solving the seismic wave equation

Solving the equation of motion, we finally produce the synthetic seismograms related to a given source-stations and velocity model set-up.

---

<sup>11</sup> Meshing can be accomplished by using an external mesher like CUBIT or by using the SPECFEM3D's own internal mesher.

The workflow is different when using internal or external mesher: the external mesher benefits from 3D mesh partitioning through the routine SCOTCH whereas the internal does not. The internal mesher partitions the mesh into slices that correspond to vertical portions of the mesh.

To build up my geometrical mesh, I have used the internal mesher and will explain in the following part the strategy employed.

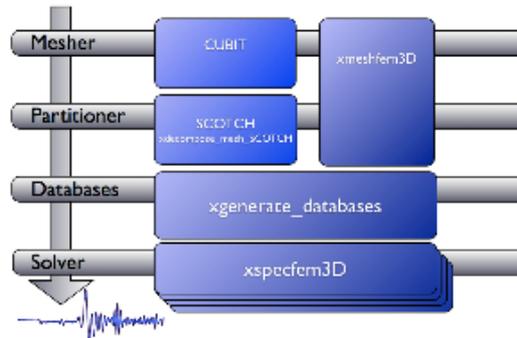


Figure 43 - SPECfEM3D workflow to compute synthetic seismograms in a 3D Earth model.

### IV.II.1.2. Build the computational mesh with the internal mesher

The internal mesher *xmeshfem3D* defines the geometry of the mesh to build. Principal information about the mesh include:

1) the geographical coordinates of the block and its depth

2) the regularity of the mesh

The size of the spectral elements has to reflect the wave speed distribution throughout the model. Using a tomographic model with velocity increasing with depth implies that elements have to become larger in depth to define a homogeneous number of grid points per wavelength in the block. For this scope, it is possible to place doubling layers in the mesh that define interfaces from where the spectral elements become twice larger (see Figure 44).

3) the physical properties of the block in terms of density, seismic velocities and anisotropy and if these properties are applied on the entire block or on given portions.

4) the number of spectral elements in the mesh along the horizontal and vertical components

This number is linked to the resolution of the mesh one wants to achieve. Indeed, the accuracy of the simulations is determined by the number of grid points per shortest wavelength that has to be homogeneous throughout the model: the smaller the spectral elements, the shorter periods (or high frequencies) can be simulated.

When set these numbers together with the size of the spectral element in the surface, one has to set the number of spectral elements in the vertical direction. This has to be done regarding the presence of eventual doubling layers so that to obtain squared elements.

5) the number of MPI processors along each horizontal direction<sup>12</sup>

If the mesh is not regular and contains mesh doublings this number is set so that one MPI processor manages the computation over a multiple of eight spectral elements (Figure 44).

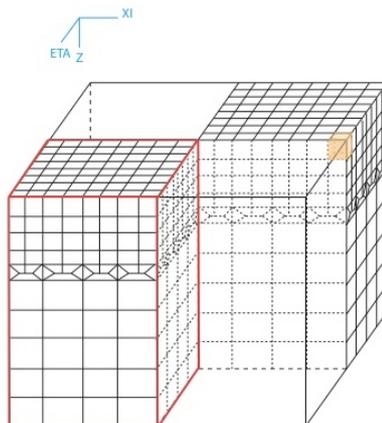


Figure 44 - Example of an irregular mesh consisting in 16 spectral elements (orange box) along the horizontal directions XI and ETA and a doubling layer. The number of MPI processors along XI and ETA is 2 ( $16 = 8 \cdot 2$ ), represented by the red box. The doubling layer makes elements twice bigger in horizontal direction and the number of elements in Z direction as to take it into account to have squared element. Thus there are 4 elements in the first interface and 5 in the second one.

## IV.II.2. FORWARD MODELING ON CINECA BLUE GENE/Q

### IV.II.2.1. The Vrancea CALIXTO set up

#### IV.II.2.1.1. Input files

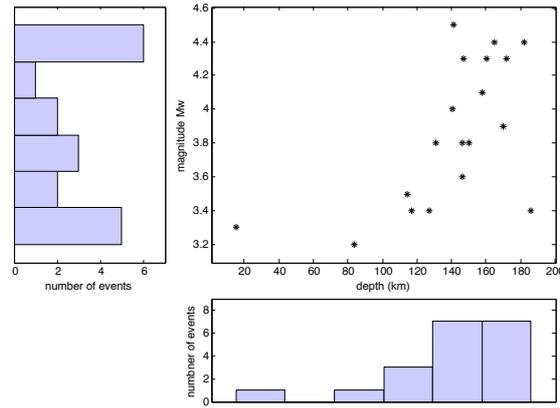
I have used the 19 local events selected from the CALIXTO99 database detailed in Table 4 in Appendix A. The distribution of this dataset is given in Figure 45, with the main part of events presenting depths between 140km and 160km and a magnitude going from Mw 3,4 to 4,5.

From fault plane parameters, I have written the moment tensor source files needed in SPECFEM3D.

The reference tomographic model has been taken from the Tondi et al. (2009) study after some corrections explained previously. It is a  $230 \times 230 \times 230$ km P-S velocity-density grid points covering the geographical area from latitude 25 to 28 and longitude 44,36 to 46,43. More details about the implementation of this model in SPECFEM3D can be found in Appendix D.

The topography has been drawn from the ETOPO topographic/bathymetric dataset.

<sup>12</sup> An MPI (Message Passing interface) system provides communication functionality between a set of processes that have been mapped to nodes/servers/computer facilities.

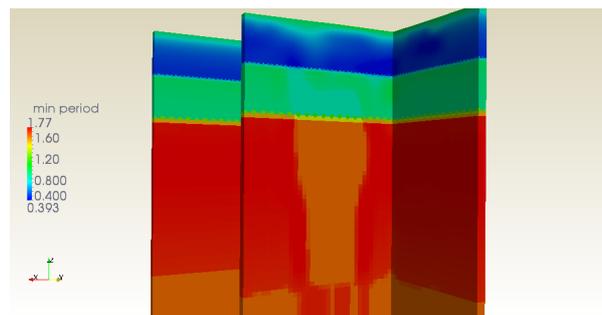


**Figure 45 - Magnitude/Depth distribution of the 19 events dataset. Nearly all the selected events stay in the depth range from 100 to 180km.**

IV.II.2.1.2. Computational mesh

The geometrical mesh has been built bigger than the reference tomographic model to avoid eventual boundary effects. It is thus a 500\*500\*500km grid - centred on the tomographic model- made of 14Millions spectral elements and two doubling layers at 40km and 80km<sup>13</sup>. On the surface, the number of spectral elements in each direction is of 512 for an element size of less than 1\*1\*1km. At the bottom of the mesh, the element size reaches 4\*4\*4km. This set-up requires a number of MPI tasks of 4096. The computational configuration I built involves 4096 computational cores (CPU, Central Processor Unit). An in-depth description about the computational mesh configuration, with an emphasis on the implementation of SPECSEM3D on the resources of the supercomputer centre CINECA, is given in Appendix D.

This mesh allows computing seismic waveforms in the frequency range imposed by the dataset that is from 1,25 to 2,5seconds, and might be accurate around 1sec period and less (Figure 46).



**Figure 46 - Distribution of the minimum period resolution through the mesh. The resolution depends on the velocity values as well as the spectral element size for a given region.**

IV.II.2.4 Computational costs

The mean elapsed time to run a forward simulation is about 3h40, leading to a computational cost of about 15.000 CPU hours (a CPU hour is defined as the

<sup>13</sup> See Figure 68 in Appendix B for a justification of the doubling layer depths.

product of the number of computational core used for the simulation with the elapsed time: 4.096\*3h40).

### IV.II.2.2. First waveforms comparisons

Before being compared, data and synthetic waveforms have to be processed. This consists of sampling signals at the same rate, correcting data from the instrument response, and filtering both signals in the same frequency range, which has been chosen to be [0,4 - 0,8]Hz. The waveforms have been then compared using the Seismic Analysis Code tool SAC (Figure 47).

Picking P and S arrivals, I observed a recurrent time shift of the synthetic body waves seismic arrivals. The mean P shift is of about 1,5second and the S shift is of about 4seconds (Figure 48).

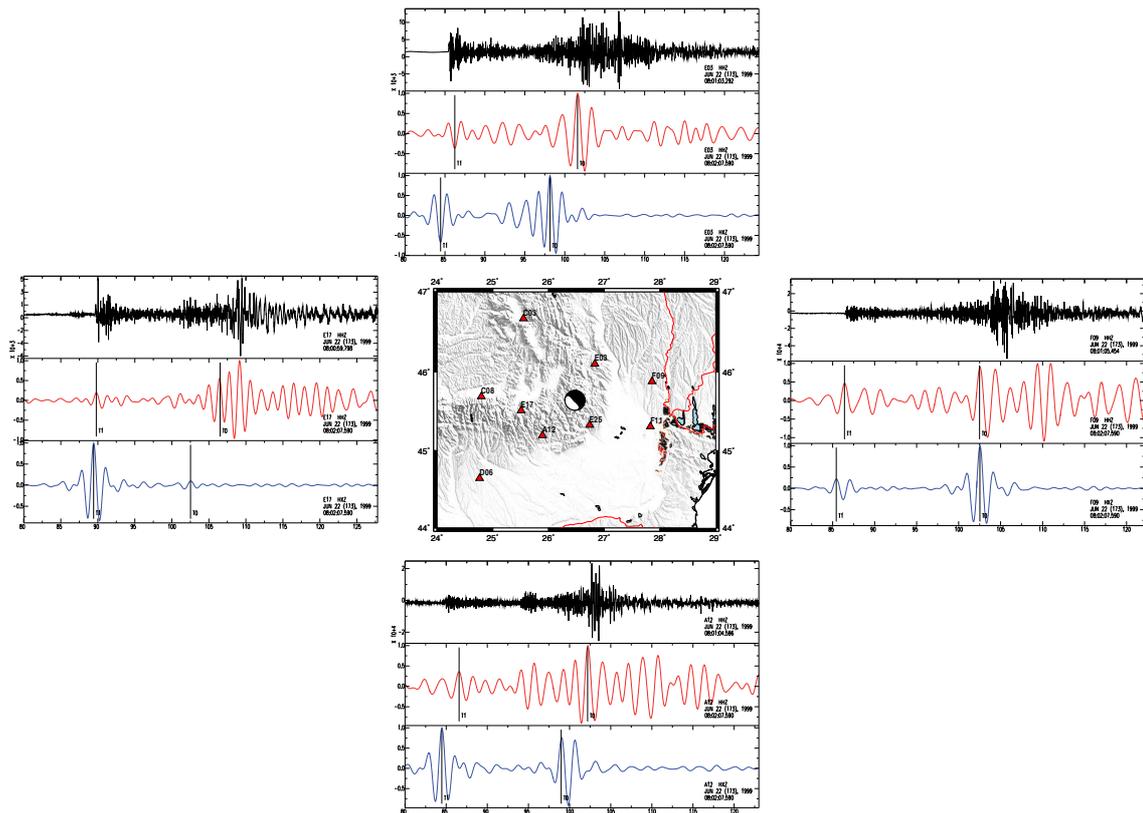


Figure 47 - Sample of forward simulations (blue signal) for the 990622 event and comparison with real data (black for rough signal and red for processed one).

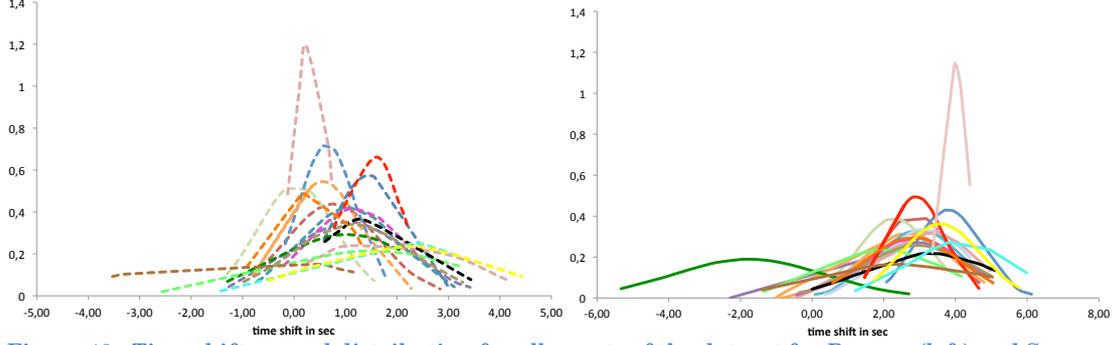


Figure 48 - Time shift normal distribution for all events of the dataset for Pwaves (left) and S waves (right). Each coloured curve represents an event distribution. The P waves time shift is around 1,5 second whereas the S waves time shift is around 4 seconds.

To help the misfit convergence I needed to start the inversion with a satisfying reference model giving good misfit results.

To reduce the time shift difference between the P and S waves arrivals in the reference tomographic model, I applied a depth correction. For each event, this correction is estimated using the mean P and S shifts measurements and the maximum velocity values in the model:

$$depth_{correction} = \frac{-v_{pmax}}{0,7} (mean_{shift P} - mean_{shift S}) \quad (81)$$

$$depth_{new} = depth + depth_{correction}$$

A second adjustment, consisting in having a mean shift around 0 is achieved by correcting the reference time for each event based on the respective P shift (Figure 49). As a result, I obtained an initial dataset where pick arrival time shift normal distribution for P and S waves are centred on zero with a standard deviation of about 1 and 2 seconds respectively (Figure 50).

Looking at the depth distribution with respect to the CALIXTO99 archive values, I note an improvement in the consistency between the different localizations (Figure 51). First of all, the CALIXTO99 archive depth values (grey line) are very consistent with the ROMPLUS catalogue (black crosses), which can be an indication of the accuracy of the given values. Looking at the NLLOC initial localization results (blue triangles), I noted a loss of this consistency with a main part of the depths between the +/-25km difference with respect to CALIXTO99. With the NLLOC re-localization (green triangles) this consistency between catalogues comes back and is relatively strong. Finally, the empirical depth correction applied to the 19 events selected for the tomographic inversion, introduces a +25km mean depth difference with respect to the archive values (orange circles and light orange line for the 25km depth difference).

In terms of reference time estimation, the empirical manual corrections led to a better consistency with the CALIXTO99 archive values for almost all the 19 selected events, with only 5 events with an increased difference going up from 2 to 3 seconds (Figure 52).

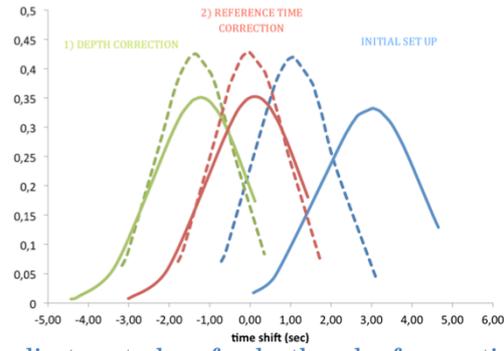


Figure 49 - Example of the adjustments done for depth and reference time on the 990622 event. The dashed lines stand for P time shift distribution whereas the plain lines stand for S measurements. Starting from an initial set up with a nearly 2 seconds shift between P and S distributions (blue curves), the depth correction (green curves) allows to zero this shift. In a second step, the reference time correction (red curves) allows to zero-centered the distributions.

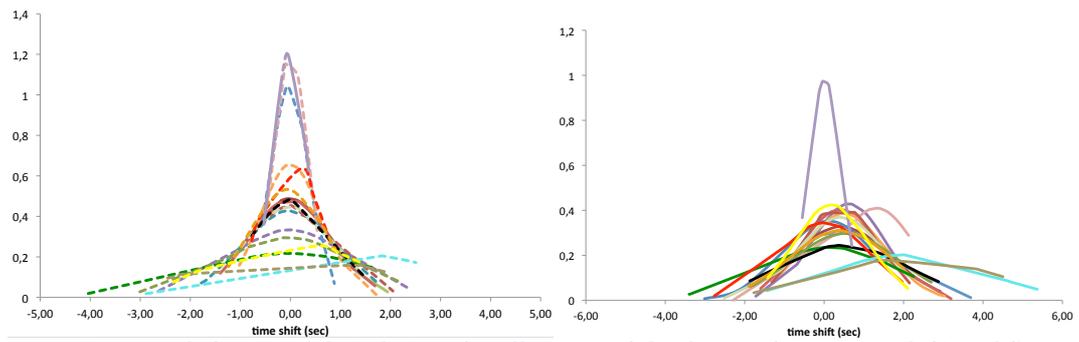


Figure 50 - Time shift normal distribution for all events of the dataset for Pwaves (left) and S waves (right) after adjustment of depths and reference times.

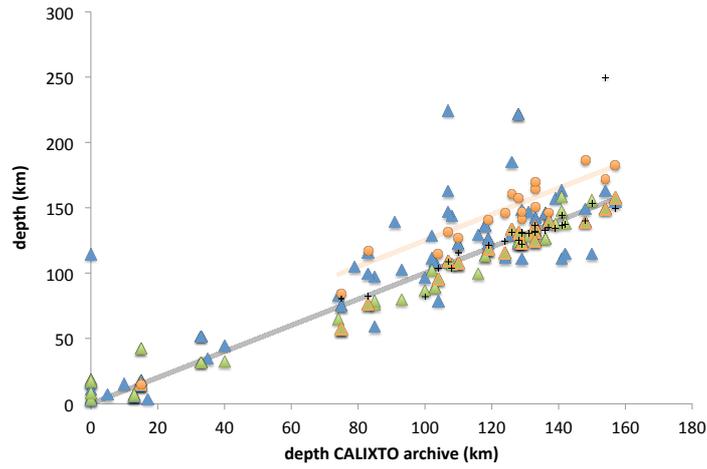


Figure 51 - Depth comparison with respect to the CALIXTO99 archive values. Blue triangles are for the initial depth values from Tondi et al. (2009), green triangles are for the depth values after the new NLLOC localization, orange circles are the 19 events of the selected dataset from which depth has been empirically corrected regarding the pick time shift observed between synthetic and real waveforms. Finally black crosses indicate the ROMPLUS catalogue values.



Figure 52 - Reference time comparison with respect to the CALIXTO99 archive values for the 19 selected events. Green bars refer to the NLLOC reference time before the correction and orange ones, after the correction. All events become more consistent with the CALIXTO99 archive and only 5 events come out with farther values.

## IV.III. STRUCTURE INVERSION

### IV.III.1. FROM MISFIT TO ADJOINT SOURCE

#### IV.III.1.1. Selecting the windows

The misfit measurement is made in the portions of the seismograms that give best correlations between real data and synthetic signals. This can be done automatically using the FLEXWIN software (Maggi et al. 2009) that selects misfit windows with respect to a set of parameters. Working with relatively low quality data and at high frequency I preferred doing this selection manually, selecting simple window boxes for P and S arrivals (Figure 53). When possible, I applied the same window selection for all three components of a given seismic records.

