Università degli Studi di Bologna

DIPARTIMENTO DI FISICA

DOTTORATO DI RICERCA IN GEOFISICA

GEO/10

Surface wave tomography in the European and Mediterranean region

Dottoranda:

Dott.ssa Renata Schivardi

Tutore:

Dott. Andrea Morelli

XIX CICLO

ANNO ACCADEMICO 2006/2007

Contents

1. Introduction

The broad European and Mediterranean region is characterized by an extremely complex tectonic setting, driven by the major convergence between Eurasian and African plates. A detailed model of the upper mantle in this region is fundamental to improve our understanding of its geodynamical evolution. Seismic tomography can help to address this problem modeling seismic speed anomalies, that can be related to different tectonic features, such continental roots, rifting areas, magmatic provinces, plumes or subducting slabs. Due to high seismicity rates and dense seismograph coverage, this region has been the subject of many tomographic studies from regional to local scale.

Traveltime high resolution models of P-wave speed anomalies [*Spakman et al.*, 1993; *Piromallo and Morelli*, 2003] have illuminated the deep structure of the Mantle, but at shallow depth they often suffer from uneven ray coverage, being strongly dependent on stations and epicenters distribution. Regional S-wave velocity models have been retrieved from the analysis of surface wave dispersion group or phase velocity [*Ritzwoller and Levshin*, 1998; *Villaseñor et al.*, 2001], from waveform inversion of surface waves mode coupling [*Marquering and Snieder*, 1996] or both surface and body waves [*Marone et al.*, 2004]. However, the non-optimal distribution of observatories and seismic sources has affected regional tomographic models. Global models derived from surface wave data image the large-scale structures of the region, but their resolution is insufficient to describe its complexity [*Shapiro and Ritzwoller*, 2002; *Boschi and Ekström*, 2002; *Ritsema et al.*, 1999; *Zhou et al.*, 2006; *Trampert and Woodhouse*, 1995]. Global models with finer parameterization on Mediterranean [*Boschi et al.*, 2004] have been proposed and recent modeling of surface waves from ambient noise gave new insights into the shallowest European upper mantle [*Yang et al.*, 2006].

The increased availability of high quality seismic records from permanent observatories and from the recent temporary deployment experiment RETREAT in northern Apennines, gave us the opportunity to exploit new data, that can provide new and finer constraints to the tomographical problem. We present a new surface waves tomography model, because of the enhanced sensitivity of those waves to shallow structure and their more uniform spatial coverage in complex sources-stations distributions.

The inverse problem of obtaining a V_S three-dimensional model from analysis of surface waves dispersion can be solved in different ways. [*Marone et al.*, 2004] use the

partitioned waveform inversion of [*Van der Lee and Nolet*, 1997], where the 1-D average S-velocity structure along each path is first determined by non-linear waveform fitting, and in a second step the 1-D path averaged structures are combined in a damped least-squares linear inversion for a 3-D S-velocity model. [*Shapiro and Ritzwoller*, 2002] in a first step estimate 2-D dispersion maps with a linear tomographic inversion of path average fundamental mode group and phase velocities, and afterward apply a Monte-Carlo method to perform the non-linear inversion of the dispersion curves at each geographical point and retrieve the 3-D shear-velocity model. [*Boschi and Ekström*, 2002] carry out a single non-linear inversion of phase anomaly measurements making use of JWKB ray-theory sensitivity kernels computed in a reference 3-D model. [*Zhou et al.*, 2006] invert long period fundamental mode phase delays with finite-frequency 3-D Born approximation kernels, calculated in a reference 1-D model. We will proceed with a 2 steps scheme, first inverting group path average speeds for a regionalized group velocity model assuming a linearized ray theory wave propagation. Then we will use the group velocity maps as data to perform an iterative linearized depth inversion for the local 1-D structure, accounting for lateral variations of the Crust.

This thesis presents the new model along with a discussion of the robustness and resolution of its main features. We will firstly present the group velocity measurement technique and an analysis of measurement errors (Chapter 2), then we will introduce the linear inversion of the regional data starting from a reference global model, with an accurate examination of the implication of different regularization constraints (Chapter 3). Group velocity maps will be then shown and discussed. Subsequently we will invert the group velocity for the V_S structure of Upper Mantle (Chapter 4). Our resulting 3-D radially anisotropic model will be discussed in detail and compared with other global and regional models.

2. Data analysis

2.1 Data set

We have analyzed long period seismograms relative to 98 medium to large size shallow events (depth$\leq$ 50.0 km and magnitude $M_s \geq 4.8$) occurred between 1998 and 2005 in the geographical frame 10N-80N 35W-80E. Earthquakes have been recorded by 90 stations placed in the same region (see Figure 2.1) and belonging to networks IRIS/IDA, IRIS/USGS, MEDNET, GEOSCOPE, GEOFON and the temporary Passcal deployment RETREAT in northern Apennines (Italy). The choice of this particular geographical frame is due to the effort of maximizing the regional path coverage, including the events from Mid Atlantic Ridge at the Western edge of the studied region, and the high seismicity zone of Middle East and Asia Minor to the East. Seismic networks are densely distributed in Europe but are very sparse on the African continent. In addition, its scarce seismicity makes North Africa poorly sampled. This results in an uneven data coverage, which restricts the region where our results are well constrained to the Mediter-

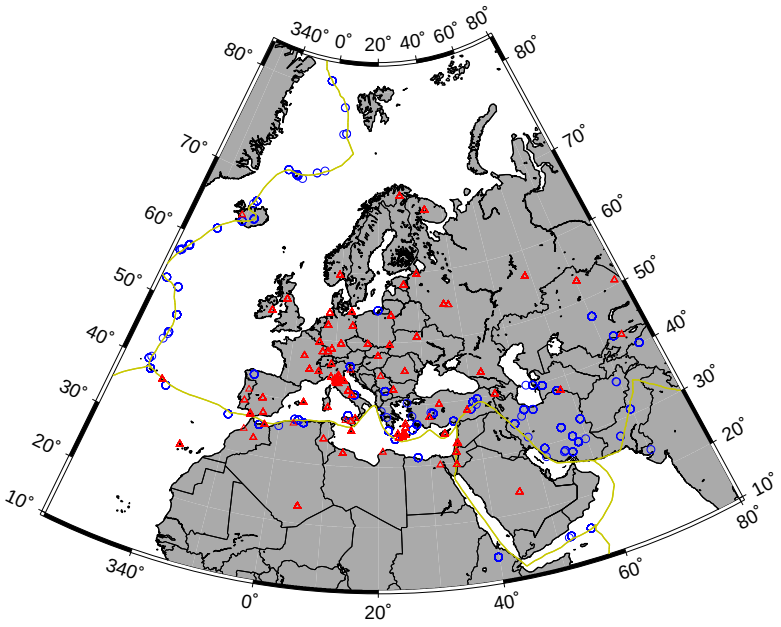


FIGURE 2.1: *Earthquakes (blue circles) and station (red triangles) distribution.*

ranean and European area. More than 3000 three components seismograms have been downloaded from the FSDN data center at IRIS DMC and from GEOFON. Each seismogram has been pre-processed with SAC to remove its mean, linear trend and to taper it, and visually inspected to detect potential anomalies such as calibration pulses, gaps, spiky signals, non-seismic signals, etc.. (see [*Laske*, 2002] for an exposition of seismic processing complications), which would prevent the seismogram from being analyzed. Only records with satisfactory signal to noise ratio has been kept for the subsequent dispersion analysis. Horizontal North and East components have been rotated to the Great Circle Path, in order to have them as Radial and Transverse components. After this first pre-processing step we had 1689 vertical and 1351 transverse traces.

2.2 Group velocity measurements

Unlike body waves, which phases can be picked on seismograms at a definite instant of time, surface waves have a longer duration in time given by their dispersion, i.e. by the different arrival times of the different frequencies packets. Therefore, surface waves characteristics are better studied with methods that allow a representation of the signal simultaneously in time and frequency domain, such as the Frequency-Time analysis described by [*Levshin et al.*, 1989, 1992]. This method permit to separate the different frequencies of the signal in accordance with their dispersive characteristics. We make use of an equivalent method implemented in the "Computer Programs in Seismology" routines package by R. Herrmann. We measure fundamental mode surface wave group speed through iterative application of Multiple Filters (MF) and Phase Matched Filters (PMF). These methods permit to isolate the fundamental mode (the most energetic) of surface waves from other arrivals and to compute its dispersion characteristics. We apply these filters on the vertical and the transverse component of the seismogram to study respectively Rayleigh and Love dispersion characteristics. Multiple Filters (firstly introduced by [*Landisman et al.*, 1969] and [*Dziewonski et. al.*, 1969]) are a set of Gaussian narrow band filters centered around different frequencies ω_i. We apply each filter $H(\omega - \omega_i)$ to the Fourier transformed seismic signal, and we successively transform back the outputs in time domain. In this way, we isolate the frequencies around ω_i, that have approximately the same group arrival time $\tau(\omega)$ corresponding to the maximum of the filtered signal. We can display side by side all the output signal envelopes in function of their central frequency and pick the arrival time of their maximum amplitude, obtaining a Frequency-Time (FT) graph of the signal spectral

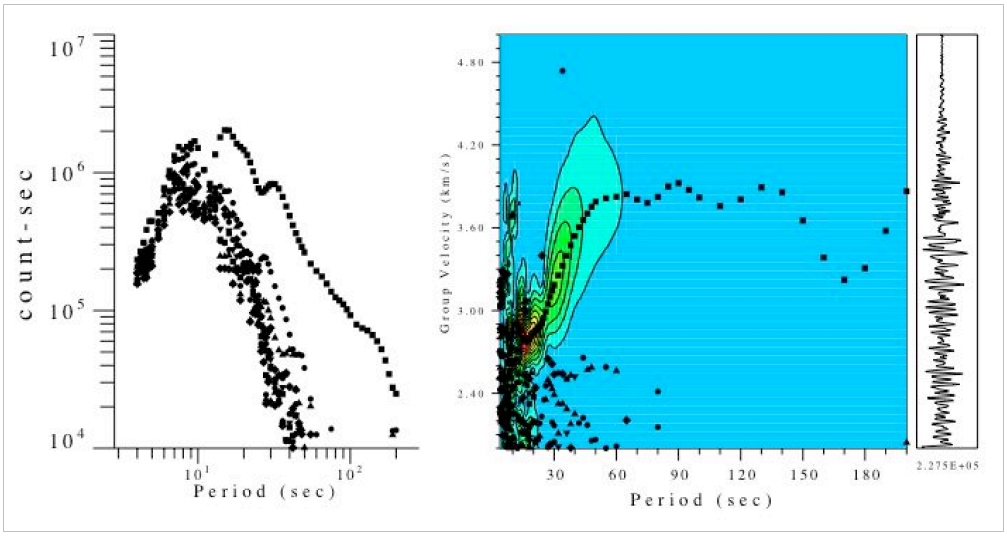


FIGURE 2.2: *Graphic interactive window of the CPS multiple filters analysis tool.* Left: *The chart shows on the x axis the central frequencies (expressed as periods) of the Gaussian filters in a logarithmic scale, on the y axis the (spectral) amplitude of the output signals in counts???. Dots are the relative maxima on the output signals.* Right: *On x axis are present the periods of the filters in a linear scale, on y axis the computed average group velocity for each period. Color scale gives the spectral amplitude of the signal, while dots indicate the peaks of maximum energy.*

amplitude. We can express the group arrival time $\tau(\omega_i)$ in terms of the average group velocity $U(\omega_i) = \Delta/\tau(\omega_i)$, where Δ is the ray length, in order to have a graph of group velocity dispersion in function of frequency (or equivalently period). In Figure 2.2 is shown the graphic interface of the MF routine. On the left the relative maxima of energy of each output signals are displayed in function of the filter central period. On the right, the spectral amplitude is given by different colors and relative maxima of energy are signed by dots, while y axis gives their group velocity. Thus, the FT graph gives a first estimate of the dispersion curve, that can be interactively picked to select the most energetic mode. The FT graph displays different ridges of maximum amplitude, corresponding to different arrivals. The ridge of absolute maximum energy generally corresponds to the fundamental mode, while other phases of minor amplitude can be related to higher modes or to the effects of crustal heterogeneities (multipathing, etc..). In order to isolate the fundamental mode from the rest of the signal we apply a Phase Matched Filter [*Herrin and Goforth*, 1977], a type of linear filter in which the Fourier phase of the filter is made equal to that of the seismic signal. The phase spectrum $\Psi(\omega)$

of the surface wave fundamental mode is related to the group delay time $\tau(\omega)$

$$\begin{aligned} K(\omega) &= |K(\omega)| \exp[-i\Psi(\omega)] \\ \frac{\mathrm{d}\Psi(\omega)}{\mathrm{d}\omega} &= \tau(\omega) \end{aligned} \tag{2.1}$$

Estimating the group velocity curve on the FT graph within the frequency range (ω_0, ω_1) where the signal is well observed, it is possible to build a phase correction

$$\Psi'(\omega) = \left(\int_{\omega_0}^{\omega_1} \tau(\omega)\mathrm{d}\omega \right) + c_1\omega = \left(\Delta \int_{\omega_0}^{\omega_1} \frac{1}{U(\omega)}\mathrm{d}\omega \right) + c_1\omega \tag{2.2}$$

to be introduced into the spectrum $K(\omega)$ by multiplying it by the factor $\exp[-i\Psi'(\omega)]$. $U(\omega)$ is obtained by a spline interpolation of the discrete values selected on the FT graph and the constant c_1 is arbitrary chosen. The inverse Fourier transform of the phase matched spectrum will be a signal compressed in time, ideally a delta function. In other words, all the energy traveling with the different speeds given by the dispersion curve $U(\omega)$ should arrive on the same instant of time, thus the signal to noise ratio will be maximized. All the non-dispersed energy (overtones, reflections, coda, other arrivals,...) is rejected as noise with a tapering time filter around the signal impulse, and the filtered signal is dispersed again to obtain the presumable single mode trace. Afterwards, the MF can be reapplied to the single mode trace, and a better estimate of the dispersion curve can be obtained. This two step procedure can be iterated to refine the dispersion curve until a stable condition is reached (usually after 2-3 times) and the final dispersion curve is saved. In Figure 2.3 we show a raw vertical seismogram and below the two iteratively filtered traces where only the fundamental mode is present. Below the same procedure is applied to the transverse component, and finally the dispersion curves for Rayleigh and Love wave are saved. We have iteratively applied MF and PMF to 1689 vertical and 1351 transverse components seismograms, and eventually we have obtained 1549 paths for Rayleigh and 906 paths for Love wave. In Figure 2.4 we show the final ray coverage for Rayleigh and Love 50 s waves.

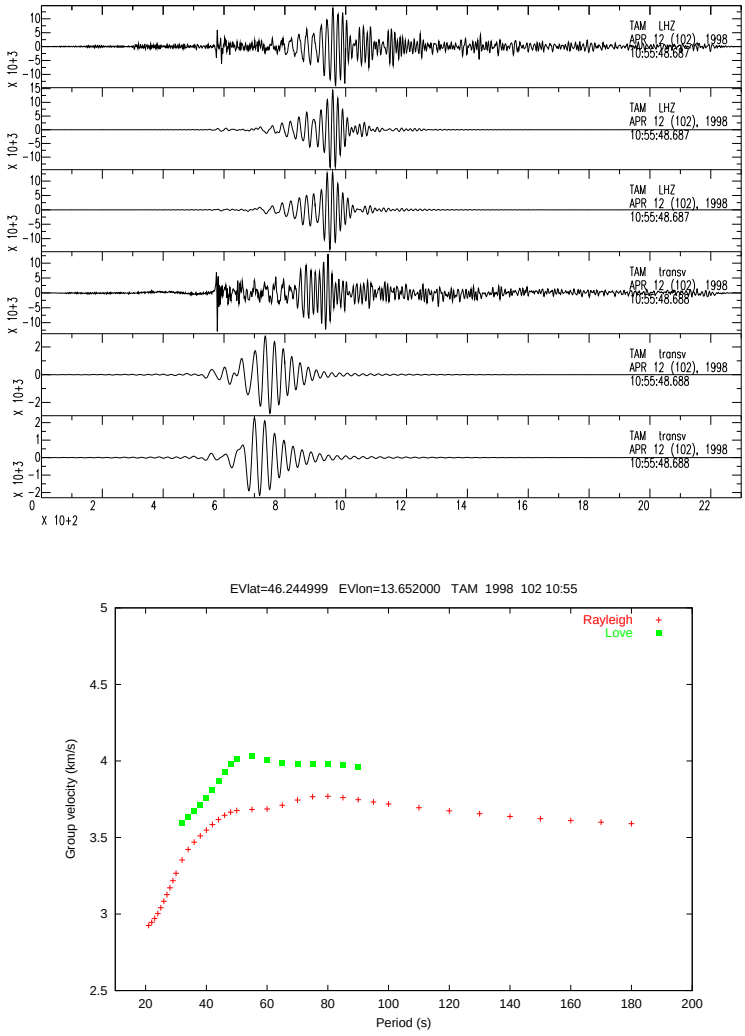


FIGURE 2.3: *An example of raw and two successively MF and PMF filtered seismograms (above) of vertical and transversal component and final dispersion curves (below) of fundamental mode Rayleigh and Love wave for an earthquake in Northern Italy recorded by II station TAM (Algeria).*

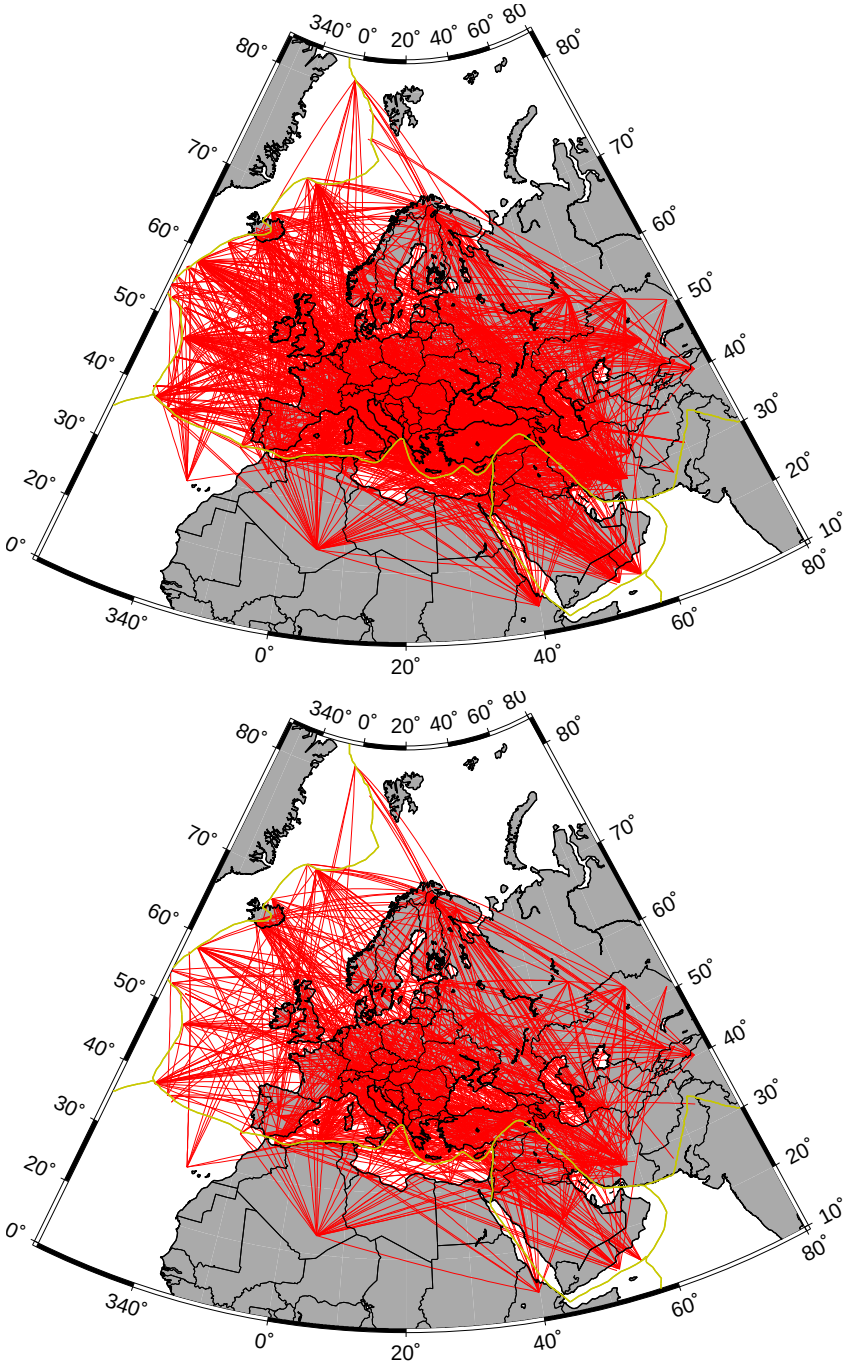


FIGURE 2.4: *Rayleigh* (top) *and Love* (bottom) *ray path coverage for the period* $T = 50$ *s.*

2.3 Dataset statistics

Data coverage is not only spatially uneven, but also period dependent, as it is shown in Figure 2.5 for the periods at which measurements are available. This is due to the difficulty in identifying the fundamental mode at short periods, where the signal is often masked by strong crustal heterogeneities effects, such as multipathing and scattering, whereas at long periods ($> 100s$) the signal amplitude decreases with respect to the noise level, making the measure unstable, especially for moderate size events which aren't able to excite the lower frequencies. Love wave measures are much less than Rayleigh owing to their higher noise level. Horizontal components of a seismogram are actually often noisier with respect to vertical one, because of their sensitivity to tilt. The difference in data distribution are shown in Figure 2.6 for 35, 50 and 100 s where measured velocities are compared to PREM values. Love waves have a wider distribution than Rayleigh wave and the same holds for short (35s) periods with respect to longer ones (50 and 100 s). This is due to the increased sensitivity of shorter periods and Love waves to crustal structures, which are much more heterogeneous with respect to Upper Mantle. The distributions of T=35 s have a strongly negative average because group velocities in this sudy are mainly measured on continental paths, hence at 35 s they are mostly sensitive to slower continental crust, whereas PREM model has a thinner crustal structure, with a 3 Km water layer on top, which results in a faster group speed.

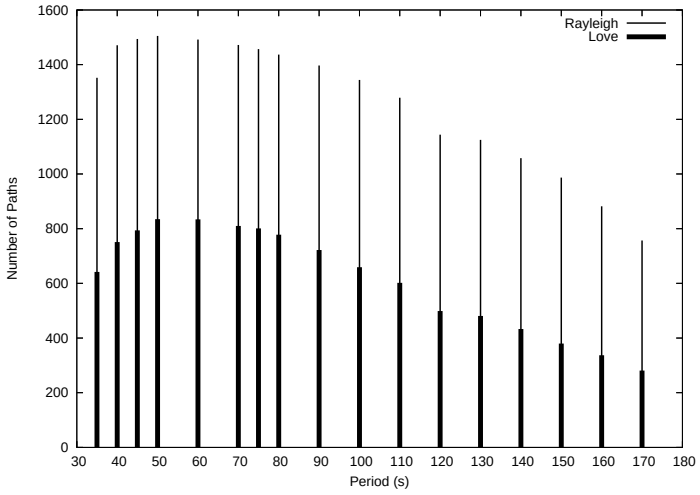


FIGURE 2.5: *Number of paths for each period for Rayleigh and Love waves*

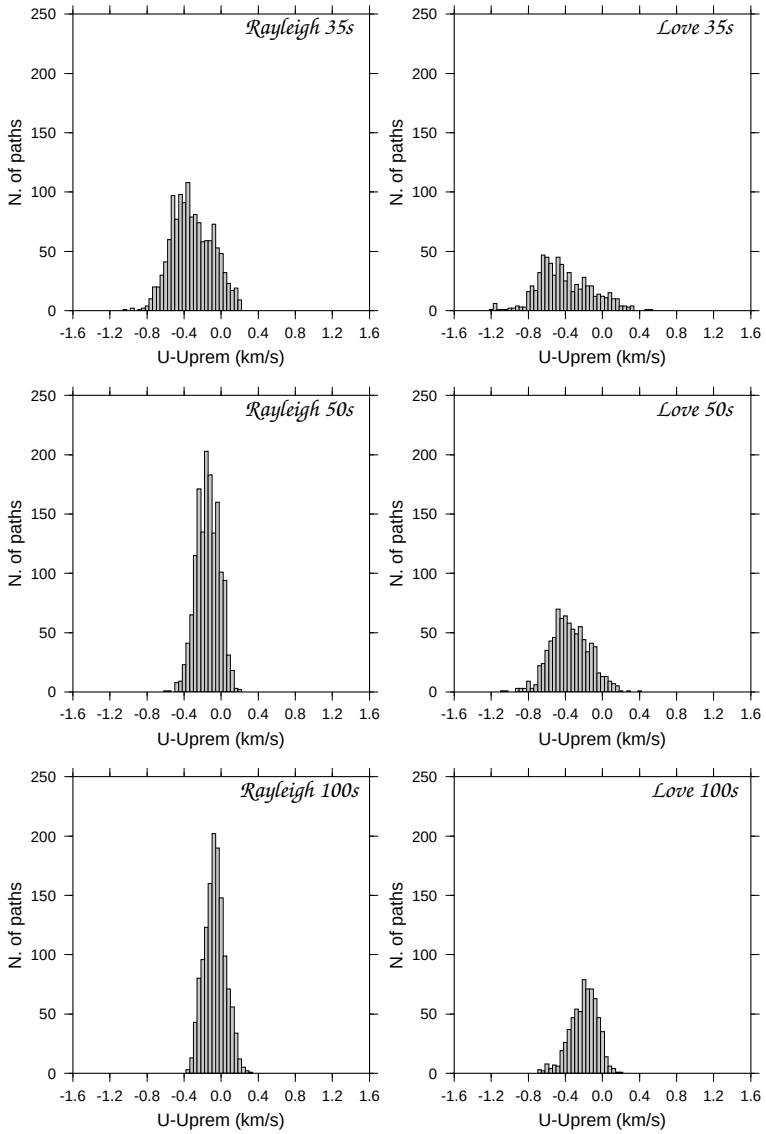


FIGURE 2.6: *Histograms showing difference between observed average group velocity along path and PREM group velocity for Rayleigh and Love waves at 35 s, 50 s and 100 s.*

2.4 Data errors

Data errors are always difficult to estimate, especially when measurements derive from rather complex analysis including filtering and operator interaction. Moreover, data errors are not due only to measurements uncertainties, but include also a wide range of different systematic theoretical sources, such as those caused by event mislocations, azimuthal anisotropy, off-great circle path propagation due to lateral scattering and refraction, and group time shifts due to the initial phase at the source. [*Ritzwoller and Levshin*, 1998] showed that of these systematic error sources, the most influent is probably azimuthal anisotropy, whose effect is particularly severe in regions of poor azimuthal coverage. We will not perform here any experiment to estimate the amplitude of the bias introduced in the solution by these effects, but we will keep them in mind when we will analyze, after the tomographic inversion, the misfit between the data predicted by our group velocity maps and the observed data. The net effect of these errors will actually be that of increase the part of data that our model is not able to explain.

2.4.1 Error in time or in velocity

It is not trivial to define what is the physical quantity that we measure when we perform the frequency time analysis. The procedure output is the group speed averaged along the path and the operator, working on a period vs. velocity plot, can manually select between different group velocities relative to different wave packets. Thus, the measurement error seems to be an error on velocity. However, what is actually measured with the the MF technique is the time delay of a frequency packet, which should lead to an error on time. Moreover, when we perform the inversion it is convenient to express the data as total traveltimes, retrieved dividing the path length by the average speed. To clear this ambiguity we analyze the residuals distribution after the inversion. Residuals are that part of the data that are not explained by the solution. Therefore, they are an indicator of the level of error present in the observed data. We plot in Figure 2.7 the absolute value of the residual of each ray versus the path length for 35, 50, 100 and 150 s: in the left graphs the residuals are expressed in velocity (km/s), while on the right ones they are in time (s). Although there is some difference between periods, we find that velocity residuals decrease with distance (on average), while time residuals do not show such a dependence. If short rays are more affected by measurement errors, we conclude that the measured quantity is time, and therefore we will calculate and express the errors in seconds in the rest of the thesis.

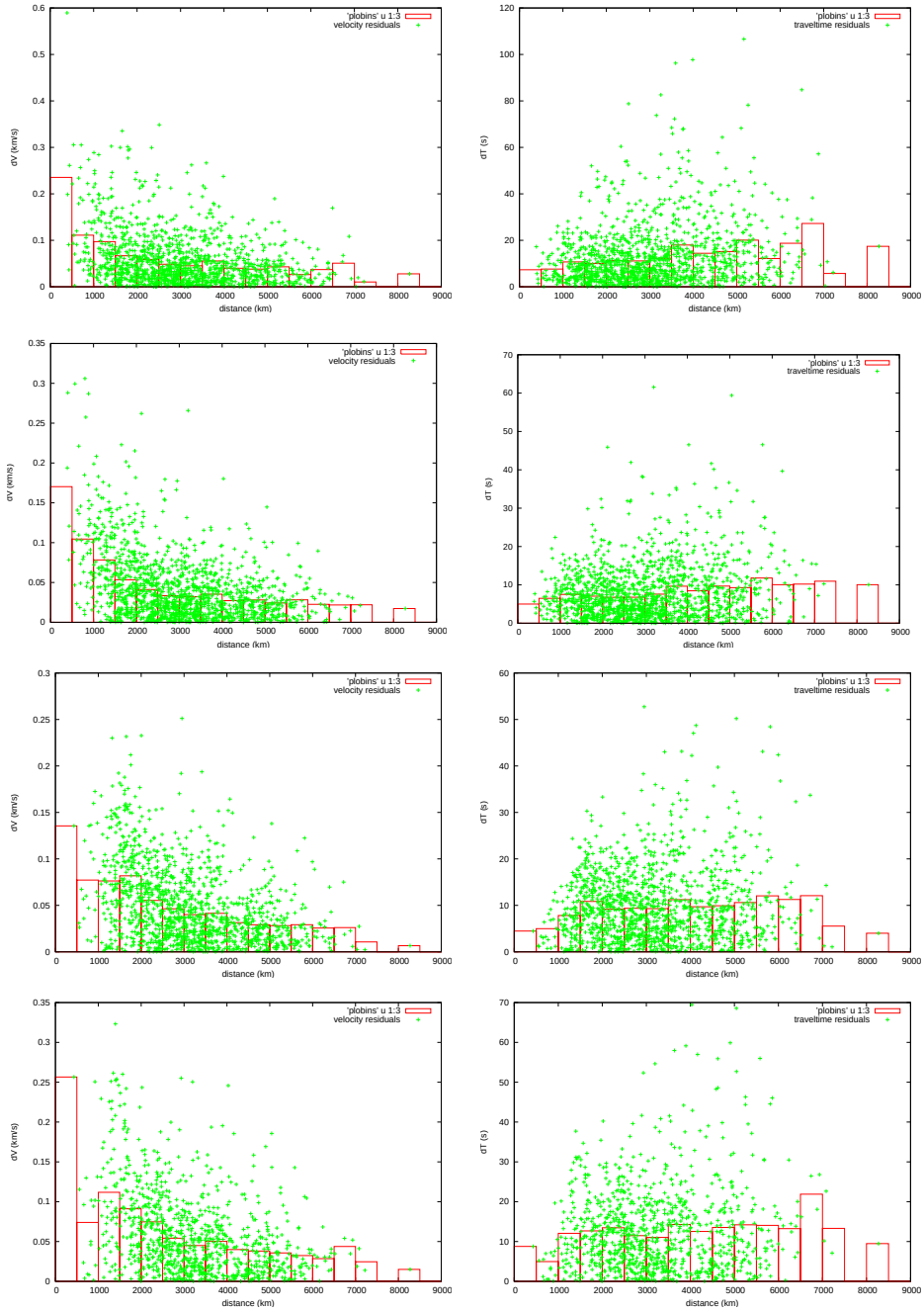


FIGURE 2.7: *Velocity (*Up*) and traveltime (*Down*) residuals for each ray in function of path length for Rayleigh wave at periods 35 s ,50 s, 100 s and 150 s. Green crosses are the singular values, red boxes are the values averaged on all the rays in each length range*

Let $\Delta t_i = t_i^{obs} - t_i^{theor}$ be the difference between the observed traveltime and that predicted by the model for $i-th$ ray, and let $\Delta\bar{t}$ be the average residual for all rays. We

define the dat error as the covariance of the residuals distribution.

$$\sigma_d^2 = \frac{\sum_{i=1}^{N}(\Delta t_i - \Delta\bar{t})^2}{N} \tag{2.3}$$

This *a posteriori* evaluation of the error is a rather conservative choice, but we feel that in this way we can account for all the theoretical errors precedently mentioned.

