Tomographie de la croûte et du manteau Équatoriens à partir des données du réseau sismologique national

Sebastián Araujo

Sebastián Araujo. Tomographie de la croûte et du manteau Équatoriens à partir des données du réseau sismologique national. Geophysics [physics.geo-ph]. Université Grenoble Alpes, 2016. English. NNT : 2016GREAU015 . tel-01530749

HAL Id: tel-01530749
https://theses.hal.science/tel-01530749

Submitted on 31 May 2017

UNIVERSITÉ GRENOBLE ALPES

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES
Spécialité : **Sciences de la Terre, de l'Univers et de l'Environnement**

Arrêté ministériel : 7 Août 2006

Présentée par

Sebastián Araujo

Thèse dirigée par **Bernard Valette**

préparée au sein de **l'Institut des Sciences de la Terre**
et de l'école doctorale **Terre Univers Environnement**

Travel time tomography of the crust and the mantle beneath Ecuador from data of the national seismic network

Thèse soutenue publiquement le **26 septembre 2016**,
devant le jury composé de :

Joseph Martinod
Université de Savoie Mont Blanc, ISTERRE, Président
Philippe Charvis
IRD Géoazur, Sophia Antipolis, Rapporteur
Dominique Gibert
OSUR Géosciences, Université de Rennes 1, Rapporteur
Etienne Jaillard
IRD ISTerre, Grenoble, Examinateur
Vadim Monteiller
CNRS LMA, Marseille, Examinateur
Mario Ruiz
Institut Géophysique de l'Equateur, Examinateur
Bernard Valette
IRD ISTerre, Chambéry, Directeur de thèse

Agradecimientos

La realización de este trabajo de investigación se financió gracias a un beca otorgada por la Secretaría Nacional de Ciencia, Tecnología y Educación Superior SENESCYT del Ecuador.

Los datos provienen en su mayoría de la Red Nacional de Sismógrafos del Instituto Geofísico de la Escuela Politécnica Nacional. Sin el abnegado trabajo de todos sus miembros esta tesis no hubiese sido posible.

Mi agradecimiento personal al Instituto de Ciencias de la Tierra ISTerre de la Universidad de Grenoble que me prestó todas las facilidades para llevar a cabo mi investigación.

Gracias a Bertrand Potin que compartió conmigo todos sus conocimientos en el manejo y desarrollo de los softwares que ambos utilizamos en nuestros trabajos de tesis.

Bernard Valette asumió el reto de dirigir esta tesis e involucrarse en todo los aspectos que implican realizar un proyecto de cooperación entre Ecuador y Francia.

Una mención especial a otro miembro de mi equipo, Christel Marchica, quien me ayudó en variados temas logísticos durante mi estadía en Grenoble.

Résumé

Bien que de nombreuses études géodynamiques et tectoniques aient été effectuées à partir l'activité sismique en Equateur, il n'existait pas à ce jour une tomographie complète utilisant l'ensemble des données du réseau sismologique Equatorien (RENSIG), mise à part une étude prélimaire sur la partie centrale de l'Equateur menée en 1994 par Prévot et coll. et de plusieurs profils sismiques déterminés à la suite des campagnes marines SALIERI et SISTEUR. Inverser les centaines de millier de temps d'arrivées d'ondes P et S, de qualité inégale, formant le catalogue RENSIG était le défi qu'a constitué le sujet de cette thèse.

Nous décrivons comment nous avons complété le catalogue RENSIG par des données provenant du Nord du Pérou et comment nous avons homogénéisé et filtré l'ensemble de données résultant, comportant plus de 800 000 temps d'arrivée correspondant à plus de 50 000 séismes. Pour inverser ces données nous avons adopté une approche Bayésienne. Nous montrons comment le problème peut être reformulé dans un contexte Gaussien par un changement de variables, tout en imposant une statistique robuste aux données, qui conduit à un problème de moindre carrés non linéaire. Nous détaillons particulièrement la régularisation du problème au travers des noyaux de covariance qui conduit à définir des paramètres de contrôle fort utils pour l'inversion. Nous montrons également qu'inverser des différences de données revient à introduire des termes spécifiques de corrélation dans la matrice de covariance des données, tout en conservant les données brutes. Nous indiquons finalement comment le calcul de l'indice de restitution permet de définir une zone de confiance du modèle résultant de l'inversion.

L'inversion a été menée pratiquement en utilisant les codes informatiques (en Fortran 2003 par B. Potin, B. Valette, V. Monteiller): LOCIN (localisation) et INSIGHT (tomographie). La région finale d'étude est constituée par une boite parallélipipédique de dimension 590×770 km^2 de base et de 252 km de hauteur qui contient la topographie de la surface. Le modèle est constitué d'une part des valeurs de v_P et v_P/v_S sur une grille ayant 5 km de pas horizontal et 2 km de pas vertical et, d'autre part, des paramètres d'identification spatiale et temporelle des séismes. Un ensemble de tests nous a permis de déterminer des valeurs raisonables de ces paramètres au travers d'un analyse de type courbe en L.

Nous avons obtenu une amélioration de la localisation de la sismicité, qui nous a permis de mieux décrire les essaims superficiels comme ceux de Pisayambo, Macas et du Reventador et d'identifier des linéaments en relation avec la Tectonique. Nous avons

également obtenu une image de la sismicité à profondeur intermédiaire qui est dominée par la présence de 4 nids sismiques, ceux de Madonaldo, La Maná et de Guayaquil à des profondeurs entre 75 et 115 km et celui de Puyo à de plus grandes profondeurs. La zone de Wadati-Benioff nous a permis de définir la profondeur du slab jusqu'à des profondeurs de 100-150 km en fonction de la latitude et d'observer la décroissance du pendage de 25° environ au nord et au centre de l'Equateur jusqu'à environ 10° au sud puis au nord du Pérou. Par ailleurs, l'analyse du champ de vitesse des ondes P suggère fortement que le slab est coupé en deux morceaux, le morceau sud passant sous le morceau nord au niveau du nid sismique de Puyo. Le modèle v_P/v_S présente une forte anomalie positive de ce rapport le long de la cordillère occidentale à des profondeurs entre 30 et 50 km qui caractérise des matériaux partiellement fondus et correspond au réservoir d'alimentation profond de l'arc volcanique. Enfin, nous avons déduit de notre modèle un modèle de profondeur de Moho en prenant la profondeur de maximum de la norme du gradient de vitesse entre les vitesses de 7.2 et 7.4 km/s et en incorporant l'information sur la profondeur de Moho provenant des campagnes SALIERI et SISTEUR dans la marge active.

Abstract

Although there have been numerous studies on the geodynamics and the tectonics in Ecuador based on the seismic activity, there has not been to date a comprehensive tomography study using the entire database of the National Seismic Network (RENSIG). Only a preliminary limited study was performed by Prevot et al. to infer a simple P velocity model in central Ecuador, and several profiles in the South-Colombian-Ecuador margin were also investigated by using travel time inversion of wide-angle seismic data obtained during the two marine experiments SISTEUR and SALIERI. Inverting the hundreds of thousands of arrival times of P and S waves of uneven quality that constitutes the RENSIG catalogue is the challenging subject of this thesis.

We describe how we complemented the RENSIG catalogue with data from the Northern Peru network and how we homogenized and filtered the resulting dataset of more than 800 000 first arrival times of P and S waves corresponding to more than 50 000 earthquakes. To invert these data for both the velocity models and the event locations we adopted a Bayesian approach. We show how the problem can be recast in the Gaussian framework by changes of variable while imposing a robust statistics to the data, and how it leads to a generalized nonlinear least squares problem. We detail in particular the regularization of the models through the smoothing and damping properties of the covariance kernels. We also show that inverting differences in data instead of the raw data amounts to the introduction of specific correlation terms in the data covariance matrix, while keeping the same set of data. We finally indicate how the computation of the averaging index allows the delimitation of a confidence region for the resulting model.

The practical inversion has been carried out by using the two Fortran 2003 codes (B. Potin, B. Valette, V. Monteiller): LOCIN (prior localization) and INSIGHT (tomography). The final study region is a parallelepipedic box of 590×770 km^2 area and 252 km height that contains the topography of the surface. The models consist of the v_P and v_P/v_S fields discretized over a grid, the spacing of which is 5 km in the horizontal directions and 2 km in the vertical one, and of the spatial and temporal parameters of the seismic events. A battery of tests allowed us to set reasonable values for these tuning parameters through an L-curve analysis.

We obtained the spatial distribution of the seismicity with an improved accuracy which allows us to describe with more details the shallow seismic clusters, as those of Pisayambo, Macas, Reventador, and to identify lineaments in the seismicity in relation

with tectonics. We obtained also a clear image of the intermediate depth seismicity wich is dominated by 4 nests, namely the Maldonado, La Maná, and Guayaquil nests, at depths ranging between 75 km and 115 km, and the Puyo nest at much deeper depths. The Wadati-Benioff zone allowed us to clearly defined the topography of the slab only to a depth to about 110-150 km, depending on the latitude, and to observe the decrease of the dip angle from about 25° in northern and central Ecuador down to about 10° in southern Ecuador and northern Peru. On the other hand, the analysis of the P velocity clearly suggests that the slab is broken in two pieces, the southern one passing under the northern at the level of the Puyo nest. The v_P/v_S model presents a high anomaly of the ratio along the western cordillera at a depth ranging between 30 km and 50 km that characterized partially melted rocks and corresponds to the feeding reservoir of the volcanic arc. Finally, we deduced the Moho depth from our model by taking the depth for which the norm of the velocity gradient is maximum between 7.2 and 7.4 km/s and by incorporating information on the Moho depth provided by the SISTEUR and SALIERI experiments in the convergent margin.

Contents

Introduction

The dynamics of the Nazca oceanic plate, which converges towards the South American continental plate in a $N83^oE$ direction and $56\,mm/year$ (Kendrick et al., 2003), generates an intense seismicity at both shallow and intermediate depths in Ecuador. There has been a series of studies on the geodynamics and the tectonics in Ecuador based on this seismic activity (*e.g.* Guiller et al., 2001; Legrand et al., 2005; Manchuel et al., 2011; Font et al., 2013; Alvarado et al., 2014; Yepes et al., 2016). However, there has not been to date a comprehensive tomography study using the entire database of the National Seismic Network (RENSIG). The main objectives of the RENSIG are to locate seismic events and monitor the activity of volcanoes, both phenomena being a source of seismic hazard for the inhabitants of the country. However, the records of up to a hundred stations of the network have given rise to hundreds of thousands of arrival times of P and S waves over time, which now constitute a huge catalogue of data of uneven quality that is available for a large-scale tomographic adventure.

A first attempt of seismic tomography at regional scale was performed in 1994 by Prévot et al. to infer a seismic velocity model in central Ecuador. Although that study was relatively limited in its approach and by the number of data inverted, it showed the possibility to use the database of the National Seismological Network of Ecuador to obtain tomographic images. Two additional local tomography studies of the Pichincha and Tungurahua volcanoes, conducted by Molina et al., in 2005 and García-Aristizábal et al. in 2007, have also demonstrated that obtaining a satisfactory tomographic resolution is challenging.

In contrast, the south Columbian-Ecuador convergent margin has been intensively studied by using travel time inversion of wide angle seismic data. These data were acquired during two major marine seismic experiments: the SISTEUR cruise in 2000 (Collot et al. 2002) and the SALIERI (South American Lithospheric Transects Across Volcanic Ridge) cruise in 2001 (Flueh et al., 2001). It allowed the study of the seismic structure and of the crustal overthickening of the Carnegie ridge as well as of the segmentation of the margin into three contrasted zones (Graindorge et al., 2004, Sallarés et al.,2005; Gailler et al.,2007).

This thesis work is pioneering the inversion of the complete RENSIG database of arrival times of waves since 1988 for imaging the crust and the mantle wedge beneath Ecuador. A challenging task was to extract an optimal core of data from this very large and heterogeneous database by using robust statistics.

The region of Ecuador and northern Peru was modeled by several interesting geodynamic phenomena that cause a structural complexity that can be unraveled by using tomographic methods. To the process of subduction of the Nazca plate beneath the South American plate must be added the change in the dynamics generated by the pronounced curvature of the continental margin between northern Peru and the Gulf of Guayaquil. The subduction of the Nazca plate in Ecuador occurs in the context of the transition between a flat slab in Peru and a dipping slab in northern Colombia. The subduction process has also been greatly influenced by the creation of the Cocos plate (Londsdale et al. 2005) with an opening ridge that was subducting until 10 million years ago, and by the Carnegie ridge that is still currently subducting. All these combined phenomena created a zone of magmatic and tectonic activity of great complexity that frames the main objectives of our research. These objectives are to obtain an image of the subduction of the Nazca plate in Ecuador and to define the image of the Moho in this region. These images result from the a priori localization of the seismicity by a grid-search method followed by the inversion of the data for both the event locations and the velocity models. These two major objectives at the regional scale are accompanied by the detailed study of the main seismicity clusters directly linked to the subduction and finally by a preliminary tomographic study of Cotopaxi volcano.

To achieve the objectives of our thesis work, we needed a deep understanding of theoretical foundations of the inverse problem from a Bayesian approach (Valette, 2011) that constitute the subject of the first part of the thesis and an essential requirement for the proper application of location and tomography softwares (Potin and Valette, 2016).

Chapter 1

Seismic data

1.1 Introduction to the studied area

From a topographical point of view, the studied region displayed in figure 1.1 can be divided in 4 areas between the Pacific ocean and the Amazonian Basin. On the west side, there is the continental margin and the low coastal region with several basins, such as the Progreso and Sechura basins, and a coastal range in Ecuador and in northest Peru that lies about 300 m above the sea level, and is separated by the opening of the Guayaquil gulf. Eastward lies the narrow Andean range with a width increasing from 140 km at the Colombian border to more than 300 km in Northern Peru. On its east side lies the Sub-Andean Belt with a lower altitude, wich overthrusts the Amazonian Basin with an altitude of about 300 m. North of $2^o S$, the Andean range reaches the highest average altitude in the studied region, and corresponds to a volcanic arc, with two cordilleras separated by the narrow interandean depression. The main active volcanoes of the Ecuadorian part of the arc are from South to North: the Sangay, the Tungurahua (active since 1999), the Cotopaxi (phreatic crises since August 2015), the Guagua Pichincha (phreato-magmatic crisis since 1998 until 2001), the Cotocachi, the Reventador, and the Chiles-Cerro Negro (active since 2014).

From a geodynamical point of view, the main feature is the obliquity of the subduction, oriented $N83^o$ with a rate of $56\,mm/yr$ (Kendrick et al. 2003), that induces, by partitionning of the slip vector, the NE motion of the North Andean Block (Nocquet et al. 2014), which is a oceanic block mainly constituted of terranes accretated more than 60 my ago (Jaillard et al., 2009), and delimited on the east by the Chingual-Cosanga-Pallatanga-Puna (CCPP) transpressive right-lateral fault system at the limit of the stable South American plate (Alvarado et al. 2016). Moreover, the presence of the Carnegie ridge, an aseismic ridge with a height of about 2000 m which is subducting under the North Andean Block. Symmetricaly, due to the convexity of the trench line south of the Guayaquil gulf, the obliquity of the subduction causes the southeast motion of the Inca sliver, a continental domain wedged between the Nazca plate and the South American plate (Nocquet et al. 2016).

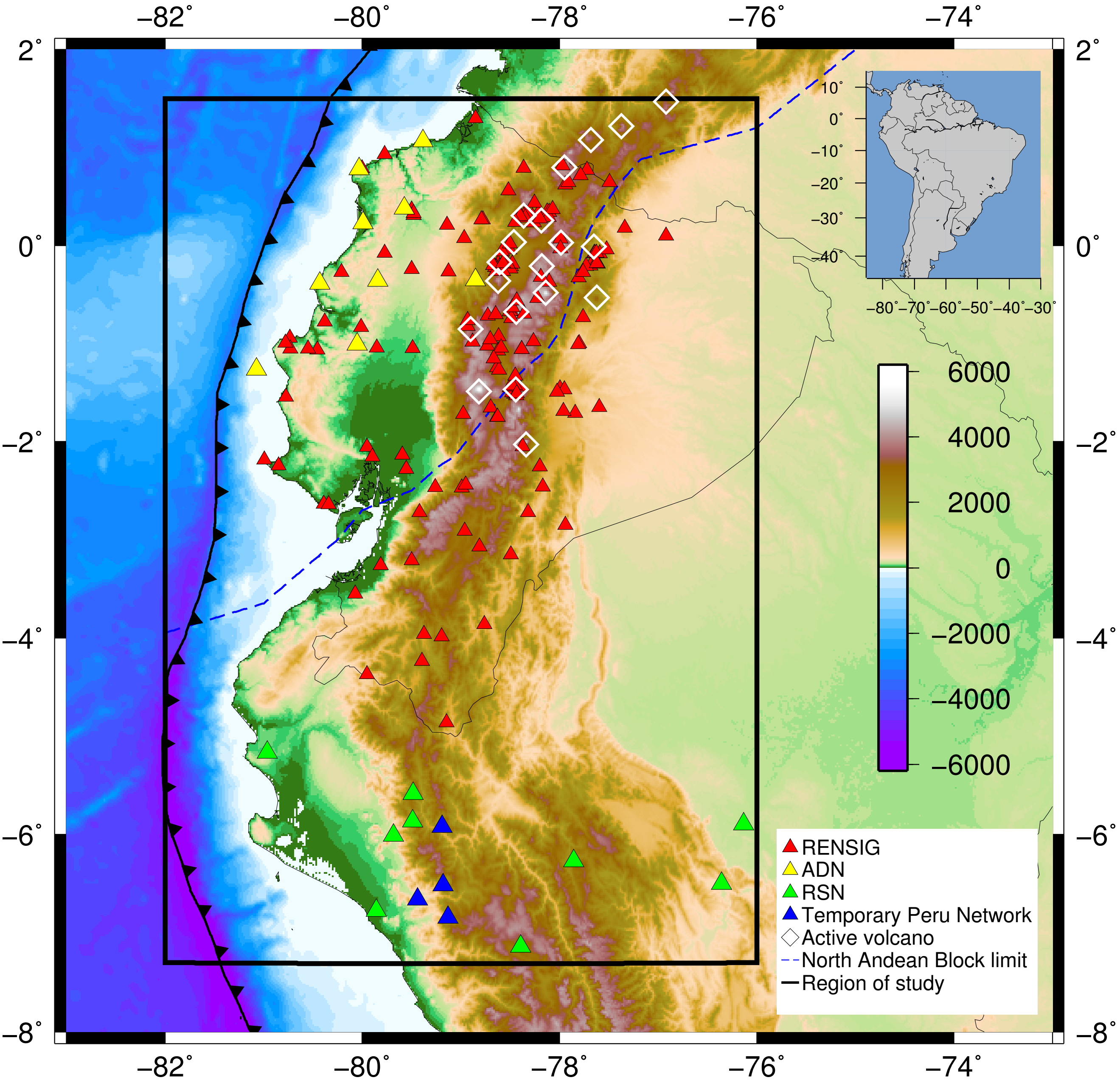

Figure 1.1: *Topographic map of the area concerned by the tomography study over the black box. The topography allows one to identify easily the trench and the Carnegie ridge at sea, as well as the four main geographical regions in Ecuador, namely the coastal region with coastal ranges on both sides of the Guyaquil gulf, the Cordillera, the Eastern sub-Andean Belt, and the Amazonian basin. The blue dotted line indicates the limit of the North Andean Block and the diamonds the active volcanoes. The four seismic networks involved in the study are represented by triangles of different colors. Clearly, the RENSIG network is the most important regarding the size of coverage and the number of stations.*

Concerning the seismicity, the studied area is relatively close to the subduction trench, and to the seismogenic zone, with a high rate of earthquakes especially in the high coupled segment corresponding to the Carnegie ridge (Font et al., 2013). The seismicity of the subduction extends down to the Puyo cluster (or nest) at a maximum depth of 250 km, with several other intermediate clusters. The crust in the region is marked by both a large set of seismic faults (CCPP) that border the Andean block and produce events down to a depth of $40 km$, and by relatively shallow clusters of eathquakes. In the central northern part, an important source of seismicity comes from the active volcanic arc that produces numerous volcano tectonics events.

1.2 Seismic networks

The data used for this tomography study are the arrival times of P and S waves at seismic stations. They mainly come from the *Red Nacional de Sismógrafos del Instituto Geofísico de la Escuela Politécnica Nacional del Ecuador* (RENSIG). They were complemented with data from the *Red Sísmica Nacional del Instituto Geofísico del Perú* (RSN). Additionally, we had access to few data coming from two seismic experiments: a seismological campaign developed in Ecuador Coast in the framework of ADN (*Andes Du Nord*) project from 2008 to 2014 and a temporary network installed in Northern Peru from 1996 to 2001 (Tavera, 2006). The location of these four networks are displayed in figure 1.1.

For this study we have taken as reference basis the catalogue of the RENSIG network. In this catalogue, the event duration magnitude ranges from $M_D = 3.5$ (Beauval et al., 2013) up to more than $M_D = 6.5$ for major events.

1.2.1 RENSING Network

The RENSIG network began operating in the late seventies and since then it has been progressively growing up to reach 161 permanent seismic stations today. The first objective of the RENSIG network was to insure the monitoring of the main volcanoes which threaten the Interandean zone of Ecuador. Until 2011, the network consisted of 30 short period seismometers (L4C-1D and L4C-3D Mark Products) and 10 broad band seismometers CMG-40T Guralp. Since 2011 the network has been complemented by 65 broad band seismometers (Trillum Compact, Trillum 120 and Guralp CMG-3T). The recorded signals of most of the stations are now transmitted in real time and are automatically processed first in order to insure seismic alerts and then manually checked to improve the picking of P and S phases. The system used for the data acquisition was successively Pickev until May 2001 (Alvarado, 2012), Earthworm (Alvarado, 2012), and SeisComP now (Alvarado, 2012).

In order to insure the monitoring of volcanic activity, 69 stations of the network have been installed over 9 volcanic edifices; namely the volcanic cone of Antisana, Cayambe, Chimborazo, Cotopaxi, Igualata, Pichincha, Quilindaña, Reventador and Tungurahua.

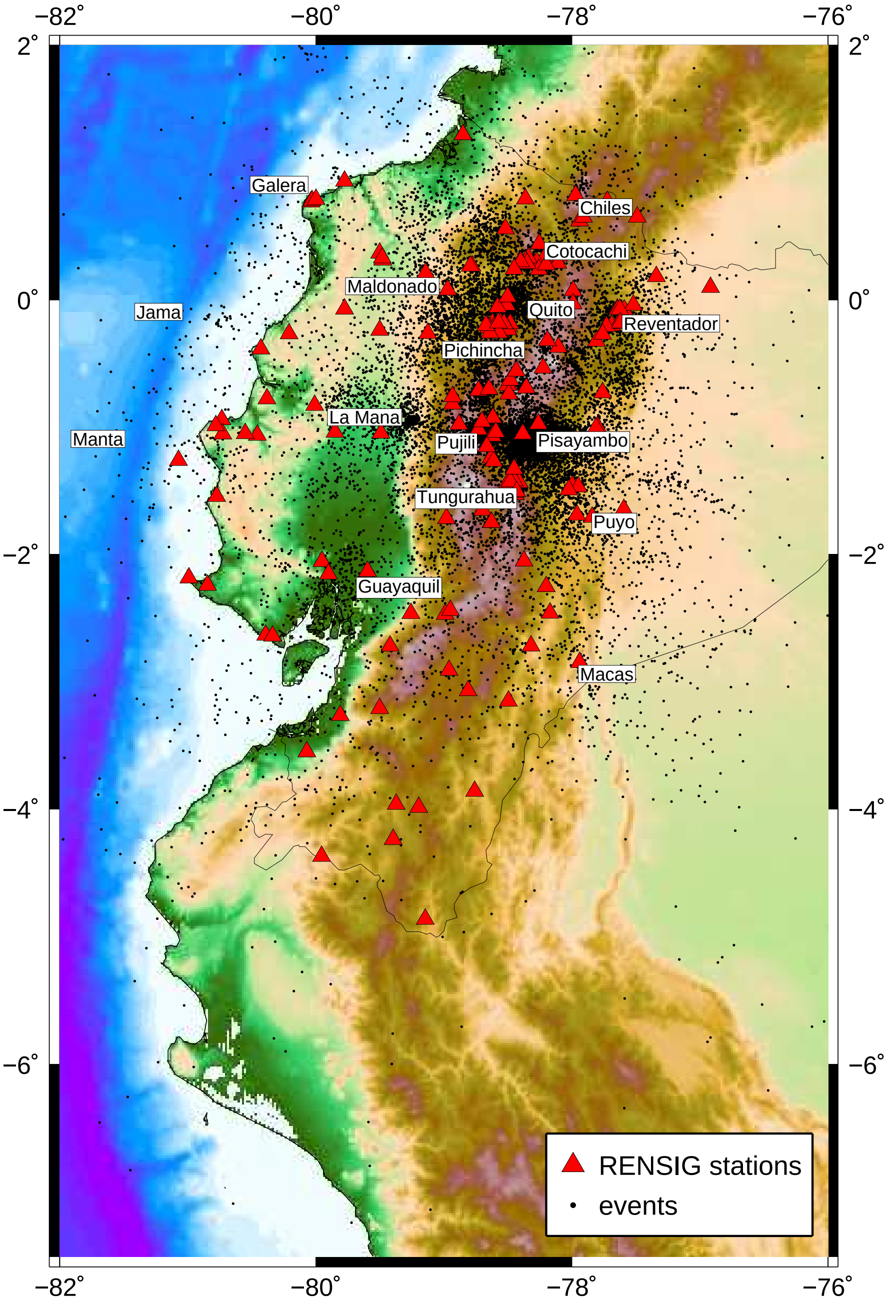

Figure 1.2: *Positions of RENSIG stations used in tomography study and location of seismicity recorded in these stations. Geographical distribution of seismic events point to the presence of clusters linked to subduction process: La Maná, Guayaquil and Puyo; to crustal active faults: North Quito, Pujilí, Pisayambo and Macas; and to volcanoes in eruption process: Guagua Pichincha, Reventador and Tungurahua.*

This point must be carefully taken into account and imposes to consider a site effect term when performing a regional tomography with data from these stations.

Including temporary experiments, The RENSIG dataset was provided by 172 seismic stations of the network in total. The location of these stations are displayed in figures 1.1, and 1.2, indicated by a red triangle. Besides the location of the network, the figure 1.2 shows the location of the seismicity recorded by this network. We can observe that this seismicity is relatively concentrated in clusters that originate from crustal activity as for the Pisayambo, Pujilí or North Quito swarms, or from volcanic processes as for the Reventador, Tungurahua, Guagua Pichincha and Cotopaxi volcanoes, or from magmatic processes. Other seismic nests are related to the subduction process either in the seismogenic zone or at intermediate depths as the La Maná, Guayaquil, or Puyo nests. The names chosen for the clusters are related to towns close to the mean location of epicenters (Instituto Geofísico, 2013).

The set of data was provided by the *Instituto Geofísico* (IG) in different files containing the first arrival time of P and S waves, picked by hand for most of them, corresponding to seismic events with magnitude greater than 2.5. Each file, which corresponds to a stage of development of the network and of the process of data picking, had first to be checked to solve the specific format problems. The period of time covered by these files are respectively: April 1988-August 2011, May 2011-October 2012, January 2011-October 2012, March 2013-December 2013, January 2013-June 2014, and

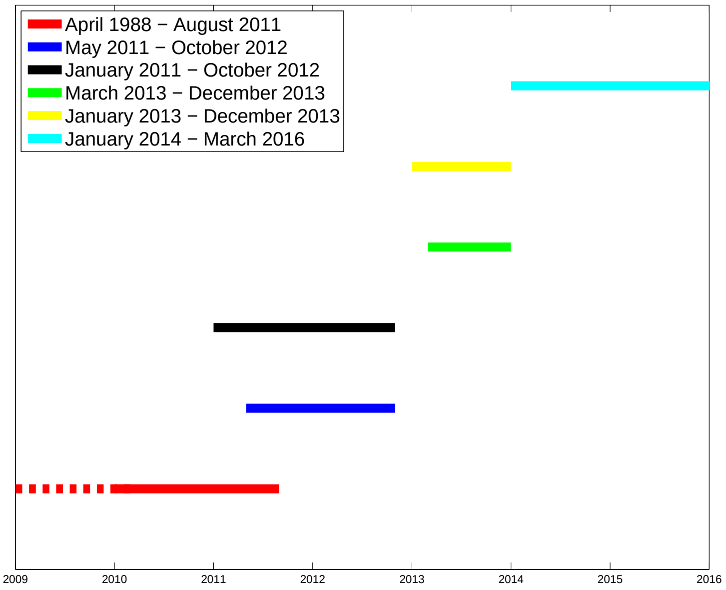

Figure 1.3: *Graphical representation of the time intervals of the six files containing the picked phases furnished by the Instituto Geofisico. We can observe time periods for which several different files are avalaible. The year of the onset of the national seismic catalogue is 1988.*

January 2014-March 2016. As can be seen in figure 1.3 different files, and data, may correspond to a same period of time between 2011 and 2014, due notably to the change of acquisition system. Consequently, we had to develop a strategy to obtain a coherent data base for periods where different data were available.

Basically, two cases arise. In the first one, the arrival time data common to the two considered files are strictly equal. Taking one of this file as reference, we complement it by the data of the other file that are not existing in the reference file, and that correspond to either a common event or to an event not recorded in the file. This operation is relatively easy since the common data are strictly the same, and consequently the identification of common events is obvious. In the second case, corresponding to a change of operating system between the two files, the common data for a same event and a same station may be not exactly the same. Consequently we have first to identify the common events in the two files by retaining those for which the differences in arrival time for the same phase at the same station in the two files are all less than one minute. We assign the value corresponding to the newest acquisition system for the arrival time in the case when two data where available for the same phase at the same station. After that, the procedure is the same as for the first case. We had also to eliminate artificial duplication of data that occured for some periods.

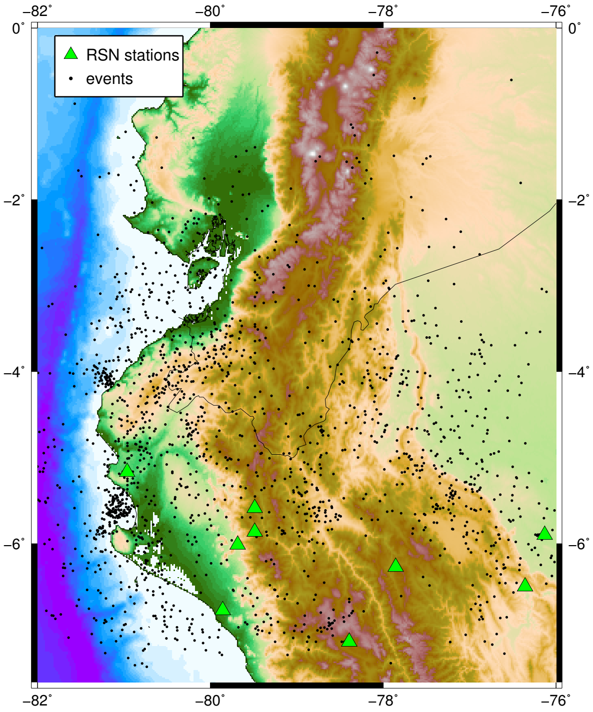

Figure 1.4: Location of stations of Peru Red Sísmica Nacional RSN and events recorded by the network that have been used in our tomography study.

1.2.2 RSN Network

The Seismological Peru Network consists, in its northern part, of short period seismometers of type: SS-1 and WR1 Ranger Kinematrics and Mark Products L4C. The transmission of recorded signal is done in real time to the *Instituto Geofísico del Perú* in Lima, where first wave arrival time are picked by hand. We have data from 9 stations of the network corresponding to 3163 seismic events recorded between January 2010 and November 2014.

The locations of those stations of the RSN network, the data of which have been used in our study, are displayed in figure 1.4. The figure also shows the *a priori* location of the seismicity determined by this network. We have defined the South and East limits of our study box by taking the location of these stations and events into account.

1.2.3 North Peru temporary network

This seismic temporary network was deployed with the specific objective to study and define the slab geometry in Northern Peru (Tavera, 2006). It was composed of 7 short period stations with Kinemetrics SS-1(vertical component) seismometers. The network has been operating from 1996 until 2001 and we obtained the picked P and S phases corresponding to 1005 seismic events that occurred between March and December 1998.

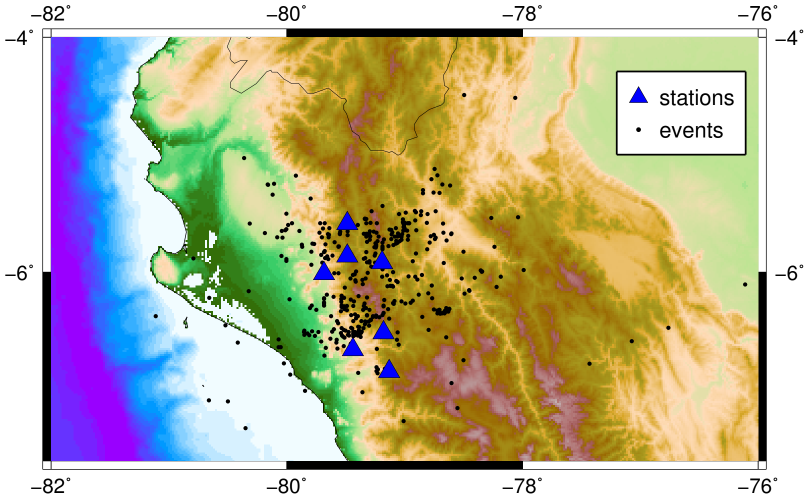

Figure 1.5: *The seven stations of the temporary seismic network in the north Peru added to RSN and the seismic events registered with the help of this temporary network.*

3 stations of this network were incorporated to the RSN network (Tavera et al., 2006). That explains why in figure 1.1 only four stations of this network are specified in

blue color to distinguish them from the other ones of the RSN network. The seismicity recorded by the temporary network is displayed in figure 1.5. The events are clearly limited to a narrow area.

1.2.4 The ADN network

The stations of this network (represented in yellow in figure 1.1) was devoted to the study of the oceanic margin seismicity and are located in the Ecuadorian coastal area. The nine seismic stations of the campaign consisted of Kephren digitizers with Guralp CMG3-ESP broadband sensors (Ponce, 2014).

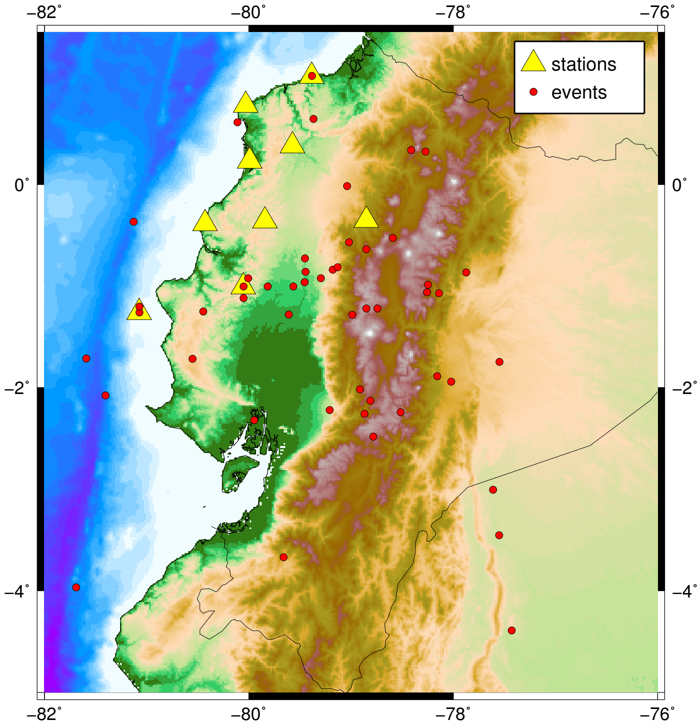

Figure 1.6: *Location of the 53 events for which we obtained arrival times of the* 9 *stations of the ADN network.*

We obtained from Marc Régnier arrival times picked picked with a good quality from the records of the ADN network corresponding to 53 events with magnitude greater than 4.5 that occurred between October 2009 and June 2013. Figure 1.6 displays the location of these events and shows that they are well distributed over the study area.

1.3 Building up the data set

To build up the data set we have to put together all the arrival times of the different networks corresponding to a same event. Merging the different data set thus imposes to identify the events that are the same in the various catalogue and to harmonize the different evaluations of the observational errors.

1.3.1 Harmonizing observational uncertainty

When inverting seismological data the observational uncertainties represent an important amount of information. But they are unfortunately very difficult to determine when picking phases arrival times, and mostly depend of the appreciation of the observer. The observational uncertainties are differently evaluated depending on the network considered, either in the form of a standard deviation in second or in the form of weight. Usually this weigh is a natural number, ranging from 0 for good quality up to 3 for poor quality, and the value 4 indicates that the indicated value must not be taken into account. This is the case for the RENSIG catalogue between 1988 and 2010 and for the temporal networks in northern Peru and in the coastal region of Ecuador. We took this way of evaluating uncertainty as reference scale in building up our data set.

The protocol of the RENSIG network was changed a first time in 2011, and a second one in 2014. For years 2011, 2012, 2013, the data are given with only two levels of quality, 0 or 1. Since 2014, the arrival times are given with three quality levels. For these data we kept the weight furnished by RENSIG for P waves, but increased by one the weight for S waves.

The data of the RSN network were furnished with only two weights: 0 or 4. We decided to replace 0 by 1 and 4 by 3 for P phases, and to keep only S data with 0 weight that we increased to 2.

Finally, as the inversion procedure requires standard deviation in seconds, we set the following correspondance (1.1) between the quality weights and the standard deviations.

Weight	$\sigma(s)$
0	0.15
1	0.3
2	0.5
3	0.8

Table 1.1: *Values of standard deviations as function of the quality weights*

1.3.2 Merging the different data sets

In order to merge two data sets we have to identify the events that have been recorded in both catalogues. Different criteria can be used to identify such events. For instance, the arrival times in the two catalogues must belong to a common time window, the two locations must be close to each other in space and in time or the stations having detected the event must belong to the same area, the size of which depends on the magnitude.

For this study we have used the following strategy. In a first step, we have retained the events for which the gap between the two sets of arrival time is less than a given value, 90 s in our case, taking account of the distance between North Peru networks and RENSIG network. In a second step, when two subsets of arrival times have been identified as corresponding to the same event by the first criterion, a new localization of this event is performed with the whole set of the corresponding arrival times in the two catalogues. The root mean square of the difference between observed and computed arrival times weighted by a modeling error allows the detection of wrong gathering of events. More precisely, the regrouping that yields a RMS greater than 2 are rejected. The location of events was performed with the HypRef code (Thouvenot and Frechet, 2006), which is derived from Hypo71 (USGS, 1975) in a layered half space down to a depth of 250 km and and with velocities given in (Tab.1.2). For the S waves velocities we took a V_P/V_S ratio of 1.74.

P wave velocity (km/s)	Depth (km)
5.3	0.0
5.92	3.0
6.60	27.0
8.0	35.0

Table 1.2: *Values of P wave velocity of the layered half space used to the preliminary localization of seismic events with HypRef*

Using the aforementioned criteria for merging northern Peruvian data with RENSIG catalogue, 163 out of 1004 events were found to correspond to an event of the RENSIG catalogue for the year 1998, and 343 out of 3163 for the years 2010-2014. Thus about 88% of the events of the Northern Peruvian catalogues correspond to new events and where directly incorporated in our data set. The data associated to the 12% of common events were added to the data of the corresponding events of the RENSIG catalogue in our data set.

For the 53 events registered by the ADN network, the data were incorporated by hand in our data set, event by event, controlling for each of them the time propagation of waves through the whole set of stations and checking the RMS corresponding to the new location determined in adding the new data.

1.4 Time and space coverage of seismic network

Figure 1.7 displays the number of data available each year in our data set composed with the four aforementioned catalogues over the period ranging from 1988 to 2016. We can observe a jump in the number of data in 1995 and 2013 corresponding to the increase of stations of the RENSIG network.

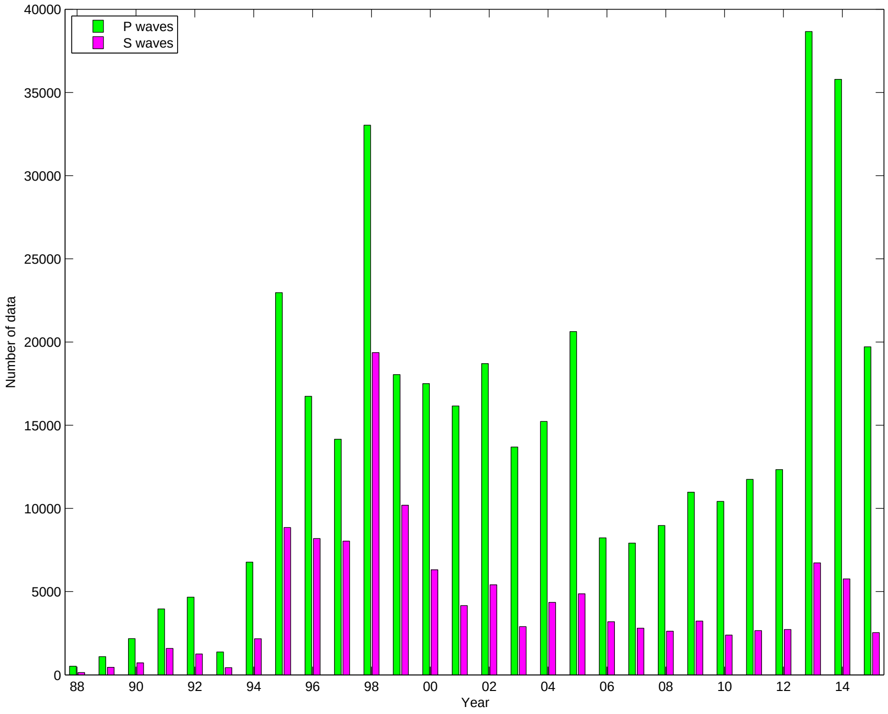

Figure 1.7: *Number of P and S arrival times for each year of the period of the data set used for the study. We may observe the low level of S picking compared to that of P, especially since 2013 due to the introduction of the automatic picking procedure*

The high number of data in year 1998 can be explained by two facts: firstly it is the year during which the Guagua Pichincha volcano started its activity, and secondly it is the only year for which we had data of the temporary Network in North Peru.

Then the volcanic activity was reinforced by the awakening of the Tungurahua volcano in 1999 and by the seismic (but not eruptive) activity of the Cotopaxi volcano in 2001-2002. This coincidence during which three main volcanoes produced seismic signals ended in 2005, and consequently the level of data dramatically decreased in 2006.

Considering now the spatial distribution of stations, figure 1.1 shows an heterogeneous coverage of the tomography area, with most of the stations concentrated in the volcanic central-north inter Andean zone and no station at all in the Amazonian basin and in the Progreso basin north of Guayaquil city. We can observe clusters of stations over the main active volcanoes Pichincha, Tungurahua and Cotopaxi due to the monitoring of volcanic activity. It is also worth noting that the seismic coverage of the northern part of the coastal seismic region and of the southern Ecuador has been improved only over the last few years, and that the Progreso and Amazonian basins remain without any station. We could incorporate in our study the data of the few stations over Northern Peru only for a restricted period of time.

The Voronoi tessellation provides an efficient way to display the density of the seismic array considered as a net (see figure 1.8). It is built up by considering the triangles linking neighbouring points of the net. The Voronoi cells are obtained by joining with straight lines the barycenters of those triangles. Thus each cell of the Voronoi tessellation is associated to a station (a nod of the net) and corresponds to the set of points that are closer to this station than to the others. Consequently, each of these cells represents, in a way, the area monitored by the corresponding station.