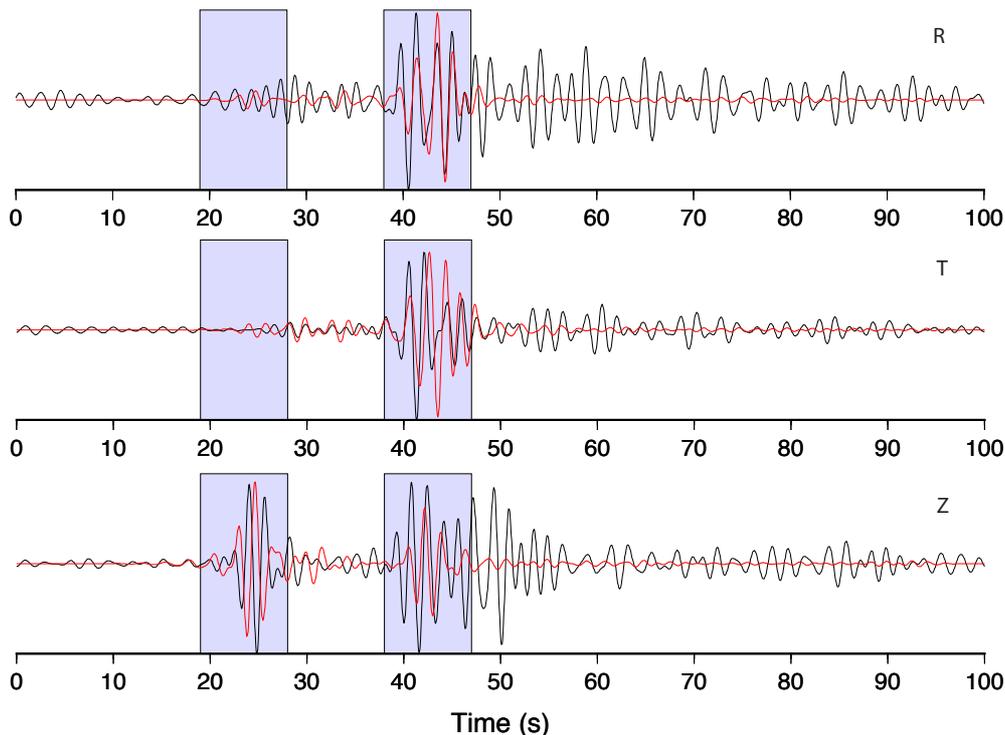


Figure 53 - Selection windows example with the event 991012\_234833 at station A12 for a cross correlation travel-time misfit. The seismograms have been scaled for a better reading. The frequency range stays between  $[0,4 - 0,8]$  Hz. In this case, I have selected a P and S window for all three components.

### IV.III.1.2. Compute the misfit and adjoint sources

I have worked with the cross-correlation travel-time misfit measurement. It is defined as the least-square difference over the station-source pair of the cross-correlation time delay between observed data and synthetic waveforms computed within a given model  $m$ :

$$\chi(m) = \frac{1}{2} \sum_{i=1}^n (T_i^{obs} - T_i^m)^2 \quad (82)$$

Where  $n$  is the number of seismic stations and  $(T_i^{obs} - T_i^m)$  is the time delay of maximum cross-correlation between the real and synthetic seismograms for the station-source pair  $i$ .

In practice, the misfit is computed individually in the selected windows of each seismic record for the windows that present a minimum cross-correlation coefficient of about 70%. Results from the cross-correlation computation in the selected windows (Figure 55) are quite similar to the preliminary study I have done by picking first body waves arrivals. These measurements present time shifts for P and S waves centred on 0 with a standard deviation until 4 seconds in some cases.

From misfit measurements, I have built the adjoint sources of each event to compute successively the misfit kernel. Adjoint sources are the velocity waveforms at each station weighted by the misfit measurements in reverse time.

To compute the misfit value regarding to a particular model, I have defined two other misfit values: the misfit for a particular component (R, T, Z) and the misfit regarding a given event.

The “component-misfit” value is taken as the sum of all “window-misfit” measurements, weighted by the number of windows. Consecutively, the “event-misfit” is taken as the sum of all “component-misfit” measurements, weighted by the number of components. Finally the “model-misfit” is the sum of all “event-misfit” values, weighted by the number of events (Figure 54).

For simplicity, I did not introduce any kind of extra weighting function in the misfit computation.

Window-misfit	$\chi_w = \frac{1}{2} (T^{obs} - T^m)^2$
Component-misfit	$\chi_c = \frac{1}{n_w} \sum_1^{n_w} \chi_w$
Event-misfit	$\chi_e = \frac{1}{n_c} \sum_1^{n_c} \chi_c$
Model-misfit	$\chi(m) = \frac{1}{n_e} \sum_1^{n_e} \chi_e$

Figure 54 - Misfit values computation. From window-misfit to model-misfit.



## IV.III.2. FROM EVENT KERNELS TO FIRST UPDATE

Owing to computational difficulties, I am able to present here first results of the tomographic inversion only for a representative sample of the dataset. It consists of five events (990620, 990622, 990713, 990807 and 990914) that have been selected regarding their signal quality and station coverage.

### IV.III.2.1. Misfit kernels

The misfit kernel is a representation of the gradient of the misfit function for a given model used for solving the minimization problem.

In SPEC3D, the computation of the misfit kernel is based on executing the following steps for each event:

- 1) a forward simulation in which all time step of the wave field path is saved
- 2) an adjoint simulation where the interaction between the previous forward and a reversed time simulation is computed for each source-station pairs. At this step, the source is replaced by the adjoint sources built with the misfit measurements and placed at stations positions. For each source-station pairs, a volume sensitivity kernel is thus computed.

The sum of all these volume sensitivity kernels related to a given event is called an event-kernel, which is the output of the SPEC3D simulation. Finally, the sum of all these events-kernel for a given model is the misfit kernel, representing the gradient of the misfit function for a given model. This sum is computed successively in parallel.

A first positive observation done from my study is that even if working at high frequencies I was still able to design fat sensitivity kernels (Figure 57) and thus there was still an advantage to work on these data using the finite frequency theory rather than the ray-based one.

A second observation dealt with the presence of some high amplitude kernels singularities near source and receivers locations (Figure 57) that make the volume kernel non-homogeneous along its length. These high amplitude values can compromise the model updating efficiency leading to incorrect model corrections. To eliminate these spurious amplitudes in the neighbourhood of the source and receivers as well as artificial effects due to unresolved features in the model a smoothing of the misfit kernel has been performed. In our case, the smoothing operation consisted in the convolution of 3D Gaussian function with the misfit kernel upon all spectral elements of the mesh.

$$G(x, y, z) = \frac{1}{\sqrt{(2\pi)^3 \sigma_x^2 \sigma_y^2 \sigma_z^2}} e^{-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} + \frac{z^2}{2\sigma_z^2}\right)} \quad (83)$$

$$Kernel_{smoothed}(x, y, z) = Kernel(x, y, z) * G(x, y, z) \quad (84)$$

The parameter  $\sigma$  defines the standard deviation of the Gaussian function for each direction  $x$ ,  $y$  and  $z$ . It can be chosen with respect to the shortest wavelength resolved in the simulation by taking into account the size of the spectral elements.

The larger is the Gaussian (for larger scale smoothing), the more spectral elements are taken into account in the processing and the heavier is the computational cost of the smoothing (Figure 56). For me, a favourite choice would have been a 4-5km scale smoothing that is a little bit larger than our bottom mesh spectral element size (4km). However, such a smoothing scale length is not implementable due to the high computational cost involved (more than 80.000 CPU: more than 20 hours computational time \* 4096 processors). Instead, I could afford to implement a 2km smoothing involving around 50.000 CPU (12 hours computational time \* 4.096 processors).

Applying this smoothing I obtained quite satisfying results with the main part of the singularities removed (Figure 57). Nevertheless this smoothing is not large enough to remove all of them and to better homogenize the whole volume kernel rendering a pre-smoothing operation necessary (Figure 57 and Figure 58).

I created a mask operator consisting in a threshold function that intends to remove the sources/receivers misfit kernel singular values by imposing a maximum value. The mask has been implemented in parallel and work at slice level for the entire mesh.

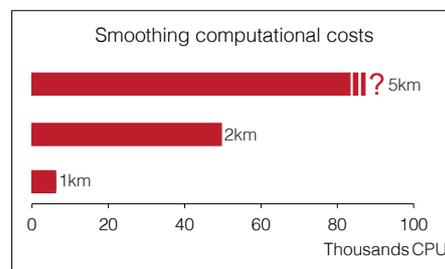
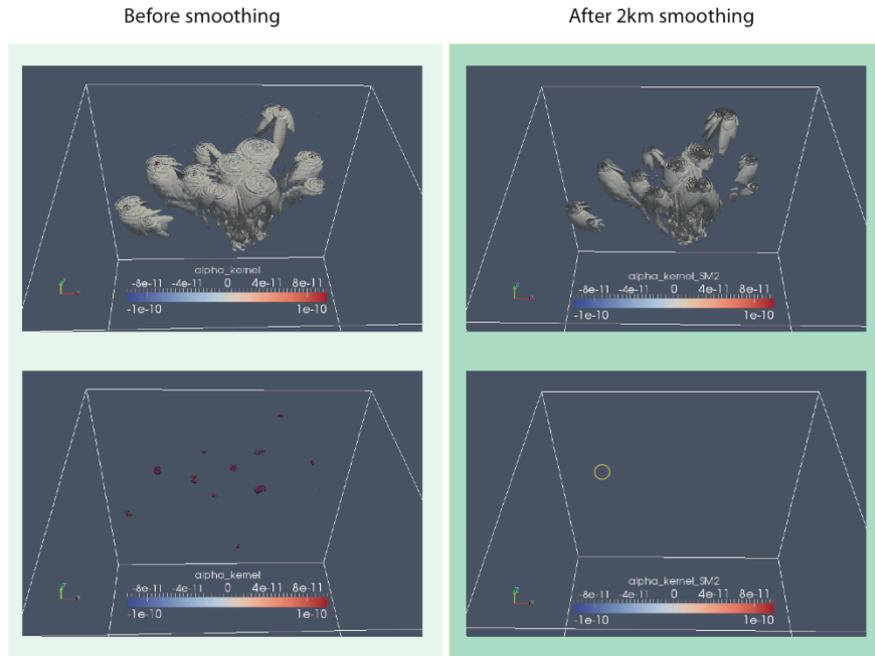
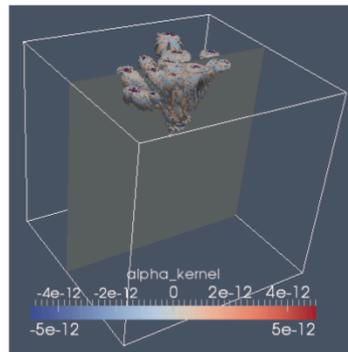


Figure 56 - Computational costs involved by the smoothing operation on misfit kernels for a 1km, 2km and 5km standard deviation.

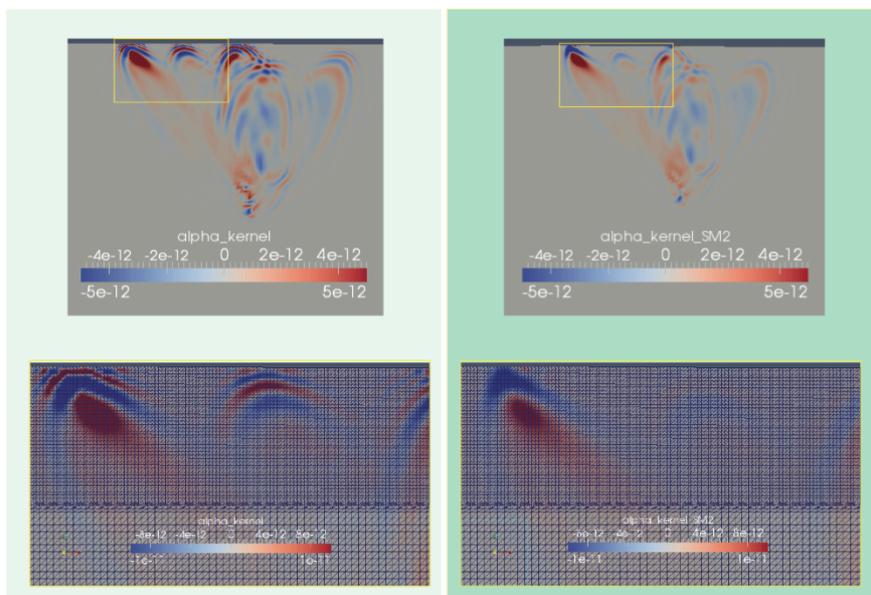


**Figure 57 - Smoothing effect on a 5events misfit kernel in 3D. The 2km Gaussian smoothing has removed main part of the amplitudes singularities.**



Before smoothing

After 2km smoothing



**Figure 58 - Details of smoothing effect on a 5events misfit kernels on 2D planes. High amplitudes still exist in the vicinity of source and receivers at lower scale.**

### IV.III.2.2. Update

The aim of the tomographic inversion is the minimization of a given misfit function  $\chi$  (in this study the cross-correlation travel-time misfit) for a model  $m$  of  $p$  parameters:

$$\min_m \chi(m) \quad (85)$$

As minimization method, I have used the steepest descent algorithm that iteratively minimizes the misfit function through model updates using its gradient  $g_n$  and a step length parameter  $l_n^{test}$ . As explained previously, I have computed the gradient under the form of misfit kernel using the adjoint method.

$$m_n^{test} = m_n - l_n^{test} g_n \quad (86)$$

For each model update, the step length parameter that defines the extent of model variations is found empirically above a set of trial values. The value giving the best misfit is taken as the best step length for the model update. To save computational time, one could work with a representative sample of the dataset to define this parameter and instead of repeating it for each model update, change it only after a given number of iterations.

## IV.III.3. FROM INITIAL MODEL TO 3<sup>rd</sup> UPDATE

### IV.III.3.1. Results

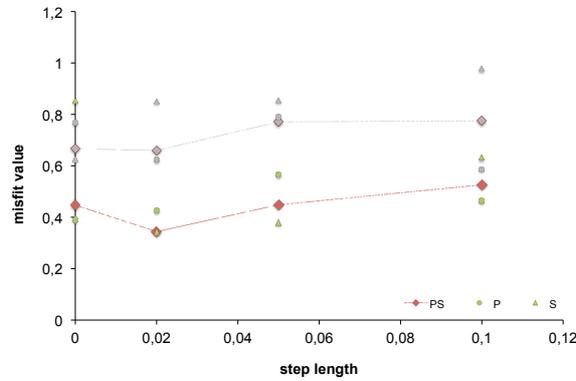
#### IV.III.3.1.1. Misfit

The cross-correlation measurements have been done in selected windows of the seismograms that passed several criteria, such as at least a 70% correlation value. Consequently, not the same windows are always taken into account in the overall comparison between synthetic from a given model update and real data. This is a good way to see how model updating enable to take into account even more windows measurements in the misfit calculation, indicating an increasing fit between the synthetics and the data. However, this makes hard any comparison between one update and another. To be able to compare the updates rigorously with respect to the initial model and with respect with each other, I looked also at a second interpretation of the measurements, imposing strictly the same measurement windows for each model. These two measurement types are defined in the following figures as “all windows” measurements and “same windows” measurements.

For the first model update, I have tried a set of step length values with the aim to find the one giving the best misfit reduction. This trial set consisted of a 2%, 5% and 10% variation allowed with respect to the initial tomographic model M0. I have then looked at the travel time cross-correlation and misfit measurements for these updates (Figure 59).

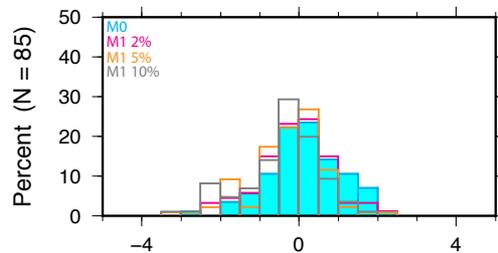
I observed that the misfit (diamonds) reaches a minimum for the 2% step length model update. This can be seen either for the “all windows” misfit measurement

(grey symbols) either for the “same windows” one (coloured symbols). Separating the measures done in P-windows and S windows (green circles and triangles respectively), I noted in a second time, that the P and S misfits have different behaviours.



**Figure 59 - Travel-time cross-correlation misfit measurements at first update for 3 different steps length (for the 5 events tomographic inversion). Coloured symbols define misfit measurement in identical windows for all models, instead grey symbols stand for measurements in all respective windows of each model. The minimum misfit value (diamonds) is found for the 2% step length.**

Looking at the travel-time cross-correlation measurements in identical windows for each model updates with respect to the initial model (Figure 60), I have had the confirmation that the 2% step length updates is the best choice for the first model update M1, as it makes the Gaussian-like measurements distribution shaper.

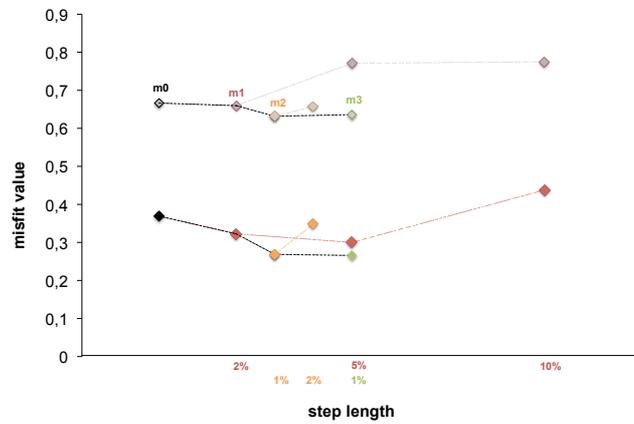


**Figure 60 - Travel-time cross-correlation distribution at first update for 3 different steps length (for the 5 events tomographic inversion) in the same identical windows.**

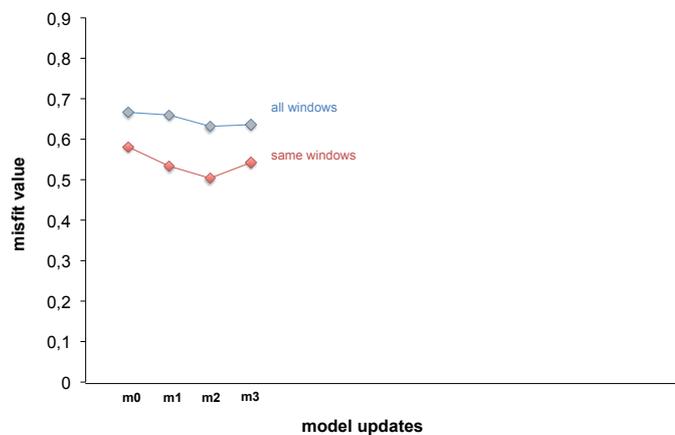
I repeated these measurements for the successive updates until the third one. Figure 61 shows in details the misfit measurements variations among the step length choices. For each model iteration trial with a given step length, the misfit value is represented with the colour assigned to the model. For example, the first model update is represented with red symbols and each misfit values obtained with the 2% 5% and 10% step length are thus given in red symbols. The best model update is finally indicated with its name (m1 for the first model update). As before, I indicate in the figure the values for an “all windows” (grey symbols) and “same windows” (coloured symbols) measurement.