2.4.2 Measurement errors

The inclusion of RETREAT stations in our dataset allow a direct estimation of measurements uncertainties through the analysis of several cluster of rays directed to these stations. Temporary experiment RETREAT stations (showed on the right in Figure 2.8) have been operating from 2003 to 2006 in a relatively small area in Northern Apennines, and they were up to 200 km apart one from each other, thus below our expected resolution. We can estimate the maximum ideal resolution by a simple geometrical consideration: our shortest period waves have a wavelength of $\lambda \simeq 35s \cdot 3.5km/s = 122.5km$. We can not expect to discriminate structures smaller than a couple of wavelengths, i.e. 250 km. Despite a complicate local crustal structure, we expect all the measurements of the same event to be reasonably similar among the different RETREAT stations, and we consider the differences as a reasonable estimate of the measurement uncertainty.

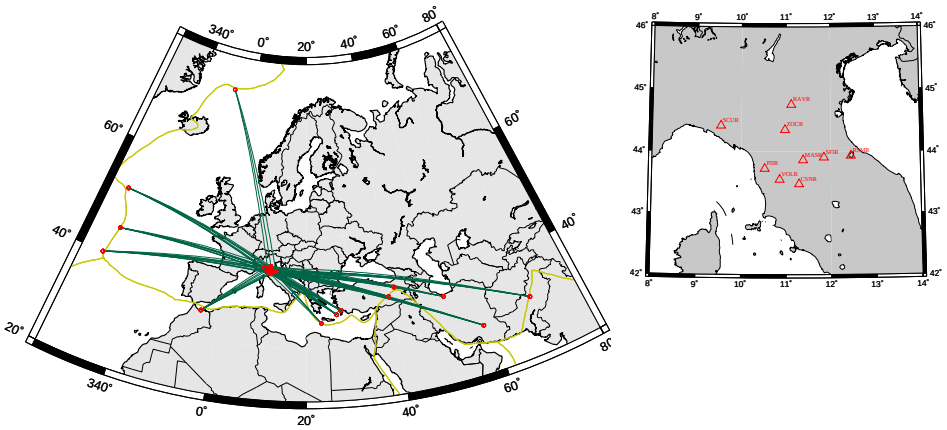


FIGURE 2.8: *RETREAT stations (triangles in the rigth panel) and ray clusters (left) used for measurements error analysis*

TABLE 2.1: *Events used for measurement error evaluation on RETREAT station records*

RETREAT EVENTS						
DATE	HOUR	LAT	LON	Ms	$\bar{\Delta}$(km)	N° of rays
2003/11/21	04:09:09.52	45.219	-28.035	5.2	3077.8	9
2003/12/23	14:02:04.13	40.135	-29.692	5.4	3367.6	9
2003/12/26	01:56:52.44	28.995	58.311	6.8	4456.6	9
2004/01/24	13:01:45.70	52.122	-30.179	5.4	3156.7	8
2004/02/07	21:17:24.20	36.040	26.910	5.1	1612.1	8
2004/02/24	02:27:46.23	35.142	-3.997	6.4	1623.9	8
2004/03/17	05:21:00.80	34.589	23.326	5.7	1469.5	6
2004/03/25	19:30:49.04	39.930	40.812	5.4	2528.9	4
2004/03/28	03:51:10.05	39.847	40.874	5.3	2445.6	5
2004/04/14	23:07:39.94	71.067	-7.747	5.6	3176.1	3
2004/05/28	12:38:44.27	36.249	51.622	6.4	3547.9	7
2004/08/04	03:01:07.57	36.833	27.815	5.2	1621.9	9
2004/08/11	15:48:26.82	38.377	39.261	5.5	2430.1	9
2004/08/13	10:48:38.93	30.922	69.769	5.1	5256.0	8

We consider 14 clusters of rays directed to 3 to 9 RETREAT stations relative to the 14 events listed in Table 2.1 and showed on the left in Figure 2.8. An example relative to the event of 11 August 2004 is shown in Figure 2.9. In the top panel are represented the 9 filtered vertical traces, and beneath are shown the corresponding group velocity dispersion curves. For each cluster we compute the average distance $\bar{\Delta}$ traveled by the rays and for each period we calculate the standard deviation of the travel times of the cluster. We estimate each traveltime as $t_i = \bar{\Delta}/v_i$ where v_i is the measured group velocity. We use the average distance instead of the real distance travelled by each ray in order to avoid traveltime discrepancies due to different path length in the same cluster. In Figure 2.10 we plot in blue the 14 traveltime standard deviations, and in green their average weighted for the number of rays in each cluster. This average values are a reliable estimate of the measurement errors in our data set, and provide an evaluation of the repeatability of our measurements. In red we show the data error calculated in 2.3. As previously remarked, this error estimate considers all the possible

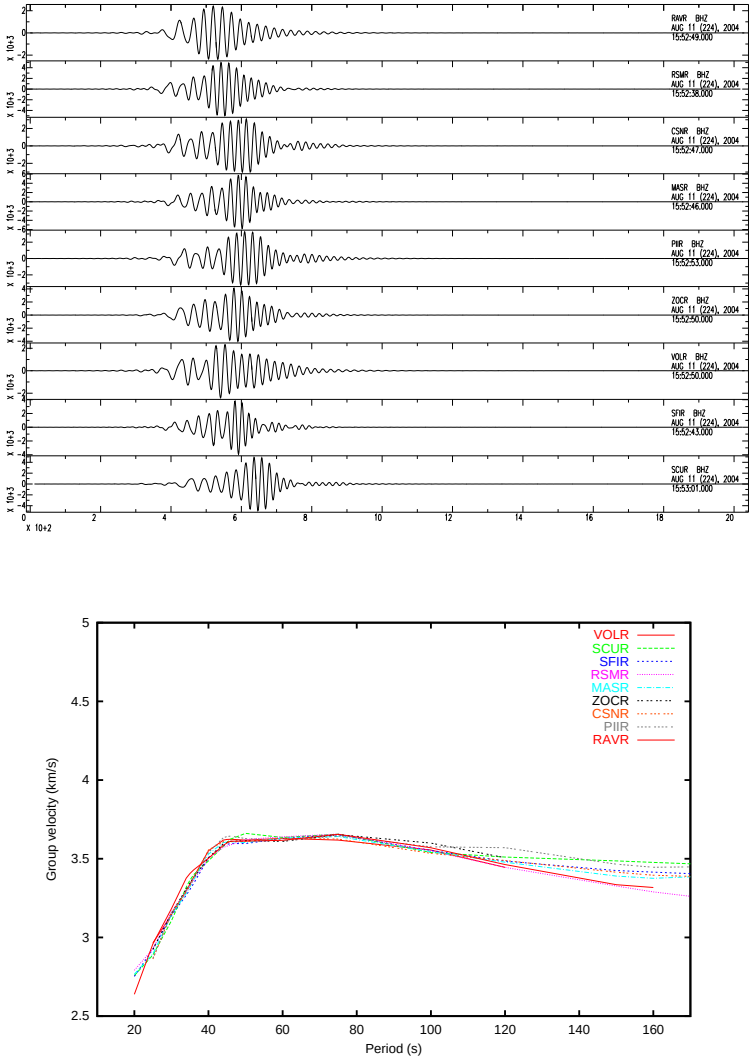


FIGURE 2.9: Top: *Filtered seismograms relative to the 11 August 2004 Turkey event recorded by 9 RETREAT stations.* Bottom: *Dispersion curves obtained from the frequency-time analysis of the traces.*

errors deriving from the theoretical approximations used in the inverse problem, such as great circle path propagation, no azimuthal anisotropy, no source group time shifts, etc... Hence, σ_d is always greater than the RETREAT average errors in green which are measurement errors. Nevertheless, σ_d remains still comparable with the clusters singular errors (the blue dots), meaning that this a posteriori error estimate is legitimate.

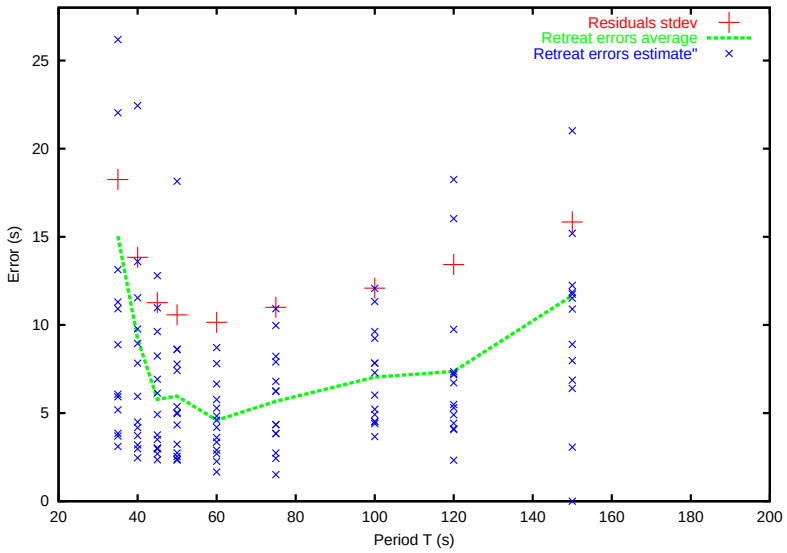


FIGURE 2.10: *Errors computed from Retreat data compared with standard deviation of the residual distribution after inversion (error a posteriori) errors in velocity, distance on y-axis*

2.5 Global dataset

A common limitation of regional tomographic models is their being poorly constrained near the borders, because of lack of data or scarce path coverage. This is due to the choice of stations and epicenters all inside the study area, to avoid the inclusion of any external anomaly inside the model. To improve on this situation and also to warrant consistency with global models, we used the global dataset of phase velocities measurements by [*Ekström et al.*, 1997] (ETL97) to build a global group velocity model. We will later invert our regional higher resolution measurements keeping the global model as an *a priori* reference model, in order to constrain the inversion with a more reliable laterally varying model. ETL97 global dataset consist of more than 50,000 high quality fundamental mode Rayleigh and Love phase dispersion curves. Measurements are the result of the application of an automatic algorithm to records from the Global Seismographic Network, using events in the Harvard centroid-moment tensor catalog. ETL97 dataset is in the form of a discrete sampling of each apparent average phase curve at a set of fixed periods (35s, 37s, 40s, 45s, 50s, 60s, 75s, 100s and 150s, but not all of the curves have the whole range). Group velocity $U(\omega)$ can be easily derived from phase

velocity $c(\omega)$ through

$$U(\omega) = \frac{c(\omega)}{1 - \frac{\omega}{c(\omega)} \frac{\mathrm{d}c(\omega)}{\mathrm{d}\omega}} \tag{2.4}$$

but correct numerical evaluation of frequency derivative of phase requires a comment: to compute $dc(\omega)/d\omega$ we need to interpolate the discrete values of $c(\omega_i)$ through splines functions (either B-splines or cubic splines). However, the phase is a rapidly decreasing function of ω, so it is preferable to spline the phase anomaly $\delta c(\omega)$ (difference between the apparent average observed and PREM phase)

$$\delta c(\omega) = c^{obs}(\omega) - c_0(\omega) \tag{2.5}$$

and to carry out the conversion to group velocity

$$\delta U + U_0 = \frac{c_0 + \delta c}{1 - \frac{\omega}{c_0 + \delta c} \left(\frac{\mathrm{d}c_0}{\mathrm{d}\omega} + \frac{\mathrm{d}\delta c}{\mathrm{d}\omega} \right)} \tag{2.6}$$

where $dc_0/d\omega$ is accurately computed from PREM normal modes expansion.
We tested the precision of this method to give the correct the group velocity by applying Multiple filters and Phase Matched filters on some of the seismograms and comparing the resulting dispersion curves with the ones given by phase to group transformation. In some cases the fit is satisfactory (see Figure 2.11 top), in some other cases we noticed some differences, that can be partially explained by the different source parameters: our group speed measurements uses epicenter location from the NEIC catalog, while the ETL97 dataset is based on CMT location, that can differ up to some kilometers, because CMT algorithm tends to mask out the local complex structure and to move the epicenter from its real position. As a result, the ray can be longer or shorter than the one estimated with NEIC location and the whole group velocity curve can be shifted in time (see Figure 2.11 down). However, this effect is not systematic (i.e. on average the rays are not shorter or longer), so the mean global group velocity is likely not to be affected by this phenomenon.

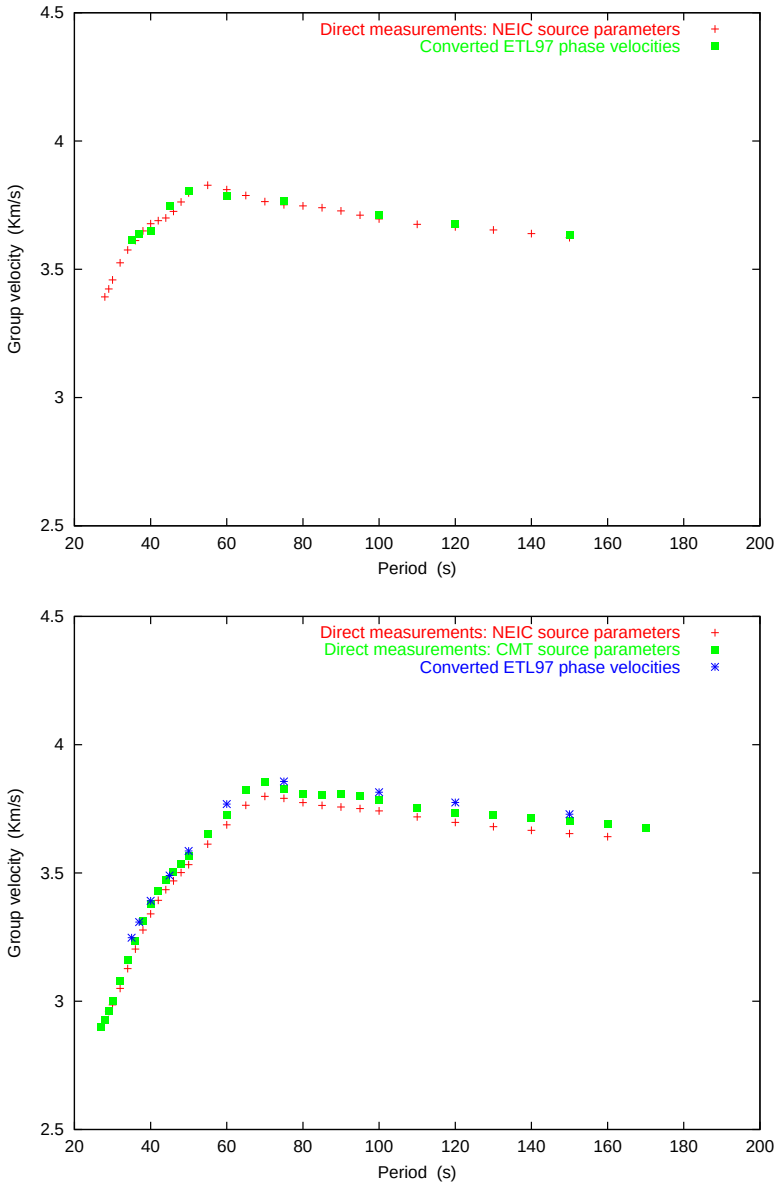


FIGURE 2.11: Top: *Dispersion curves given by direct measurements applying multiple filters and phase matched filters on the record (red) and group velocity converted from phase velocity (blue)* Bottom: *Dispersion curves given by direct group speed measurements (red) using NEIC source parameters, direct measurements using CMT source parameters (green) and group velocity derived from phase dispersion.*

3. Surface wave tomography

We give some insights in the Linear Inverse Problem theory, starting from a Least-Squares approach and afterwards including it in a more general probabilistic view. The tomographic inversion is divided into two steps: we first invert the ETL97 dataset to build a background reference global model, and then we proceed with the regionalization of our dispersion curves keeping the global model as an a priori model. We show our resulting group velocity maps and calculate the associated a posteriori errors. We treat carefully the implications of the regularization conditions applied to the inverse problem and finally we discuss and compare their main features with other works present in literature.

3.1 Linear inverse problems theory

In an inverse problem the aim is to retrieve the parameters of an unknown model $\mathbf{m}$ once we are given some experimental observations $\mathbf{d}^{obs}$, a theory that relates parameters and observations $\mathbf{d}^{obs} = \mathbf{g}(\mathbf{m})$ and some constraints on the possible solutions (a priori information, or information about the model which cannot be derived from data).

3.1.1 Least-Squares solution

In travel time seismic tomography the central point is the solution of the linear system

$$\mathbf{d}^{obs} = \mathbf{G}\,\mathbf{m} \tag{3.1}$$

where $\mathbf{d}^{obs}$ are travel time observations (either P-wave or other body wave phases arrivals, as well as surface waves group or phase delays) and $\mathbf{m}$ are model slowness or speeds. Owing to noise in seismic data this problem is generally inconsistent, and therefore it does not have an exact solution. The standard procedure is to find its least squares solution, which is the vector $\mathbf{m}$ such that $||\mathbf{d}^{obs} - \mathbf{G}\,\mathbf{m}|| = min$. It can be shown that

$$\mathbf{m} = (\mathbf{G}^T\,\mathbf{G})^{-1}\mathbf{G}^T\,\mathbf{d}^{obs}$$

However, due to inhomogeneity in ray coverage, most tomographic inverse problems are ill-conditioned, meaning that the matrix $(\mathbf{G}^T\,\mathbf{G})$ is singular, or very close to being singular. In this cases the solution can be better constrained adding further conditions

which reflect our a priori information about the model. We can modify the linear system (3.1) by adding, for example, two conditions on norm and roughenss

$$\begin{cases} \mathbf{d} - \mathbf{G}\,\mathbf{m} &= 0 \\ \gamma\,\mathbf{I}\,(\mathbf{m} - \mathbf{m}_0) &= 0 \\ \lambda\,\mathbf{H}\,\mathbf{m} &= 0 \end{cases} \tag{3.2}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{H}$ is a suitable smoothing operator (Gradient or Laplacian). The first Equation is the previous data misfit minimization. The second condition tends to keep the solution close to a reference model m_0, while the third acts on the model roughness minimizing the gradient (or the Laplacian) between adjacent model cells. λ and γ are two arbitrary trade-off parameters which balance the damping effects with the data fit requirement. Similarly to (3.1), a solution to (3.2) is only possible in a least-squares sense, and is given by

$$\mathbf{m} = (\mathbf{G}^T\,\mathbf{G} + \gamma^2\,\mathbf{I} + \lambda^2\,\mathbf{H}^T\,\mathbf{H})^{-1}\,(\mathbf{G}^T\,\mathbf{d}^{obs} + \gamma^2\,\mathbf{m}_0) \tag{3.3}$$

3.1.2 Statistical - Bayesian approach

We follow the approach of [*Tarantola*, 2005] based on the concept of "states of information". In this point of view, it is postulated that the more general way to describe such a state of information (the level of knowledge we have of a parameters set) is to define a "probability density" over the parameter space. Consequently, the results of measurements of the observable parameters (data), the a priori information on model parameters, and the information on the physical correlations between observable and model parameters (theory) can all be described using probability densities. The general inverse problem can then be developed combining all of this information. We restrict our attention to a very particular (and simple) case: the physical theory is linear and exact and all the probability density functions (p.d.f.) are Gaussian. We are therefore supposing that:

- the forward problem relation $d = g(m)$ is exact, i.e. the p.d.f. $\theta(d/m)$ to have the data set $\mathbf{d}$ from the model $\mathbf{m}$ is a Dirac's δ function. Thus, we neglect the modelling errors related to the theory.

  $$\theta(d/m) = \delta(d - g(m)) \tag{3.4}$$

  In addition, we suppose that the forward problem is linear

  $$\mathbf{d} = \mathbf{g}(\mathbf{m}) = \mathbf{G}\,\mathbf{m} \tag{3.5}$$

where **G** is a linear operator relating model parameters **m** to data **d**.

- The p.d.f. of the experimental information (data) is Gaussian. This means that the distribution $\nu(d_{obs}/d)$ (i.e. the probability to have the value d_{obs} as output from the experimental instrument if the real datum in input is d) is a Gaussian

$$\nu(d_{obs}/d) = K \cdot \exp\Big\{-\frac{1}{2}(d - d_{obs})^T C_d^{-1}(d - d_{obs})\Big\} \tag{3.6}$$

where K is a constant and C_d is the covariance operator which describes the experimental uncertainties. Making use of the Bayes theorem it is possible to define the a priori information on data as the conditional probability to have the datum d if the observed value is d_{obs}

$$\rho_D(d) = \frac{\nu(d_{obs}/d)\,\mu_D(d)}{\int_D \mathrm{d}\,d \quad \nu(d_{obs}/d)\,\mu_D(d)} \tag{3.7}$$

where $\mu_D(d)$ is an homogeneous p.d.f. for data. In the Gaussian hypothesis Equation (3.7) becomes

$$\rho_D(d) = \mu_D(d)\,K \cdot \exp\Big\{-\frac{1}{2}(d - d_{obs})^T C_d^{-1}(d - d_{obs})\Big\} \tag{3.8}$$

- The p.d.f. of the a priori information about the model is Gaussian

$$\rho_M(m) = K \cdot \exp\Big\{-\frac{1}{2}(m - m_0)^T C_M^{-1}(m - m_0)\Big\} \tag{3.9}$$

where C_M is the a priori model covariance operator, and m_0 is the mean of the distribution. Thus, from our a priori information, we expect the model to be gaussianly distributed around m_0 with a covariance given by C_M .

The p.d.f. of the solution shall be given by the posterior information on the model

$$\sigma_M(m) = \rho_M(m) \int \mathrm{d}\,d \quad \frac{\rho_D(d)\,\theta(d/m)}{\mu_D(d)} \tag{3.10}$$

Substituting Equations (3.4), (3.7) and (3.9) into Equation (3.10) we get

$$\begin{aligned}\sigma_M(m) = K \cdot \exp\Big\{-\frac{1}{2}\,\big[&(g(m) - d_{obs})^T C_d^{-1}(g(m) - d_{obs}) + \\ &+ (m - m_0)^T C_M^{-1}(m - m_0)\big]\Big\}\end{aligned} \tag{3.11}$$

which is also a Gaussian distribution if the forward problem is linear $d = g(m) = Gm$. The mean of the (3.11) is given by the following three equivalent expressions

$$\begin{cases} \langle m\rangle = (G^T\,C_d^{-1}\,G + C_M^{-1})^{-1}\,(G^T\,C_d^{-1} d_{obs} + C_M^{-1}\,m_0) \\ \langle m\rangle = m_0 + (G^T\,C_d^{-1}\,G + C_M^{-1})^{-1}\,G^T\,C_d^{-1}(d_{obs} - G\,m_0) \\ \langle m\rangle = m_0 + C_M\,G^T(G\,C_M\,G^T + C_d)^{-1}\,(d_{obs} - G\,m_0) \end{cases} \tag{3.12}$$

It is important to remark that, although we calculate "the solution" of the inverse problem with Equations (3.12), in fact this is just the *mean* model in the set of models that form the a posteriori probability distribution $\sigma(m)$. It should be more correct to show a random sample of models from this distribution, for example obtained by applying Monte-Carlo methods. Comparing the first of the (3.12) with the Least-Squares solution (3.3) it is evident that the two approaches are perfectly compatible, provided we express the a priori information (3.9) as a sum of two different conditions

$$\begin{aligned} \rho_M(m) &= K \cdot \exp\Big\{-\frac{1}{2}\big[(m-m_0)^T C_M^{-1}(m-m_0)\big]\Big\} = \\ &= K \cdot \exp\Big\{-\frac{1}{2}\big[(m-m_0)^T C_m^{-1}(m-m_0) + \\ &+ (Hm-0)^T C_L^{-1}(Hm-0)\big]\Big\} \end{aligned} \tag{3.13}$$

As C_m is the covariance operator which rules the condition on the closeness to m_0, C_L is the covariance operator which acts on the smoothing condition, stating how much the curvature (or the gradient) of the model should be close to zero. Thus, the solution becomes

$$\boxed{\langle m \rangle = (G^T C_d^{-1} G + C_m^{-1} + H^T C_L^{-1} H)^{-1} (G^T C_d^{-1} d + C_m^{-1} m_0)} \tag{3.14}$$

3.2 Parameterization

The surface of the Earth is parameterized into a global grid of squared cells of size 250x250 km^2 approximately. We project the 12 edges of the inscribed cube on the Earth surface (we neglect the Earth's ellipticity) as in figure 3.1. Each face projected on the sphere is then divided into 40x40 equal size cells, and we apply a trigonometric correction to minimize the cell stretching near the angles. We apply an Euler rotation to the reference cube in order to have one face centered on Eurasia an Mediterranean area. When we perform the regional inversion we consider only the face centered on Eurasia, divided into a finer grid (80x80 cells), which results in a 120x120 km^2 cell size.

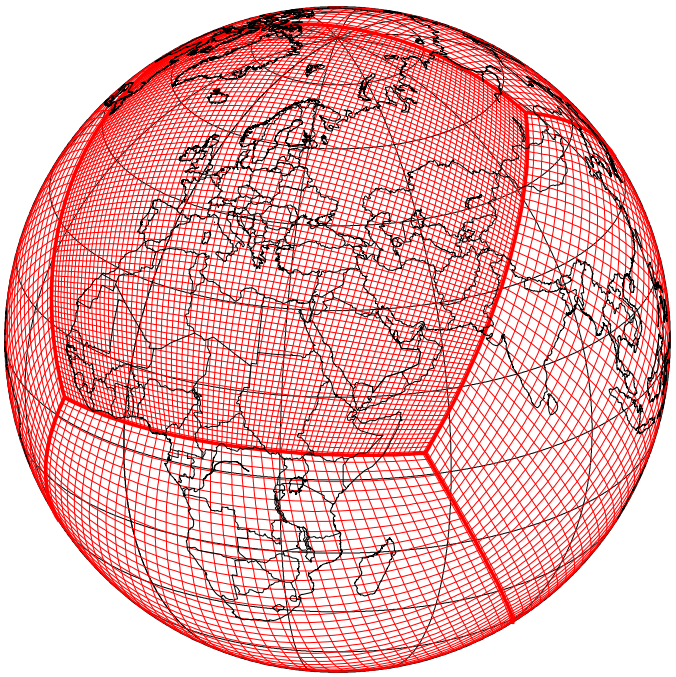


FIGURE 3.1: *Rotated cubed sphere, the Earth surface parameterization used in this study.*

3.3 Forward and inverse problem

We assume that seismic waves propagate along the great-circle arc that connects sources and receivers, (i.e. we assume a linearized ray theory, where lateral velocity anomalies are small enough so as to not deviate the ray from its path in a reference laterally homogeneous model) and we neglect second order effects as azimuthal anisotropy, source mechanism, scattering, finite frequency effects, multipathing, etc... We set up a separate linear inverse problem for each period at which group velocity measurements are available, and we solve it to obtain two-dimensional group velocity maps of the studied region. For each ray $i = 1, \ldots, N$ of length L_i we express the total travel time t_i as the sum of the traveltime t_0 (employed by the ray in a reference laterally homogeneous model of slowness $c_0 = 1/u_0$) plus a perturbation δt_i due to the local perturbation in slowness $\delta c(\mathbf{x}) = 1/\delta u(\mathbf{x})$. We approximate the perturbed traveltime with the discretized sum of all the partial times employed by the ray to cross each cell, considering the unknown group slowness perturbation δc_j constant inside each one of the $j = 1, \ldots, M$ cells which form the model.

$$t_i = \int_{L_i} c_i(\mathbf{x})\,\mathrm{d}l = \int_{L_i} \big(c_0 + \delta c_i(\mathbf{x})\big)\,\mathrm{d}l = \int_{L_i} c_0\,\mathrm{d}l + \int_{L_i} \delta c_i(\mathbf{x})\,\mathrm{d}l =$$
$$= t_0 + \int_{L_i} \delta c_i(\mathbf{x})\,\mathrm{d}l \;\simeq\; t_0 + \sum_{j=1}^{M} \frac{\partial t_i}{\partial c_j}\,\delta c_j = t_0 + \sum_{j=1}^{M} G_{ij}\; m_j$$

where the matrix element G_{ij} is the distance travelled by the *i-th* ray in the *j-th* cell of perturbed slowness m_j. The forward problem is then set as

$$t_i - t_0 = d_i = \sum_{j=1}^{M} G_{ij}\; m_j$$

Following the formulation by [*Tarantola and Valette*, 1982] we solve the inverse problem making use of the (3.14). If we associate the same error σ_d to all data (i.e. we do not weight data), the same a priori error σ_m to all model parameters and the same value σ_L to the laplacian of each model parameter, the three covariance matrices become a multiple of the identity matrix I, and we can simplify Equation (3.14) with

$$m = (G^T G + \frac{\sigma_d^2}{\sigma_m^2} I + H^T \frac{\sigma_d^2}{\sigma_L^2} H)^{-1} (G^T d + \frac{\sigma_d^2}{\sigma_m^2} m_0) \tag{3.15}$$

As a result, only two of the three covariance parameters are independent and the solution is influenced by the two ratios $K_m = \sigma_d^2/\sigma_m^2$ and $K_L = \sigma_d^2/\sigma_L^2$. As explained in the previous chapter, we evaluate σ_d as the standard deviation of the residuals distribution

after inversion, considered to be a conservative estimate of data errors. The choice of σ_m and σ_L is instead more subjective. We do not know of any procedure to derive quantitative constraints on the allowed variation from the a priori model, or on the allowed range of Laplacian or Gradient. These values are often a matter of qualitative evaluation. Hence we give to K_m and K_L two arbitrary values, we solve the inverse problem and then we compute the residuals Δd_i distribution along with its r.m.s. $\sigma_r.m.s.$ and standard deviation σ_d.

$$\begin{cases} \Delta d_i = d_i^{obs} - \sum_{j=1}^{M} G_{ij}\, m_j & \text{residual} \\ \langle \Delta d \rangle = \sum_{i=1}^{N} \dfrac{\Delta d_i}{N} & \text{mean} \\ T^2_{r.m.s.} = \sum_{i=1}^{N} \dfrac{\Delta d_i^2}{N} & \text{r.m.s.} \\ \sigma_d^2 = T^2_{r.m.s.} - \langle \Delta d \rangle^2 & \text{covariance} \end{cases} \tag{3.16}$$

Finally, σ_m e σ_L are retrieved simply by calculating $\sigma_m^2 = \sigma_d^2/K_m$ and $\sigma_L^2 = \sigma_d^2/K_L$. The implications of different choices of σ_m and σ_L will be exhaustively investigated in the next section.