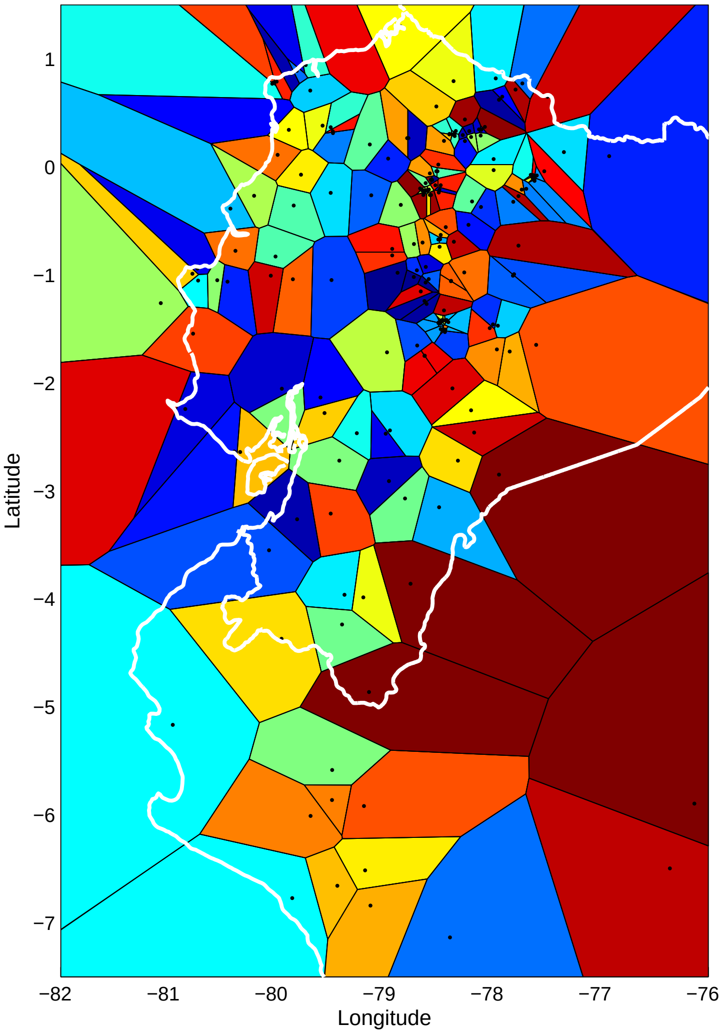

Figure 1.8: *Voronoi tessellation of the whole set of seismic stations used for the study. The positions of stations are indicated by black dots. The Voronoi cells are drawn in various colors over the white coast and border lines.*

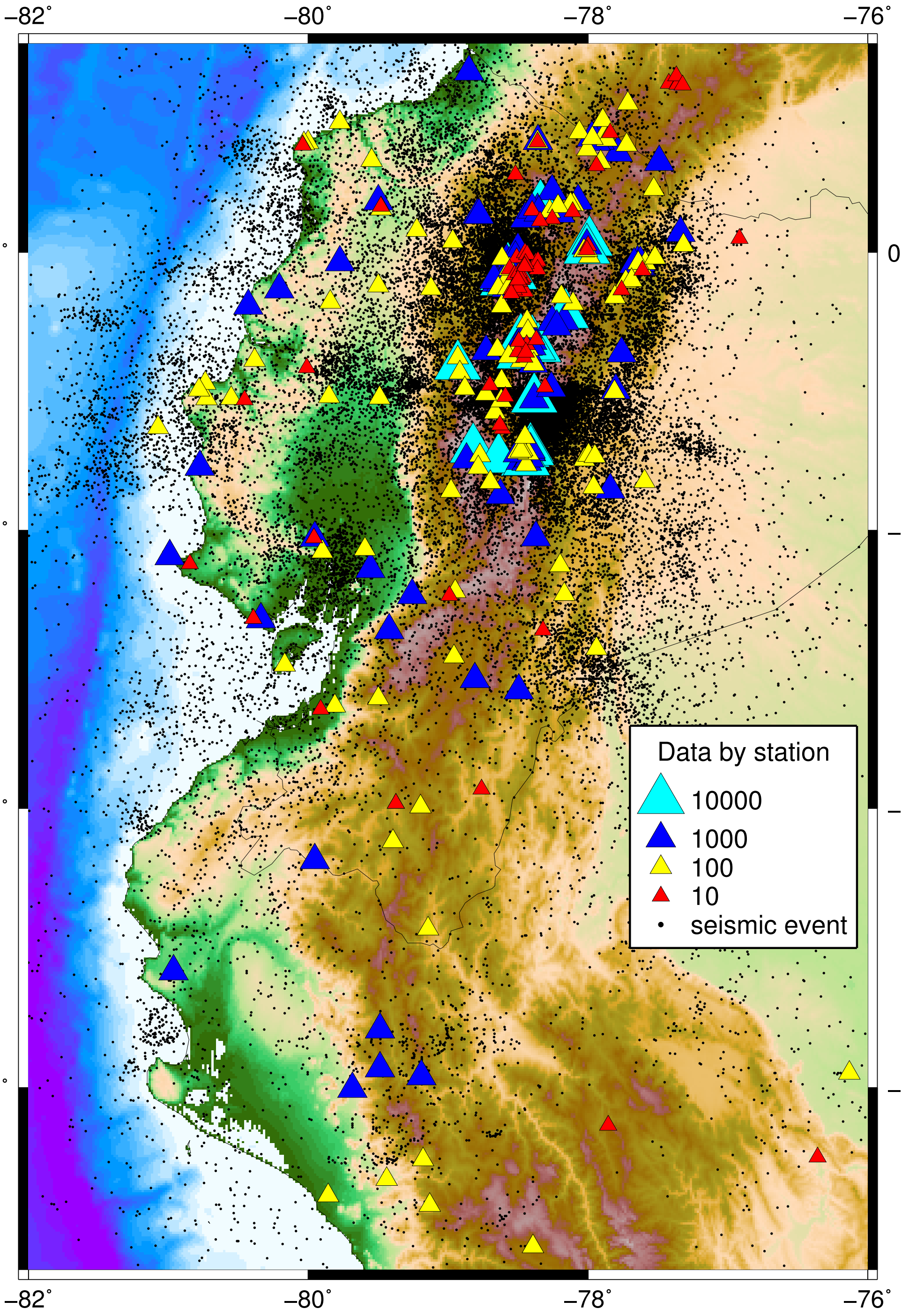

Figure 1.9: *Location of the whole set of stations and of the seismicity as a priori determined by the HypRef code. The epicenters are represented by black dots, and the station by triangles the color and size of which depend on the logarithm of the number of data (P and S phases) recorded at the station.*

The temporal and spatial coverage of the study region by the seismic networks gives the density of data represented in figure (1.9). It confirms that the maximum concentration of data is reached by the set of stations settled in the Interandean volcanic region of Ecuador, notably by the stations represented by blue triangles in figure 1.9). This fact determines that is the north central region of Ecuador where we expect the best resolution of the tomographic study.

1.5 Filtering the data base

The data set that we obtained by merging different catalogues of unequal quality contains some outliers. A first tool to find these blunders is to perform a Wadati diagram. This diagram (figure 1.10) is obtained by comparing for each event and each pair of stations the difference in arrival times of P and S waves between the stations. Moreover it gives an estimation of the mean value of V_P/V_S in the domain of wave propagation.

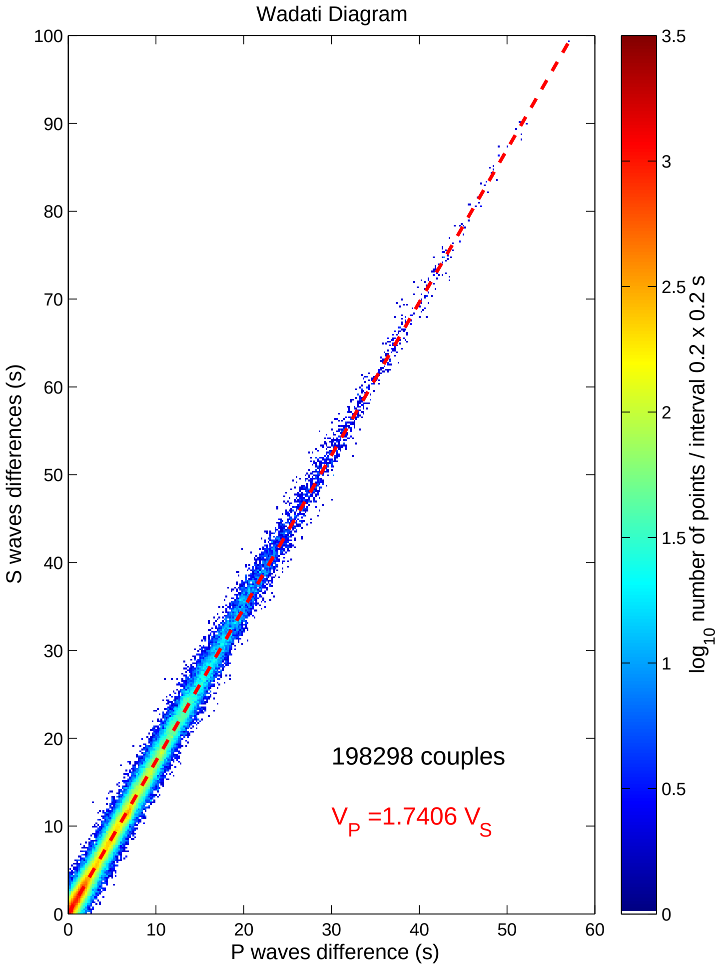

Figure 1.10: *Wadati diagram of S arrival time differences versus P differences between two stations. The color scale indicates the density of pairs of stations for which P and S phases were both measured. The straight line and its slope correspond to the linear regression of the point cloud.*

The arrival time of a wave at a station may be expressed as :

$$t = t_0 + \int_{\mathcal{R}} \frac{ds}{v(s)} \tag{1.1}$$

where t_0 is the initial time of the event, s is the curvilinear abscissa of the ray linking the source to the station, and v is the local velocity of the waves. Thus the difference of arrival times t_1 and t_2 of a same wave between two stations is expressed as :

$$t_1 - t_2 = \int_{\mathcal{R}_1} \frac{d(s)}{v(s)} - \int_{\mathcal{R}_2} \frac{ds}{v(s)} \tag{1.2}$$

Under the assumption that the ratio v_P/v_S is constant, which may be justified by the fact that this ratio remains relatively close to $\sqrt{3}$ in the lithosphere, the rays of P and S waves are identical. It then follows from (1.2) that :

$$t_1^S - t_2^S = \int_{\mathcal{R}_1} \frac{ds}{v_S(s)} - \int_{\mathcal{R}_2} \frac{ds}{v_S(s)} \simeq \frac{v_P}{v_S}\left[\int_{\mathcal{R}_1} \frac{ds}{v_P(s)} - \int_{\mathcal{R}_2} \frac{ds}{v_P(s)}\right] = \frac{v_P}{v_S}(t_1^P - t_2^P) \tag{1.3}$$

Consequently, representing for all events the values of $t_1^S - t_2^S$ versus the ones of $t_1^P – t_2^P$ for all pairs of stations that have recorded the P and S waves of this event, we obtain a point cloud in the neighborhood of a straight line passing through the origin. Moreover the slope of this line provides an estimation of the ratio v_P/v_S in the zone of wave propagation.

This representation also allows outliers to be detected. Such an outlier yields a set of points along a line that is parallel to the mean line of the cloud, but that does not pass by the origin. Moreover, each point of this set is associated to a pair containing the same station which is the one associated to the outlier. Then the picked phases must be revisited one by one and corrected when possible, for example when it corresponds to an error of minutes or to a change of day. The Figure 1.10 displays the Wadati diagram obtained after removing all the outliers and shows that the corresponding estimation of the ratio v_P/v_S is 1.74.

After this analysis, we have performed a localization of the events with the HypRef code and removed all the data (P or S) with a RMS greater than a value that we took equal to 7 after some tests. We remove all the events located outside the inversion box (Fig. 1.1), with a margin of tolerance of 0.1^o in each direction. We also removed all the events with less than 5 data or with less than 3 P data. Finally the data set consists of 404 097 P phases and 167 245 S phases corresponding to 45 941 events.

Chapter 2

The Bayesian approach of inverse problems

The probabilistic approach of the inversion uses the concepts of probability measure and of random elements to describe the level of available information on the observable (*i.e* physically measurable) parameters that constitute the data space, and on the non-measurable parameters that constitute the model space. The inversion of the new data, *i.e.* taking account of obervation of new physical quantities depending of the model through a physical theory, is achieved by using the concept of conditionnal probability. The probability measures are generally defined by probability density functions (*pdf*) in the case of finite dimentional spaces; which is always the case for the data space.

2.1 The various interpretations of the Bayesian approach of the inverse problem

There are basically two ways to interpret these probability measures (Valette 2011): the one named subjective and the other one objective.

In the subjective point of view, the probability measure is considered as strictly representing the state of information over the considered space at a given stage of knowledge. The probability density then represents the probability of occurence of the true value for the physical quantity we consider. It accounts either for the inaccuracy of the physical measurements in the case of data, or for the level of information on parameters of the model. The probability density function is then centered either on a mean physical measure in the case of a datum, or on an a priori value in the case of a parameter. Since the true value of the physical quantity is unknown, it is considered as a random variable associated with the probability measure. This does not mean that we consider the value of the physical quantity as the result of a random draw; this random variable simply represents the statistics of the possible real value around the measured or a priori value, at the stage of information considered. In this approach, the concept of probability measure representing the current state of information is highlighted rather than that of random variable representing the possible value of the physical quantity at the given stage of information.

In the objective point of view, the probability measure is considered as the law of a statistical estimator, that is to say a random variable, or more generally a random element that is associated to a random process of physical measurement or to a random draw, the expectation of which is the true value of the physical quantity. That means that the deterministic true value of the physical quantity is given by the expected value of the estimator, *i.e* that the estimator is unbiased. Of course, this true value is unknown and the only information available is the conditionnal law of the estimator of the physical quantity given its true value, along with a particular realization of this estimator. This realization is either the result of a physical measurement, considered as a random process, or the result of a prior inversion in the case of a chain of processing.

Thus the subjective interpretation highlights the concept of measure to represent the currently available information. Consequently, the associated random vector is centered on a mean value without any specific relation with the true value, whereas the objective interpretation prefers the concept of random estimator which is assumed to be unbiased, that is to say, centered on the unknown (but deterministic) true value.

As a result of these two points of view there are three ways to address the stochastic inverse problem that we will detail in the following paragraphs. We will denote by $\mathbb{D}$ the data space and by $\mathbb{M}$ the model space, *i.e.* the space of physical parameters that are involved in the considered experiment and that we want to infer, though they are not directly observable. Both of these spaces will be first assumed to be finite dimensional. Then we will consider the case of an infinite dimensional model. The corresponding generic data and model vectors will be denoted by $\mathbf{d} \in \mathbb{D}$ and $\mathbf{m} \in \mathbb{M}$.

2.1.1 The classical Bayesian approach

In the Bayesian standard approach $\mathbf{d}$ is considered as an estimator and $\mathbf{m}$ as a random vector associated to a probability measure of information. Thus we have the probability density $\rho_{prior}(\mathbf{m})$ that describes the a priori information on the model, on the one hand, and a particular realization $\mathbf{d}^{obs}$ of an estimator $\mathbf{d}$ with its conditional probability law $\rho(\mathbf{d}|\mathbf{d}^t)$ given the true data $\mathbf{d}^t$, on the other hand. If we consider that the theoretical mapping that links the model vector $\mathbf{m}$ to the true data vector $\mathbf{d}^t$ is exact,*i.e.* that it can be written as $\mathbf{d}^t = g(\mathbf{m})$, where g is a mapping from $\mathbb{M}$ into $\mathbb{D}$, then we get:

$$\rho(\mathbf{d}|\mathbf{m}) = \rho(\mathbf{d}|g(\mathbf{m})) \tag{2.1}$$

and this *pdf* only represents the uncertainty related to the process of physical measurement.

If now we consider that the theory is not exact, but we can represent it in the form of a conditional *pdf* $\Theta(\mathbf{d}'|\mathbf{m})$ given the true model $\mathbf{m}$, we get:

$$\rho(\mathbf{d}|\mathbf{m}) = \int_{\mathbb{D}} \rho_{obs}(\mathbf{d}|\mathbf{d}^t)\Theta(\mathbf{d}^t|\mathbf{m})d\mathbf{d}^t \tag{2.2}$$

It follows that in these two cases, we can deduce the joint probability density function of $(\mathbf{d}, \mathbf{m}) \in \mathbb{D} \times \mathbb{M}$:

$$\rho(\mathbf{d}, \mathbf{m}) = \rho(\mathbf{d}|\mathbf{m})\rho_{prior}(\mathbf{m}) \tag{2.3}$$

and the Bayes rule yields the expression of the posterior *pdf* of $\mathbf{m}$, *i.e* the conditionnal *pdf* of $\mathbf{m}$ given the particular realization $\mathbf{d}^{obs}$ of $\mathbf{d}$:

$$\rho_{post}(\mathbf{m}) \propto \rho(\mathbf{d}^{obs}|\mathbf{m})\rho_{prior}(\mathbf{m}) \tag{2.4}$$

where the proportionality accounts for the normalization to 1 of the *pdf*.

2.1.2 The approach by merging probability measures

In this approach, the data and the parameters are treated in the same manner, and are all considered as physical parameters with specific information. We have a priori measure of information over $\mathbb{D}$ and over $\mathbb{M}$ in the same way. The one over $\mathbb{D}$, with *pdf* $\rho_{obs}(\mathbf{d})$, results from a physical measurement made during the experiment we are studying; the other one over $\mathbb{M}$, with *pdf* $\rho_{prior}(\mathbf{m})$, represents the information available before considering the new dataset. As the measurement uncertainty of these data is clearly independent of the a priori information on the model, the two laws are independent in probability, and the joint *pdf* of $(\mathbf{m}, \mathbf{d})$ can be written as :

$$\rho(\mathbf{m}, \mathbf{d}) = \rho_{prior}(\mathbf{m})\rho_{obs}(\mathbf{d}) \tag{2.5}$$

Note that here $\mathbf{d}$ and $\mathbf{m}$ are the true data and model parameters, which are unknown and thus considered as random vectors at this stage of information.

If we assume that the theoretical mapping g connecting $\mathbf{m}$ to $\mathbf{d}$ is exact, then it can be written as $\mathbf{t} = \mathbf{d} - g(\mathbf{m}) = 0$. This amounts to considering that $\mathbf{t} \in \mathbb{D}$, which is a priori a random vector because it depends of the random vectors $\mathbf{d}$ and $\mathbf{m}$, is actually exactly null. It leads to consider the conditionnal *pdf* of $\mathbf{m}$ given that $\mathbf{t} = 0$ as posterior *pdf* over $\mathbb{M}$; which yields (Valette, 2011) :

$$\rho_{post}(\mathbf{m}) = \rho(\mathbf{m}|\mathbf{t} = 0) = \rho_{prior}(\mathbf{m})\rho_{obs}(g(\mathbf{m})) \tag{2.6}$$

which is equivalent to expression (2.4).

If we now consider that the theoretical mapping is not exact, but can be written as $\mathbf{t} = \mathbf{d} - g(\mathbf{m}) - \mathbf{n} = 0$, where $\mathbf{n}$ is a noise vector, with a *pdf* $\rho_T(\mathbf{n})$, which is independent in probability from $\mathbf{m}$ and $\mathbf{d}$, the joint *pdf* of $(\mathbf{m}, \mathbf{d}, \mathbf{n})$ is:

$$\rho(\mathbf{m}, \mathbf{d}, \mathbf{n})) = \rho_{prior}(\mathbf{m})\rho_{obs}(\mathbf{d})\rho_T(\mathbf{n}) \tag{2.7}$$

from which we can deduce by a change of variable that the joint *pdf* of $(\mathbf{m}, \mathbf{t}, \mathbf{n})$ is:

$$\rho(\mathbf{m}, \mathbf{t}, \mathbf{n})) = \rho_{prior}(\mathbf{m})\rho_{obs}(\mathbf{t} + g(\mathbf{m}) + \mathbf{n})\rho_T(\mathbf{n}) \tag{2.8}$$

It shows finally that the posterior *pdf* over $\mathbb{M}$, which is defined as the conditionnal *pdf* given that $\mathbf{t} = 0$ can be written as :

$$\rho_{post}(\mathbf{m}) = \rho(\mathbf{m}|\mathbf{t} = 0) = \rho_{prior}(\mathbf{m}) \int_{\mathbb{D}} \rho_{obs}(g(\mathbf{m}) + \mathbf{n})\rho_T(\mathbf{n})d\mathbf{n} \tag{2.9}$$

This latter expression is equivalent to the equations (2.2), (2.4), but here $\mathbf{d}^{obs}$ corresponds to the centre of ρ_{obs}.

If, more generally, we consider that the theory is represented by a *pdf* $\Theta(\mathbf{d}, \mathbf{m})$ in the product space $\mathbb{D} \times \mathbb{M}$, we are faced with two independent *pdf* in this product space: the one coming from the a priori information on parameters and data, the other one representing the physical theory. Tarantola and Valette (1982b) have proposed a method for merging these two pieces of information that is comparable in its results to the classical Bayesian approach.

2.1.3 The estimation approach

In the third approach, $\mathbf{d}^{obs}$ and $\mathbf{m}_{prior}$ are both considered as unbiased estimators, and a new estimator of the true model $\mathbf{m}$ is searched for, assuming that it is only function of $\mathbf{d}^{obs}$ and $\mathbf{m}_{prior}$, and it is unbiased with a law as much centered as possible. From a pratical point of view, this approach can be easily followed only in the case of a linear theoretical mapping, and of Gaussian estimators $\mathbf{d}^{obs}$ and $\mathbf{m}_{prior}$. The new posterior estimator $\mathbf{m}_{post}$ of $\mathbf{m}$ is then determined by supposing that: (i) it depends linearly of $\mathbf{d}^{obs}$ and $\mathbf{m}_{prior}$; (ii) it is unbiased; (iii) it minimizes the expected value of its squared deviation from the expected true model $E\left(||\mathbf{m}_{post} - \mathbf{m}_{true}||^2\right)$ among all the linear estimators. It is named the Best Linear Unbiased Estimator (BLUE) and is obtained through the same least squares algebra as the two previous approaches in this very restrictive case. Only the interpretation of $\mathbf{d}^{obs}$ and $\mathbf{m}_{prior}$ differs.

2.2 Bayesian approach and data differences

In many fields, it is more interesting to inverse the differences of data rather than the raw data. This may indeed lead to weaken and even sometimes to remove the influence of some parameters, while strengthening the influence of others, and consequently their resolution.

Let us consider n physical quantities, and an estimator $\mathbf{d}$ associated to a measurement process of these quantities, $\mathbf{d} \in \mathbb{R}^n \equiv \mathbb{D}$, with conditional *pdf* given the true vector $\mathbf{d}^t$ as:

$$\rho(\mathbf{d}|\mathbf{d}^t) = \varphi(\mathbf{d} - \mathbf{d}^t) \tag{2.10}$$

where φ is a function from $\mathbb{R}^n$ to $\mathbb{R}^+$.

Let us now suppose that we are only interested in the data differences, instead of the raw data, and let us define the new vector $\mathbf{d}' \in \mathbb{R}^{n-1}$ by:

$$d'_i = d_{i+1} - d_i \qquad\qquad 1 \leq i \leq n-1 \tag{2.11}$$

so that:

$$d_i = d_1 + \sum_{k=1}^{i-1} d'_k \qquad\qquad 1 \leq i \leq n \tag{2.12}$$

We can easily deduce the image *pdf* of $\mathbf{d}$ through the mapping (2.11) to obtain the conditionnal *pdf* of $\mathbf{d}'$ given $\mathbf{d}^t$. It yields:

$$\begin{aligned} \rho(\mathbf{d}'|d^t) =& \int_{\mathbb{R}} \rho(s, ..., s + \sum_{k=1}^{i-1} d'_k, ...|\mathbf{d}^t) ds \\ =& \int_{\mathbb{R}} \varphi(s - d^t_1, ..., s + \sum_{k=1}^{i-1} d'_k - d^t_i, ...) ds \end{aligned} \tag{2.13}$$

and, by making the change of variable $t = s - d^t_1$:

$$\begin{aligned} \rho(\mathbf{d}'|d^t) =& \int_{\mathbb{R}} \rho(t + d_1, ..., t + d_i, ..., t + d_n|\mathbf{d}^t) dt \\ =& \int_{\mathbb{R}} \varphi(t + d_1 - d^t_1, ..., t + d_i - d^t_i, ..., t + d_n - d^t_n|\mathbf{d}^t) dt \end{aligned} \tag{2.14}$$

This shows that the *pdf* of the data differences $\mathbf{d}'$ given the true value of $\mathbf{d}^t$ is deduced by a simple integration of the one of $\mathbf{d}$

Now let us suppose that the vector $\mathbf{d}$ depends on the model $\mathbf{m} \in \mathbb{M}$ we want to infer, through an exact theoretical mapping g. When considering raw data, the standard Bayesian approach leads to (see 2.4):

$$\rho_{post}(\mathbf{m}) \propto \rho(\mathbf{d}^{obs}|\mathbf{m})\rho_{prior}(\mathbf{m}) = \varphi(\mathbf{d}^{obs} - g(\mathbf{m}))\rho_{prior}(\mathbf{m}) \tag{2.15}$$

where $\mathbf{d}^{obs}$ is the observed realization of the estimator $\mathbf{d}$. Inverting now data differences, the Bayesian approach yields :

$$\begin{aligned}\rho_{post}(\mathbf{m}) &\propto \rho_{prior}(\mathbf{m})\rho(\mathbf{d}'^{obs}|\mathbf{m})\rho_{prior}(\mathbf{m}) = \rho(\mathbf{d}'^{obs}|\mathbf{d}^t = g(\mathbf{m})) \\ &\propto \rho_{prior}(\mathbf{m})\int\limits_{\mathbb{R}} \varphi(d_1^{obs} - g_1(\mathbf{m}) + t, ..., d_i^{obs} - g_i(\mathbf{m}) + t, ..., d_n^{obs} - g_n(\mathbf{m}) + t)dt\end{aligned} \tag{2.16}$$

in taking the relation (2.9) between $\mathbf{d}'^{obs}$ and $\mathbf{d}^{obs}$ into account.

Thus we see that considering data differences, when tranferring the probability measure, leads to a new posterior *pdf* on $\mathbf{m}$ which can be easily deduced by a simple integration of the *pdf* obtained from the raw data.

Adopting the point of view of measure of information for the data, *i.e.* considering a *pdf* $\rho_{obs}(\mathbf{d}) = \varphi(\mathbf{d} - \mathbf{d}^{obs})$ representing the possible true values of $\mathbf{d}$ around the mean vector $\mathbf{d}^{obs}$ instead of the conditionnal *pdf* of an unbiased estimator given the true vector, we can obtain in the same way that :

$$\begin{aligned}\rho_{obs}(\mathbf{d}') &= \int\limits_{\mathbb{R}} \rho_{obs}(t + d_1, ..., t + d_i, ..., t + d_n)dt \\ &= \int\limits_{\mathbb{R}} \varphi(t + d_1 - d_1^{obs}, ..., t + d_i - d_i^{obs}, ..., t + d_n - d_n^{t}|\mathbf{d}^{obs})dt\end{aligned} \tag{2.17}$$

and :

$$\rho_{post}(\mathbf{m}) \propto \rho_{prior}(\mathbf{m})\int\limits_{\mathbb{R}} \varphi(g_1(\mathbf{m}) - d_1^{obs} + t, ..., g_i(\mathbf{m}) - d_i^{obs} + t, ..., g_n(\mathbf{m}) - d_n^{obs} + t)dt \tag{2.18}$$

which is equivalent to the previous results. Only the interpretation of the ingredients differs.

2.3 Exploring the a posteriori density function

We have seen that the Bayesian approach of the inverse problem yields an a posteriori *pdf* over the model space once given the prior information on the model and data. This posterior *pdf* may be considered as the final result of the inversion; nevertheless it remains to explore it. In the case of a model space of small dimension, it is very easy to display the various marginal density of physical interest. For intermediate dimensions, random Markov walks or simulated anhealing techniques can be used to describe the *pdf* or to determine maximum likelihood model, *i.e.* the set of values of the model parameters for which the *pdf* reaches its maximum (see *e.g.* Mosegaard and Tarantola, 2002, for a review). In the case of very large dimensions, the use of analytical techniques is the only possible way to search for local maximum likelihoods. Considering Gaussian laws is very useful in this approach.

2.3.1 Gaussian random vectors

A random vector $\mathbf{x}$ is Gaussian, or has a Gaussian *pdf* $\rho(\mathbf{x})$, when:

$$\rho(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n det(\mathbf{C})}} \exp\left\{ -\frac{1}{2}(\mathbf{C}^{-1}(\mathbf{x}-\mathbf{x}_0)|\mathbf{x}-\mathbf{x}_0) \right\} \tag{2.19}$$

where $\mathbf{x}_0$ is the mean vector, *i.e.* the expectation of $\mathbf{x}$, and $\mathbf{C}$ its covariance operator in the n-dimensional space $\mathbf{X}$, and where *det* denotes the determinent. Thus, a Gaussian vector is fully described by its expectation and its covariance.

2.3.2 Searching for maxima of the posterior *pdf*

Of particular interest are the area in $\mathbb{M}$ where the values of the a posteriori *pdf* are large. A first step in the study of the posterior density is thus to identify the models $\hat{\mathbf{m}}$, commonly refered to as maximum likelihood, for which the density is locally maximum. Taking advantage of the positivity of the density, we can rewrite it as:

$$\rho_{post}(\mathbf{m}) \propto \exp(\text{- E}(\mathbf{m})) \tag{2.20}$$

where the scalar function $\mathrm{E}(\mathbf{m})$ will be assumed twice continuously differentiable, *i.e.* g twice continuously differentiable. A local maximum of ρ_{post} corresponds to a local minimum of E, for which the gradient of E vanishes. Let us consider a model $\hat{\mathbf{m}}$ that makes vanish grad(E($\mathbf{m}$)), and write the Taylor expansion of E at the second order around it. It leads to:

$$\mathrm{E}(\mathbf{m}) - \mathrm{E}(\hat{\mathbf{m}}) = \frac{1}{2}(\mathbf{H}(\hat{\mathbf{m}})(\mathbf{m}-\hat{\mathbf{m}})|\mathbf{m}-\hat{\mathbf{m}}) + o(\|\mathbf{m}-\hat{\mathbf{m}}\|^2) \tag{2.21}$$

where $\mathbf{H}$ denotes the Hessian of E, *i.e.* the derivative of the gradient, which can be written in matrix notation as:

$$\mathrm{H}_{i,j} = \frac{\partial^2 \mathrm{E}}{\partial m_i \partial m_j}(\hat{\mathbf{m}})$$

and where $o(\|\mathbf{m}-\hat{\mathbf{m}}\|^2)$ denotes a function such that $o(\|\mathbf{m}-\hat{\mathbf{m}}\|^2)/\|\mathbf{m}-\hat{\mathbf{m}}\|^2$ vanishes with $\|\mathbf{m}-\hat{\mathbf{m}}\|$. $\mathbf{H}(\hat{\mathbf{m}})$ is a self-adjoint (symmetrical) operator, which is positive when E is minimum at $\hat{\mathbf{m}}$, as can be deduced from (2.21). In that case, substituting (2.21) into (2.20) yields :

$$\rho_{post}(\mathbf{m}) \propto \exp(o(\|\mathbf{m}-\hat{\mathbf{m}}\|^2))\exp\left\{\frac{1}{2}(\mathbf{H}(\hat{\mathbf{m}})(\mathbf{m}-\hat{\mathbf{m}})|\mathbf{m}-\hat{\mathbf{m}})\right\}$$

which shows that the posterior *pdf* is locally tangent to a Gaussian density, the expectation and covariance of which are respectively the maximum likelihood $\hat{\mathbf{m}}$ and the inverse $\mathbf{H}(\hat{\mathbf{m}})^{-1}$ of the Hessian. Thus, when the posterior *pdf* is sufficiently regular to present only one maximum, or a few maxima, it can be correctly approximate by a few Gaussian distributions. The point is then to determine the models for which these maxima are reached.

2.4 The Gaussian case

We will now assume that both the data and the model vector $\mathbf{d}$, and $\mathbf{m}$ are Gaussian with expectations $\mathbf{d}_{obs}$ and $\mathbf{m}_{prior}$, and with covariances $\mathbf{C}_d$ and $\mathbf{C}_m$, respectively. By a change of variable it is always possible to set the inverse problem in this framework.

Taking expressions (2.19), (2.21), and (2.4) or (2.6) into account, we deduce that:

$$\mathrm{E}(\mathbf{m}) = \frac{1}{2}\left\{(\mathbf{C}_d^{-1}(\mathbf{g}(\mathbf{m})-\mathbf{d}_{obs})|\mathbf{g}(\mathbf{m})-\mathbf{d}_{obs}) + (\mathbf{C}_m^{-1}(\mathbf{m}-\mathbf{m}_{prior})|\mathbf{m}-\mathbf{m}_{prior})\right\} \quad (2.22)$$

Assuming that the theoretical function $\mathbf{g}$ is continuously differentiable, and denoting by $\mathbf{G}_m$ its derivative operator at model parameter $\mathbf{m}$, the stationnarity equation that corresponds to the nullity of the gradient :

$$\nabla_m(\mathrm{E}) = \mathbf{G}_m^*\mathbf{C}_d^{-1}(\mathbf{g}(\mathbf{m})-\mathbf{d}_{obs}) + \mathbf{C}_m^{-1}(\mathbf{m}-\mathbf{m}_{prior}) \quad (2.23)$$

can be written as the following implicit equation:

$$\mathbf{m}-\mathbf{m}_{prior} = \mathbf{C}_m\mathbf{G}_m^*\mathbf{C}_d^{-1}(\mathbf{d}_{obs}-\mathbf{g}(\mathbf{m})) \quad (2.24)$$

where $\mathbf{G}_m^*$ denotes the adjoint of $\mathbf{G}_m$ with respect to the usual scalar products in $\mathbb{D}$ and $\mathbb{M}$, which verifies :

$$(\mathbf{G}_m\mathbf{v}|\mathbf{u})_{\mathbb{D}} = (\mathbf{v}|\mathbf{G}_m^*\mathbf{u})_{\mathbb{M}} \quad \forall(\mathbf{v},\mathbf{u}) \in \mathbb{M}\times\mathbb{D}$$

As shown by (Tarantola and Valette 1982a, Valette 2011) this equation may be equivalently written in the following form:

$$\mathbf{m}-\mathbf{m}_{prior} = \mathbf{C}_m\mathbf{G}_m^*(\mathbf{C}_d+\mathbf{G}_m\mathbf{C}_m\mathbf{G}_m^*)^{-1}(\mathbf{d}_{obs}-\mathbf{g}(\mathbf{m})+\mathbf{G}_m(\mathbf{m}-\mathbf{m}_{prior})) \quad (2.25)$$

which becomes explicit in the linear case, for which $\mathbf{g}(\delta\mathbf{m}) \equiv \mathbf{G}_m\delta\mathbf{m} \equiv \mathbf{G}\delta\mathbf{m}$. Thus, for weakly non linear $\mathbf{g}$, the equation can be easily solved by means of a fixed point algorithm that takes the following form :

$$\mathbf{m}_{k+1} - \mathbf{m}_{prior} = \mathbf{C}_m\mathbf{G}_k(\mathbf{C}_d + \mathbf{G}_k\mathbf{C}_m\mathbf{G}_k a)^{-1}(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_k) + \mathbf{G}_k(\mathbf{m}_k - \mathbf{m}_{prior})) \quad (2.26)$$

where $\mathbf{G}_k \equiv \mathbf{G}_{m_k}$, and which can be initialized at different $\mathbf{m}_0$, and in particular at $\mathbf{m}_{prior}$. Each iteration imposes to solve a linear system in the data space $\mathbb{D}$. But, by taking the following standard formulae into account:

$$\mathbf{C}_m\mathbf{G}^*(\mathbf{C}_d + \mathbf{G}\mathbf{C}_m\mathbf{G}^*)^{-1} = (\mathbf{C}_m^{-1} + \mathbf{G}^*\mathbf{C}_d^{-1}\mathbf{G})^{-1}\mathbf{G}^*\mathbf{C}_d^{-1} \quad (2.27)$$

$$\mathbf{C}_m - \mathbf{C}_m\mathbf{G}^*(\mathbf{C}_d + \mathbf{G}\mathbf{C}_m\mathbf{G}^*)^{-1}\mathbf{G}\mathbf{C}_m = (\mathbf{C}_m^{-1} + \mathbf{G}^*\mathbf{C}_d^{-1}\mathbf{G})^{-1} \quad (2.28)$$

the algorithm can be rewritten as :

$$\mathbf{m}_{k+1} - \mathbf{m}_k = -(\mathbf{C}_m^{-1} + \mathbf{G}_k^*\mathbf{C}_d^{-1}\mathbf{G}_k)^{-1}(\mathbf{G}_k^*\mathbf{C}_d^{-1}(\mathbf{g}(\mathbf{m}_k) - \mathbf{d}_{obs}) + \mathbf{C}_m^{-1}(\mathbf{m}_k - \mathbf{m}_{prior})) \quad (2.29)$$

From the expression (2.23) of the gradient of E, we deduce by derivation that the Hessian of E is approximately :

$$\mathbf{H}(\mathbf{m}) \simeq \mathbf{G}_m^*\mathbf{C}_d^{-1}\mathbf{G}_m + \mathbf{C}_m^{-1} \quad (2.30)$$

in neglecting the term of second order related to the derivative of $\mathbf{G}_m^*$. This shows that, in its formulation (2.29), the algorithm is a quasi Newton algorithm, implying the resolution of a linear system in the model space $\mathbb{M}$ at each iteration. We can also conclude that if $\mathbf{H}(\mathbf{m})$ is definite positive the posterior *pdf* is locally tangent to the Gaussian distribution, the expectation of which is the model $\hat{\mathbf{m}}$, limit of the algoritm (2.26), (2.29), and solution of (2.24) or (2.25), and the covariance of which is the inverse of the Hessian at $\hat{\mathbf{m}}$:

$$\begin{aligned}\hat{\mathbf{C}}_m = \mathbf{H}(\hat{\mathbf{m}})^{-1} &\simeq (\mathbf{C}_m^{-1} + \mathbf{G}_{\hat{m}}^*\mathbf{C}_d^{-1}\mathbf{G}_{\hat{m}})^{-1} \\ &= \mathbf{C}_m - \mathbf{C}_m\mathbf{G}_{\hat{m}}^*(\mathbf{C}_d + \mathbf{G}_{\hat{m}}\mathbf{C}_m\mathbf{G}_{\hat{m}}^*)^{-1}\mathbf{G}_{\hat{m}}\mathbf{C}_m\end{aligned} \quad (2.31)$$

It is important to remark that this evaluation of the uncertainty through the a posteriori covariance of the model is correct only if the a priori covariance truly corresponds to physical properties, and has been correctly estimated. In the case of tomography problems, the a priori covariance operator is designed in order to smooth the posterior resulting model and to avoid numerical oscillations, while obtaining a good fit of the data. It must therefore be clearly distinguished from the physical covariance operator from which it is deduced through a renormalization operation depending on the length of smoothing, as we will see later on. Furthermore, the computation of the posterior covariance implies the inversion of a matrix in the model space which is very time expensive, or even impossible, in the case of a very large dimension. In that case of tomography problems, it is much more judicious to use the concept of resolution operator, introduced by Backus and Gilbert (1970), in order to evaluate the degree of confidence in the reliability of the posterior model.

2.4.1 The resolution operator

This operator links the a posteriori model $\hat{\mathbf{m}}$ to the true one $\mathbf{m}$ at first order, in the case of a non linear mapping $\mathbf{g}$. It is easily obtained through a first order development of $\mathbf{d} = g(\mathbf{m})$ in the neighbourhood of $\mathbf{m} = \hat{\mathbf{m}}$ and by remarking that the true data $\mathbf{d}$ (or $\mathbf{d}^t$ when adopting the classical Bayesian approach) is equal to $g(\mathbf{m})$ where $\mathbf{m}$ is the true model, since the physical theory is assumed to be perfectly exact. It yields:

$$\mathbf{d}_{obs} - \mathbf{g}(\hat{\mathbf{m}}) + \mathbf{G}_{\hat{m}}(\hat{\mathbf{m}} - \mathbf{m}_{prior}) = \mathbf{d}_{obs} - \mathbf{d} + \mathbf{G}_{\hat{m}}(\mathbf{m} - \mathbf{m}_{prior}) + o(\|\mathbf{m} - \hat{\mathbf{m}}\|)$$

and, upon substituting it into the stationnarity equation (2.25):

$$\hat{\mathbf{m}} - \mathbf{m}_{prior} = \mathbf{R}_{\hat{m}}(\mathbf{m} - \mathbf{m}_{prior}) - \mathbf{K}_{\hat{m}}(\mathbf{d} - \mathbf{d}_{obs}) + o(\|\hat{\mathbf{m}} - \mathbf{m}\|) \tag{2.32}$$

where:

$$\mathbf{K}_{\hat{m}} = \mathbf{C}_m \mathbf{G}^*_{\hat{m}} (\mathbf{C}_d + \mathbf{G}_{\hat{m}} \mathbf{C}_m \mathbf{G}^*_{\hat{m}})^{-1}$$

and where

$$\begin{aligned} \mathbf{R}_{\hat{m}} = \mathbf{K}_{\hat{m}} \mathbf{G}_{\hat{m}} &= \mathbf{C}_m \mathbf{G}^*_{\hat{m}} (\mathbf{C}_d + \mathbf{G}_{\hat{m}} \mathbf{C}_m \mathbf{G}^*_{\hat{m}})^{-1} \mathbf{G}_{\hat{m}} \\ &= (\mathbf{C}_m^{-1} + \mathbf{G}^*_{\hat{m}} \mathbf{C}_d^{-1} \mathbf{G}_{\hat{m}})^{-1} \mathbf{G}^*_{\hat{m}} \mathbf{C}_d^{-1} \mathbf{G}_{\hat{m}} \end{aligned} \tag{2.33}$$

is the resolution operator.

Equation (2.32) shows that, correct to first order, the improvement in the prior model, $\hat{\mathbf{m}} - \mathbf{m}_{prior}$, is equal to the true one, $\mathbf{m} - \mathbf{m}_{prior}$, filtered by the resolution operator $\mathbf{R}_{\hat{m}}$, plus a noise term proportional to the observational errors. Thus, we can see that the posterior model is the result of a trade of between the discrepency of the resolution operator with the identity, on the one hand, and the amplitude of the noise term coming from the errors in the data, in the other hand :

$$\hat{\mathbf{m}} - \mathbf{m} = (\mathbf{R}_{\hat{m}} - \mathbf{I}_d)(\mathbf{m} - \mathbf{m}_{prior}) - \mathbf{K}_{\hat{m}}(\mathbf{d} - \mathbf{d}_{obs}) + o(\|\hat{\mathbf{m}} - \mathbf{m}\|)$$

It is also important to notice that, in the case of several maxima of the posterior *pdf*, the resolution operator associated to each of this maxima only describes the resolving quality of the corresponding possible ”solution”.