For the second update, the best step length seems to be for a 1% variation, as the 2% one makes the misfit increase. For the third update, I tried a 1% step length but it seems it does not improve the misfit measurement, which remains nearly constant between M2 and M3. At this point, decreasing again the step length

would not be efficient anymore for the whole iteration process. One solution to help the misfit convergence may be to add some events to the dataset. Indeed it is to remind that I am presenting here a reduced tomographic set of 5 events, among the 19 selected for the inversion.



**Figure 61 - Travel-time cross-correlation misfit measurements at first three iterations. Coloured symbol define misfit measurements in identical windows for all models, instead grey symbols stand for measurements in all respective windows of each model: black for M0, red for M1 and orange for M2. Results are relative to PS windows measurements. The measurements obtained at each trial steps length is also indicated with the colour relative to the model.**



**Figure 62 -Misfit measurements among model updates.**

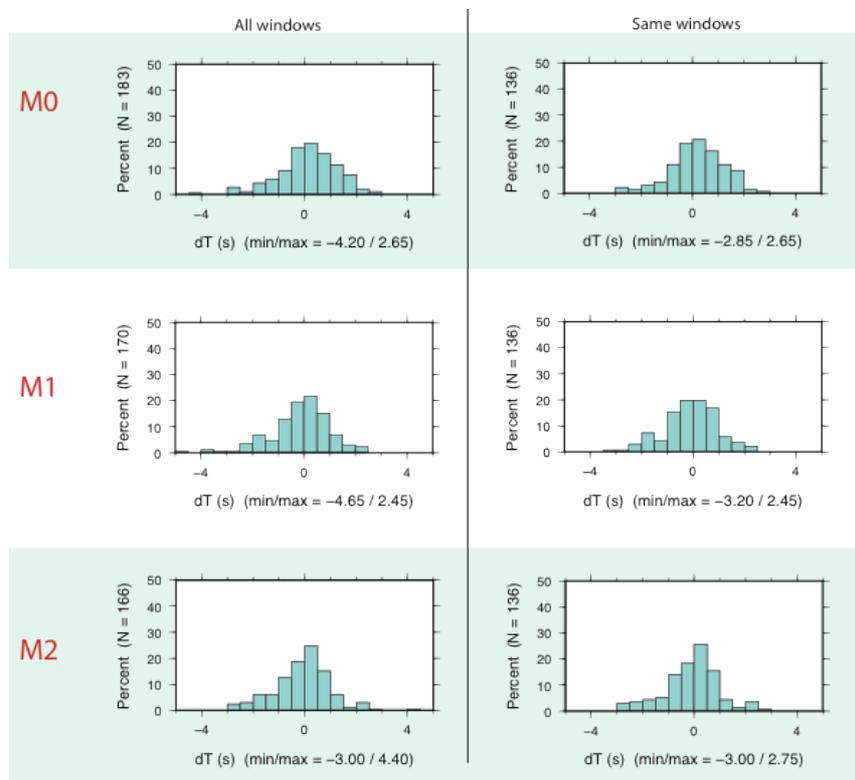
The misfit decrease is very slow (Figure 62), announcing the need for several more iterations to obtain a satisfying threshold.

For the following discussion, I look more in details in the first two updates, as the third one need more improvements.

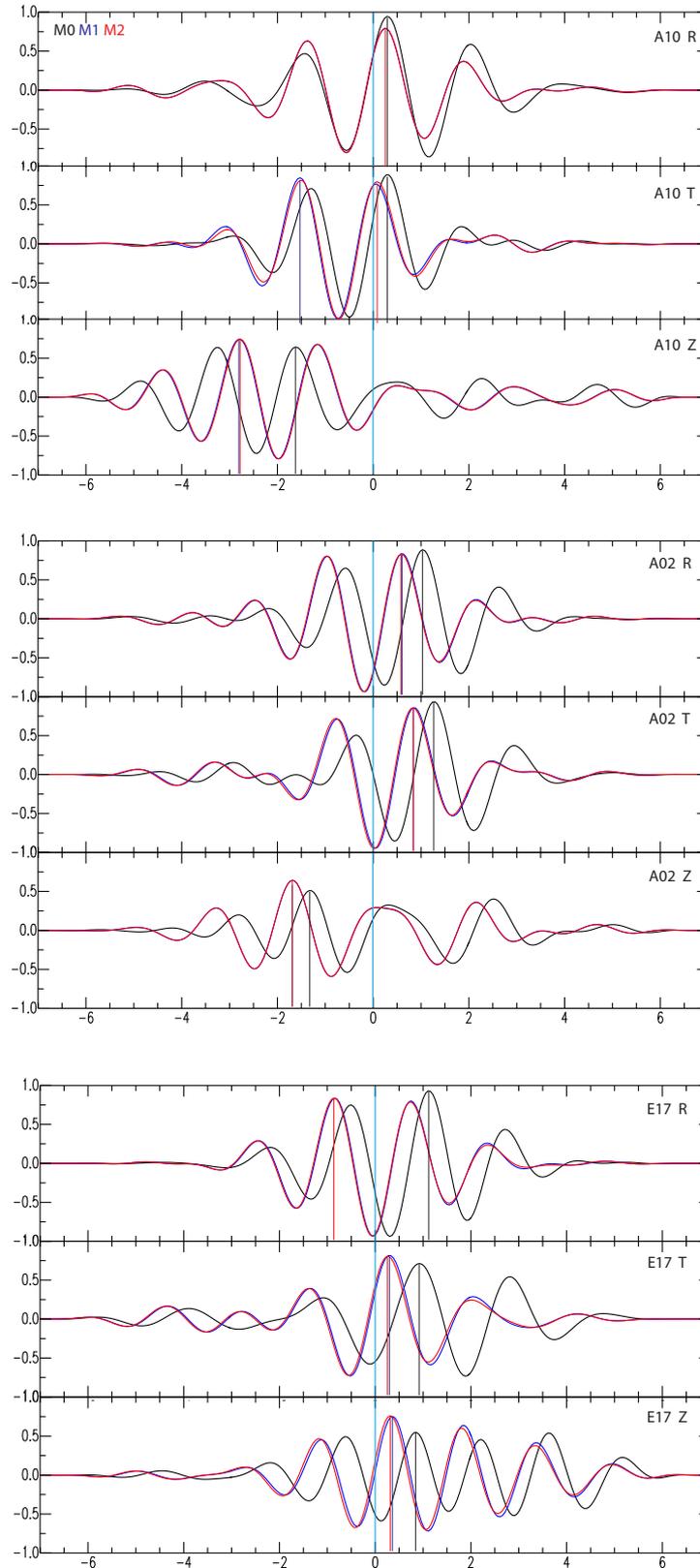
The cross-correlation measurements present a very light improvement from M0 to M2, either in an “all windows” case measurement, either in a “same windows” case (respectively left and right columns in Figure 63). However, one can note that considering measures in the “all windows” case, the number of selected windows

to compute the cross-correlation values decreases. This might indicate a worse fit over all the real and synthetic signals.

At such frequency, here between 0,4 and 0,8Hz, some instability behaviour can happen. This can be seen on cross-correlated seismograms. In Figure 64, I show the comparison between the data and synthetic cross-correlated seismograms at three stations for three models: the reference one and the first two updates. Inversion instability can be suspected for the A10T and E17R signals, where the maximum cross-correlation measurement passes from the positive side to the negative one between different models. Such an instable behaviour would lead to the incapacity of the algorithm to approach efficiently a minimum misfit.



**Figure 63 - Cross-correlation measurements from M0 to M2. Two measurement cases are compared, as explained in the text: one considering all the selected measurement windows for each model (left), the other considering exactly the same windows for each model (right). A small improvement of the Gaussian shape can be seen in both cases. However, in the first case, a decrease of the measurable windows has to be noted.**



**Figure 64 - Cross-correlation examples for first two model updates at three stations. The maximum cross-correlation pic indicates the travel-time cross-correlation measurement between real and synthetic signals with respect to the zero. Placed at zero, this would indicate a perfect fit. Instability can be suspected for A10T and E17R records, as the measure passes from on side to another with respect to zero (changes in the variations sign).**

#### IV.III.3.1.1. Models

Our tomographic inversion starts from an initial reference model based on the 3D Vp, Vs and density model of Tondi et al. (2009). This model, defined in a 25-28 latitude degrees and 44,36-46,43 longitude degrees area for a 230km depth, contained details features of the Vrancea lithosphere structure.

The model variations from the reference model to the second update are discreet. Detailed slice views of the Vp and Vs model updates can be found in Figure 69 in Appendix B.

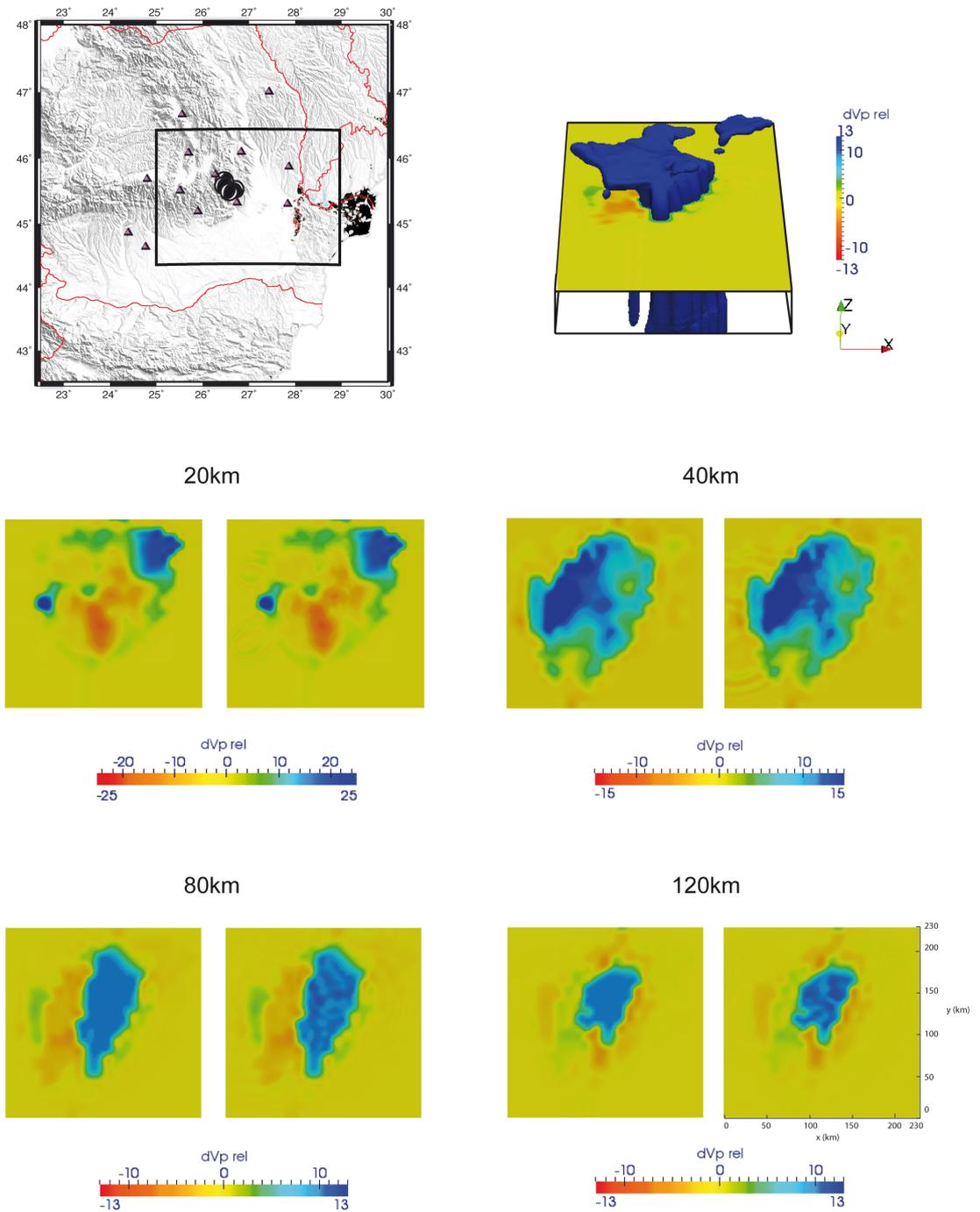
Figure 65 (for Vp) and Figure 66 (for Vs) show the relative velocity variations with respect to the mean velocity at different depths (20, 40, 80 and 120km), for the reference model and the second model update. The velocities corrections from reference model to second model update do not benefit from a large homogeneous extension but instead follow the kernel fringes as a consequence of the high-frequency range inversion. More iterations might be necessary to achieve bigger variations that lead to an appreciable decrease of the misfit measurements.

Another consequence of this high-frequency inversion stays in the fact that most of the variations focus in the station neighbours, as no efficient smoothing was implementable at such computational costs.

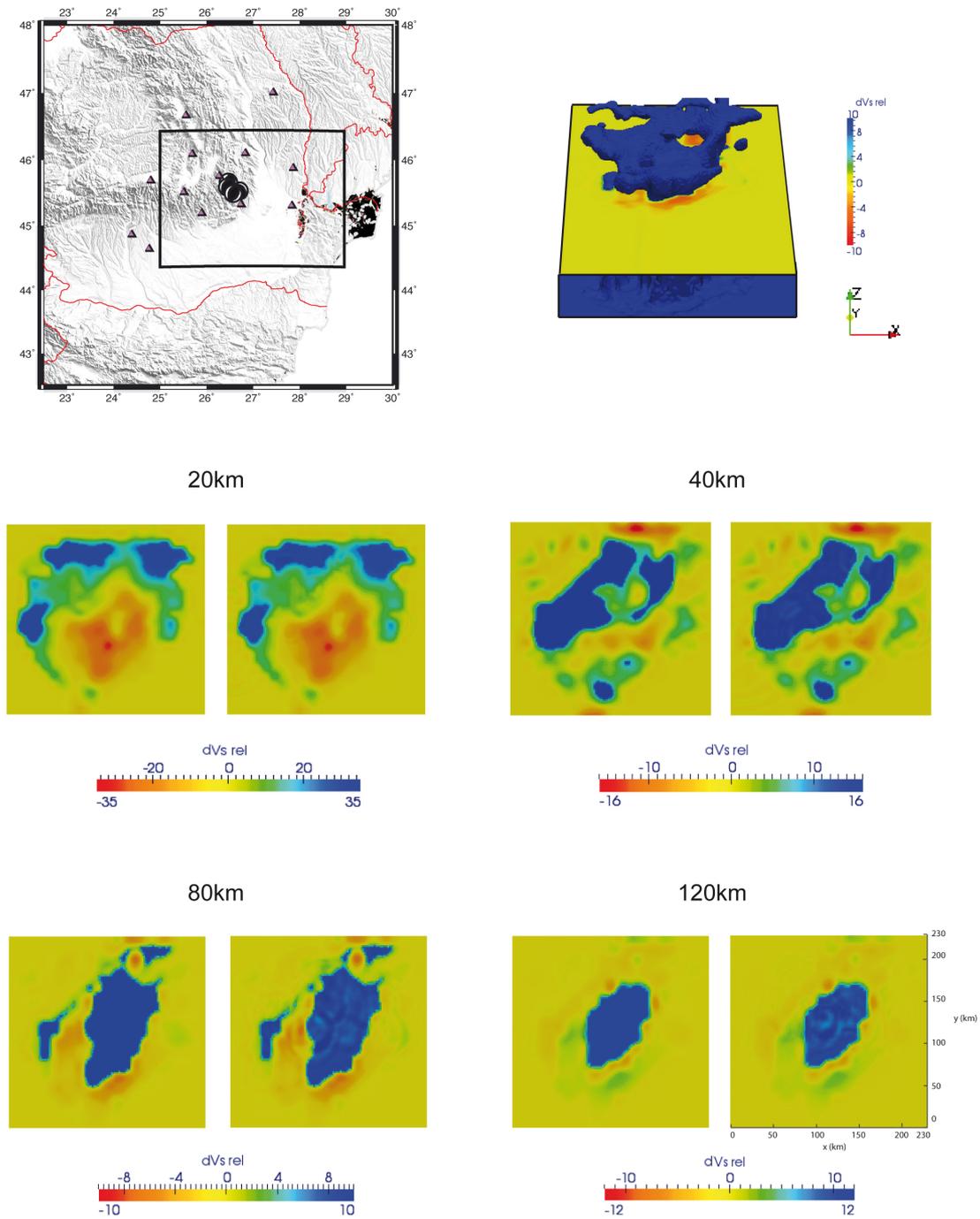
These preliminary inversion results show two main geophysical distributions. First, a low velocity anomaly in the upper part of the model, at shallow depth until 35km, that presents an arc shape. This is visible in the detailed slices depth of the model in Appendix B as well as in Figure 65 and Figure 66 at the 20km depth slice. The low anomaly is the red ring-like feature at the center of the model. The second geophysical distribution is a well-defined downgoing slab shape high velocity anomaly, composed of a N-S horizontal anomaly in the depths between 40 and 70km, linked to a nearly vertical NE-SW anomaly from 70 to 180km (see the blue anomaly in these depths and in the 3D representation of the model in Figure 65 and Figure 66, as well as in Appendix B for more details). The sign's change from the low velocity anomaly in the shallower depths to the high velocity anomaly stays around 40km depth (Figure 65, Figure 66 and Appendix B). Moreover, in the depths range characteristic of the Vrancea seismic gap, between 40 and 70km, a low velocity anomaly lies beside the high anomaly. Tondi et al. (2009) have interpreted his particular feature as partial melting of the upper mantle due to a delamination process.

The features observed in these first updated models, remain consistent with the initial model used for the inversion, as well as with the results obtained from other tomographic studies (Martin et al., 2006; Koulakov et al., 2010; Ren et al., 2012).