3.3.1 Regularization constraints

The condition on model smoothness is imposed by minimizing either the squared norm of the gradient or the Laplacian of the solution $\mathbf{m}$. On Earth's surface they are defined as

$$\int_\Omega \|\boldsymbol{\nabla}_1\, \mathbf{m}(\theta,\phi)\|^2 \mathrm{d}\Omega \qquad \text{and} \qquad \int_\Omega \nabla_1^2\, \mathbf{m}(\theta,\phi)\mathrm{d}\Omega \tag{3.17}$$

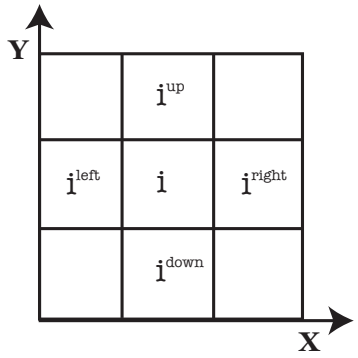
where θ is the co-latitude and ϕ is the longitude, with $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$ unit vectors tangent to Earth's surface. The surface gradient on the unit sphere is defined

$$\boldsymbol{\nabla}_1 = \hat{\boldsymbol{\theta}} \frac{\partial}{\partial \theta} + \frac{1}{\sin\theta} \hat{\boldsymbol{\phi}} \frac{\partial}{\partial \phi}$$

while the surface Laplacian is

$$\nabla_1^2 = \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \big(\sin\theta \frac{\partial}{\partial \theta}\big) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2}$$

In a cell model, we can approximate locally the cell parameterization with a cartesian plane, with the two axes $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ oriented to the East and North, respectively. Let's consider the i-th cell, with slowness m_i and its surrounding cells i^{left} , i^{right}, i^{up} and i^{down}.



Gradient minimization

We can approximate the Gradient of m_i with a finite-difference operator

$$\boldsymbol{\nabla} m_i = \left(\frac{\partial m_i}{\partial \mathbf{x}}, \frac{\partial m_i}{\partial \mathbf{y}}\right) \simeq \left(\frac{m_i^{right} - m_i}{x_i^{right} - x_i}, \frac{m_i^{up} - m_i}{y_i^{up} - y_i}\right) \tag{3.18}$$

Since the distance between adjacent cells is approximatively constant throughout all the M cells of the model , the Gradient in (3.18) can be expressed as a 2MxM linear operator acting on $\mathbf{m}$

$$\boldsymbol{\nabla}\mathbf{m} = k_g \begin{pmatrix} \mathbf{H_x} \\ \mathbf{H_y} \end{pmatrix} \mathbf{m} \tag{3.19}$$

where k_g is a constant depending on the distance between cells.

$$\mathbf{H_x}(i,j) = \begin{cases} -1 & \text{for } j = i \\ 1 & \text{for } j = i^{right} \\ 0 & \text{elsewhere} \end{cases} \qquad \mathbf{H_y}(i,j) = \begin{cases} -1 & \text{for } j = i \\ 1 & \text{for } j = i^{up} \\ 0 & \text{elsewhere} \end{cases} \tag{3.20}$$

The squared norm of the Gradient is

$$\|\boldsymbol{\nabla}\mathbf{m}\|^2 = \mathbf{m}^T \left(\mathbf{H_x}^T | \mathbf{H_y}^T\right) \begin{pmatrix} \mathbf{H_x} \\ \mathbf{H_y} \end{pmatrix} \mathbf{m} = \mathbf{m}^T \left[\left(\mathbf{H_x}^T \mathbf{H_x} + \mathbf{H_y}^T \mathbf{H_y}\right)\right] \mathbf{m} \tag{3.21}$$

Substituting the previous condition into the (3.15) we obtain the solution for the gradient minimization

$$\mathbf{m} = \left(\mathbf{G}^T\mathbf{G} + \frac{\sigma_d^2}{\sigma_m^2}\mathbf{I} + \frac{\sigma_d^2}{\sigma_{grad}^2}(\mathbf{H_x}^T\mathbf{H_x} + \mathbf{H_y}^T\mathbf{H_y})\right)^{-1} \left(\mathbf{G}^T\,\mathbf{d} + \frac{\sigma_d^2}{\sigma_m^2}\mathbf{m}_0\right) \tag{3.22}$$

Laplacian minimization

We can approximate the Laplacian of m_i with a finite-difference operator

$$\nabla^2 m_i = \frac{\partial^2 m_i}{\partial x^2} + \frac{\partial^2 m_i}{\partial y^2} \simeq \frac{m_i^{right} + m_i^{up} + m_i^{left} + m_i^{down} - 4m_i}{4\Delta x} \tag{3.23}$$

As before, we write $\nabla^2\mathbf{m}$ as a MxM linear operator

$$\nabla^2\mathbf{m} = \mathbf{H}_L\mathbf{m} \qquad \text{with} \qquad \mathbf{H}_L(i,j) = \begin{cases} -4 & \text{for } j = i \\ 1 & \text{for } j = i^{right} \\ 1 & \text{for } j = i^{left} \\ 1 & \text{for } j = i^{up} \\ 1 & \text{for } j = i^{down} \\ 0 & \text{elsewhere} \end{cases}$$

and we substitute its squared norm in (3.15)

$$\mathbf{m} = \left(\mathbf{G}^T\mathbf{G} + \frac{\sigma_d^2}{\sigma_m^2}\mathbf{I} + \frac{\sigma_d^2}{\sigma_{lapl}^2}\mathbf{H}_L^T\,\mathbf{H}_L\right)^{-1}\left(\mathbf{G}^T\,\mathbf{d} + \frac{\sigma_d^2}{\sigma_m^2}\mathbf{m}_0\right) \tag{3.24}$$

3.4 Global inversion

We invert the ETL97 global group velocity dataset on the global grid of cells of size 250x250 km^2 following the least square approach explained in the previous section. Since at this stage we are building our reference model, we do not have any particular expectation on the norm of the solution, and therefore we neglect any a priori information $\mathbf{m}_0$: we simplify Equation (3.15) posing $\sigma_m = \infty$ and we consider only a smoothing condition on the Laplacian (minimum curvature). After a number of tests we chose $\sigma_L = 1.04\ s/km^3$. We performed 11 separated inversions for Rayleigh and 11 for Love wave at the periods (35 s, 40 s, 45 s, 50 s, 60 s, 75 s, 100 s, 120 s, 130 s, 140 s and 150s). We show the resulting global dispersion maps at three selected periods in Figure 3.2 in terms of percent deviations of group velocity from the Prem value. The resulting maps are perfectly compatible with those of [*Larson and Ekström*, 2001], obtained from the same dataset but with a slightly different inversion procedure and parameterization (gradient minimization).

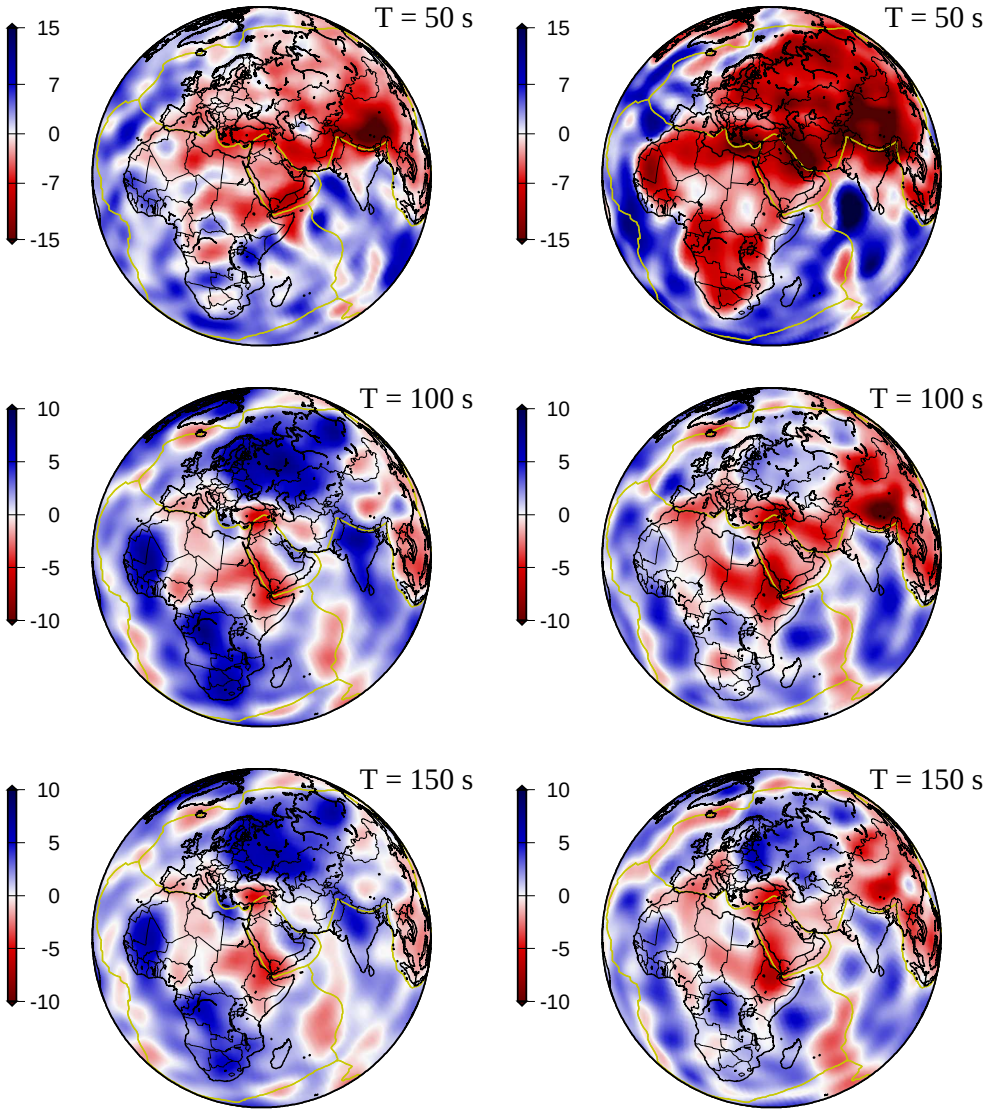


FIGURE 3.2: Left: *Global group velocity anomalies derived from the inversion of the ETL97 phase velocity dataset for Rayleigh wave at three selected periods. Color scale is percent variation of group velocity with respect to PREM value.* Right: *Global group velocity anomalies for Love wave. Same color scale as for Rayleigh*

3.5 Regional inversion

We invert the regional dataset on the finer grid centered on Eurasia with cells of 120x120 km^2 (see figure 3.1). We apply Equation (3.15) separately for each period (35 s, 40 s, 45 s, 50 s, 60 s, 75 s, 100 s, 120 s , 130 s, 140 s, 150 s, 160 s and 170 s) and each

wave type. As m_0 we use the slowness of the global model (each large cell of the global model is divided into 4 cells with the same slowness in the regional grid). For the 160 s and 170 s data inversion we use the global model at 150 s as reference model, since there is little difference between the resulting models at these periods. We tested different values for the two ratios σ_d^2/σ_m^2 e σ_d^2/σ_L^2 and we performed inversions both with the Gradient and the Laplacian minimization. We will exhaustively analyze the implications of different regularization constraints in the next section. However, we retaine as our preferred model the solution obtained through the Laplacian minimization with $\sigma_L = 4.59\ s/km^3$ and $\sigma_m = 0.012\ s/km$ (which corresponds to a covariance in group velocity of about 4.5%). In Figure 3.3 we show a comparison between the global model for Rayleigh waves at T=100s and the regional solution in Eurasia and Mediterranean (bounded by the green square) superimposed on the the global model: the inclusion of our shorter regional paths show finer details missing in the the global model. Anomalies tend to be stronger, mostly in slow areas. This could be ascribed to the slow tectonic region of Western Europe and Tethys collision zone along the Alpine-Himalayan belt, sampled by most regional rays, whereas the global dataset is more equally distributed between slow and fast ares of the Earth. As a result the average of the regional model

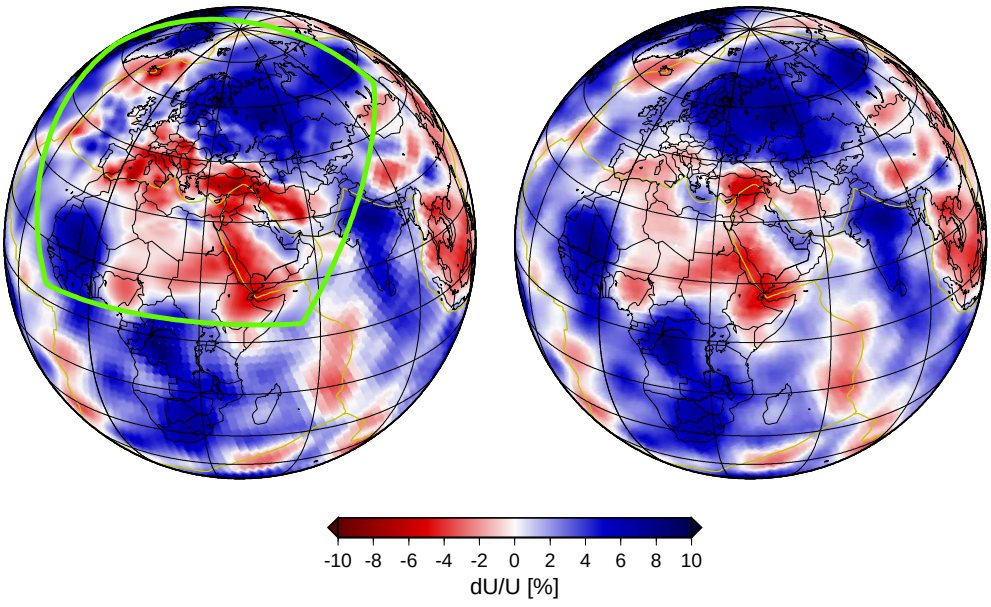


FIGURE 3.3: *Regional (left) and Global (right) 100 s Rayleigh group velocity map. The regional model (inside the green square) is superimposed on the global one. Color scale represents relative group velocity anomaly with respect to PREM*

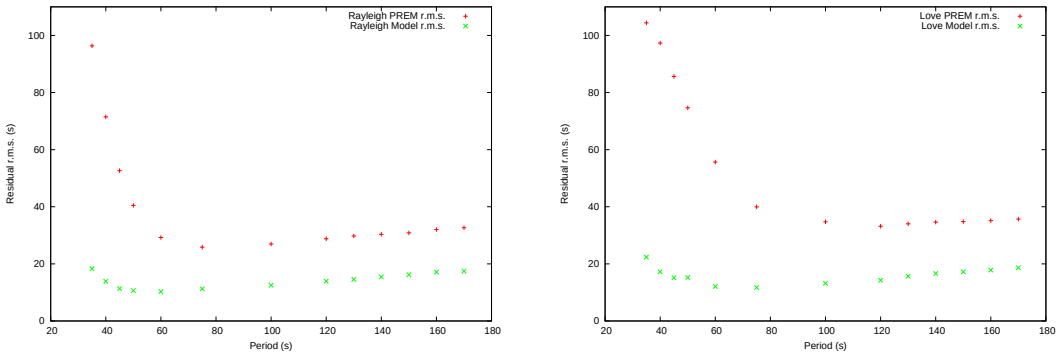


FIGURE 3.4: Left: *Rayleigh time residuals r.m.s. (green) for different periods compared to time residuals with respect to PREM model (red).* Right: *Love time residuals r.m.s. (green) for different periods compared to time residuals with respect to PREM model (red).*

is probably slightly slower than the global one. However, near the borders, where we do not have data, the solution is perfectly compatible with the global model.

In Figure 3.4 we show the time residuals r.m.s. of the Rayleigh and Love resulting group velocity model at different periods, compared with the time residuals with respect to PREM model. Our model explains observed data much better than PREM especially for shorter periods, whose sensitivity to strongly heterogeneous crustal structures evidences the inadequacy of the laterally homogeneous PREM crust. Data are better fitted by our model for longer periods as well, revealing that in such complex regions like Europe and Mediterranean the one dimensional PREM Upper Mantle is a poor approximation of the real structure.

3.5.1 Group velocity maps

Figures 3.5 and 3.6 display our best regional group velocity model for Rayleigh wave at periods 35 s, 40 s, 50 s, 60 s, 75 s, 100 s, 120 s and 150 s. Figures 3.7 and 3.8 show the regional model for Love wave at the same periods. We do not show other periods results because they are very similar to those here displayed. We plot anomalies as relative variations with respect to PREM group velocity. Maps have been plotted after having removed their average anomaly for legibility.

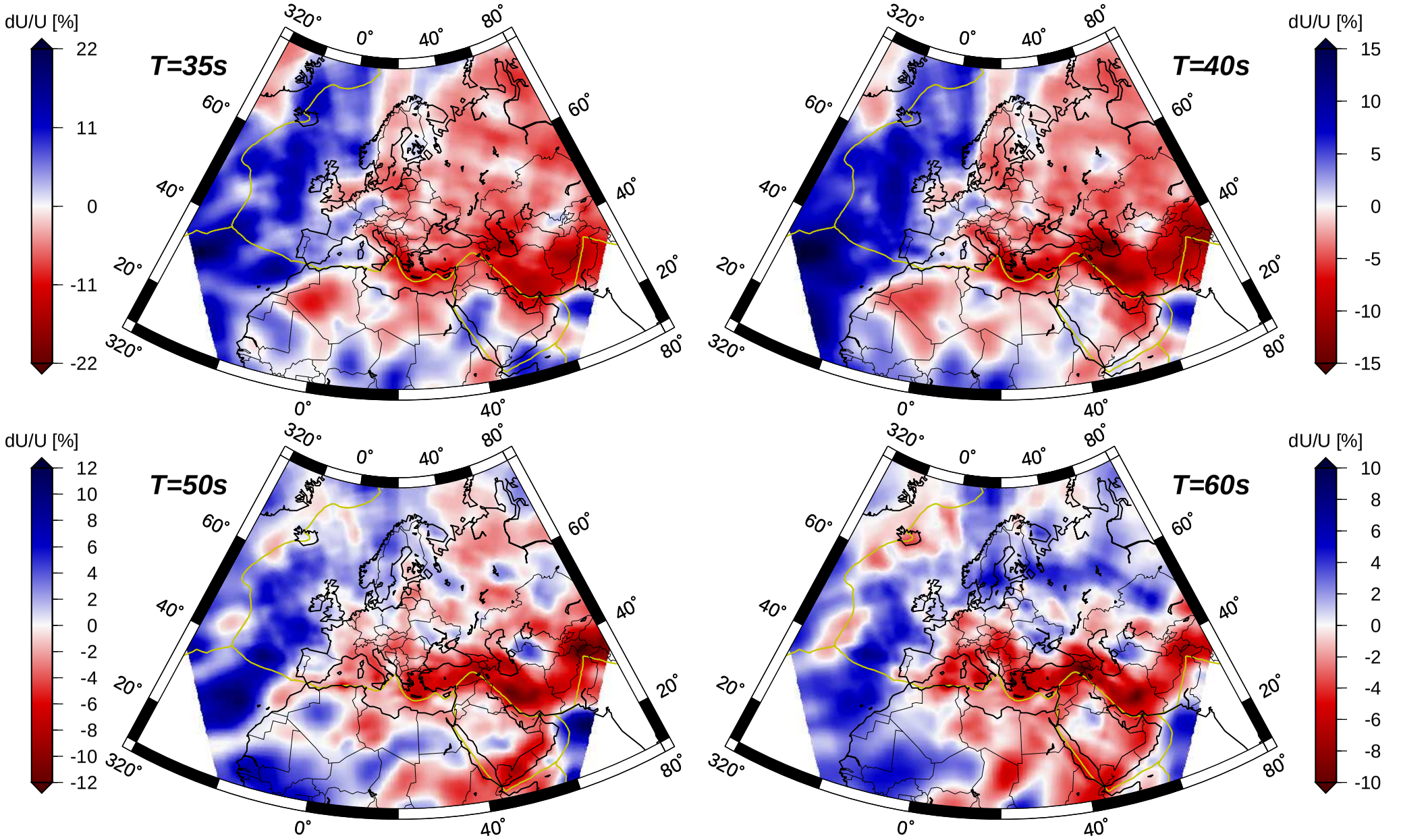


FIGURE 3.5: *Rayleigh group velocity maps at the indicated periods. Color scale represents relative group velocity anomaly with respect to PREM*

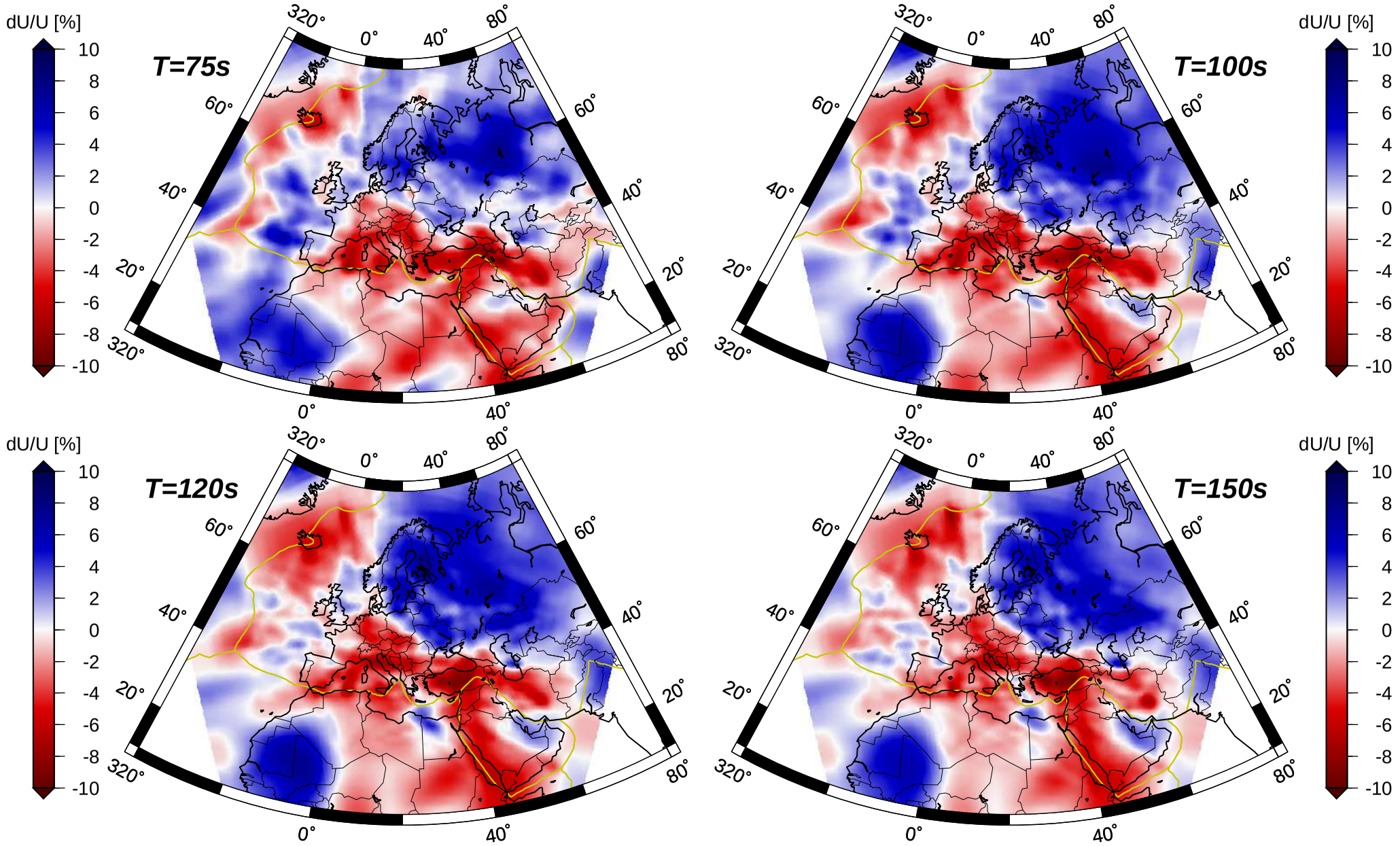


FIGURE 3.6: *Rayleigh group velocity maps at the indicated periods. Color scale represents relative group velocity anomaly with respect to PREM*

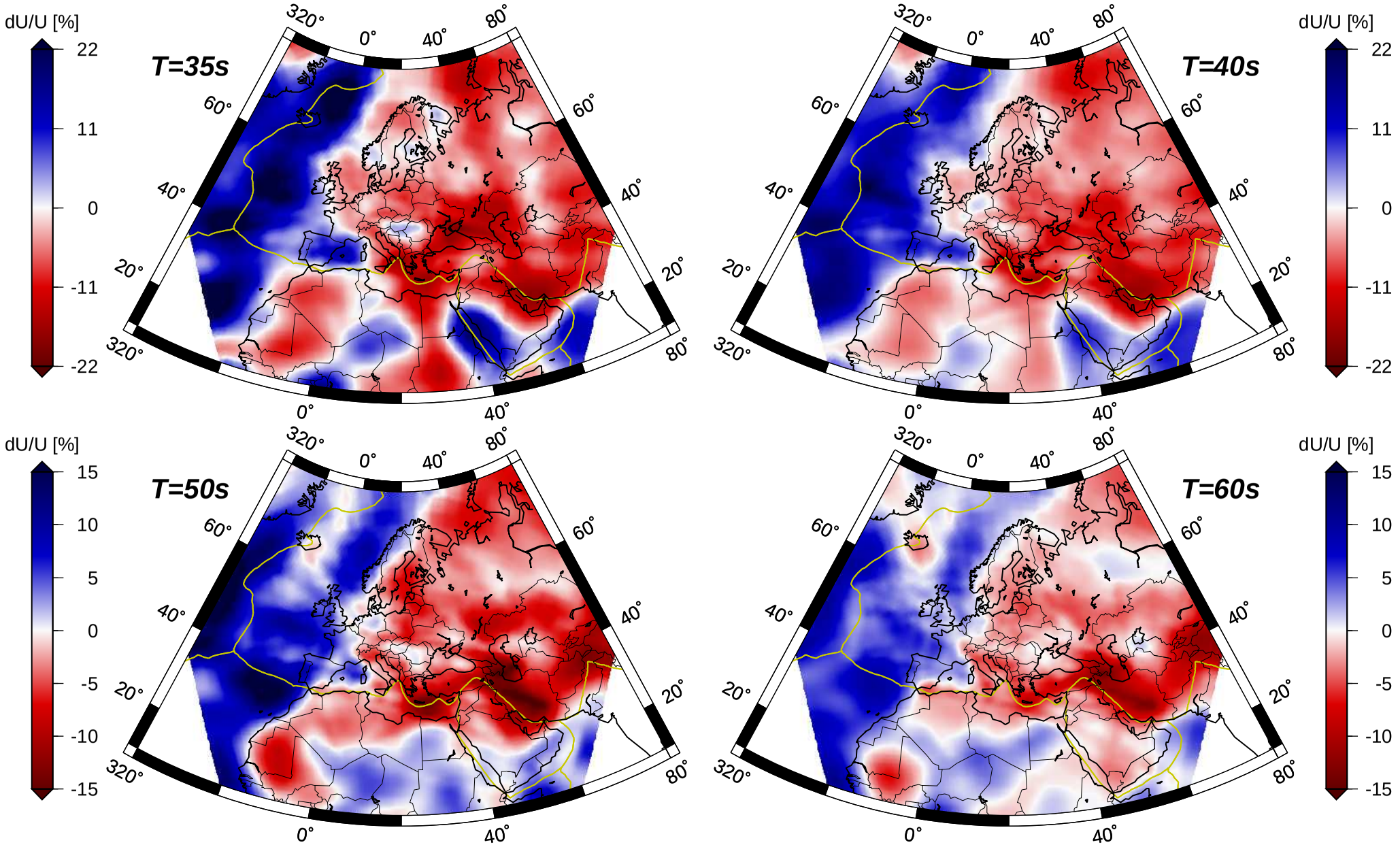


FIGURE 3.7: *Love group velocity maps at the indicated periods. Color scale represents relative group velocity anomaly with respect to PREM*

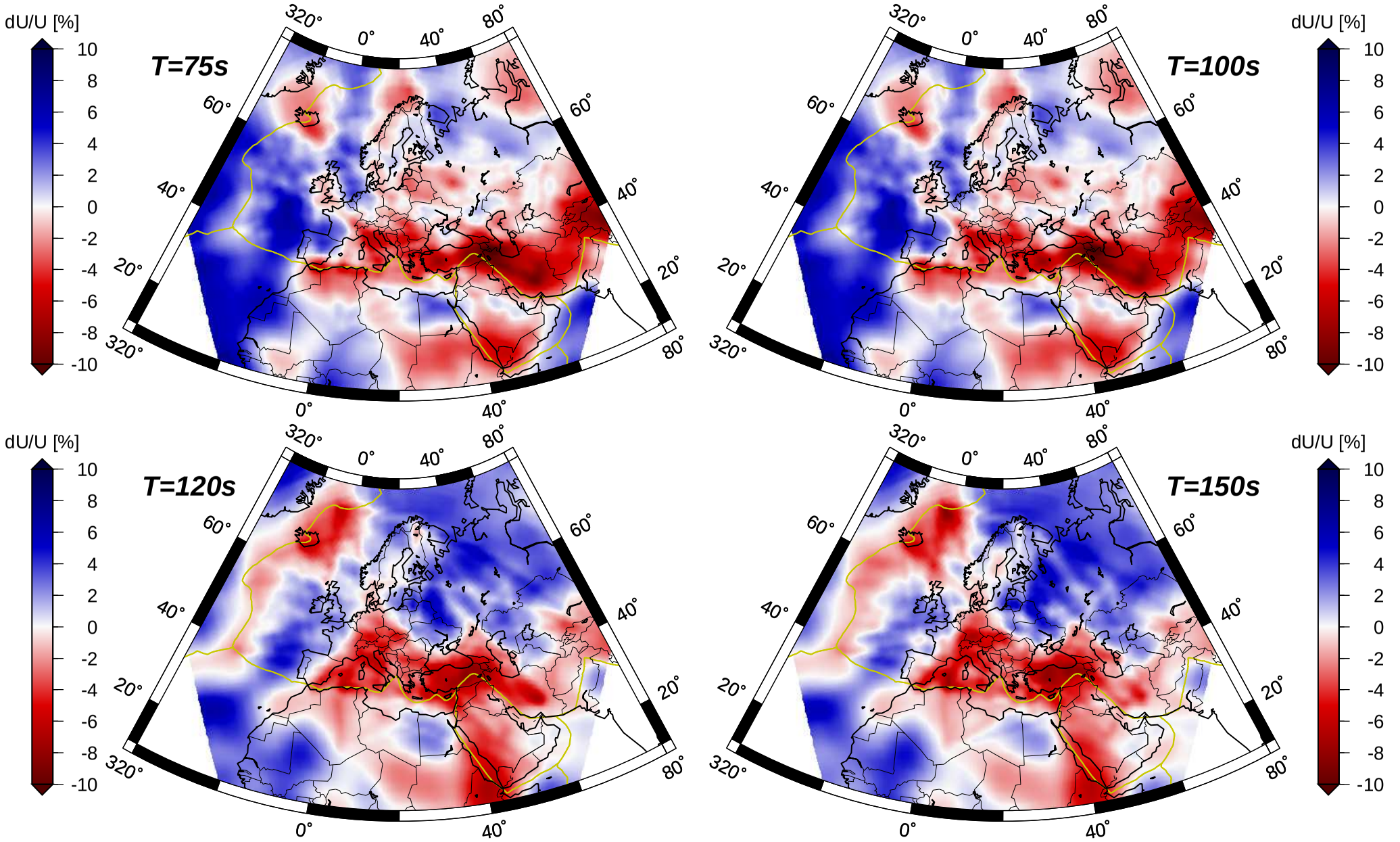


FIGURE 3.8: *Love group velocity maps at the indicated periods. Color scale represents relative group velocity anomaly with respect to PREM*