2.4.2 Inverting data differences in the Gaussian framework

We have seen in section 2.2 that inverting differences of data $\mathbf{d}'$, when transporting correctly the statistics of errors, amounts to the change of the *pdf* of errors in the data through an integration, while keeping the raw data $\mathbf{d}$ in the inversion process (equation 2.17, or 2.14 for the classical Bayesian approach). In the framework of a Gaussian law for the data, we can explicitely compute the *pdf* accounting for the data differences :

$$\rho_{obs}(\mathbf{d}') = \int_{\mathbb{R}} \rho_{obs}(t + d_1, ..., t + d_i, ..., t + d_n) dt$$

Assuming that the vector $\mathbf{d}$ is gaussian in the n dimensional space $\mathbb{D}$, with expectation $\mathbf{d}^{obs}$ and covariance operator $\mathbf{C}$, (2.34) can be rewritten as:

$$\rho_{obs}(\mathbf{d}') \propto \int_{\mathbb{R}} \exp\left[-\frac{1}{2}\left(\mathbf{C}^{-1}(\boldsymbol{\delta} - t\mathbf{u})\middle|\boldsymbol{\delta} - t\mathbf{u}\right)\right] \mathrm{d}t$$

where $\boldsymbol{\delta} = \mathbf{d} - \mathbf{d}^{obs}$ and $\mathbf{u} \in \mathbb{D}$, with : $\mathbf{u}_i = 1$ $(i = 1, ..., n)$. Following a calculation similar to that of the appendice of (Tarantola and Valette, 1982b), it successively leads to:

$$\begin{aligned}\rho_{obs}(\mathbf{d}') &\propto \int_{\mathbb{R}} \exp\left[-\frac{1}{2}\left(at^2 - 2b(\boldsymbol{\delta})t + c(\boldsymbol{\delta})\right)\right] \mathrm{d}t \\ &= \int_{\mathbb{R}} \exp\left[-\frac{1}{2}\left(a\left(t - \frac{b(\boldsymbol{\delta})}{a}\right)^2 + c(\boldsymbol{\delta}) - \frac{b^2(\boldsymbol{\delta})}{a}\right)\right] \mathrm{d}t\end{aligned}$$

with $a = (\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})$, $b(\boldsymbol{\delta}) = (\mathbf{C}^{-1}\mathbf{u}|\boldsymbol{\delta})$, $c(\boldsymbol{\delta} = (\mathbf{C}^{-1}\boldsymbol{\delta}|\boldsymbol{\delta})$, and to:

$$\rho_{obs}(\mathbf{d}') \propto \exp\left[-\frac{1}{2}\left(c(\boldsymbol{\delta}) - \frac{b^2(\boldsymbol{\delta})}{a}\right)\right]$$

which can be rewritten as:

$$\rho_{obs}(\mathbf{d}') \propto \exp\left[-\frac{1}{2}\left((\mathbf{C}^{-1}\boldsymbol{\delta}|\boldsymbol{\delta}) - \frac{(\mathbf{C}^{-1}\boldsymbol{\delta}|\mathbf{u})^2}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})}\right)\right] \tag{2.34}$$

The cost function which appears in (2.34):

$$\mathcal{J}(\boldsymbol{\delta}) = \left(\mathbf{C}^{-1}\boldsymbol{\delta}\middle|\boldsymbol{\delta}\right) - \frac{(\mathbf{C}^{-1}\boldsymbol{\delta}|\mathbf{u})^2}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})} \tag{2.35}$$

can be rewritten in two more significant ways. Firstly, a simple development yields:

$$\mathcal{J}(\boldsymbol{\delta}) = \left(\mathbf{C}^{-1}\left(\boldsymbol{\delta} - \frac{(\mathbf{C}^{-1}\boldsymbol{\delta}|\mathbf{u})}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})}\mathbf{u}\right)\middle|\boldsymbol{\delta} - \frac{(\mathbf{C}^{-1}\boldsymbol{\delta}|\mathbf{u})}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})}\mathbf{u}\right) \tag{2.36}$$

But, we can write it in an alternative way. Denoting by $\mathbf{v}\otimes\mathbf{v}$ the operator acting in $\mathbb{D}$, defined by:

$$\forall\, \mathbf{w} \in \mathbb{D},\ \mathbf{w} \longrightarrow \mathbf{v}\otimes\mathbf{v}(\mathbf{w}) = (\mathbf{w}|\mathbf{v})\,\mathbf{v}$$

and putting :

$$\mathbf{v} = \frac{\mathbf{C}^{-1/2}\mathbf{u}}{\|\mathbf{C}^{-1/2}\mathbf{u}\|}$$

we can deduce that:

$$\frac{(\mathbf{C}^{-1}\boldsymbol{\delta}|\mathbf{u})^2}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})} = \left(\mathbf{C}^{-1/2}(\mathbf{v}\otimes\mathbf{v})\mathbf{C}^{-1/2}\boldsymbol{\delta}\middle|\boldsymbol{\delta}\right) \tag{2.37}$$

since

$$\mathbf{C}^{-1/2}(\mathbf{v}\otimes\mathbf{v})\mathbf{C}^{-1/2}\boldsymbol{\delta} = \mathbf{C}^{-1/2}\frac{\mathbf{C}^{-1/2}\mathbf{u}\otimes\mathbf{C}^{-1/2}\mathbf{u}}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})}\mathbf{C}^{-1/2}\boldsymbol{\delta} = \mathbf{C}^{-1}\mathbf{u}\frac{(\mathbf{C}^{-1}\boldsymbol{\delta}|\mathbf{u})}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})}$$

Taking (2.35) and (2.37) into account it follows that $\mathcal{J}(\boldsymbol{\delta})$ may be rewritten as:

$$\mathcal{J}(\boldsymbol{\delta}) = \left(\mathbf{C}^{-1/2}(\mathrm{I} - \mathbf{v}\otimes\mathbf{v})\mathbf{C}^{-1/2}\boldsymbol{\delta}\middle|\boldsymbol{\delta}\right) \tag{2.38}$$

The operator:

$$\mathbf{P} = \mathbf{I} - \mathbf{v}\otimes\mathbf{v} \tag{2.39}$$

is an orthogonal projector in $\mathbb{D}$ since $\mathbf{P}^2 = \mathbf{P}$ and it is selfadjoint (or symmetrical) for the usual scalar product of $\mathbb{D}$. Moreover, its null space, which corresponds to its direction of projection, is the 1-D subspace $\mathbb{R}\mathbf{C}^{-1/2}\mathbf{u}$ in $\mathbb{D}$, since:

$$\mathbf{P}\mathbf{w} = 0 \Leftrightarrow \mathbf{w} = \mathbf{v}\otimes\mathbf{v}(\mathbf{w}) = (\mathbf{v}|\mathbf{w})\,\mathbf{v} = \frac{\left(\mathbf{C}^{-1/2}\mathbf{u}\middle|\mathbf{w}\right)}{(\mathbf{C}^{-1/2}\mathbf{u}|\mathbf{u})}\mathbf{C}^{-1/2}\mathbf{u}$$

Thus, the operator $\mathbf{C}^{-1/2}\mathrm{P}\mathbf{C}^{-1/2}$ is self-adjoint and positive, but not definite positive as its null space is the 1-D subspace $\mathbb{R}\mathbf{u}$:

$$\mathbf{C}^{-1/2}\mathbf{P}\mathbf{C}^{-1/2}\mathbf{w} = 0 \Leftrightarrow \mathbf{P}\mathbf{C}^{-1/2} = 0 \Leftrightarrow \mathbf{C}^{-1/2}\mathbf{w} \in \mathbb{R}\mathbf{C}^{-1/2}\mathbf{u} \Leftrightarrow \mathbf{w} \in \mathbb{R}\mathbf{u}$$

This shows that considering differences of the data only amounts to the assignation of a new inverse covariance operator to the raw data, such that adding a constant term to each component of the data vector $\mathbf{d}$ does not change the cost function. This can be simply understood since adding such a term does not change the data differences. Actually, it amounts evaluating the original cost function only within the hyperplan orthogonal to vector $\mathbf{u}$ in the data space $\mathbb{D}$, as can be shown by:

$$\mathbf{P}^2 = \mathbf{P} \quad \mathbf{P}^* = \mathbf{P} \Longrightarrow \mathcal{J}(\boldsymbol{\delta}) = \left(\mathbf{C}^{-1/2}\mathbf{P}\mathbf{C}^{-1/2}\boldsymbol{\delta}\middle|\boldsymbol{\delta}\right) = \left(\mathbf{P}\mathbf{C}^{-1/2}\boldsymbol{\delta}\middle|\mathbf{P}\mathbf{C}^{-1/2}\boldsymbol{\delta}\right)$$

In summary, the posterior *pdf* of $\mathbf{d}'$ is Gaussian, degenerate in $\mathbb{D}$:

$$\rho_{obs}(\mathbf{d}') \propto \int\limits_{\mathbb{R}} \exp\left[-\frac{1}{2}\left(\mathbf{D}(\mathbf{d}-\mathbf{d}_{obs})|\mathbf{d}-\mathbf{d}_{obs}\right)\right] \mathrm{d}t$$

where the initial operator $\mathbf{C}^{-1}$ relevant for the raw data is replaced by the operator :

$$\mathbf{D} = \mathbf{C}^{-1/2}\mathbf{P}\mathbf{C}^{-1/2} = \left(\mathbf{P}\mathbf{C}^{-1/2}\right)^* \mathbf{P}\mathbf{C}^{-1/2}\text{with: } \mathbf{P} = \mathbf{I} - \frac{\mathbf{C}^{-1/2}\mathbf{u}\otimes\mathbf{C}^{-1/2}\mathbf{u}}{(\mathbf{C}^{-1}\mathbf{u}|\mathbf{u})} \tag{2.40}$$

2.4.3 Accounting for non Gaussian statistics in the data and robustness

Let us now show how we can recast non Gaussian statistics into the Gaussian framework. It is especially important for data. Indeed, in some specific cases the *pdf* corresponding to a physical quantity is obtained through repeated measurements, and is thus the interpolation of a sampling. It may also present two maxima when some uncertainty exists either between two possible mean values depending on other data, or in the interpretation of the measurement. In the case of large datasets, there exist surely outliers due to errors in measurement, and copy or transcription errors in files. A gaussian *pdf* cannot account for these blunders.

Moreover, from a practical perspective, the least squares approach, which consists in minimizing the cost function (2.22) in the Gaussian framework, is not robust in the sense that an outlier strongly affects the inverted model. This point can be simply undertstood since the gradient (2.23) of the cost function is proportionnal to the misfit between the theoretical value corresponding to a component of $\mathbf{g}(\mathbf{m})$ (2.23) and the outlier corresponding to the same component in $\mathbf{d}_{obs}$. Thus, the outliers play a significant role in the equation of stationnarity imposing the nullity of the gradient, and consequently bias its solution.

There exist a lot of solution to enhance the robustness of the inversion process. A well known solution consists in using Huber M-estimators (Huber 1964), which amounts to minimizing the following cost for a single datum d:

$$\mathrm{E}(d) = \begin{cases} (d - d_{obs})^2/2 & \text{if } |d - d_{obs}| \leq k \\ & \\ k(|d - d_{obs}| - k/2) & \text{if } |d - d_{obs}| \geq k \end{cases} \qquad \text{with } k \geq 0 \tag{2.41}$$

instead of the usual quadratic cost everywhere. In this way, the derivative of the cost remains finite and constant for large value of $|d - d_{obs}|$. It is also well known that the use of exponential law yields more robust estimation. For instance, in the case of 1-D distributions it yields the median instead of the mean value. An alternative way is thus to consider *pdf* that are tangeant to a Gaussian one near the origin and to an exponential one at infinity, as the hyperbolic secant *pdf* (Crase et al., 1990), defined as:

$$f(d) = \frac{1}{\pi\sigma} \frac{1}{\cosh\left(\frac{d - d_{obs}}{\sigma}\right)} \tag{2.42}$$

More generally, let us consider that the information on a particular physical quantity x is represented by the *pdf* $\rho(x)$, and let us search for (Valette 2011) a new variable $y(x)$ such that the pushforward *pdf* of y is Gaussian centered. Assuming that y is a non decreasing function of x, it leads to:

$$\rho(x)\,\mathrm{d}x = \frac{1}{\sqrt{\pi}} exp(-y^2)\,\mathrm{d}y \tag{2.43}$$

then, by integrating:

$$\int_{x_0}^{x} f(t)dt = \frac{1}{\sqrt{\pi}} \int_0^{y(x)} \exp(-t^2)dt = \frac{1}{2}\mathrm{Erf}(y) \tag{2.44}$$

where x_0 is the median of the x distribution and Erf denotes the error function defined as:

$$\mathrm{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^{x} \exp(-t^2)\,\mathrm{d}t$$

We finally obtain that:

$$y = \mathrm{Erf}^{-1}\left[2\int_{x_0}^{x} f(t)dt\right] \tag{2.45}$$

and check a posteriori that this change of variable is strictly increasing with x. More precisely (2.43) yields:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \sqrt{\pi}\exp(y^2)\rho(x) \tag{2.46}$$

The inverse problem can then be recast with this new variable $y = y(x(\mathbf{m}))$, the *pdf* of which is Gaussian, and that corresponds in probability to the exact law of the raw datum x. The derivative of y with respect to the model $\mathbf{m}$ is readily obtained by using the chain rule and (2.46).

2.5 Infinite dimensionnal model spaces in the Gaussian framework

As noted by (Valette 2011) the issue of inverse problems follows various lines according to the kind of physical problems that are considered. For instance, in the case of dynamical sytems, the theory of control provides means to compute easily the adjoint operator that is associated to an adjoint problem and to back propagation. In the meteorology forecasting problem, the data must be incorporated time after time. This leads to a sequential analysis of the data space, which is the topic of data assimilation. We are here interested in tomography problems, for which the model parameters that we want to identify depend on the usual space ($\mathbb{R}^3$). The model space thus consists of random functions of space. Let us first recall briefly the definition and main properties of Gaussian random functions (Loève, 1965; Neveu, 1968). We will follow Valette (2011) .

2.5.1 Gaussian random functions

A random function is a set of random variables m($\mathbf{r}$) depending on the position $\mathbf{r}$ within a domain V $\subset \mathbb{R}^q$. The random function is Gaussian if for any integer n and any positions $\mathbf{r}_1$,$\mathbf{r}_2$,..., $\mathbf{r}_n$, the joint *pdf* of the variables m($\mathbf{r}_1$),m($\mathbf{r}_2$),...,m($\mathbf{r}_n$) is Gaussian. A Gaussian function can thus be regarded as the generalisation of a Gaussian vector to a continuous index. Given a Gaussian function, we can define the covariance function (or kernel) $C(\mathbf{r}$,$\mathbf{r}$') which is the covariance of m($\mathbf{r}$) and m($\mathbf{r}$') when $\mathbf{r}$ and $\mathbf{r}$' vary within V. Clearly, a covariance function is symmetric with respect to ($\mathbf{r}$,$\mathbf{r}$'), and is characterized by the fact it is a positive definite function, *i.e.*:

$$\forall n \in \mathbb{N},\ \forall(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n) \in \mathrm{V}^n,\ \forall(v_1, v_2, ..., v_n) \in \mathbb{R}^n :\ \sum_{i,j=1}^{n} C(\mathbf{r}_i, \mathbf{r}_j)v_i v_j \geq 0 \tag{2.47}$$

It shows that if C is a covariance kernel, (i) the restriction of C to any subset of V is a covariance kernel, and (ii) for any function f defined over **V** C($\mathbf{r}$,$\mathbf{r}$')f($\mathbf{r}$)f($\mathbf{r}$') is also a covariance kernel. Consequently, it is sufficient to consider covariance kernel in $\mathbb{R}^q$ to deduce covariance kernel in any subset V $\in \mathbb{R}^q$, and we can only considered correlation kernel such that $Cor(\mathbf{r}, \mathbf{r}) = 1$ for any point $\mathbf{r}$, and for which:

$$C(\mathbf{r}, \mathbf{r}') = Cor(\mathbf{r}, \mathbf{r})\sigma(\mathbf{r})\sigma(\mathbf{r}')$$

is a covariance function for any function $\sigma(\mathbf{r})$ (here assumed positive), and then:

$$\sigma^2(\mathbf{r}) = C(\mathbf{r}, \mathbf{r})$$

is the variance of the corresponding random function at point $\mathbf{r}$. The Cauchy-Schwartz inequality applied to the definite positive matrix obtained by restricting $Cor(\cdot, \cdot)$ to the points $\mathbf{r}$ and $\mathbf{r}$' shows that $Cor(\mathbf{r}, \mathbf{r}')$ ranges in $[-1, 1]$.

From a practical point of view, it is important to characterise positive definite functions, or positive symmetric matrices. Simple criterions can be obtained only for particular classes of functions or matrices.

2.5.1.1 Continuous homogeneous correlation kernels and self-correlation operators

A correlation function is homogeneous (or stationnary) if it only depends on the difference in position, and can thus be written as :

$$Cor(\mathbf{r}, \mathbf{r}') = \phi(\mathbf{r} - \mathbf{r}')$$

where ϕ is an even (since Cor is symmetric) function defined over $\mathbb{R}^q$. The function ϕ, which verifies the property equivalent to (2.47), is also referred to as definite positive. We have the following characterisation (Valette 2011):

A (real) function ϕ defined over $\mathbb{R}^q$, which is continuous, even, and absolutely integrable ($\in \mathbb{L}^1(\mathbb{R}^q)$), is positive definite if, and only if, its Fourier transform is positive.

This result corresponds to a particular case of Bochner's theorem which characterises positive definitiveness for even, continuous functions in general. It allows an easy calculation of correlation functions in dimension $q = 1$. Simple examples of correlation functions are :

$$\phi(r) = e^{-|r|} \quad \mathcal{F}(\phi)(k) \propto \frac{2}{1+k^2} \qquad \phi(r) = e^{-r^2} \quad \mathcal{F}(\phi)(k) \propto \sqrt{\pi} e^{-k^2/4} \tag{2.48}$$

or

$$\phi(r) = \frac{1}{ch(r)} \quad \mathcal{F}(\phi)(k) \propto \frac{\pi}{ch(k\pi/2)} \tag{2.49}$$

when taking the following definition of the Fourier transform :

$$\mathcal{F}(f)(k) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ikx} f(x) dx \tag{2.50}$$

It is important to note that a function of positive type is not necessarely positive everywhere. For instance the sinc (cardinal sine) function is of positive type since its Fourier transform is a box type function, whereas a box type function is not. Given correlation function in 1D, we can construct correlation functions in upper dimension by tensorial product :

$$\phi(\mathbf{r}) = \prod_{i=1}^{q} \phi_i(r_i) \tag{2.51}$$

where the functions ϕ_i $(i = 1, ..., q)$ are one dimensional correlation functions. Given a stationary correlation function $Cor(\mathbf{r}, \mathbf{r}') = \phi(\mathbf{r} - \mathbf{r}')$ in $\mathbb{R}^q$, by applying the Fourier transform we can remark that the function

$$\phi\left(\frac{\mathbf{r} - \mathbf{r}'}{\xi}\right)$$

is also a correlation function. The positive scalar ξ is a length that gives the scale of the correlation zone around a point.

Finally, the operator of self-correlation **Cor** is defined by convolution in the case of a positive definite function $\phi \in \mathbb{L}^1(\mathbb{R}^q)$. It follows from standard results on convolution that this operator can be defined in the space of square integrable functions $\mathbb{L}^2(\mathbb{R}^q)$ as :

$$\forall f \in \mathbb{L}^2(\mathbb{R}^q), \quad f \longrightarrow \mathbf{Cor}(f)(\mathbf{r}) = \int_{\mathbb{R}^q} \phi(\mathbf{r} - \mathbf{r}') f(\mathbf{r}') \mathrm{d}V(\mathbf{r}') \in \mathbb{L}^2(\mathbb{R}^q) \tag{2.52}$$

and is bounded and self-adjoint positive.

2.5.1.2 Continuous isotropics correlation kernels

A particular case of stationnary correlation functions are the isotropic ones that are defined as :

$$Cor(\mathbf{r}, \mathbf{r}') = \phi(\mathbf{r} - \mathbf{r}') = \psi(\|\mathbf{r} - \mathbf{r}'\|)$$

where ψ is assumed to be even (by convention) over $\mathbb{R}$ and $\|\cdot\|$ denotes the usual euclidean norm in $\mathbb{R}^q$. Restricting oneself to the usual space (q=3) and taking account of the following result that holds in $\mathbb{R}^3$:

$$\mathcal{F}(\phi)(\mathbf{k}) = -\frac{2\pi}{k}\frac{\mathrm{d}}{\mathrm{d}k}(\psi)(k) \qquad \text{for } k = \|\mathbf{k}\| \neq 0 \tag{2.53}$$

the preeceding characterization of stationnary kernels takes the following formulation for isotropic ones in $\mathbb{R}^3$ (Gaspari and Cohn, 1999; Valette, 2011):

A (real) function $\phi(\mathbf{r}) = \psi(\|\mathbf{r}\|)$ *defined over* $\mathbb{R}^3$*, where* ψ *is continuous, even, and* $r^2\psi$ *is absolutely integrable over* $\mathbb{R}$ $(\in \mathbb{L}^1(\mathbb{R}))$*, is positive definite if, and only if, the Fourier transform of* ψ *is a decreasing fonction of the wave number k, for* $k \geq 0$*.*

We can remark that then ψ is surely absolutely integrable since it is continuous and $r^2\psi(r)$ is absolutely integrable. Thus $\mathcal{F}(\psi)$ is continuous and vanishes at ∞ by the Riemann-Lebesgue lemma. This shows that $\mathcal{F}(\psi)$ is also positive if ϕ is positive definite.

We can also remark that the one dimensional functions given above as examples may be used for isotropic correlation function in dimension 3, and thus also in dimension 2.

Actually, we can use such isotropic functions to account for anisotropic correlation. Indeed, considering a continuous even function ϕ from $\mathbb{R}^3$ into $[-1, 1]$ which is definite positive (and thus associated to a stationnary kernel), we define the function ϕ_ξ, where $\xi = (\xi_1, \xi_2, \xi_3)$ is a 3-tuple of positive numbers, by :

$$\phi_\xi(\mathbf{r}) = \phi\left(\frac{r_1}{\xi_1}, \frac{r_2}{\xi_2}, \frac{r_3}{\xi_3}\right) \tag{2.54}$$

By taking its Fourier transform in $\mathbb{R}^3$, and through a change of variable, we obtain:

$$\mathcal{F}(\phi_\xi)(\mathbf{k}) = \xi_1\xi_2\xi_3\mathcal{F}(\phi)(\mathbf{k}') \tag{2.55}$$

where:

$$\mathbf{k}' = (\xi_1 k_1, \xi_2 k_2, \xi_3 k_3)$$

This shows that the Fourier transform of ϕ_ξ is also everywhere positive, and consequently that the kernel:

$$Cor(\mathbf{r}, \mathbf{r}') = \phi\left(\frac{r_1 - r_1'}{\xi_1}, \frac{r_2 - r_2'}{\xi_2}, \frac{r_3 - r_3'}{\xi_3}\right)$$

is also a stationnary correlation kernel. In particular, for an isotropic kernel with $Cor(\mathbf{r}, \mathbf{r}') = \phi(\mathbf{r} - \mathbf{r}') = \psi(\|\mathbf{r} - \mathbf{r}'\|)$ we deduce that the kernel

$$Cor(\mathbf{r}, \mathbf{r}') = \psi\left[-\left(\sum_{i=1}^{3}\frac{(r_i - r_i')^2}{\xi_i^2}\right)^{1/2}\right] \tag{2.56}$$

is also a correlation one. The lengths ξ_1, ξ_2, ξ_3 now give the scale of correlation in each direction.

2.5.1.3 The exponential kernel in $\mathbb{R}^3$

The exponential kernel $\phi(\mathbf{r}) = exp(-\|\mathbf{r}\|)$ is of particular interest in $\mathbb{R}^3$ since we may explicitely compute the square root of the inverse correlation operator (Tarantola , Valette

2011). Applying the preeceding results, we see, by using successively (2.48), (2.53), and (2.55), that :

$$\mathcal{F}(\phi)(\mathbf{k}) = \frac{8\pi}{(1+\|\mathbf{k}\|^2)^2}$$

and

$$\mathcal{F}(\phi_\xi)(k) = \frac{8\pi\xi_1\xi_2\xi_3}{(1+\xi_1^2k_1^2+\xi_2^2k_2^2+\xi_3^2k_3^2)^2} \tag{2.57}$$

Considering now the correlation operator (2.52) associated to ϕ_ξ, defined as:

$$\forall f \in \mathrm{L}^2(\mathbb{R}^3)\ f \to g = \mathbf{Cor}(f) = \phi_\xi * f$$

where the star denotes the convolution, we get by taking the Fourier transform:

$$\mathcal{F}(g)(\mathbf{k}) = \mathcal{F}(\phi_\xi)(\mathbf{k})\mathcal{F}(f)(\mathbf{k})$$

Then, by using (2.57) we deduce:

$$\begin{aligned}\mathcal{F}(f(\mathbf{k})) &= \frac{1}{8\pi\xi_1\xi_2\xi_3}\left(1+\sum_{j=1}^{3}k_j^2\xi_j^2\right)^2\mathcal{F}(g)(k)\\ &= \frac{1}{8\pi\xi_1\xi_2\xi_3}\mathcal{F}\left((\mathbf{I}-\Delta_\xi)^2 g\right)(k)\end{aligned}$$

where Δ_ξ denotes the modified Laplacian defined as:

$$\Delta_\xi = \sum_{j=1}^{3}\xi_j^2\partial_j^2$$

and thus:

$$f = \frac{1}{8\pi\xi_1\xi_2\xi_3}\left(\mathbf{I}-\Delta_\xi\right)^2 g \tag{2.58}$$

This gives the expression of the inverse of the correlation operator as well as its square root :

$$\mathbf{Cor}^{-1} = \frac{1}{8\pi\xi_1\xi_2\xi_3}\left(I-\Delta_\xi\right)^2 \qquad \mathbf{Cor}^{-1/2} = \frac{1}{\sqrt{8\pi\xi_1\xi_2\xi_3}}\left(I-\Delta_\xi\right) \tag{2.59}$$

Finally, the covariance operator $\mathbf{C}$ associated to the covariance kernel $C(\mathbf{r},\mathbf{r}') = \sigma(\mathbf{r})\sigma(\mathbf{r}')Cor(\mathbf{r})\mathbf{r}')$ can be written as the following composition of operators :

$$\mathbf{C} = \mathbf{\Sigma}\,\mathbf{Cor}\,\mathbf{\Sigma} \tag{2.60}$$

where $\mathbf{\Sigma}$ is the multiplication operator by σ: $f \to \mathbf{\Sigma} f \,:\, (\mathbf{\Sigma} f)(\mathbf{x}) = \sigma(\mathbf{x}) f(\mathbf{x})$

It follows that :

$$\mathbf{C}^{-1} = \mathbf{\Sigma}^{-1}\mathbf{Cor}^{-1}\mathbf{\Sigma}^{-1} = \left(\mathbf{Cor}^{-1/2}\mathbf{\Sigma}^{-1}\right)^{*}\mathbf{Cor}^{-1/2}\mathbf{\Sigma}^{-1} \tag{2.61}$$

with:

$$\mathbf{Cor}^{-1/2}\mathbf{\Sigma}^{-1}(f) = \frac{1}{\sqrt{8\pi\xi_1\xi_2\xi_3}}\left(I - \Delta_\xi\right)\left(\frac{f}{\sigma}\right)$$

2.5.2 Regularization of tomography problems

Let us now consider an inverse problem where $\mathbb{D} = \mathbb{R}^n$ and where the model space consists of square integrable scalar functions defined in a domain $\mathrm{V} \in \mathbb{R}^3$ representing a physical quantity depending on the position. Let us assume that the theoretical mapping g is continuously differentiable in $\mathbb{L}^2(\mathrm{V})$, thus the derivative operateur $\mathbf{G}_m$ at model $\mathbf{m}$ can be written as:

$$\forall i \in \{1, ..., n\} \qquad \delta \mathrm{d}_i = (\mathbf{G}_m \delta\mathbf{m})_i = \int_V \mathrm{h}^i_m(\mathbf{r})\delta\mathrm{m}(\mathbf{r})\mathrm{d}V \tag{2.62}$$

where the n functions $\mathrm{h}^i_m \in \mathbb{L}^2(\mathrm{V})$ are the Fréchet kernels related to the data. From :

$$(\mathbf{G}_m\delta\mathbf{m}|\mathbf{d}) = \sum_{i=1}^{n} \mathrm{d}_i \int_V \mathrm{h}^i_m(\mathbf{r})\delta\mathrm{m}(\mathbf{r})\mathrm{d}V = (\delta\mathbf{m}|\sum_{i=1}^{n}\mathrm{d}_i\mathrm{h}^i_m)_{L^2}$$

we deduce the expression of the adjoint $\mathbf{G}^*$:

$$\mathbf{G}^*_m\mathbf{d} = \sum_{i=1}^{n}\mathrm{d}_i\mathrm{h}^i_m \tag{2.63}$$

We have seen that in the finite dimensional case the Bayesian approach leads to search for the mimimum of a cost function in the model space:

$$E(\mathbf{m}) = \frac{1}{2}\left(||\mathbf{C}_d^{-1/2}(\mathbf{d}^{obs} - g(\mathbf{m}))||^2_{\mathbb{D}} + ||\mathbf{C}_m^{-1/2}(\mathbf{m} - \mathbf{m}_{prior})||^2_{\mathbb{M}}\right) \tag{2.64}$$

that corresponds to the maximum likehood estimator. In the case when the model is infinite dimensional, this type of optimisation problem is called a Tikhonov problem.

The operator $\mathbf{C}_m^{-1/2}$ is the regularisation operator and $\mathcal{R}\mathbf{C}_m^{1/2}$ (the range, or image of $\mathbf{C}_m^{1/2}$) is the space of regularisation, *i.e.* the space where the problem may be set.

If we assume that the correlation kernel ϕ_ξ for the model space is in the form (2.48) where ϕ is a positive definite function in $\mathbb{R}^3$ and $\xi = (\xi_1, \xi_2, \xi_3)$ are the 3 correlation lengths, the covariance operator $\mathbf{C}_m$ defined by :

$$\forall f \in \mathbb{L}^2(V), \quad f \longrightarrow \mathbf{C}_m(f)(\mathbf{r}) = \sigma(\mathbf{r}) \int_{\mathbf{V}} \phi_\xi(\mathbf{r}-\mathbf{r}')\sigma(\mathbf{r}')(f(\mathbf{r}')\mathrm{d}V' \in \mathbb{L}^2(\mathbf{V}) \tag{2.65}$$

is bounded, self-adjoint, and positive definite. Its range, and that of its square root, consist of functions more or less smooth depending of the lengths of correlation; the greater are the correlation lengths the smoother are the functions of the regularization space.

As in the case of a finite dimensional model space, the stationarity equation of the cost function gradient ($\nabla_m E(\mathbf{m}) = 0$) takes the form :

$$\mathbf{m} - \mathbf{m}_{prior} = \mathbf{C}_m\mathbf{G}_m^*(\mathbf{C}_d + \mathbf{G}_m\mathbf{C}_m\mathbf{G}_m^*)^{-1}(\mathbf{d}^{obs} - g(\mathbf{m}) + \mathbf{G}_m(\mathbf{m} - \mathbf{m}_{prior})) \tag{2.66}$$

since we are in a Hibert space setting. Writing explicitly this equation, a model $\hat{\mathrm{m}}$ that minimizes the cost function verifies :

$$\hat{\mathrm{m}}(\mathbf{r}) - \mathrm{m}_{prior}(\mathbf{r}) = \sigma(\mathbf{r}) \sum_{i=1}^{n} v_i \int\limits_{\mathrm{V}} \phi_\xi(\mathbf{r}-\mathbf{r}')\,\sigma(\mathbf{r}')\mathrm{h}_{\hat{m}}^i(\mathbf{r}')\mathrm{d}V' \tag{2.67}$$

where the vector $\mathbf{v} = (\mathrm{v}_i)_{i=1,n}$ defined as :

$$\mathbf{v} = \mathbf{M}^{-1}(\mathbf{d}_{obs} - \mathbf{G}\mathbf{m}_{prior})$$

with :

$$\mathbf{M} = \mathbf{C}_d + \mathbf{G}_m\mathbf{C}_m\mathbf{G}_m^*$$

verifies :

$$\sum_{j=1}^{n} \mathrm{M}^{ij} v_j = \mathrm{d}_{obs}^i - g(\hat{\mathrm{m}}) + \int_{\mathrm{V}} \mathrm{h}_m^i(\mathbf{r})\,(\hat{\mathrm{m}} - \mathrm{m}_{prior})(\mathbf{r})\mathrm{d}V,$$

with

$$\mathrm{M}^{ij} = \mathbf{C}_d^{ij} + \int\limits_{\mathrm{V}\times\mathrm{V}} \sigma(\mathbf{r})\mathrm{h}_{\hat{m}}^i(\mathbf{r})\,\phi_\xi(\mathbf{r}-\mathbf{r}')\,\sigma(\mathbf{r}')\mathrm{h}_{\hat{m}}^j(\mathbf{r}')\mathrm{d}V\mathrm{d}V' \tag{2.68}$$

Furthermore, as in the case of finite dimensional model space, the equation (2.66) can be solved through a fixed point algorithm, the formal formulation of which is strictly the same as (2.26), and which can be explicitely formulated as (2.67), (2.68). This algorithm implies at each iteration the solution of a linear system, the matrix of which is analogous to (2.68), in the finite dimensional data space $\mathbb{D}$.

2.5.3 Smoothing, damping and a priori Information

The experience shows that if we assign reasonable values to the a priori variances and correlation lengths ξ_i, we obtain for most datasets excessively oscillating models. Consequently, in order to regularize the resulting models we must take greater ξ and/or smaller variances. It is thus important to clearly distinguish the true physical covariance $\mathbf{C}_{phys}$, related to the real level of prior information, from the smoothing covariance $\mathbf{C}_m$ used to calculate the model $\hat{\mathbf{m}}$. Once a type of correlation function has been chosen according to the width of the scale range we want to consider, we must take a reference length ξ_0 which corresponds more or less to the smallest scale we have to consider, or to the numerical mesh. It also corresponds to the scale for which it is reasonable to assign a value for the standard variation of the physical parameter. The physical a priori covariance is thus defined by the reference correlation length ξ_0 and the physical standard deviation field $\sigma_{phys}(\mathbf{r})$. We have now to specify $\mathbf{C}_m$, *i.e.* the correlation (smoothing) lengths ξ_i and the effective standard deviation $\sigma(\mathbf{r})$. These values cannot be taken independently of each other. A first point on this concern is to observe that by making $\xi = (\xi_i)_{i=1,3}$ tends toward 0 in the expression (2.59) of $\mathbf{C}_m$, we can deduce:

$$\mathbf{C}_m(f)(\mathbf{r}) \underset{\xi\to 0}{\sim} \xi_1\xi_2\xi_3\, \sigma^2(\mathbf{r})f(\mathbf{r}) \int_{\mathbb{R}^q} \phi(\mathbf{r})\mathrm{d}V$$

as can be seen by using the change of variable $y_i = (\mathrm{r'}\text{- r})_i/\xi_i\,(i = 1, ...q)$. It suggests that $\sigma^2(\mathbf{r})$ must be taken as proportional to the inverse of $\xi_1\xi_2\xi_3$ in the limit of small correlation lengths, to make converge $\mathbf{C}_m$ towards a white noise covariance. Furthermore, the inspection of equation (2.67) :

$$(\hat{\mathrm{m}} - \mathrm{m}_{prior})(\mathbf{r}) = \int_V \underbrace{\sigma(\mathbf{r})\phi_\xi(\mathbf{r}-\mathbf{r}')\,\sigma(\mathbf{r}')}_{regularization} \overbrace{\left\{\sum_{i=1}^{n} v_i\, \mathrm{h}^i_{\hat{m}}(\mathbf{r}')\right\}}^{data\ influence} \mathrm{d}V'$$

shows that for homogeneous inference of the data (constant h^i_m) and a constant prior standard deviation σ, the resulting improvement in the model at point $\mathbf{r}$:

$$(\hat{\mathrm{m}} - \mathrm{m}_{prior})(\mathbf{r}) \propto \int\limits_{\mathrm{V}} \phi_\xi(\mathbf{r}-\mathbf{r}')\mathrm{d}V$$

is proportional to the volume of the correlation zone around $\mathbf{r}$ from which the information is gathered at $\mathbf{r}$, and which is proportional to $\xi_1\xi_2\xi_3$. This shows that, all other factors being equal, the estimated improvement in the model is roughtly proportional to $\xi_1\xi_2\xi_3$. Consequently, in order to correct this artifact and to maintain the amplitude of the correction of the prior model indepently of the correlation lengths, we have to renormalize the standard deviation in function of the smoothing lengths by putting:

$$\sigma^2(\mathbf{r}) = \left(\frac{\xi_0^3}{\xi_1\xi_2\xi_3}\right)\sigma^2_{phys}(\mathbf{r}) \tag{2.69}$$

In this way, the effective standard deviation corresponds to the physical one when considering correlation lengths equal to ξ_0. From a numerical point of view, ξ_0 controls the damping of the regularization, *i.e.* the amplitude of the second term of the cost function $\mathrm{E}(\mathbf{m})$ (2.64) for a fixed value of the lengths of correlation; and these lengths control the smoothing of the resulting model for a fixed ξ_0.

2.5.3.1 Resolution Kernel and averaging index

The expression of the resolution operator is derived from (2.66) in the same way as for the finite dimensional case, and yields the same expression:

$$\hat{\mathbf{m}} - \mathbf{m}_{prior} = \mathbf{R}_{\hat{m}}(\mathbf{m} - \mathbf{m}_{prior}) - \mathbf{K}_{\hat{m}}(\mathbf{d} - \mathbf{d}_{obs}) + o(\|\hat{\mathbf{m}} - \mathbf{m}\|) \tag{2.70}$$

with:

$$\mathbf{R}_{\hat{m}} = \mathbf{K}_{\hat{m}}\mathbf{G}_{\hat{m}} \quad \text{and} \quad \mathbf{K}_{\hat{m}} = \mathbf{C}_m\mathbf{G}^*_{\hat{m}}(\mathbf{C}_d + \mathbf{G}_{\hat{m}}\mathbf{C}_m\mathbf{G}^*_{\hat{m}})^{-1}$$

Taking (2.62), (2.63) and (2.65) into account, the expression of $\mathbf{R}$ reads:

$$\begin{aligned}(\mathbf{R}_{\hat{m}}\delta\mathbf{m})(\mathbf{r}) &= \sigma(\mathbf{r})\sum_{i,j=1}^{n}\int_{\mathrm{V}\times\mathrm{V}}\phi_\xi(\mathbf{r}-\mathbf{r}'')\sigma(\mathbf{r}'')\,\mathrm{h}^i_{\hat{m}}(\mathbf{r}'')(\mathrm{M}^{-1})^{ij}\,\mathrm{h}^j_{\hat{m}}(\mathbf{r}')\,\delta\mathrm{m}(\mathbf{r}')\mathrm{d}V'\mathrm{d}V'' \\ &= \int_{\mathrm{V}\times\mathrm{V}} R(\mathbf{r},\mathbf{r}')\delta\mathrm{m}(\mathbf{r}')\mathrm{d}V'\end{aligned} \tag{2.71}$$

with:

$$R(\mathbf{r},\mathbf{r}') = \sigma(\mathbf{r})\sum_{i,j=1}^{n}(\mathrm{M}^{-1})^{ij}\mathrm{h}^j_{\hat{m}}(\mathbf{r}')\int_{\mathrm{V}}\phi_\xi(\mathbf{r}-\mathbf{r}'')\sigma(\mathbf{r}'')\,\mathrm{h}^i_{\hat{m}}(\mathbf{r}'')\mathrm{d}V'' \tag{2.72}$$

where the matix M is defined as (2.68):

$$\mathrm{M}^{ij} = \mathbf{C}_d^{ij} + \int\limits_{\mathrm{V}\times\mathrm{V}} \sigma(\mathbf{r})\mathrm{h}_{\hat{m}}^{i}(\mathbf{r})\, \phi_\xi(\mathbf{r}-\mathbf{r}')\, \sigma(\mathbf{r}')\mathrm{h}_{\hat{m}}^{j}(\mathbf{r}')\mathrm{d}V\mathrm{d}V'$$

The resolution kernel can be simply rewritten as:

$$R(\mathbf{r},\mathbf{r}') = \sigma(\mathbf{r}) \sum_{i=1}^{n} \mathrm{s}^i(\mathbf{r})\mathrm{h}_{\hat{m}}^{i}(\mathbf{r}') \tag{2.73}$$

by defining the vector fields $\mathbf{k}(\mathbf{r})$, $\mathbf{s}(\mathbf{r})$ respectively by:

$$\mathrm{k}^i(\mathbf{r}) = \int\limits_{\mathrm{V}} \phi_\xi(\mathbf{r}-\mathbf{r}')\, \sigma(\mathbf{r}')h_{\hat{m}}^{i}(\mathbf{r}')\mathrm{d}V' \tag{2.74}$$

and by :

$$\sum_{j=1}^{n} \mathrm{M}^{ij}\mathrm{s}^j(\mathbf{r}) = \mathrm{k}^i(\mathbf{r}) \tag{2.75}$$

With these notation, the equation (2.70) can also be explicitely reformulated as:

$$(\hat{\mathrm{m}} - \mathrm{m}_{prior})(\mathbf{r}) \simeq \sigma(\mathbf{r}) \sum_{i=1}^{n} \mathrm{s}^i(\mathbf{r}) \left\{ \int\limits_{\mathrm{V}} \mathrm{h}_{\hat{m}}^{i}(\mathbf{r}')(\mathrm{m} - \mathrm{m}_{prior})(\mathbf{r}'))\mathrm{d}V' - (\mathrm{d}^i - \mathrm{d}_{prior}^{i}) \right\} \tag{2.76}$$

Equations (2.70) and (2.71) show that the improvement in model, $\hat{\mathrm{m}} - \mathrm{m}_{prior}$, is equal to the true model correction, m - m_{prior}, filtered through the resolution kernel, and contaminated by a noise term coming from the discrepancy between true and a priori data. More precisely the value of $\hat{\mathrm{m}} - \mathrm{m}_{prior}$ at point $\mathbf{r}$ is an average of the values of the true model correction all around, weighted by the resolution kernel. Clearly, the resolution would be perfect if this kernel was equal to the Dirac distribution $\delta(\mathbf{r}-\mathbf{r}')$. Given a point $\mathbf{r}$, $R(\mathbf{r},\mathbf{r}')$ is generally a peaked function of $\mathbf{r}'$, centered near $\mathbf{r} = \mathbf{r}'$, with possibly few negative small lobes. In dimension 1 or 2, it can be very informative to display the resolution kernel at different positions within V.

A useful concept is the concept of averaging (or restitution) index (Rodgers, 2000; Vergely et al., 2010; Valette, 2011) which is defined as the total weight of the resolution kernel at a given point in V:

$$\mathrm{I}(\mathbf{r}) = \int\limits_{\mathrm{V}} \mathrm{R}(\mathbf{r},\mathbf{r}')\mathrm{d}V' = \sigma(\mathbf{r}) \sum_{i=1}^{n} \mathrm{s}^i(\mathbf{r}) \int\limits_{\mathrm{V}} \mathrm{h}^i(\mathbf{r}')\mathrm{d}V' \tag{2.77}$$

If the resolution at a point $\mathbf{r}$ is good, the resolution kernel is strongly centered around $\mathbf{r}$, and (2.71) yields:

$$(\hat{\mathrm{m}} - \mathrm{m}_{prior})(\mathbf{r}) \simeq (\overline{\mathrm{m} - \mathrm{m}_{prior}})\, \mathrm{I}(\mathbf{r})$$

where $(\overline{m - m_{prior}})$ denotes the (spatial) mean value of the true model correction around $\mathbf{r}$. In area in which the data provide poor information, $\mathrm{I}(\mathbf{r})$ takes very low values, and hence, the resulting model remains close to the prior one, independently of the value of the true model correction. In contrast, in area in which the index is close to one, the resulting model corresponds at each point to the mean value of the true model over the neighbourhood of the point. Thus, this index provides an efficient tool to delimit the zone outside which its value is lower than a given value, and consequently outside which the resolution is surely poor.

Chapter 3

Earthquake localisation

To obtain an acceptable tomography model, accurate *a priori* locations of seismic events are necessary. In this study, events locations are determined by a grid search method based on a stochastic approach, as described in chapter 2. To proceed, two problems have to be solved. The first one concerns the tracing of seismic rays and the computation of travel times using an heterogeneous media. The second one is the choice of an *a priori* velocity model over the studied region. Both of these points are discussed in the present chapter.

3.1 Numerical methods for travel times calculation

Most algorithms used to compute travel times of seismic waves through non homogeneous media are based on a finite differences solution of the seismic wave equation. A simplification of this problem consist of considering only plane front waves propagating in the medium. Rays joining seismic hypocenter and station are deduced from the perpendicular direction of these plane waves. This approximation is equivalent to considering an infinite frequency $\omega \to \infty$ for the seismic wave, which allows to despise scattering and diffraction phenomena (Nolet, 2008).

Following (MIT, 2008), lets consider the solution of the wave equation $\ddot{\phi} = \alpha^2 \nabla^2 \phi$ as:

$$\phi(\mathbf{x}, t) = A(\mathbf{x}) e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$

which can be written as a function of the wave travel time $T(\mathbf{x})$:

$$\phi(\mathbf{x}, t) = A(\mathbf{x}) e^{-i\omega T(\mathbf{x})}$$

In order to use this solution, we calculate:

$$\begin{aligned} \nabla \phi &= \nabla A e^{i\omega T} - i\omega A \nabla T e^{-i\omega T} \\ \nabla^2 \phi &= \left(\nabla^2 A - \omega^2 A |\nabla T|^2 - i(2\omega \nabla A \cdot \nabla T + \omega A \nabla^2 T) \right) e^{-i\omega T} \end{aligned}$$

$$\ddot{\phi} = -\omega^2 A e^{-i\omega T}$$

and the wave equation becomes:

$$\nabla^2 A - \omega^2 A |\nabla T|^2 - i(2\omega \nabla A \cdot \nabla T + \omega A \nabla^2 T) = -\frac{A\omega^2}{\alpha^2}$$

Dividing by ω^2:

$$\frac{\nabla^2 A}{\omega^2} - A|\nabla T|^2 - \frac{i}{\omega}(2\nabla A \cdot \nabla T + A\nabla^2 T) = -\frac{A}{\alpha^2}$$

By taking the limit for the infinity frequency $\omega \to \infty$, we obtain:

$$|\nabla T|^2 = \frac{1}{\alpha^2} \tag{3.1}$$

Commonly, this equation is parametrized by the slowness $s = 1/\alpha$ which is more practical than the speed. In order to clarify this introduction, we will develop the numerical method for the two dimensions scheme but as shown by Belayouni (2013), the problem is the same for the three dimensions case.