The potential of the full waveform inversion and finite frequency method is not yet perceptible at this stage of the inversion with a reduced amount of seismic paths. However this inversion framework is not ended. Pursuing the iterations and adding later more events to the inversion will help to understand whether such a high frequency range inversion is stable and furnishes reliable and valuable outcomes.



**Figure 65 -  $V_p$  model relative variations at different depths for reference model (left side) and second model update (right side). 2D velocity maps represent the relative  $V_p$  variations with respect to the mean  $V_p$  reference velocity in each slice at depth 20,40,80 and 120km. The 3D representation gives the  $V_p$  20% relative variations with respect to the mean reference velocity in all the model, together with the 80km slice. Scales are given in percent.**



**Figure 66 -  $V_s$  model relative variations at different depths for reference model (left side) and second model update (right side). 2D velocity maps represent the relative  $V_s$  variations with respect to the mean  $V_s$  reference velocity in each slice at depth 20,40,80 and 120km. The 3D representation gives the  $V_s$  20% relative variations with respect to the mean reference velocity in all the model, together with the 80km slice. Scales are given in percent.**

### III.3.2. Discussion

The recipe for an efficient tomographic inversion is made of several ingredients: a good quality data, a good spatial coverage of the area, robust tools and method to solve for the seismic waves equation and compute synthetic seismograms as well as to compute the misfit and its gradient and finally computational resources able to sustain the whole minimization process.

If nowadays it is easier to deal with powerful methods that support the computational part of a tomographic inversion, the latter cannot fill the lack of the primary and essential material needed for such a study: the data.

The data, in terms of spatial coverage, number and quality of observed seismograms, spread its own qualities or defaults among the whole inversion process and influence its efficiency.

It has been illustrated with these first results of full waveform tomographic inversion for the Vrancea region using the CALIXTO99 temporary experiment dataset. The results show the complexities in implementing such a full waveform inversion at a high frequency range, in a configuration where the amount of data is limited. These difficulties are characterised by 1) a modest spatial coverage of the Vrancea region, related to the limited amount of eligible data for inversion and the proper characteristic of the region seismicity with events clustered in a small depth volume; 2) a high and short frequency range in which to carry out the inversion that is another consequence of the local seismicity and the data quality; 3) a high computational cost that depends directly from the previous statement and does not allow to fasten the inversion; 4) a slow misfit convergence.

The slowness of the convergence has to be seen as the final consequence of the enumerated points.

Tomographic inversion is based on minimization algorithms, which one key of success remains in the capacity to find the right gradient direction that points to a minimum. In my case the gradient is built from full waveform seismic seismograms in a high frequency range (0,4-0,8Hz) imposed by the quality of the observable data. At such frequency, confronting synthetic and real signals becomes a hard task and thus the gradient computation can be unstable, slowing the minimization process.

Another point that could influence the efficiency of the minimization is the smoothing procedure. Indeed, as underlined previously, due to too high computational costs, I have not been able to implement a smoothing operation larger than 2km. Thus, high sensitivity kernel values might remain in the model and this has a strong influence in model updating, as the model correction is based on the maximum sensitivity kernel values found in the model.

Even if these results of full waveform inversion are quite modest in this particular context, they also show the great power of such an inversion method as, as we have seen, even though working at high frequency it is still possible to retrieve information from the seismic signal.

To carry out such kind of inversion, it is of importance to start with a relatively well-defined tomographic model that solves the large-scale heterogeneities and to

work with a satisfying amount of data that allow a good spatial coverage of the area. Finally the computational resource is of primary importance for such an inversion. With the fast improvements in this field, it might be less hard to carry out this kind of study in the near future.

# CONCLUSIONS

This study was aimed to evaluate the application of a full waveform inversion (FWI) tomography at regional scale for the Vrancea area in Romania using data from the 1999 six months temporary local network CALIXTO99. Starting from a detailed 3D  $V_p$ ,  $V_s$  and density model, built on classical travel-time tomography together with gravity data, my goal was to evaluate the improvements that could derive from using such a full waveform approach in the frame of an optimization algorithm.

Full waveform inversion implies working with full numerical solution of the wave equation and the finite-frequency theory that takes into account the frequency dependency of the travel-time and amplitude of seismic waves. In principle, this allows the use of more information from seismic waveform and the better illumination of the seismic path, leading to more detailed seismic tomographic Earth's models. However these advantages come at the cost of very large computations, necessary for example for the computation of synthetic seismic waves or for the gradient of the misfit through the adjoint method. I have applied the method using the SPECFEM3D software that supports the computation of synthetic waveforms as well as sensitivity kernels for model updating. Computations have been done on a Blue Gene Q (BG/Q) architecture at the CINECA supercomputing centre in Bologna, Italy, for a total computational cost of about 5Millions CPU hours. This represents a high computational cost. The computational facilities used for this study have been reached on the basis of national calls for proposal.

Several complexities came out from this study. Main part takes roots from the data.

The particularity of the Vrancea seismic properties, with intermediate magnitude events concentrated in a nearly vertical volume, makes any regional tomography study using local data constrained by the receivers' coverage. A dense broad network of receivers is thus of primary importance to unveil the lithospheric structure of Vrancea.

In my case, the CALIXTO99 Broad-Band data, whose stations gathered in a small area above the seismic nest, present a sporadic coverage that could be otherwise compensated by the region enlightening offered by the finite frequency theory with respect to classical ray tomography.

The low quality of the data, in terms of archiving and noise level, is to underline as it has reduced a lot the whole invertible tomographic dataset.

Working with relatively old data (1999) from a temporary experiment made it difficult to collect the entire original dataset consisting in seismic waveforms, instrumental responses and experiment archives. As a consequence, several seismic records have been thrown out as unreadable or excluded from the dataset for a lack of available instrumental responses. This appears as a relevant drawback of temporary experiments, which database tend to be sparse or lost with time passing, when they are not entrusted rapidly to public seismological data archive.

The usable database has been screened to select only the seismic signal with best signal-to-noise ratio. This cut a lot the number of seismic waveforms as

## Conclusions

main part of the local events recorded belongs to small magnitudes range at intermediate depth and are hidden by noise level.

As a consequence of the noise level and the distribution of the Vrancea seismic sources, the selected waveforms were finally studiable only in a high and narrow frequency range between 0,4 and 0,8Hz. These are high values for full waveform inversion application and introduced several computational challenges.

Any tomographic inversion presents a strong dependency on seismic sources physics in terms of locations and fault plane mechanism (moment tensor). It is thus better to start the inversion with a dataset of seismic source defined with respect to the reference tomographic model, and to consider a re-evaluation of the seismic sources during the inversion at some stage of/ or parallelly to model updating. For the data of Vrancea, I benefited from several information given by the CALIXTO99 archive as well as from the previous study of Tondi et al. (2009) that presents a joint ray-based tomographic/gravimetric inversion for Vrancea using the CALIXTO99 short-period data together with gravity measurements. To remain coherent with my inversion process, I decided to correct and invert again for the seismic sources using the same protocol as in the Tondi et al. (2009) study: first by doing a relocation of the events with NLOC and in a second time by computing the fault plane mechanisms with FPFIT. The uncertainty above these estimations, from event depths to fault plane mechanisms, is to point out and is due to the low azimuthal coverage offered by the cone-like event-stations geometry. Proceeding with the study of the overall cross-correlation travel time measurements between synthetic waveforms and real data, it appeared necessary to adjust the depth and reference time of the events to start the inversion with zero-centred Gaussian like measurements distribution. The latter is required to start the non-linear inversion algorithm in the vicinity of a minimum and to facilitate its convergence.

Implementing a full waveform inversion in a high frequency range might be a problem to recover a comprehensive model updating, as the low frequency discontinuities - that correspond to large-scale heterogeneities - might not been well solved. It can be though in such cases to solve first for the large-scale heterogeneities by incrementing the dataset with low-frequency waveforms. For example in my case, it could have been done with some recent data from the permanent Romanian network. However, as I am using a 3D reference tomographic model -from Tondi et al. (2009)- that already solved for the low frequency discontinuities this limitation might not come out. Instead, the goal was to add some details to the model in terms of small-scale heterogeneities.

The synthetic waveform computation has been done with SPEC3D in parallel through a computational grid of spectral elements reflecting the geophysical properties of the Earth as well as the frequency content to be reached. This grid consists in hexahedral elements of Gauss-Lobatto Legendre points, whose size increases with depth to keep the number of point per wavelength stable<sup>14</sup>.

To reach the high frequency range of [0,4 – 0,8]Hz, a grid of small spectral elements has to be built implying heavy computational costs in terms of effective computational time and number of processors used to manage the computation (4096 processors). The 14 millions spectral elements grid built for Vrancea on a

---

<sup>14</sup> There is a general increase of PREM seismic velocities with depth. The size increase is done in SPEC3D by defining doubling layers at given depths of seismic velocity discontinuity.

## Conclusions

BG/Q architecture required around 15.000 CPU hours (3h effective time \* 4096 processors) to solve the seismic waves equation for each seismic source, and more than the double for an event-kernel computation using the adjoint method. The heaviest operation implemented in this inversion appeared to be the 3D Gaussian smoothing of the misfit kernel that required nearly 50.000 CPU hours (12h effective time \* 4096 processors) for a 2km smoothing.

Working on a supercomputing centre, these computations are not managed instantaneously as on a dedicated computer. Instead, queue rules are applied to deal with all the users jobs management, implying latency time, between job submission and its effective run, to be added to the effective computational time.

This heavy computational mesh introduced in consequence heavy outputs in terms of storage capacity. Thus a particular care had to be taken to manage the amount of outputs from all the simulations. My study implied nearly 30TB of file storage for the inversion presented. This weight also complicates the visualization step, as common visualization softwares (e.g. Paraview) struggle to manage such heavy files (4Gb).

Difficulties in managing the computations schedule time together with the deadline of the thesis, imposed me to reduce the size of the inversion, for the time being, considering 5 events above the 19 intended. The complete inversion with the 19 events is in process and will be completed in the future. Using a representative sample of the whole dataset can be used at some stage of the inversion, for example when looking for the best step length for model update. Applied for the whole inversion process this could be reductive and deprive of crucial information about the velocity model. I can however believe that, working with very similar events concentrated in the same small area and producing similar seismic ray paths, the reduction of seismic sources number involved in the inversion may not mean a reduction of the information retrieved. This is a guess that can be checked later, confronting the results of this study with the 19 events inversion.

I have presented here the first three updates of the Vrancea full-waveform tomographic inversion. The latter is based on travel-time cross-correlation misfit measurements. Although the frequency range for the inversion is quite high and narrow, the resulting misfit kernel- representing the area of the model to be corrected- is appreciable, considering the particular geometry of the station-event ray coverage. The amount of variation is however conditioned by the smoothing parameterization. In my case, the computational impossibility to carry out large smoothing leads to very small variations in delimited areas that follow mainly the kernel fringes. The misfit decrease is thus slow and much more iterations might be necessary to obtain a satisfying change. This decrease might be also slowed down by possible inversion instabilities coming from the difficulty to measure the goodness of fit of the synthetic and real data signals. The addition of other events seismic records in the inversion frame may bring other valuable information and help to stabilize the inversion process, favouring the misfit convergence.

At this stage of the iteration, my tomographic model presents several recognized features. The nearly vertical NE-SW high velocity anomaly, defined in several other studies, is clearly visible in the depth range from 70 to 180km. This anomaly is linked to a N-S horizontal high velocity anomaly in the shallower depths between 40 and 70km, forming a downgoing slab shape. The upper part of

## Conclusions

the model instead is characterised by a low velocity anomaly until 40km depth. The inversion of the anomaly, passing from low to high velocity values, occurs in the depths around 40km. Between 40 and 70km, in the depth interval characteristic of the Vrancea seismic gap, the model presents a low velocity anomaly western of the high anomaly. This feature has been interpreted in the study of Tondi et al. (2009) as a zone of partial melting of the upper mantle.

The features observed in these first updated models, remain consistent with the initial model used for the inversion, as well as with the results obtained from other tomographic studies (Martin et al., 2006; Koulakov et al., 2010; Ren et al., 2012).

The preliminary results of the full waveform inversion of the Vrancea region have a relatively limited extent and show the limits of application of such a highly detailed inversion method. Today, even with one of the best available computational infrastructure FERMI<sup>15</sup>, this type of inversion method cannot be applied efficiently to studies involving a high frequency range resolution. This would imply a higher computational power and capacity storage.

Although they can in principle use more information contained in seismic waveforms, more elaborate computational approaches, such as the full waveform inversion, cannot make up for limitations in the number of seismograms or their quality. These kinds of approach are actually more demanding on data quality. Citing Fichtner (2011) from his book about full waveform modelling and inversion: “While FWI is still a comparatively young method, there is already an important conclusion to be drawn: its application is highly problem dependent”. In a nutshell, the key of success for such an inversion remains in the data: the data availability, distribution and quality.

As evocated previously, a seismic tomography inversion is more consistent if one solve both for the Earth structure and the seismic source parameters. As this joint inversion involves linear and non-linear parameters, it is common to treat each inverse problem sequentially. The Variable Projection Method I have tested in a simple application appears to be an efficient tool to solve least squares minimization problems involving linear and non-linear parameters. Absent in the tomographic field, it could be seen as an alternative to carry out such joint inversion. In its simplest form it implies to use the waveform misfit and thus might not be applicable for some case of noisy seismic signals. Nevertheless, it can have a huge range of applications and could be further developed to implement other misfit functions.

---

<sup>15</sup> The Fermi system at CINECA has been ranked as the 15<sup>th</sup> most powerful system worldwide on the 42<sup>st</sup> TOP500 list released in November 2013. (<http://www.top500.org/list/2013/11/>)

# APPENDIX A

## TABLES

Appendix A

Table 1 - Instrumental characteristics of the CALIXTO Broadband seismic stations.

STA	lat	lon	high (m)	Sensor		natural period (s)	sensitivity (V/m/s)	Data logger
				manufacturer	model			model
A02	47,01	27,43	414	Guralp	G-3T	120	1500	RT130
A05	46,49	27,00	187	Guralp	3ESP	30	2000	
A10	45,76	26,27	1026	Guralp	G-3T	120	1500	RT130
A12	45,19	25,89	446,66	Streickensen	STS2	120	1500	RT130
B04B	45,69	26,08	675,5	Streickensen	STS2	120	1500	6NT
C03	46,67	25,55	716,83	Guralp	G-3T	120	1500	RT130
C08A	45,69	24,79	687,33	Guralp	G-3T	120	1500	RT130
D04	44,86	24,39	399,66	Guralp	G-3T	120	1500	RT130
D06	44,64	24,76	208,5	Guralp	G-40T	30	800	RT130
D15	44,20	25,50	122,5	Streickensen	STS2	120	1500	RT130
D15A	44,20	25,50	122,5	Guralp	3ESP	30	2000	
E03	46,10	26,83	370,33	Guralp	G-40T	30	800	RT130
E08	45,87	26,73	408	Guralp	3ESP	30	2000	
E09	45,78	26,72	419,11	Streickensen	STS2	120	1500	6NT
E17	45,51	25,51	1036,3	Guralp	G-3T	120	1500	RT130
E21	45,49	25,95	1361	Guralp	3ESP	30	2000	
E25	45,33	26,74	303,2	Guralp	G-3T	120	1500	RT130
F01	46,77	26,39	242	Guralp	3ESP	30	2000	
F02	46,51	26,65	88	Guralp	3ESP	30	2000	
F09	45,88	27,86	222,16	Streickensen	STS2	120	1500	RT130
F11	45,31	27,83	60	Guralp	G-40T	30	800	RT130
L01	45,82	26,38	1759	Lennartz	L3D/20	20	1000	3XT
L02	45,77	26,51	1103,2	Lennartz	L3D/20	20	1000	3XT
L03	45,66	26,58	369	Lennartz	L3D/20	20	1000	3XT
S07	46,09	25,69	553	Guralp	G-3T	120	1500	RT130
Z05	44,68	26,25	166	Guralp	3ESP	30	2000	
Z08	44,74	27,29	66	Guralp	3ESP	30	2000	

Appendix A

Table 2 - Source parameters from the CALIXTO99 archive for the 61 events, whose moment tensor have been computed in the Tondi et al. (2009) study

date	time	lat	lon	depth(km)	Mw
990520	19:55:33.00	45,450	26,260	128	3,6
990525	09:35:53.00	45,550	26,540	129	4,5
990606	12:01:31.00	45,560	26,350	139	3,7
990611	08:20:04.00	45,743	26,667	107	3,8
990612	07:08:12.00	45,740	26,760	100	3,1
990620	00:09:07.00	45,540	26,600	129	4,3
990622	08:02:06.00	45,650	26,520	157	4,4
990622	20:10:42.00	45,640	26,590	141	3,4
990626	17:26:25.00	44,650	26,030	17	2,2
990629	20:04:07.00	45,550	26,630	126	4,3
990712	23:47:52.00	45,780	26,780	75	3,2
990713	10:09:25.00	45,660	26,810	125	3,2
990715	02:52:03.00	45,790	26,870	110	3,4
990715	07:36:23.00	45,580	26,460	137	3,8
990722	10:07:50.00	45,440	26,300	133	3,8
990731	23:42:31.00	45,630	26,670	40	2,2
990801	03:53:51.00	45,740	26,560	141	3
990801	05:56:50.00	45,530	26,600	124	3,6
990802	15:16:22.00	45,300	25,400	0	2,2
990803	07:18:39.00	45,460	26,160	33	2,3
990803	08:07:05.00	45,650	26,490	148	3,4
990806	11:00:04.00	45,770	26,850	103	2,8
990806	11:55:14.00	45,750	26,710	104	3,5
990807	02:25:49.00	45,610	26,490	128	4,1
990809	07:16:16.00	45,730	26,700	133	3,9
990809	23:24:30.00	44,950	27,060	15	3,3
990810	06:32:14.00	45,690	26,790	107	2,7
990810	06:35:06.00	45,690	26,780	93	2,6
990810	06:36:55.00	45,710	26,820	102	2,4
990813	05:21:13.00	45,790	26,760	85	2,5
990823	20:57:40.00	45,570	27,180	15	2,6
990824	16:20:07.00	45,780	26,860	118	2,9
990826	03:47:16.00	45,790	26,840	129	3,4
990826	13:06:51.00	44,890	27,140	13	2,8
990827	21:53:19.00	45,660	26,530	150	3,5
990828	05:32:28.00	45,700	26,630	142	3,6
990828	09:42:59.00	45,470	26,880	0	2
990830	00:52:06.00	46,000	27,200	35	2,8
990830	13:53:38.00	45,540	26,590	119	3,1
990831	02:11:12.00	45,570	26,510	133	3,6
990901	23:30:28.00	45,770	26,840	126	3
990910	03:12:30.00	45,750	26,590	116	2,7