3.5.2 A posteriori errors

A good estimator of the dispersion of the distribution σ_M in (3.11) is the a posteriori covariance, defined as

$$\mathbf{C'_M} = \left(\mathbf{G^T}\,\mathbf{C_D^{-1}}\,\mathbf{G} + \mathbf{C_M^{-1}} + \mathbf{H^T}\,\mathbf{C_L^{-1}}\,\mathbf{H}\right)^{-1} \tag{3.25}$$

whose diagonal elements give the absolute error on the solution parameters, while the off-diagonal terms are related to the correlation between different cells. Since the solution is sought with the dimension of a slowness C, its covariance as well is an error in slowness δC. We therefore plot the diagonal elements of (3.25) divided by the reference PREM slowness $\mathrm{C_{PREM}}$ as relative errors on model parameters. In Figure 3.9 we show $\delta\mathrm{C}/\mathrm{C_{PREM}}$ for our best inversion of T= 100s Rayleigh and Love wave. For small δC, $\delta\mathrm{C}/\mathrm{C_{PREM}} \simeq \delta\mathrm{U}/\mathrm{U_{PREM}}$ so Figure 3.9 can be compared with 3.5,3.6, 3.7 and 3.8. It is however formally correct to carry on the error analysis in slowness, as we will need it in the next Chapter when we will invert for V_S in depth. As it was to be expected, errors are smaller in good coverage areas and increase toward the borders. Other periods show similar patterns, the most remarkable difference being substantially only in the average error value. This is strictly related to the total number of paths and their geometrical distributions, as well as to errors on data at different periods. These error maps give an idea of the region where the model is better constrained by data, and therefore an

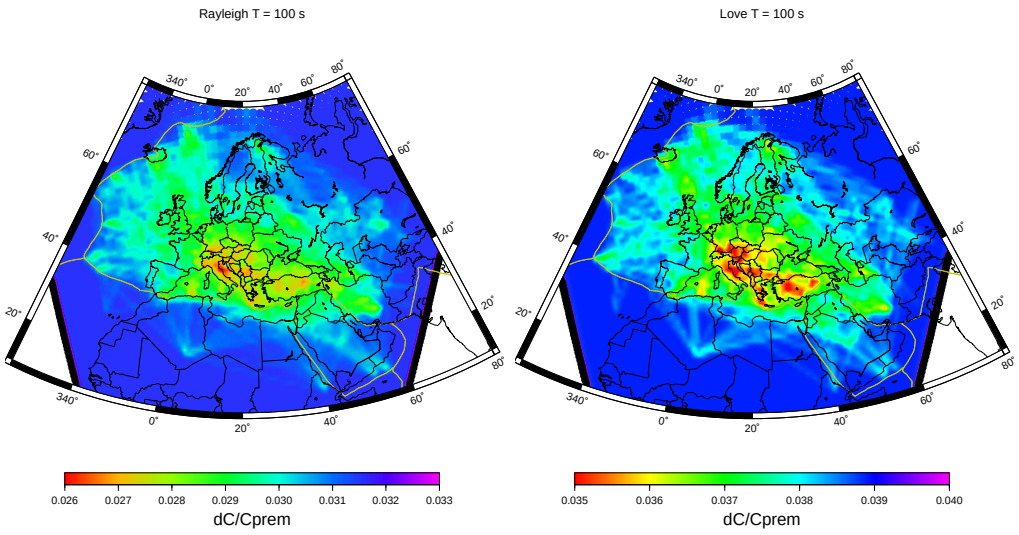


FIGURE 3.9: *Relative errors* $\delta\mathrm{C}/\mathrm{C_{PREM}}$ *on group slowness for Rayleigh* (left) *and Love* (right) *wave at* $T = 100$ *s with respect to PREM slowness* $\mathrm{C_{PREM}}$.

interpretation of the small scale length anomalies in terms of geodynamical structure is more reliable. However, since our reference model is a good constrained global model, also longer scale anomalies in areas of higher error levels are trustworthy.

3.6 Discussion

Analysis of lateral variations in group velocity showed in Figures 3.5, 3.6, 3.7 and 3.8 is an optimal tool for an evaluation of the different tectonic features present in the complex Eurasian and Mediterranean area. Group velocity is actually related to crustal and Upper Mantle structures with a dependence that varies with periods and wave type. Generally, shorter periods bring informations on shear speed anomalies at shallow depths, whereas longer periods sample deeper into the Earth's Upper Mantle. Moreover, Love waves are sensitive to shallower structures than Rayleigh waves of the same period. Exact calculation of depth sensitivity kernels allows the inversion of group velocity anomalies into elastic parameters anomalies at depth. We will present the inversion of such group velocity anomalies for the three-dimensional V_s structure of the Upper Mantle in the next Chapter, but for the moment we can discuss the resulting tomographic maps from a qualitative point of view. The group velocity depth sensitivity is a complicate function of radius, both positive and negative. However , in this discussion we can legitimately make a simplification by assuming that fast group speed anomalies at short (longer) periods are due to fast shear anomalies at shallower (deeper) depths and vice versa. We calculate the amplitude of lateral variations in the resulting maps with twice the standard deviation of the group speed anomalies with respect to PREM. This value ranges from 21% (35 s) to 5% (170 s) for Love waves and from 14% (35 s) to 6% for Rayleigh waves. For a correct interpretation of smaller scale anomalies (which derive from our regional dataset inversion) it is important to remind that better resolved areas are those with lower errors in Figure 3.9, while higher error areas are less reliable and should be considered with more precautions.

In Figure 3.10 are plotted the Rayleigh and Love sensitivity kernels at two representative periods of 35 s and 100 s for a thin oceanic (left) and for a thick continental (right) crust. Shorter periods (35-40 s for Rayleigh, 35-60 s for Love) depicts primarily the differences in crustal thickness, i.e. slow group speed is mainly related to thicker crust, whose shear-wave velocity is much slower than speed in the mantle underlying the thin oceanic crust at depths where sensitivity of those waves is maximum. 35s Love and Rayleigh maps report a strong gradient between the thinner oceanic crust in Atlantic ocean and the

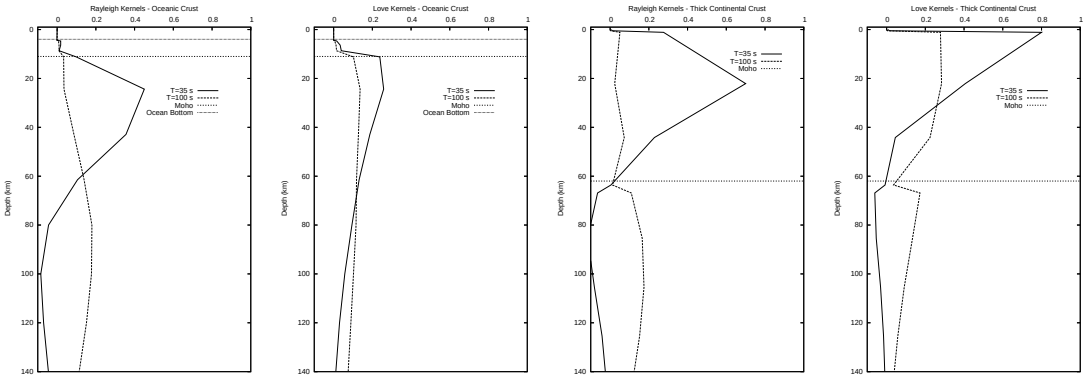


FIGURE 3.10: *Fundamental mode Rayleigh and Love depth sensitivity kernels plotted vs depth, showing sensitivity to different crustal structures at periods T=35 s and T= 100 s. Oceanic crust (*left*) is adapted from model CRUST2.0 tile located at (50N,20W), continental crust (*right*) from tile (36N,80E). Underlying Mantle is model PREM.*

thicker slower continental crust of Eurasia and Africa. Slow anomalies are also present in areas of thickened continental crust, such as the orogenic belts of Caucasus, Zagros Mountains, and Hindu Kush. Fast anomalies are instead present under the Iberian peninsula and the Sardinia-Corse Block, the Red Sea Rift area, and a small fast anomaly in 50 s Love maps correlates very well with the Carpathians. For sublithospheric and Upper Mantle structures investigation it is better to focus the attention on Rayleigh wave maps, which have higher resolution at this depth (see Figure 3.10) and provide more reliable information than Love maps. At 60s and 75 s Rayleigh waves show a persistent slow anomaly that runs from Hellenic-Cyprean Arc along Anatolian plate and Southern Iran (Zagros mountains belt), following well the suture between Eurasian and African plate. This is an extremely active tectonic zone, whose geodynamical asset is still evolving under the collision of the two continents. This area is well distinct from the older and tectonically stable Precambrian Eastern European Craton (EEC), which shows a fast anomaly at 75 s and longer periods. The most striking lineament in our longer periods maps is the Tornquist-Tesseyre Zone (TTZ), i.e. the junction between faster EEC and slower thinner lithosphere of Western Europe that runs from Southern Scandinavia through Poland to the Black Sea. The TTZ has been imaged in other group velocity studies of the region (such as [*Ritzwoller and Levshin*, 1998] and [*Pasyanos*, 2005]), but the particularly dense ray coverage at regional length of this study makes its lineament perfectly delineated. Other thick and cold continental cratons that show fast wave speed in our long period maps are the Western Africa and the Baltic shields.

Slow anomalies with thermal origin are present in correspondence of the Mid Atlantic Ridge and Iceland, getting wider with increasing periods, as well as beneath the Red Sea rift. Upper mantle thermal anomalies of the Western Mediterranean basin and the Southern Thyrrenian magmatic province [*Piromallo and Morelli*, 2003] show up quite clearly at small scale length in the 60 s and 75 s maps. At T = 150 s we see a fast anomaly related to the Hellenic arc subduction zone. As well, we see a trace of the European Cenozoic Rift System on the slow Rhine Graben in Rayleigh 75-150s maps. In the 100-150 s Rayleigh map it is also evident a sharp distinction between the slow Southern Caspian Block and the Fast Northern, divided by the Ashgabat fault. This sharp gradient continues along the Caucasus up to the Northern part of the Black Sea, merging there with the TTZ.

We compare our Love T = 60 s and Rayleigh T = 100 s maps with the recently published maps of group velocity dispersion in Eurasia and North Africa by [*Pasyanos*, 2005], reported in Figure 3.11. We found very good agreement between the large scale features imaged by the two studies, but Western Europe smaller scale details are better imaged in our model: for instance, in 60 s Love map the Iceland Hot Spot is better delineated by a slow anomaly, and in Rayleigh 100 s map we can associate a small anomaly with the Rhine graben, which is instead missing in their model. In general, our model is also less affected by artifacts like streaks and stripes near the border of the study region. This is due to our use of a global reference model which constrains better the solution in scarcely covered areas.

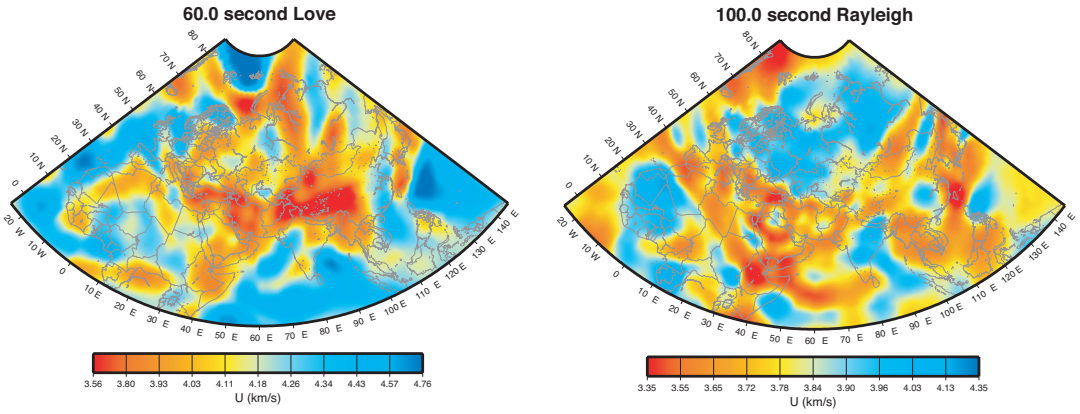


FIGURE 3.11: Left: *T = 60 s group velocity anomalies map for Love wave from the model of* [Pasyanos, *2005*]. Right: *T = 100 s Rayleigh map from the same model. Anomalies are with respect to the map average group velocity.*

Concluding, our group velocity model for the European and Mediterranean region appears to be an improvement with respect to precedent studies, because of its higher resolution and simultaneous compatibility with global models. Dispersion maps can be useful for predicting the arrival time of surface waves, which has applications in earthquake location and regional CMT determination. These maps could also be used as reference starting models in future smaller scale studies to obtain even higher resolution regional models. Group velocity anomalies correlate well with the known tectonic features of this area, providing new constraints on crustal and Upper Mantle structure. Moreover, these group velocity maps can be inverted for the shear velocity at depth to investigate the three-dimensional structure of the Upper Mantle, which will be the subject of Chapter 4.

3.7 Norm and roughness

The choice of the best parameters σ_m and σ_L is always a matter of debate. Models are sensibly dependent on the values of these parameters, but it is not easy to associate the right level of uncertainty σ_m to the a priori model, which in our case is itself result of an inversion, derived from a different dataset on a global scale; in addition, the choice of a particular value of the smoothing parameter σ_L is difficult to explain on the base of physical considerations. In Equation (3.15) it is evident that the two parameters are also related to the requested level of data fit, represented by the estimate of error σ_d, with higher data fit corresponding to higher curvature and/or higher norm of the model. Although some criteria have been proposed to address this problem (such as the Akaike criterion [*Akaike*, 1974]), it appears to be a rather subjective choice, depending on what are, in our expectations, the model features that our data are reliably able to image.

We performed two set of tests for the regional inversion to investigate the effects of different values of the parameters on the resulting model for Rayleigh wave T=100 s data. In the first test we minimize the Laplacian of the model while in the second we minimize the Gradient. In both experiments we select 8 models that have the same data fit as our preferred model PM (the model presented and discussed so far) but show different characteristics in terms of smoothness and closeness to the reference global model $\mathbf{m}_0$. By fitting the data equally well, these 8 models are formally all equally valid.

3.7.1 Laplacian minimization

We apply Equation (3.24) to minimize the curvature of the solution . We let $K_m = \sigma_d^2/\sigma_m^2$ and $K_L = \sigma_d^2/\sigma_L^2$ vary from 10^4 to 10^8 with steps set on a logarithmic scale. At each inversion we compute $\sigma_m^2 = \sigma_d^2/K_m$ and $\sigma_L^2 = \sigma_d^2/K_L$ for all the models, where $\sigma_d = 12.09$ s is the data error for T = 100s. Figure 3.12 shows the time residual r.m.s. $T_{r.m.s.}$ as a function of (σ_L, σ_m) for each inversion, which define a curved surface (whose z value is evidenced with different colors). We cut this surface at $T_{r.m.s.} = 12.50$ s (the $T_{r.m.s.}$ of the PM) and we select 8 models (green circles in Figure 3.12 and following) which have the same data fit as the PM. We analyze the time residual distribution $(d_i^{obs} - \sum_{j=1}^{M} G_{ij}\ m_j)$, the $(m_i - m_i^{prior})$ and the laplacian $(H_{ij}m_j)$ distribution for these models in terms of varying σ_m and σ_L. The latter two distributions are made up only of cells touched by rays, so as not to introduce any bias from the part of the a priori model which is not changed by the data. In Figure (3.15) and (3.16) we show the 3 distributions and maps for the 8 selected models, plotted with the same color scale of Figure 3.6 for T=100s.

In Figures 3.13 and 3.14 we show the mean, standard deviation and r.m.s. of the 3 distributions for each inversion.

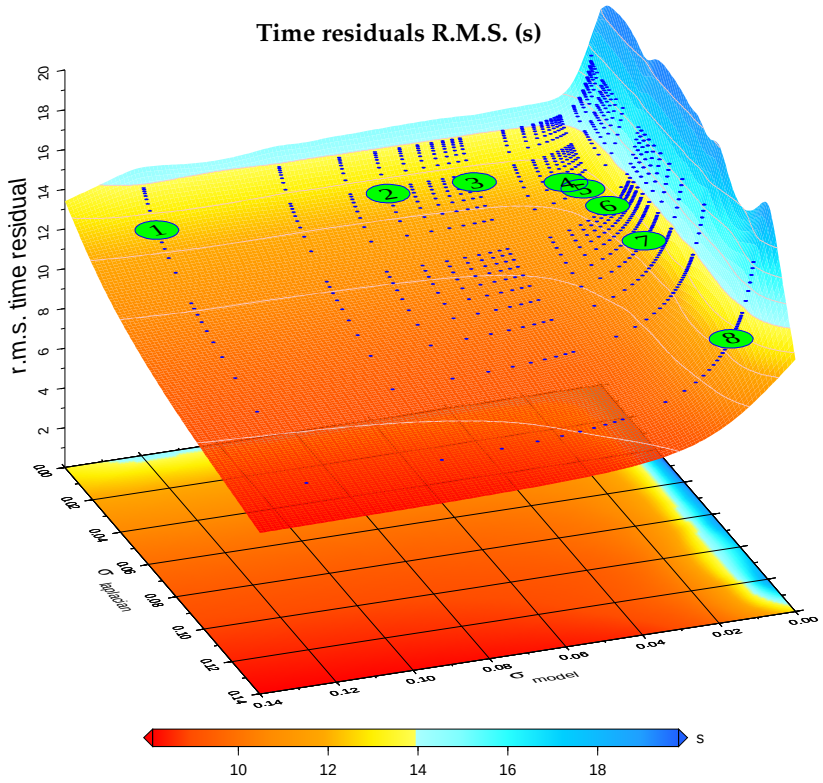


FIGURE 3.12: *Residual misfit surface $T_{r.m.s.}$ as a function of regularization parameters (σ_L and σ_m). Blue dots are the $(\sigma_L, \sigma_m, T_{r.m.s.})$ triplets actually calculated, that are then interpolated with a colored surface. We project on the (σ_m, σ_L) plane the color of the surface, related to the z-value ($T_{r.m.s.}$). Green circles are the 8 selected models whose $T_{r.m.s.} \simeq 12.50$ s (the $T_{r.m.s.}$ of the PM).*

Analysis

- **Time residuals ($\mathbf{d}^{obs} - \mathbf{G}\,\mathbf{m}$) distribution:** The effect of σ_m (model constrained to be close to $\mathbf{m}_0$) is evident on the mean of the residuals distribution, as we can see in Figure 3.13 (left column, top). While for large values of σ_m the mean is approximatively constant, it increases sharply for $\sigma_m \lesssim 0.02$ s/km. In Table 3.1 are reported the values mean, σ_d and r.m.s. for the selected models. Going from model 1 to model 8 the average time residual increase from 1.73 s to 3.72 s. This

means that the model becomes faster (on average) when it is constrained to be closer to $\mathbf{m}_0$. This behaviour is probably due to a different mean in ETL97 and our regional dataset, the first being somewhat faster and therefore tending to decrease the average computed traveltime. As we requested, models fit data equally well (r.m.s. is circa constant) and actually the shape of the residuals distribution (the width of the Gaussian bell) does not change much from model 1 through 8. On the other hand, the effect of σ_L on the mean of the residuals appears to be of minor importance, with only a light decrease of the mean when σ_L increases. σ_L acts on time residual standard deviation and r.m.s., mainly when σ_m weights less (compare for example a branch for constant $\sigma_m \simeq 0.12$ and $\sigma_m \simeq 0.01$ in Figure 3.13). Thus, when the model is constrained to be smooth, as one could expect, the data fit decreases.

- $(\mathbf{m} - \mathbf{m_0})$ **distribution:** The main features of the residuals distributions are reflected also in the distribution of the solution around the a priori model $\mathbf{m}_0$ In Figure 3.13 (right column) are represented the mean, standard deviation and r.m.s. of the distributions as a function of (σ_m, σ_L). As we pointed out for the time residuals mean, the mean of $(\mathbf{m} - \mathbf{m_0})$ decreases sharply for $\sigma_m \lesssim 0.02$ s/km, while it is almost not affected by σ_L. Obviously, since a smaller σ_m keeps the model closer to $\mathbf{m}_0$, this reduces the standard deviation and r.m.s. of the distribution but the shape of such distribution changes dramatically, as we can see in model 1 to 8 (see central column in Figures 3.15 and 3.16). When σ_m approaches 0.02 s/km the distribution begins to loose symmetry (and consequently it is not really Gaussian anymore), with a longer tail on the right (larger slowness $\rightarrow$ slower model). This is due to the trade-off between data-fit (which requires a slower model, as discussed previously) and a priori model. The solution tries to satisfy the vicinity to the a priori model with a narrower distribution around $\mathbf{m_0}$, but it has to yield to (under?) data fit increasing the number of slower cells with respect to faster ones (ma si capisce?). The effect of σ_L on the mean of the distribution is of minor importance compared to σ_m, while on standard deviation and r.m.s. we can note that, mainly for large values of σ_m, a small value of σ_L tends to decrease the standard deviation and the r.m.s.. Thus, we can deduce that the smoothing condition tends to bring the solution closer to $\mathbf{m}_0$, because $\mathbf{m}_0$ is already a model with minimum curvature.

- **H m distribution:** In Figure 3.14 are plotted the the mean, standard deviation and r.m.s. of the distributions of the laplacian of the model **H m**. As we expected, the standard deviation decreases when σ_L gets smaller, and this is particularly evident for large σ_m (i.e. when the model is quite far from $\mathbf{m}_0$). The same characteristics are present in Figures 3.15 and 3.16 where the Gaussian bell becomes wider particularly from model 4 to 8, in correspondence with rougher models . However, in Figure 3.14 we notice also a standard deviation decrease for small values of σ_m when σ_L is large: this is likely to be an effect of the smoothness of $\mathbf{m}_0$, which is a model of minimum curvature. The r.m.s. is very similar to the standard deviation for all the inversion and the mean is consequently very close to zero, although a bit negative. The mean , at constant σ_L, has a minimum for $\sigma_m \simeq 0.02$ s/km. Having a mean value very close to zero and a fairly symmetric distribution for each one of the 8 selected models, we can argue that points with negative curvature are balanced by points with positive curvature.

TABLE 3.1: *Mean, standard deviation and r.m.s. of the distributions for models 1 to 8*

	time residuals (s)			$m - m_0$ (s/km)			H m		
n°	mean	σ	r.m.s.	mean $\cdot 10^{-4}$	σ_{m-m_0} $\cdot 10^{-4}$	r.m.s. $\cdot 10^{-4}$	mean $\cdot 10^{-4}$	σ_{Hm} $\cdot 10^{-4}$	r.m.s. $\cdot 10^{-4}$
1	1.7305	12.362	12.482	32.70	63.61	71.52	-0.22	30.45	30.45
2	1.8047	12.400	12.531	31.95	58.08	66.29	-0.38	29.91	29.91
3	1.9196	12.379	12.527	31.00	53.63	61.95	-0.48	31.12	31.12
4	2.5491	12.243	12.505	27.70	43.22	51.33	-0.72	41.92	41.93
5	2.8212	12.188	12.510	26.65	40.69	48.64	-0.82	48.98	48.98
6	3.1902	12.090	12.504	25.34	38.22	45.86	-0.98	63.54	63.54
7	3.5240	11.989	12.496	24.24	36.89	44.14	-1.18	83.08	83.09
8	3.7231	11.889	12.458	23.56	36.94	43.81	-1.34	103.54	103.55

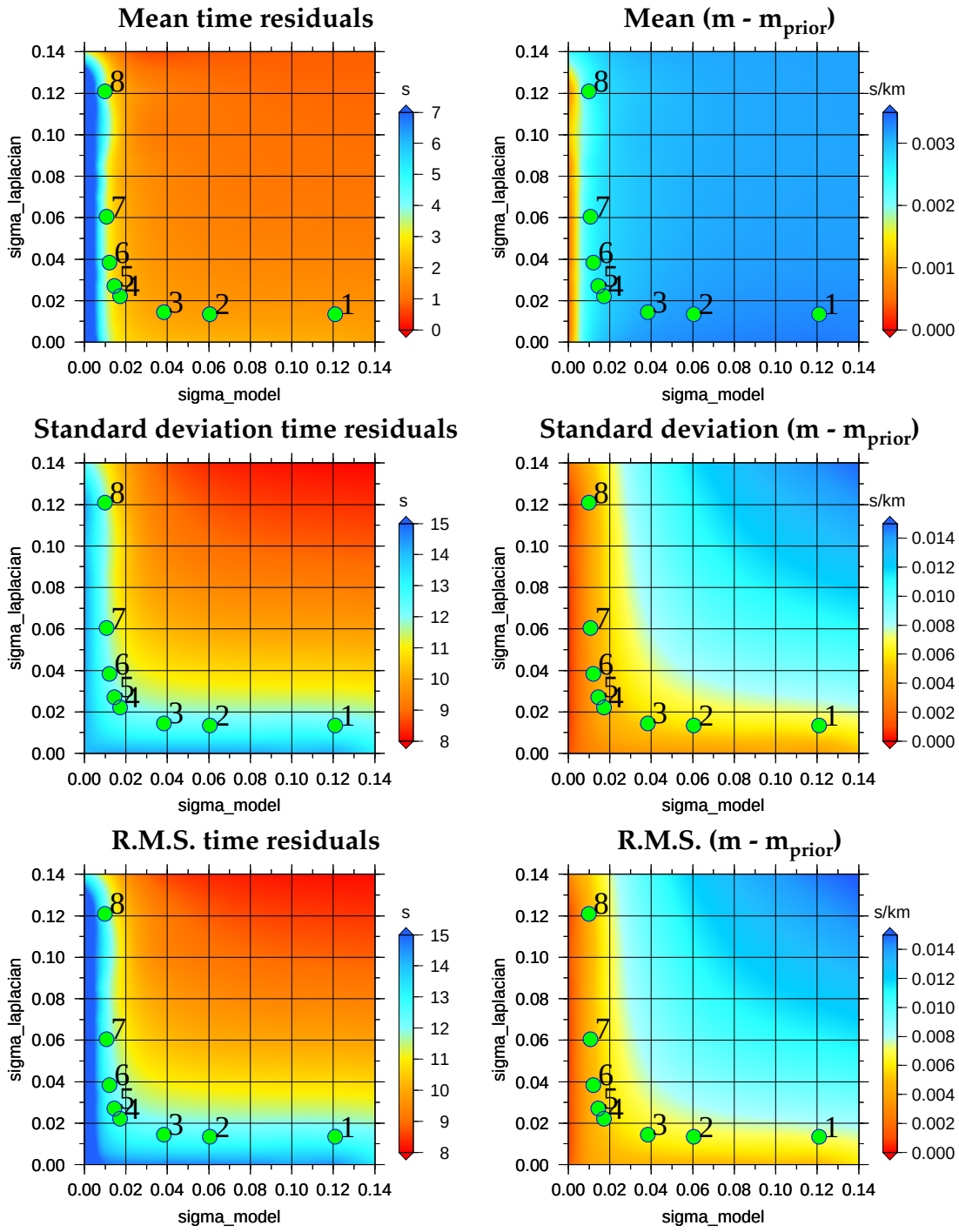


FIGURE 3.13: Left: *Mean* (top), *standard deviation* (center) *and r.m.s.* (bottom) *of the time residuals distribution. Green circles are the selected models whose* $T_{r.m.s.} \simeq 12.50$ *s. Model 6 is the PM.* Right: *Mean* (top), *standard deviation* (center) *and r.m.s.* (bottom) *of the a posteriori distribution around the reference model* $(\mathbf{m} - \mathbf{m_0})$.