The two dimensional Eikonal equation can be written as:

$$\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2 = s(x,z)^2 \tag{3.2}$$

3.1.1 TIME3D software

In order to numerically solve this equation we use the TIME3D algorithm developed by Podvin and Lecomte (1991). This algorithm computes the travel times at each point of a regularly spaced grid, given a discrete velocity field. TIME3D is based on a Cartesian coordinate system. This algorithm makes use of three different methods to compute the propagation of the travel times.

3.1.1.1 2-D transmission

The first method of propagation is called the 2-D transmission. The travel time of a seismic wave at a grid knot is determined using the travel times previously determined at three neighbour knots as shown in figure 3.1.

On the figure 3.2, we present the method used by TIME3D to compute propagation times of a front wave in the grid element $HCDV$. In the case presented here, the front wave reaches the point D first, passes through the point H and then the point C. Angles α and β represented on figure 3.2 let us define two congruent triangles. The length of

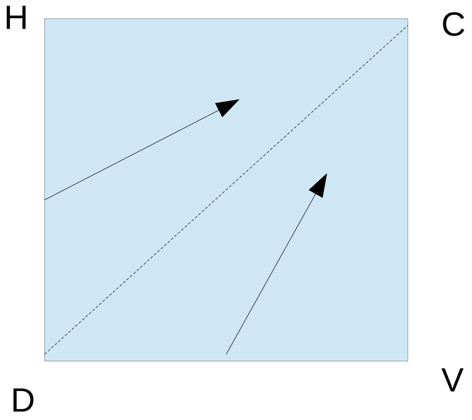

Figure 3.1: *Scheme of transmission in 2D for TIME3D. Time at point C is determined from times at points H, D and V. Two times are possibles.*

catheti can be expressed as functions of a constant slowness of the front wave s and the propagation times:

$$minor\ cathetus = \frac{T_H - T_D}{s}$$

$$major\ cathetus = \frac{T_C - T_H}{s}$$

The length of the hypotenuse is equal to the grid size length h:

$$h^2 = \left(\frac{T_C - T_H}{s}\right)^2 + \left(\frac{T_H - T_D}{s}\right)^2$$

Which leads to rewrite the equation under the following form:

$$\left(\frac{T_C - T_H}{s}\right)^2 + \left(\frac{T_H - T_D}{s}\right)^2 = s^2 \tag{3.3}$$

representing the discretized eikonal equation in the triangle CDH.

If the following condition, called illumination criterion, is stated:

$$0 \leq (T_H - T_D) \leq \frac{hs}{\sqrt{2}} \tag{3.4}$$

then the time T_C at point C is deduced from equation 3.3:

$$T_C = T_H + \sqrt{(hs)^2 - (T_H - T_D)^2} \tag{3.5}$$

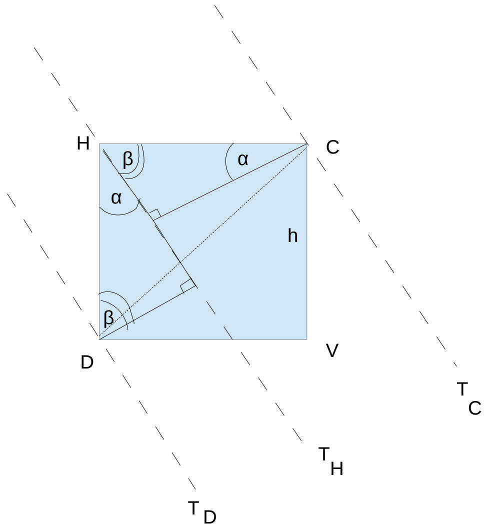

Figure 3.2: *Front plane wave transmission for the Podvin-Lecomte two dimensional scheme. h is the grid size length.* T_D*,* T_H *and* T_C *are the times when the front wave reaches the vertex* D*,* H *and* C*.*

Similarly, for the HVC triangle of the figure 3.1:

$$T_C = T_V + \sqrt{(hs)^2 - (T_V - T_D)^2} \tag{3.6}$$

With the illumination condition:

$$0 \leq (T_V - T_D) \leq \frac{hs}{\sqrt{2}} \tag{3.7}$$

3.1.1.2 1-D Transmission

The Podvin-Lecomte algorithm also introduces the possibility of a wavefront displacing along cells interface. This type of displacement corresponds to a refraction wave propagating at the interface between two different media. For the Podvin-Lecomte algorithm, the slowness is supposed constant in each cell as it is represented on figure 3.3.

The $1D$ transmission operator (figure 3.3) computes the time T_C at point C from the times T_V and T_H and the slowness (s, s') and (s, s'') respectively, as:

$$T_C = T_V + h \cdot min(s, s')$$

or:

$$T_C = T_H + h \cdot min(s, s'')$$

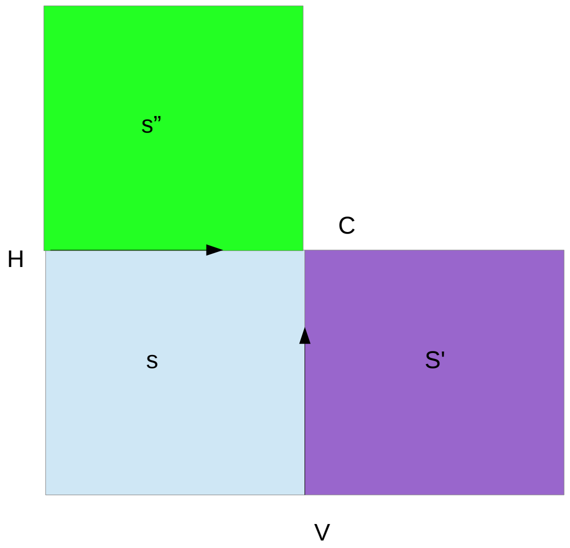

Figure 3.3: *1D transmission along interfaces for Podvin-Lecomte algorithm. Time at point C is calculated from time at point H or point V. Colored squares represent the cells with different slowness s, s' and s''. Arrows represent the front waves displacing at interfaces.*

3.1.1.3 2-D Diffraction

Finally, a last operator is used when illumination conditions of the $2-D$ transmission are not verified and $1-D$ transmission does not correspond the front wave. This operator is the $2-D$ diffraction represented on figure 3.4. This operator is introduced to treat the case when point D behaves like a secondary source from a wave reaching point C with slowness s.

$$T_C = T_D + \sqrt{2}hs$$

3.1.2 TIME3D implementation

The TIME3D algorithm propagate the times by the mean of concentric rings centred at source point, but the source is not necessary on a grid point. These rings gradually cover the entire domain, as represented on figure 3.5. Times are computed from the source to each point of the smaller ring, and then to the next ring and so on. As different times can be calculate at each point, the "correct" time is obtained through a minimum time criteria.

An important information which as to be taken into account at each step of calculation is the direction of propagation: operators can be considered in two directions in the $2D$ case (four in the $3D$ case). The algorithm proceed as follow to compute the time propagation:

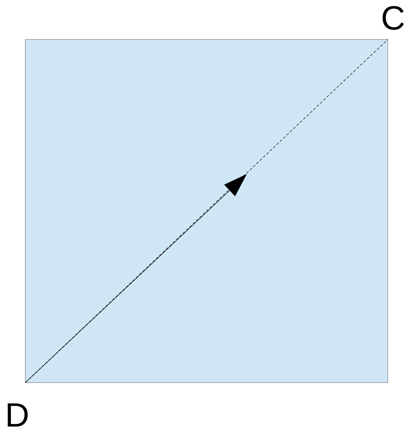

Figure 3.4: *Podvin-Lecomte 2D diffraction operator. Time at point C is calculated from the time at point D.*

1. Transmission $1D$ and $2D$ from inner to outer ring. Minimum times are retained.

2. Sweeping directions (right to left) for $1D$ and $2D$ and uses diffraction operators. Minimum time criterion again.

3. Same as step 2 but in the opposite direction (left to right). Time are actualized only if they are smaller than the precedent ones.

The Podvin-Lecomte code is written as a subroutine named TIME3D which is used by the localisation program LOCIN (Potin, 2016). Time precision computed by the Podvin-Lecomte algorithm can be improved after completion of this algorithm. These improved times are computed by numerical integration of the slowness field along the rays (Latorre et al., 2004). The rays are deduced from the gradient of the times obtained using Podvin-Lecomte algorithm. The Latorre et al. (2004) approach reduces the error on travelling time but is quite expensive in computation time. For this reason, this approach is used only in case of tomography computation to calculate travel time in the *a posteriori* model. For the localisation determination, the times are determined for a quasi-$1D$ uniform model so the accuracy of the Latorre et al. (2004) algorithm is unnecessary.

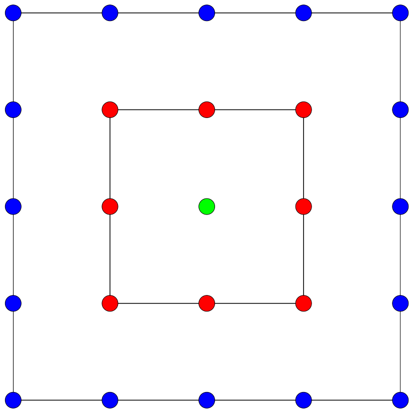

Figure 3.5: *Concentric rings to compute time propagation by the Podvin-Lecomte algorithm. First time defined is at the green point, then times at red points are computed from the green one, and then times at blue points are computed from the red ones, and so on.*

3.2 Seismicity localisation by probability density function

An approach to locate seismic events consists in determining a probability density function of its spatial location. In order to locate an event, one can proceed by "grid search" which consist in determining this function at each node of a grid covering a spatial domain. This approach replaces classical location methods which define an earthquake location as a point in space without a correct uncertainties estimation. A grid search is more consistent with modern physics idea that we need not only to know the position of an physical phenomenon as a point, but also a complete probability density function of this phenomenon in space. The computation we adopted for the a posteriori probability density function of seismic event location is perform by mean of the stochastic approach theory of the inverse problem we developed in Chapter 2.

3.2.1 The inverse problem of localisation

In the particular case of seismic localisation, we note $\mathbb{M} \equiv \mathbb{R}^4$ the model space and $\mathbf{m} = (\mathbf{x}, t_0)$ the vector of $\mathbb{M}$ containing both space parameters $\mathbf{x}(x, y, z)$ and time parameter t_0 of an event location. The data consist of a set of P-wave arrival times t_p and delay times $\Delta t = t_s - t_p$ observed at each station. The benefit of using these delay times instead of the S-wave arrival times t_s is double: delays do not depend on origin time t_0 and also eliminate an eventual time shift that might occur at the station. We note $\mathbf{d}$ the data vector defined in $\mathbb{D} \equiv \mathbb{R}^{m+n}$ and consisting of vectors $\mathbf{d}^{S-P} = (t_s - t_p)_{i=1,m} \in \mathbb{R}^n$ and $\mathbf{d} = (t_p)_{i=m+1,m+n} \in \mathbb{R}^n$. The ideal theoretical relationship g between data and model takes the form:

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_i^{S-P} \\ \mathbf{d}_i^{P} \end{bmatrix} = g(\mathbf{x}, t_0) = \begin{bmatrix} T_i^S(\mathbf{x}) - T_i^P(\mathbf{x}) \\ T_i^P(\mathbf{x}) + t_0 \end{bmatrix} \qquad \begin{matrix} i \leq i \leq m \\ m+1 \leq i \leq m+n \end{matrix} \tag{3.8}$$

and the conditional law of estimator $\mathbf{d}$ (cf. 2.10):

$$\begin{aligned} \rho(\mathbf{d}'|\mathbf{m}) =& \varphi(\mathbf{d} - g(\mathbf{x}, t_0)) \\ =& \varphi\big(\underbrace{..., \mathit{\Delta} t_i - T_i^S(\mathbf{x}) + T_i^P(\mathbf{x}), ...,}_{i \leq i \leq m} \underbrace{..., t_j^p - T_j^P(\mathbf{x}) - t_0 + t, ...}_{m+1 \leq j \leq m+n}\big) \end{aligned} \tag{3.9}$$

which leads, by a similar calculation of 2.16, to:

$$\rho_{post}(\mathbf{x}) \propto \int_{\mathbb{R}} \varphi\left(..., \mathit{\Delta} t_i - T_i^S(\mathbf{x}), ..., ..., t_j^P - T_j^P(\mathbf{x}) - t_0 + t, ...\right) \mathrm{d}t \tag{3.10}$$

where the variable t occurs only in the n last entries of function φ. Moreover, making the variable change $t = t_0 + t$ leads to:

$$\rho_{post}(\mathbf{x}) \propto \int_{\mathbb{R}} \varphi\left(..., \mathit{\Delta} t_i - T_i^S(\mathbf{x}) + T_i^P(\mathbf{x}), ..., ..., t_j^P - T_j^P(\mathbf{x}) - t, ...\right) \mathrm{d}t \tag{3.11}$$

which is the marginal density in $\mathbf{x}$ of the a posteriori law of $(\mathbf{x}, t_0)$ given the physical measured data. Considering data differences provides a mean to eliminate the origin time t_0 which is generally less important than the spatial location. This approach is equivalent to considering an a posteriori marginal density in $\mathbf{x}$, as long as the probabilities are properly transported. However, it should be noted that in some cases, it is possible to directly measure differences between arrival times of same wave front at two nearby stations by intercorrelation. In this case, the data are differences for which the accuracy is better than in our case. In the case of intercorrelation, the approach developed above does not apply.

The probability density function 2.10 has to be chosen in order to localize seismic events. It can either be chosen through equation 2.2 for which we have to distinguish the measurement uncertainties and the uncertainties resulting from the theoretical modeling, or fixed immediately. The first approach is easily applicable only with Gaussian laws. Therefore, it is most suitable for localisation in a $3-D$ model of v_P and v_P/v_S fields, produced by a tomographic study. In this particular case, a Gaussian modelling of errors is reasonable as we consider a relatively accurate model. At this point, we only want to determine the a priori location of the seismic events in order to initiate a tomographic study. As we do not yet have an accurate model, we compute these locations using a smooth $1-D$ velocity model. Therefore, it is more reasonable to directly fix φ, firstly because the theoretical uncertainties might be large and secondly because a Gaussian modelling of these errors is not suitable.

3.2.2 Localisation in an a priori velocity model

In the case of the a priori localisation, we adopted a conditional density on $\mathbf{d}$ of the form:

$$
\begin{aligned}
\rho(\mathbf{d}|\mathbf{x}, t_0) = & \prod_{i=1}^{m} \frac{1}{2\sigma(\mathbf{x})} exp\left[-\frac{|\Delta t_i - T_i^S(\mathbf{x}) + T_i^P(\mathbf{x})|}{\sigma_i(\mathbf{x})} \right] \\
& \times \prod_{i=m+1}^{m+n} \frac{1}{2\sigma(\mathbf{x})} exp\left[-\frac{|\Delta t_i - T_i^S(\mathbf{x}) - t_0|}{\sigma_i(\mathbf{x})} \right]
\end{aligned}
\tag{3.12}
$$

where $\sigma_i^2(\mathbf{x}) = \sigma_{i\,obs}^2 + \sigma_{i\,th}^2$, $\sigma_{i\,obs}^2$ representing the observational uncertainty and $\sigma_{i\,th}^2$ the uncertainty on travel times $T_i(\mathbf{x})$ which is due to the approximation of the $1-D$ velocity model in comparison with the true velocity of the waves.

Here, we assume the data independent from each other. For observational data, this is justified as both picks and uncertainties only depends on the signal observed at the station considered. However, in the case of theoretical errors, it is only true at first order since the uncertainties are correlated when the seismic stations are close from each other in comparison to the event location. In this particular configuration, seismic rays between this earthquake and the two stations are very similar, implying the theoretical uncertainties due to velocity model approximation being roughly identical. Such a configuration leads to a strongly correlated theoretical error in comparison with data corresponding to distant stations. However, the approach in which we consider independent errors seems sufficient in the case of the a priori localisation, for which the velocity model is quite simple. Moreover, the exponential form of the law is sufficient to overcome this approximation and is also well suitable to deal with possible outlier data. We now have to determine the shape of $\sigma_{th}(\mathbf{x})$.

Let us suppose that the uncertainty on the slowness $n(\mathbf{x}) = 1/v(\mathbf{x})$ can be modelled by a stationary Gaussian random function, with a null mean and a covariance kernel of the form:

$$
C(\mathbf{x}, \mathbf{x}') = \sigma_n^2 f\left(\frac{||\mathbf{x} - \mathbf{x}'||}{\xi} \right) \tag{3.13}
$$

where ξ is the characteristic length of the autocorrelation function f. The variance of the travel time $T(\mathbf{x})$ between a source $\mathbf{x}$ and the station corresponding to the data we consider is:

$$
\begin{aligned}
\sigma_{th}^2 = & \int_0^L \int_0^L C(\mathbf{x}(s), \mathbf{x}(s'))\, \mathrm{d}s\, \mathrm{d}s' \\
= & \sigma_n^2 \int_0^L \int_0^L f\left(\frac{||\mathbf{x} - \mathbf{x}'||}{\xi} \right) \mathrm{d}s\, \mathrm{d}s'
\end{aligned}
\tag{3.14}
$$

where L is the ray length.

Let us analyse the asymptotic behavior of this expression as a fonction of the length ratio ξ/L. for the large values of ξ/L, the correlation function remains close to 1 when

$\mathbf{x}(s)$ and $\mathbf{x}(s')$ describe the ray, so σ_{th}^2 is of the order of $\sigma_n^2 L^2$. Now, for small values of ξ/L, if we write:

$$\sigma_{th}^2 = \sigma_n^2 \int_0^L ds \int_0^L f\left(\frac{||\mathbf{x}(s) - \mathbf{x}(s')||}{\xi}\right) \mathrm{d}s' \tag{3.15}$$

and if we make the variable change $t = \frac{||\mathbf{x}(s)-\mathbf{x}(s')||}{\xi}$, we obtain, regarding:

$$\frac{dt}{ds'} = \frac{1}{\xi}\frac{(\mathbf{x}(s') - \mathbf{x}(s)) \cdot \tau(s)}{||\mathbf{x}(s') - \mathbf{x}(s)||} = \frac{1}{\xi} cos(\mathbf{x}(s') - \mathbf{x}(s), \tau(s)) \tag{3.16}$$

where $\tau(s)$ is the unit tangent vector to the ray at point $\mathbf{x}(s)$, by replacing in 3.15:

$$\sigma_{th}^2 = \sigma_n^2 \xi \int_0^L \mathrm{d}s \int_{(x(0)-x(s))/\xi}^{(x(1)-x(s))/\xi} f(t) \frac{\mathrm{d}t}{cos(\mathbf{x}(s') - \mathbf{x}(s), \tau(s))} \tag{3.17}$$

and $||\mathbf{x}(s') - \mathbf{x}(s)|| = \xi t$. Then, for the limit $\xi/L \to 0$:

$$\sigma_{th}^2 \underset{\xi \to 0}{\sim} \sigma_n^2 \xi L \int_{-\infty}^{+\infty} f(t) \mathrm{d}t \tag{3.18}$$

In conclusion, $\sigma_{th} \propto L$ for the small values of L and $\sigma_{th} \propto \sqrt{L}$ for the large values of L. This implies that σ_{th} is proportional to T or $\sqrt{T}$ depending on the value of T.

Regarding this result, we adopt for σ_{th} the following function:

$$\sigma_{th}(T) = \begin{cases} kT & T \leq T_c \\ k(2\sqrt{TT_c} - T_c) & T > T_c \end{cases} \tag{3.19}$$

where k and T_C are control parameters. In the case of a first localisation study in the Western Alps region (Potin, 2016), these parameters were set as $k = 0.05$ and $T_c = 6$ s. The choice of these values was based on the geographic configuration of earthquakes and station, data quality analysis, and series of test samples. Our study area is geographically more extended than the one of Potin (2016) resulting on average in larger travel times. Based on this information and some test we performed, we fixed these values to $k = 0.04$ and $T_c = 8$ s, which correspond to a relative error of 4 % for travel times lower than 8 s. We represented the corresponding function on figure 3.6.

Regarding the differences $t_S - t_P$, the theoretical error on the $\Delta T(\mathbf{x}) = T_S(\mathbf{x}) - T_P(\mathbf{x})$ data depends not only on the v_P velocity model but also on the v_P/v_S model. In the case of the a priori location, this error must be estimated based on a model where v_P/v_s is constant, with:

$$\Delta T(\mathbf{x}) = \int_0^L \left(\frac{v_p}{v_s}(\mathbf{x}(s)) - 1\right) \frac{ds}{v_p(\mathbf{x}(s))} \tag{3.20}$$

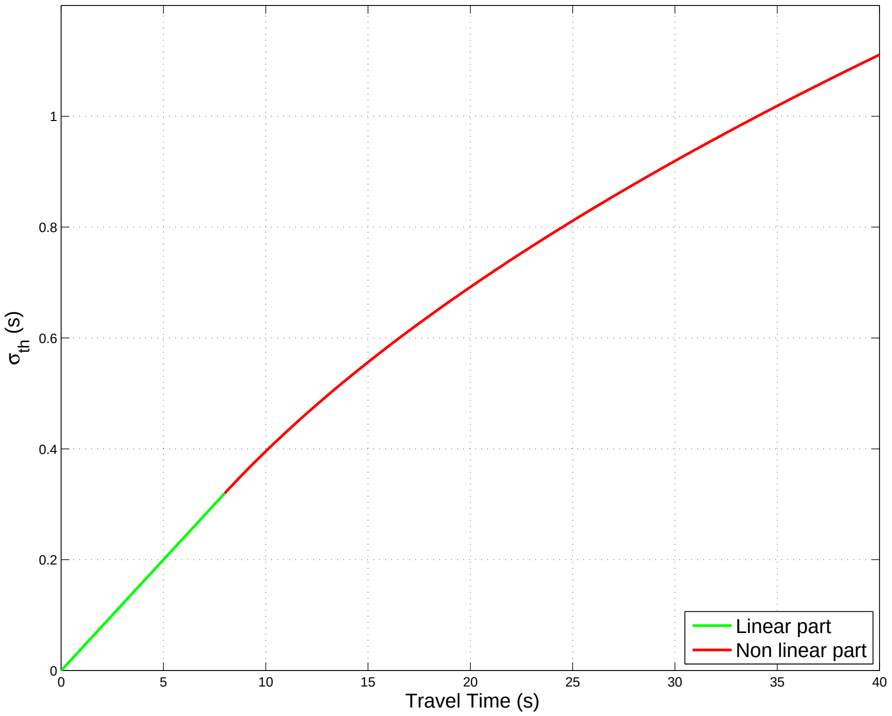

Figure 3.6: *σ_{th} as function of travel time T.*

Let us suppose that the error on $v_P/v_S(\mathbf{x})$ is modelled by a random, stationary and Gaussian function, with a null mean and a covariance $\sigma^2_{v_P/v_S} f\left(\frac{||\mathbf{x}-\mathbf{x}'||}{\xi}\right)$. Let us suppose as we that this function has the same correlation kernel $f\left(\frac{||\mathbf{x}-\mathbf{x}'||}{\xi}\right)$ that of $n_p = 1/v_P(\mathbf{x})$. We can deduce the variance on $\Delta T(\mathbf{x})$ is:

$$\sigma^2_{\Delta T} = \int_0^L\int_0^L \left(\sigma^2_{n_P}\left(\frac{v_P}{v_S}-1\right)^2 + \frac{\sigma^2_{v_P/v_S}}{v_P(\mathbf{x}(s))v_P(\mathbf{x}(s'))}\right) f\left(\frac{||\mathbf{x}(s)-\mathbf{x}(s')||}{\xi}\right) \mathrm{d}s\,\mathrm{d}s' \quad (3.21)$$

leading to:

$$\sigma^2_{\Delta T} \sim \left(\left(\frac{v_P}{v_S}-1\right)^2 + \frac{\sigma^2_{v_P/v_S}}{\sigma^2_{n_P}}\overline{n_P}^2\right)\sigma^2_{T_P} \quad (3.22)$$

where $\overline{n_P}$ is a mean value of n_P in our smooth model and σ_{T_P} is the theoretical error of the travel time of the P-wave. For $v_P/v_S \sim 1.7$ and $\sigma_{v_P/v_S} \sim \frac{\sigma_{n_P}}{n_P}$, we obtain:

$$\sigma^2_{\Delta T} \sim \frac{3}{2}\sigma^2_{T_P} \quad (3.23)$$

In order to be conservative, we adopted for $\sigma_{\Delta T}$ the value corresponding to σ_{T_S}, which is of the order of $\sqrt{3}\sigma_{T_P}$. This value is slightly greater than the estimate obtained above but it seems reasonable since S-wave data are generally of poor quality.

Now, let us come back to the computation of the a posteriori density on $\mathbf{x}$. From equation 3.12, we can deduce:

$$\begin{aligned}\rho_{post}(\mathbf{x}) \propto \left(\prod_{k=1}^{m+n} \frac{1}{\sigma_k(\mathbf{x})}\right) exp\left[-\sum_{j=1}^{m} \frac{|\Delta t_j^{obs} - T_j^S(\mathbf{x}) + T_j^P(\mathbf{x})|}{\sigma_j(\mathbf{x})}\right] \\ \times \int_{\mathbb{R}} exp\left[-\sum_{i=m+1}^{m+n} \frac{|t_i^{obs} - T_i^P(\mathbf{x}) - t|}{\sigma_i(\mathbf{x})}\right] dt\end{aligned} \tag{3.24}$$

The last factor of this expression which contains the integral can be analytically calculated. By taking $t_i = T_i^P(\mathbf{x}) - t_i^{obs}$, this factor becomes:

$$\int_{\mathbb{R}} exp\left[-\sum_{i=m+1}^{m+n} \frac{|t-t_i|}{\sigma_i}\right] \mathrm{d}t \tag{3.25}$$

Let us suppose that the t_i are in an increasing order: $t_{m+1} \leq t_{m+2} \leq ... \leq t_{m+n}$. Position of t compared to t_i allows to explicit the absolute values $|t - t_i|$, leading to:

- For $t \leq t_{m+1}$:

$$\frac{\prod \frac{1}{\sigma_i}}{\sum \frac{1}{\sigma_i}} exp\left(\sum_{i=m+n}^{m+n} \frac{t_{m+1} - t_i}{\sigma_i}\right) \tag{3.26}$$

- For $t_{m+j-1} \leq t \leq t_{m+j}$, with $j = 2, ..., n$:
  - If $\sum_{i=m+j}^{m+n} \frac{1}{\sigma_i} - \sum_{i=m+1}^{m+j} \frac{1}{\sigma_i} \neq 0$:

$$\frac{\prod \frac{1}{\sigma_i}}{\sum_{i=m+j}^{m+n} \frac{1}{\sigma_i} - \sum_{i=m+1}^{m+j} \frac{1}{\sigma_i}} exp\left(-\sum_{i=m+1}^{m+j-1} \frac{t-t_i}{\sigma_i} + \sum_{i=m+j}^{m+n} \frac{t-t_i}{\sigma_i}\right) \tag{3.27}$$

  - If $\sum_{i=m+j}^{m+n} \frac{1}{\sigma_i} - \sum_{i=m+1}^{m+j} \frac{1}{\sigma_i} = 0$:

$$(t_{m+j} - t_{m+j-1}) \prod \frac{1}{\sigma_i} exp\left(-\sum_{i=m+1}^{m+j-1} \frac{t_i}{\sigma_i} + \sum_{i=m+j}^{m+n} \frac{t_i}{\sigma_i}\right) \tag{3.28}$$

- For $t \geq t_{m+n}$:

$$\frac{\prod \frac{1}{\sigma_i}}{\sum \frac{1}{\sigma_i}} exp\left(\sum_{i=m+n}^{m+n-1} \frac{t_i - t_{m+1}}{\sigma_i}\right) \tag{3.29}$$

3.3 Definition of the computation grid

In order to solve the Eikonal equation and determine the travel times needed to locate a seismic event, we have seen in section 1 that we chose to use the TIME3D algorithm (Podvin and Lecomte, 1991). This algorithm uses a Cartesian grid to propagate the travel time. Each cell of this grid is a cube in which we assume a constant velocity. The computational grid is a parallelepiped box constituted by these cells and if the cells are small enough, the velocity model can be relatively smooth. As the area of study is several hundred kilometres wild, the true shape of the Earth can not be ignored. Therefore, we need to project the Earth shape on this grid and also to identify every node by both its local Cartesian coordinates and its geographical coordinates.

The shape of the Earth is generally modelled as a revolution ellipsoid, which is close to the shape of the Earth geoid. In this study, we chose to base our work on the GRS80 ellipsoid. On the figure 3.7, we represented schematically a cross-section of the Earth on which the computational grid is drawn. As we can see, our approach preserves the true shape of the Earth as we do not introduce any geometrical deformation. The box is centred on a point M which is located on the ellipsoid. M geographical coordinates are (ϕ, λ, h). It is important to note that the geographical latitude λ depends on the local vertical at point M. This latitude differs from the geocentric latitude θ which is the angle between the equatorial plane and the line between the center of the ellipsoid and M (see figure 3.8).

The Cartesian computational grid is a parallelepiped tangent the the ellipsoid at point M. Local Cartesian system is centred on M and its three dimensions are oriented towards the East (x), the North (y) and to the depth (z).

3.3.1 About the reference revolution ellipsoid

Reference ellipsoid GRS80 is a revolution ellipsoid around the North-South axis of the Earth. Two semiaxes define this ellipsoid[1]: the Equatorial semiaxe $a = 6\,378,137$ km and the Polar semiaxe $b = 6\,356,752\,314\,14$ km.

The expression of the ellipsoid equation using an orthonormal generic coordinate system takes the form:

$$\frac{X^2}{a^2} + \frac{Y^2}{a^2} + \frac{Z^2}{b^2} = 1 \tag{3.30}$$

The generic coordinates of the point M, located on the ellipsoid, are (X_M, Y_M, Z_M) and its geocentric coordinates are $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and $\phi \in [\pi, \pi]$. The difference between geocentric latitude θ and geographic latitude λ is clearly represented on the figure 3.8.

At point M, two different curvature radius exist: one following meridians direction and one following parallels direction. Let us express these two radius. The equation of

---

[1] The a constant is exact as it is part of the definition of the GRS80 ellipsoid. The b constant derives from a and other definition constants, so its precision is rounded to twelve digits

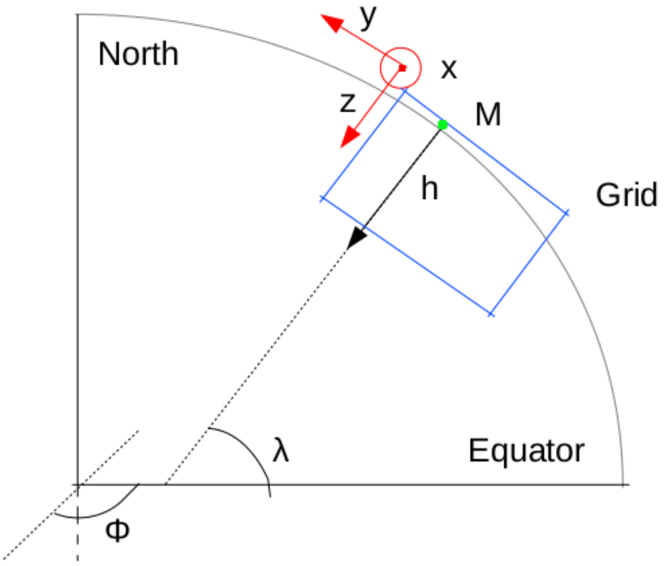

Figure 3.7: *Cross-section of the earth through the reference point* M*.* (ϕ, λ, h) *represent the geographic coordinate system and* (x, y, z) *the Cartesian coordinate system, which is centred on* M *and is used to define the computational grid, represented by the blue box.*

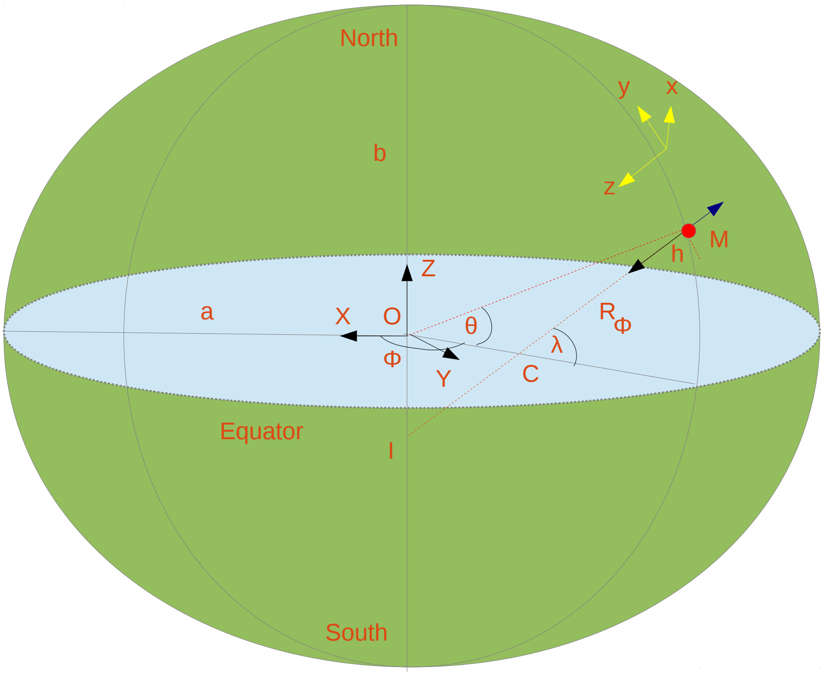

Figure 3.8: *The ellipsoid The GRS80 reference ellipsoid. Magnitudes used in transformation from geographic* (ϕ, λ, h) *to grid computing system* (x, y, z) *are introduced.*

the ellipse corresponding to the meridian plan $\phi = 0$ takes the form:

$$\frac{X^2}{a^2} + \frac{Z^2}{b^2} = 1 \tag{3.31}$$

If we search the intersection of this ellipsoid with the straight line $Z = X\,tan\theta$, we find the coordinates of the generic point of this ellipse as a function of θ:

$$\overrightarrow{OM_0} = \frac{ab}{\sqrt{b^2 + a^2 tan^2\theta}} \begin{cases} 1 \\ 0 \\ tan\theta \end{cases} \tag{3.32}$$

and the components of the normal $\overrightarrow{n_0}$ to this ellipse at point M_0 is:

$$\overrightarrow{n_0} = \frac{1}{\sqrt{b^4 + a^4 tan^2\theta}} \begin{cases} b^2 \\ 0 \\ a^2 tan\theta \end{cases} \tag{3.33}$$

A rotation of this ellipse by an angle ϕ around the Polar axis let us deduce the coordinates of point $M(\theta, \phi)$ and the components of the corresponding normal $\overrightarrow{n}$:

$$\overrightarrow{OM} = \frac{ab}{\sqrt{b^2 + a^2 tan^2\theta}} \begin{cases} cos\phi \\ sin\phi \\ tan\theta \end{cases} \qquad \overrightarrow{n} = \frac{1}{\sqrt{b^4 + a^4 tan^2\theta}} \begin{cases} b^2 cos\phi \\ b^2 sin\phi \\ a^2 tan\theta \end{cases} \tag{3.34}$$

The principal curvatures of a surface are the eigenvalues of Weingarten symmetric operator, defined as the derivative operator following the tangent plane Σ of the normal $\overrightarrow{n}$. Base vectors of this tangent plane to the ellipsoid are $\frac{\partial M}{\partial \phi}$ and $\frac{\partial M}{\partial \theta}$, and by definition:

$$D\overrightarrow{n}\left(\frac{\partial M}{\partial \phi}\right) = \frac{\partial \vec{n}}{\partial \phi} \qquad and \qquad D\vec{n}\left(\frac{\partial M}{\partial \theta}\right) = \frac{\partial \vec{n}}{\partial \theta} \tag{3.35}$$

From expressions 3.34, we deduce:

$$\frac{\partial \vec{n}}{\partial \phi} = \frac{b}{a}\sqrt{\frac{b^2 + a^2 tan^2\theta}{b^4 + a^4 tan^2\theta}}\frac{\partial M}{\partial \phi} \tag{3.36}$$

and:

$$\frac{\partial M}{\partial \theta} = \frac{ab}{(b^2 + a^2 tan^2\theta)^{3/2}} \begin{cases} -a^2 tan\theta cos\phi \\ -a^2 tan\theta sin\phi \\ b^2 \end{cases} \qquad \frac{\partial \vec{n}}{\partial \theta} = \frac{a^2 b^2}{(b^4 + a^4 tan^2\theta)^{3/2}} \begin{cases} -a^2 tan^2\theta cos\phi \\ -a^2 tan^2\theta sin\phi \\ b^2 \end{cases} \tag{3.37}$$

Therefore:

$$\frac{\partial \vec{n}}{\partial \theta} = ab\left(\frac{b^2 + a^2 tan^2\theta}{b^4 + a^4 tan^2\theta}\right)^{3/2} \frac{\partial M}{\partial \theta} \tag{3.38}$$

We deduce from this development that the vectors $\frac{\partial M}{\partial \phi}$ and $\frac{\partial M}{\partial \theta}$ are the eigenvectors of the Weingarten operator and the radius of curvature in the meridian direction is:

$$R_\theta = \frac{1}{ab}\left(\frac{b^4 + a^4 tan^2\theta}{b^2 + a^2 tan^2\theta}\right)^{3/2} \tag{3.39}$$

and the one in the parallel direction:

$$R_\phi = \frac{a}{b}\sqrt{\frac{b^4 + a^4 tan^2\theta}{b^2 + a^2 tan^2\theta}} \tag{3.40}$$

3.3.2 Radius of curvature

For the calculations, it is natural to use the geocentric latitude as it is part of the spherical coordinate system. However, coordinates on the Earth are generally given using the geographical latitude, which differs from the geocentric latitude as we shown earlier (see figure 3.8). Taking advantage of the symmetry of revolution, let us come back to the ellipse of meridian plane $\phi = 0$ ($Y = 0$). Considering the λ correspond to the slope of the normal $\overrightarrow{n}$ at point M, we deduce intermediately from 3.33 that:

$$tan\lambda = \frac{a^2}{b^2} tan\theta$$

and therefore:

$$cos\lambda = \frac{b^2}{\sqrt{b^4 + a^4 tan^4\theta}} \tag{3.41}$$

The components of $\overrightarrow{OM_0}$ (equation 3.32) can then be written as:

$$\overrightarrow{OM_0} = R_\phi \begin{cases} cos\lambda \\ 0 \\ \frac{a^2}{b^2} sin\lambda \end{cases} \tag{3.42}$$

from which we deduce the components of $\overrightarrow{OM}$ by rotation around the OZ axis of an angle $\phi \in [-\pi, \pi]$:

$$\overrightarrow{OM} = R_\phi \begin{cases} cos\lambda cos\phi \\ cos\lambda sin\phi \\ \frac{a^2}{b^2} sin\lambda \end{cases} \tag{3.43}$$

Moreover, in the meridian plan $\phi = 0$ and its ellipse, we see that:

$$\|CM_0\| = R_\phi \frac{b^2}{a^2}$$

and:

$$\|IM_0\| = R_\phi$$

from which we deduce:

$$\begin{aligned}\|IC\| =& \frac{a^2-b^2}{a^2}R_\phi \\ \|OC\| =& R_\phi\frac{a^2-b^2}{a^2}cos\lambda = \frac{b^2}{a^2}\frac{a^2-b^2}{\sqrt{b^4+a^4tan^2\theta}}\end{aligned}$$

3.3.3 Conversion between geographic and local coordinates

3.3.3.1 From geographic to Cartesian coordinates

Passing the geographic coordinates of a point $m(\phi_m, \lambda_m, h_m)$ to the local Cartesian coordinates system $M(x, y, z)$ is achieved through the use of a Cartesian geocentric system (O, X, Y, Z) (see figure 3.8). As it is more common to manipulate geographic latitude λ than geocentric latitude θ, we write the curvature radius R_ϕ from the equation 3.40:

$$R_\phi = \frac{a}{cos\lambda\sqrt{1+(b^2/a^2)\,tan^2\lambda}} \tag{3.44}$$

The coordinates of the origin of the Cartesian system M, located on the surface of the ellipsoid, and expressed in the geocentric system are:

$$M\begin{cases}\frac{a}{\sqrt{1+(b^2/a^2)tan^2\lambda_M}}cos\phi_M \\ \frac{a}{\sqrt{1+(b^2/a^2)tan^2\lambda_M}}sin\phi_M \\ \frac{b^2/a}{\sqrt{1+(b^2/a^2)tan^2\lambda_M}}tan\lambda_M\end{cases} \tag{3.45}$$

and the point in space m, defined by its geographic coordinates (ϕ_m, λ_m, h_m):

$$m\begin{cases}(R_\phi + h_m)cos\phi_m cos\lambda_m \\ (R_\phi + h_m)sin\phi_m cos\lambda_m \\ (R_\phi(b^2/a^2) + h_m)sin\lambda_m\end{cases} \tag{3.46}$$

From the coordinate differences $X' = X_M - X_m$, $Y' = Y_M - Y_m$ and $Z' = Z_M - Z_m$, we define the coordinates of $m(x_m, y_m, z_m)$ in the local Cartesian system (M, x, y, z):

$$m\begin{cases}-X'sin\phi_M + Y'cos\phi_M \\ -X'sin\lambda_M cos\phi_M - Y'sin\lambda_M sin\phi_M + Z'cos\lambda_M \\ X'cos\lambda_M cos\phi_M + Y'cos\lambda_M sin\phi_M + Z'sin\lambda_M\end{cases} \tag{3.47}$$

3.3.3.2 From Cartesian to geographic coordinates

The inverse operation, consisting in the conversion of coordinates in the local Cartesian system (M, x, y, z) to the geographic coordinates is a bit more complex. From the coordinates of point m (equation 3.47) in the local Cartesian system (M, x, y, z), we deduce its coordinates in the geocentric system (O, X, Y, Z):

$$m \begin{cases} X_M - x_m sin\phi_M - y_m sin\lambda_M cos\phi_M + z_M cos\lambda_M cos\phi_M \\ Y_M + x_m cos\phi_M - y_m sin\lambda_M sin\phi_M + z_m \lambda_M sin\phi_M \\ Z_M + y_m cos\lambda_M + z_m sin\lambda_M \end{cases} \tag{3.48}$$

We obtain the longitude ϕ_m from the components the X_m and Y_m with a subtlety:

$$\phi_m \begin{cases} arcos\frac{X_m}{\sqrt{X_m^2+Y_m^2}} & Y_m \geq 0 \\ -arcos\frac{X_m}{\sqrt{X_m^2+Y_m^2}} & Y_m < 0 \end{cases} \tag{3.49}$$

The geographic latitude λ_m is more delicate to determine because we do not know the local normal n at the ellipsoid surface directly at the geographical vertical of m. To determine λ_m, we make use of an iterative process which converge to its value. As the ratio $\frac{a}{b}$ is very near to 1 for the Earth, we take as first approximation:

$$\lambda_m \simeq \theta_m = \frac{Z_m}{\sqrt{X_m^2 + Y_m^2}} \tag{3.50}$$

To determine λ_m value, we define $l_{k=0} = \theta_m$ and we recursively calculate l_{k+1} as function of l_k, from $k = 0$:

$$l_{k+1} = \frac{Z_m}{\sqrt{X_m^2 + Y_m^2}} \left[1 - \frac{ae^2}{\sqrt{X_m^2 + Y_m^2}\sqrt{1 + (b^2/a^2) l_k^2}} \right] \tag{3.51}$$

with e the ellipsoid eccentricity $\frac{\sqrt{a^2+b^2}}{a}$. The latitude λ_m is:

$$\lambda_m = arctan\, l_\infty \tag{3.52}$$

In practice, the first approximation is already close to the solution and, therefore, less than ten iterations are sufficient to converge to a reasonable solution, as the increment $l_{k+1} - l_k$ is of the order of 10^{-15} when $k \to 10$.