Appendix A

990910	14:24:08.00	45,480	26,450	102	2,6
990914	02:59:09.00	45,790	26,910	91	2,3
990914	14:12:54.00	45,730	26,660	74	2,4
990914	23:48:10.00	45,530	26,580	119	4
990915	06:01:09.00	45,750	26,800	108	2,4
990916	19:00:19.00	45,820	26,820	85	2,9
990919	21:32:46.00	45,400	26,410	107	2,7
990921	11:29:55.00	45,300	25,340	0	2,7
990922	05:13:00.00	45,700	26,710	83	2,4
990927	19:16:51.00	45,500	25,890	5	1,9
990930	00:55:57.00	44,880	25,740	10	2,6
991003	00:25:05.00	45,780	26,800	136	3,1
991011	03:16:02.00	45,620	26,630	79	2,3
991012	19:23:39.00	45,700	26,610	83	3,4
991012	23:48:33.00	45,650	26,430	154	4,3
991016	09:33:18.00	45,520	26,450	131	3,5
991017	12:12:47.00	45,480	26,880	0	2
991018	09:33:30.00	45,300	25,400	0	2,6
991028	11:01:59.00	45,360	25,370	0	2,9

Appendix A

Table 3 - Source parameters from NLLOC for the 61 events from which Moment Tensor have been computed in the Tondi et al. 2009 study.

date	time	lat	lon	depth (km)
990520	19:54:48.79	45,462	26,262	221,63
990525	09:34:57.59	45,548	26,501	148,5
990606	12:00:57.44	45,583	26,282	157,15
990611	08:19:53.78	45,747	26,570	162,96
990612	07:07:58.78	45,796	26,586	96,41
990620	00:09:00.15	45,571	26,517	124,27
990622	08:01:59.97	45,667	26,448	154,22
990622	20:09:57.69	45,689	26,487	163,27
990626	17:25:57.29	44,589	26,061	3,61
990629	20:04:00.25	45,603	26,484	129,47
990712	23:46:59.16	45,814	26,765	74,85
990713	13:10:00.14	45,669	26,427	135,92
990715	02:51:58.59	45,795	26,857	122,33
990715	07:35:59.19	45,574	26,379	140,81
990722	10:06:59.26	45,495	26,244	142
990731	23:41:59.18	45,612	26,639	44,76
990801	03:53:05.73	45,737	26,512	110,93
990801	05:56:00.6	45,585	26,506	111,59
990802	15:15:58.71	45,291	25,380	2,52
990803	07:17:57.37	45,381	26,147	51,7
990803	08:06:58.57	45,697	26,414	148,94
990806	10:59:58.00	45,777	26,779	108,74
990806	11:55:01.74	45,774	26,496	78,58
990807	02:25:00.63	45,610	26,411	122,45
990809	07:15:59.32	45,722	26,632	132,45
990809	23:24:00.23	44,930	27,021	18,11
990810	06:31:46.15	46,000	27,414	224,28
990810	06:34:58.22	45,668	26,669	102,67
990810	06:35:57.05	45,705	26,757	128,75
990813	05:21:03.60	45,782	26,725	59,07
990823	20:57:00.55	45,579	27,145	13,79
990824	16:19:58.32	45,793	26,804	136,51
990826	03:47:02.21	45,744	26,656	110,69
990826	13:05:59.63	44,853	27,123	5,36
990827	21:53:03.98	45,700	26,427	114,54
990828	05:32:03.07	45,739	26,547	114,77
990828	09:41:59.71	45,456	26,890	4,93
990830	00:52:00.34	45,987	27,171	34,46
990830	13:52:59.76	45,511	26,577	122,66
990831	02:11:01.72	45,630	26,383	124,94
990901	23:29:54.42	45,820	26,785	184,82
990910	03:11:57.49	45,808	26,561	129,14

*Appendix A*

990910	14:23:58.42	45,483	26,419	111,41
990914	02:58:54.34	45,868	26,929	139,01
990914	14:11:58.66	45,716	26,640	82,68
990914	23:47:58.94	45,542	26,538	126,49
990915	06:00:56.32	45,762	26,743	143,68
990916	18:59:58.38	45,866	26,795	97,16
990919	21:31:55.55	45,425	26,252	146,96
990921	11:29:00.42	45,309	25,430	3,78
990922	05:12:57.50	45,697	26,665	99,29
990927	19:15:59.14	45,476	25,891	7,01
990930	00:54:58.80	44,814	25,666	15,09
991003	00:24:58.39	45,833	26,638	145,88
991011	03:15:55.94	45,636	26,615	105,04
991012	19:22:55.31	45,851	26,563	115,26
991012	23:47:58.01	45,653	26,418	162,81
991016	09:32:58.11	45,565	26,383	146,77
991017	12:11:58.97	45,466	26,876	16,41
991018	09:32:58.00	45,277	25,380	10,79
991028	11:01:00.10	45,476	25,584	114,4

Appendix A

Table 4 - Source parameters from NLLOC for the 61 events from which Moment Tensor have been computed in the Tondi et al. 2009 study after the corrections (velocity model and reference time). The fill lines represent the events that we will finally use for the inversion.

date	time	lat	lon	depth (km)
990520	19:55:34.14	45,285	26,277	123,77
990525	09:35:54.91	45,441	26,552	123,2
990606	12:01:32.26	45,490	26,210	137,86
990611	08:20:05.10	45,658	26,633	109,28
990612	07:08:13.11	45,768	26,566	86,48
990620	00:09:08.72	45,532	26,704	122,87
990622	08:02:07.59	45,641	26,475	157,49
990622	20:10:43.31	45,662	26,535	147,71
990626	17:26:23.18	44,621	26,093	18,41
990629	20:04:08.57	45,564	26,575	133,49
990712	23:47:53.63	45,757	26,787	56,96
990713	13:10:59.05	45,664	26,467	135,7
990715	02:52:04.29	45,746	26,946	107,29
990715	07:36:24.14	45,522	26,367	139,46
990722	10:07:51.71	45,447	26,144	131,4
990731	23:42:31.14	45,581	26,627	32,25
990801	03:53:53.42	45,876	26,544	158,55
990801	05:56:51.90	45,561	26,509	115,09
990802	15:16:20.84	45,299	25,402	3,77
990803	07:18:37.86	45,411	26,167	31,76
990803	08:07:06.23	45,705	26,335	139,18
990806	11:00:05.26	45,668	26,829	89,18
990806	11:55:15.28	45,759	26,591	95,4
990807	02:25:50.53	45,589	26,450	130,81
990809	07:16:17.86	45,677	26,744	124,1
990809	23:24:30.18	44,933	26,997	15,07
990810	06:32:15.38	45,658	27,098	106,95
990810	06:35:07.60	45,607	26,778	79,83
990810	06:36:55.81	45,627	26,849	102,22
990813	05:21:16.80	45,845	26,836	75,83
990823	20:57:41.70	45,452	27,155	42,22
990824	16:20:08.74	45,742	26,954	113,08
990826	03:47:17.58	45,731	26,855	129,22
990826	13:06:51.54	44,887	27,019	6,76
990827	21:53:20.63	45,634	26,460	155,53
990828	05:32:30.63	45,715	26,561	138,01
990828	09:42:58.26	45,443	26,868	8,23
990830	00:52:07.59	45,997	27,158	29,48
990830	13:53:39.55	45,418	26,748	118,4
990831	02:11:13.95	45,566	26,438	138,83

Appendix A

990901	23:30:29.32	45,688	27,034	130,97
990910	03:12:30.84	45,725	26,671	111,2
990910	14:24:08.90	45,449	26,468	99,51
990914	02:59:10.46	45,617	26,920	84,85
990914	14:12:55.29	45,673	26,650	64,72
990914	23:48:11.42	45,491	26,573	117,91
990915	06:01:10.29	45,666	26,883	106,81
990916	19:00:20.24	45,820	26,836	78,56
990919	21:32:47.04	45,344	26,380	109,55
990921	11:29:55.10	45,303	25,477	1,19
990922	05:13:00.97	45,619	26,689	72,75
990927	19:16:48.60	45,498	25,737	47,88
990930	00:55:55.49	44,807	25,707	16,01
991003	00:25:07.09	45,767	26,837	126,46
991011	03:16:02.70	45,523	26,696	65,91
991012	19:23:39.81	45,703	26,572	75,94
991012	23:48:34.62	45,582	26,408	148,99
991016	09:33:19.95	45,475	26,449	128,33
991017	12:12:45.46	45,442	26,814	18,42
991018	09:33:27.99	45,277	25,397	17,48
991028	11:02:00.96	45,436	25,562	4,72

Appendix A

Table 5 - Availability of the dataless-files from Potsdam and erroneous station records ignored by msfix correction procedure (grey).

				<b>sensor</b>			<b>Potsdam Resp</b>
<b>STA</b>	<b>lat</b>	<b>lon</b>	elevation(m)	manufacturer	model	model	<b>available</b>
A02	47,01	27,43	414,00	Guralp	G-3T	RT130	<b>yes</b>
A05	46,49	27,00	187,00	Guralp	3ESP	CD24	<b>yes</b>
A10	45,76	26,27	1026,00	Guralp	G-3T	RT130	<b>yes</b>
A12	45,19	25,89	446,66	Streickensen	STS2	RT130	<b>yes</b>
<b>B04B</b>	45,69	26,08	675,50	Streickensen	STS2	6NT	
C03	46,67	25,55	716,83	Guralp	G-3T	RT130	<b>yes</b>
C08A	45,69	24,79	687,33	Guralp	G-3T	RT130	<b>yes</b>
D04	44,86	24,39	399,66	Guralp	G-3T	RT130	<b>yes</b>
D06	44,64	24,76	208,50	Guralp	G-40T	RT130	<b>yes</b>
D15A	44,20	25,50	122,50	Guralp	3ESP	CD24	<b>yes</b>
D15	44,20	25,50	122,50	Streickensen	STS2	RT130	<b>yes</b>
E03	46,10	26,83	370,33	Guralp	G-40T	RT130	<b>yes</b>
E08	45,87	26,73	408,00	Guralp	3ESP	CD24	<b>yes</b>
<b>E09</b>	45,78	26,72	419,11	Streickensen	STS2	6NT	
E17	45,51	25,51	1036,30	Guralp	G-3T	RT130	<b>yes</b>
E21	45,49	25,95	1361,00	Guralp	3ESP	CD24	<b>yes</b>
E25	45,33	26,74	303,20	Guralp	G-3T	RT130	<b>yes</b>
F01	46,77	26,39	242,00	Guralp	3ESP	CD24	<b>yes</b>
F02	46,51	26,65	88,00	Guralp	3ESP	CD24	<b>yes</b>
F09	45,88	27,86	222,16	Streickensen	STS2	RT130	<b>yes</b>
F11	45,31	27,83	60,00	Guralp	G-40T	RT130	<b>yes</b>
<b>L01</b>	45,82	26,38	1759,00	Lennartz	L3D/20	3XT	
<b>L02</b>	45,77	26,51	1103,20	Lennartz	L3D/20	3XT	
<b>L03</b>	45,66	26,58	369,00	Lennartz	L3D/20	3XT	
S07	46,09	25,69	553,00	Guralp	G-3T	RT130	<b>yes</b>
Z05	44,68	26,25	166,00	Guralp	3ESP	CD24	<b>yes</b>
Z08	44,74	27,29	66,00	Guralp	3ESP	CD24	<b>yes</b>



# APPENDIX B

## **DETAILS ON VELOCITY MODELS**

## Appendix B

The 1D reference velocity model used for model correction and preliminary studies comes from the Tondi et al. (2009) tomographic study, where it has been used as initial reference model. It was obtained from a ray tracing optimization from the 1D model of Popa et al. (2001).

1D model			
Depth(km)	Vp(km/s)	Vs(km/s)	Rho(kg/m <sup>3</sup> )
-1	3,627	2,304	2,510
1	4,167	3,572	2,510
5	5,111	3,078	2,510
10	6,092	3,584	2,510
15	6,656	3,932	2,590
20	7,045	4,171	2,510
35	7,608	4,569	3,180
45	7,816	4,699	3,340
70	8,109	4,892	3,410
100	8,143	4,893	3,380
150	8,215	4,862	3,440
230	8,573	5,017	3,470

Figure 67 - Reference 1D velocity model.

From the 3D initial tomographic model I defined the two doubling layers depth for the SPECFEM3D mesh that correspond to geophysical discontinuities. In this case, these limits are at 40 and 80km depth, as it can be seen on Figure 68.

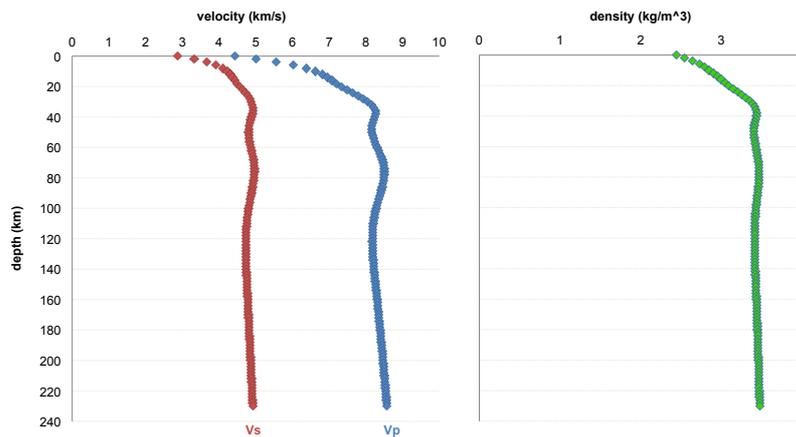


Figure 68 - Geophysical discontinuities in the reference 3D model.

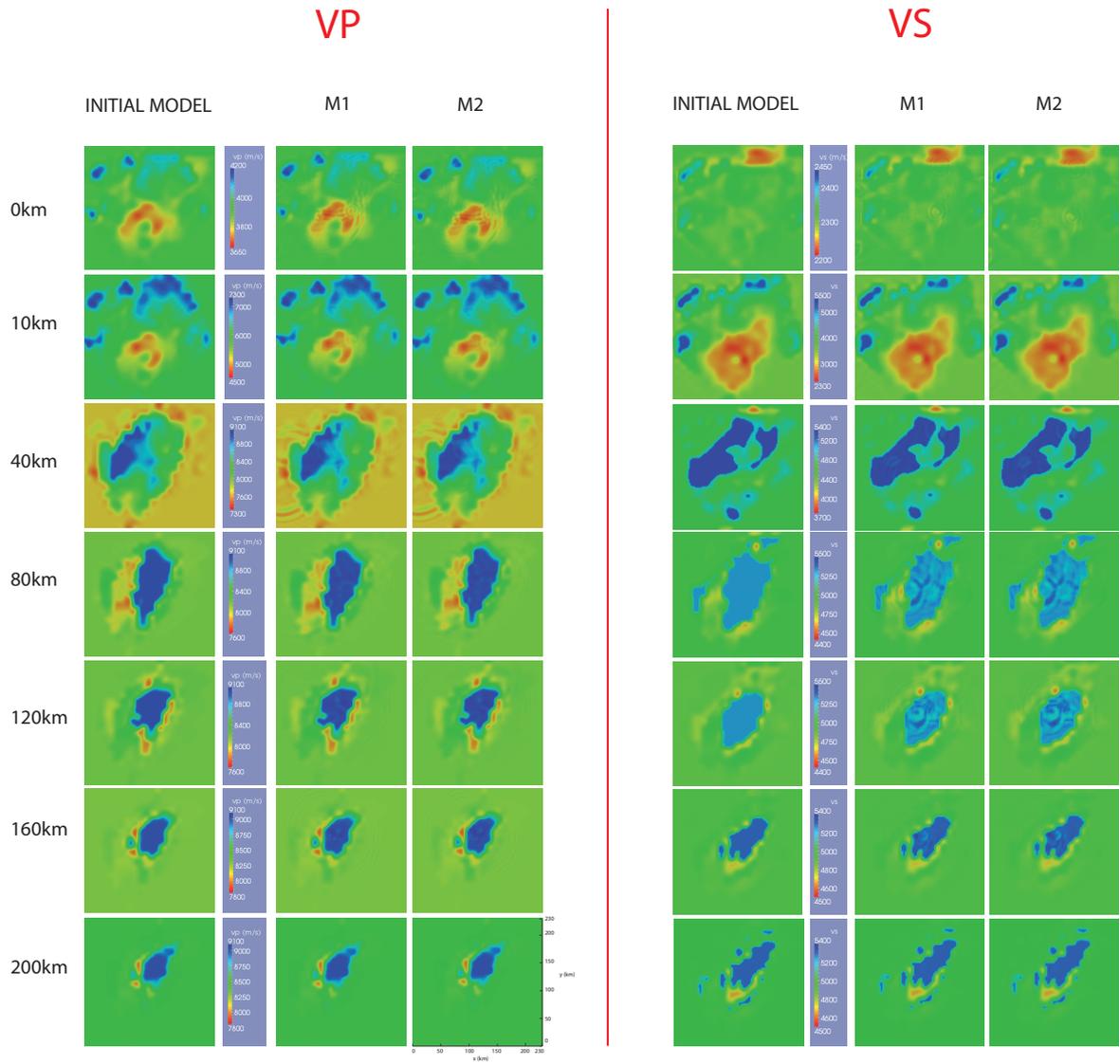


Figure 69 -  $V_p$  and  $V_s$  first two updates velocity models at slice depths.



# APPENDIX C

## **DETAILS ON EVENTS RELOCATION**

## Appendix C

The study of Tondi et al. (2009), taken as reference for my own study, is based on a tomographic inversion of the CALIXTO99 experiment, using local event recorded at the short period stations of the temporary network. During the experiment, the events source parameters have been estimated in a 1D reference model and reported in the archive of the experiment (Table 2 in Annexes). These parameters have been then revised in the 3D velocity model obtained from their inversion using the NLLOC algorithm (Lomax, 2000). The latter uses synthetic pick arrival times of seismic waves to estimate the reference time and location of an event. The results presented in Table 4 in Annexes have been then used to get the moment tensor parameters of each event using the FPFIT code.

The reference times given by this study are quite different from the times saved in the CALIXTO99 archive -built during the experiment with a 1D model-, as well as from the times indicated in a recent study of European arrival time picking catalogue done by Amaru et al. in 2008 (Table 6 and Figure 70).

It occurs that the NLLOC input P and S picking file was reporting erroneous picking time, giving a time automatically shifted from one minute to few seconds from the real picking time. As a consequence, I automatically corrected this error adjusting the picking times. With these new values, I computed again the event location (Table 4) and the moment tensor of the CALIXTO99 events in my reference 3D model. The resulting reference time, in comparison with the CALIXTO99 archive one is definitely better (Figure 18).