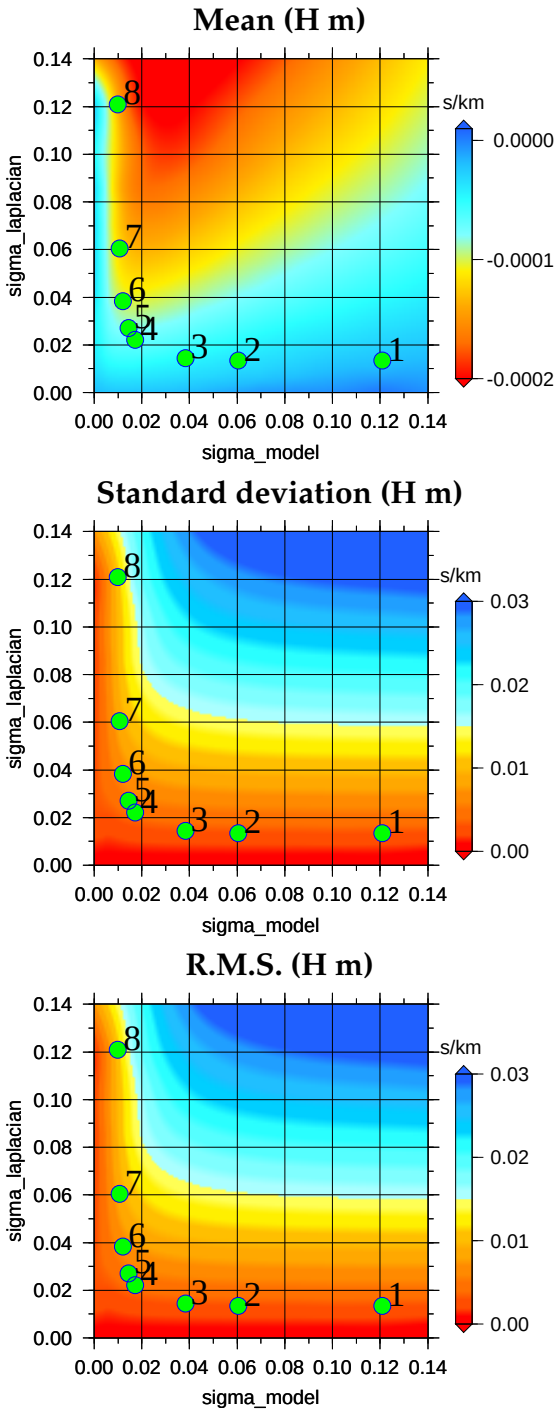


FIGURE 3.14: *Mean* (top), *standard deviation* (center) *and r.m.s.* (bottom) *of the laplacian distribution* **H m** . *Green circles are the selected models whose* $T_{r.m.s.} \simeq 12.50$ *s.*

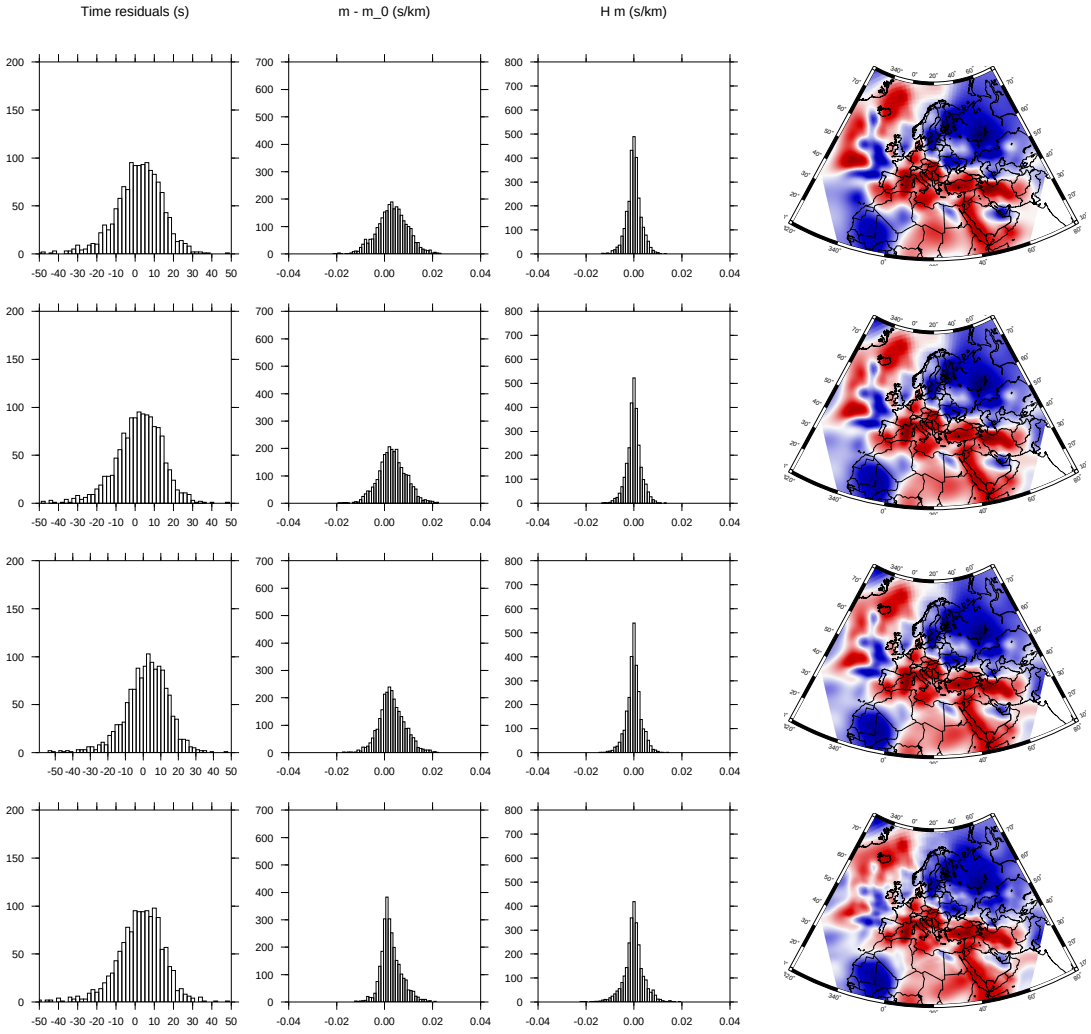


FIGURE 3.15: From left to right : *Time Residuals distribution* ($\mathbf{t}^{obs} - \mathbf{t}^{calc}$), ($\mathbf{m} - \mathbf{m_0}$) *distribution,* $\mathbf{H}\,\mathbf{m}$ *distribution. Models number 1, 2 3, 4 (from top to bottom). Color scale go from -10% to 10% as in Figure 3.6*

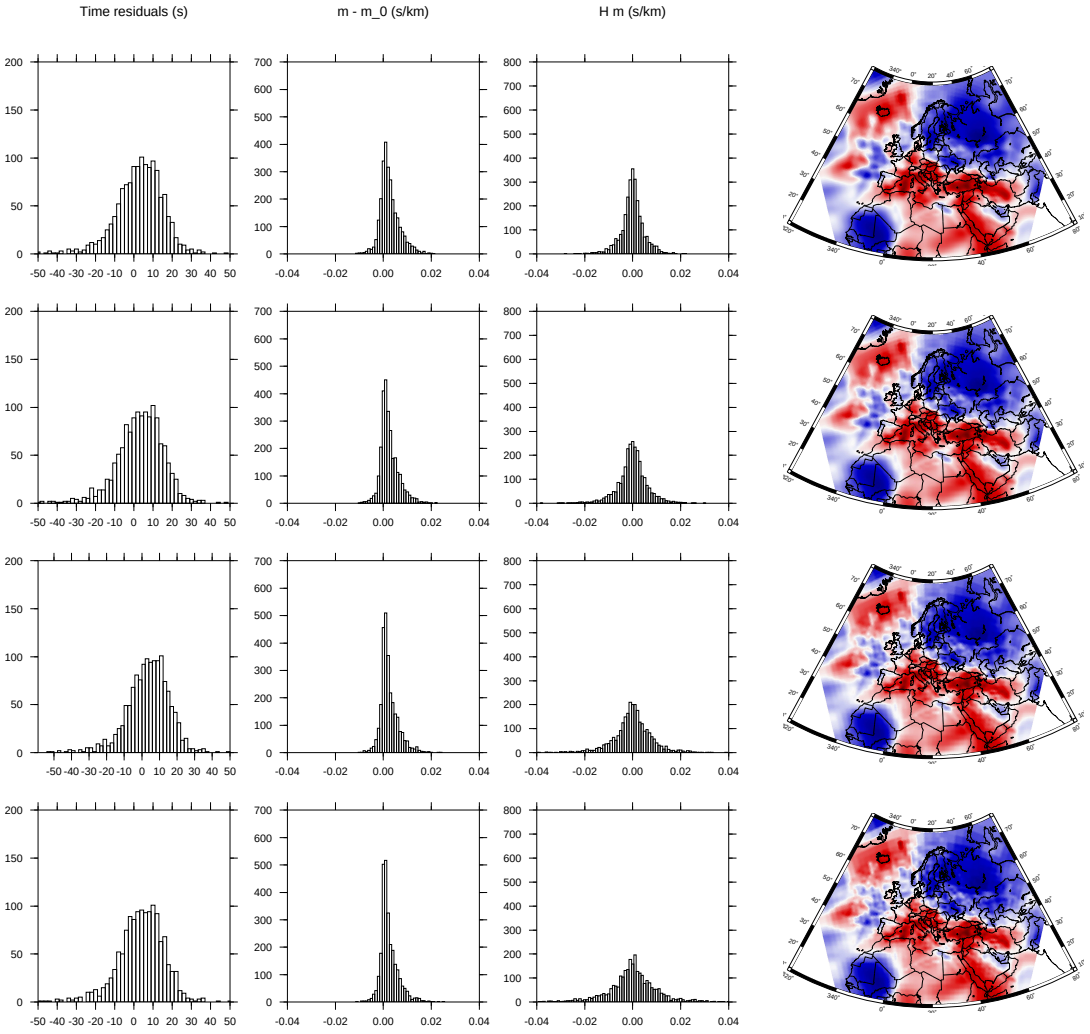


FIGURE 3.16: From left to right : *Time Residuals distribution* ($\mathbf{t}^{obs} - \mathbf{t}^{calc}$), *(*$\mathbf{m} - \mathbf{m_0}$*) distribution,* $\mathbf{H}\,\mathbf{m}$ *distribution. Models number 5, 6 (PM), 7, 8 (from top to bottom). Color scale go from -10% to 10% as in Figure 3.6*

3.7.2 Gradient minimization

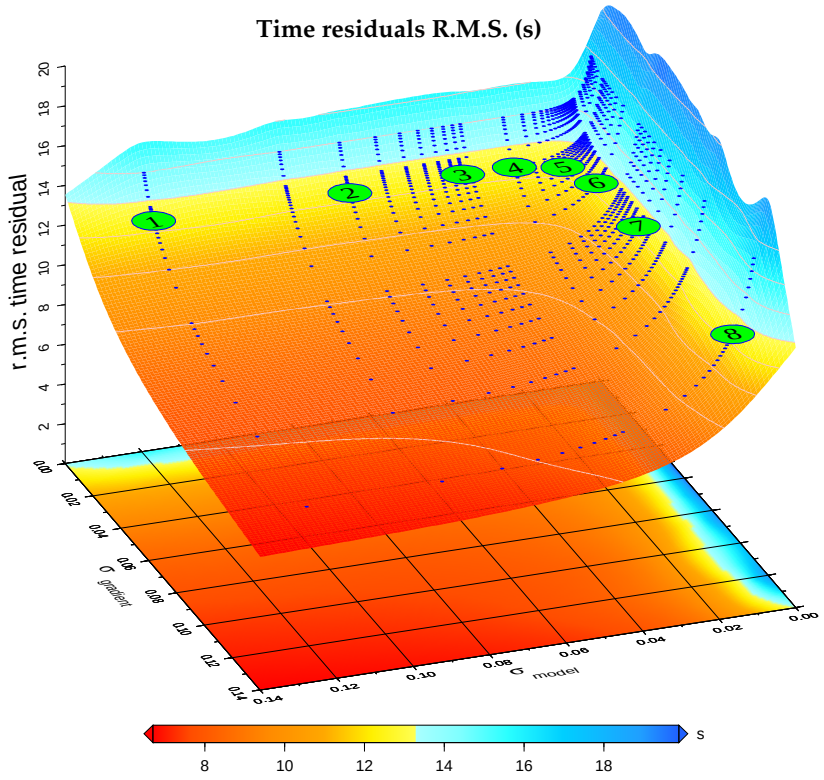


FIGURE 3.17: *Residual misfit surface $T_{r.m.s.}$ as a function of regularization parameters (σ_g and σ_m). Blue dots are the $(\sigma_g, \sigma_m, T_{r.m.s.})$ triplets actually calculated, that are then interpolated with a colored surface. We project on the (σ_m, σ_g) plane the color of the surface, related to the z-value ($T_{r.m.s.}$). Green circles are the 8 selected models whose $T_{r.m.s.} \simeq 12.50$ s (the $T_{r.m.s.}$ of the PM, obtained with the laplacian minimization).*

In the second test we apply Equation (3.22) to minimize the Gradient of the solution, to test the effect of a different smoothing condition. As in the Laplacian case, we let $K_m = \sigma_d^2/\sigma_m^2$ and $K_g = \sigma_d^2/\sigma_g^2$ vary from 10^4 to 10^8 on a logarithmic scale. We then calculate $\sigma_m^2 = \sigma_d^2/K_m$ and $\sigma_g^2 = \sigma_d^2/K_g$ for all the models, whit $\sigma_d = 12.09$ s. For each inversion we compute the r.m.s. of the time residuals $T_{r.m.s.}$. Figure 3.17 shows with blue dots the $T_{r.m.s.}$ as a function of the (σ_g, σ_m) regularitazion parameters. The values actually calculated are then interpolated by the colored surface. We carry out the analysis on the time residuals distribution $(d_i^{obs} - \sum_{j=1}^{M} G_{ij}\ m_j)$, the $(m_i - m_i^0)$ and the

gradient $(H_{x_{ij}} m_j, H_{y_{ij}} m_j)$ distribution for each model in terms of varying σ_m and σ_g. In Figures 3.19 and 3.20 we show with different colors the mean, the standard deviation and the r.m.s. of each distribution. We select 8 models (green circles) that have the same $T_{r.m.s.}$ of the Preferred Model ($T_{r.m.s.} = 12.50$ s), and we display the three distributions for these 8 models in Figures 3.21 and 3.22.

Analysis

- **Time residuals $(\mathbf{d}^{obs} - \mathbf{G}\,\mathbf{m})$ distribution:** Figure 3.19 resembles closely the Laplacian case, with the mean (top) that remains approximatively constant throughout the (σ_m, σ_g) plane with a sharp increase for $\sigma_m \lesssim 0.02$ s/km, i.e. when the model approaches $\mathbf{m}_0$. Also the standard deviation and the r.m.s. show similarity with the Laplacian case, although with slightly lower values. This could lead to the conclusion that, since the Gradient minimization provide a better data fit, it should be a preferable smoothing condition. However, we must keep in mind that, if we choose $\sigma_L = \sigma_g$ and we invert with the same σ_m, the model with the minimum gradient will be rougher with respect to the minimum laplacian model. Therefore, since the subjective choice of the "best" model does not involve only data fit, but also acceptable smoothness, the two smoothing methods are completely equivalent. For example we plot in Figure 3.18 model A (top) obtained with laplacian minimization and $\sigma_L = 0.054$, $\sigma_m = 0.054$, and model B (down) obtained with gradient minimization and $\sigma_g = 0.054$, $\sigma_m = 0.054$. Model A has time residual r.m.s. = 10.32 s and model B has r.m.s. = 9.07 s.

- **$(\mathbf{m} - \mathbf{m_0})$ distribution:** The right column of Figure 3.19 shows the mean, standard deviation and r.m.s. of the models distributions around the reference model. The pattern are perfectly comparable with the Laplacian case, with the mean that is extremely sensitive to small values of σ_m and almost not influenced by σ_g. The standard deviation and r.m.s. still increase when the solution is less constrained (for larger values of both σ_g and σ_m), and even the gradient minimization alone is able to bring the solution toward $\mathbf{m}_0$, because the reference model, although is not a minimum gradient model, is nevertheless a smooth model. Since time residuals standard deviation and r.m.s. are generally lower than in the Laplacian minimization, consequently $(\mathbf{m} - \mathbf{m_0})$ st.dev and r.m.s. are slightly higher, at least for medium to large σ_g and σ_m. As previously remarked, as the mean decreases, the asymmetry in the $(\mathbf{m} - \mathbf{m_0})$ distribution for the selected models increases going

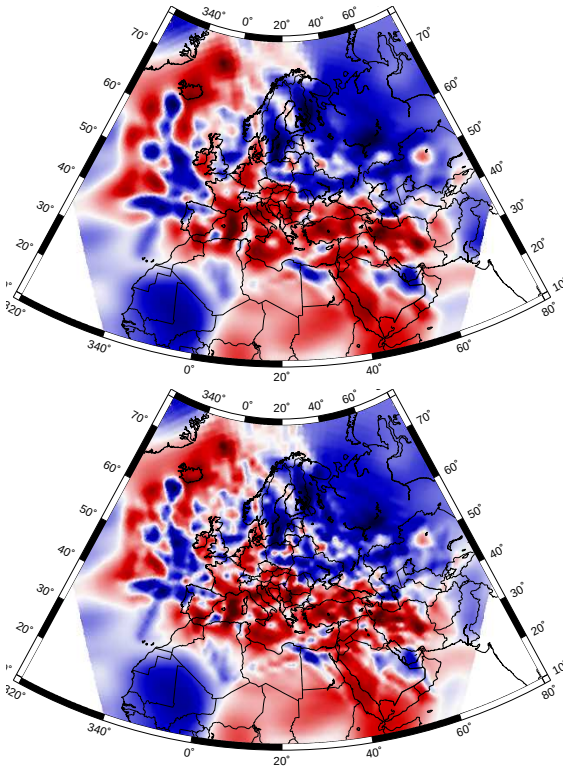


FIGURE 3.18: Top: *Model A, obtained with laplacian minimization and* $\sigma_L = 0.054$, $\sigma_m = 0.054$ *has time residuals r.m.s. = 10.32 s.* Down: *Model B, obtained with gradient minimization and* $\sigma_g = 0.054$, $\sigma_m = 0.054$ *has time residuals r.m.s.=9.07s.*

from model 1 to 8 (see Figures 3.21 and 3.22).

- **$\mathbf{H_x\,m}$ and $\mathbf{H_y\,m}$ distribution:** In Figure 3.20 we show $\mathbf{H_x\,m}$ (left) and $\mathbf{H_y\,m}$ (right). As with the Laplacian, the mean is extremely close to zero, although a bit negative. There is a strong difference in pattern between the **x** and **y** direction: whereas the **x** gradient mean goes close to zero both for small σ_g and σ_m (always because $\mathbf{m}_0$ is smooth), the **y** gradient remains further from zero for most values of σ_g when σ_m is small, but when σ_m increases there is a band of σ_g (roughly $0.03 \lesssim \sigma_g \lesssim 0.10$) for which the mean moves toward zero. This discrepancy could derive from the preferential orientation EW of most rays, that sample the southern Europe and Mediterranen area and from the same EW orientation of the slow tectonic Tethydean region. For what concerns standard deviation and r.m.s. the two directions are similar, and resemble fairly well the laplacian pattern, but with lower values: apparently the Gradient minimization is better satisfied than the Laplacian one.

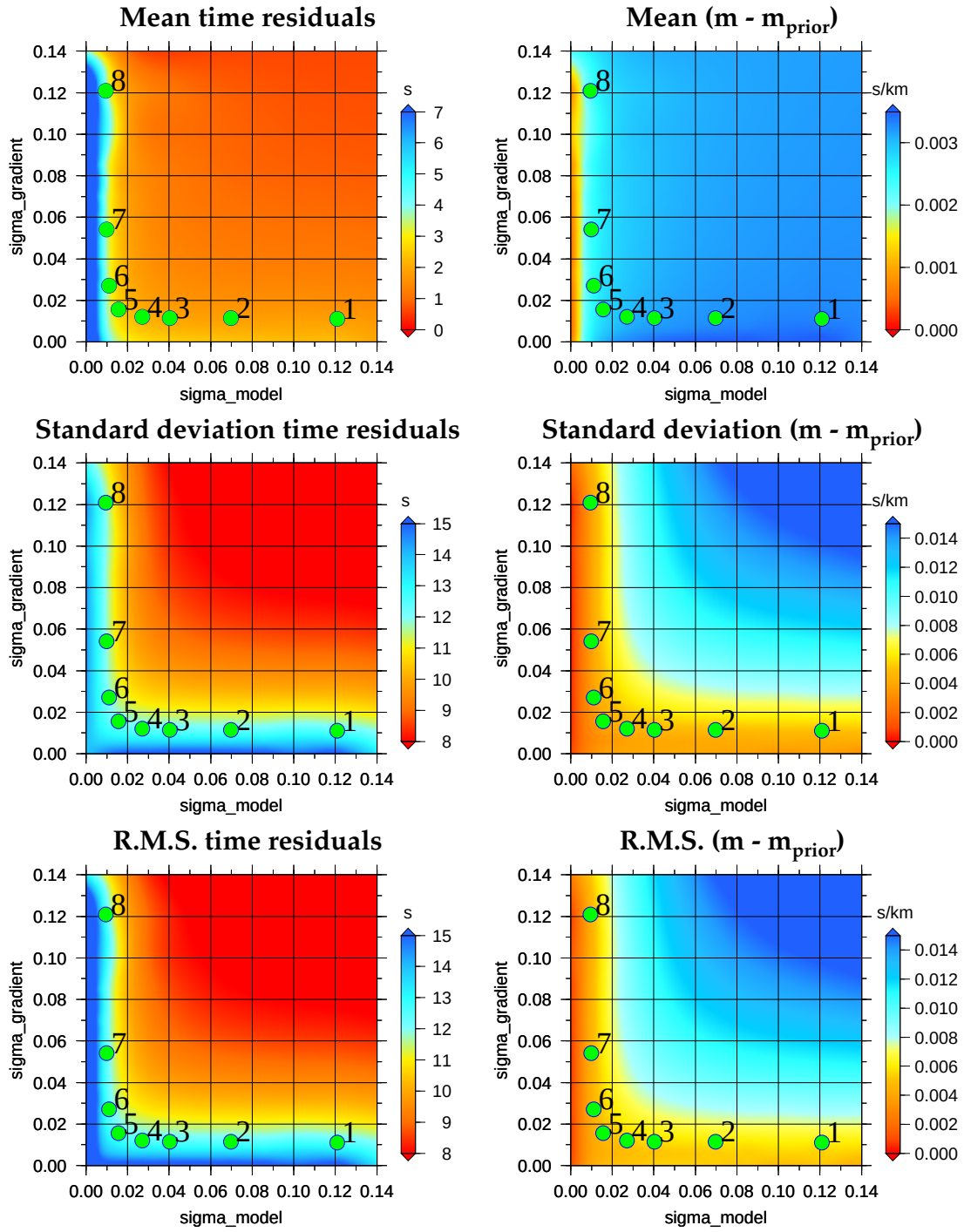


FIGURE 3.19: Left: *Mean* (top), *standard deviation* (center) *and r.m.s.* (bottom) *of the time residuals distribution. Green circles are the selected models whose* $T_{r.m.s.} \simeq 12.50$ *s.* Right: *Mean* (top), *standard deviation* (center) *and r.m.s.* (bottom) *of the a posteriori distribution around the reference model* $(\mathbf{m} - \mathbf{m_0})$.

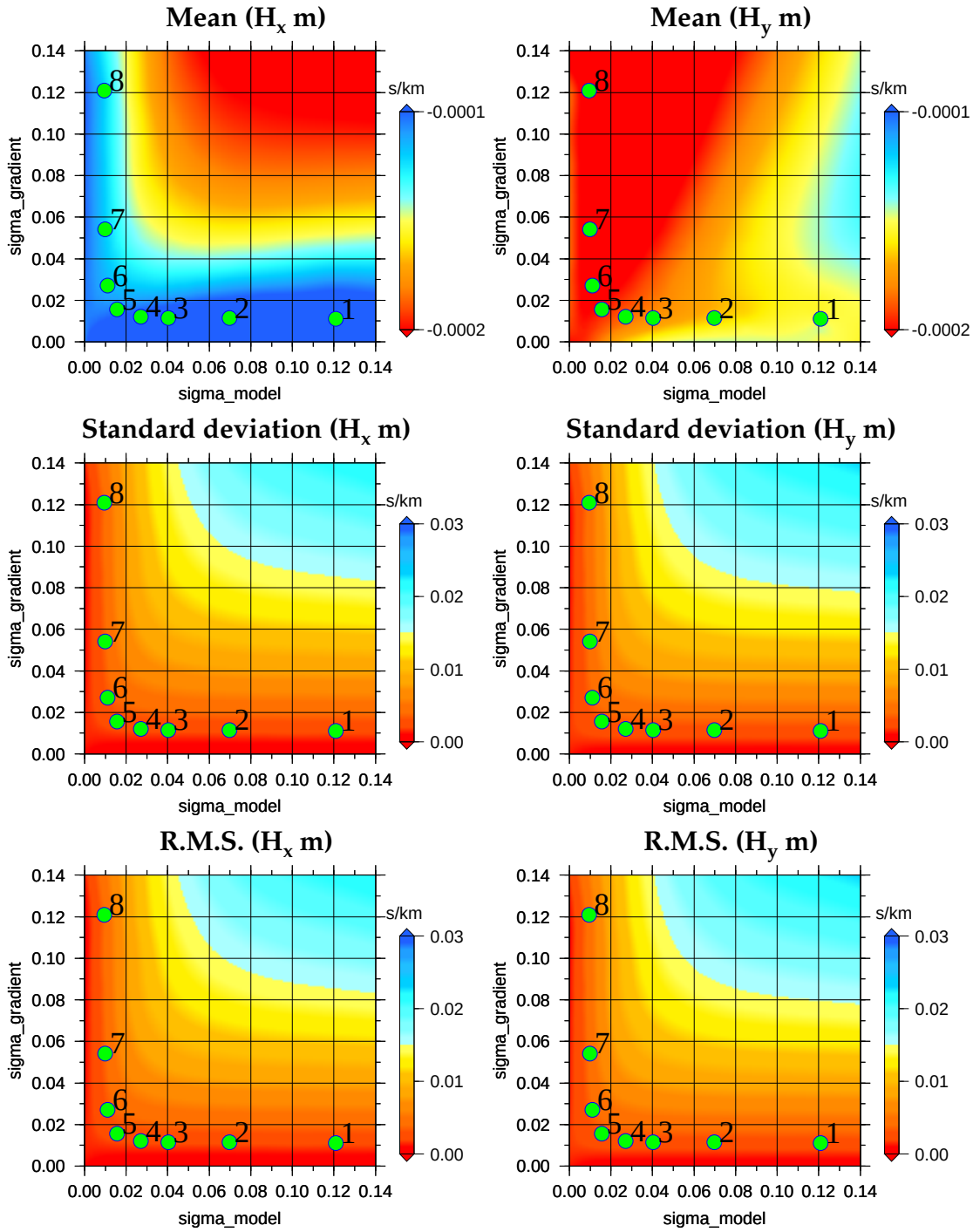


FIGURE 3.20: *Mean* (top)*, standard deviation* (center) *and r.m.s.* (bottom) *of the gradient distribution* $\mathbf{H_x}\,\mathbf{m}, \mathbf{H_y}\,\mathbf{m}$. *Green circles are the selected models whose* $T_{r.m.s.} \simeq 12.50$ *s.*

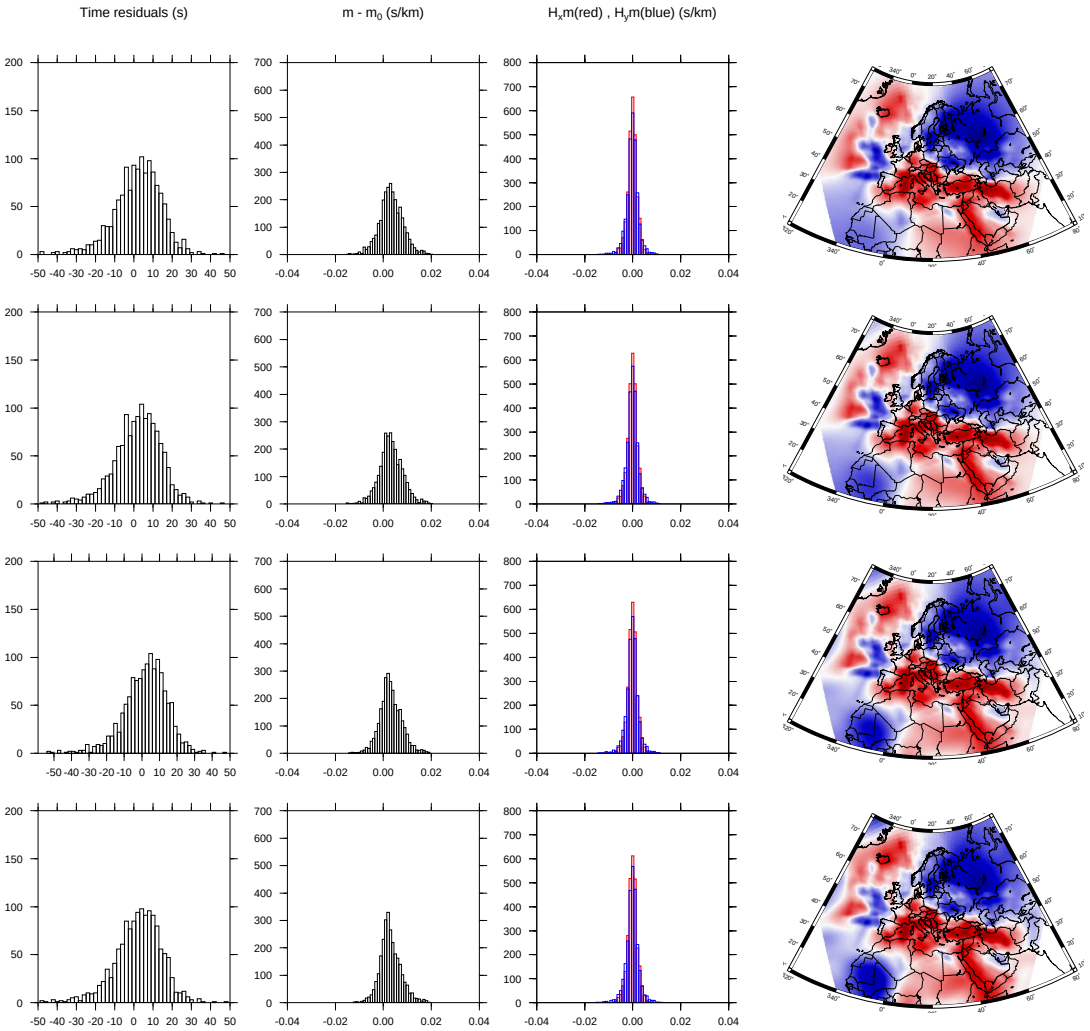


FIGURE 3.21: Left: *Time Residuals distribution* ($t^{obs} - t^{calc}$). Center: $(\mathbf{m} - \mathbf{m_0})$ *distribution (NB: only the cells where I have rays)* Right: $\mathbf{H_x}\,\mathbf{m}$*(red) and* $\mathbf{H_y}\,\mathbf{m}$ *(blue) distribution (NB: only the cells where I have rays). Most right column: Models number 1, 2 3, 4 (from top to down). Color scale goes from -10% to 10% as in Figure 3.6*