Finally, the altitude h_m is given by:

$$h_m = \frac{1}{cos\lambda_m} \left[\sqrt{X_m^2 + Y_m^2} - \frac{a}{\sqrt{1 + (b^2/a^2) tan^2\lambda}} \right] \tag{3.53}$$

3.4 Localisation of the seismicity of Ecuador and Northern Peru

3.4.1 About the study region

A first insight on the localisation of the seismicity of Ecuador was introduced in Chapter 1, showing the main features of the region. We defined the study area dimension based

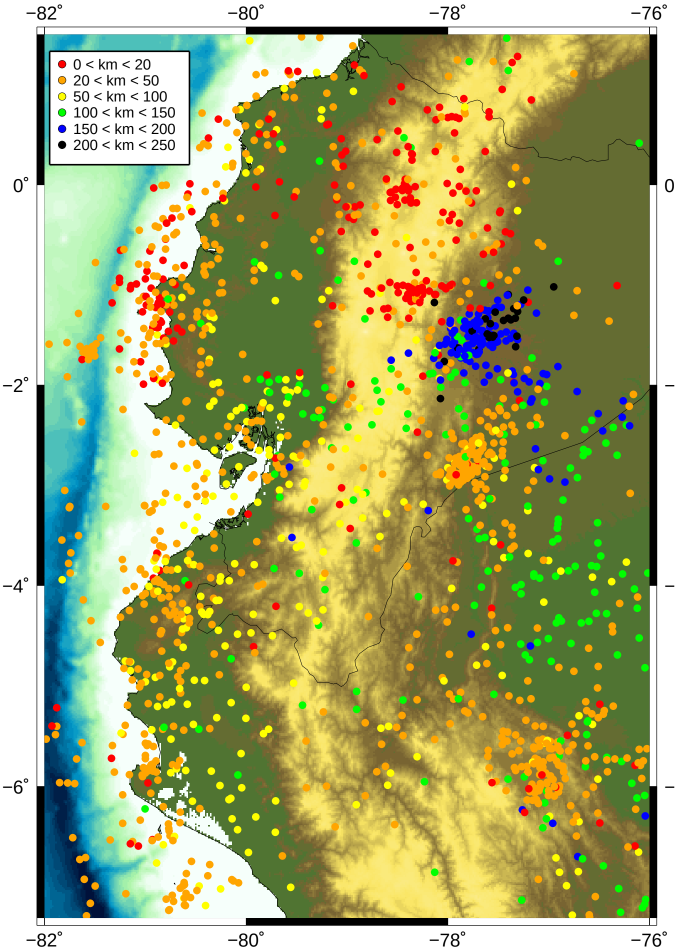

Figure 3.9: *Seismicity location in the study region over the period 1988-2014, as established by the Incorporated Research Institutions for Seismology (IRIS) which is obtained by the analysis of data coming from the Global Seismological Network. Earthquakes of this data base are of magnitude* $M \geq 3$.

on the geographical repartition of the seismicity and the seismological station. The parallelepiped box used for this study encompass a region extending from 76^oW to 82^oW and from 1.5^oN to 7.3^oS (figure 3.9).

3.4.1.1 Extension of the domain

The center of the parallelepiped box (reference point M as described in the previous paragraph) is located at 79.0^oW, 2.9^oS. South-North extension is -338 km $\leq x \leq$ 337 km and West-East extension is -490 km $\leq y \leq$ 490 km.

In a second stage of our study we define another parallelepiped centered at 79.4^oW,

1.7^oS. This new parallelepiped has the same extension in depth that the first one, that is 244 km in the z axis. In the West-East direction its extension is -295 km $\leq x \leq 295$ km and its South-North extension is -385 km $\leq y \leq$ 385 km. We will use this smaller inversion box in the next sections.

Concerning the vertical extension of the parallelepiped, the origin point is located on the ellipsoid. The altitude of the surface topography above this point is 2530 m. As earthquakes can be located anyFunderneath the surface topography, we must insure that this entire surface is in the parallelepiped box. As we can see on figure 3.7, the height of a point in the parallelepiped box depends on its geographic altitude and its distance to the reference point M. Therefore, in order to fix the vertical extension of the box, we had to convert the altitude of the entire topography surface to the local Cartesian coordinate system. The highest point we obtained was the Chimborazo volcano with a geographical altitude of 6 268 m, and a height in the box of $z = 4\,241$ m. As a precaution, a security margin has to be added on the sides of the box in order to avoid side effects in the TIME3D algorithm. Therefore, the highest limit of the box has been fixed 8 km above the reference point M.

In order to set the lower bound of the study box, we considered the locations obtained by previous studies which identified deep events around $z = 240$ km at maximum in the study box coordinate system. These events correspond to the Puyo seismic swarm clearly identified on figure 3.9. Considering the same security margin than for the top bound of the study box, we fixed the lower bound to 244 km deep. Therefore, the vertical extension of the parallelepiped box is -8 km $\leq z \leq$ 244 km as we take the vertical axis oriented toward the center of the Earth.

One can notice that we did not take into account the shape of the seismic rays for the determination of the lower bound of the parallelepiped box. Of course, not only the seismic events and the stations have to be in the domain to be studied, but also the corresponding seismic rays as they represent the time-shortest path followed by the waves to propagate. Therefore, we need to make sure no rays goes deeper than $z = 240$ km in the box. The curvature of a seismic ray is expressed in general as the radius of curvature:

$$R = \frac{1}{p}\frac{dz}{dv} \tag{3.54}$$

Where p is the ray parameter and $v = v(z)$ is the velocity of the seismic wave. This radius has a simple analytically meaning only when the velocity $v(z)$ is a simple velocity gradient function. In general case, a simple numerical test may be used to determine the seismic rays deepest point (Potin, 2016). However, a useful simple first approximation based on a linear velocity model $v(z) = v_0 + kz$ can be computed. Such a model is relatively accurate in the crust and the upper mantle (Udias, 2000). In order to estimate the parameters v_0 and k, we took the AK135 reference model (Kennet at al., 1995). In this model, $v(0) \sim 6$ km/s, which is the velocity near the Earth surface and $v(200) \sim 8$ km/s around 200 km depth. Then:

$$k \sim \frac{8-6}{200-0} \sim \frac{1}{50} \tag{3.55}$$

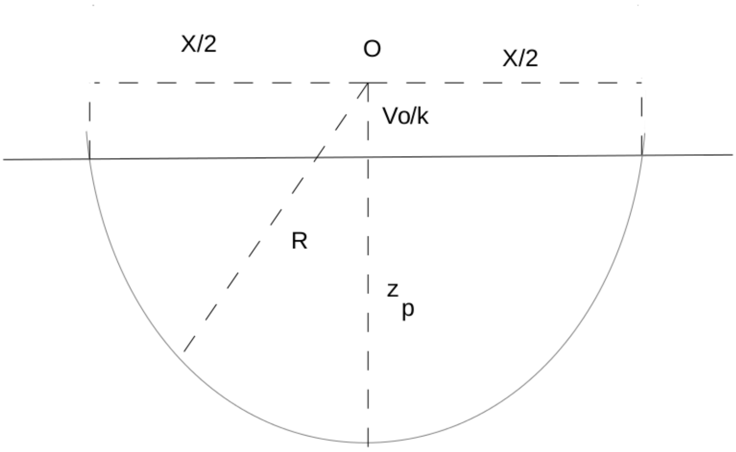

Figure 3.10: *A circular seismic ray in a medium with linear variation of velocity with depth.*

The maximum depth z_p of the ray, corresponding to its turning point let determine the ray parameter:

$$p = \frac{1}{v_0 + kz_p} \tag{3.56}$$

According to 3.54 the curvature radius is:

$$\begin{aligned} R =& \frac{1}{p}\frac{1}{k} \\ R =& \frac{v_0}{k} + z_p \end{aligned} \tag{3.57}$$

R is then constant and the ray is a circular arc as seen in figure 3.10.

Concerning the domain of study, we can express the curvature radius corresponding to a ray crossing the entire domain:

$$R = \left[\left(\frac{d}{2}\right)^2 + \left(\frac{v_0}{k}\right)^2 \right]^{1/2} \tag{3.58}$$

with d the horizontal distance between the source and the receiver. The corresponding maximum depth is then:

$$z_p = R - \frac{v_0}{k} \tag{3.59}$$

Using these expressions, we find that the maximum depth for a ray crossing the domain is $z_p \approx 124$ km in the East-West direction, $z_p \approx 232$ km in the North-South direction. Both these depth are below the lower bound which reaches 244 km. In the extreme case of a ray crossing the box in a diagonal direction we find $z_p \approx 317$ km for this first estimation and therefore we need to obtain a numerical solution of the seismic rays to ensure our preliminary calculations for the inversion box depth.

The numerical test takes account also the problem of rays for deep events and we get this using the tomography code described in the next section. With some iterations of the tomography code, we can look that the number of seismic rays which reach the lower limit of the inversion box is negligible for the established of $244km$. In the eventually that a ray reaches the inversion box borders this will be artificially deformed and LOCIN and tomography codes do not take it in account for the calculations.

3.4.1.2 Discretization of the domain

The last parameter needed to define the computation grid is the length of the cubic cells composing it. In those cells, the Podvin-Lecomte algorithm considers a constant velocity so in order to optimize the quality of computation, their size should be as small as possible. The only limitation concerns the computation itself as the number of cells increases rapidly when their size decreases. Moreover, the LOCIN algorithm (Potin, 2016) virtually duplicate this grid for each data of a single event the memory limitation not only depends on the number of cells but also the maximum number of data per event. Regarding these constraints, we fixed the cell size to $h = 1$ km, which leads to a number of grid cells of $676 \times 981 \times 253 = 167\,778\,468$.

3.4.2 A priori velocity model

The a priori velocity model we chose is a 1-D model, taken as a function of the geographical depth. As the domain of study covers both the crust and the upper mantle, this velocity model is constituted of three parts: one for the crust, one for the upper part of the mantle et one for the rest of the mantle which is simply the model proposed by Kennet et al. (1995). We did not use any Moho topography model in the definition of the velocity model, as we do not want to input any a priori information about its topography. Therefore, the velocity jump between the crust and the upper mantle is reproduced by a smooth change of velocity along a reasonable depth range. The model presented here is a P-wave velocity model.

Concerning the crust, The velocity between the surface topography and the GRS80 ellipsoid is defined by a linear increment from a surface velocity $v_{surf} = 4$ km/s to an intermediate velocity at the ellipsoid surface $v_0 = 4.5$ km/s. Below the ellipsoid, we adopt a model determined by an exponential decay in function of depth d:

$$v(d) = v_0 + \delta v_c \left[1 - exp\left(\frac{d}{d_0}\right)\right] \tag{3.60}$$

Parameters d_0 and δv_c are control parameters. The first one has an impact on the decay of the exponential function and the second one is set in order to obtain a velocity $v_0 + \delta v_c$ close to the value expected at the bottom of the lower crust. In our case, $\delta v_c = 2.6$ km/s and $d_0 = 25$ km. This crustal velocity model is represented on the figure 3.11 as a blue line.

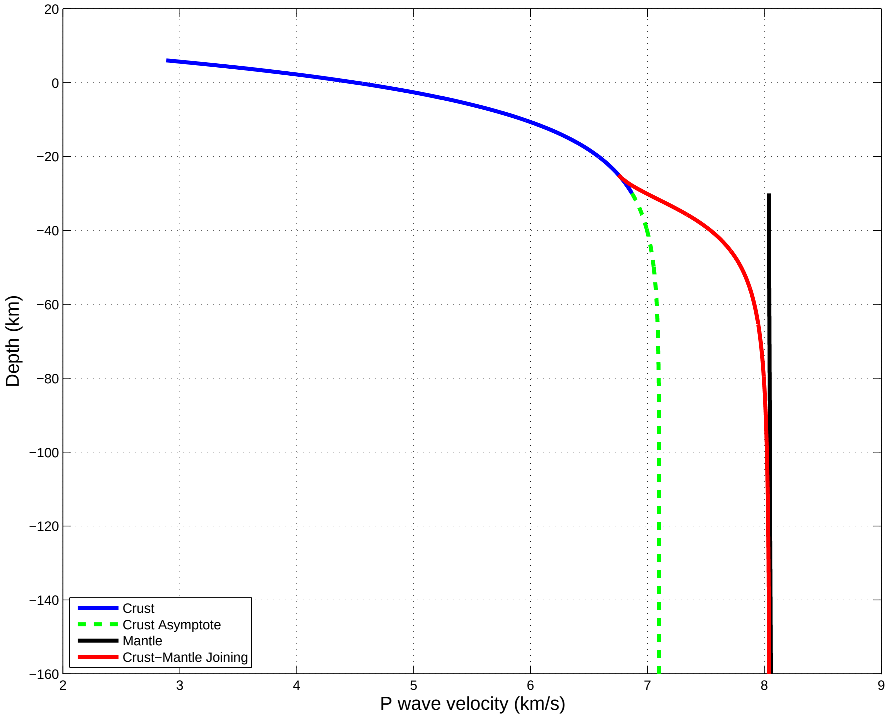

Figure 3.11: *1-D a priori velocity model defined after physical a priori parameters and the AK135 model (Kennet et al., 1995). Here, we represent only the upper part of the model in order to highlight the variation close to the surface.*

Deeper, at a depth range corresponding to the crust-mantle transition, the velocity given by our model progressively rises to reach upper mantle velocities and match the AK135 model (Kennet et al., 1995). This AK135 model is represented on the figure 3.11 as a black line. The depth range at which the transition occurs correspond to the depth range of the Moho beneath Ecuador and Northern Peru. There, the thickness of the crust is highly variable: it reaches about 60 km underneath the Andes Cordillera but decreases down to 5 km near the oceanic plate subduction zone. As the continental wedge is a particular zone underneath which we do not directly encounter mantle material, we chose to joint the crust and the upper mantle velocity model at a depth $d'_0 = 25km$. The transition function take the form:

$$\Delta v(d) = \delta v_m \frac{(d - d'_0)^2}{l^2 + (d - d'_0)^2} \tag{3.61}$$

with $\delta v_m = 0.95$ km/s was set so $v_0 + \delta v_c + \delta v_m$ is close to the velocity expected in upper mantle, between 8.04 km/s and 8.05 km/s between 25 km and 120 km depth (AK135, Kennet et al., 1995). The parameter $l = 13$ km correspond to the transition zone thickness. This transition zone is represented on the figure 3.11 as a red line.

The third part of our 1-D velocity model, between 120 km and 244 km depth, is the

AK135. In this depth range, the AK135 model is constituted by two linear functions corresponding to a velocity change between 8.05 km/s to 8.3 km/s between 120 km and 210 km, and 8.3 km/s and 8.67 km/s between 210 km and -244 km.

Finally, the velocity model of S-waves is obtained from the model of P-waves divided by the mean v_P/v_S ratio obtained through the Wadati diagram analysis (figure 10). As a reminder, the mean value obtained for the Ecuador and Northern Peru region was $v_S = 1.7406 v_P$.

3.4.3 Surface topography

The regional surface topography is an important feature of the model as it has a role in the definition of the velocity model and also because it represent the upper limits for earthquake location. For this study, the surface topography was extracted from the ETOPO1 model provided by Amante and Eakins, (2009) for the entire Earth. In order to take the surface topography into account in the model, this surface was interpolated on the computational grid. The interpolation was performed in the geographic coordinate system and then transported to the Cartesian one. This approach is almost mandatory because in the box, the geographical vertical orientation changes, depending on the relative position to the reference point M. This geometrical problem makes the interpolation very tricky in the box. In order to proceed with this interpolation, we followed the algorithm:

- Cartesian coordinates $(x, y, z = 0)$ for x and y describing the entire grid are converted to the geographic coordinate system. For a random point of the grid, the new set of coordinates is (ϕ, λ, h_0).

- The altitude of the surface topography h_1 is interpolated at the random point (ϕ, λ). The interpolation we performed is based on a barycenter determination:

$$h_1 = \frac{1}{\sum_{k=1}^{4} \frac{1}{d_k}} \sum_{k=1}^{4} \frac{H^{ETOPO\,1}}{d_k} \tag{3.62}$$

where k successively represents the four vertices of the ETOPO1 model cell containing the current random point, d_k is the horizontal distance between this point and each of the ETOPO1 cell summits, and $H_k^{ETOPO\,1}$ is their respective altitude.

- The new coordinates (ϕ, λ, h_1) of the random point is converted back to the Cartesian system, giving a new point $(x + \delta_x, y + \delta_y, z_1)$.

- This operation is repeated but with the new random point (x, y, z_1), and so on until the increment $z_i - z_{i-1} \simeq 0$, which corresponds to the iteration where $\delta_x \simeq 0$ and $\delta_y \simeq 0$.

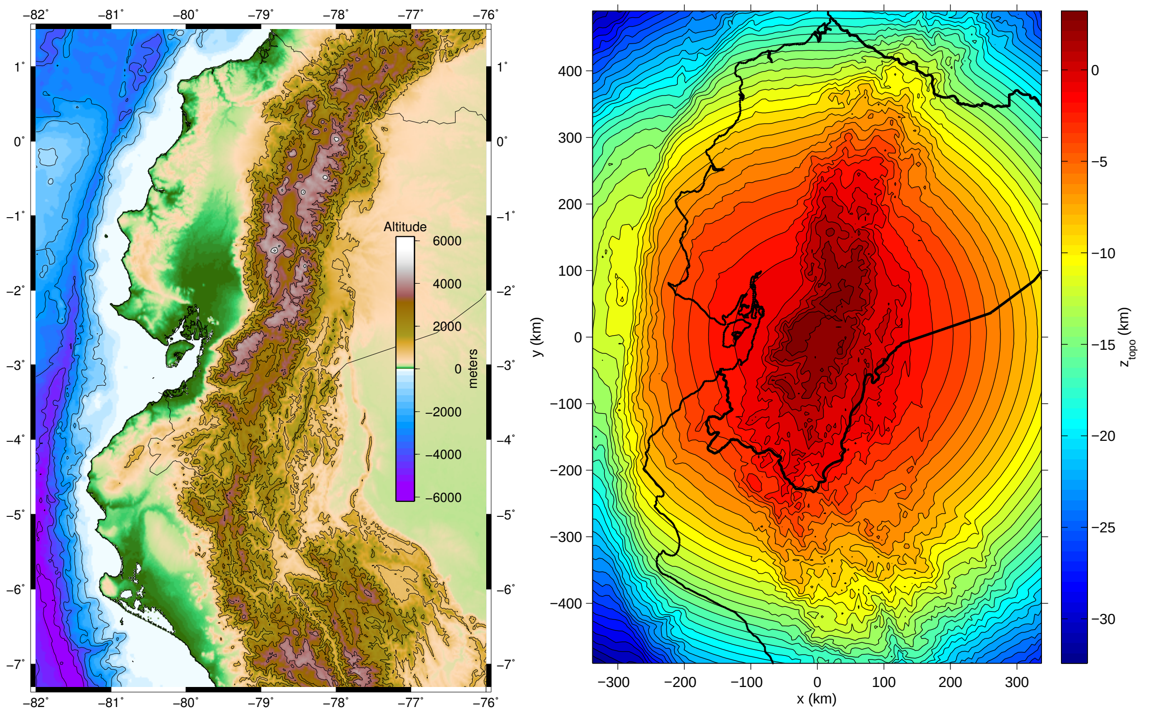

Figure 3.12: *To the left: topography of the studied region. To the right: surface topography interpolated from the ETOPO1 model on the computation grid.* z_{topo} *represent the depth of the topography in the computational box*

In our case, the convergence to the machine precision is reached in nearly fifty iterations, as the first approximation h_0 is very bad for the points located on the edges of the model. The resulting surface topography represented in the parallelepiped box is shown on figure 3.12.

3.4.4 Computational application: the LOCIN algorithm

The theoretical developments presented in the chapter 2 are implemented in the LOCIN software (Potin, 2016), which is a localisation program that follows the approach presented in this chapter. The locations are performed in two steps:

- First, we calculate a "time table" for each couple phase-station present in the dataset. These time tables are files containing the travel time of a front wave propagating from the station to each node of the computational grid. The computation grid has 167 778 468 nodes, and each value is stored as a single precision real number as the precision of these times do not need more than 6 or 7 significant digits. These time tables represent in term of memory usage $167\,778\,468 \times 4 \simeq 671$ Mb for each time table, and 671 Mb × 194 stations × 2 phases $\simeq 260$ Gb in total. The computation of these time tables is only performed once for a velocity model, as the different files will be directly used during the localisation process.

- The LOCIN software then computes the probability density function (equation 3.24) of source location at each grid point, using the time tables. Has it was introduced earlier, this density takes into account both the observational and the theoretical uncertainties.

In practice, the algorithm adopts an optimisation strategy and runs only roughly through the grid (one point out of three in each direction) to approximately determine the density (3.24). Then, the algorithm determines the density in a sub-domain around the maximum previously identified and the final location is determined by interpolation of the density between the nodes. By considering the extension and the shape of the zone of significant values of the density, we can estimate a spatial uncertainty of the localisation.

3.4.5 Earthquake localisation in the a priori velocity model

localisations of the 44, 316 seismic events was performed with the LOCIN software. The resulting locations are presented on figure 3.13.

Earthquake Locations determined by mean of the grid search approach implemented in the LOCIN software are in good agreement with the previous knowledge of the seismicity in the region, as for example the locations provided by the Global Seismology Network (figure 3.9). The algorithm produces high precision locations in both horizontal and vertical directions, allowing us to observe shallow seismicity clusters with depth lower than 35 km which are due to volcanoes activity during the study period of 1988-2014 (Chiles, Reventador, Guagua Pichincha and Tungurahua). The Cotacachi-Cuicocha volcanic complex presents as well a clear activity, although this complex has not been in eruption during the period of study. Other important shallow seismicity clusters are related to crustal faults activity (North Quito, Pujilí, Pisayambo and Macas).

The trench region seismicity produced by the subduction of the Nazca plate underneath the South America Plate is ordered around three clusters: Galera, Jama and Manta-Puerto López (Font et al., 2013).

Three mid-depth clusters are also identifiable and correspond to normal manifestation of the subduction process: Guayaquil (50 km to 100 km), Maldonado (80 km to 100 km) and La Maná (80 km to 100 km).

Finally, the deep Puyo cluster is also well identifiable with depths greater than 150 km and some events reaching 240 km.

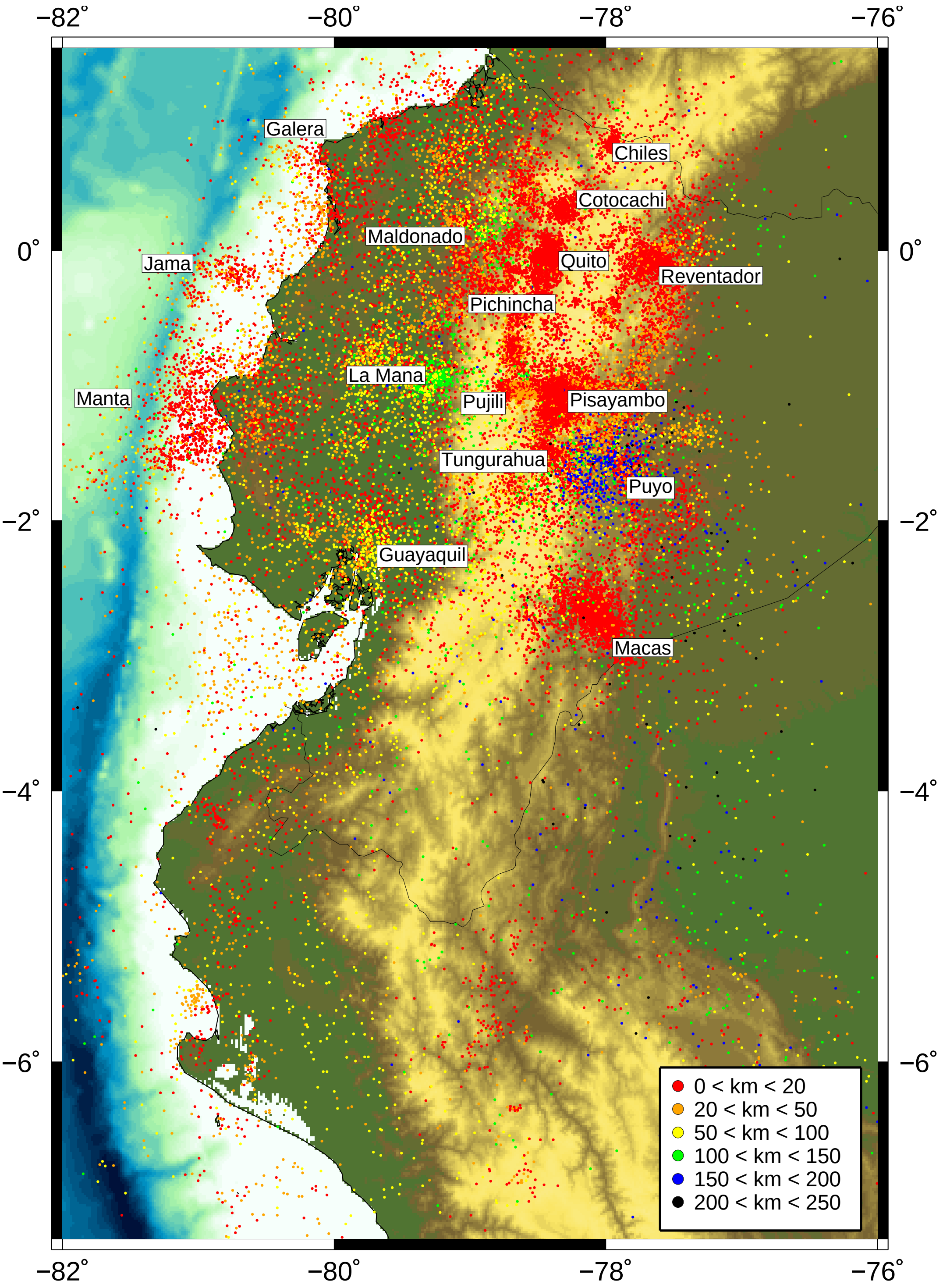

Figure 3.13: *Location map of the 44,316 earthquakes used in this study. Location were determined by LOCIN software. The names correspond to the seismic clusters. Shallow clusters due to volcanoes are Chiles, Reventador, Cotacachi, Guagua Pichincha and Tungurahua. Shallow clusters due to crust faults are North Quito, Pujilí, Pisayambo and Macas. The three clusters of trench are Galera, Jama and Manta-Puerto López. The three subduction mid-depth clusters are Maldonado, La Maná and Guayaquil. The deepest cluster is Puyo.*

Chapter 4

The inverse problem of travel time tomography

Our approach of the travel time tomography is similar to the one we have followed for the localization of seismic events. It is a Bayesian approach in which we consider both models and data as Gaussian random elements. The introduction of the velocity fields in the model space leads us to consider Gaussian random functions over the Euclidean space. Thus, the formal results of section 2.5 can be applied in the inversion. We will detail in this chapter how it is practicaly achieved for very large datasets with up to one million of data, as that of Ecuador, which strategy is followed for the discretization, and how it was implemented in the INSIGHT codes (B. Potin, B. Valette, and V. Monteiller, 2016), in the development of which I was partly involved. Let us firstly set the travel time tomography problem.

4.1 Setting the travel time tomography problem

4.1.1 The data space

The data are provided by the catalogue of wave arrival times that we have built up, and which is described in the first chapter. Practically, we have used as data, the P-waves arrival times t_{P}, and the difference of arrival times of S and P waves, $\Delta t = t^S - t^P$, at a same station. We prefer to use as data these differences instead of the S arrival time to eliminate the possible errors of clock drift at stations. Thus, the generic data vector $\mathbf{d}$ is defined in $\mathbb{D} = \mathbb{R}^{m+n}$ and can be split into two subvectors $\mathbf{d}^{\mathrm{S-P}} = (t^S - t^P)_{i=1,m} \in \mathbb{R}^m$ and $\mathbf{d}^P = (t^P)_{i=m+1,m+n} \in \mathbb{R}^n$.

4.1.2 The model space

The model space $\mathbb{M}$ comprises the functional parameters v_P and v_p/v_s defining the velocity fields in the domain V and the scalar parameters (localisations and initial times) identifying the seismic events $(\mathbf{x}, \mathbf{t}_0)$. We have also included as scalar parameters in the model space some time delays $(\Delta^P, \Delta^{\mathrm{S-P}})$ that are attached to each station, and are added to the theoretical travel times T_P or Δt. The number of scalar parameters is thus equal to $4n_e + 2n_s$ in noting by n_e the number of events and n_s the number of stations.

The use of v_P/v_S rather than that of v_S is justified by the fact that this ratio is relatively well constrained and that an a priori value can be easily obtained through a Wadati diagram analysis, as shown in the first chapter. The station delays Δ^P and $\Delta^{\mathrm{S-P}}$ account for possible site effects in the close vicinity of stations. Indeed, the local resolution of velocity is very poor just beneath a station due to a lack of intersection of rays in this zone, within which all the rays joining the station are more or less parallel to each other. Thus the inversion does not allow the identification of possible velocity anomalies localized just below the stations, such as a thin layer of unconsolidated sediment or volcanic materials. Introducing station delays provides a way to take such possible concentrated anomalies into account, in avoiding its spreading deeper into the model resulting of the inversion.

We will denote by $\mathbf{m} = \left[v_P, v_P/v_S, \mathbf{x}, \mathbf{t_0}, \boldsymbol{\Delta}^P, \boldsymbol{\Delta}^{\mathrm{S-P}}\right] \in \mathbb{L}^2 \times \mathbb{L}^2 \times \mathbb{R}^{4n_e+2n_s}$ the generic vector of the model space $\mathbb{M}$.

4.1.3 The travel time equations

The theoretical mapping $\mathbf{d} = g(\mathbf{m})$ can be written as:

$$
\begin{aligned}
\mathbf{d} &= \begin{bmatrix} \mathbf{d}^{\mathrm{S-P}} \\ \\ \mathbf{d}^P \end{bmatrix} = g(\mathbf{m}) \\
&= \begin{bmatrix} \Delta T_i(v_P, v_P/v_S, \mathbf{x}_{j(i)}) + \Delta^{\mathrm{S-P}}_{k(i)} \\ \\ T_i^{\mathrm{P}}(v_P, \mathbf{x}_{j(i)}) + t_{0i} + \Delta^{\mathrm{P}}_{k(i)} \end{bmatrix} \qquad \begin{array}{c} 1 \le i \le m \\ \\ m+1 \le i \le m+n \end{array}
\end{aligned}
\tag{4.1}
$$

where the travel times T_i^{P} verify:

$$
T_i^{\mathrm{P}} = \int_{\mathcal{R}_i^{\mathrm{P}}} \frac{\mathrm{d}s}{v_P(s)} \tag{4.2}
$$

and the difference of travel times ΔT_i:

$$
\Delta T_i = \int_{\mathcal{R}_i^S} \frac{v_P}{v_S}(s) \frac{\mathrm{d}s}{v_P(s)} - \int_{\mathcal{R}_i^{\mathrm{P}}} \frac{\mathrm{d}s}{v_P(s)} \tag{4.3}
$$

In these equations :

- $\mathcal{R}_i^{\mathrm{P}}$ or $\mathcal{R}_i^S$ denotes respectively the ray of the P or S wave linking the event with index $j(i)$ localized at $\mathbf{x}_{j(i)}$ to the station with index $k(i)$, and with station delay $\Delta_{k(i)}^{\mathrm{P}}$ or $\Delta_{k(i)}^{\mathrm{S-P}}$, where both the station and the event correspond to the ith datum.
- s denotes the curvilinear abcissa along the rays.

g is a non-linear mapping since the rays depend on the velocity model. But, as a consequence of the so called Fermat principle, a first order perturbation of the ray does not change the travel time. The derivative operator $\mathbf{G}_m$ of g can thus be easily evaluated by perturbation from (4.1, 4.2, 4.3)). More precisely, it yields:

$$\delta g(\mathbf{m}) = \begin{bmatrix} \delta\mathbf{t}^{\mathrm{S-P}} \\ \delta\mathbf{t}^{\mathrm{P}} \end{bmatrix} = (\mathbf{G}_m(\delta v_{\mathrm{P}}, \delta v_{\mathrm{P}}/v_{\mathrm{S}}, \delta\mathbf{x}, \delta\mathbf{t_0}, \delta\boldsymbol{\Delta}^P, \delta\boldsymbol{\Delta}^{\mathrm{S-P}}) =$$

$$\begin{array}{ll} \displaystyle\int\limits_{\mathcal{R}_i^{\mathrm{S}}} \frac{1}{v_{\mathrm{P}}}\left(\delta\left(\frac{v_{\mathrm{P}}}{v_{\mathrm{S}}}\right) - \frac{v_{\mathrm{P}}}{v_{\mathrm{S}}}\frac{\delta v_{\mathrm{P}}}{v_{\mathrm{P}}}\right)\mathrm{d}s + \int\limits_{\mathcal{R}_i^{\mathrm{P}}} \frac{\delta v_{\mathrm{P}}}{v_{\mathrm{P}}^2}\mathrm{d}s + (\boldsymbol{\tau}_i^{\mathrm{P}} - \boldsymbol{\tau}_i^{\mathrm{S}})\cdot\delta\mathbf{x}_{j(i)} + \delta\Delta_{k(i)}^{\mathrm{S-P}} & 1 \le i \le m \\ \displaystyle -\int\limits_{\mathcal{R}_i^{\mathrm{P}}} \frac{\delta v_{\mathrm{P}}}{v_{\mathrm{P}}^2}\mathrm{d}s - \boldsymbol{\tau}_i^{\mathrm{P}}\cdot\delta\mathbf{x}_{j(i)} + \delta t_{0j(i)} + \delta\Delta_{k(i)}^{\mathrm{P}} & m+1 \le i \le m+n \end{array} \tag{4.4}$$

with the same notation as in (4.1) (4.2) (4.3), and where

- δ denotes the perturbation of a scalar or vectorial quantity,
- $\boldsymbol{\tau}_i^{\mathrm{P}}$ and $\boldsymbol{\tau}_i^{\mathrm{S}}$ respectively denotes the unit vector tangent to the P or S ray corresponding to datum i in the direction of increasing propagating time, respectively divided by the P or S velocity, $v_{\mathrm{P}}(\mathbf{x}_{j(i)})$ or $v_{\mathrm{S}}(\mathbf{x}_{j(i)})$, at the hypocenter $\mathbf{x}_{j(i)}$ of the event $j(i)$,
- $\cdot$ denotes the usual Euclidean scalar product in $\mathbb{R}^3$

4.1.4 The a priori information

The vector $\mathbf{d}_{obs}$ of data, provided by our catalogue, has to be complemented by a *pdf* representing the errors in measurement. The corresponding standard deviations are firstly set according to the quality weight as indicated in Table 1.1. For difference data of S and P arrival times, the standard deviation is the square root of the sum of the corresponding S and P data, since the two measurements may be considered as independent.

But, as all the data are the results of picking of P or S phases in seismograms, there exist surely outliers among them, mainly due to misidentified P or S phases (see *e.g.*

Nolet, 2008), and consequently, as already indicated in paragraph 2.4.3, the corresponding *pdf* cannot be Gaussian. We have also shown in the paragraph 2.4.3 that a way to enhance the robustness of the inversion, in avoiding the bias induced by the outliers, is to take for each of the data, d, an hyperbolic secant *pdf* (2.42) that can be rewritten as:

$$\rho_{obs}(d) = \frac{1}{2\sigma_{obs}} \frac{1}{\cosh\left(\frac{\pi}{2}\frac{d-d_{obs}}{\sigma_{obs}}\right)} \tag{4.5}$$

where σ_{obs} is the standard deviation. This *pdf* is equivalent to either a Gaussian or to a decreasing exponential function according to $|d - d_{obs}|$ is large or small :

$$\rho_{obs}(d) \underset{\infty}{\sim} \frac{1}{\sigma_{obs}} \exp\left(-\frac{\pi}{2}\frac{|d-d_{obs}|}{\sigma_{obs}}\right) \qquad \rho_{obs}(d) \underset{0}{\sim} \frac{1}{2\sigma_{obs}} \exp\left(-\frac{\pi^2}{8}\frac{(d-d_{obs})^2}{\sigma_{obs}^2}\right)$$

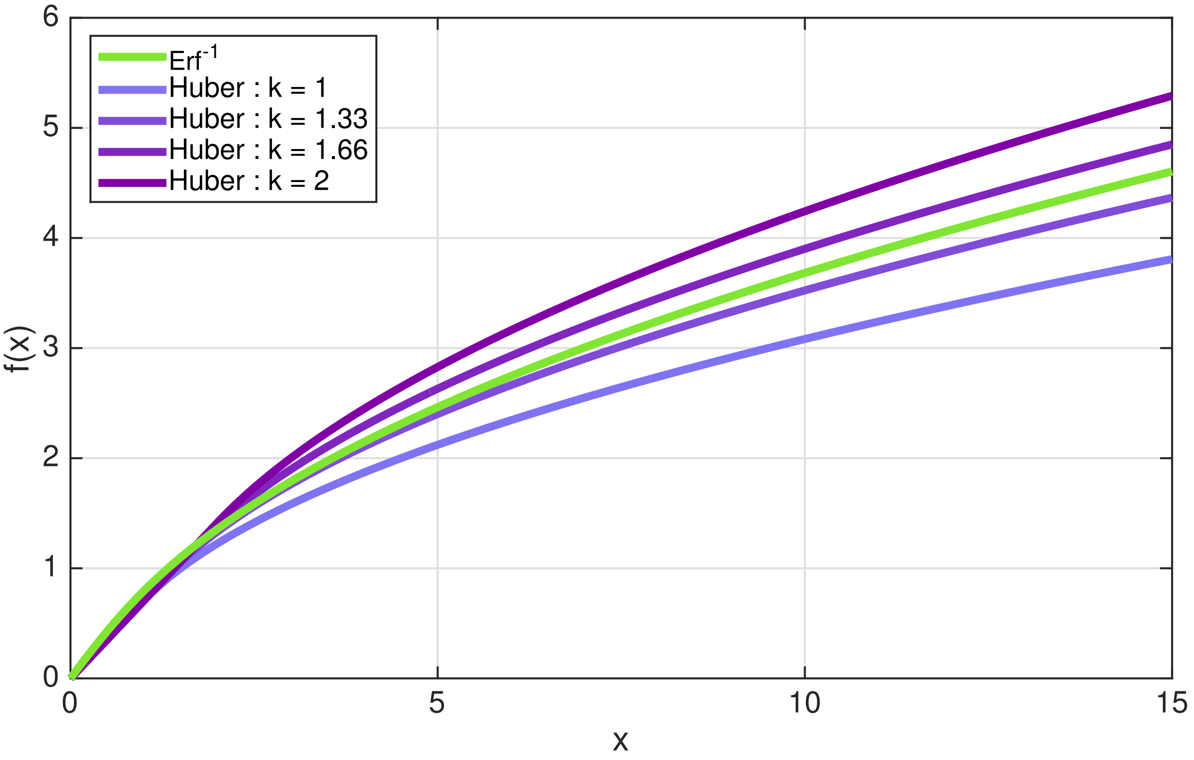

Figure 4.1: *Comparison of the secant hyperbolic pdf with the Huber M-estimator (2.41). In green the least-square variable corresponding to the secant hyperbolic pdf for a datum d,* $y(x) = Erf^{-1}\left({}^{2}/_{\pi} arctan(sinh({}^{\pi x}/_{2}))\right)$ *as function of the variable* $x = |d - d_{obs}|/\sigma_{obs}$*; in blue-purple, the least-square variable corresponding to the Huber M-estimator (2.41) for various values of k:* $y(x) = x/\sqrt{2}$ *for* $x \leq k$ *and* $y(x) = \sqrt{k(x - k/2)}$ *for* $x \geq k$*.*

The corresponding change of variable $y(d)$ (2.45, 2.46) then reads :

$$y(d) = \mathrm{Erf}^{-1}\left[\frac{2}{\pi}\mathrm{arctan}\left(\sinh\left(\frac{\pi}{2}\frac{d-d_{obs}}{\sigma_{obs}}\right)\right)\right] \tag{4.6}$$

with:

$$\frac{\mathrm{d}y}{\mathrm{d}d} = \frac{\sqrt{\pi}}{2\sigma_{obs}} \frac{\exp(y^2)}{\cosh\left(\frac{\pi}{2}\frac{d-d_{obs}}{\sigma_{obs}}\right)} \tag{4.7}$$

By using standard asymptotic developments (*e.g.* NIST handbook of mathematical functions, Olver et al., 2010), an easy calculation yields :

$$y \underset{0}{\sim} \frac{\sqrt{\pi}}{2}\frac{d-d_{obs}}{\sigma_{obs}} \qquad y \underset{\infty}{\sim} \sqrt{\frac{\pi}{2}}\,\mathrm{sgn}(d-d_{obs})\sqrt{\frac{|d-d_{obs}|}{\sigma_{obs}}} \tag{4.8}$$

and:

$$\frac{\mathrm{d}y}{\mathrm{d}d} \underset{\infty}{\sim} \frac{1}{2}\sqrt{\frac{\pi}{2\sigma_{obs}}}\frac{\mathrm{sgn}(d-d_{obs})}{\sqrt{|d-d_{obs}|}} \tag{4.9}$$

where sgn denotes the sign function (equal to 1 or -1 according to the sign). It shows that using these new variables instead of the initial variables is equivalent to using Huber M-estimator (Huber, 1964, and equation 2.41), while remaining in the Gaussian framework. We can also remark that the product $y\frac{\mathrm{d}y}{\mathrm{d}d}$ converges toward the finite limit $\frac{\pi}{4\sigma_{obs}}$ for large $|d-d_{obs}|/\sigma_{obs}$. This confirms that the outliers then do not play an important role in the equation of stationnarity of the cost function, which is now quadratic in y.

The two random fields v_P and $v_\mathrm{P}/v_\mathrm{S}$ are assumed to be Gaussian, with mean prior model corresponding to the one used for the a priori localisation described in chapter 3. The correlation kernel of each of these two random functions, assumed to be independent, is the exponential kernel, described in paragraph 2.5.3. The a priori physical standard deviation of the two fields are assumed to be constant over the study domain, thus the effective standard deviations are also constant. As a consequence, the covariance kernel of each of the two fields can be written as:

$$C(\mathbf{r},\mathbf{r}') = \sigma^2\exp\left[-\left(\sum_{i=1}^{3}\frac{(r_i-r_i')^2}{\xi_i^2}\right)^{1/2}\right] \tag{4.10}$$

with the same various lengths of correlation ξ_1, ξ_2, ξ_3 that corresponds to the three directions of the Euclidean space, and where, in taking (2.69) into account:

$$\sigma^2 = \left(\frac{\xi_0^3}{\xi_1\xi_2\xi_3}\right)\sigma_{phys}^2 \tag{4.11}$$

The other scalar parameters, *i.e.* location and initial time of the events, and time delays attached to the stations, are all assumed to be Gaussian and pairwise independent.