**Table 6 - Sample of the outlined reference time differences between the values given by the Amaru et al. (2008) catalogue, the CALIXTO99 experiment and the Tondi et al. (2009) study.**

	Amaru et al. GJI 2008	CALIXTO99	Tondi et al. 2009
Event ID	Reference time		
991012	23:48:32.88	23:48:33.00	23:47:58.01
lat	45,65	45,65	45,653
lon	26,451	26,43	26,418
depth (km)	149,6	154	162,81
991012	19:23:38.72	19:23:39.00	19:22:55.31
lat	45,715	45,7	45,851
lon	26,599	26,61	26,563
depth (km)	82,3	83	115,26
990525	09:35:54.37	09:35:53.00	09:34:57.59
lat	45,526	45,55	45,548
lon	26,451	26,54	26,501
depth (km)	123,8	129	148,5

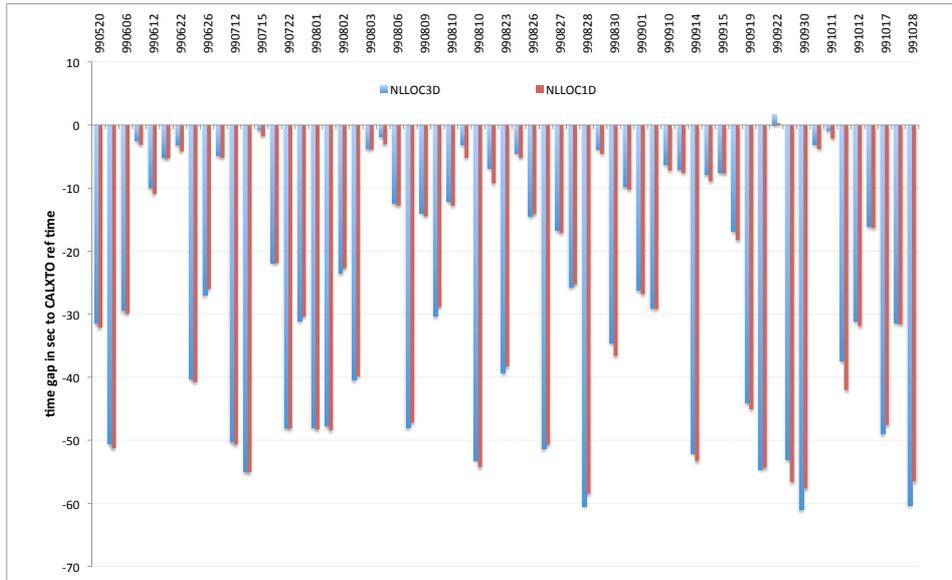


Figure 70 - Histograms representing the reference time shift between the NLLOC estimations computed with the original picking times from Tondi et al. (2009) and the CALIXTO99 archive times. The comparison is done for the computation through a reference 1D model and the Tondi et al. (2009) 3D model. Several events present a gap of about one minute from the CALIXTO99 reference time, which cannot be explained by the model.

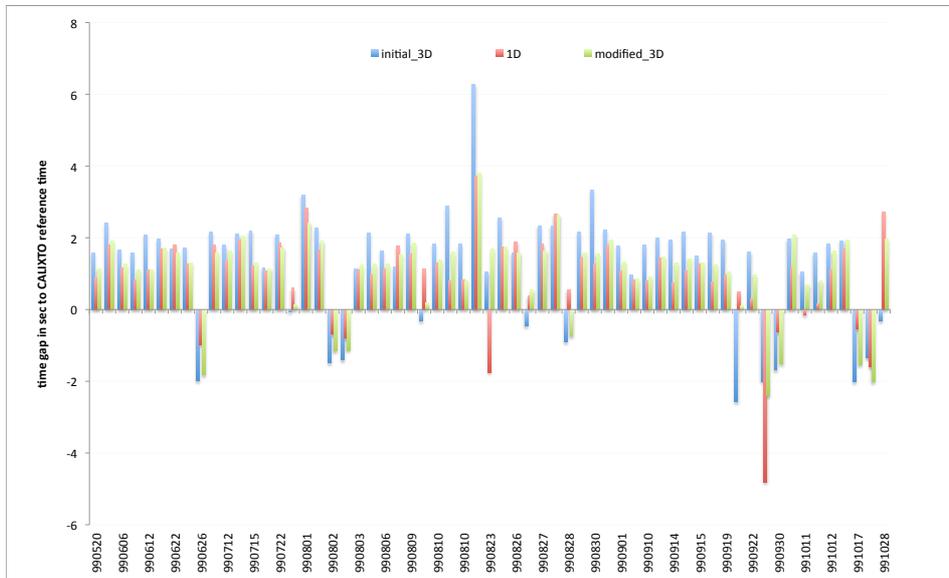


Figure 71 - Histograms representing the reference time shift between the NLLOC estimations computed with the corrected picking times from Tondi et.al (2009) and the CALIXTO99 archive. The comparison is done for the computation through a reference 1D model (red) and the Tondi et al. (2009) 3D model in two versions: the original one (blue) and the corrected one (green).



# APPENDIX D

## DETAILS ON SPECFEM3D IMPLEMENTATION

A <sub>D</sub> .I. IMPORT IT'S OWN TOMOGRAPHIC MODEL .....	130
A <sub>D</sub> .I.1. Modifying the internal mesher code .....	130
A <sub>B</sub> .I.2. The tomographic model format .....	130
A <sub>D</sub> .II. IMPLEMENTING SPECFEM ON CINECA HPC CLUSTERS .....	133
A <sub>D</sub> .II.1. Access to the computational resources.....	133
A <sub>D</sub> .II.2. Working on the Sp6 system .....	134
A <sub>B</sub> .II.3. Working on the FERMI system .....	136

## A<sub>D</sub>.I. IMPORT IT'S OWN TOMOGRAPHIC MODEL IN SPECFEM3D

### A<sub>D</sub>.I.1. Modifying the internal mesher code

The option to import its own tomographic model using the internal mesher is not given in SPECFEM3D, as this way of constructing the computational mesh has been abandoned in favour of external meshers such as CUBIT.

To be able to build the computational mesh, the executable `xgenerate_databases` need to read in mesher output files “`proc***Databases`” that contain indications related to the mesh geometry and geophysical properties.

When importing its own model, these files are written by default, and point to model properties that do not correspond to its own model characteristics. Thus we have modified the mesher code into be able to use our external tomographic model.

The modifications concern the following scripts, written in FORTRAN90, and are indicated in the simplified call tree of the code in Figure 74:

1) for the Mesher part:

```
read_parameter_file.f90
read_value_parameters.f90
meshfem3D.f90
save_databases.f90
create_regions_mesh.f90
```

2) for the Generate\_databases part:

```
model_tomography.f90
```

3) for the text parameter file:

```
Mesh_par_file
```

### A<sub>B</sub>.I.2. The tomographic model format

The tomographic model is a `.xyz` file containing P and S velocities as well as density values given at each incremental point from a given Cartesian starting point. The architecture of the file is shown in Figure 72.

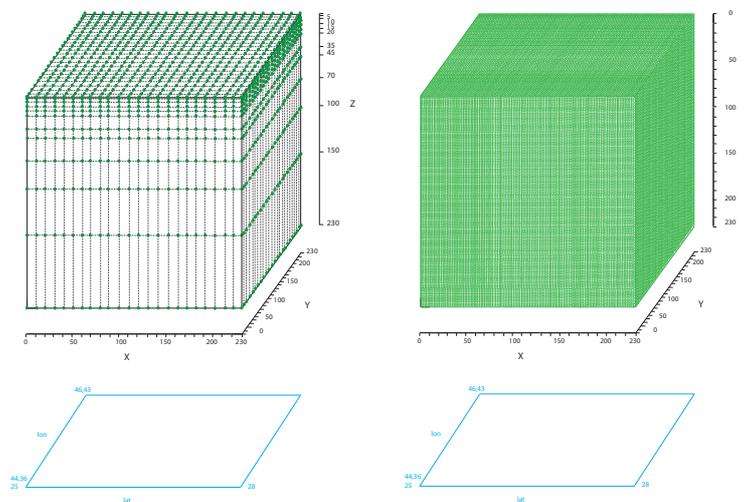
```
ORIG_X ORIG_Y ORIG_Z END_X END_Y END_Z # origin and end points in meters, depth
in negative direction
SPACING_X #spacing of given tomographic points in meters
SPACING_Y
SPACING_Z
NX #number of cell increment
NY
NZ
VP_MIN VP_MAX VS_MIN VS_MAX RHO_MIN RHO_MAX #velocities in m/s
Xi i=0:END_X Y0 Z0 Vpi,0,0 Vsi,0,0 rho0,0,0
Xi i=0:END_X Y1 Z0 Vpi,1,0 Vsi,1,0 rho0,1,0
" " " " " " " "
Xi i=0:END_X YEND_Y Z0 Vpi,END_Y,0 Vsi,END_Y,0 rho0,END_Y,0
Xi i=0:END_X Y0 Z1 Vpi,0,1 Vsi,0,1 rho0,0,1
" " " " " " " "
XEND_X YEND_Y ZEND_Z VPEND_X,END_Y,END_Z VSEND_X,END_Y,END_Z rhoEND_X,END_Y,END_Z
```

Figure 72 - External tomographic model file format for SPECFEM3D.

## Appendix D

The initial tomographic model from Tondi et al. (2009) is a .xyz bloc file with values for P and S velocities as well as density. The tomographic values are regularly 10km spaced on the horizontal components and irregularly spaced on the vertical one with values for the depths at: -1 1 5 10 15 20 35 45 70 100 150 and 230km. The starting point is placed at latitude/longitude/depth 25/44,36/-1 and the whole model covers the region (25-28) latitude degree - (44,36-46,43) longitude degree for a 230km depth.

As the SPECFEM3D tomographic model file has to be regularly spaced in each direction I had to resample the original model in the vertical direction (see Figure 73). Besides, to work with a finer spaced tomographic model, I interpolated it to get a 2km regularly spaced tomographic model as starting reference model. This initial model is finally a 116\*116\*117 .xyz file.



**Figure 73 - From an irregular spaced tomographic model .xyz file to a regularly spaced model. Velocity values indicated by green points are given for a given regular/non regular spaced grid in a 230km<sup>3</sup> Cartesian volume covering the geographical area from latitude 25 to 28 and longitude 44,36-46,43 (blue projection at the bottom). The left representation chooses the initial model format, built up on a non-regular spaced grid of points. The spacing chosen for the regular model given on the left is about 2km for each direction.**

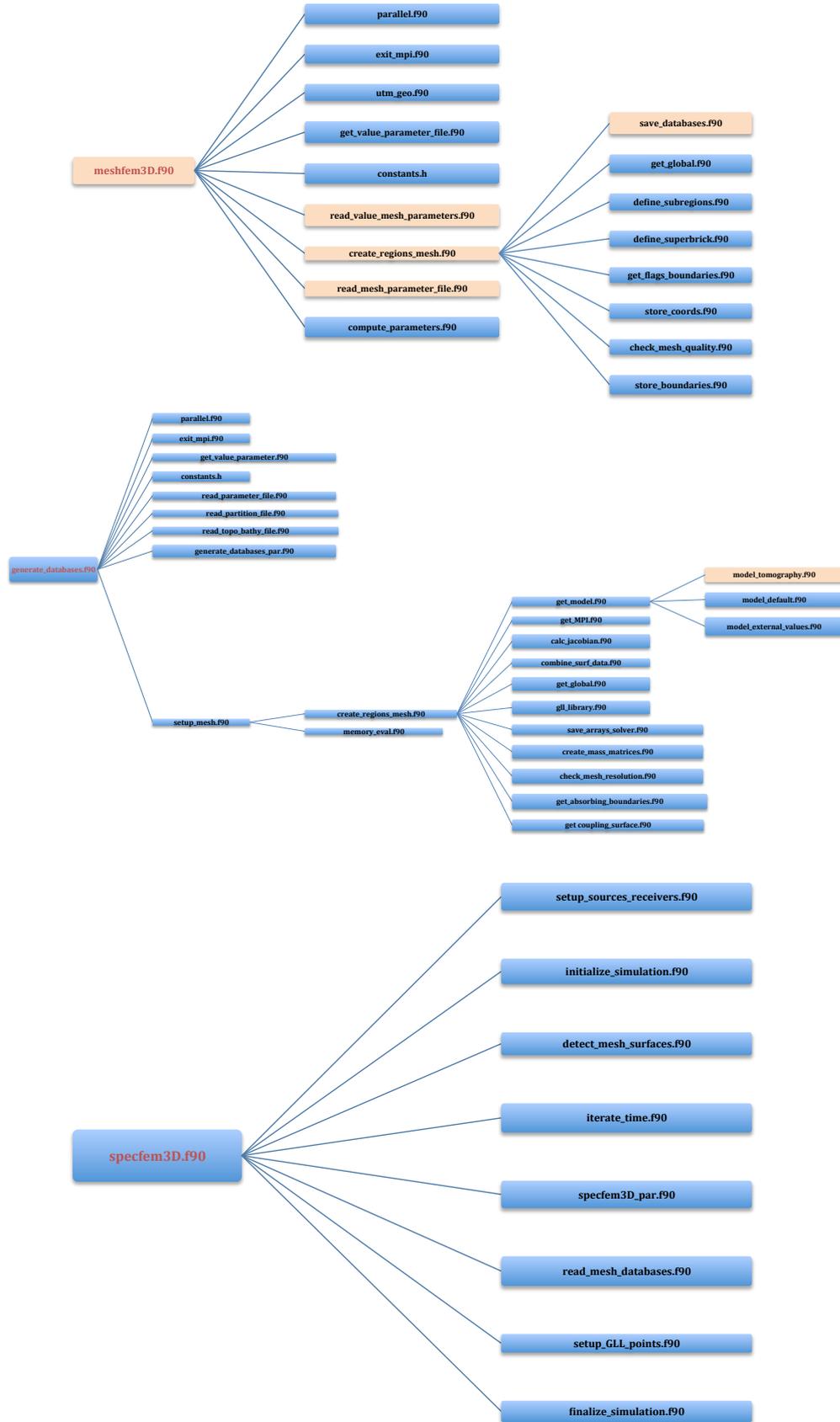


Figure 74 - Simplified call trees of the three SPECfEM executables and modified routines (orange).

## **AD.II. IMPLEMENTING SPECFEM ON CINECA HPC CLUSTERS**

The advantage of using software on clusters is to benefit from a larger amount of computational cores than on our personal laptop or local cluster.

High performance computing (HPC) facilities are reachable on the basis of national and European calls for proposals. During these three years, I have used the facilities of the CINECA HPC centre, in Bologna, Italy. At the beginning of the study the resource available at CINECA on which to run SPECFEM3D was the Sp6 machine. It was replaced in summer 2012 by FERMI with a Blue Gene/Q architecture, involving setting up again the code implementation and meshing.

### **AD.II.1. Access to the computational resources**

The programs ISCRA (Italian SuperComputing Resources Allocation) and PRACE (Partnership for Advanced Computing in Europe) manage the access to the Tier-0 supercomputer by the way of international peer-review procedures ensuring world-class research. Whereas the ISCRA program opens access to the CINECA resources, the PRACE one offers access to any European HPC centre. The resource allocation is defined in terms of CPU<sup>16</sup> (Central Processing Unit) hours through different kind of projects, which are on the FERMI BG/Q:

Class B: standard project requiring between 1M and 10Million CPU hours. The project has a duration of 12 months and the proposal selection comes two times per year.

Class C: small project requiring up to 1Million CPU hours. This class of project is made as “Test and Development” project, where the resources are used to implement, test, optimize applications in the high performance computing environment as well as scale it up to bigger platforms. The project has a duration of 9 months through monthly selection and one can have only one class C project approved each year.

During this study, I have had access to CINECA computing resources through ISCRA projects on the Sp6 and FERMI (substitute of SP6 in summer 2012) machine. On Sp6, one C project of 20.000 CPU hours has been used whereas on FERMI I have worked with two C projects of 2Millions hours and one B project of 10Millions CPU hours, from which I have used more than 2Millions CPU hours for the presented results. The resources of this B project will be used until November 2014 to pursue the inversion. In total, for the presented work, nearly 5 Millions CPU hours have been involved.

---

<sup>16</sup> A CPU hour is defined as the product of the number of core used times the elapsed time involved for the computation.

## AD.II.2. Working on the Sp6 system

Sp6 was an IBM P6-575 Infiniband cluster made of 168 compute nodes, delivering a total of 5.376 compute cores with a memory of 128 Gbytes per node (4Gbytes per core).

It gave the option to work with the computational cores (ST mode) or with virtual CPU (SMT mode), dividing twice the number of cores really used in the run (as well as the memory), and so the cost of the simulation. The point is that each task is managed by a given core, or by half a core in the case of virtual mode (see Figure 75). One has to ask the cluster the number of cores that corresponds to the number of tasks to run (MPI tasks=number of CPU). This management changes with the BG/Q architecture.

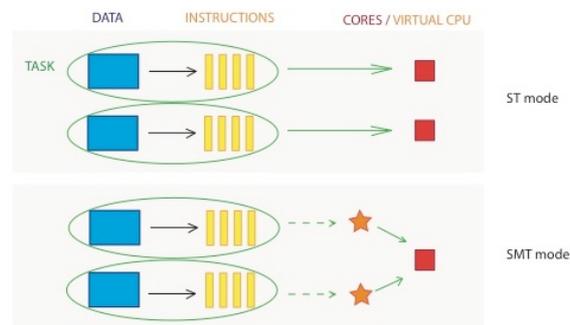


Figure 75 - Tasks management with core and virtual CPU.

In SPECFEM3D, the MPI tasks number can vary in function of the computational geometry built. It depends on the number of computational slices, or MPI processors, on each horizontal components of the mesh. Thus, for a given mesh geometry, different computational set up are possible.

On Sp6, I built a mesh of 500km\*500\*500km with 995.328 spectral elements and a doubling layer. It consisted of:

- 144 spectral elements along the two horizontal components
- 48 spectral elements in depth

Where 144 can be written as:  $144 = 8 \cdot 18 = 8 \cdot 9 \cdot 2 = 8 \cdot 3 \cdot 3 \cdot 2 = 8 \cdot 3 \cdot 6$ . That is four possible combinations<sup>17</sup> for the computational configurations. Thus, the number of MPI processors along the horizontal component can be 18, 9, 6, 3, 2, giving a total of MPI tasks of 324, 81, 36, 9 and 4 respectively.

Figure 76 on the left shows the SPECFEM3D behaviour as the number of MPI tasks increases. This means that I am comparing the different elapsed times needed to run a simulation for a fixed mesh and different computational set up (I show the case of 81 and 324 MPI tasks).