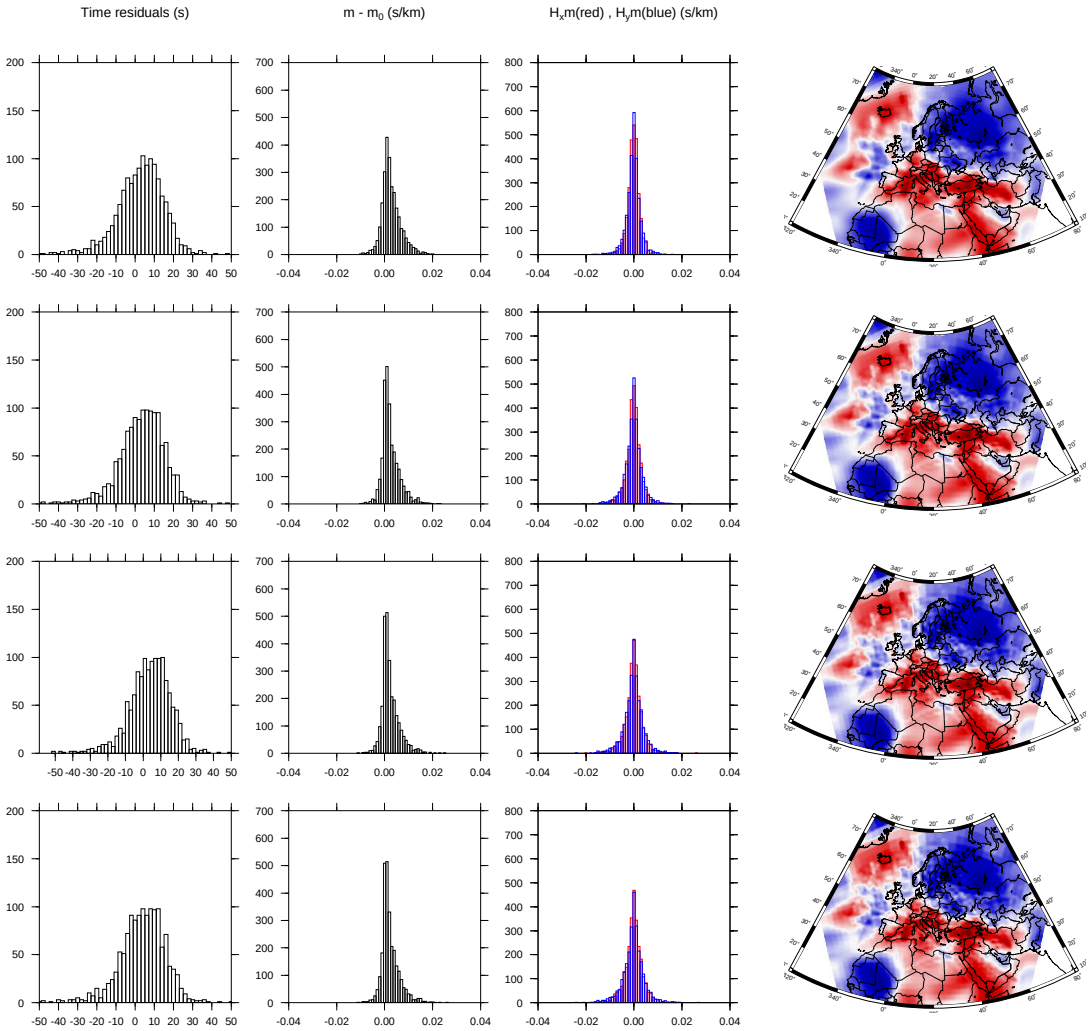


FIGURE 3.22: Left: *Time Residuals distribution* ($t^{obs} - t^{calc}$). Center: $(\mathbf{m} - \mathbf{m_0})$ *distribution (NB: only the cells where I have rays)* Right: $\mathbf{H_x}\,\mathbf{m}$*(red) and* $\mathbf{H_y}\,\mathbf{m}$ *(blue) distribution (NB: only the cells where I have rays). Most right column: Models number 5, 6 (PM), 7, 8 (from top to down). Color scale goes from -10% to 10% as in Figure 3.6*

TABLE 3.2: *Mean, standard deviation and r.m.s. of the distributions for models 1 to 8*

	time residuals (s)			$m - m_0$ (s/km)			H_y m		
n°	mean	σ	r.m.s.	mean $\cdot 10^{-2}$	σ_{m-m_0} $\cdot 10^{-2}$	r.m.s. $\cdot 10^{-2}$	mean $\cdot 10^{-2}$	$\sigma_{H_y m}$ $\cdot 10^{-2}$	r.m.s. $\cdot 10^{-2}$
1	1.8119	12.4129	12.5445	0.3306	0.4868	0.5884	-0.0152	0.2524	0.2529
2	1.8442	12.3403	12.4773	0.3239	0.4853	0.5834	-0.0166	0.2614	0.2619
3	1.9702	12.3800	12.5358	0.3135	0.4600	0.5567	-0.0177	0.2578	0.2584
4	2.1576	12.3608	12.5477	0.3015	0.4339	0.5284	-0.0186	0.2622	0.2629
5	2.7240	12.2028	12.5031	0.2738	0.3909	0.4772	-0.0197	0.2961	0.2967
6	3.4280	12.0074	12.4871	0.2466	0.3671	0.4423	-0.0207	0.3645	0.3651
7	3.7310	11.9045	12.4755	0.2358	0.3671	0.4364	-0.0211	0.4135	0.4141
8	3.8287	11.8803	12.4820	0.2326	0.3678	0.4352	-0.0212	0.4295	0.4301

H_x m			
n°	mean $\cdot 10^{-2}$	$\sigma_{H_y m}$ $\cdot 10^{-2}$	r.m.s. $\cdot 10^{-2}$
1	-0.0073	0.2062	0.2063
2	-0.0073	0.2140	0.2142
3	-0.0080	0.2094	0.2095
4	-0.0091	0.2129	0.2131
5	-0.0105	0.2447	0.2449
6	-0.0114	0.3105	0.3107
7	-0.0119	0.3589	0.3591
8	-0.0120	0.3749	0.3751

4. Inversion for shear-wave velocity

4.1 Group velocity and depth structure

Group and phase velocities of surface waves are functions of the speed and density profile of the top few hundreds of kilometers of the Earth. In particular, different frequencies sample different depth ranges, with longer periods sensitive to deeper structures. Consequently, surface wave dispersion anomalies provide useful information on the elastic properties at depth, and permit to set up an inverse problem where one retrieves crustal and upper mantle 3D structure from observed surface wave dispersion data. There are different approaches in literature to address this 3-D inverse problem. [*Marone et al.*, 2004] use the partitioned waveform inversion of [*Van der Lee and Nolet*, 1997], where the 1-D average S-velocity structure along each path is first determined by non-linear waveform fitting, and in a second step the 1-D path averaged structures are combined in a damped least-squares linear inversion for a 3-D S-velocity model. [*Shapiro and Ritzwoller*, 2002] in a first step estimate 2-D dispersion maps with a linear tomographic inversion of path average fundamental mode group and phase velocities, and afterward apply a Monte-Carlo method to perform the non-linear inversion of the dispersion curves at each geographical point and retrieve the 3-D shear-velocity model. [*Boschi and Ekström*, 2002] carry out a single non-linear inversion of phase anomaly measurements making use of JWKB ray-theory sensitivity kernels computed in a reference 3-D model. [*Zhou et al.*, 2006] invert long period fundamental mode phase delays with finite-frequency 3-D Born approximation kernels, calculated in a reference 1-D model. Our approach will consist in using the group velocity maps as data and to perform an iterative linearized depth inversion for the local 1-D structure, accounting for the lateral variations of the Crustal model.

4.1.1 Depth sensitivity kernels

The solution of the equation of motion for Love or Rayleigh waves in a vertically heterogeneous elastic medium bounded by a free-surface is an eigenvalue-eigenfunction problem, that can generally be solved through numerical methods such as Runge-Kutta integration. Another approach is to approximate the radially varying medium with a stack of isotropic homogeneous flat layers over an half-space and then apply the Thomson-Haskell

propagator matrix method. Finally, once eigenvalues and eigenfunctions are found, application of variational principle to the the energy integral permits to retrieve the phase $C(\omega)$ and group velocity $U(\omega)$ and their partial derivatives

$$\left.\frac{\partial C}{\partial m}(r)\right|_{\omega} \qquad \text{and} \qquad \left.\frac{\partial U}{\partial m}(r)\right|_{\omega}$$

with respect to the medium elastic parameter $m(r)$ (see Chapter 7 in [*Aki and Richards*, 2002] for an exposition in isotropic vertically heterogeneous media). Such partial derivatives that relate model elastic parameters to group (or phase) velocity are the so-called depth sensitivity kernels. For an isotropic, vertically heterogeneous Earth, characterized by density, P- and S-velocities $\rho(r), \alpha(r), \beta(r)$, the following integral expresses Rayleigh group velocity anomalies predicted by the perturbation of the 1-D model at the period T_i.

$$\delta U_R(T_i) = \int_0^R \Big[\frac{\partial U_i}{\partial \rho}(r)\,\delta\rho(r) + \frac{\partial U_i}{\partial \alpha}(r)\,\delta\alpha(r) + \frac{\partial U_i}{\partial \beta}(r)\,\delta\beta(r)\Big]dr \tag{4.1}$$

while for the Love group velocity there is no dependance from $\alpha(r)$.

$$\delta U_L(T_i) = \int_0^R \Big[\frac{\partial U_i}{\partial \rho}(r)\,\delta\rho(r) + \frac{\partial U_i}{\partial \beta}(r)\,\delta\beta(r)\Big]dr \tag{4.2}$$

However, surface waves are mostly sensitive to shear-velocity and provide only limited information on the density and P-wave velocity structure, so it is common to invert only for $\beta(r)$ and keep $\rho(r)$ and $\alpha(r)$ fixed in the model. In the solution of the inverse problem, we will compute partial derivatives and group velocity with the Haskell-Thomson propagator method [*Herrman*, 2005]. Hence, our model is locally approximated by a stack of flat, homogeneous and isotropic layers. Therefore we will approximate the integral over Earth radius in (4.1) and (4.2) with a discrete sum over M layers, each one of constant shear velocity m_j.

$$\delta U(T_i) \simeq \sum_{j=1}^{M} \frac{\partial U_i}{\partial m_j}\,\delta m_j \tag{4.3}$$

Kernels shape changes much with period: shorter periods kernels focus the sensitivity at shallower depths while longer periods illuminate deeper structures but their peak broadens and flattens out. In Figure 4.1 we show fundamental mode group velocity kernels for Rayleigh and Love wave as a function of depth at different periods computed for model PREM. The 35 s period Rayleigh kernel, for example, has a peak at about 45 km depth, turns negative below 80 km and is almost to zero at 200 km. Instead the 150 s kernel is maximum at about 160 km, has still a positive amplitude at 300 km

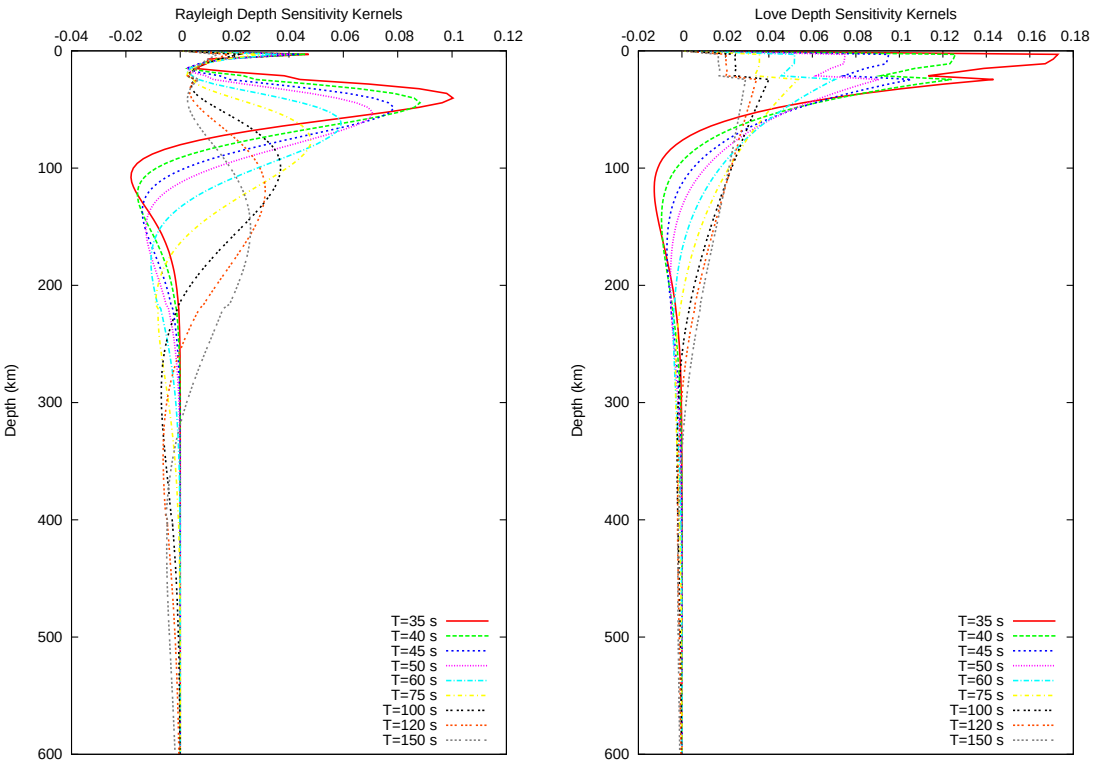


FIGURE 4.1: *Fundamental mode Rayleigh* (left) *and Love* (right) *depth sensitivity kernels as a function of depth at a set of different periods. The partial derivatives of group velocity with respect to shear velocity are computed in a layered flat model corrected for Earth Flattening that approximates the radially anisotropic model PREM.*

and a negative sensitivity down to 600 km. As a result, the simultaneous presence of positive and negative sensitivity at each depth in the Upper Mantle permits to limit the trade-off between structures in different layers in a joint inversion of different periods. These considerations allow us to expect a good resolution at least down to 250 km with the periods in our dataset, but the solution will depend also on the structure down to 500-600 km. Unlike Rayleigh, Love wave kernels are prominently sensitive to crustal and top mantle layers for all the periods, so are less able to discriminate between different depth structures. However, a joint inversion of Love and Rayleigh waves is crucial to retrieve a radially anisotropic model of Earth's Upper Mantle. .

4.2 Non-linear inversion

The linear Least-Squares or the Bayesian formulas used in the regionalization (see Chapther 3) can not be applied to solve the depth inversion problem, due to the non-linearity of the problem. The shear velocity kernels are in facts non-linear functions of the model parameters (i.e. they change when the model parameters change), therefore the forward problem is non linear

$$\delta U(T_i) \simeq \sum_{j=1}^{M} \frac{\partial U_i}{\partial m_j}\,\delta m_j \quad \Longrightarrow \quad \mathbf{d} = \mathbf{g}(\mathbf{m}) \neq \mathbf{G}\cdot\mathbf{m} \tag{4.4}$$

In case of a non-linear relation between parameters and data, $\sigma_M(m)$ in (3.11) is not Gaussian anymore and we must find an alternative formulation. This problem can be resolved with the method proposed by [*Tarantola and Valette*, 1982], which leads to an iterative algorithm that can easily converge if the non-linearity is not too strong. The solution is then given by three equivalent expressions:

$$\mathbf{m}_{n+1} = \begin{cases} \mathbf{m_n} + \Big(\mathbf{G_n^T C_d^{-1} G_n} + \mathbf{C_M^{-1}}\Big)^{-1}\Big(\mathbf{G_n^T C_d^{-1}}\big(\mathbf{g(m_n)} - \mathbf{d_{obs}}\big) + \mathbf{C_M^{-1}}\big(\mathbf{m_n} - \mathbf{m_0}\big)\Big) \\ \mathbf{m_0} - \Big(\mathbf{G_n^T C_d^{-1} G_n} + \mathbf{C_M^{-1}}\Big)^{-1}\mathbf{G_n^T C_d^{-1}}\Big(\big(\mathbf{g(m_n)} - \mathbf{d_{obs}}\big) - \mathbf{G_n}\big(\mathbf{m_n} - \mathbf{m_0}\big)\Big) \\ \mathbf{m_0} - \mathbf{C_M G_n^T}\Big(\mathbf{G_n C_M G_n^T} + \mathbf{C_d}\Big)^{-1}\Big(\big(\mathbf{g(m_n)} - \mathbf{d_{obs}}\big) - \mathbf{G_n}\big(\mathbf{m_n} - \mathbf{m_0}\big)\Big) \end{cases} \tag{4.5}$$

where $\mathbf{g(m_n)} = \mathbf{U(m_n)} = \mathbf{U_n}$ are the group velocities computed for model $\mathbf{m_n}$ and

$$\mathbf{G_n} = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{m}}\right)\Bigg|_{\mathbf{m_n}} = \frac{\partial \mathbf{U_n}}{\partial \mathbf{m_n}}$$

are the partial derivatives. Thus, at each iteration we compute with the Haskell-Thomson routines the group velocities and partial derivatives predicted by the starting model $\mathbf{m_n}$, and we invert minimizing the residuals with the observed data and keeping the solution close to the a priori reference model $\mathbf{m_0}$. Once a solution is found we use it as the new starting model, thus we calculate again group velocities and partial derivatives and iterate the inversion, until a stable condition is reached. At each step, however, the a priori condition continues to keep the model close to the reference model $\mathbf{m_0}$, to guarantee consistency with a priori information. As in the linear case, $\mathbf{C_d}$ and $\mathbf{C_M}$ are the covariance matrices that express our confidence on data and a priori model.

4.3 Transverse isotropy

Transverse isotropy (or radial anisotropy) is the property of a material in which the elastic stress-strain relations are invariant under rotations about a symmetry axis. In a

stratified Earth, where properties vary along the z-axis, the speeds of horizontally and vertically propagating P- and S-waves differ from the isotropic velocities $\alpha^2 = (\lambda+2\mu)/\rho$ and $\beta^2 = \mu/\rho$. A transverse isotropic medium is characterized by 5 independent elastic moduli A, C, F, L, N , using the notation of [*Love*, 1927]. A and C can be determined from measurements of horizontally and vertically propagating P-waves

$$A = \rho V_{PH}^2$$
$$C = \rho V_{PV}^2$$

while in general the shear speeds depend both on wave polarization and direction of propagation. For horizontally traveling wave with horizontal or vertical polarization

$$N = \rho V_{SH}^2$$
$$L = \rho V_{SV}^2$$

whereas for vertically propagating S-waves there is no splitting and the speed is controlled by L independently from the polarization. F is usually expressed by the dimensionless parameter $\eta = F/(A-2L)$ and is related to different angles of incidence of the propagating waves. In the isotropic case we have $A = C = (\lambda + 2\mu)$, $N = L = \mu$ and $\eta = 1$. Transverse isotropy is required to explain the observed Love-Rayleigh discrepancy and has been accounted for in the compilation of the global reference model PREM of [*Dziewonski and Anderson*, 1981], which exhibits a 2-4% radial anisotropy down to 220 km of depth. Love waves are sensitive only to N and L depth variations, while Rayleigh waves depend on all the 5 elastic moduli. A complete derivation of eigenfunctions and kernels for Love and Rayleigh waves is given by [*Takeuchi and Saito*, 1972]. Theoretically, it is thus possible to invert for all the 5 elastic parameters and density, but fundamental modes surface waves show only a limited sensitivity to A,C and F. Rayleigh waves respond mostly to L variations while Love waves to N, so we account for radial anisotropy with a joint inversion of Rayleigh and Love local dispersion curves to retrieve the V_{SV} and V_{SH} structure of the upper mantle. To simplify the calculations, we used averaged isotropic group velocity kernels $\partial U_L/\partial V_{SH}$ and $\partial U_R/\partial V_{SV}$ computed for anisotropic PREM (improved with a laterally varying crustal structure, as explained further in the text) and we keep P-wave speed fixed during the inversion. As [*Boschi and Ekström*, 2002] pointed out, petrological observations are not yet sufficiently numerous to determined a scaling factor relating P-and S-wave anomalies. We thus performed a test where we simply set $\delta V_{PH}/V_{PH} = \delta V_{SH}/V_{SH}$ and $\delta V_{PV}/V_{PV} = \delta V_{SV}/V_{SV}$ in each iteration, but we did not found any significant difference in the results, so we decide to keep them fixed.

4.4 Reference model and shallow layer correction

From the kernels displayed in Figure 4.1 it is evident that surface waves are strongly influenced by crustal structures. This effect is more evident at shorter periods and for Love waves, however even long period data can be affected by crustal thickness and velocity. Our data set is dominantly medium to long period and therefore it is not able to resolve variations in crustal layers. We need to introduce an a priori constraint on the crust to avoid any biased anomaly to be imaged into the upper mantle. We use at this scope the high resolution global crustal model CRUST2.0 of [*Bassin et al.*, 2000] (available on the web at http://mahi.ucsd.edu/Gabi/rem.html). This model is parameterized by a set of 1-D depth profiles on a 2°X2° global grid. Each column profile is divided into two layers of sediments and three layers of crystalline crust. Oceans and polar regions have water and ice layers. Topography and bathymetry of ETOPO5 are added. For each layer in each grid point thickness, isotropic P-and S-velocity and density are specified. In Figure 4.2 we show the European and Mediterranean crustal thickness map derived from CRUST2.0. Thickness vary mostly between ocean and continents, but also at a smaller scale length there are large differences that must be taken into account for a reliable inversion of surface waves. However, for a densely sampled region as Eurasia, we would like to be able to use an higher resolution model for crustal correction. At the moment, crustal structure in Europe takes advantage of many regional studies, but we feel that a consistent continental scale european model is still missing. There are two possible procedures to perform a crustal correction: one is to compute the predicted group velocity anomalies starting from a mantle model with a crustal model on top at every grid point and to integrate them along each ray path of the dataset. Afterwards, the path-averaged anomaly due to the crust is subtracted from the measured group velocities. Once the whole data set has been corrected it is possible to invert them for mantle structure. This scheme is used for instance by [*Silveira et al.*, 1998], [*Zhou et al.*, 2006], and [*Ekström and Dziewonski*, 1998]. In this approach the real three-dimensional crustal structure is averaged out along the path and the sensitivity kernels are those of a 1-D model (the mantle model). The second technique consists in taking the crustal structure into account directly as one of the models constraints. The observed measures are not corrected for shallow layers, they are instead inverted keeping the whole crust+mantle model as starting model. The crustal parameters (thickness and velocity) can be kept fixed during the inversion or can be allowed to some limited variation. As [*Boschi and Ekström*, 2002] remarked, this

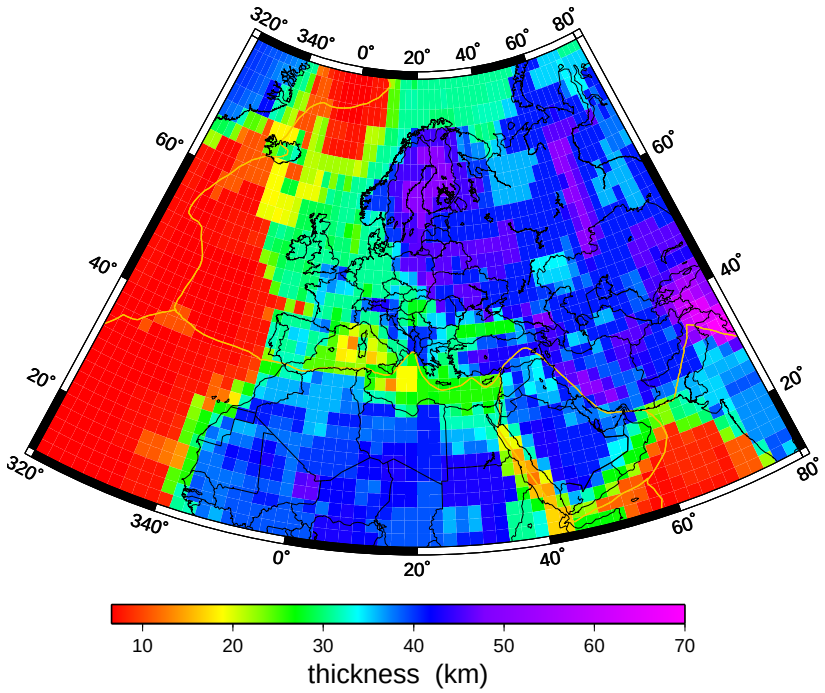


FIGURE 4.2: *Map of crustal thickness in km in the European region derived from model CRUST2.0*

procedure has the apparent drawback of being computationally more expensive, because it is necessary to recompute the sensitivity kernels and the theoretical group velocity at each grid point , but it has the advantage to use laterally varying sensitivity kernels, so the reference model is treated as fully 3-D.

4.5 Inversion

We parameterize the three-dimensional crust and upper mantle structure with a set of 80x80 depth profiles. Each profile corresponds to one cell of the previously derived regional group velocity model and the non-linear inversion is run separately for each grid point. For each cell we find the corresponding tile of CRUST2.0. We smoothed CRUST2.0 with a moving average technique, in order to avoid sharp lateral discontinuities. The reference model is locally approximated with a stack of flat layers, precisely the layers of CRUST2.0 (replacing with the right bathymetry and topography the crust of PREM) on top of model PREM, subdivided in a set of $\simeq$ 20 km layers, down to 771 km of depth. Beneath 771 km we consider an homogeneous half-space. Each mantle layer

is characterized by constant radially anisotropic elastic parameters $V_{PH}, V_{PV}, V_{SH}, V_{SV}$, and ρ calculated from model PREM at the depth corresponding to the centre of the layer. In each grid point we consider as observed data the group slowness derived from the group velocity maps at different periods. We prefer to invert group slowness **C** instead of velocities **U** because they are the real result of the tomographic inversion, and it is to slowness, and not to velocity, that the a posteriori error calculated in section 3.5.2 is related. Therefore, we set up an inverse problem where data are

$$d_i = C_i = \frac{1}{U_i} \qquad i = 1, \ldots, N \tag{4.6}$$

for the set of N periods and we transform consequently the group velocity partial derivatives for each one of the M layers into group slowness kernels

$$\frac{\partial U_i}{\partial m_j} \Longrightarrow \frac{\partial C_i}{\partial m_j} = -\frac{1}{U_i^2}\frac{\partial U_i}{\partial m_j} \qquad j = 1, \ldots, M \tag{4.7}$$

As previously explained, we use both Rayleigh and Love data to inspect the transverse isotropy of the upper mantle. A natural choice of the parameters to image should then fall on V_{SV} and V_{SH}. However, when variations in anisotropy are large, such as in the Pacific upper mantle, neither V_{SV} nor V_{SH} alone reflect thermal anomalies [*Ekström and Dziewonski*, 1998]. Instead, the Voigt average $V_{Voigt} = (2/3V_{SV} + 1/3V_{SH})$ of the two quantities is better related to isotropic variations correlated to temperature. So we choose to invert directly for Voigt average V_{Voigt} and radial anisotropy $V_{ani} = (V_{SH} - V_{SV})$, which are the two parameters which we are interested in. This approach has the advantage of allowing a direct estimation of the isotropic and anisotropic part with a coupled inversion of Love and Rayleigh waves, as well as more control on the anisotropy amplitude through the a priori covariance matrices. We calculate the isotropic partial derivatives $\partial U_L/\partial V_{SH}$ and $\partial U_R/\partial V_{SV}$ [*Herrman*, 2005], which need to be transformed into partial derivatives with respect to V_{Voigt} and V_{ani}.

$$\begin{cases} V_{Voigt} = (\frac{2}{3}V_{SV} + \frac{1}{3}V_{SH}) \\ V_{ani} = (V_{SH} - V_{SV}) \end{cases} \Longrightarrow \begin{cases} V_{SV} = (V_{Voigt} - \frac{1}{3}V_{ani}) \\ V_{SH} = (V_{Voigt} + \frac{2}{3}V_{ani}) \end{cases} \tag{4.8}$$

Hence the partial derivatives are

$$\begin{aligned}
\frac{\partial U_R}{\partial V_{Voigt}} &= \frac{\partial U_R}{\partial V_{SV}}\frac{\partial V_{SV}}{\partial V_{Voigt}} = \frac{\partial U_R}{\partial V_{SV}} \\
\frac{\partial U_R}{\partial V_{ani}} &= \frac{\partial U_R}{\partial V_{SV}}\frac{\partial V_{SV}}{\partial V_{ani}} = -\frac{1}{3}\frac{\partial U_R}{\partial V_{SV}} \\
\frac{\partial U_L}{\partial V_{Voigt}} &= \frac{\partial U_L}{\partial V_{SH}}\frac{\partial V_{SH}}{\partial V_{Voigt}} = \frac{\partial U_L}{\partial V_{SH}} \\
\frac{\partial U_L}{\partial V_{ani}} &= \frac{\partial U_L}{\partial V_{SH}}\frac{\partial V_{SH}}{\partial V_{ani}} = +\frac{2}{3}\frac{\partial U_L}{\partial V_{SH}}
\end{aligned} \tag{4.9}$$

Summarizing, we have a 2xN data vector $\mathbf{d}$ made up of N group slowness observations for Love and N for Rayleigh, and a 2xM model vector $\mathbf{m}$ consisting of M Voigt average $\mathbf{m}_V$ and M radial anisotropy $\mathbf{m}_A$ parameters.

$$\begin{aligned}
\mathbf{d^T} &= \begin{pmatrix} C_{L_1}, & C_{L_2}, & \cdots & C_{L_N}, & C_{R_1}, & C_{R_2}, & \cdots & C_{R_N} \end{pmatrix} \\
\mathbf{m^T} &= \begin{pmatrix} m_{V_1}, & m_{V_2}, & \cdots & m_{V_M}, & m_{A_1}, & m_{A_2}, & \cdots & m_{A_M} \end{pmatrix}
\end{aligned} \tag{4.10}$$

Thus, the 2Nx2M partial derivatives matrix $\mathbf{G}$ is

$$\mathbf{G} = \begin{pmatrix}
-\frac{1}{U_{L_1}^2}\frac{\partial U_{L_1}}{\partial m_{V_1}} & \cdots & -\frac{1}{U_{L_1}^2}\frac{\partial U_{L_1}}{\partial m_{V_M}} & -\frac{1}{U_{L_1}^2}\frac{\partial U_{L_1}}{\partial m_{A_1}} & \cdots & -\frac{1}{U_{L_1}^2}\frac{\partial U_{L_1}}{\partial m_{A_M}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
-\frac{1}{U_{L_N}^2}\frac{\partial U_{L_N}}{\partial m_{V_1}} & \cdots & -\frac{1}{U_{L_N}^2}\frac{\partial U_{L_N}}{\partial m_{V_M}} & -\frac{1}{U_{L_N}^2}\frac{\partial U_{L_N}}{\partial m_{A_1}} & \cdots & -\frac{1}{U_{L_N}^2}\frac{\partial U_{L_N}}{\partial m_{A_M}} \\
-\frac{1}{U_{R_1}^2}\frac{\partial U_{R_1}}{\partial m_{V_1}} & \cdots & -\frac{1}{U_{R_1}^2}\frac{\partial U_{R_1}}{\partial m_{V_M}} & -\frac{1}{U_{R_1}^2}\frac{\partial U_{R_1}}{\partial m_{A_1}} & \cdots & -\frac{1}{U_{R_1}^2}\frac{\partial U_{R_1}}{\partial m_{A_M}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
-\frac{1}{U_{R_N}^2}\frac{\partial U_{R_N}}{\partial m_{V_1}} & \cdots & -\frac{1}{U_{R_N}^2}\frac{\partial U_{R_N}}{\partial m_{V_M}} & -\frac{1}{U_{R_N}^2}\frac{\partial U_{R_N}}{\partial m_{A_1}} & \cdots & -\frac{1}{U_{R_N}^2}\frac{\partial U_{R_N}}{\partial m_{A_M}}
\end{pmatrix} \tag{4.11}$$

Since we have 2xN=26 data and 2xM$\simeq$100 model parameters, we choose to use the third of (4.5), because the $(G_n C_M G_n^T + C_d)$ matrix is easier to invert with respect to $(G_n^T C_d^{-1} G_n + C_M^{-1})$. We construct a data covariance matrix C_D in the form of a diagonal matrix with the a posteriori error on slowness calculated for each period in Section 3.5.2. We therefore neglect the correlation between different periods. The a priori model covariance matrix C_M is a diagonal matrix with the same errors associated to all the layers. We tried different values of the two errors σ_V and σ_A, that acts like damping parameters, but finally we chose $\sigma_V = \sigma_A = 0.12$ km/s, which correspond to a 2-3% error on isotropic shear velocity and 1-5% on anisotropy.