4.2 Inversion strategy

As we have shown in chapter 2, the Bayesian approach leads to a Tikhonov problem, *i.e.* searching for the minimum over a regularization space of a cost function which can be written in the general form:

$$||\mathbf{C}_d^{-1/2}(\mathbf{d}^{obs} - g(\mathbf{m}))||_{\mathbb{D}}^2 + ||\mathbf{T}(\mathbf{m} - \mathbf{m}_{prior})||_{\mathbb{M}}^2 \tag{4.12}$$

where $\mathbf{T}$ is the regularization operator, $\mathcal{D}(\mathbf{T}) \subset \mathbb{M}$ is the regularization space (*i.e.* the domaine of the operator $\mathbf{T}$ in $\mathbb{M}$), and $\mathbf{m}_{prior}$ a reference model. In the stochastic case, $\mathbf{m}_{prior}$ corresponds to the a priori expectation, and $\mathbf{T}$ to the square root of the inverse of the prior covariance operator $\mathbf{C}_m^{-1/2}$ of the model. It leads to the stationnarity equation (2.66) and the algorithm (2.67) that involves the solution of a linear system with matrix M (2.68) in the finite dimensional data space $\mathbb{D}$. A relatively long, but easy calculation that takes (4.4) into account and that the kernels h_m are actually distributions over the rays, allows the explicitation of the entries of the matrix M:

- for $1 \leq i \leq m$ and $1 \leq j \leq m$:

$$\begin{aligned}
\mathrm{M}_{ij} = \mathrm{C}_{ij}^d &+ \sigma_{\mathrm{P}}^2 \int\limits_{\mathbb{R}_i^{\mathrm{S}}} \int\limits_{\mathbb{R}_j^{\mathrm{S}}} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{S}}(\mathbf{x})v_{\mathrm{P}}(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{S}}(\mathbf{x}')v_{\mathrm{P}}(\mathbf{x}')} \\
&+ \sigma_{\mathrm{P}}^2 \int\limits_{\mathbb{R}_i^{\mathrm{P}}} \int\limits_{\mathbb{R}_j^{\mathrm{P}}} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{P}}^2(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{P}}^2(\mathbf{x}')} \\
&- \sigma_{\mathrm{P}}^2 \int\limits_{\mathbb{R}_i^{\mathrm{S}}} \int\limits_{\mathbb{R}_j^{\mathrm{P}}} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{S}}(\mathbf{x})v_{\mathrm{P}}(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{P}}^2(\mathbf{x}')} \\
&- \sigma_{\mathrm{P}}^2 \int\limits_{\mathbb{R}_i^{\mathrm{P}}} \int\limits_{\mathbb{R}_j^{\mathrm{S}}} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{P}}^2(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{S}}(\mathbf{x}')v_{\mathrm{P}}(\mathbf{x}')} \\
&+ \sigma_{v_{\mathrm{P}}/v_{\mathrm{S}}}^2 \int\limits_{\mathbb{R}_i^{\mathrm{S}}} \int\limits_{\mathbb{R}_j^{\mathrm{S}}} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{P}}(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{P}}(\mathbf{x}')} \\
&+ \sum_{\alpha=1}^{3} ((\tau_i^{\mathrm{P}})^\alpha - (\tau_i^{\mathrm{S}})^\alpha)((\tau_j^{\mathrm{P}})^\alpha - (\tau_j^{\mathrm{S}})^\alpha)(\sigma_{\ell(i)}^\alpha)^2 \delta_{\ell(i)}^{\ell(j)} + (\sigma_{\Delta_{k(i)}^{\mathrm{S-P}}})^2\, \delta_{k(i)}^{k(j)}
\end{aligned} \tag{4.13}$$

- for $m+1 \leq i \leq m+n$ and $m+1 \leq j \leq m+n$:

$$\begin{aligned} \mathrm{M}_{ij} = \mathrm{C}^d_{ij} &+ \sigma_{\mathrm{P}}^2 \int\limits_{\mathbb{R}^{\mathrm{P}}_i} \int\limits_{\mathbb{R}^{\mathrm{P}}_j} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{P}}^2(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{P}}^2(\mathbf{x}')} \\ &+ \left(\sum_{\alpha=1}^{3} (\tau_i^{\mathrm{P}})^\alpha (\tau_j^{\mathrm{P}})^\alpha (\sigma^\alpha_{\ell(i)})^2 + \left(\sigma^{t_0}_{\ell(i)}\right)^2\right) \delta^{\ell(j)}_{\ell(i)} + (\sigma^{\Delta^{\mathrm{P}}}_{k(i)})^2\, \delta^{k(j)}_{k(i)} \end{aligned} \tag{4.14}$$

- for $m+1 \leq i \leq m+n$ and $1 \leq j \leq m$:

$$\begin{aligned} \mathrm{M}_{ij} = \mathrm{C}^d_{ij} &+ \sigma_{\mathrm{P}}^2 \int\limits_{\mathbb{R}^{\mathrm{P}}_i} \int\limits_{\mathbb{R}^{\mathrm{P}}_j} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{P}}^2(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{P}}^2(\mathbf{x}')} \\ &+ \sigma_{\mathrm{P}}^2 \int\limits_{\mathbb{R}^{\mathrm{P}}_i} \int\limits_{\mathbb{R}^{\mathrm{S}}_j} \varphi\left(\frac{\|\mathbf{x}-\mathbf{x}'\|}{\xi}\right) \frac{\mathrm{d}s_i(\mathbf{x})}{v_{\mathrm{P}}^2(\mathbf{x})} \frac{\mathrm{d}s_j(\mathbf{x}')}{v_{\mathrm{S}}(\mathbf{x}')v_{\mathrm{P}}(\mathbf{x}')} \\ &+ \left(\sum_{\alpha=1}^{3} (\tau_i^{\mathrm{P}})^\alpha ((\tau_j^{\mathrm{P}})^\alpha - (\tau_j^{\mathrm{S}})^\alpha)(\sigma^\alpha_{\ell(i)})^2\right) \delta^{\ell(j)}_{\ell(i)} \end{aligned} \tag{4.15}$$

where:

- $\delta_k^{k'}$ is the usual Kronecker symbol,
- the index $\ell(i)$, in (4.14) is the index of the event corresponding to the ith datum (to avoid any confusion with index j),
- $\mathcal{R}_i^{\mathrm{P}}$ and $\mathcal{R}_i^{\mathrm{S}}$ represents respectively the P, and possibly S, ray involved in the ith datum,
- σ_{P}, $\sigma_{v_{\mathrm{P}}/v_{\mathrm{S}}}$, $\sigma^\alpha_{\ell(i)}$, $\sigma^{t_0}_{\ell(i)}$, $\sigma^{\Delta^{\mathrm{P}}_{k(i)}}$, $\sigma^{\Delta^{\mathrm{S-P}}_{k(i)}}$ respectively denote the prior standard deviation of the v_P and v_P/v_S fields, of the 3 components of the position vector and of the initial time corresponding to the $\ell(i)$th event, and of times delays attached to the $k(i)$th station.

When using the variables y_i associated to a secant hyperbolic *pdf* for the data, all the terms at the right hand side of (4.13), (4.14), (4.15), excepting C^d_{ij}, must be multiplied by $\frac{\mathrm{d}y_i}{\mathrm{d}d_i}\frac{\mathrm{d}y_j}{\mathrm{d}d_j}$ (4.7).

Thus, each entry of the matrix needs the computation of a double integral over rays within the study domain; which is relatively time expensive. A first attempt was made in this way by (Nercessian et al., 1984) with a very small dataset of 480 P-phases, by adopting a first order approach, and by assuming Gaussian correlation kernels and straight rays that allow for analytical integration of the matrix entries. But, despite the dramatic increase in computer power, this approach cannot be practically used for large datasets, such as seismic catalogue with up to one million data. All the more so because it would be currently almost impossible to solve a full linear system of about 1 million equations in a reasonable time.

An alternative way is to use the algorithm in its quasi-Newton formulation :

$$\mathbf{m}_{k+1}-\mathbf{m}_k = -(\mathbf{C}_m^{-1}+\mathbf{G}_k^*\mathbf{C}_d^{-1}\mathbf{G}_k)^{-1}(\mathbf{G}_k^*\mathbf{C}_d^{-1}(\mathbf{g}(\mathbf{m}_k)-\mathbf{d}_{obs})+\mathbf{C}_m^{-1}(\mathbf{m}_k-\mathbf{m}_{prior})) \quad (4.16)$$

which needs, that time, the solution of a linear inversion in an infinite dimensional space. This involves discretizing the model space, and consequently the study domain, while keeping in background the functional character of the fields, especially for correlations. Furthemore, it leads to a linear system, the matrix of which is of huge order (greater than the data number and than twice the number of grid nodes within the study domain). But, the key point is that this matrix is made very sparse when using exponential kernel for the covariance, since $\mathbf{C}_m^{-1}$ (2.59, 2.61) is then a differential operator that is easily sparsely discretized by finite difference, and since the acting of operator $\mathbf{G}_k$, and of its adjoint, only involves the model values at the grid nodes in the vicinity of the ray corresponding to each datum. This approach is thus the only workable one when considering tomography problems with a large number of data, except for very specific cases where some simplification may be done (Vergely et al., 2010).

4.3 The INSIGHT computer code package

There exist numerous computer codes to perform regional travel time tomography. One of the most commonly used is the SIMULPS code (Thurber, 1983; Um and Thurber, 1987, Eberhart and Phillips, 1986, 1990; Evans et al., 1994). But other codes may be used as the LOTOS code (koulakov, 2009), or the code of (Benz et al., 1996). The algorithms of these codes generally amount to minimizing a cost function in the neighbourhood of an initial or prior model, *i.e.* to a Tikhonov problem. The cost function (eqn 4.12) contains a term corresponding to the adjustment of the data and most of the time another one corresponding to the regularization of the model, i.e. a term that depends on the discrepancy of the model with respect to a prior one and/or on the smoothness of the model through an operator $\mathbf{T}$. From a practical point of view, it may be thought of as introducing additional regularization data corresponding to the vanishing of this term $\mathbf{T}(\mathbf{m}-\mathbf{m}_{prior})$. The minimum of the cost is usually obtained through a quasi Newton algorithm in a discretized model space.

In the case of the SIMULPS code, there is no regularization term in the cost function. The stability of the quasi-Newton algorithm is insured by adding a positive term in the diagonal of the quasi-Hessian matrix, in a way analogous to the Levenberg-Marquardt algorithm. The value of this term is determined by trade off between the fit of the data and the smoothness of the model. Moreover the level of smoothness of the model is basically controlled by the grid spacing of the study domain. Each cell of the grid must be sufficiently large to be crossed by a minimum number of rays. It raises a number of problems; from a practical perspective, the variability in size of the grid can make its geometry very complex, and consequently, accuracy in the forward computation of the traveltimes may be lost. Since there is no regularisation term in the cost function, there is no prior model. Only an initial model is needed to start the algorithm. That makes this choice crucial. This initial model is generally a 1-D model determined along with the location of events through a prior inversion (Kissling et al., 1994). In this case, it is very important to take into account the modelling error due to the 1-D assumption in order to avoid over-fitting of the data; which would bias the model.

In the case of the LOTOS code, the regularization term corresponds to the H^1 norm of the velocity model. $(\int v^2 \mathrm{d}V + \int \|\nabla\|^2 \mathrm{d}V$. That means that the additional regularizing data are the values of the velocity, and of their differences in each direction and between adjacent points of a regular grid.

The algoritm of Benz is relatively hybrid. Assuming that a reference model is given, the events are firstly localized by a standard least squares approach. Then the problem of tomography is linearized with respect to both velocity fields and event locations around these reference values. Following (Pavlis and Booker, 1980) a set of new data is determined as linear combinations of the data in order to eliminate the first order influence of the earthquake locations. This is theoretically possible if the number of independent data is greater than the number of event location parameters, but it requires the computation of a vector basis of the null space of the adjoint of the operator of partial derivative of the data with respect to the event parameters, which is very time consuming for only a first order approximation. These new data are complemented by regularizing data that correspond to the Laplacian of the velocity fields discretized by finite difference over the grid. From a Tikhonov point of view, it corresponds to the first iteration of a minimizing process of a Tikhonov cost, the regularization term of which is the L^2 norm of the Laplacian of the velocity field (which corresponds to only a semi norm for the velocity).

All these codes do not allow for managing large datasets of variable quality. They are commonly used for dataset containing at most a few thousands of data carefullly selected. Unfortunately, there are no perfect dataset in seismology, and as shown in the first chapter, outliers or errors in files may be very hard to detect. We have thus used the new package of computer codes INSIGHT, which has been designed by B. Valette and B. Potin, and written by B. Potin in Fortran 2003 during his PHD, and in the development of which I have been involved as well as V. Monteiller.

The goal of the codes INSIGHT is to perform the travel time tomography from regional earthquakes. They are designed to invert large (up to one million) catalogue of data, with variable quality, for local structure. It allows the use of all the data available over the study zone, which offers two advantages. Firstly, it provides maximal coverage of the study domain by the rays, and secondly, we may raisonably expect that the information is stabilized by the "large numbers" of data. The origin of the code comes back to a double difference tomography code developped by V. Monteiller (Monteiller et al., 2005; Monteiller, 2005), initialy dedicated to the Hawaï volcanic dome. The density of the Hawaï seismic network and the quality of the data allow the mesurement of arrival time differences between two events at a same station, by cross correlation of the signals, which considerably increases the accuracy of measurement (Poupinet et al., 1984; Frechet, 1985; Got et al., 1994) with repect to hand picking. Thereafter, a variant of this code was developped by V. Monteiller and B. Valette to perform direct travel time tomography. This variant was used for regional tomography of the western Alps (Potin et al., 2012, 2014, 2015a,b), and of Ecuador (Araujo et al., 2014). The main points of this code is the use of the analytical form of the inverse covariance operator for an exponential kernel, the introduction of clear tuning parameters for the smoothing and damping, the use of an hyperbolic secant *pdf* for data, and the decorrelation of the crust with the mantle. The code was written by V. Monteiller, and upgraded with B. Valette, in Fortan 77. It was completly restructured, and rewritten in Fortan 2003 by B. Potin (2016), with a good level of efficiency in the parallelisation, based on the OpenMP and MPI libraries, and an improvement of the accuracy and stability, due to the introduction of the "double precision". A variant of this code allows for inverting differences of data, related either to a same event or to a same station. It allows the relocalization of a seismic swarm or the precise tomography of a sub-domain related to an object of interest, as the sub-structure of a volcano, below a dense sub-network.

Let us now described the discretization of the model space, and of the tomography process, implemented in these codes.

4.3.1 Discretization of the functional parameters, and model space

As described in chapter 3, the study zone is a parallelipipedic domain, the base direction of which is parallel with the tangent plan to the reference ellipsoid GRS80 at a reference point for the zone, and the upper plan of which is a few kilometers higher than this tangent plan in order to include the topograhy within the domain. Each point of the domain is identified by either its geographical coordinates (latitude, longitude, altitude) or by its cartesian coordinates in the parallelipipedic box, as shown in chapter 3. A regular grid in defined within the box, with parallelipipedic cells. A finer grid is also defined for the Podvin Lecompte algorithm, usualy with an exact number of Podvin-Lecomte cubic cells in each reference cell, to insure a good accuracy of the travel time calculation. The fields v_P and $v_\mathrm{P}/v_\mathrm{S}$ are dicretized into the finite dimensional vectors of their values at each nodes of the reference grid. Since we need their values at any point inside the parallelipidic box to compute the travel times, or to evaluate the derivative operator, the values at the grid nodes must be interpolated. The trilinear interpolation (linear in each direction) yields simple formula, and is sufficiently precise when using a tigh grid spacing. Given the set of values $f(x_i, y_j, z_k)$ of a scalar function at each nodes (i, j, k) of this regular grid, the trilineral interpolation yields

$$
\begin{array}{rccccc}
f(x,y,z) & = & f(x_i,y_j,z_k) & (1-r_x) & (1-r_y) & (1-r_z) \\
 & + & f(x_{i+1},y_j,z_k) & r_x & (1-r_y) & (1-r_z) \\
 & + & f(x_i,y_{j+1},z_k) & (1-r_x) & r_y & (1-r_z) \\
 & + & f(x_{i+1},y_{j+1},z_k) & r_x & r_y & (1-r_z) \\
 & + & f(x_i,y_j,z_{k+1}) & (1-r_x) & (1-r_y) & r_z \\
 & + & f(x_{i+1},y_j,z_{k+1}) & r_x & (1-r_y) & r_z \\
 & + & f(x_i,y_{j+1},z_{k+1}) & (1-r_x) & r_y & r_z \\
 & + & f(x_{i+1},y_{j+1},z_{k+1}) & r_x & r_y & r_z
\end{array}
\tag{4.17}
$$

with:

$$
r_x = \frac{x - x_i}{x_{i+1} - x_i} \qquad r_y = \frac{y - y_j}{y_{j+1} - y_j} \qquad r_z = \frac{z - z_k}{z_{k+1} - z_k}
$$

and where (i,j,k), $(i+1,j,k)$, $(i,j+1,k)$, $(i+1,j+1,k)$, (i,j,k), $(i+1,j,k+1)$, $(i,j+1,k+1)$, $(i+1,j+1,k+1)$ are the 8 3-tuples corresponding to the vertices of the parallelipidic cell in wich lies the point (x,y,z). It can be more compactly rewritten as:

$$
f(x,y,z) = \sum_{\alpha=1}^{8} c_\alpha^i(x,y,x) f_\alpha \tag{4.18}
$$

where the functions c_α are the interpolating functions attached at each vertice, independently of the function f. These functions c_α take different expressions in the different cells admitting the considered node as vertex, but they can be continuously defined over the reunion of these (generally) 8 cells.

Then, given the geometry of the rays, the wave travel times (4.2) (4.3) may be easily computed by interpolating the slowness which can be evaluated at each node from the values of v_{P} and $v_{\mathrm{P}}/v_{\mathrm{S}}$. For example, in the case of P-waves, it yields:

$$
T_{\mathrm{P}} = \sum_{i=1}^{n_c} \sum_{\alpha=1}^{8} \frac{1}{(v_{\mathrm{P}})_{i,\alpha}} \int\limits_{\mathcal{R}_{\mathrm{P}}^i} c_\alpha(s) \mathrm{d}s \tag{4.19}
$$

where n_c is the number of cells and $\mathcal{R}_{\mathrm{p}}^i$ is the intersection of the ray with the ith cell, and c_α is the interpolating function attached to the vertex $\alpha(i)$. But, as the model vector is indexed by the grid nodes rather than the cells, the expression may be rewritten as:

$$
T_{\mathrm{P}} = \sum_{i=1}^{n_n} \frac{1}{v_{\mathrm{P}\,i}} \int\limits_{R_{\mathrm{P}}^i} c_i(s) \mathrm{d}s \tag{4.20}
$$

where n_n is the number of nodes, $\mathcal{R}_{\mathrm{p}}^i$ is, this time, the intersection of the ray with the reunion of the cells having the ith node as vertex, and where c_i is the interpolation function attached to this node.

In the same way, we obtain that the travel time difference (4.3) may be computed through:

$$\Delta T = \sum_{i=1}^{n_n} \frac{1}{v_{\mathrm{P}\,i}} \left[\left(\frac{v_\mathrm{P}}{v_\mathrm{S}}\right)_i \int\limits_{\mathcal{R}_\mathrm{S}^i} c_i(s)\mathrm{d}s - \int\limits_{\mathcal{R}_\mathrm{P}^i} c_i(s)\mathrm{d}s \right] \tag{4.21}$$

where $\mathcal{R}_\mathrm{P}^i$, $\mathcal{R}_\mathrm{S}^i$, and c_i are repectively the part of the rays and the interpolation functions attached to the ith node.

From a practical perspective, the model $\mathbf{m}$ is now represented by a vector in $\mathbb{R}^{2n_n+4n_e+2n_s}$:

$$\mathbf{m} = \begin{bmatrix} (v_\mathrm{P})_i \\ (v_\mathrm{P}/v_\mathrm{S})_i \\ \mathbf{x}_i \\ t_i^0 \\ \Delta_i^\mathrm{P} \\ \Delta_i^\mathrm{S-P} \end{bmatrix} \quad \begin{array}{c} 1 \leq i \leq n_n \\ n_n + 1 \leq i \leq 2n_n \\ 2n_n + 1 \leq i \leq 2n_n + 3n_e \\ 2n_n + 3n_e + 1 \leq i \leq 2n_n + 4n_e \\ 2n_n + 4n_e + 1 \leq i \leq 2n_n + 4n_e + n_s \\ 2n_n + 4n_e + n_s + 1 \leq i \leq 2n_n + 4n_e + 2n_s \end{array} \tag{4.22}$$

where n_n is the number of nodes of the inversion grid, n_e is the number of seismic events in the box and n_s the number of stations.

Actually, the code allows for defining a sub-grid of the reference grid as inversion grid, in the case when we want only to infer the structure in a specific area, while taking into account the seismic events over a larger domain. In that case, the propagating times and the rays may be computed in the whole domain through the interpolation of the values of the reference grid, but these values are modified by the inversion only inside the inversion box, and all the events, inside or outside the inversion box, may be relocated by the inversion process. The code also allows for introducing the Moho discontinuity, and decoupling the crust from the mantle. In this case the reference grid and the corresponding velocity values are duplicated, the ones corresponding to to the crust, down to the bottom of the box, the other ones to the mantle, up to the top of the box. The inversion is performed on both the boxes at the same time, but for sure, the travel times, the rays and the derivatives are computed by considering the values of the crust grid above the moho discontinuity, and the values of the mantle grid below the Moho discontinuity.

4.3.2 Discretization of the derivative operator $\mathbf{G}_m$

The discretization of the derivative operator could be done as above for the direct travel time computation involved in the expression of the mapping g, by starting directly from (4.4). But for consistency with the evaluation (4.20) of T_P and (4.21) of ΔT, we replace in the expressions (4.4), the derivative of T_P and ΔT with respect to the fields v_P and $v_\mathrm{P}/v_\mathrm{S}$ by the derivatives of the dicretized expressions (4.20) and (4.21). This leads to the following expressions of the coefficients of a raw of the matrix $\mathbf{G}_m$ (Monteiller, 2005):

- for the ith t_P datum :

$$\frac{\partial t_P}{\partial v_{P_j}} = -\frac{1}{v_{P_j}^2}\int_{\mathcal{R}_P^j} c_j(s)\mathrm{d}s \qquad \frac{\partial t_P}{\partial (v_\mathrm{P}/v_\mathrm{S})_j} = 0 \tag{4.23}$$

for $1 \leq j \leq n_n$,

$$\frac{\partial t_P}{\partial x_\alpha} = -(\tau_{\ell(i)}^\mathrm{P})_\alpha = -\frac{1}{v_\mathrm{P}(\mathbf{x})} u_\alpha^\mathrm{P} \tag{4.24}$$

in the column corresponding to the location $\mathbf{x} = (x_\alpha)_{\alpha=1,2,3}$ of the seismic event $\ell(i)$ corresponding to the P arrival time, and where u_α^P are the components of the unit vector tangent to the starting ray .

$$\frac{\partial t_P}{\partial t_0} = 1 \tag{4.25}$$

in the column corresponding to the initial time t_0 of the corresponding $\ell(i)$th event ,

$$\frac{\partial t_P}{\partial \Delta^P} = 1 \qquad \frac{\partial t_\mathrm{P}}{\partial \Delta^{\mathrm{S-P}}} = 0 \tag{4.26}$$

in the column corresponding to the time delays attached to the $k(i)$th station of the P datum,

- for the ith raw ($t_\mathrm{S} - t_\mathrm{P}$ delay):

$$\begin{aligned} \frac{\partial (t_\mathrm{S} - t_\mathrm{P})}{\partial v_{\mathrm{P}_j}} &= \frac{1}{v_{\mathrm{P}_j}^2}\int_{\mathcal{R}_\mathrm{P}^j} c_j(s)\mathrm{d}s - \frac{1}{v_{\mathrm{P}_j} v_{\mathrm{S}_j}}\int_{\mathcal{R}_\mathrm{S}^j} c_j(s)\mathrm{d}s \\ \frac{\partial (t_\mathrm{S} - t_\mathrm{P})}{\partial v_\mathrm{P}/v_\mathrm{S}} &= \frac{1}{v_{\mathrm{P}_j}}\int_{\mathcal{R}_\mathrm{S}^j} c_j(s)\mathrm{d}s \end{aligned} \tag{4.27}$$

for $1 \leq j \leq n_n$,

$$\frac{\partial (t_\mathrm{S} - t_\mathrm{P})}{\partial x_j} = (\tau_{\ell(i)}^\mathrm{P})_\alpha - (\tau_{\ell(i)}^\mathrm{S})_\alpha = \frac{1}{v_\mathrm{P}(\mathbf{x})} u_j^\mathrm{P} - \frac{1}{v_\mathrm{S}(\mathbf{x})} u_j^\mathrm{S} \tag{4.28}$$

in the columns corresponding to the locations $\mathbf{x} = (x_\alpha)_{\alpha=1,2,3}$ of the seismic event $\ell(i)$ corresponding to the S − P datum, and where u_α^P and u_α^S are respectively the components of the unit vector tangent to the starting P and S ray,

$$\frac{\partial(t_{\mathrm{S}}-t_{\mathrm{P}})}{\partial t_0}=0 \tag{4.29}$$

in the column corresponding to the initial time t_0 of the corresponding $\ell(i)$th event,

$$\frac{\partial(t_{\mathrm{S}}-t_{\mathrm{P}})}{\partial\Delta^{P}}=0 \qquad \frac{\partial(t_{\mathrm{S}}-t_{\mathrm{P}})}{\partial\Delta^{S-P}}=1 \tag{4.30}$$

in the column corresponding to the time delays attached to the $k(i)$th station of the S-P datum.

4.3.3 The discretization of the covariance operator

From a practical point of view, the discretization of the covariance operator for the velocity fields is made in considering $\mathbf{Cor}^{-1/2}$ (2.59), because this operator is involved in the cost function $E(\mathbf{m})$ (4.12), (2.64). The inspection of the regularisation term of $E(\mathbf{m})$ (2.64) yields :

$$\|\mathbf{Cor}^{-1/2}\Sigma^{-1}(\mathbf{m}-\mathbf{m}_{prior})\|^2=\frac{1}{8\pi\xi_1\xi_2\xi_3}\int\limits_V\left[(\mathbf{I}-\Delta_\xi)\left(\frac{\mathbf{m}-\mathbf{m}_{prior}}{\sigma}\right)\right]^2\mathbf{d}V \tag{4.31}$$

$$=\frac{1}{8\pi}\prod_{\alpha=1}^{3}\left(\frac{h_\alpha}{\xi_\alpha}\right)\sum_{i=1}^{n_n}\left[(\mathbf{I}-\Delta_\xi)\left(\frac{\mathbf{m}-\mathbf{m}_{prior}}{\sigma}\right)\right]_i^2 \tag{4.32}$$

where h_α is the grid length in the αth space direction. This shows that the σ values of the fields must be renormalized by the square root of the volume element $h_1h_2h_3$:

$$(\sigma_{eff})_i=\frac{\sigma_i}{\sqrt{h_1h_2h_3}}=\prod_{\alpha=1}^{3}\left(\frac{\xi_0}{\xi_\alpha h_\alpha}\right)^{1/2}(\sigma_{phy})_i$$

in taking account of the renormalization (4.11) as functions of the smoothing lengths ξ_1,ξ_2,ξ_3. This finally yields:

$$(\sigma_{eff})_i=\mu(\sigma_{phy})_i \quad \text{with } \mu=\prod_{\alpha=1}^{3}\left(\frac{\xi_0}{\xi_\alpha h_\alpha}\right)^{1/2} \tag{4.33}$$

where μ is a parameter for the tuning of the damping of the fields v_{P} and $v_{\mathrm{P}}/v_{\mathrm{S}}$.

The discretization of the operator $\mathbf{I}-\Delta_\xi=\mathbf{I}-\sum_{j=1}^{3}\xi_j^2\partial_j^2$, which is involved in the expression (2.59) of $\mathbf{Cor}^{-1/2}$, can be simply obtained by finite differences. It yields the following expression when applied to a generic function f:

$$
\begin{aligned}
[(\mathbf{I}-\Delta_\xi)(f)]_{i,j,k} = f_{i,j,k} &- \left(\frac{\xi_1}{h_1}\right)^2 (f_{i+1,j,k} - 2f_{i,j,k} + f_{i-1,j,k}) \\
&- \left(\frac{\xi_2}{h_2}\right)^2 (f_{i,j+1,k} - 2f_{i,j,k} + f_{i,j-1,k}) \\
&- \left(\frac{\xi_3}{h_3}\right)^2 (f_{i,j,k+1} - 2f_{i,j,k} + f_{i,j,k-1})
\end{aligned}
\tag{4.34}
$$

The global covariance matrix is thus block diagonal. Concerning the block $\mathbf{C}_m^{-1/2}$ corresponding to the scalar parameters, as they are assumed to be independent with each other, it is purely diagonal. Consequently:

$$
\left(\mathrm{Cor}^{-1/2}\Sigma^{-1}\right)_{i,i} = 1/\sigma_i \quad 2n_n + 3n_e + 1 \leq i \leq 2n_n + 4n_e + 2n_s
$$

where σ_i is the physical standard deviation of the parameter, except for the location of the events, for which we introduce a tuning parameter λ such that:

$$
\left(\mathrm{Cor}^{-1/2}\Sigma^{-1}\right)_{i,i} = \lambda/\sigma_i \quad (2n_n + 1 \leq i \leq 2n_n + 3n_e)
$$

and which, as μ or ξ_0 for the velocity fields, controls the damping in the inversion process.

4.3.4 Implementation of the quasi Newton algorithm

The quasi-Newton algorithm (4.16) that is written as:

$$
\mathbf{m}_{k+1} - \mathbf{m}_k = -(\mathbf{C}_m^{-1} + \mathbf{G}_k^*\mathbf{C}_d^{-1}\mathbf{G}_k)^{-1}(\mathbf{G}_k^*\mathbf{C}_d^{-1}(\mathbf{g}(\mathbf{m}_k) - \mathbf{d}_{obs}) + \mathbf{C}_m^{-1}(\mathbf{m}_k - \mathbf{m}_{prior}))
$$

requires the solution of a linear system in the model space. Following (Monteiller et al., 2005), it is performed by using the LSQR algorithm (Paige and Saunders, 1982). This in turn requires rewriting each iteration in normal form by using the decomposition (2.60), (2.61) of the covariance operator:

$$
\mathbf{m}_{k+1} - \mathbf{m}_k = (\mathbf{A}_k^*\mathbf{A}_k)^{-1}\mathbf{A}_k^*\mathbf{v}_k \tag{4.35}
$$

with:

$$
\mathbf{A}_k = \begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{G}_k \\ \mathbf{Cor}^{-1/2}\mathbf{\Sigma}^{-1} \end{bmatrix} \quad \mathbf{v}_k = \begin{bmatrix} \mathbf{C}_d^{-1/2}(g(\mathbf{m}_k) - \mathbf{d}_{obs}) \\ \mathbf{Cor}^{-1/2}\mathbf{\Sigma}^{-1}(\mathbf{m}_k - \mathbf{m}_{prior}) \end{bmatrix} \quad \text{and} \quad \mathbf{C}_m = \mathbf{\Sigma}\,\mathbf{Cor}\,\mathbf{\Sigma}
$$

The linear system solved by the LSQR code thus takes the form:

$$\mathbf{A}^*\mathbf{A}\mathbf{x} = \mathbf{A}^*\mathbf{v} \tag{4.36}$$

It is numerically uneasy since the orders of magnitude of the various parameters strongly differ, and consequently the system is relatively ill-conditioned. To improve the conditoning, the matrix $\mathbf{A}$ is decomposed as $\mathbf{A} = \mathbf{A}'\,\mathbf{\Lambda}$ where $\mathbf{\Lambda}$ is a matrix diagonal. Thus (4.4.35) becomes:

$$\mathbf{A}'^*\,\mathbf{A}'\,\mathbf{\Lambda}\mathbf{x} = \mathbf{A}^*\mathbf{v} \tag{4.37}$$

It amounts to replacing $\mathbf{A}$ by $\mathbf{A}'$ as input into the LSQR code, and to applying the diagonal matrix $\mathbf{\Lambda}^{-1}$ to the returned vector.

The INSIGHT code allows two possible choices for the matrix $\mathbf{\Lambda}$. The first choise consists in putting:

$$\Lambda_{ii} = \sqrt{\sum_{k=1}^{n} \mathbf{A}_{ki}^2} \tag{4.38}$$

in analogy with the computation of the standard deviations of a covariance matrix, for the positive-definte matrix $\mathbf{A}^*\mathbf{A}$. The second one consists in putting:

$$\Lambda_{ii} = \max_{j\in\mathcal{P}_i}\left(\sqrt{\sum_{k=1}^{n} \mathbf{A}_{kj}^2}\right) \tag{4.39}$$

where $\mathcal{P}_i$ is a subset of indices, which the index i belongs to, and that corresponds to a same family of physical parameters (velocity, events, or time delays parameters). For the tomography of Ecuador-Northern-Peru that we have conducted, we have used this second block pre-conditioning.

The INSIGHT code also allows the inversion of differences in data, either corresponding to a same event at different stations, or to a same station for different events (in that case the so called 'double differences', Waldhauser and Ellsworth, 2000). The prior data must be assumed to be Gaussian. We have seen in paragraph 2.4.2 that the only modification for the inversion consists in replacing the inverse of the covariance operator $\mathbf{C}_d^{-1}$ by the operator (2.40):

$$\mathbf{C}_d^{-1/2}\mathbf{P}\mathbf{C}_d^{-1/2} = (\mathbf{P}\mathbf{C}_d^{-1/2})^*\,\mathbf{P}\mathbf{C}_d^{-1/2}$$

for the block of corresponding data, *i.e.* the set of arrival times of a same phase corresponding to an event, or the set of events for which a phase of a same kind has been recorded at a same station. In this expression the vector $\mathbf{u}$ verifies $\mathrm{u}_i = 1$ or 0 according to the ith datum belongs to this set or not. The only resulting modification in the normal decomposition (4.34) consists in replacing $\mathbf{C}_d^{-1/2}$ by $\mathbf{P}\mathbf{C}_d^{-1/2}$ (2.40). In that case, the first pre-conditionning column by column must be used.

4.4 Implementing the INSIGHT code for the tomography of Ecuador

As described in chapter 3 we have considered two domains, one including Northern Peru, the other one centered on Ecuador. In both cases the size of the inversion cells is 5×5 $km \times km$ horizontally in the box and 2 km in the vertical direction. The size of the cubic Podvin-Lecomte mesh is 1 km.

The values assigned to the physical standard deviations of the fields v_P and $v_\mathrm{P}/v_\mathrm{S}$ are $\sigma_\mathrm{P} = 750\ m/s$ and $\sigma_{v_\mathrm{P}/v_\mathrm{S}} = 0.15$ that represent relative uncertainties of about 10% of the a priori model, a little more for the P velocity in the crust, a little less in the mantle. The corresponding reference length ξ_0, which may also be thought of as a damping parameter for these fields, is of the order of a few kilometers.

Concerning the standard deviation of the location of events, the assigned values to the horizontal components are $\sigma_x = \sigma_y = 30km$ to let a great degree of freedom in longitude and latitude. The vertical standard deviation σ_z depends on the depth h of the event with respect to the topographic surface according to:

$$\sigma_z = \begin{cases} \sigma_{min} + (\sigma_{max} - \sigma_{min}) \left[a \left(\frac{h}{h_c}\right)^3 + b \left(\frac{h}{h_c}\right)^2 + c \frac{h}{h_c} \right] & \text{for} \quad 0 \le h \le h_c \\ \\ \sigma_{max} & \text{for} \quad h > h_c \end{cases} \tag{4.40}$$

with:

$$a = 8p - 6 \qquad b = -16p + 11 \qquad c = 8p - 4 \qquad \text{with:} \quad \frac{1}{2} \le p \le \frac{7}{8}$$

We can remark that this function verifies:

$$\sigma_z(0) = \sigma_\mathrm{min} \ , \quad \sigma_z\left(\frac{h_c}{2}\right) = (1-p)\,\sigma_\mathrm{min} + p\,\sigma_\mathrm{max} \ , \quad \sigma_z(h_c) = \sigma_\mathrm{max} \ , \quad \sigma_z'(h_c) = 0$$

and is increasing from σ_min to σ_max for $1/2 \le p \le 7/8$. For $p = 1/2$:

$$\sigma_z'(0) = 0 \ , \quad \sigma_z\left(\frac{h_c}{2}\right) = \frac{\sigma_\mathrm{min} + \sigma_\mathrm{max}}{2}$$

and the point $\left(\frac{h_c}{2}, \frac{\sigma_\mathrm{min}+\sigma_\mathrm{max}}{2}\right)$ is a center of symmetry of the graph. For $p = 7/8$, $\sigma_z''(h_c) = 0$.

This function accounts for the simple fact that the events a priori lie within the Earth. The non-null value of the function at the surface allows for a minimal degree of freedom for the depth of the events a priori located at very shallow depth. Moreover in the INSIGHT code, the events that jump over the surface during an iteration are put back to the surface while keeping the same a priori location. The values chosen for the parameters of the function are: $\sigma_{min} = 0.5\ km$, $\sigma_{max} = 30\ km$, $h_c = 30\ km$ and $p = 0.6$.

Concerning the initial times t_0 of the seismic events the value set for the standard deviation is $\sigma_{t_0} = 1000s$ to let a complete degree of freedom.

For the time delays attached to the stations, the standard deviations depend on the number of recorded data at the station. For the stations having recorded a small number of data this value must be very small in order to avoid any bias in the inversion. On the contrary, for stations having recorded a large number of data we may expect that this delay is correctly averaged, and consequently we may assign a greater standard deviation to it. The form chosen is:

$$\sigma = \sigma_{max} \sqrt{\frac{\min(n, n_c)}{n_c}} \tag{4.41}$$

where n is the number of P or S data, depending on the type of delay considered, recorded at the the station, n_c is a limit number of data, which is set at 20 000, and σ_{max} is the maximum value of the standard deviation which is reached for $n \geq n_c$, and set at $5 \cdot 10^{-2}s$. (in taking into account that the number of data S-P is actually equivalent to that of data S for most of the catalogues). For the P data, n ranges between 1 and 37 226 and consequently σ between 3.54 $10^{-4}s$ and 5 $10^{-2}s$ according to the station. For the S (or S-P) n ranges between 1 and 13 183 and σ between 3.54 $10^{-4}s$ and 4.1 $10^{-2}s$.

4.4.1 Iterations and indicators of convergence

We have performed numerous tomographies with various datasets, and differents damping and smoothing parameters in order to obtain realistic models. At each iteration, the data with a misfit greater than 20 standard deviations or 3 seconds are temporarly eliminited; they may be put back in the dataset at a next iteration according to their misfits. The events may also be localized outside the box during an iteration. If it is above the surface, as already indicated in the previous paragraph, they are relocated at the surface, while keeping the same a priori. If it is sideways or below the bottom of the box, up to three times they are then relocated at their prior location. If these 3 relocations are not sufficient enough to stabilize them, they are definitely eliminated.

To follow the convergence of the algorithm, the code computes the following indicators that are displayed in figure 4.2:

- RMS of the filtered data misfit in hyperbolic secant variable
- RMS of the normalized filtered misfit of the arrival times (filtered and unfiltered)
- RMS of the arrival times in seconds;
- Number of effective data for the iteration (unfiltered data);
- Effective number of earthquakes taken into account for the iteration

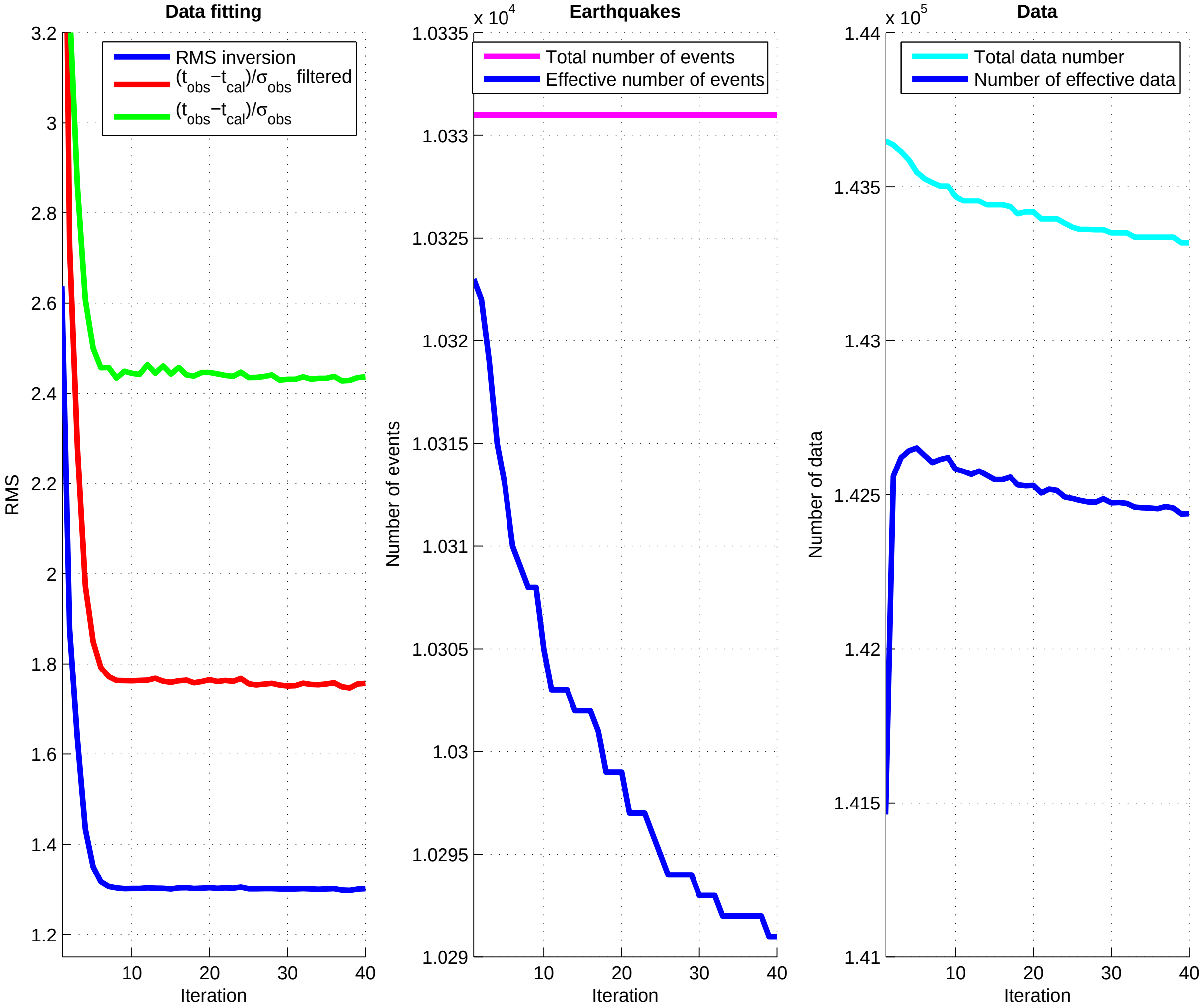

Figure 4.2: *Exemple of sequence of values of the inversion indicators throughout the iterations of the tomography process. On the left: convergence of the RMS of the normalized data adjustment (unfiltered in green, filtered in red, in Sech variables (4.6) in blue); in the middle: effective number of events throughout the iterations; on the right: number of data corresponding to the remaining events, before (light blue) and after (blue) filtering according to their misfits (20 standard deviation or 3 seconds), throughout the iterations.*

The code computes also the L_1, L_2 and L_∞ norms of the difference between the current v_P and v_P/v_S models with the prior ones, as well as the median, mean and RMS of the distances between the current and prior locations of the earthquakes (figure 4.3).

These indicators allow checking the convergence and the determination of the tuning parameters for which the process is stable. A representative example of convergence of these indicators of data fitting and of behaviour of model is displayed in figures (4.2), (4.3). It concerns a set of 143 785 data (112 931 P and 30 854 S) for a subset of 10 332 events well localized in the box centered on Ecuador. Generally, 10 iterations are sufficient to obtain a good fit of the data, but a few dozen of iterations are needed to obtain a negligible increment in the model at each step.

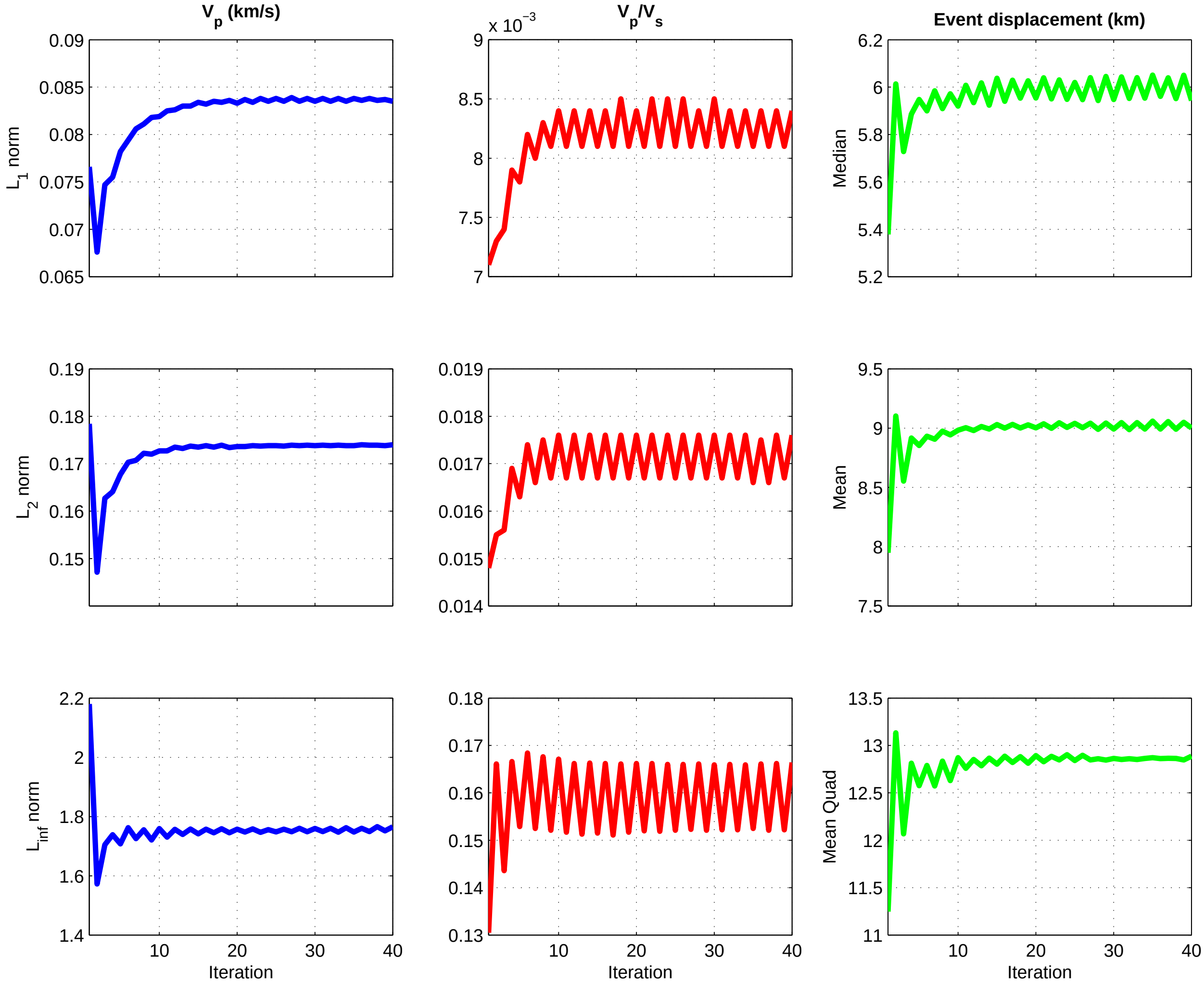

Figure 4.3: *Convergence of the norms L_1, L_2 and L_∞ of the perturbation of the model with respect to the a priori, throughout the iterations. On the left for the V_P fields; in the middle, for the V_P/V_S field; on the right, median, mean and RMS distances of the hypocenters to their a priori locations.*

All the computations were performed on the FROGGY cluster of the centre of *Calcul Intensif, Modélisation, Expérimentation Numérique et Technologique* (CIMENT) at Grenoble University. The parallelization of the INSIGHT code over 64 cores allows an iteration involving up to 850 000 data to be run in approximatively one hour.