The more the MPI tasks / number of cores increases, the less time is involved in the simulation. This happens because each MPI task / core manages less spectral elements. This is clearly visible in Figure 76 on the right, where the number of MPI tasks / cores is fixed whereas the number of spectral elements in the computational mesh is increased.

<sup>17</sup> As the mesh contains doubling layers, the computational combination has to be written in terms of multiples of 8 spectral elements. Each computational core has to manage the computations regarding a multiple of 8 spectral elements.

## Appendix D

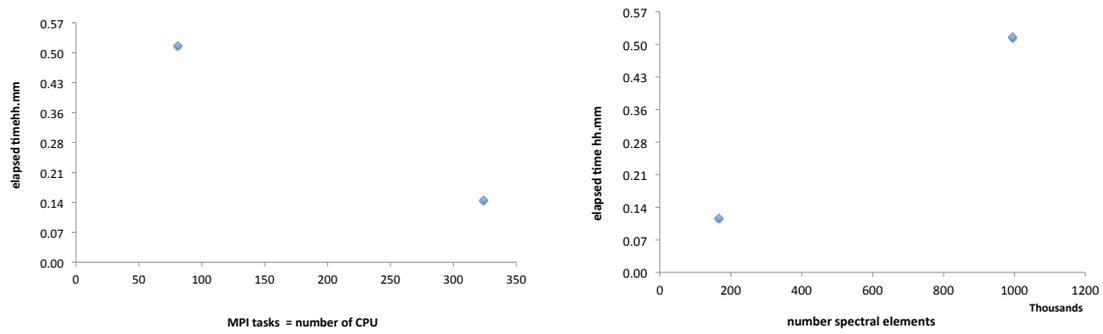


Figure 76 - SPEC-FEM3D efficiency on CINECA Sp6 architecture.

The point is that the finer is the mesh the higher frequencies can be simulated. After having studied the CALIXTO99 database, this was a crucial step in the inversion. As the data were studiable in a high frequency range, corresponding to periods between 1,25 to 2,5 seconds, I had to be able to generate synthetic seismograms with SPEC-FEM3D in the same frequency range. However, increasing the mesh resolution means an increase of the number of spectral elements and thus also an higher computational cost.

From different mesh geometries, I built a period resolution curve to see how to reach nearly 1sec period synthetic waveforms (Figure 77). The minimum period resolution of a given mesh is retrieved from the comparison of the Fourier transforms of synthetic seismograms computed in the mesh and in a finer one. These Fourier transformed signals are identical until a given corner frequency where the largest mesh diverged from the finest one. As the finer mesh is supposed to reach higher frequencies than the larger one, I took this corner frequency as an indication of the maximum frequency accuracy of the largest mesh. This has been done for four different mesh resolutions. I present the details of the geometry and computational configuration for three of them in Figure 77. For the finest mesh, mesh432, I have not been able to check the period resolution, as I could not run an even finer mesh due to computational costs. A logarithmic equation of the other results led to a 0,71sec period resolution estimate for this mesh.

The important conclusion from these results was that I was able to reach a 1,25 second period resolution with a mesh (mesh320) that involves 3 millions of spectral elements and 400 compute cores on Sp6.

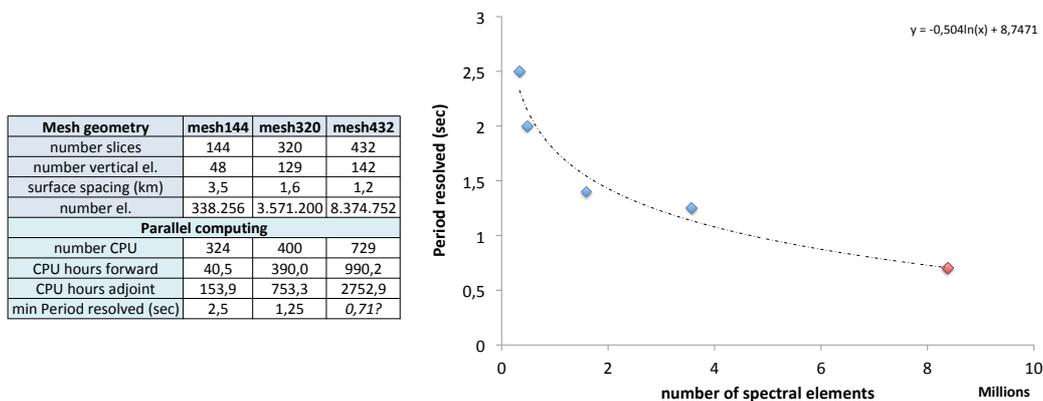


Figure 77 - Mesh set up for different resolution tests on Sp6.

### AB.II.3. Working on the FERMI system

In summer 2012, FERMI has replaced the Sp6 system with a new Blue Gene/Q architecture.

FERMI is composed of 10.240 compute nodes, with 16 cores each, totalling 163.840 compute cores. Each processor comes with 16 Gbytes of RAM (1Gbyte per core). BG/Q is different from Sp6 in terms of MPI tasks management and cores architecture:

- the computational set up modulation is expressed in terms of number of requested nodes (*bg\_size*) and number of MPI tasks allocated per node (*rank\_per\_node*).
- jobs are sent in terms of number of compute nodes, for a multiple of 16 number of cores (it was multiple of 2 on Sp6)
- the minimum number of requested compute nodes for job submission is of 64 (1.024 cores)
- the number of MPI tasks do not match the number of compute cores anymore
- the virtual CPU option for job submission is no more available

I thus had to build up a new mesh, equivalent to the previous one, where the number of MPI tasks is equal to the product of the *bg\_size* with the *rank\_per\_node*. This had to be done under the consideration that the *bg\_size* has to be a multiple of 64, which is the minimum set up requested by the system:

$$MPI\ tasks = rank_{per\ node} * bg_{size} = rank_{per\ node} * B * 64 \quad (87)$$

In SPECfem3D, the number of MPI tasks is linked to the number of spectral elements in horizontal directions:

$$MPI\ tasks = NPROC_{XI} * NPROC_{ETA} \quad (88)$$

$$\text{where } NEX_{XI} = C * 8 * NPROC_{XI} \text{ (resp. } ETA)$$

$$\text{in our case } NEX_{XI} = NEX_{ETA} \text{ thus we can write } NPROC_{XI} = \sqrt{MPI\ tasks} \quad (89)$$

The number of spectral elements in horizontal directions can be then written as:

$$NEX_{XI} = C * 8 * \sqrt{MPI\ tasks} = C * 8 * \sqrt{rank_{per\ node} * B * 64} \quad (90)$$

From these relations, all the possible combinations set up for the BG/Q architecture can be retrieved (see a sample in Table 7). The computations are done on node level, involving a multiple of 16 number of cores. For a fixed number of MPI tasks, different configurations can be applied. For example, for the mesh with  $NEX_{XI}=512$ , involving 4096 MPI tasks (orange highlighted lines in the table), one can build the computational environment in these ways:

<i>bg_size</i> = 64	&	<i>rank_per_node</i> =64	involving 64*16=1024 compute cores
<i>bg_size</i> =128	&	<i>rank_per_node</i> =32	involving 128*16=2048 compute cores
<i>bg_size</i> =256	&	<i>rank_per_node</i> =16	involving 256*16=4096 compute cores
<i>bg_size</i> =512	&	<i>rank_per_node</i> = 8	involving 512*16=8192 compute cores

## Appendix D

The `rank_per_node` parameter shares out a given number of MPI tasks to a single node (between 16 cores). This setting plays a role on the memory allocated to each task. The more filled is a node with MPI tasks, the less memory is allocated to each of these tasks. Allocating 16 MPI tasks to a single node means sharing out 16 MPI tasks to 16 cores (each of them having a 1GByte memory chip) and thus assigning 1GByte of memory space for each single MPI task.

Table 7 - Sample of possible combinations for building a mesh on FERMI-BG/Q.

C	B	rankpernode	MPI_tasks	NEX_XI	NPROC_XI	Mem/task	bg_size
1	1	1	64	64	8	16	64
		4	256	128	16	4	
		16	1024	256	32	1	
		64	4096	512	64	0,25	
	2	2	256	128	16	8	128
		8	1024	256	32	2	
		32	4096	512	64	0,5	
	4	1	256	128	16	16	256
		4	1024	256	32	4	
		16	4096	512	64	1	
		64	16384	1024	128	0,25	
	8	2	1024	256	32	8	512
		8	4096	512	64	2	
		32	16384	1024	128	0,5	
	9	1	576	192	24	16	576
		4	2304	384	48	4	
		16	9216	768	96	1	
		64	36864	1536	192	0,25	
	16	1	1024	256	32	16	1024
		4	4096	512	64	4	
16		16384	1024	128	1		
64		65536	2048	256	0,25		

To choose the best setup, I looked at the behaviour of SPECFEM3D for a fixed number of MPI tasks and an increasing number of compute cores. Results, presented in Figure 78, are related to the mesh with `NEX_XI=512` (orange highlighted lines in the table). I tested the four computational configurations detailed previously (the points from left to right stand for a `bg_size` of 64, 128, 256 and 512). These tests have been done in the base of 10 minutes simulations and the estimation time given by SPECFEM3D for a complete simulation. The simulations started being stable in terms of running time from the third set up (`bg_size` of 256 for 4096 compute cores). The stability is thus reached when the number of MPI tasks corresponds exactly to the number of compute cores involved in the simulation. This corresponds to a computational environment equivalent to the Sp6 one. With these settings, the memory allocated for each single task is of almost 1Gbyte. This appeared to be the best computational setup for my study.

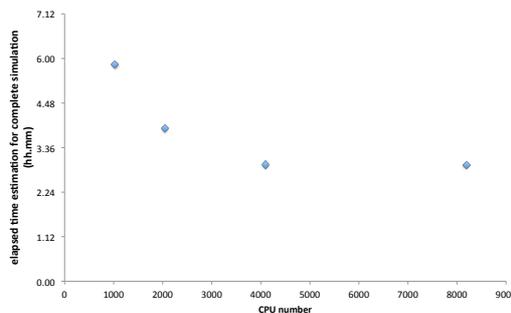


Figure 78 - SPECFEM3D efficiency on CINECA BG/Q architecture.



# BIBLIOGRAPHY

Aki K. and Richards P.G., 1980, Quantitative seismology. Freeman, San Francisco, Vol. I and II

Ardeleanu L., Radulian M., Leny J. and Panza G. F., 2005, Source Parameters of Weak Crustal Earthquakes of the Vrancea Region from Short-period Waveform Inversion, *Pure Applied Geophysics* 162, 495–513 doi: 10.1007/s00024-004-2618-y

Bala A., Radulian M. and Popescu E., 2003, Earthquakes distribution and their focal mechanism in correlation with the active tectonic zones of Romania, *Journal of Geodynamics* 36, 129–145

Bamberger A., Chavent, G., Lailly, P. and Hemon, C., 1982, Inversion of normal incidence seismograms, *Geophysics* 47, 737–770

Besutiu, L., 2001, Vrancea active seismic area: a continental unstable triple junction? *Rev. Roum. Geophysique* 45, 59-72

Bozdog E., Trampert J. and Tromp J., 2011, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophysical Journal International* 185, 845–870  
doi: 10.1111/j.1365246X.2011.04970.x

Burchfiel B.C., 1980, Eastern European Alpine system and the Carpathian orocline as an example of collision tectonics, *Tectonophysics* 63, 31-61

Carbunar O. F., Radulian M., 2011, Geometrical constraints for the configuration of the Vrancea (Romania) intermediate-depth seismicity nest, *Journal of Seismology* Vol. 15, Issue 4, pp. 579-598

Carcione, J.M., 1990, Wave propagation in anisotropic linear viscoelastic media: Theory and simulated wavefields, *Geophysical Journal International* 101, 739-750.

Cauchy A., 1847, Méthode générale pour la résolution des systèmes d' équations simultanées, *Comptes Rendus Hebdomadaires des séances de L'Académie des Sciences de Paris* 25, 536-538

Chaljub E., 2000, Modélisation numérique de la propagation d'ondes sismiques en géométrie sphérique: application à la sismologie globale, Ph.D. thesis, Université Paris VII Denis Diderot, Paris, France.

Chalot-Prat F., Girbacea R., 2000, Partial delamination of continental mantle lithosphere, uplift-related crust-mantle decoupling, volcanism and basin formation: a new model for the Pliocene-Quaternary evolution of the southern East-Carpathians, Romania, *Tectonophysics*

## *Bibliography*

Chapman C. H., 2004, *Fundamentals of seismic waves propagation*, Cambridge University Press ISBN 13:9780521815383

Chavent G., 1974, Identification of function parameters in partial differential equations in *Identification of parameter distributed systems*, eds Goodson R.E., Polis New-York, ASME 1974

Dahlen F. A., Hung S.-H. and Nolet G., 2000, Fréchet kernels for finite-frequency traveltimes -I. Theory, *Geophysical Journal International* 141, 157-174

Dahlen F. A. and Baig A.M., 2002, Fréchet kernels for body-wave amplitudes, *Geophysical Journal International* 150, 440–466

Dando B. D. E., Stuart, G. W., Houseman G. A., Hegedus E., Bruckl E. and Radovanovic S., 2011, Teleseismic tomography of the mantle in the Carpathian–Pannonian region of central Europe, *Geophysical Journal International* 186, 11–31 doi: 10.1111/j.1365-246X.2011.04998.x

Ducea M. N., 2011, Fingerprinting orogenic delamination, *Geology* 39, 191–192; doi: 10.1130/focus022011.1.

Dumbser M. and Käser M., 2006, An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes II: The Three-Dimensional Isotropic Case, *Geophysical Journal International* 167(1), 319-336

Dziewonski A. M., and Woodhouse J. H., 1983, An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981, *Journal of Geophysical Research* 88(B4), 3247–3271, doi: 10.1029/JB088iB04p03247

Ekström G., Nettles M., Dziewonski A.M., 2012, The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Physics of the Earth and Planetary Interiors* 200–201, 1–9

Enciu D. M., Knapp C. C. and Knapp J. H., 2009, Revised crustal architecture of the southeastern Carpathian foreland from active and passive seismic data, *Tectonics* Vol. 28, tc4013 doi: 10.1029/2008tc002381

Fan G., Wallace T. C. and Dapeng Z., 1998, Tomographic imaging of deep velocity structure beneath the eastern and southern Carpathians, Romania: implications for continental collision, *Journal of Geophysical Research* 103, 2705–2724

Fichtner A., Bunge H.-P., Igel H., 2006a, The adjoint method in seismology I. Theory, *Physics of the Earth and Planetary Interiors* 157, 86–104

Fichtner A., Bunge H.-P., Igel H., 2006b, The adjoint method in seismology—II.Applications: traveltimes and sensitivity functionals, *Physics of the Earth and Planetary Interiors* 157, 105–123

## *Bibliography*

Fichtner A., Kennett B. L. N., Igel H. and Bunge H.-P., 2008, Theoretical background for continental and global scale full-waveform inversion in the time-frequency domain, *Geophysical Journal International* 175, 665–685 doi: 10.1111/j.1365-246X.2008.03923.x

Fichtner A., Igel H., Bunge H.-P. and Kennett, B. L. N., 2009, Simulation and Inversion of Seismic Wave Propagation on Continental Scales Based on a Spectral-Element Method, *Journal of Numerical Analysis, Industrial and Applied Mathematics* Vol. 4, no. 1-2, pp. 11-22 ISSN 1790–8140

Fichtner A., Kennett B. L. N., Igel H. and Bunge H.P., 2009, Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophysical Journal International* 179, 1703–1725 doi: 10.1111/j.1365-246X.2009.04368.x

Fichtner A., 2009, Full seismic waveform inversion for structural and source parameters, PhD thesis

Fichtner A., 2011, Full seismic waveform modelling and inversion, advances in geophysical and environmental mechanics and mathematics series, ed. Springer

Fichtner A. and Trampert J., 2011, Hessian kernels of seismic data functionals based upon adjoint techniques, *Geophysical Journal International* 185, 775–798 doi: 10.1111/j.1365-246X.2011.04966.x

Fillerup M.A., Knapp J.H., Knapp C.C. and Raileanu V., 2010, Mantle earthquakes in the absence of subduction? Continental delamination in the Romanian Carpathians. *Lithosphere* 2, 333-340 doi: 10.1130/L102.1

Fletcher R. and Reeves C. M., 1964, Function minimization by conjugate gradients. *Comput, J.* 7, 149-154.

Foldvary G. Z., 1988, *Geology of the Carpathian Region*, World Scientific, Singapore.

Fornberg B., 1987, The pseudospectral method: Comparisons with finite differences for the elastic wave equation, *Geophysics* Vol. 52, no. 4; p. 483-501,

Fuchs K., Bonjer K.-P., Bock G., Cornea I., Radu C, Enescu D., Jianu D., Nourescu A., Merkle G., Moldoveanu T., and Tudorache, G., 1979, The Romanian earthquake of March 4, 1977, II, Aftershocks and migration of seismic activity, *Tectonophysics* 53, 225–247

Girbacea R. and Frisch W., 1998, Slab in the wrong place: lower lithospheric mantle delamination in the last stage of the Eastern Carpathian subduction retreat, *Geology* 26, 611–614

Golub G.H. and Pereyra V., 1973, The differentiation of the pseudoinverse nonlinear least square problems whose variable separate. *SIAM Journal on Numerical analysis* 10, 2

## *Bibliography*

Golub G.H. and Pereyra V., 2003, Separable nonlinear least-squares: the variable projection method and its applications, *Inverse Problems* 19, R1-R26

Heidbach O., Ledermann P., Kurfeß, G. Peters D., Buchman T., Matenco L., Negut M., Sperner B., Müller B., Nuckelt A. and Schmitt G., 2007, Attached or not attached: slab dynamics beneath Vrancea, Romania, International symposium on strong Vrancea earthquakes and risk mitigation, Bucharest, Romania

Helfrich G., Wookey J., and Bastow I., 2013, *The Seismic Analysis Code*. 1st ed. Cambridge: Cambridge University Press, 2013. Cambridge Books Online. Web. <http://dx.doi.org/10.1017/CBO9781139547260>

Hurukawa N., Popa M. and Radulian M., 2008, Relocation of large intermediate-depth earthquakes in the Vrancea region, Romania, since 1934 and a seismic gap, *Earth Planets Space* 60, 565–572,

Igel H., Nissen-Meyer T. and Jahnke G., 2002, Wave propagation in 3D spherical sections: effects of subduction zones, *Physics of the Earth and Planetary Interiors* 132, 219–234

Ismail-Zadeh A., Sokolov V. and Bonjer K.-P., 2007, Tectonic stress, seismicity, and seismic hazard in the southeastern Carpathians, *Natural Hazards* 42, 493–514 doi: 10.1007/s11069-006-9074-1

Ismail-Zadeh A., Matenco L., Radulian M., Cloetingh S. and Panza G., 2012, Geodynamics and intermediate-depth seismicity in Vrancea (the south-eastern Carpathians): Current state-of-the art, *Tectonophysics* 530–531, 50–79

Kikuchi M. and Kanamori H., 1991, Inversion of complex body wave-III, *Bulletin of the Seismological Society of America* 81, 2335-2350.