4.6 Results

We show in Figures 4.3, 4.4, 4.5, 4.6, 4.7 and 4.8 our resulting shear velocity model SM07. Voigt average V_S anomalies are plotted as percentage variation with respect to PREM. Each map is plotted after having removed the mean value.

The most dominant large scale feature of SM07 over a wide range of depths is the difference between the fast Precambrian East European Platform and the low velocity signature of Central Europe and Western Mediterranean, in good accord with imaged with previous works [*Marquering and Snieder*, 1996; *Shapiro and Ritzwoller*, 2002; *Boschi et al.*, 2004]. The fast roots of the East European and Baltic Shields are visible at least to 250 km of depth, which is our maximum resolvable depth. The extensional basin of Western Mediterranean is marked by a strong slow anomaly, particularly well correlated at shallow depths with the Tyrrhenian back arc region, interpreted as the evidence of an asthenospheric wedge [*Mele et al.*, 1998], and the rifting area of the Sicily Channel. We observe only a weak and not well resolved fast anomaly in Southern Italy at depths greater than 120 km possibly related to the subduction of the Ionian lithosphere beneath the Calabrian arc [*Spakman et al.*, 1993; *Piromallo and Morelli*, 2003; *Lucente et al.*, 1999]. The Alpine-Himalayan collision belt is also evidenced well on the 60 km map by a continuos low velocity band that runs from the Carpathians through Anatolia and Zagros belt to Hindu Kush. The sharp gradient of the boundary between the East and Central Europe, running along the ancient Tornquist-Tesseyre Zone is impressively well delineated in our maps, at least down to 240 km depth. At a smaller scale, in the 60 km map we can recognize a low velocity anomaly marking the Pannonian Basin. Well represented at least down to 160 km is also the European Cenozoic Rift System, with low speed anomalies beneath the Eifel region and Rhine Graben. Eastern Mediterranean is instead characterized by overall high velocities. The Hellenic Arc anomaly, generally interpreted as old Tethydean oceanic crust trapped in the collision between Africa and Eurasia and subducting beneath the Aegean [*Spakman et al.*, 1993], is clearly visible in our maps, justapposed to a shallow slow anomaly in correspondence to the Aegean back arc basin. In Africa the main features are the fast West Africa Craton, and the slow Afar triple junction and Red Sea Rift. There are also traces of thermal anomalies in the rifting zone of Central Africa, maybe related to Tibesti, Hoggar and Darfur hotspots, but our resolution does not permit to individuate such narrow features. At depths greater than 160 km we observe also a narrow fast band running SW-NE , imaged also by [*Sebai et al.*, 2006] and [*Ritsema et al.*, 2000]. Eastern Atlantic Ocean is

characterized by a strong slow anomaly following the Ridge, particularly pronounced at shallow depths beneath Iceland and Azores Triple Junction, while the plate boundary between Azores and the Iberian Peninsula has higher velocities. The low velocity anomaly becomes wider and looses intensity with increasing depth, including the whole oceanic basin. Similar images of the Atlantic Ocean Upper Mantle have been obtained by [*Silveira and Stutzmann*, 2002] and [*Marone et al.*, 2004].

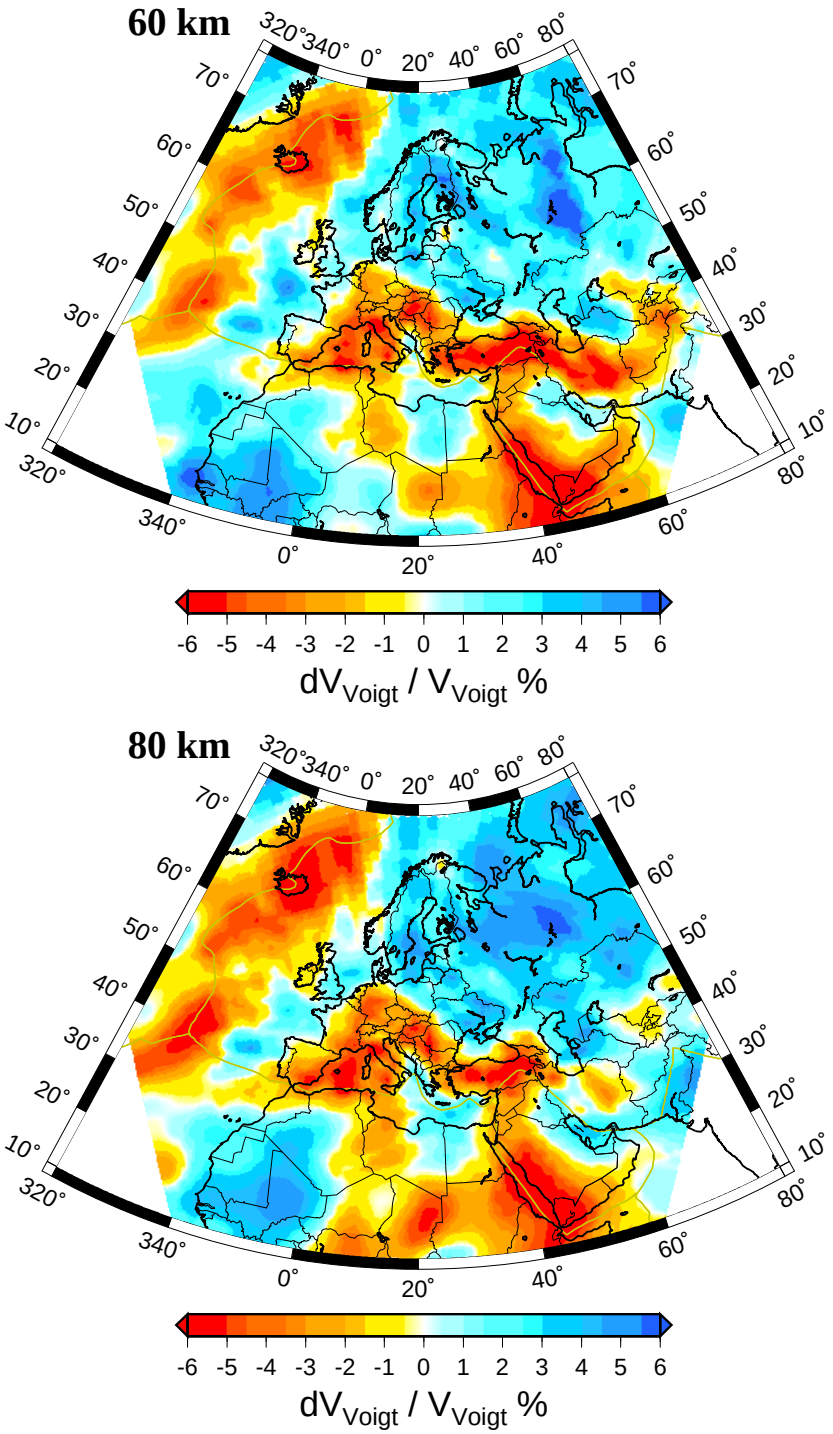


FIGURE 4.3: *Voigt average* $V_{Voigt} = (\frac{2}{3}V_{SV} + \frac{1}{3}V_{SH})$ *percentage anomaly at different depths. Color scale represents relative shear velocity anomaly with respect to PREM value. Map means have been removed before plotting.*

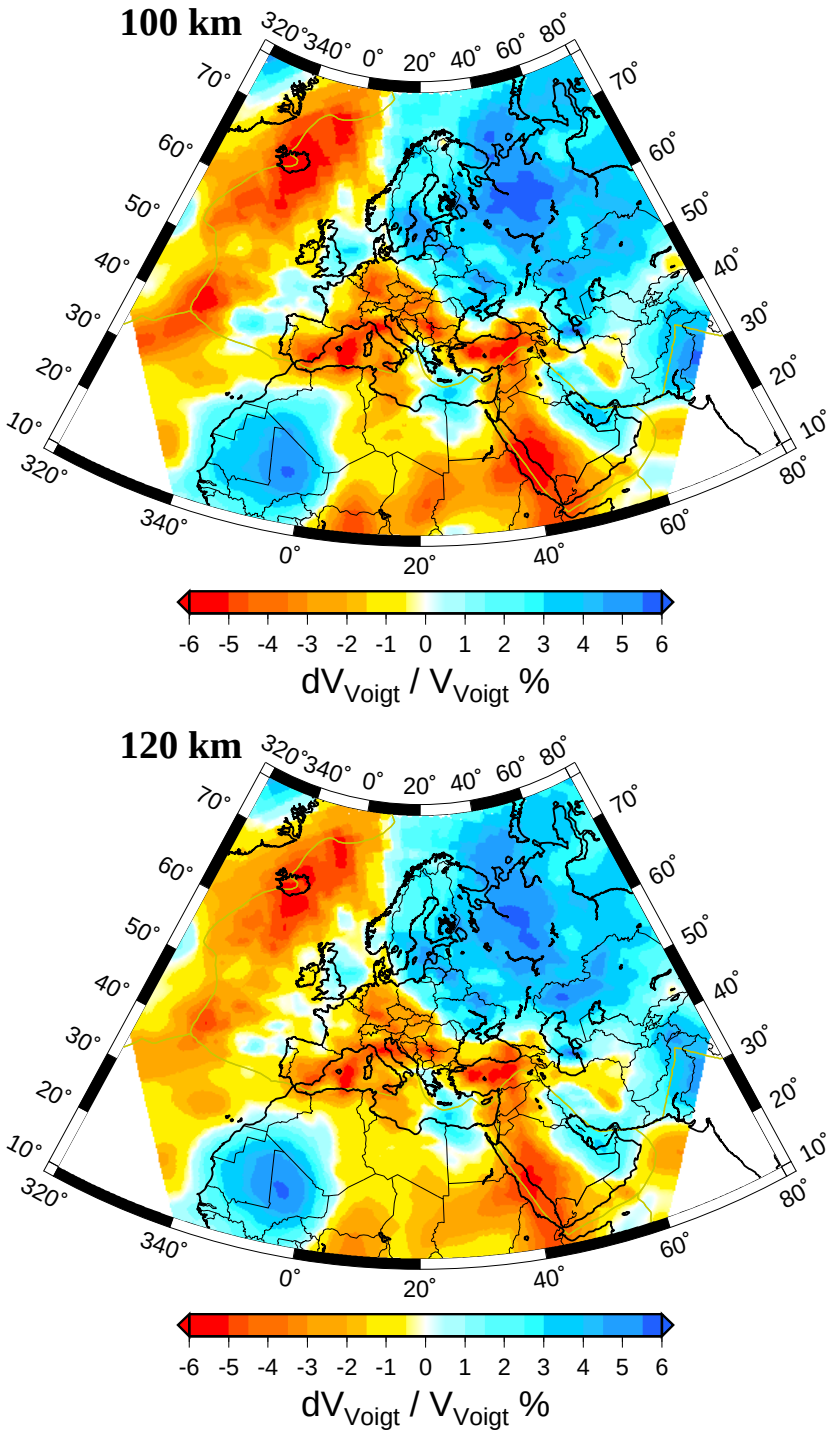


FIGURE 4.4: *Voigt average* $V_{Voigt} = (\frac{2}{3}V_{SV} + \frac{1}{3}V_{SH})$ *percentage anomaly at different depths. Color scale represents relative shear velocity anomaly with respect to PREM value. Map means have been removed before plotting.*

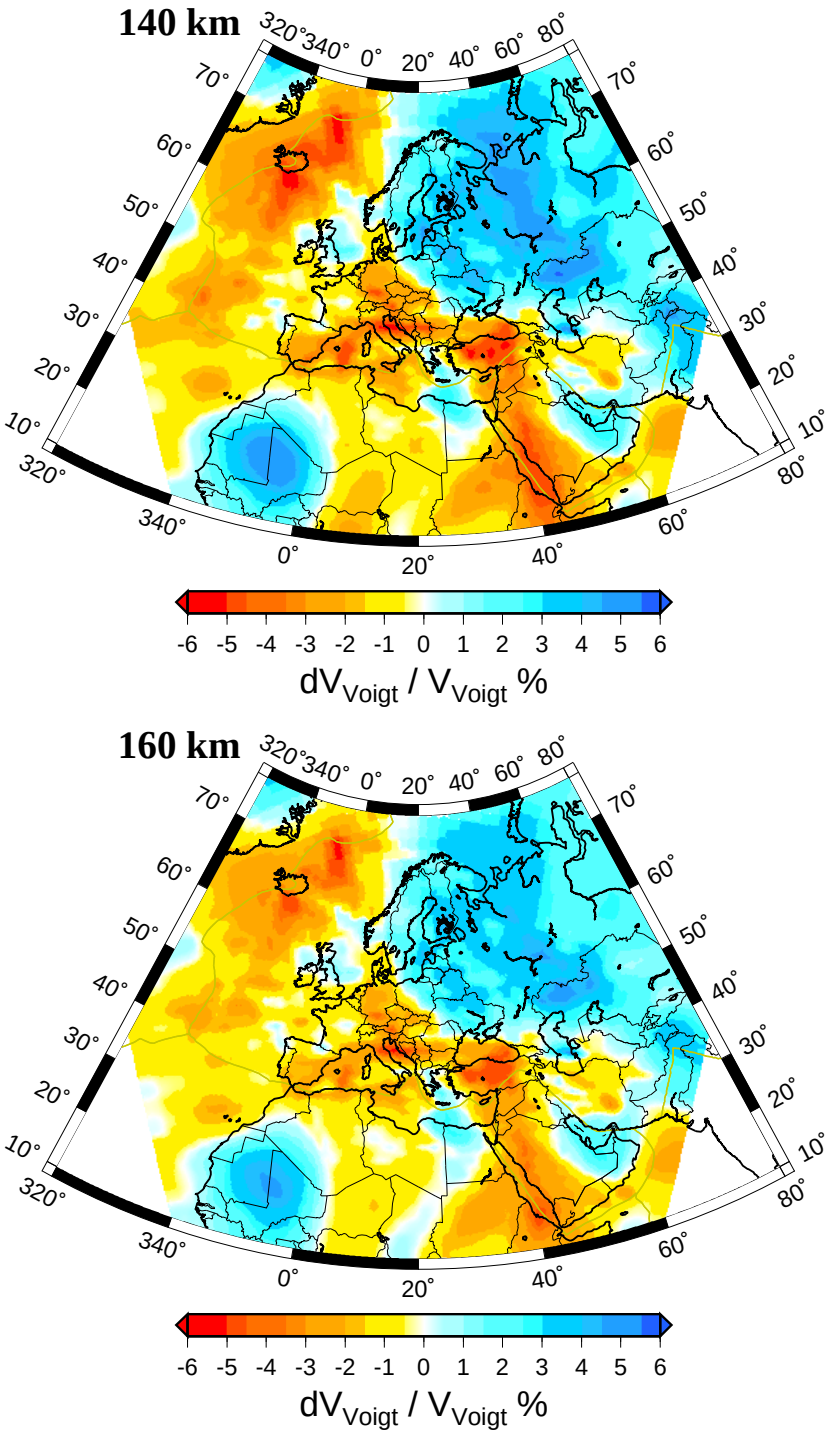


FIGURE 4.5: *Voigt average* $V_{Voigt} = (\frac{2}{3}V_{SV} + \frac{1}{3}V_{SH})$ *percentage anomaly at different depths. Color scale represents relative shear velocity anomaly with respect to PREM value. Map means have been removed before plotting.*

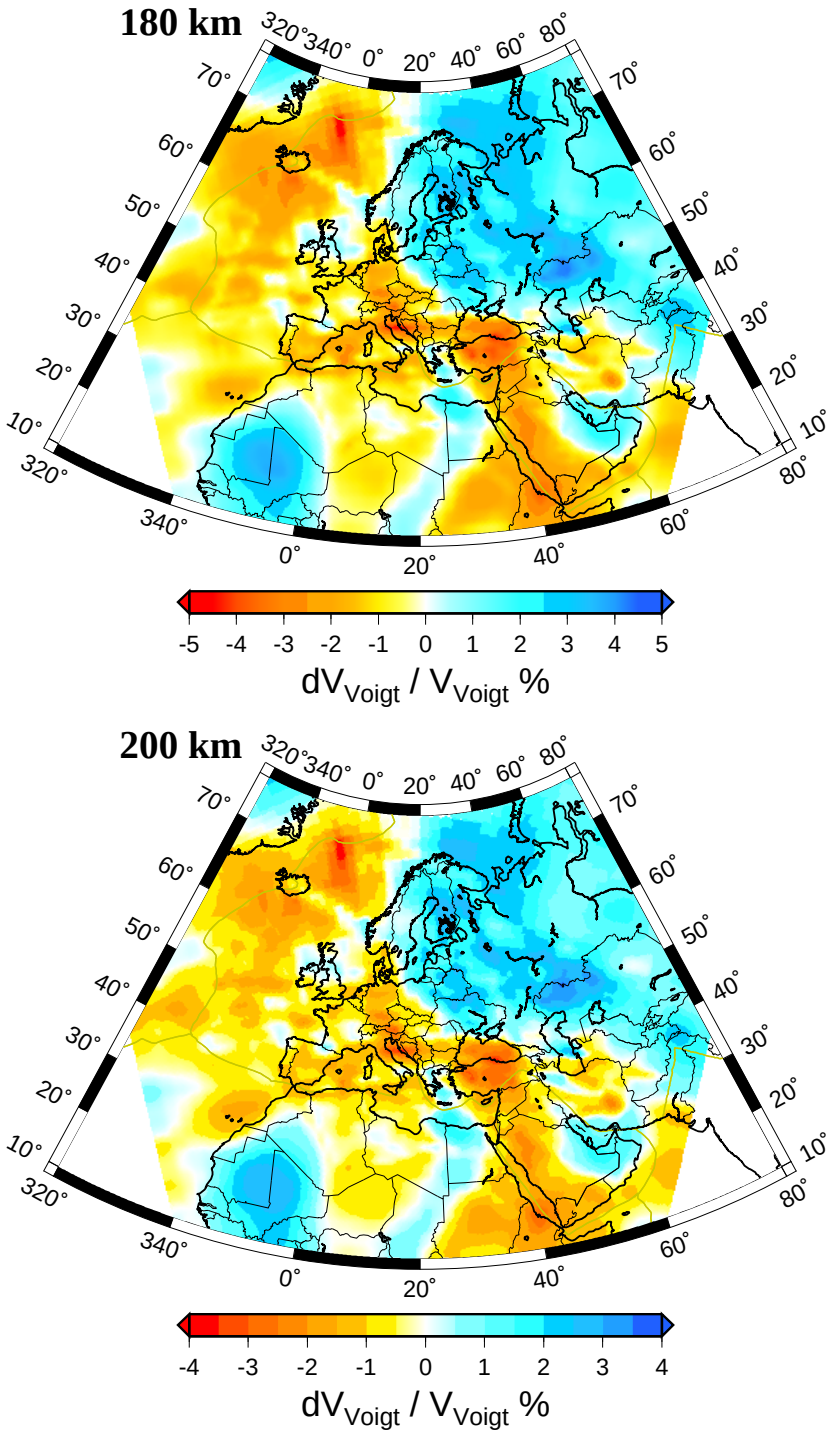


FIGURE 4.6: *Voigt average* $V_{Voigt} = (\frac{2}{3}V_{SV} + \frac{1}{3}V_{SH})$ *percentage anomaly at different depths. Color scale represents relative shear velocity anomaly with respect to PREM value. Map means have been removed before plotting.*

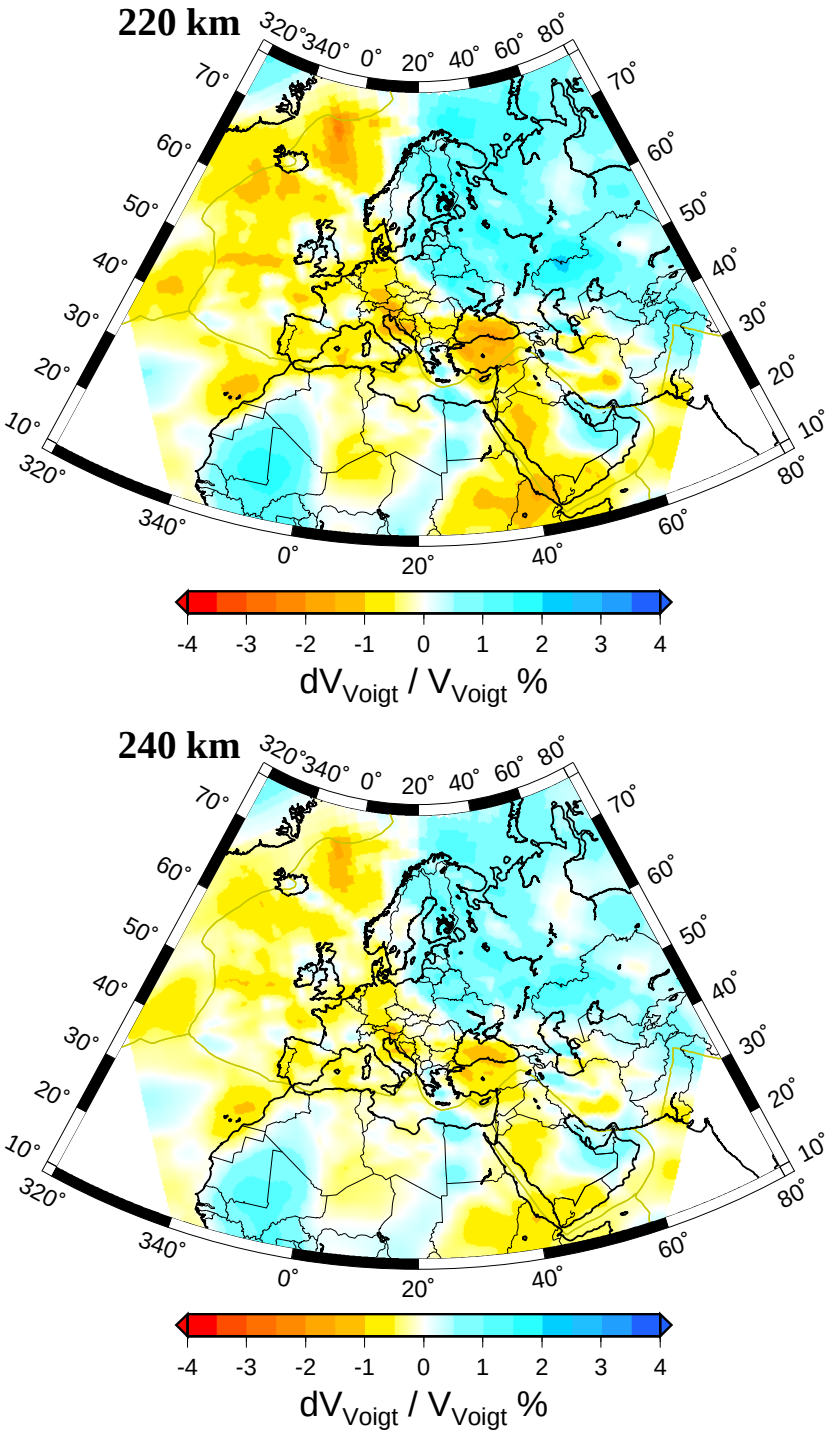


FIGURE 4.7: *Voigt average* $V_{Voigt} = (\frac{2}{3}V_{SV} + \frac{1}{3}V_{SH})$ *percentage anomaly at different depths. Color scale represents relative shear velocity anomaly with respect to PREM value. Map means have been removed before plotting.*

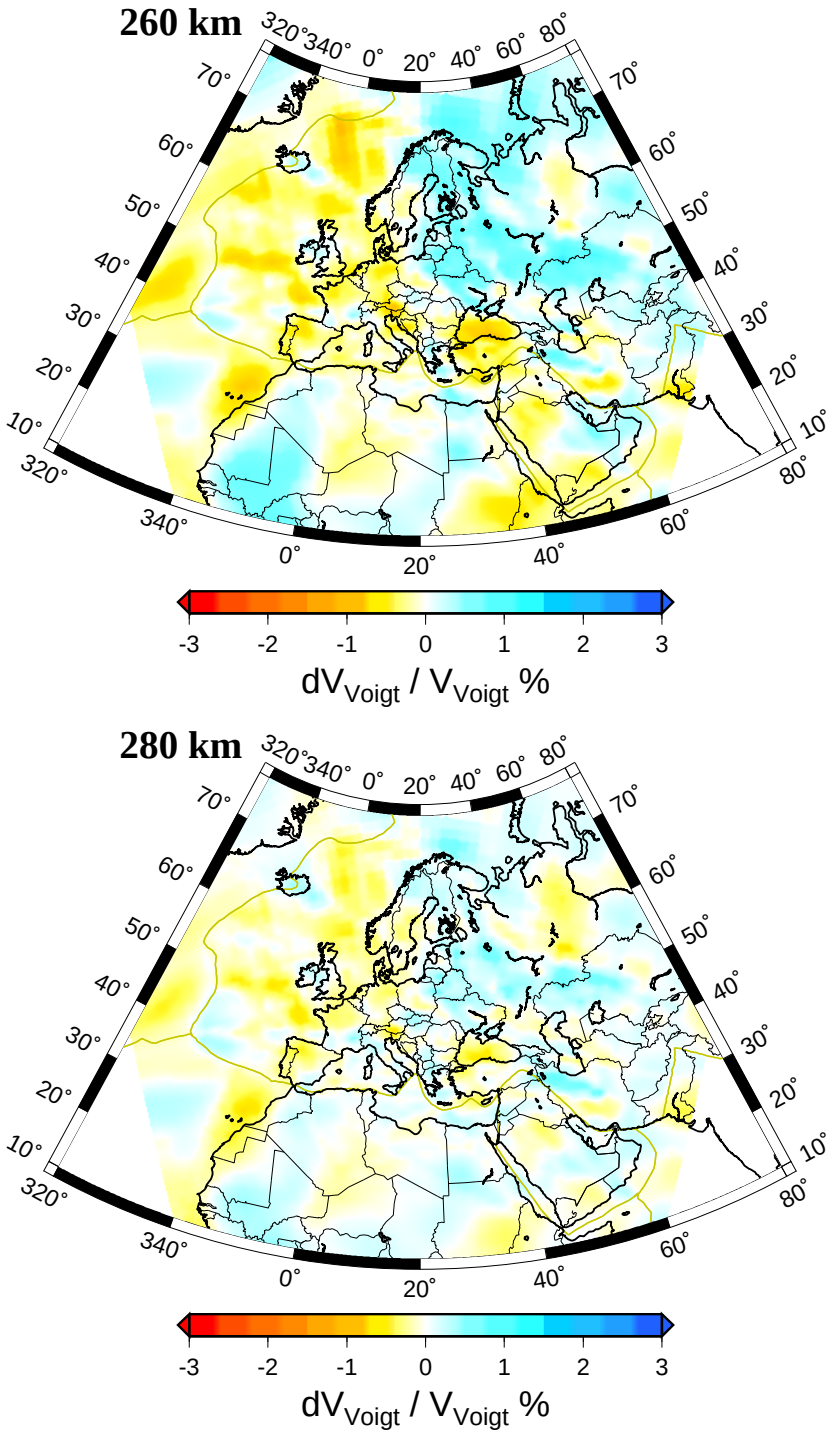


FIGURE 4.8: *Voigt average* $V_{Voigt} = (\frac{2}{3}V_{SV} + \frac{1}{3}V_{SH})$ *percentage anomaly at different depths. Color scale represents relative shear velocity anomaly with respect to PREM value. Map means have been removed before plotting.*

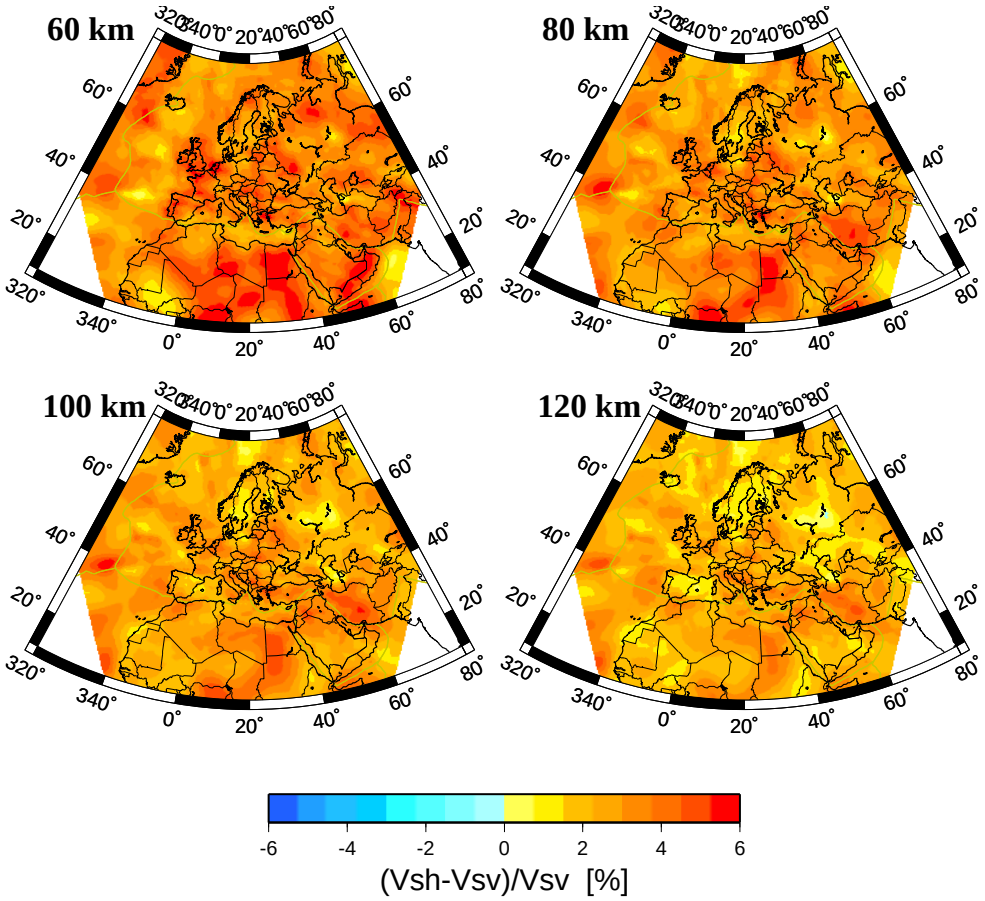


FIGURE 4.9: *Transverse isotropy in model SM07 expressed as percentage* $(V_{SH} - V_{SV})/V_{SV}$.

4.6.1 Radial anisotropy

As previously mentioned, we carried out the inversion of Rayleigh and Love dispersion curves for the anisotropic structure of the Upper Mantle, keeping the difference between V_{SH} and V_{SV} close to the PREM value. In Figure 4.9 we show maps of $(V_{SH} - V_{SV})/V_{SV}$ radial anisotropy at four depths in SM07. Our results confirm a moderate radial anisotropy over the wide European and Mediterranean area , generally close to the PREM value of 4% at the top of the mantle and decreasing with depth. We do not observe significant negative anisotropy, though some areas have values lower than those of PREM. Strong positive ($V_{SH} \gg V_{SV}$) anomalies are found in Central Africa and West of Red Sea, in the Azores and in Iran. Lower values are observed in correspondence of more stable tectonic features, such as West African Craton at 60 km

depth, and Eastern European Platform from 100 km depth. Most lateral variations, however, do not show such a remarkable long spatial wavelength and are likely to be the result of the different quality and coverage of Rayleigh and Love wave observations, the first being sensible to V_{SV}, the second to V_{SH} [*Shapiro and Ritzwoller*, 2002; *Boschi et al.*, 2004]. Anisotropy in Upper Mantle, though evidenced by a number of independent observations (e.g. from receiver function amplitude versus azimuth and SKS shear-wave splitting) and widely accepted, is still particularly difficult to resolve, and shows the need of further seismological research.