4.4.2 Tuning parameters and L-curves

As detailed in the paragraph 4.3.3, several tuning parameters are available to control the regularization of the inversion process: the 3 correlation lengths ξ_1, ξ_2, ξ_3 for the smoothing of the velovity fields and 2 damping parameters, λ for the event localization and (4.32):

$$\mu = \prod_{\alpha=1}^{3} \left(\frac{\xi_0}{\xi_\alpha h_\alpha} \right)^{1/2}$$

for the velocity fields. From a practical perspective, once the smoothing lengths has been chosen, the change of the value of μ amounts to the change of that of ξ_0. On the contrary, if the value of ξ_0 has been fixed, then a change in the smoothing lengths varies the damping μ in such a way that the order of magnitude of the resulting velocity fields is globaly the same. For most of the inversion, we don't consider the moho discontinuity, thus we have only one set of parameters to fix. After several tests, we chose to put $\xi_V = 20\ km$ and $\xi_1 = \xi_2 = \xi_H$ for most of the runs.

In the case of the largest study domain (box) that contains Northern Peru, we found by varying the value of the location damping λ for different sets of reasonable values assigned to the other tuning parameters, that an efficient (and constant with respect to the set of other chosen values) value for the damping λ is 0.7. This leaves two unassigned tuning parameters: the horizontal correlation length ξ_H and the reference length ξ_0 (in the order of magnitude of a few kilometers) for the smoothing and damping of the velocity fiels.

To illustrate how ξ_H acts on the resulting model, the figure 4.4 displays the East-West cross-section, at approximately the latitude of Quito, of 4 models of P-velocity resulting of a common value of $\xi_0 = 6\ km$ and of 4 different values of ξ_H (30, 35, 50, 100 km). The dramatic change in thickness of the crust at the west of the Andean range, marking the transition with the coastal area, is all the more smoothed out that the horizontal smoothing length is increased, while the global order of magnitude of the velocity is maintained. On the other hand, if we now keep unchanged the value of $\xi_H = 35\ km$ and vary the value of the velocity damping ($\xi_0 = 2, 4, 8, 10\ km$), we obtain 4 new cross-sections, displayed in figure 4.5, that can be compared to the one on the top right of the figure 4.4 corresponding to $\xi_0 = 6\ km$ for the same $\xi_H = 35\ km$. It shows that lower values of ξ_0 reduce the oscillation pattern, whereas greater ones induce oscillations of large amplitude in the model. Taking much greater value of ξ_0 would make the inverse process become unstable. We can also remark a relative correlation between the effects of a small value of ξ_0 and those of a great length of smoothing ξ_H on the resulting model.

A useful tool to determine reasonable values of the tuning parameters consists in representing the various models in the diagram of data adjustment versus norm of the incremental model with respect to the prior one (figure 4.6). For a given value of ξ_H we obtain classical L-curves (Hansen, 1992) for the models obtained by varying ξ_0 (dots with the same color on the left or in the middle of figure 4.6). For small values of ξ_0, the adjustement of the data is poor, whereas the model is smooth and close to the prior one. On the contrary, for large values of ξ_0 the adjustment of the data is better, while the model is less smooth. The figure (4.6) also shows that for a given value of the damping parameter ξ_0, the adjustment of the data increases when the smoothing length decreases, while the norm of the model keeps the same order of magnitude (in virtue of the renormalization of the damping by the smoothing described in paragraph 2.5.3).

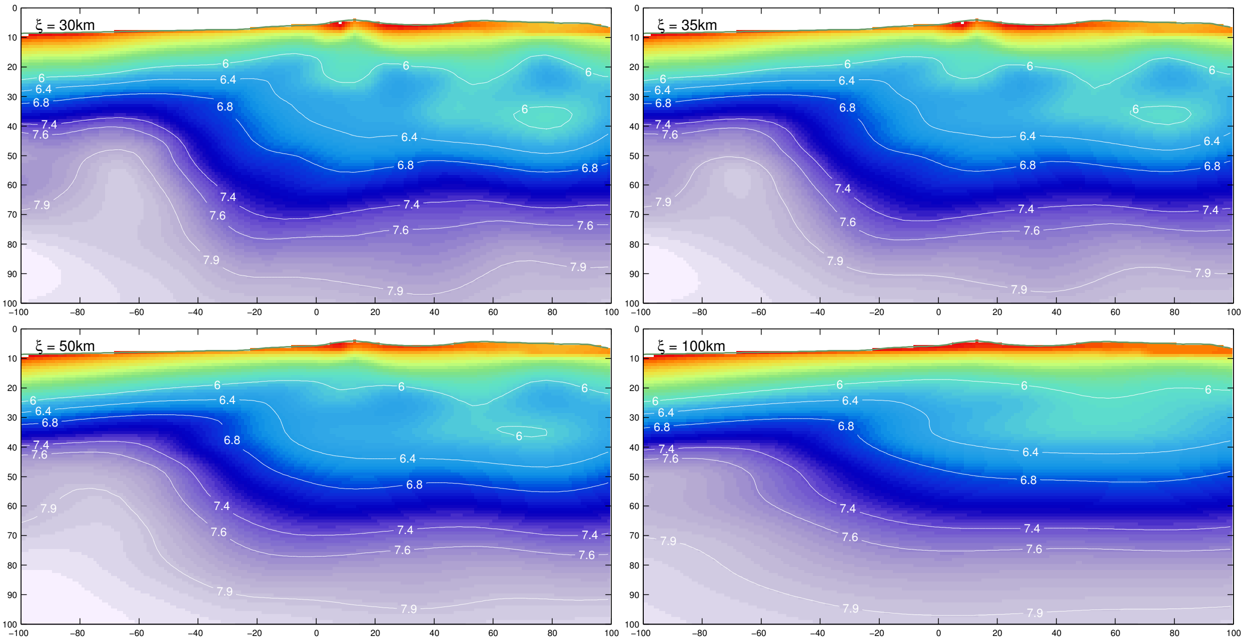

Figure 4.4: *West-East cross sections at the latitude of Quito of the P velocity of 4 models obtained with different horizontal smoothing lengths ξ_H, and the same damping parameter $\xi_0 = 6$ km (and ξ_V = 20 km). On the top left $\xi_H = 30$ km, on the top right $\xi_H = 35$ km, on the bottom left $\xi_H = 50$ km, and on the bottom right $\xi_H = 100$ km.*

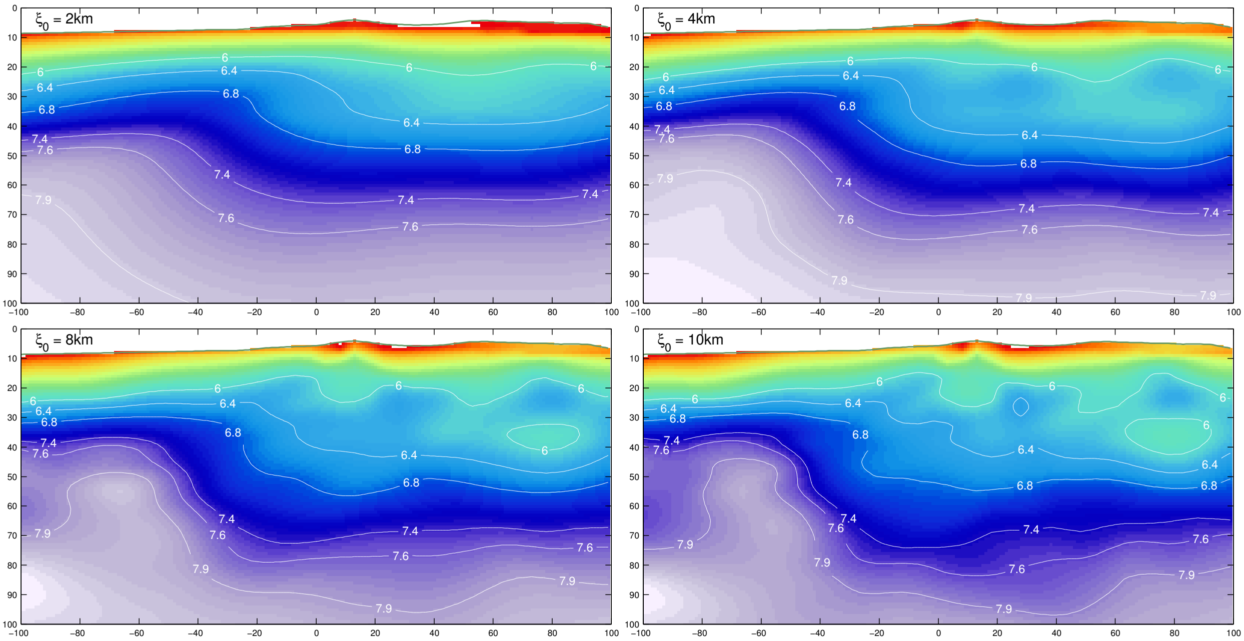

Figure 4.5: *West-East cross sections at the same latitude of Quito as figure 4.4 of the P velocity of 4 models obtained with different damping parameter $\xi_0 = 6$ km, and the same horizontal smoothing lengths ξ_H = 35 km (and ξ_V = 20 km). On the top left $\xi_0 = 2$ km, on the top right $\xi_0 = 4$ km, on the bottom left $\xi_0 = 8$ km, and on the bottom right $\xi_0 = 10$ km.*

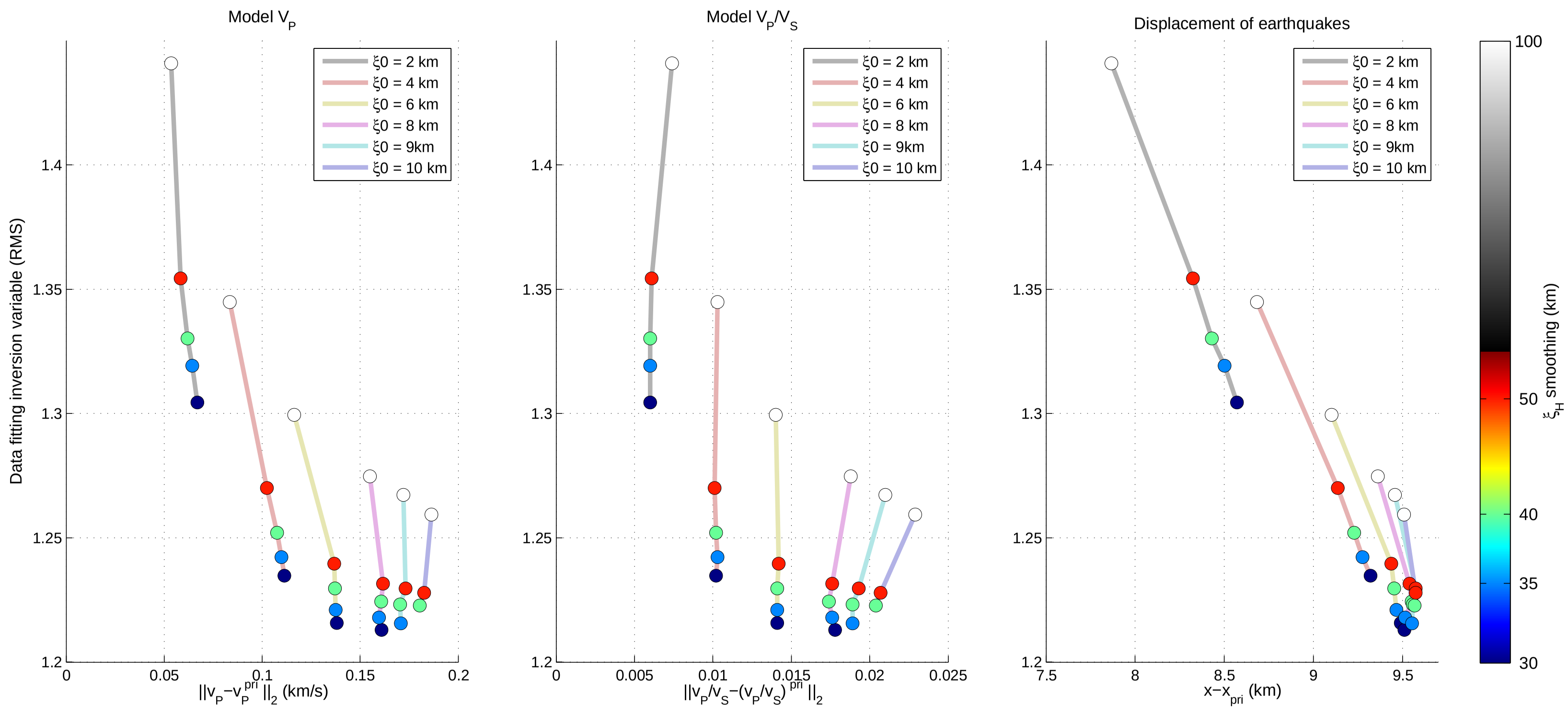

Figure 4.6: *diagram of data adjustment versus norm of incremental model with respect to the prior one in the case of the largest box with a subset of 143 785 data corresponding to the most accurately events. The data adjustment is the RMS in Sech variable (eqn. 4.6.) The norm is (left) the* L_2 *norm of the P velocity, (middle) of the ratio* v_P/v_S*, (right) the mean distance of the epicenter to the prior location. Each model, which corresponds to a pair of values* ξ_0, ξ_H*, is represented by a dot, the color of which identified the value of* ξ_H *according to the scale on the right side. The color of each line gives the value of* ξ_0 *for all the models (dots) that it links according to the legend of the figures. In other words, the colored lines are the trajectories of models with a given* xi_0 *and varying* ξ_H*; the vitual lines linking the dots of same color are the L-curves corresponding to a value of* ξ_H *(the color) and a varying* ξ_0

Thus the diagram allows the determination of a good trade off between the data adjustment and the smoothness of the model. A reasonable choice in this example consists in choosing $\xi_0 = 6\ km$ and $\xi_h = 30$ or $35\ km$, or values of these parameters yielding a model close to these points in the diagram.

4.5 Evaluation of the uncertainty in the model, subsets of data

The more common way to test the resolving power of the dataset in imaging study is to perform a checkboarder test, though it is not infallible (Leveque et al., 1993). Another simple approach consists in delimiting the region where the intersection of rays is sufficiently dense. A more sophisticated way consists in evaluating the restitution (or averaging) index introduced in the paragraph (2.5.3.1) and which provides an efficient and fast way to delimit the volume outside which the resolution is surely very poor. We can also theoretically evaluate the kernel of resolution at each point of the domain, but since it involves an amount of computation at least equivalent to that of an iteration, it may only be envisioned for specific points of interest. But, for sure, the improvement in data adjustment must also be taken into account when evuating the model (or data !) quality.

4.5.1 Adjustment of arrival times

The adjustments of the arrival times t_P and t_S, before and after the tomography for the whole set of data and the events within the box centered on Ecuador, is displayed in figure (4.7). The a priori adjustments are obtained with the prior model and the LOCIN code described in chapter 2. The data set consists of 641 036 P and 215 134 S arrival times, and corresponds to 59 883 events localized by the code LOCIN within the box.

The histogram of prior S adjustments is very asymmetric, with a mean shift toward the negative values. This shift can be explained by the fact that the initial times are determined by the LOCIN code only from the P-arrival times, independently of the S ones, and is completely rectified by the tomography. The adjustment of the P arrival times is globaly good with about 31% smaller than 0.05 s, 63% smaller than 0.15 s and 77% smaller than 0.25 s. For the S-arrival times the global fit is far from being so good, with 31% smaller than 0.15 s, 47% smaller than 0.25 s and 54% smaller than 0.35 s.

4.5.2 Filtering a posteriori the events and new data sets

Due to the great heterogeneity of the spatial repartition of the stations, many event localizations are inaccurate, even those resulting from the tomography. We thus defined a geometrical filter to eliminate the poorly localized events. It consists in imposing that

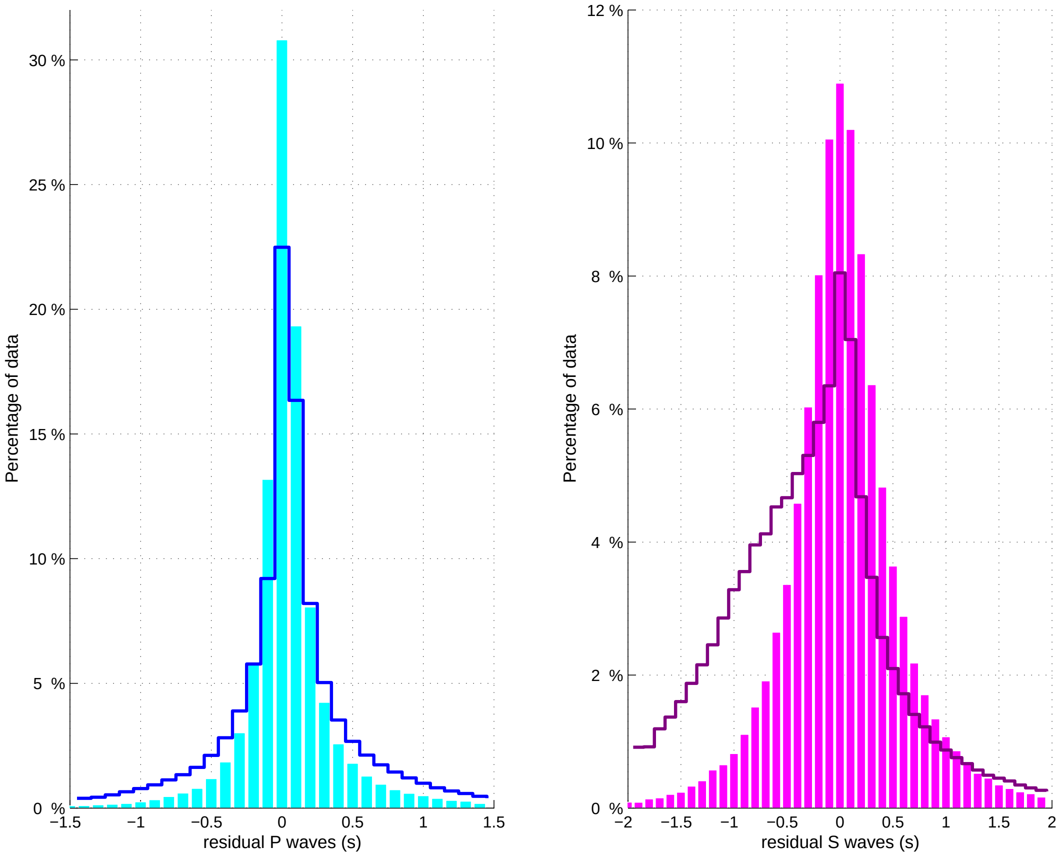

Figure 4.7: *Histogram of the adjustment of arrival times for P waves (on le left) and S waves (on the right) after removing the misfits greater than 3 seconds or 20 standard deviations considered as correponding to outliers. The continuous lines for the prior adjustments and the bar charts for the adjustments after tomography.*

among at least 4 data better fitted than k times their standard deviations, there exist 3 data for which the angles between the associated rays at the event location are all greater than α, and for which at least one of these angles is greater than β. We have performed many tests to define the values of k and of the critical angles α and β allowing the elimination of obviously poorly localized events while keeping an acceptable number of events for the two boxes and sets of data we have considered.

For the largest box including Northern Peru and after having removed the data recorded on volcanic edifices, the data set consists of 345 052 P and 118 801 S data corresponding to 44 317 events. Setting $k = 1.25$, $\alpha = 21^o$ and $\beta = 25^o$ yields a new set consisting of 112 931 P and 30 854 S data associated to 10 332 events. This latter dataset has been used for figures 4.4, 4.5, 4.6.

For the smallest box centered on Ecuador, with all the data available until April 2016, the prior set consists of 641 036 P and 215 134 S data corresponding to 58 060 events. The histograms displayed in figure 4.7 correspond to this whole data set. Setting $k = 1.4$, $\alpha = 17^o$ and $\beta = 20^o$ yields a new set consisting of 336 012 P and 111 536 S data associated to 25 462 events. The figure (4.8), which displays the resulting sismicity located in the vicinity of an East-West cross-section at about 1.5^{o}S, exemplifies the effect of the filter in this case. We can observe that most of the obviously poorly localized hypocenters are removed by the filter.

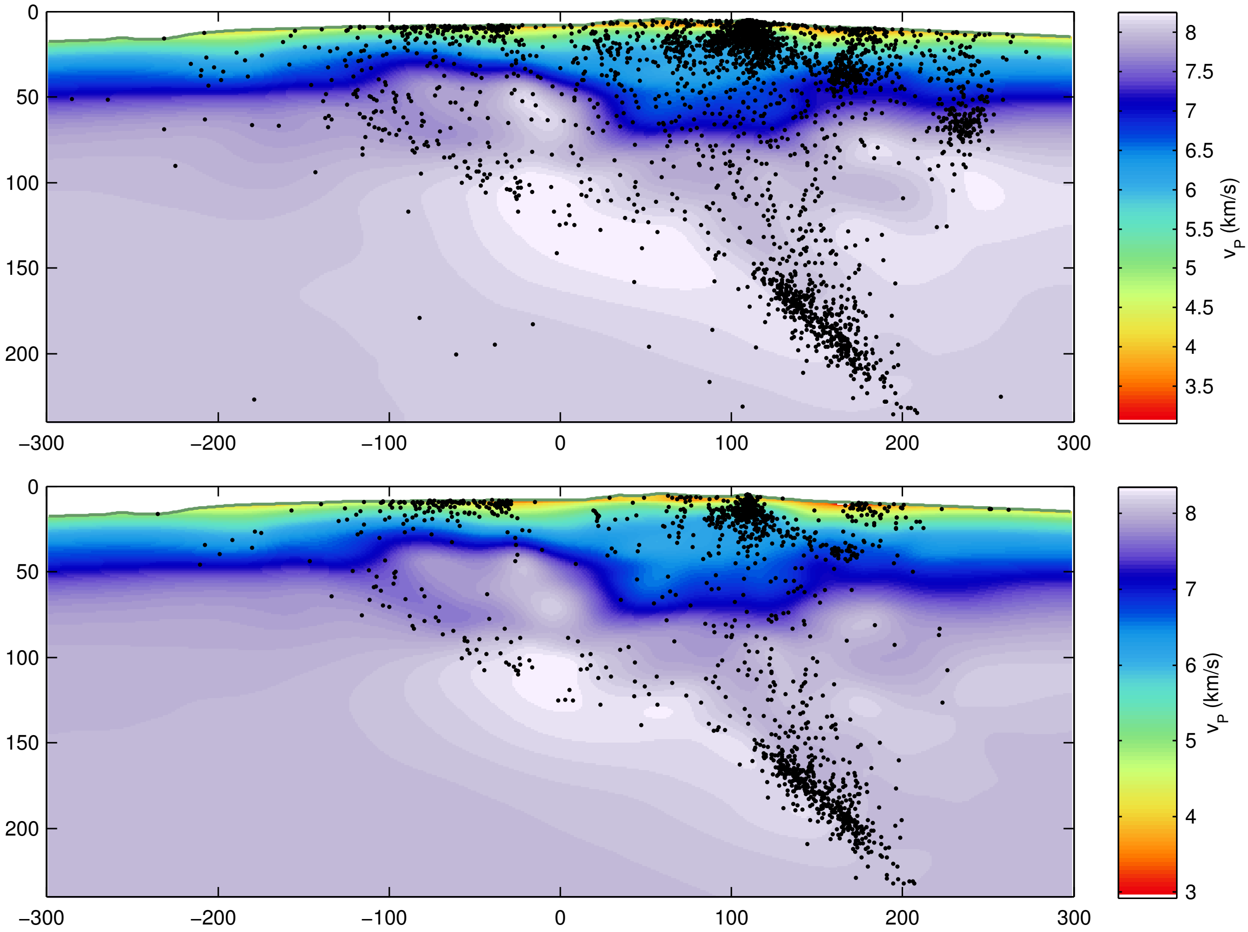

Figure 4.8: *Location of the seismic events in the vicinity of a East-West cross section at about 1.5^{o}S (top) as resulting of the tomography in the Ecuadorian box with the set of 58 060 events, (bottom) after applying the geometrical filter. We can observe that the deep Puyo nest is not affected by filtering.*

With this new set of 447 548 data we have computed new L-curves (Figure 4.9) to adapt the values of the two damping parameters ξ_0 and λ, while keeping the same smoothing lengths $\xi_V = 20\ km$ and $\xi_H = 35\ km$. The figure 4.9 shows that the various L-curves corresponding to a given value of ξ_0 can be approximately deduced by translation from each other, and allows us to assign the new values of $\xi_0 = 9\ km$ and $\lambda = 0.5$ to the damping parameters.

In summary, we have obtained two models :

- the first one over the largest box (including Northern Peru) by inverting 143 785 data (without any stations on volcanic edifice) with $\xi_V = 20\ km$, $\xi_H = 35\ km$, $\xi_0 = 6\ km$ and $\lambda = 0.7$

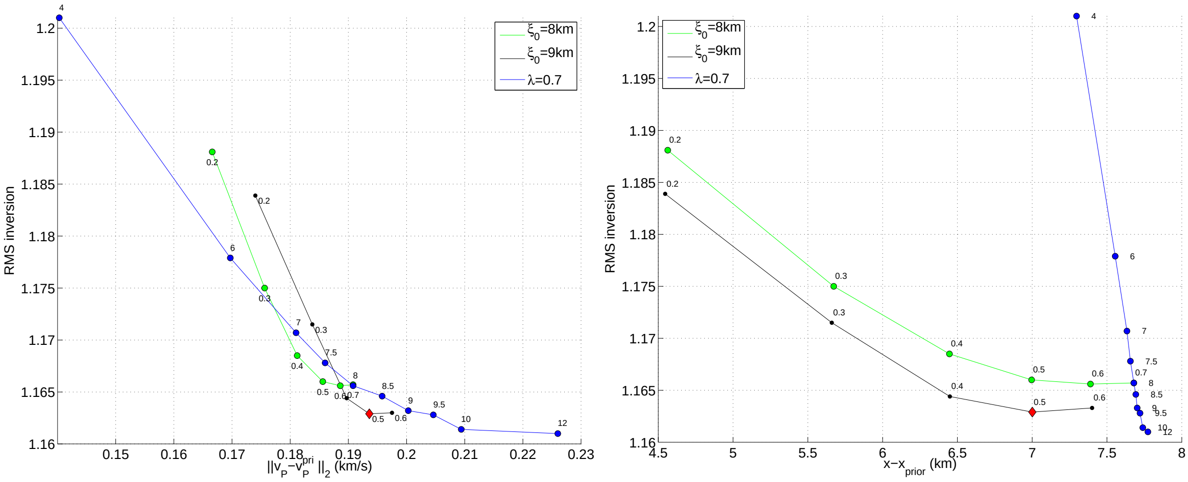

Figure 4.9: *L-curves corresponding to the filtered set of 447 548 data over the smallest box centered on Ecuador. The corresponding values (in km) of* ξ_0 *are indicated near the blue dots, and those of* λ *near the green dots. The red diamond corresponds to the chosen model, with* $\lambda = 0.5$ *and* $\xi_0 = 9$ *km.*

- the second one over the box centered on Ecuador by inverting 447 548 data (until April 2016) with $\xi_V = 20\ km$, $\xi_H = 35\ km$, $\xi_0 = 9\ km$ and $\lambda = 0.5$

4.5.3 Computing the restitution index

We have introduced the restitution index in the paragraph (2.5.3.1) equation (2.77). When discretizing, the resolution operator becomes finite dimentionnal, the expression of which is (2.33):

$$\mathbf{R}_{\hat{m}} = (\mathbf{C}_m^{-1} + \mathbf{G}_{\hat{m}}^* \mathbf{C}_d^{-1} \mathbf{G}_{\hat{m}})^{-1} \mathbf{G}_{\hat{m}}^* \mathbf{C}_d^{-1} \mathbf{G}_{\hat{m}} = \mathbf{I} - (\mathbf{C}_m^{-1} + \mathbf{G}_{\hat{m}}^* \mathbf{C}_d^{-1} \mathbf{G}_{\hat{m}})^{-1} \mathbf{C}_m^{-1} \tag{4.42}$$

The restitution index is defined for the parameters v_{P} and $v_{\mathrm{P}}/v_{\mathrm{S}}$. The inspection of equation (2.33) shows that its expression for one a these parameters at a given node that correspond to the ith component of the model space is:

$$\mathcal{I}_i = (\mathbf{R}\mathbf{w})_i$$

where $\mathbf{w}$ is the model vector defined by $w_i = 1$ or 0 according to the component i corresponds whatever the node to this parameter or not in the model space. Taking the decomposition (4.35) and (4.42) into account, we deduce that:

$$\mathbf{R} = \mathbf{I} - (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{C}_m^{-1}$$

where $\mathbf{A}$ is defined as in (4.35). It follows that:

$$\mathcal{I}_i = 1 - x_i$$

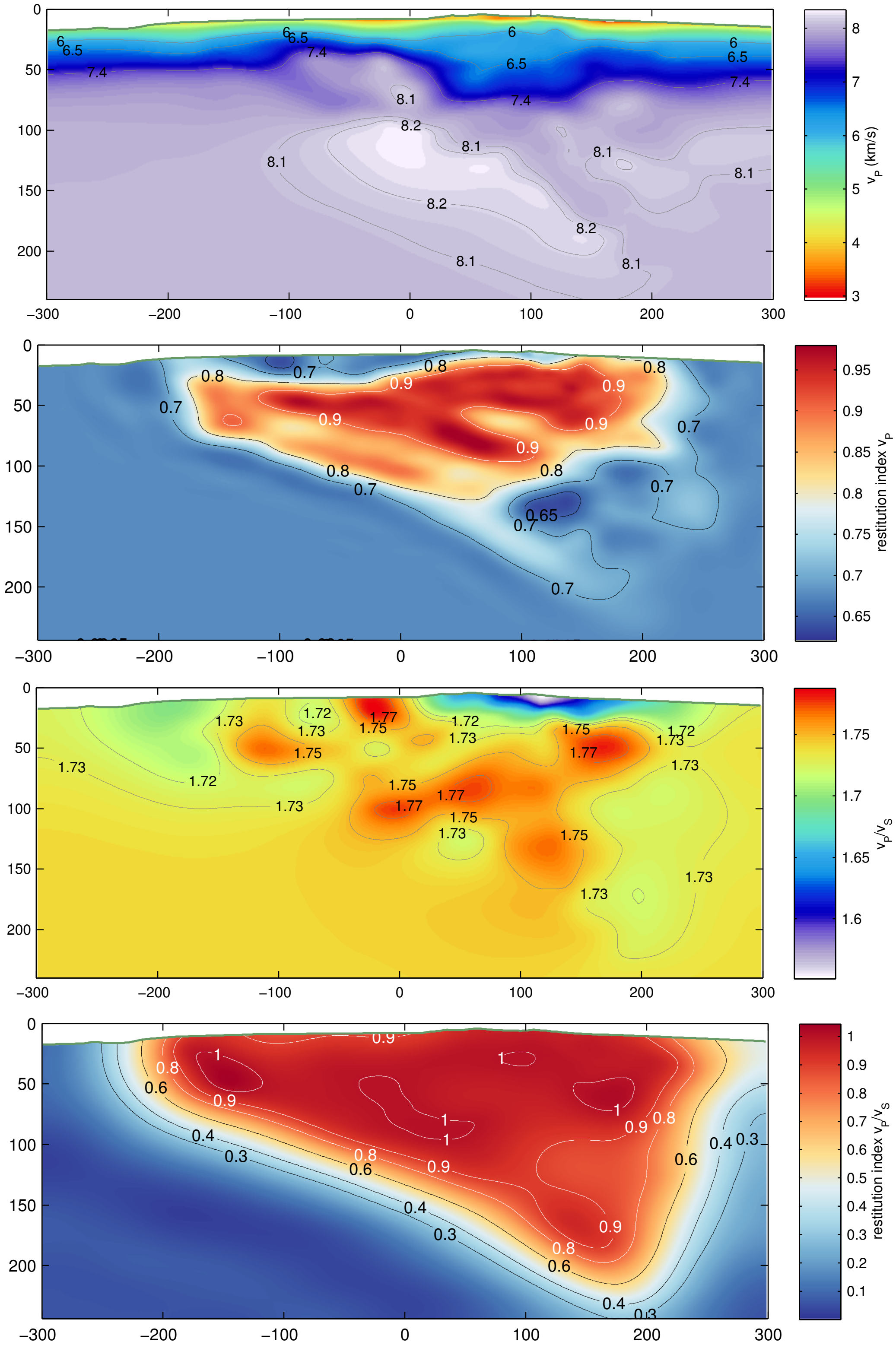

Figure 4.10: *East West cross section at 1.5°S (as figure 4.8) of restitution indices and velocity fields of the resulting model of the inversion of the filtered dataset of the Ecuadorian box. From top to bottom: v_P, restitution index for v_P, v_P/v_S, restitution index for v_P/v_S. The restitution indices allow delimiting the area outside which the resolution is surely very poor*

with $\mathbf{x}$ solution of (4.36)

$$\mathbf{A}^*\mathbf{A}\mathbf{x} = \mathbf{A}^*\mathbf{v}$$

and:

$$\mathbf{v} = \begin{bmatrix} 0 \\ \mathbf{Cor}^{-1/2}\mathbf{\Sigma}^{-1}(\mathbf{w}) \end{bmatrix}$$

This shows that the practical computation of the restitution index at each point of the domain is equivalent to performing an iteration of the inversion process.

The figure (4.9) displays the restitution index of v_{P} and $v_{\mathrm{P}}/v_{\mathrm{S}}$ over the same cross section at 1.5^{o}S as figure (4.8) and the corresponding velocity models. We can observe that the Puyo nest generates a zone where the index is greater than 0.7.

Chapter 5

Geodynamical context

Ecuador is located above one of the world's major subduction zones east of the Pacific Ocean, where the Nazca oceanic plate and the South American continental plate converge. This feature is noteworthy because western South America provides a prime case where the subduction of an oceanic lithosphere under a major continental plate can be observed in detail.

The studied region is situated in the northern limit of this wide subduction zone as is evident in figure (5.1). The major regional geodynamic elements are shown in figure (5.1) and described hereafter.

The Nazca Plate resulted from the breakup of the Farallon Plate approximately $20Ma$ ago. At that time, the new oceanic crust of both Cocos and Nazca plates started to form along the Cocos-Nazca spreading ridge and the East Pacific Rise (Szary, 2014). As an example of oceanic crust, the Nazca Plate is in average $\sim 7km$ -thick and is formed by a very thin layer of non-consolidated sediments, a volcanic layer with a marked seismic velocity gradient, and an oceanic layer which consists of mafic rocks where seismic velocities increase slowly with depth (Daniel et al., 2006). Toward the south, the Nazca plate meets the Antarctic Plate, while in the north the boundary with the Cocos plate is currently the Panama fracture zone.

The Nazca Plate moves in a $N80^oE$ direction and subducts beneath the South American Plate from -45^oS to 5^oN with a convergence velocity that varies slightly along the margin. This convergence velocity is about $79mm/year$ along the Chilean margin, $77mm/year$ along the southern and central Peruvian margin and $65mm/year$ along the northern Colombian margin (Rhea et al., 2010). In the study area, this velocity is $73mm/year$ along the North Peru margin, and $68mm/year$ along the Ecuadorian margin (Rhea et al., 2010). More precise measurements within the study area point to convergence velocities between $56mm/year$ at 1^oN and $60mm/year$ at -6^oS , in the $N82^oE$ direction (Nocquet et al., 2015).

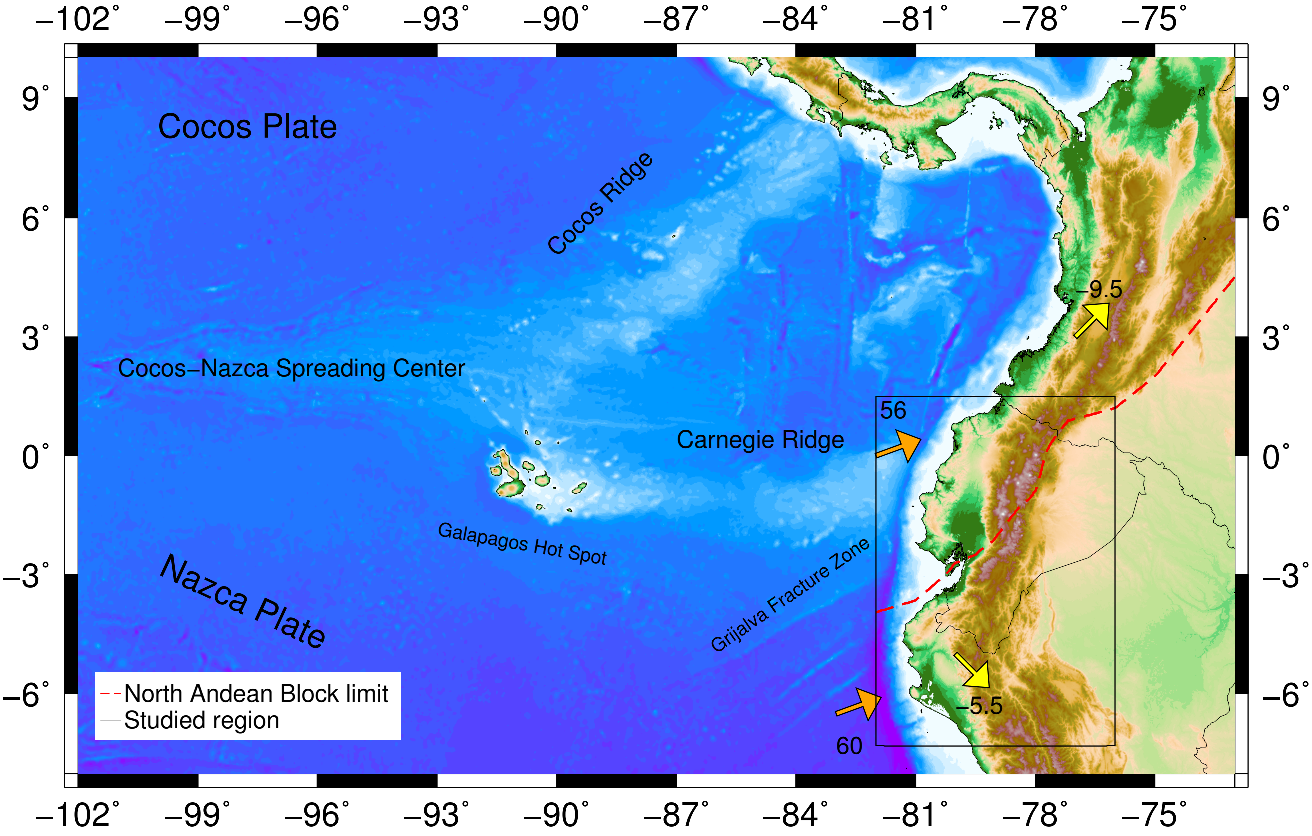

Figure 5.1: *Geodynamics and tectonics features in the study region showing velocity of convergence as orange arrows and velocities of sliver escapes as yellow arrows. The velocities are in mm/year . The Nazca and Cocos plates are shown as well as the corresponding spreading center.*

5.1 Peculiarities of the Nazca Plate

An important structure to take into account in the geodynamics of region is the Carnegie Ridge, which is an aseismic volcanic ridge that was progressively generated above the Galapagos Hot Spot. The Carnegie Ridge is transported by the Nazca plate (fig. 5.1), which makes it enter subduction along a portion of the trench located west of Ecuador. The Carnegie Ridge trends E-W, being located between latitudes 0^oN and -2.5^oS, and longitudes -91^oW and $-80.5W$. It is $282km$ across strike and $1045km$ along strike, and occupies an area of approximately $325\,000km^2$. The ridge comprises two segments, to the west and to the east, where its topography is prominent, and a depressed segment inbetween (Pazmiño and Michaud, 2009). The top of the ridge culminates $657m$ below sea level. Its southern flank is marked by a morphological lineament trending $N85^oE$ (Collot et al., 2009). This southern flank includes a terrace that is $3000m$ deep and $100 - 150km$ wide (Collot et al., 2009).

Another morphological accident of the sea floor is the Grijalva Fracture Zone (GFZ) (fig. 5.1), which marks the boundary between depth lower than $3000m$ to the north, made of Neogene oceanic crust associated with the Cocos center expansion and the Galapagos hot spot, and older oceanic crust of Oligocene age in the south, with a depth greater than $3000m$. The GFZ appears as a scarp trending $N60^oE$ and slightly concave southwards (Collot et al., 2009).

The subduction trench separates the Nazca Plate from the continental plate. The morphology of the Ecuadorian Trench (i.e., the segment of the trench that extends west of Ecuador) and its western wall reflects the large-scale segmentation of the Nazca plate imposed by the existence of the GFZ and Carnegie Ridge (Collot et al., 2009). The depth of the Ecuadorian Trench is $5000m$ west of the Gulf of Guayaquil (-3^oS) but decreases down to $2000m$ northwards due to the subduction of the Carnegie Ridge. The Ecuadorian Trench has a very steep wall along its continental side; it has an approximate width of $20km$ in the Gulf of Guayaquil, and $8km$ west of Manta (-1^oS). The trench receives terrestrial sediments discharged by the coastal rivers, especially through the canyons of the Esmeraldas River to the north and the Guayas River to the south (Goyes, 2009) .

The Cocos-Nazca spreading center forms the boundary between the Cocos and Nazca plates (fig. 5.1). The spreading center originated when the Farallon Plate split into the Cocos and Nazca plates, $\sim 23My$ ago in the Oligocene (Meschede and Barckhausen, 2000).

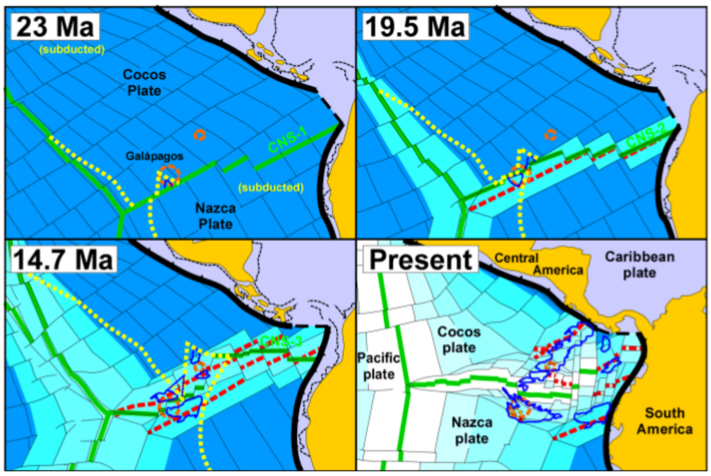

Figure 5.2: *Changes in the directionality of CNS from Oligocene to the present. Direction changes from NE to WE and it is stopped in the Panama Fracture Zone nowadays. Figure from Meschede and Barckhausen, 2000.*

A study of the past history of the Cocos-Nazca spreading center (CNS) concluded that spreading initially occurred along a NE-trending ridge, but later switched to the current W-E ridge, which abruptly stops against the Panamá fracture zone at about -83^oW,4^oN (Meschede and Barckhausen, 2000) (figure 5.2). This results in that what remains of the original CNS is currently mainly located south of the Equator; in particular, the NE-trending Grijalva Fracture Zone separates a younger oceanic crust in the northwest, which started to form in the earliest Miocene, from an older crust in the southeast (Meschede et al., 1998; Londsdale, 2005) (figure 5.3).

Activity of the CNS contributed to thicken the oceanic crust. To the south, this influence is clearly recognizable where the GFZ exhibits a sudden change of almost $700m$ in the bathymetry, which reflects a contrast in crustal thickness (Londsdale, 2005). Because this thickened crust is also subducted, two dissimilar subduction regime have to be distinguished: (1) north of the GFZ, where the subducting oceanic crust is $22Ma$-old; and (2) south of GFZ, where this crust is $30Ma$-old (Lonsdale, 2005; Yepes et al., 2015).