Kim Y., Liu Q. and Tromp J., 2011, Adjoint centroid-moment tensor inversions, *Geophysical Journal International* 186, 264–278  
doi: 10.1111/j.1365-246X.2011.05027.x

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, 89–101

Knapp J. H., Knapp C. C., Raileanu V., Matenco L., Mocanu V. and Dinu C., 2005, Crustal constraints on the origin of mantle seismicity in the Vrancea Zone, Romania: The case for active continental lithospheric delamination, *Tectonophysics* 410, 311–323

Knapp J.H., Knapp C.C., 2012, Active Continental Lithospheric Delamination in the Southeastern Carpathians, American Geophysical Union (AGU), San Francisco, California

Komatitsch D. and Tromp J., 1999, Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophysical Journal International* 139, 806-822

## *Bibliography*

Komatitsch D. and Tromp J., 2002a, Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophysical Journal International* 149, 390–412

Komatitsch D. and Tromp J., 2002b, Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophysical Journal International* 150, 303–318

Komatitsch D., Liu Q., Tromp J., Suss P., Stidham C. and Shaw J.H., 2004, Simulations of Ground Motion in the Los Angeles Basin Based upon the Spectral-Element Method, *Bulletin of the Seismological Society of America* Vol. 94, No. 1, pp. 187–206

Komatitsch D., Tsuboi S. and Tromp J., 2005, The Spectral-Element Method in Seismology, *Geophysical Monograph Series* 157, 205-227 doi: 10.1029/156GM13

Komatitsch D., Erlebacher G., GÖddeke D. and Michéa D., 2010, High-order finite-element seismic wave propagation modeling with MPI on a large GPU cluster, *Journal of Computational Physics* 229, 7692–7714 doi:10.1016/j.jcp.2010.06.024

Koulakov I., Zaharia B., Enescu B., Radulian M., Popa M., Parolai S. and Zschau J., 2010, Delamination or slab detachment beneath Vrancea? New arguments from local earthquake tomography, *Geochemistry Geophysics Geosystem* 11, Q03002, doi:10.1029/2009GC002811

Linzer H.-G., Frisch W., Zweigel P., Girbacea R., Hann H.-P. and Moser F., 1998, Kinematic evolution of the Romanian Carpathians, *Tectonophysics* 297, 133–156

Lysmer J. and Drake L. A., 1972, A finite element method for seismology, Ch. 6 in *Methods in Computational Physics* 11, Seismology, Alder, B., Fernbach, S., and Bolt, B. A., Editors, Academic Press

Liu Q., Polet J., Komatitsch D. and Tromp J., 2004, Spectral-element moment tensor inversions for earthquakes in southern California, *Bulletin of the Seismological Society of America* 94, 1748–1761.

Liu Q. and Tromp J., 2006, Finite-Frequency Kernels Based on Adjoint Methods, *Bulletin of the Seismological Society of America* Vol. 96, No. 6, pp. 2383–2397, doi: 10.1785/0120060041

Liu Q. and Tromp J., 2008, Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophysical Journal International* 174, 265–286 doi: 10.1111/j.1365-246X.2008.03798.x

Liu Q., and Gu Y.J, 2012, Seismic imaging: From classical to adjoint tomography, *Tectonophysics*, [http://dx.doi.org/ 10.1016/j.tecto.2012.07.006](http://dx.doi.org/10.1016/j.tecto.2012.07.006)

## *Bibliography*

Lomax A., Virieux J., Volant P. and Berge C., 2000, Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations, *Advances in Seismic Event Location*, Thurber, C.H., and N. Rabinowitz (eds.), Kluwer, Amsterdam, 101-134

Lorinczi P., Houseman G.A., 2009, Lithospheric gravitational instability beneath the southeast Carpathians, *Tectonophysics* 474, 322–336, doi: 10.1016/j.tecto.2008.05.024

Luo Y., Zhu H., Nissen-Meyer T., Morency C. and Tromp J., 2009, Seismic modeling and imaging based upon spectral-element and adjoint methods, *SPECIAL SECTION: seismic modelling*

Maggi A., Tape C., Chen M., Chao D. and Tromp J., 2009, An automated time-window selection algorithm for seismic tomography, *Geophysical Journal International* 178, 257–281 doi: 10.1111/j.1365-246X.2009.04099.x

Manea V.C and Manea M., 2009. Thermally induced stresses beneath the Vrancea area, Integrated research on the intermediate depth earthquake genesis within Vrancea zone, In Besutiu, L. (Ed.), Vergiliu Publishing House pp.172-183. ISBN 978-973-7600-59-22009

Marquering H., Dahlen F. A. and Nolet G., 1999, Three-dimensional sensitivity kernels for finite-frequency traveltimes: the bananadoughnut paradox, *Geophysical Journal International* 137, 805-815

Martin M., Wenzel F. and the CALIXTO working group, 2006. High-resolution teleseismic body wave tomography beneath SE-Romania – II. Imaging of a slab detachment scenario, *Geophysical Journal International* 164, 579–595 doi: 10.1111/j.1365-246X.2006.02884.x

McKenzie D., 1972, Active Tectonics of the Mediterranean Region. *Geophysical Journal of the Royal Astronomical Society* 30, 109–185. doi: 10.1111/j.1365-246X.1972.tb02351.x

Morelli, A. and Dziewonski, A., 1991, Joint determination of lateral heterogeneity and earthquake location, in *Glacial Isostasy, Sea-Level and Mantle Rheology*, Kluwer Academic Publishers, Dordrecht.

Müller B., Heidbach O., Negut M., Sperne B., Buchmann T., 2010, Attached or not attached—evidence from crustal stress observations for a weak coupling of the Vrancea slab in Romania, *Tectonophysics* 482, 139–149

Oncescu M.C., 1984, Deep structure of Vrancea region, Romania, inferred from simultaneous inversion for hypocenters and 3-D velocity structure, *Annal of Geophysics* 2, 23–28.

Patera A.T., 1984, A spectral element method for fluid dynamics: Laminar flow in a channel expansion, *Journal of Computational Physics* 54, 468–488.

Peter D., Komatitsch D., Luo Y., Martin R., Le Goff N., Casarotti E., Le Loher P., Magnoni F., Liu Q., Blitz C., Nissen-Meyer T., Basini P. and Tromp J., 2011,

## *Bibliography*

Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophysical Journal International* 186, 721–739 doi: 10.1111/j.1365-246X.2011.05044.x

Piromallo C. and Morelli A., 2003, P-wave tomography of the mantle under the Alpine-Mediterranean area, *Journal of Geophysical Research* 108 (B2)

Plessix R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophysical Journal International* 167, 495–503

Popa M., Kissling E., Radulian M., Bonjer K.-P., Enescu D., Dragan S. and the CALIXTO Research Group, 2001, Local source tomography using body waves to deduce a minimum 1D velocity model for Vrancea (Romania) zone, *Romanian Report in Physics* 53, 519-536

Prieto G. A., Beroza G. C., Barrett S. A., López G.A. and Florez M., 2012, Earthquake nests as natural laboratories for the study of intermediate-depth earthquake mechanics, *Tectonophysics* 570–571 42–56

Qin Y., Capdeville Y., Montagner J.-P., Boschi L. and Becker T.W., 2009, Reliability of mantle tomography models assessed by spectral element simulation, *Geophysical Journal International* 177, 125–144 doi: 10.1111/j.1365-246X.2008.04032.x

Radu, C., 1979, Catalogue of strong earthquakes occurred on the Romanian territory. Part I – before 1901; Part II – 1901-1979, in Cornea si C. Radu (eds.), *Seismological Studies on the March 4, 1977 Earthquake*, Bucuresti, 723-752.

Radu, C., 1991, Strong earthquakes occurred on the Romanian territory in the period 1901-1990, *Vitralii* 3, 12-13.

Rawlinson N. and Sambridge M., 2003, Seismic traveltime tomography of the crust and lithosphere, *Advances in Geophysics*, 46, 81-197.

Rawlinson N., Pozgay S. and Fishwick S., 2010, Seismic tomography: A window into deep Earth, *Physics of the Earth and Planetary Interiors* 178, 101–135

Raykova R.B and Panza G.F., 2006, Surface waves tomography and non-linear inversion in the southeast Carpathians, *Physics of the Earth and Planetary Interiors* 157, 164–180 doi:10.1016/j.pepi.2006.03.019

Reasenber P.A., and Oppenheimer D., 1985, FPFIT, FPLOT, and FPPAGE: Fortran computer programs for calculating and displaying earthquake fault-plane solutions, U.S. Geological Survey Open-File Report 85-739

Ren Y., Stuart G.W., Houseman G.A., Dando B., Ionescu C., Hegedus E., Radovanovic S., Shen Y. and South Carpathian Project Working Group, 2012, Upper mantle structures beneath the Carpathian–Pannonian region: Implications for the geodynamics of continental collision, *Earth and Planetary Science Letters* 349–350, 139–152

## *Bibliography*

Sandu I., Zaicenco A., 2008, Focal mechanism solutions for Vrancea seismic area, A. Zaicenco et al. (eds.), *Harmonization of Seismic Hazard in Vrancea Zone*, Springer Science + Business Media B.V., pp. 17-46

Schmid S., Bernoulli D., Fügenschuh B., Matenco L., Schefer S., Schuster R., Tischler M. and Ustaszewski K., 2008, The Alpine-Carpathian-dinaridic orogenic system: correlation and evolution of tectonic units, *Swiss Journal of Geosciences* 101, 139–183 doi 10.1007/s00015-008-1247-3

Seriani G., 1998, 3-D large-scale wave propagation modeling by a spectral element method on a Cray T3E multiprocessor, *Comput. Methods Appl. Mech. Engrg.* 164, 235–247.

Sieminski A., Liu Q., Trampert J. and Tromp J., 2007a, Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods, *Geophysical Journal International* 168, 1153–1174 doi: 10.1111/j.1365-246X.2006.03261.x

Sieminski A., Liu Q., Trampert J. and Tromp J., 2007b, Finite-frequency sensitivity of body waves to anisotropy based upon adjoint methods, *Geophysical Journal International* 171, 368–389 doi: 10.1111/j.1365-246X.2007.03528.x

Spencer C. and Gubbins D., 1980, Travel-time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media, *Geophysical Journal of the Royal Astronomical Society* 63, 95-116

Sperner B., Lorenz F., Bonjer K., Hettel S., Muller B. and Wenzel F., 2001, Slab break-off: abrupt cut or gradual detachment? New insights from the Vrancea Region (SE Carpathians, Romania), *Terra Nova* 13, 172-179

Tape C., Liu Q. and Tromp J., 2007, Finite-frequency tomography using adjoint methods— Methodology and examples using membrane surface waves, *Geophysical Journal International* 168, 1105–1129 doi: 10.1111/j.1365-246X.2006.03191.x

Tape C., Liu Q., Maggi A. and Tromp J., 2009, Adjoint tomography of the southern California crust, *Science* 325

Tape C., 2009, *Seismic Tomography of Southern California Using Adjoint Methods*, PhD thesis

Tape C., Liu Q., Maggi A. and Tromp J., 2010, Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophysical Journal International* 180, 433–462 doi: 10.1111/j.1365-246X.2009.04429.x

Tarantola A., 1984a, Linearized inversion of seismic data, *Geophysical Prospecting* 32, 998-1015

Tarantola A., 1984b, Inversion of seismic reflection data in the acoustic approximation, *Geophysics* Vol. 49, No. 8, pp. 1259-1266

## *Bibliography*

Tarantola A., 1987a. Inverse problem theory, methods for data fitting and model parameter estimation, Elsevier, Amsterdam.

Tian Y., Montelli R., Nolet G., Dahlen F.A., 2007, Computing traveltimes and amplitude sensitivity kernels in finite-frequency tomography, *Journal of Computational Physics* 226, 2271–2288

Tondi R., de Franco R. and Barzaghi R., 2000, Sequential integrated inversion of refraction and wide-angle reflection traveltimes and gravity data for two-dimensional velocity structures, *Geophysical Journal International* 141, 679 – 698doi:10.1046/j.1365-246x.2000.00104.x.

Tondi R., Achauer U., Landes M., Davì R., and Besutiu L., 2009, Unveiling seismic and density structure beneath the Vrancea seismogenic zone, Romania, *Journal of Geophysical Research* 114, B11307 doi: 10.1029/2008JB005992

Trampert J., Van der Hilst R.D., 2005, Towards a Quantitative Interpretation of Global Seismic Tomography, *Earth's Deep Mantle: Structure, Composition, and Evolution*, Geophysical Monograph Series 160 doi: 10.1029/160GM05

Tromp J., Tape C. and Liu Q., 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophysical Journal International* 160, 195-216 doi: 10.1111/j.1365-246X.2004.02453.x

Tromp J., Komatitsch D. and Liu Q., 2008, Spectral-Element and Adjoint Methods in Seismology, *Communications in Computational Physics* Vol. 3, No. 1, pp. 1-32

Tromp J., Komatitsch D., Hjørleifsdóttir V., Liu Q., Zhu H., Peter D., Bozdag E., McRitchie D., Friberg P., Trabant C. and Hutko A., 2010, Near real-time simulations of global CMT earthquakes, *Geophysical Journal International* 183, 381–389 doi: 10.1111/j.1365-246X.2010.04734.x

Valentine A. P. and Woodhouse J.H., 2010, Reducing errors in seismic tomography: combined inversion for sources and structure, *Geophysical Journal International* 180, 847–857 doi: 10.1111/j.1365-246X.2009.04452.x

Van Leewen T. and Mulder W. A., 2009, A variable projection method for waveform inversion, 71st EAGE conference & Exhibition, Amsterdam

Virieux J. and Operto S., 2009, An overview of full-waveform inversion in exploration geophysics, *Geophysics* vol. 74, no. 6 10.1190/1.3238367

Wenzel F., Achauer U., Enescu D., Kissling E., Russo R., Mocanu V. and Musacchio G., 1998, Detailed look at final stage of plate break-off is target of study in Romania, *EOS* 79, 589–594

Wortel M.J.R. and Spakman W., 2000, Subduction and Slab Detachment in the Mediterranean-Carpathian Region, *Science* 290, 1910-1917 doi: 10.1126/science.290.5498.19103

## *Bibliography*

Zarifi Z. and Havskov J., 2003, Characteristics of dense nests of deep and intermediate-depth seismicity, *Advances in Geophysics* 46, 237–278.

Zhao D. and Kayal J.R., 2000, Impact of seismic tomography on Earth sciences, *Current Science* 79, no. 9, Special section: Seismology, 1208-1214

Zhu H., Bozdag E., Peter D. and Tromp J., 2012, Structure of the European upper mantle revealed by adjoint tomography, *Nature Geoscience* 5, 493-498  
doi: 10.1038/NGEO1501

# Acknowledgments

*My PhD. took part of the European Marie curie project QUEST that involved researchers and students from all Europe. Thus, this has implied for me, apart from the scientific work, to move from France to Italy, learn a new language and build new roots in Bologna. This work would not have been the same without the support of the people that I want to thank here. In this personal part, I would like to address my acknowledgments in the three languages that have animated my three years study: the language of Dante, Shakespeare and Molière.*

*Nella lingua di Dante...*

*Vorrei ringraziare il mio tutor Andrea Morelli per avermi dato l'opportunità di partecipare al progetto QUEST e per avermi sostenuta e incoraggiata nei passaggi difficili della tesi, cercando sempre di farmi vedere il lato positivo delle cose.*

*Ringrazio Peter Danecek che mi è stato di grande aiuto all' inizio della tesi per farmi familiarizzare con SPECFEM e il mondo della computazione in parallelo, e per le sue valutazioni e consigli durante la tesi.*

*Ringrazio Rosaria Tondi per la sua collaborazione e per i suoi consigli nella fase iniziale del mio trasferimento a Bologna.*

*Ringrazio i miei colleghi di lavoro, Adriano, Brunella, Iolanda, Letizia, Paola, Pablo e Simone, per la loro simpatia, il loro buon umore e il loro contributo nel farmi imparare tutte le parolacce dell' italiano, dalla Sicilia all' Emilia, durante le pause pranzo nella cucina INGV.*

*Ringrazio Federica per i suoi consigli essenziali e la sua grande disponibilità.*

*Al di fuori del lavoro, ringrazio: i miei coinquilini Nicola e Michele con cui la convivenza è stata molto facile e piacevole; Yamilet la mia radiosa maestra di salsa cubana; Sophie che mi fa riavvicinare al Tedesco; Maria-Ausilia che mi ha iniziata al fango del rugby nel mio primo anno a Bologna.*

*Nella lingua di Shakespeare....*

*I would like to thank the team of Schlumberger Cambridge with which I have worked for three months: Colin Thompson and James Rickett. I am very grateful of the opportunity they gave me to work in an industrial research frame and their warm welcome. I thank also Philip, Emma and Juerg with whom I have shared the office room, lunchtimes and even a typical college dinner.*

*I thank my QUEST colleagues Anna-Lucia, Costas, Laura, Lubiza, Lucia and Rafael, as well as Jean-Paul Montagner, Eléonore Stutzmann and Anna Ferreira for their good advices.*

*I thank Carl Tape for his advices and his availability to answer my demands.*

## *Acknowledgments*

### *Dans la langue de Molière*

*La difficulté de vivre loin de ses racines a été allégée par les nouvelles rencontres et amitiés tissées à Bologne, mais surtout par le soutien et l'intérêt constant de mes proches.*

*Je remercie mes parents qui m'ont soutenue et encouragée tout au long de cette aventure; mes grands-parents qui, bien qu'étrangers à mes études, se sont révélés de grands supporteurs et ont démontré un vif intérêt; mes soeurs Marie, Charlotte et Laure ainsi que mes amis Graziella et Florent, qui m'ont rendu visite à Bologne et avec qui j'ai partagé des instants ensoleillés et joyeux.*

*Un merci particulier à Hervé, Carmine, Marie et un correcteur anonyme, qui ont participé, avec leurs moyens, à la correction de ma thèse.*

*Pour finir, la dernière mais pas des moindres, je remercie Cristina Donà qui m'a permis de rencontrer la personne qui m'a soutenue à chaque instant de ce doctorat, et partage chaque jour de ma nouvelle vie italienne.*