4.6.2 Comparison with global and regional models

We compare **SM07** with other recently published global and regional models:

- global model **S20RTS** of [*Ritsema et al.*, 1999]. This is a degree 20 isotropic shear wave velocity model of Earth's Mantle down to the CMB. It has been derived from fundamental and higher modes Rayleigh phase velocity measurements, teleseismic traveltimes and normal mode splitting data. (Model parameters are available on the web at http://oshi.geo.lsa.umich.edu/ jritsema/Research.html)

- global model **BEK04** of [*Boschi et al.*, 2004]. This is a global radially anisotropic model of the Upper Mantle derived from inversion of fundamental mode Rayleigh and Love phase anomalies (mainly the ETL97 dataset expanded with longer periods measurements). It is parameterized with a variable resolution grid, which is more dense on the Mediterranean basin. (Model parameters are available on the web at http://www.spice-rtn.org/research/planetaryscale/tomography/).

- the global model **CU_SDT1.0** of [*Shapiro and Ritzwoller*, 2002]. This is a global transversely isotropic Upper Mantle model, derived from the inversion of global fundamental mode phase velocity measurements of [*Trampert and Woodhouse*, 1995] and ETL97, as well as regional and teleseismic fundamental mode group velocity observations in Eurasia of [*Ritzwoller and Levshin*, 1998]. (Model parameters are available on the web at http://ciei.colorado.edu/ nshapiro/MODEL/).

- Eurasia-Africa plate boundary region model **EAV03** of [*Marone et al.*, 2004]. This is an isotropic regional model of the Upper Mantle derived from waveform fitting of regional S-wave arrivals and Rayleigh surface waves.

In Figure 4.10 we show SM07, S20RTS, BEK04 and CU_SDT1.0 at 100 km depth plotted on the same projection and with the same color scale ($\pm$ 6% max amplitude). Model

EAV03 is shown with a different color scale, corresponding to a maximum amplitude of $\pm$ 15%. The known large scale features of Eurasia and Mediterranean area reported in global models are extremely well imaged by SM07: the large high velocity anomalies beneath the cratons of Eastern Europe (Precambrian Baltic shield and East European) and West Africa, the low velocity related to the Atlantic Mid Ocean Ridge and Red Sea and the high velocity in Eastern Mediterranean mirrored by the slow anomaly in Western Mediterranean. Nevertheless, SM07 achieves higher resolution with respect to global models, comparable maybe only with the resolution of CU_SDT1.0 or EAV03, which derive also from regional measurements in this area. With respect to EAV03, SM07 present a more coherent pattern of anomalies, deriving probabily from our better ray coverage. Our model anyway show features that are not imaged neither in CU_SDT1.0, such as the Rhine Graben slow speed, the Southern Caspian Block slower than Northern, an unprecedented imaging of the Ashgabat fault, and the slower anomaly in Western Mediterranean in correspondence with Magmatic and extensional area of Thyrrenian sea. Tournquist Tesseyre Zone has also never been imaged so sharp as in our model. At very shallow depth Northern Appenines show a very slow speed, in disaccord with previous studies. This could be related to the complexity of subduction of the Adriatic plate under Thyrrenian, but a complete geodynamical interpretation of these results is beyond the scope of this thesis, and will be given elsewhere.

Figure 4.11 shows the same comparison at 200 km depth, where our model has a considerably smaller amplitude with respect to other models. This results from the use of only fundamental surface wave data in period range 35-170 s, which provide an optimal resolution for the uppermost mantle, but loose discrimination power at greater depths, that are better illuminated by longer periods, higher modes or body waves data. We found our model in good agreement with global models in imaging the fast roots of the continents (Eastern Europe and West African Cratons), the widening of the low velocity anomaly around Mid Atlantic Ridge and the slower Western Mediterranean. However our model exhibits new features: a much thinner fast anomaly related to the Hellenic slab and the Eastern European Craton still well separated from the slow active tectonic region of Western Europe: other models see these two fast anomalies (craton and hellenic slab) joined together: but they are related to two different geological features, cratons and old oceanic crust trapped during the closure of Tethyde. Moreover, we continue to see the strong gradient between fast Northern and Southern Caspian Sea , dividev by the Ashgabat fault.

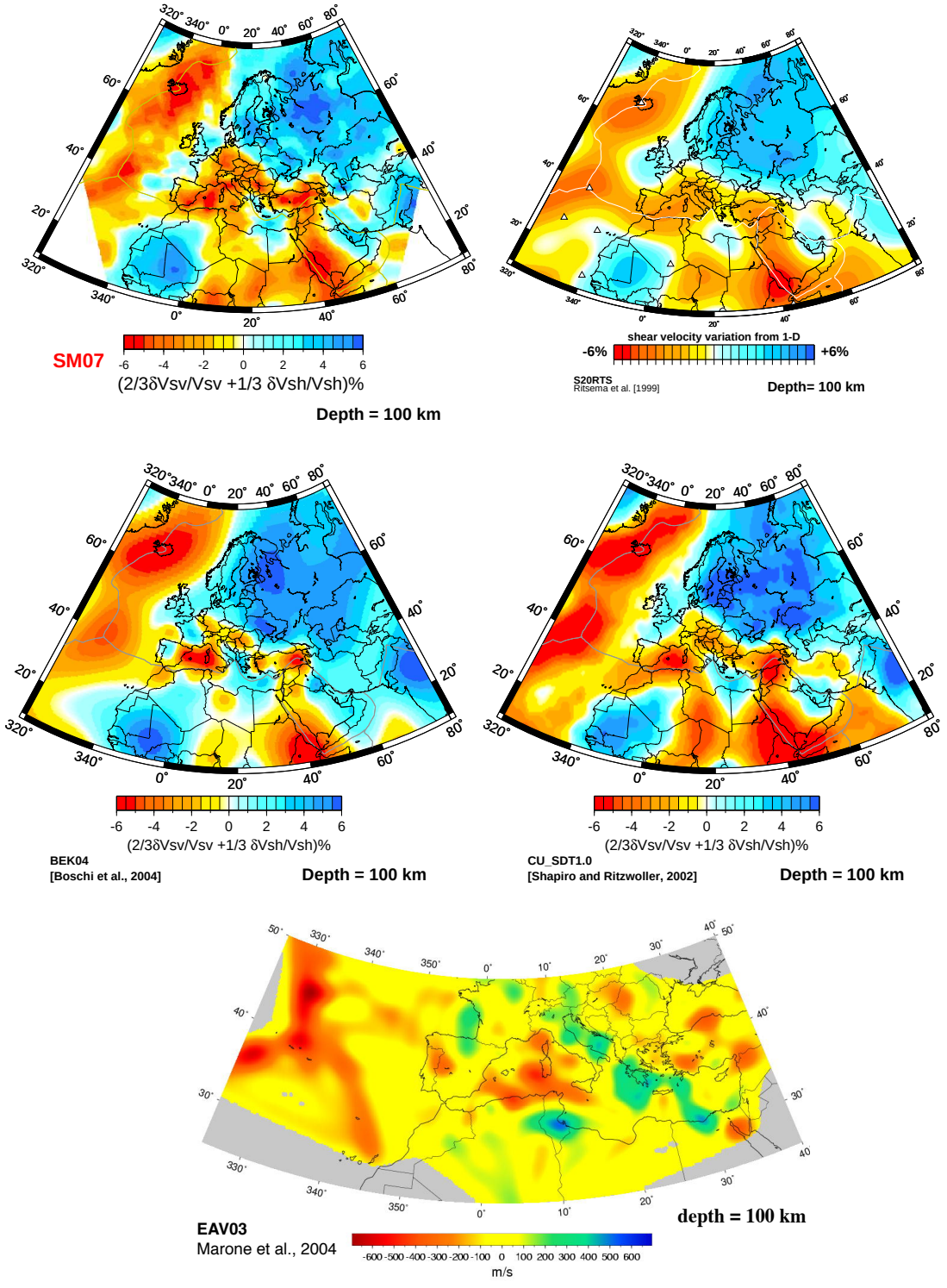


FIGURE 4.10: *Comparison between SM07* (top left) *and other global and regional models at a depth of 100 km: global model SM07, BEK04 and CU_SDT1.0 display the Voigt average anomaly, while models S20RTS are EAV03 isotropic models.*

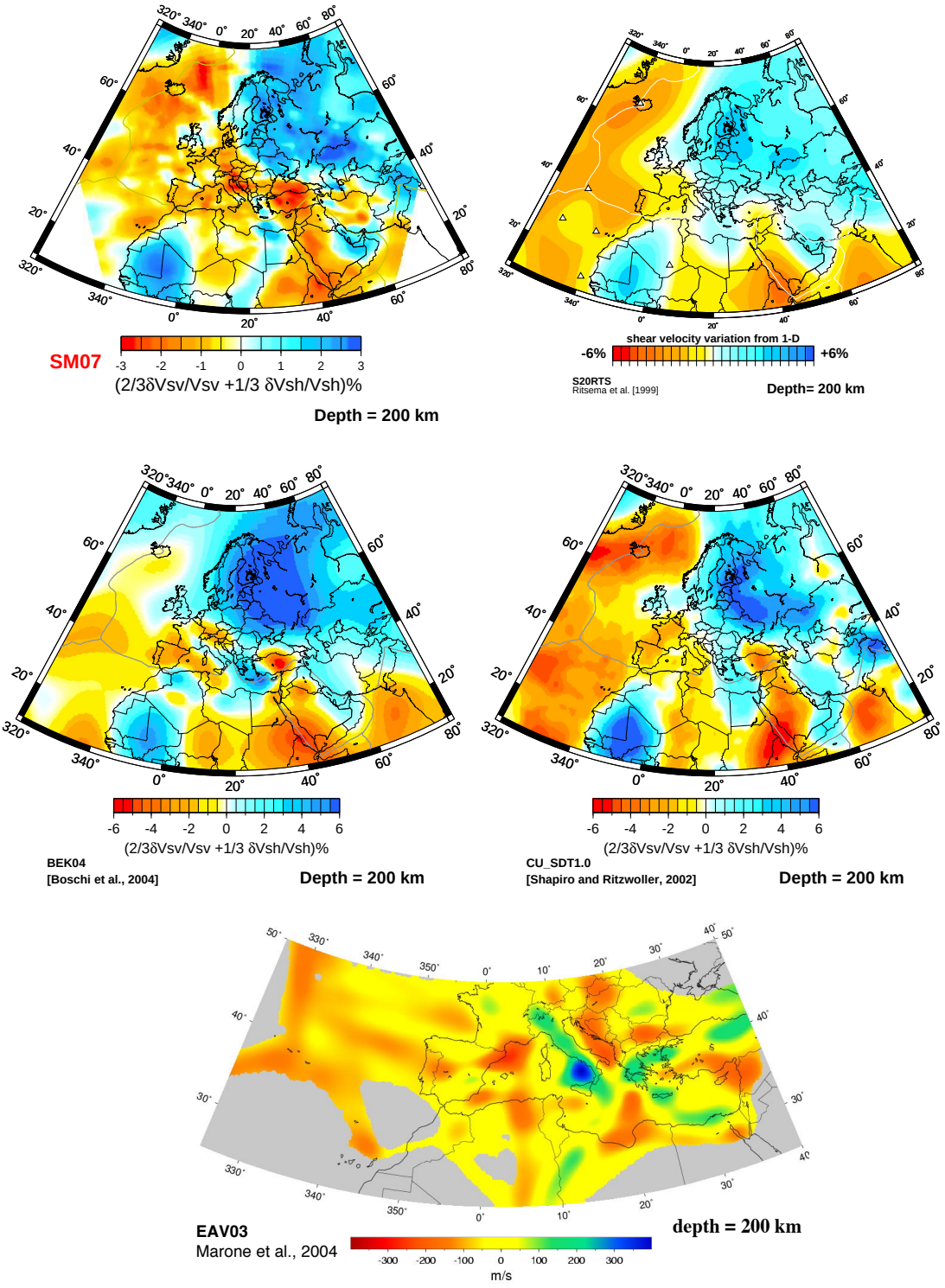


FIGURE 4.11: *Comparison between SM07* (top left) *and other global and regional models at a depth of 200 km.*

4.6.3 How well do we know the seismic structure of the European Upper Mantle?

To investigate the quality of our knowledge of subsurface structure, we may wonder how well results from different modeling approaches, with different sources of bias and indetermination, compare. We chose to compare therefore results from perhaps the two most distant modeling approaches:

- travel time tomography, sensitive to the P wave speed structure, based on inversion of high-frequency, hand-picked, first-arrival times retrieved from Bulletins
- surface wave tomography, sensitive to S wave structure, based on two-stage inversion of measurements made on intermediate and long period waveforms

We show in Figure 4.12 our model SM07 compared with a filtered version of P-wave model PM0.5 of [*Piromallo and Morelli*, 2003]. PM0.5 is a European and Mediterranean model derived from fitting travel times retrieved from the ISC Bulletins for regional and teleseismic rays. Travel times of high-frequency P waves ($\sim$ 1 Hz) can image structures with high resolution. PM0.5 shows many features with geodynamic relevance, such as traces of collisional belts, back-arc extensional basins, cenozoic volcanism. Data coverage at shallower depths is much influenced by the distribution of seismic sources and observatories, and leave peripheral areas blank. Surface waves reach a much more uniform spatial coverage even where epicenters and stations are sparse. However, because of their long wavelength they have limited spatial resolution. A comparison of our model with high-resolution tomography can thus be misleading, so we choose to compare it with a low-passed version of the P wave speed model, simply obtained with a horizontal boxcar filter with 3° diameter. The qualitative resemblance shown by the visual comparison is very good. The scale of variation is approximately double for shear wave than for P speed, as expected, and the sign of anomalies correlates well in many regions. Worth noting are the consistent low speeds of the Alpine tectonic belt, the Western Mediterranean and Central Europe. As well, there is good correlation with fast East European platform, Eastern Mediterranean, Ionian Sea. There are of course reasons why relative P and S speed variations can be uncorrelated (e.g. chemical heterogeneity, presence of fluids,...) and significant geodynamic considerations may follow ([*Trampert and Van der Hilst*, 2005]); but before a truly significant quantitative analysis can be carried on we need to ascertain in detail which components in model spaces are well constrained by data, and which are not.

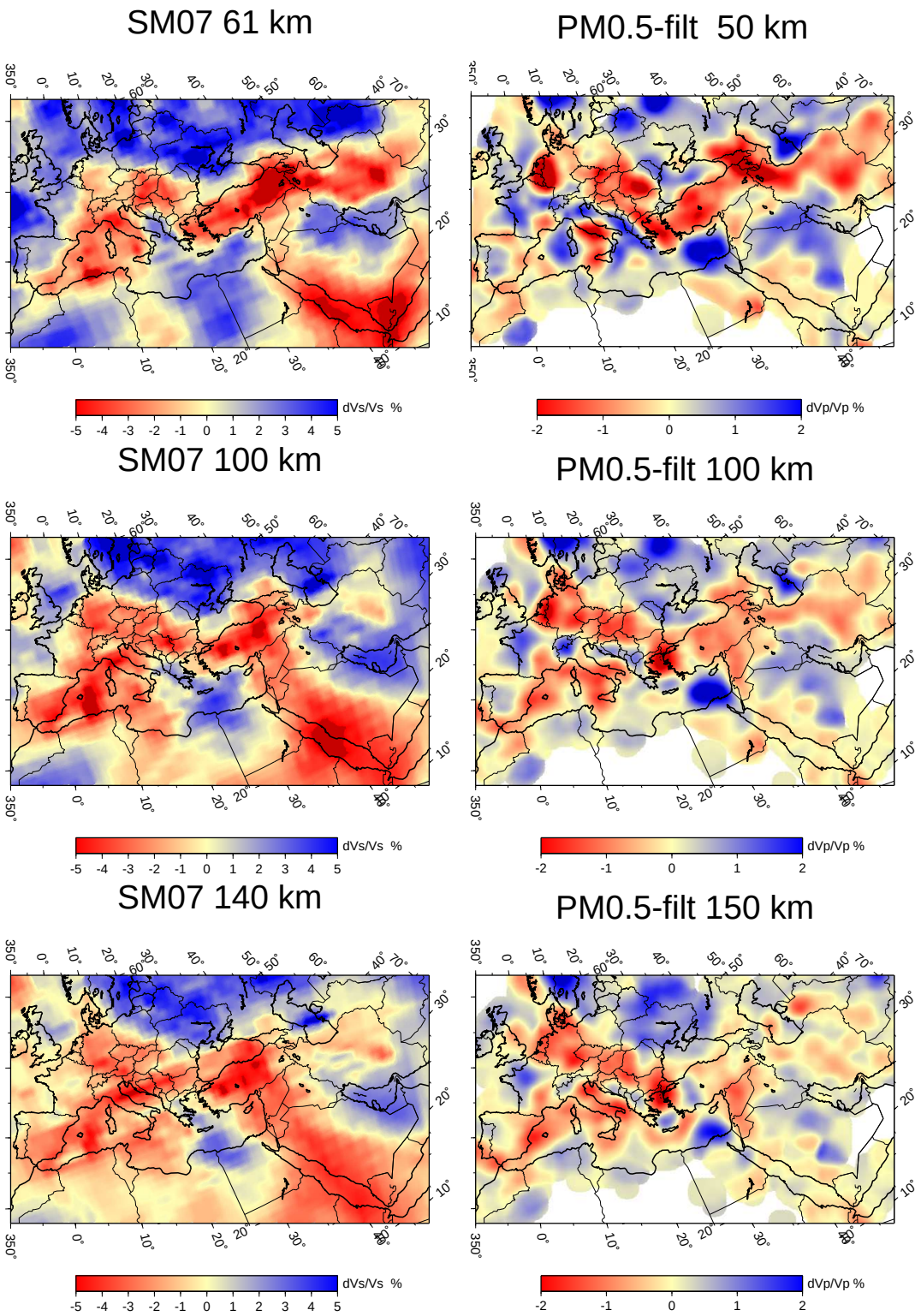


FIGURE 4.12: *Comparison between shear velocity model SM07 and a filtered version of P-velocity model PM0.5 of* [Piromallo and Morelli, *2003] at three selected depth.*

5. Conclusions

In this thesis we have presented a new high resolution tomographic model of the Upper Mantle under the broad European and Mediterranean area in the geographical frame 10N-80N 35W-80E. This region had been previously object of several regional studies with smaller geographical extent [*Spakman et al.*, 1993; *Piromallo and Morelli*, 2003; *Marone et al.*, 2004] or only partially overlapping with this area [*Ritzwoller and Levshin*, 1998; *Silveira and Stutzmann*, 2002]. We felt the need for a comprehensive and detailed model of the entire Eurasian and Mediterranean region.

We have modeled surface waves dispersion, because of their their ability to reach a more uniform spatial coverage compared to body wave phases, even where epicenters and seismographs are sparse. The model has been obtained from inversion of an high quality dataset of Rayleigh and Love group speeds, resulting in a dense and uniform coverage of the study area. Despite a rather classical measurement method (Multiple Filters and Phase Matched Filter), this study has brought significant advancement in current understanding of tomographic inversion schemes. We suggested a direct measurement error evaluation based on the analysis of RETREAT stations ray clusters. We carried out a linear inversion of group speed data applying regularization constraints on smoothness and on the vicinity of the solution to an a priori reference model. The innovative approach of this work lies in the use of a reference laterally heterogeneous global model, which we derived from a new inversion of group speeds obtained from the global phase velocity observations of [*Ekström et al.*, 1997]. This expedient, in addition to provide a more realistic starting model with respect to commonly used 1-D models, reduces the emergence of artifacts in areas of non-optimal path coverage. Another original scheme adopted in this thesis is the expression of the a posteriori model covariance matrix in terms of slowness errors; this guarantees the formal consistency with the linear problem solution, and legitimize the Gaussian error analysis.

Various attempts to assess the stability of the solution are commonly performed through synthetic tests such as checkerboard tests, but their results have been shown to have only a limited significancy [*Lévêque et al.*, 1993]. We have instead investigated the effects of different regularization constraints on the shape of the solution with a systematic exploration of the model space and a statistical analysis of different solutions obtained varying the regularization parameters. By selecting a set of formally equivalent

models which have the same data fit but differ in terms of smoothness and closeness to the reference model, we have shown that large-scale features of our model are generally robust and stable, whereas care must be taken in the interpretation of finer details, some of them depending on the particular regularization choice.

Our resulting group speed dispersion maps show enhanced resolution with respect to other surface waves studies and exhibit an excellent correlation with the known tectonic features of the region. Particularly worth mentioning are the contrast between the slower Central Europe and Western Mediterranean basin and the higher speed associated to the Eastern European Platform. These two different tectonic regimes are separated by a sharp gradient corresponding to the Tornquist-Tesseyre Zone, a consistent feature on our long period maps. We resolved with unprecedented detail several small scale features, such as the difference between the fast Northern and slow Southern Caspian block divided by the Ashgabat fault, and the narrow fast signature of the Carpathian Mountains.

The resulting group velocity maps have been inverted for the shear-wave speed at depth to image the three-dimensional radially anisotropic structure of the Upper Mantle. A non-linear inversion algorithm [*Tarantola and Valette*, 1982] has been implemented for the local inversion of group slowness curves, starting from the reference model PREM [*Dziewonski and Anderson*, 1981] corrected with the high resolution laterally varying crustal structure of CRUST2.0 [*Bassin et al.*, 2000]. Again, we inverted slowness instead than velocity (as it is commonly done) to guarantee a formal consistency with the Gaussian error analysis. Joint inversion of coupled Rayleigh and Love wave group velocity has lead to a transversely isotropic V_S model of the upper 280 km of the mantle. Our innovative parameterization of the model in terms of Voigt isotropic mean and radial anisotropy gives a better control on the amplitude of anisotropy with respect to the usual inversion for V_{SH} and V_{SV}. We compared our SM07 model with other global and regional S-wave speed models [*Boschi et al.*, 2004; *Ritsema et al.*, 1999; *Shapiro and Ritzwoller*, 2002; *Marone et al.*, 2004], derived from different types of data. The large-scale features in SM07 are in very good agreement with those reported in global models, and also small-scale robust anomalies are imaged with a resolution never before attained by inversion of surface wave data. Group velocity maps and deriving shear-wave speed model SM07 are constantly refined with the inclusion of new data, which leads to a continuous increase in path coverage and consequently to a better resolution. Dispersion maps could be helpful in regional epicentral location, and provide a good starting model for smaller-scale studies. Further investigation is however necessary to assess the effects

of underlying theoretical assumptions used in this study. Radial anisotropy in particular shows the need of more research, to overcome the limitations due to uneven distribution between Rayleigh and Love wave paths. A new inversion including azimuthal anisotropic parameters could provide new information.

Bibliography

Akaike, H., 1974. A new look at the statistical model identification, *IEEE Trans. Autom. Contr.*, **6**, 716 - 723.

Aki, K. and P.G. Richards, 2002. Quantitative seismology, Second Edition. University Science Books, Sausalito, California, USA

Bassin, C., Laske, G. and G. Masters, 2000. The Current Limits of Resolution for Surface Wave Tomography in North America, EOS Trans AGU, 81, F897.

Boschi, L. and G. Ekström, 2002. New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. geophys. Res.*, **107**, doi:10.129/2000JB000059.

Boschi, L., Ekström G. and B. Kustowski, 2004. Multiple resolution surface wave tomography: the Mediterranean basin, *Geophys. J. Int*, **157**, 293 - 304.

Dziewonski, A. M., and D. L. Anderson, 1981. Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, **25**, 297 - 356.

Dziewonski, A., Bloch, S. and M. Landisman, 1969. A technique for the analysis of transient seismic signals, *Bull. Seismol. Soc. Am.*, **59**, 427 - 444

Ekström, G. and A. M. Dziewonski, 1998. The unique anisotropy of the Pacific upper mantle, *Nature*, **394**, 168 -172.

Ekström, G., Tromp, J. and Larson, E.W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.*, **102**, 8137 - 8157.

Herrin, E. and T. Goforth, 1977. Phase Matched Filters: application to the study of Rayleigh waves, *Bull. Seismol. Soc. Am.*, **67**, 1259 - 1275.

Herrmann, R. B., 1973. Some aspects of band-pass filtering of surface waves, *Bull. Seismol. Soc. Am.*, **63**, 663 - 671.

Herrmann, R. B., 2004. Computer Programs in Seismology.

Landisman, M., Dziewonski, A. and Y. Satô, 1969. Recent improvements in the Analysis of Surface wave Observations, *Geophys. J. R. astr. Soc.*, **17**, 369 - 403.

Larson, E.W.F. and G. Ekström, 2001. Global Models of Surface Wave Group Velocity, *Pure Appl. Geophys.* **158**(8), 1377 - 1400.

G. Laske, Feb. 19, 2002. Instrumental Effects as seen by a Data Analyst, IGPP, Seismometry Seminar, at www.iris.edu/stations/seisWorkshop04/PDF/seismometry.pdf

Lévêque, J.J., Rivera, L. and G. Wittlinger, 1993. On the use of the checkerboard test to assess the resolution of tomographic inversions, *Geophys. J. Int*, **115**, 313 - 318.

Levshin, A. L., T. B. Yanovskaya, A. V. Lander, B. G. Bukchin, M. P. Barmin, L. I. Ratnikova, and E. N. Its, 1989. Seismic surface waves in a laterally inhomogeneous Earth, edited by V. I. Keilis-Borok, Kluwer Publ., Dordrecht.

Levshin, A. L., Ratnikova, L. and J. Berger, 1992. Peculiarities of surface-wave propagation across central Eurasia, *Bull. Seismol. Soc. Am.*, **82**, 2464 - 2493.

Love, A. E. H., 1927. A treatise on the mathematical theory of elasticity, 4th Edn., Cambridge Univ.

Lucente, F.P., Chiarabba, C., Cimini, G.B. and D. Giardini, 1999. Tomographic constraints on the geodynamic evolution of the Italian region, *J. Geophys. Res.* , **104**, 20307 - 20327.

Marone, F., Van der Lee, S. and D. Giardini, 2004. Three-dimensional upper mantle S-velocity model for the Eurasia-Africa plate boundary region, *Geophys. J. Int.*, **158**, 109 - 130.

Marquering, H., and R. Snieder, 1996. Surface-wave velocity structure beneath Europe, the northeastern Atlantic and western Asia from waveform inversion including surface-wave mode coupling, *Geophys. J. Int.*, **127**, 283 - 304.

Mele, G., A. Rovelli, D. Seber, T. M. Hearn, and M. Barazangi, 1998. Compressional velocity structure and anisotropy in the uppermost mantle beneath Italy and surrounding regions, *J. Geophys. Res.*, **103**, 12529 - 12543.

Pasyanos, M.E. , 2005. A variable resolution surface wave dispersion study of Eurasia, North Africa and surrounding regions, *J. geophys. Res.*, **110**, doi:10.1029/2005JB003749.

Piromallo, C. and A. Morelli, 2003. P wave tomography of the mantle under the Alpine-Mediterranean area, *J. geophys. Res.*, **108**, NO. B2, 2065, doi:10.1029/2002JB001757

Ritsema, J. and H. J.van Heijst, 2000. New seismic model of the upper mantle beneath Africa, *Geology*, **28**, 63 - 66.

Ritsema, J., van Heijst, H. J. and J. H. Woodhouse, 1999. Complex Shear Wave Velocity Structure Imaged Beneath Africa and Iceland, *Science*, **286**, 1925 - 1928.

Ritzwoller, M.H. and A. L. Levshin, 1998. Eurasian surface wave tomography: Group velocities, *J. Geophys. Res.* , **103**, N. B3, 4839 - 4878.

Sebai, A., Stutzmann, E., Montagner, J-P., Sicilia, D. and E. Beucler, 2006. Anisotropic structure of the African upper mantle from Rayleigh and Love wave tomography, *Phys. Earth Planet. Inter.*, **155**, 48 - 62.

Shapiro, N.M. and M. H. Ritzwoller, 2002. Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int* **151**, 88 - 105.

Silveira,G., Stutzmann, E., Griot, D.A., Montagner, J.P. and L.M. Victor, 1998. Anisotropic tomography of the Atlantic Ocean from Rayleigh surface waves, *Phys. Earth Planet. Inter.*, **106**, 257 - 273.

Silveira,G. and E. Stutzmann, 2002. Anisotropic tomography of the Atlantic Ocean, *Phys. Earth Planet. Inter.*, **132**, 237 - 248.

Spakman, W., Van der Lee, S. and R.D. Van der Hilst, 1993. Travel-time tomography of the European-Mediterranean mantle down to 1400 km, *Phys. Earth Planet. Inter.*, **79**, 3 - 74

Takeuchi, M., and M. Saito, 1972. Seismic surface waves, in *Seismology: Surface Waves and Free Oscillations; Methods in Computational Physics*, volume **11**, edited by B. A. Bolt, pages 217 - 295, Academic Press, New York.

A. Tarantola, 2005. Inverse Problem Theory and Model Parameter Estimation, SIAM, Philadelphia

Tarantola, A., and B. Valette, 1982. Generalised nonlinear inverse problems solved using the least square criterion, *Rev. Geophys.*, **20**, 219 - 232

Trampert, J. and R. D. Van der Hilst, 2005. Towards a quantitative interpretation of Global Seismic Tomography, in: Earth's Deep Mantle: Structure, Composition and Evolution. *Geophysical Monograph Series*, **160**, AGU.

Trampert, J. and J.H. Woodhouse, 1995. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 s period, *Geophys. J. Int.* , **122**, 675 - 690.

Trampert, J. and J.H. Woodhouse, 2003. Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s, *Geophys. J. Int.*, **154**, 154 - 165.

Van der Lee, S. and G. Nolet, 1997. Upper mantle S-velocity structure of North America, *J. geophys. Res.*, **102**, 22815 - 22838.

Villaseñor, A., Ritzwoller, M.H., Levshin, A.L., Barmin, M.P., Engdahl, E.R., Spakman, W., and J. Trampert, 2001. Shear velocity structure of central Eurasia from inversion of surface wave velocities, *Phys. Earth Planet. Inter.*, **123**, 169 - 184.

Wessel, P., and W. H. F. Smith, 1995. New version of the Generic Mapping Tools released, Eos Trans. AGU, 76, 329

Yang, Y., Ritzwoller, M. H., Levshin A. L. and N. M. Shapiro, 2006. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.*, **168**, 259 - 274.

Zhou, Y., Dahlen, F. A., Nolet, G. and G. Laske, 2005. Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.*, **163**, 1087 - 1111.

Zhou, Y., Nolet, G., Dahlen, F. A., and G. Laske, 2006. Global upper mantle structure from finite-frequency surface-wave tomography, *J. geophys. Res.*, **111**, B04304, doi:10.1029/2005JB003677.