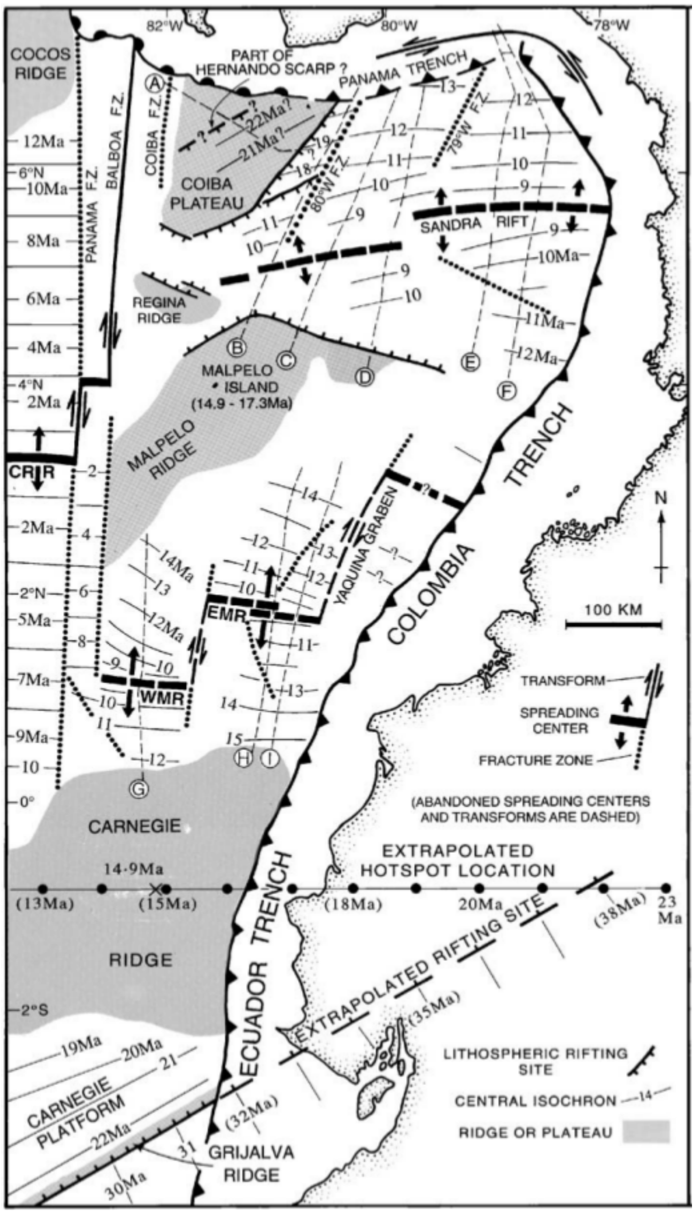

__Figure 5.3:__ Abrupt discontinuity in magnetic profiles at Gijalva Fracture Zone from $22Ma$ *to* $30Ma$ *due to the original CNS. Possible extrapolation of GFZ coincides with the position of Puyo seismic cluster. Figure from Londsdale,2010.*

5.2 Effects of subduction along the margin of Ecuador

Subduction of the Nazca Plate has had a significant bearing on the morphology of the continental margin, along which it caused the formation of mountain ranges that are generally parallel to the margin: the Coastal Cordillera, Western Cordillera, Interandean Zone, and Eastern Cordillera. East of the latter extends the Guyana Shield or the Brazilian Shield (Daniel et al., 2006). In the study region, these megastructures are shown in figure (5.4). The Ecuadorian Coastal Cordillera has a maximum height of $\sim 800m$ and a length of $600km$ (Reyes, 2011). The two main Cordilleras (Western and Eastern) are parallel and trend roughly NS: they are topographically elevated, between $4000-4400m$, and separated by the inter-Andean Depression $(2000-2700m)$ (Alvarado, 2012).

Another important characteristic of Ecuador north of -2^oS is the existence of a volcanic arc, which is $\sim 120km$ wide. It is divided into three areas: the volcanic front along the Western Cordillera, the magmatic arc proper along the Eastern Cordillera, and a backarc area in the Subandean, eastern lowlands. The volcanoes of the main volcanic arc and the backarc area show mainly andesitic products, but volcanoes along the active volcanic front are rather dacitic (Alvarado, 2012).

The volcanic arc of Ecuador is well-developed, to the point that the question of whether an arc orogeny has occurred is worth investigating. A relation of this arc orogeny with the subduction regime has been suggested by the steepness of the slab from 1^oN to 1.5^oS (Guillier et al., 2001). The main reservoirs of the magma produced by subduction-related melting are estimated to be located approximately at the crust-mantle boundary and about $20-40km$ depth (Michaud et al., 2009).

5.3 Characteristics of the continental crust along the margin

Ecuador and northern Peru regions are characterized by the presence of the North Andean Block (NAB). This is a crustal sliver formed by fragments of the Caribbean Plate that were accreted to the South American Plate 85 to 60 millions years ago (Jaillard et al., 2009).

The main accreted terranes are the Guaranda, Piñón-Naranjal oceanic terranes and the Macuchi island arc (Jaillard et al., 2009). They serve as a basement for the sedimentary basin of Borbón in Colombia-Ecuador north-west border (Borrero et al., 2011), the Manabi and Progreso sedimentary basins in Ecuador (Witt and Bourgois, 2010), the Guayaquil-Tumbes basin in the Guayaquil Gulf (Witt and Bourgois, 2010), the Lancones basin in the Ecuador-Peru border area (Jaillard et al., 2009), and the Zorritos basin in northern Peru (Witt and Bourgois, 2010). The Talara and Sechura basins extend south of 4^oS, and continue southwards of 7^oS into the Trujillo basin (Higley, 2004). The location of some of these sedimentary basins is shown in figure (5.4).

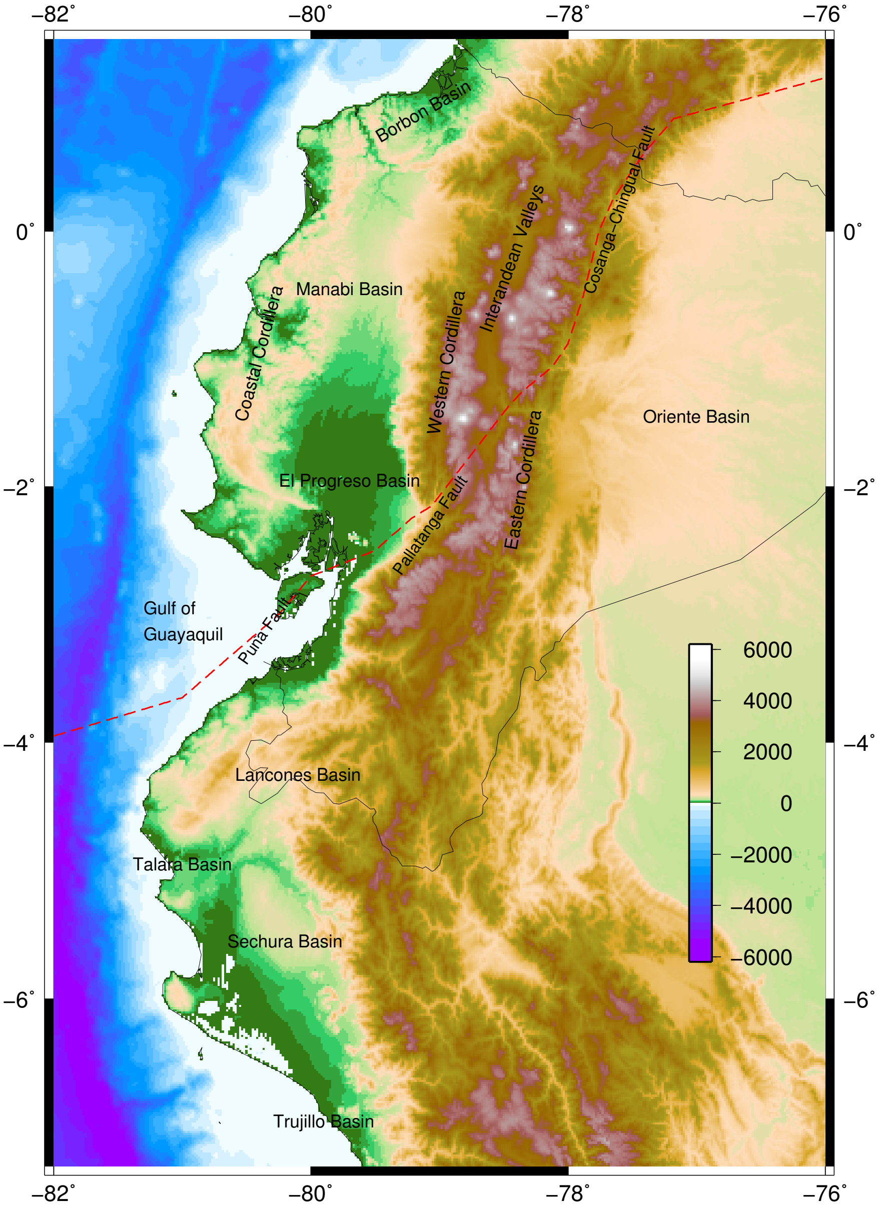

Figure 5.4: *Cordilleras, sedimentary basins and seismic faults in our study region.*

The main geological structures are parallel to the oceanic trench and trend approximately N-S. From west to east, the Coastal area, the Western Cordillera, the narrow Inter-Andean valleys, the Eastern Cordillera, the Sub-Andean belt, and finally the Oriente Basin, are usually distinguished (Baldock, 1982). There is a clear opposition between (1) the Western region, which is characterized by a basement comparable in composition to ocean or island arcs, and includes the Coastal area and Western Cordillera, and (2) the Amazonian region, which is characterized by a sialic basement and comprises the Eastern Cordillera, the Sub-Andean belt and the Amazonian region (Mégard, et al., 1987). The eastern limit of the accreted oceanic terranes is still debated, because it is probably located in the Inter-Andean valleys, which is filled by Neogene to recent volcanic and volcaniclastic deposits that obscure the basement geology. However, the Eastern Cordillera and Sub-Andean belt appear to be well delimited by two eastern Andean major faults (Guillier et al., 2001).

The boundary between the NAB and South America is unclear in detail and might be diffuse, but overall it runs from the Gulf of Guayaquil to the Eastern Cordillera, extending into the northeastern area of the Eastern Cordillera of Ecuador (figure5.1). The NAB does not coincide with a block formed by the oceanic accreted terranes, since it presently involves part of the sialic Eastern Cordillera (Alvarado et al., 2016). This fault system (figure (5.4)) runs from the Puná island in the Gulf of Guayaquil and prolongates into the continent through the Pallatanga fault following a $N60^oE$ direction. In northern Ecuador and at the Colombian border, the Cosanga-Chingual fault system strikes $N20^oE$ (Alvarado et al., 2016).

5.4 Motion of the North Andean Block as a consequence of oblique subduction

The oblique subduction regime is responsible for the current escape of the NAB toward the northeast, the rate of which is $9.5mm/year$ (Nocquet el al., 2014). This oblique subduction is verified by GPS measurements (Trenkamp et al., 2002), the moment tensor analysis of the seismicity in the area (Corredor, 2003) and the presence of undoubted tectonic peculiarities such as the pull-apart-like basin active in the Gulf of Guayaquil (Deniaud et al., 1999). The oblique subduction has been active since at least the Miocene, and provides a new kinematic model for the NAB (Alvarado et al., 2016).

Oblique subduction occur when the angle φ between the convergence direction and the trench is different from 90^o, as outlined in figure (5.5). This phenomenon, which implies an oblique displacement on the motion plane is not very common, and therefore most of the faults are only reverse, only normal, or only strike-slip (Jolivet and Nataf, 2001). But if oblique subduction exists, two parallel lithospheric structures are developed: the subduction trench absorbs the normal component, and the strike-slip area, shown in blue lines in fig. 5.5, absorbs the parallel component (Jolivet and Nataf, 2001)

The oblique velocity of the subducting plate V_c is transmitted to the continental plate because there is a friction stress τ_n between the two plates. When a friction force F_f is generated, it can be decomposed into 3 components: horizontal, tangential, and vertical. It is the tangential component of this friction force F_{f_t} that is responsible for the the sliver scape. If β is the dip subduction angle, the components of the friction force are (Chemenda et al., 2000):

$$\begin{aligned} F_{f_t} =& F_f sin\varphi \\ F_{f_v} =& F_f cos\varphi sin\beta \\ F_{f_h} =& -F_f cos\varphi cos\beta \end{aligned} \tag{5.1}$$

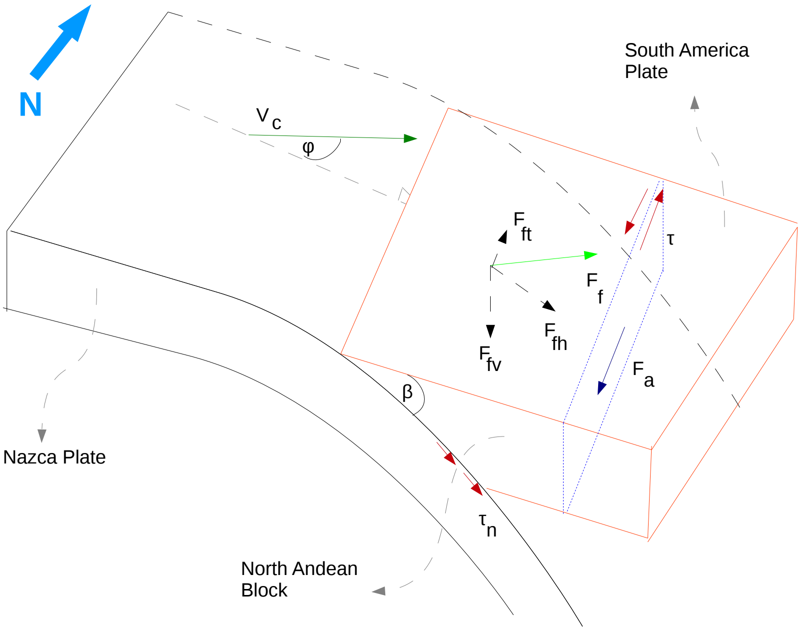

Figure 5.5: *Schema of the northward escape of the crustal sliver formed by the North Andean Block (NAB), thanks to oblique subduction of the Nazca plate beneath South America. Figure adapted from Chemenda, A., S. Lallemand and A. Bokun, 2000.*

The condition for sliver escape is that the tangential component must beat the force F_a generated when the block tries to resist the strike-slip displacement τ:

$$F_{f_t} \geq F_a \tag{5.2}$$

Analog modeling of oblique subduction (Chemenda et al., 2000), shows that slip partitioning is able to move a sliver parallel to trench if there is high interplate friction and previous weakening of the upper plate. In the case of the NAB, the first requirement is accomplished thanks to the known seismic asperities in the subduction zone (Chlieh et al., 2014). Crustal weakening is related to contact of the NAB with the South American Plate through the Guayaquil-Dolores Megashear (Egbue and Kellog, 2010).

The motion of the NAB generated a system of crustal faults parallel to the trench and to the Western and Eastern cordilleras (Veloza et al., 2012), the Pallatanga fault being located at its southern tip (Baize et al., 2015).

5.5 The Carnegie Ridge and the dynamics of slab

It is necessary to address the entire process of subduction before focusing on the effects of the subduction of the aseismic Carnegie Ridge. The Earth's mantle behaves like a fluid submitted to a convective regime. Therefore equations of fluid mechanics are needed to describe its motion (Turcotte and Schubert, 2002).

Three kinds of equations are needed to describe this fluid mechanics (Billen, 2008). In the case of an incompressible fluid, the equation of mass conservation leads to the well-known continuity equation of fluid velocity **u**:

$$\nabla \cdot \mathbf{u} = 0 \tag{5.3}$$

The equation for the momentum conservation relates the fluid stress σ to the external forces **f** and the internal fluid pressure P:

$$\nabla \cdot \sigma + \mathbf{f} + P = 0 \tag{5.4}$$

Finally, the equation for the heat-energy conservation provides the evolution on fluid temperature T in space and time. This equation has the general form:

$$\frac{\partial T}{\partial t} = -\mathbf{u} \cdot \nabla T + \kappa \nabla^2 T + Q \tag{5.5}$$

Q is the internal heating of slab; it can be neglected because it is very small in comparison to heat exchanges between the slab and surrounding mantle. κ is the coefficient of thermal conductivity.

These equations must be resolved in a three-dimensional space and numerical methods should be used. An acceptable approximation of these equations are the two-dimensional case cartooned in figure (5.6).

The simplification of fluid motion equations to $2D$ allows us to obtain in first approximation a comprehension of why the slab dip angle is not 90^o (Turcotte and Schubert 2002). Two fluid pressures are developed in the upper part and in the lower part of the slab. These two pressures contribute to lift the slab against the gravity force, and to give it its particular dip angle.

In a more advanced modeling of the slab in $2D$, the effects of the aseismic ridge can be considered (Gerya et al., 2009). Results of this modeling show that the slab dip angle is not directly influenced by changes in the slab composition and density caused by the ridge. A noticeable effect of an aseismic ridge in the slab subduction only appears when subduction is nearly flat, the positive buoyancy provided by the ridge acting in this case to reinforce the flat subduction already under development but not triggering it.

On the other hand, subduction of an aseismic ridge indubitably uplifts the topography of the overriding plate, at least near the trench. This uplift can be transient if the ridge is subducted under the accretionary wedge, but it may remain as a permanent modification of the topography if the overriding plate is a continental plate (Gerya et al., 2009). Uplifts produced by ridge subduction are also observed in analogue models (Martinod et al., 2012).

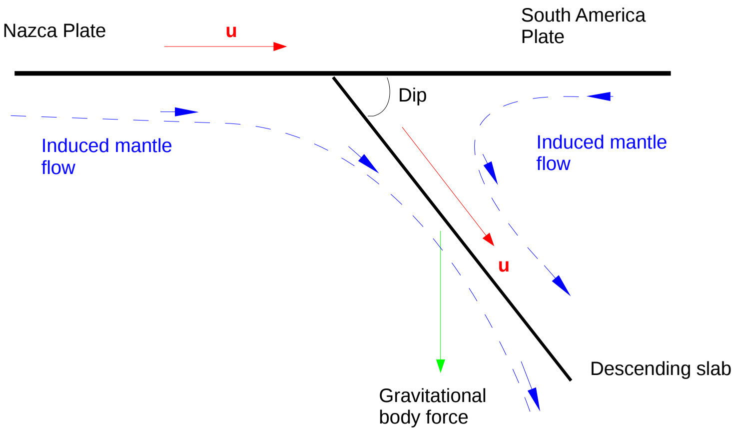

Figure 5.6: *Two-dimensional simplification for the slab subduction dynamics. The problem is to solve the fluid equations for the mantle flow. Modified from Turcotte and Schubert, 2002.*

In Ecuador, subduction of the Carnegie Ridge clearly caused the uplift of the Manta Peninsula and La Plata Island (Pedoja et al., 2006 a). Similar uplifts extended throughout the Talara region from $2^o N$ to $6.5^o S$ (Pedoja et al., 2006 b). The subduction of the Carnegie Ridge is however not the only cause of these uplifts: another important parameter is the concave shape of the subduction trench, as illustrated in numerical models (Bonnardot et al., 2008). This aspect has been used to explain uplifts along the entire Ecuadorian margin (Dumont et al., 2014).

Therefore the results of aseismic ridge subduction modeling (Gerya et al., 2009) imply that the Carnegie Ridge cannot be held responsible for the initiation of flat subduction west of Ecuador. The idea of the flat slab was proposed on the basis of seismicity localization using Global Network (Gutscher et al., 1999) but could not be confirmed with a more refined seismic experiment using a dense network situated directly above the region where the slab dip angle is $25^o - 35^o$ down to $200km$ depth (Guillier et al., 2001). Seismic evidence of flat-slab subduction is however confirmed in southern Ecuador, south of $3^o S$ (Beathe et al., 2001), and in northern Peru (Tavera et al., 2006).

A wider description of the slab dip angle in Ecuador has been possible by using the complete RENSIG catalog to localize seismicity more precisely (Yepes et al., 2015). The dip angle is 20^o at $-1^o S$ latitude, according to the information provided by the La Maná seismic cluster. It dips to 35^o at $-2^o S$ latitude according to the Puyo cluster, and decreases to 12^o at $-3^0 S$. North of the Equator, Yepes et al.'s (2015) re-localization does not provide a good definition on the slab dip angle. This change in the subduction slope might be related to a different subduction regime across the Grijalva boundary, as explained above, rather than to the influence of the Carnegie Ridge subduction (Yepes et al., 2015).

The Carnegie Ridge is characterized by a a thickened oceanic crust, which reaches almost $14km$ in the subduction zone (Graindorge, 2004) but, as stressed above, this over-thickness is not sufficient to create a buoyancy capable to produce flat-slab subduction.

The main reason to dismiss the Carnegie Ridge as a major factor in the general dynamics of the NAB is that transmission of vertical strain generated by the ridge subduction are inefficient, and that all features that may be attributed to its influence can find other explanations (Michaud, 2009).

5.6 The Mohorovičić discontinuity

The Mohorovičić discontinuity, simply called Moho, is the boundary between crust and mantle. It was discovered by examining the refraction patterns of seismic waves. This refraction is due to marked changes in Earth's density structure, i.e. changes in its refraction index. Other geophysical techniques, like seismic reflexion, electrical conductivity, and gravity potential, concur to confirm the existence and location of the Moho boundary (Cook et al. 2010).

The physical behavior of the Moho boundary is not well understood yet. If the crust is envisioned as a brittle elastic medium (Scholz, 2002) and the mantle as a visco-elastic medium, the Moho is expected to exhibit an intermediate solid-fluid phase (Machetel, 2008). Another idea, known as the metamorphic (or metasomatic) front hypothesis, posits that the Moho is overprinted by a phase transformation (Eaton, 2006). However, this hypothesis may be an oversimplification as it requires to recognize the Moho discontinuity to have been produced by other more complex geophysical phenomena, such as: the relict Moho hypothesis posits that the oceanic Moho is preserved during continental assembly; the magmatic underplating hypothesis posits that a new Moho is formed by episodic emplacements of sill-like intrusive bodies; and the regional décollement hypothesis posits that the Moho behaves as a structural detachment (Eaton, 2006).

One of the surprising Moho characteristics of the tremor generated in its neighborhood (Katsumata and Kamaya, 2003). This seismic phenomenon is a low-frequency signal that is not directly associated with fracture episodes. Therefore a fluid flow model was proposed to explain it (Katsumata and Kamaya, 2003), but a shear slip motion of the crust over the mantle has later appeared as a more plausible explanation (Shelly et al. 2006).

In Ecuador, no direct experiment has been undertaken yet in order to elucidate the Moho issue. In spite of this lack of information, three previous studies indirectly address the question. A summary of the main characteristics of the Ecuadorian Moho is presented hereafter, in chronological order.

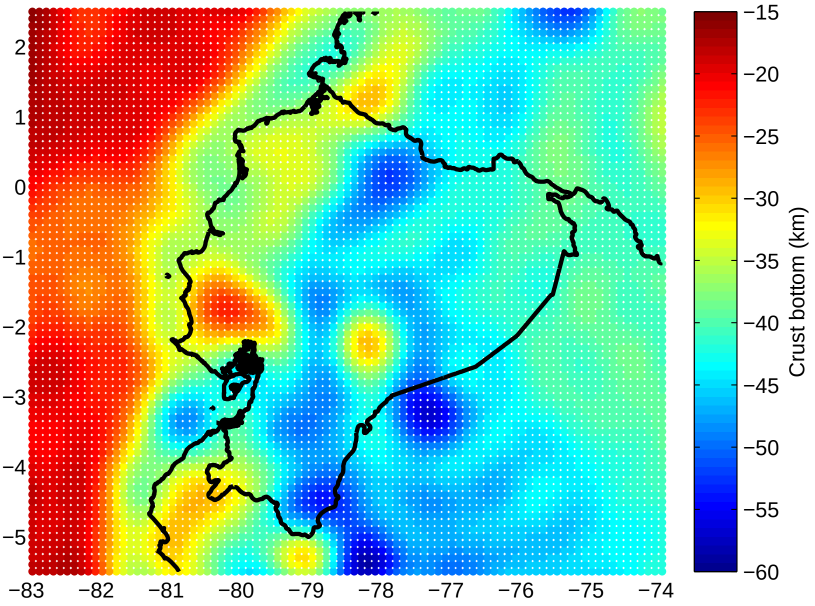

Figure 5.7: *Moho model from gravity 0.1 degree model. The main problem for this model is the excessive oscillation in the Andes.*

5.6.1 Gravimetric Moho in Ecuador

This study was based on geopotential and topography data inversion with a 0.1^o resolution (Chambat, 1996). The model is global and a window over the study region has been chosen.

In this model, the Moho is $15km$-deep beneath the oceanic crust, but deepens down to $20km$ below the Carnegie Ridge. The crust of the North Andean Block is $35km$-thick in the coastal region, but this thickness decreases to $30km$ beneath the Borbón and Manabí basins. Two apparent gravity anomalies are observed, one in the Gulf of Guayaquil, due to presence of sediments, and another one in the El Progreso Basin. A narrow, $45km$-deep, transition belt exists along the Andean Cordillera, where the Moho appears to be $55 - 60km$-deep. In the Amazonian region, the Moho is about $40km$-deep beneath the Guiana Shield. Between the Central Ecuadorian Andes and the Guiana Shield, a marked anomaly apparently suggests a locally shallow Moho, which is however unexplained and might reflect a numerical artefact. All these features are visible in figure (5.7).

5.6.2 GEMMA Moho

This global Moho model is based on GOCE satellite measurements of the Earth's gravity field (Reguzzoni and Sampietro, 2012). These data were inverted and interpolated until

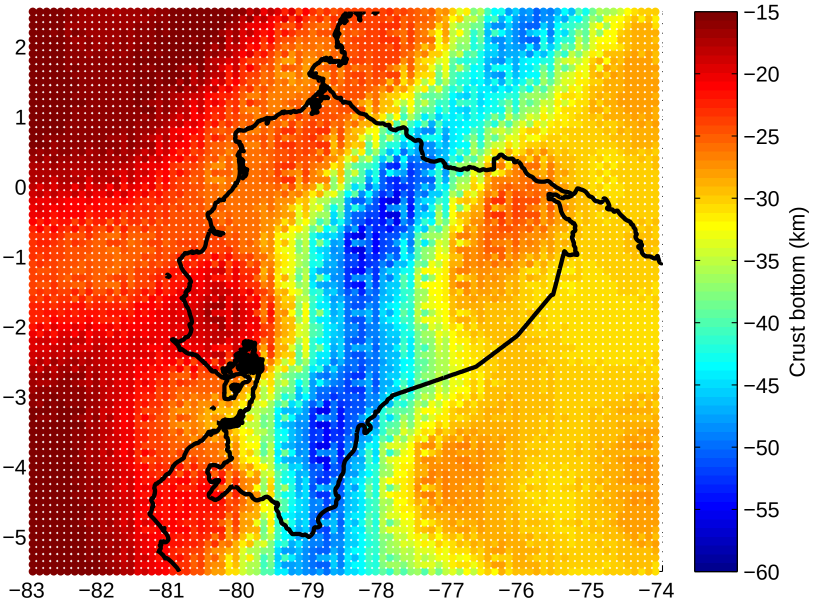

Figure 5.8: *GEMMA model of the Moho in Ecuador. An advantage of this model is the reduction of anomalous oscillations during the inversion process. Moho depths for the Coast and Amazonian Basin seem however to be overestimated.*

obtaining a 0.1 degree resolution grid of crustal thickness. The resulting models, called GEMMA MODELS, are in free access.

A window encompassing the Ecuador region is presented in figure (5.8). The base of the oceanic crust is $15km$-deep but no contrast is apparent between the Carnegie Ridge and the North Andean Block because the area where the Moho is $20-25km$-deep is continuous over these two domains. Gravity anomalies corresponding to the Gulf of Guayaquil and El Progreso Basin are visible. Along the transitions to the Western and Eastern cordilleras, crustal thickness reaches $32-45km$. The Moho depth is about $60km$ beneath the Ecuadorian Andes, and $30km$ beneath the Oriente Basin.

5.6.3 3D a-priori seismic Moho

This Moho model is based on a compilation of previous gravity studies specifically designed for the Ecuador region in order to provide a three-dimensional velocity model for P seismic waves (Font et al., 2013), and thus not a Moho representation. However, it has been the most detailed synthesis on this issue until now, providing an explicit description of each geodynamic area.

Starting from the north, the Moho beneath the oceanic crust is $5km$-deep in southern Colombia, deepens down to $19km$ below the Carnegie Ridge, and decreases to $14km$

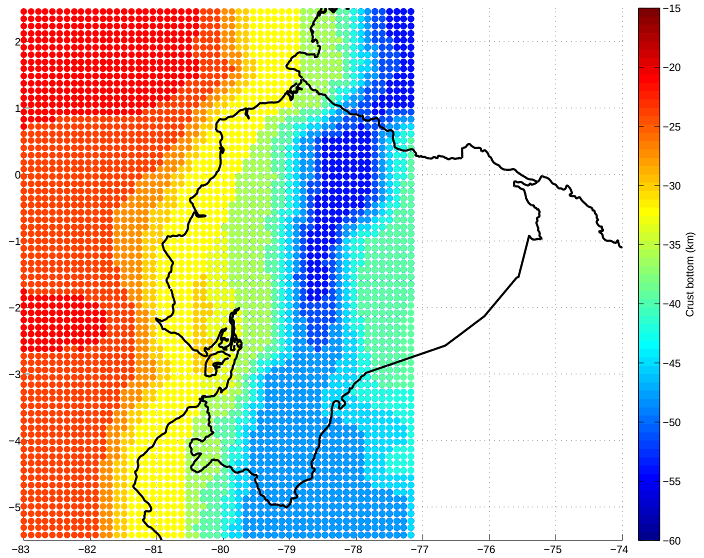

Figure 5.9: *Moho model obtained from a 3D ad-hoc velocity model. The model originally published does not extends east of* $77°W$*. From west to east, the oceanic crust, North Andean Block, Andes Cordillera, and Oriente Basin can be distinguished.*

south of the ridge. The oceanic Moho is defined here in a general sense, i.e. as an abrupt velocity gradient from 7 to $7.8km/s$ over $1km$.

More to the east, the Moho is $22-30km$ deep in the coastal region of the North Andean Block, where it is characterized by an interface at about $7km/s$. In the Ecuadorian Andes (Western and Eastern cordilleras, and inter-Andean region), it is suggested that the crustal thickness is $55-65km$ and the crust-mantle interface about $7-8km/s$. On both sides of the Andean Cordillera, the lowlands present a $30km$-deep Moho, with an interface at $7-7.8km/s$. Beneath the Oriente basin, the crustal thickness of the Guiana Shield is apparently $30-35km$, and a change in P-waves velocity between crust and mantle occur at $6.8km/s$.

Relying on these data, a model of Moho was obtained using the following parameters for 3DVM filtering: P-wave maximum velocity: $8km/s$; P-wave minimum velocity: $6.8km/s$; minimum crustal thickness: $5km$; maximum crustal thickness: $60km$. The resolution from this 3DVM model is $12km$ in latitude, $12km$ in longitude, and $6km$ in depth. The result is presented in figure (5.9).

5.6.4 Discussion of the three models presented above

The three models presented above may be scrutinized under the scope of the known geodynamic context in Ecuador (Jaillard et al., 2002; Jaillard et al., 2009). In the Pacific Ocean, the Moho is about 10 km-deep. Along the coastal lowlands, the Moho depth is almost the same, but volcanic-sedimentary rocks and Cenozoic sediments crop out. The Western Cordillera, where crustal thickness can reach $40km$, consists of accreted oceanic crust and exposes volcanic-sedimentary rocks. Crustal thickness along the Inter-Andean valleys is apparently $45-50km$; in these areas, Cenozoic and Mesozoic sediments overlie pre-Mesozoic, partly metamorphosed rocks and accreted oceanic crust. The Eastern Cordillera, where the Moho apparently rises up to $40km$, displays Mesozoic sediments, Paleozoic rocks and the same Precambrian basement. Crustal thickness is $35km$ in the Sub-Andean hills, where the upper consists of Precambrian basement, Paleozoic rocks, and Cenozoic and Mesozoic deposits. Finally, the Moho is only $30km$-deep beneath the Amazonian lowlands, which include the same strata and rocks as the Sub-Andean hills.

Combining these geodynamical characteristics, it appears that the three models fairly coincide regarding the oceanic and Andean regions. However, in the coastal and Amazonian regions, the GEMMA model seems to underestimate the Moho depth, whereas the Gravity model and the 3D models are in agreement, and more compatible with the regional structure.

Given that the current knowledge regarding the North Andean Block and Guiana Shield is relatively good and consensual (Cediel et al., 2003; Taboada et al., 2000), their geostructural characteristics can be used in order to complete information about the Moho in the coastal and Amazonian regions. This allows a comparison between the depths of the Ecuadorian Moho with those of other regions that have been studied in more detail, such as Colombia, where crustal thickness in the northern Andes can reach almost $45km$, and $40km$ in the Guiana Shield (Cediel et al., 2003). Differences are however apparent, since in Ecuador the Moho shows depths of only $5-15km$ beneath the Borbón Basin, of $15-20km$ beneath the Manabí Basin, and of $20-25km$ beneath the El Progreso Basin (Cediel et al., 2002).

Regarding the Guiana Shield in Ecuador, the Moho is deeper than $30km$ (Cediel et al., 2002). Once again Chambat-Valette and 3DVM models agree with Moho depths lower in the Coast and Amazonian Basins, whereas GEMMA does not give accurate values in the Progreso Basin. Regarding the Amazonian Basin, Chambat-Valette and 3DVM give again more accurate depths.

A preliminary model of the Ecuadorian Moho could be established by combining the best characteristics of these three models (Araujo, 2013).

5.6.5 A priori seismic information regarding the Moho in Ecuador

Previous seismological experiments have obtained information about the Moho by means of seismic reflection, at least along the margin of Ecuador. Two important marine seismic experiments were achieved: SISTEUR (Seismic Reflection-Refraction Experiment) in 2000, and SALIERI (South American Lithospheric Transects Across Volcanic Ridges), which allowed to obtain many seismic profiles (Gailler, 2005). Among them, three across-strike transects were acquired west of Ecuador, and one west of southern Colombia. A trench-parallel seismic transect was also recorded over the Carnegie Ridge.

Profiles across the Ecuadorian coastline were located at Esmeraldas, Carnegie Ridge, and Gulf of Guayaquil and they provided an invaluable knowledge about the Moho depth in these three regions (Gailler et al., 2007; Graindorge et al., 2004). The corresponding data contributed to assess more accurately the Moho depth in the region comprised between the trench and the North Andean Block. The trench-parallel seismic profile clearly showed that, west of the trench, the Moho deepens under the Carnegie Ridge to almost $20km$ (Sallarés et al., 2004).

Bathymetry measurements and a posteriori mapping of the margin of southern Colombia and northern Ecuador (Collot et al., 2006 a), and within the margin of central and southern Ecuador (Collot et al., 2006 b), provide valuable knowledge about the topography and depth of the trench and adjacent oceanic plate. These topographical features can be used to estimate the Moho depth using isostasy as a principle.

Chapter 6

The Tomographic Results

Previous local tomographic inversions over this study area are very scarce. Actually, only one very preliminar 3-D tomography (Prevot et al., 1996) was performed over continental Ecuador by using 14 684 P-arrival times obtained by the RENSIG network between 1990 and 1993, and by using a prior 1-D model derived from the models of Ocola et al. (1975), Fluh et al. (1981) and Leeds (1977). Despite the poor level of resolution of this tomography, for which parallelipipedic cells 60 × 60 km wide and with thickness ranging from 5 to 25 km were used, the authors were abble to identify high P velocities beneath the Western Cordillera, and low velocities beneath the Interandean depression and the Eastern cordillera. They also identified a low velocity zone beneath the Pisayambo nest that they interpreted as a possible upwelling of hot materials.

In contrast, the south Colombian-Ecuador convergent margin has been intensively studied by using travel time inversion of wide angle seismic data. These data were acquired during two major marine seismic experiments: the SISTEUR cruise in 2000 (Collot et al. 2002) and the SALIERI (South American Lithospheric Transects Across Volcanic Ridge) cruise in 2001 (Flueh et al., 2001). In particular, Sallares et al. (2005) investigated the seismic structure and the crustal overthickening of the Carnegie ridge along two profiles of the SALIERI experiment, and deduced some constraints on the process for the formation of this aseismic ridge. One of this profile, running alongside the Ecuadorian coast, West of the trench, provides crucial informations at the western limit of our study region. Moreover, three transects across the trench, acquired during these experiments, SAL-6 North of Ecuador, SIS-4 at 1.4°S and SAL-2 off the Guayaquil gulf, allow Graindorge et al., 2004 to investigate the crustal velocity structure of the margin wedge and of the Carnegie Ridge across the trench. They also allowed Gailler et al. (2007) to detail the segmentation of the margin into three contrasted zones (see figure 6.39).

By using the Prevot et al. model, Guiller et al. (2001) analysed the spatial distribution of seismicity beneath central Ecuador from a temporary network of 53 stations deployed between 77.4 - 80 W and 0.1 - 1.5 S in the beginning of 1995. Complementing these localizations by the locations of the 1964-1995 seismic events ($M_w \leq 5.2$) determined by Engdahl et al. (1998), they concluded that the oceanic plate is plunging continuously down to a depth of about 200km (actually 150 km, except for the Puyo nest), contrary to Gutscher et al. (1999) proposition of a flat slab extending eastwards at a depth of 80-100 km as far as 500 km from the trench. They also interpreted the parallel features dipping eastward with an angle of about 35° that exhibits the seismic activity around 1°S and 78.5°W as the reactivation of the sutures related to the Late Jurassic to Early Tertiary accretions of oceanic terranes.

Manchuel et al. (2011) analysed the seismicity recorded by a combined onshore offshore temporary seismic network (ESMERALDAS) involving 26 OBS and 31 seismic stations deployed during 3 months in 2005 from the trench up to the cordillera in Northern Ecuador. They found a Wadati-Benioff plane well defined down to a depth varying between 100 km and 140 km with a dip angle of about 25° that also clearly contradicts Gutscher's hypothesis in Northern Ecuador. Font et al. (2013) also constructed a 3-D velocity model based on the integration of geophysical and geographical data to upgrade the image of the seismicity distribution over the margin in the interseismic period.

Two local tomographies have also focussed on volcano structures by using the Benz et al. (1996) method that was intialy developped for studying the Redoubt volcano and that we briefly described at the beginning of chapter 4. On the one hand, the tomography of the Tungurahua (Molina et al., 2005) identified a high velocity zone within the central base of the edifice under the vertically aligned hypocenters between 1 and 4 km below the summit, and was interpreted as the feeding conduit system. The authors also suggested the existence of an old lateral dike system due to the identification of two other high anomalies under the lower northeastern and southern flanks of the edifice, one them connected to the central conduit. On the other hand, the tomography of the Guagua Pichincha volcano (Garcia Aristizabal et al., 2007) was performed with few stations to characterize the crisis of 1999. In contrast with the other study, a low velocity anomaly was identified beneath the caldera and was interpreted to be an active volcanic conduit.

6.1 Comparing results in the two inversion boxes

As already detailed at the end of chapter 4, we have performed tomographic inversions in two different parallelipipedic domains (boxes) with two different data set. We firstly considered a large boxe containing Northern Peru in order to incorporate data from the RSN and of a temporary network that are essential for identifying the geometry of the subduction in the southern region of Ecuador. We removed all the data acquired by stations settled on volcano edifices to avoid any bias due to site effects, especially for very long travel times associated to Peruvian events. After filtering, the remaining data, which are all prior to 2014, correspond to 10 332 events. Then, we decided to invert the increasing amount of data, that we received each year from the RENSIG, in a smaller box centered on Ecuador (see Figure 6.1). We also realized that there is no clear bias in the travel times for stations over volcanic edifices, except for few events within the edifice, due to the high velocity usually localized in the central part of the edifices crossed by most of the rays. We therefore took all the stations of this second box into account.

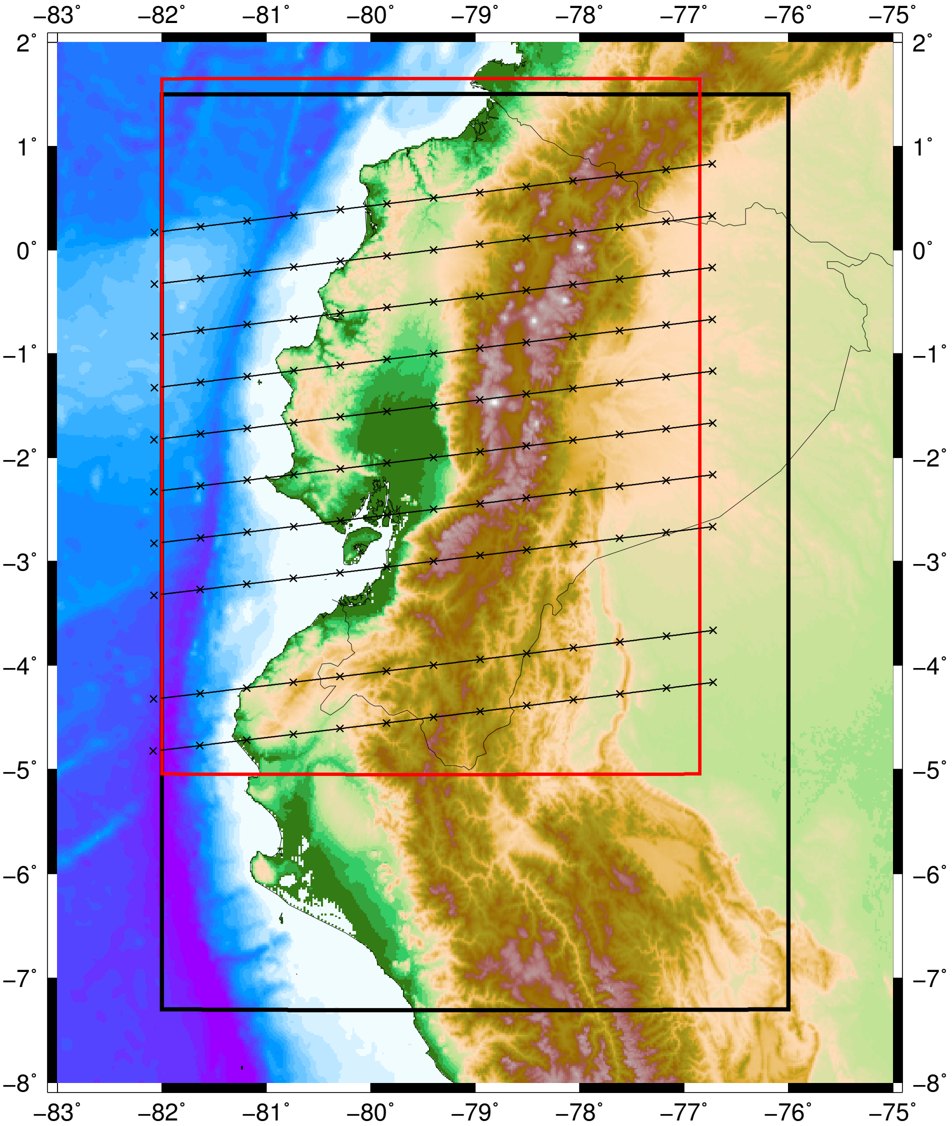

Figure 6.1: *The two inversion boxes are delimited by black or red lines. The traces of the cross-sections are indicated by thin lines with crosses every* $50km$*. The cross sections, with an azimuth of* $N83^oE$*, are centered at* -79.4^oW *every* 0.5^o *in latitude.*

Thus, after filtering, the resulting dataset (until April 2016) corresponds to 25 462 events.

To compare the results of tomography in the two boxes, we have first computed the restitution indices corresponding to these two types of inversion. Figures 6.2 displays these restitution indices for Vp over cross sections centered on the longitude of -79.4° W (Figure 6.1) with the direction of convergence of the subduction ($N83^oE$, Kendrick et al. 2003) as azimuth, and with a length of 600 km over all the width of the smallest inversion box (figure 1). The results, which are very similar for the other indices, clearly show that the resolving power of the data is greater for the smallest box north to 3^oS, and conversely for the largest box.

We can also see in Figure 6.2 that the models resulting of the inversion in the two boxes are very consistent at the latitudes for which they are both correctly resolved, that is between 0^oN and -3^oS down to $100 - 150km$ depth, eventhough the length of resolution is shorter for the more accurately determined model in the small box with three times more data.

The main contribution of the large box is south to -4°S where the resolution of the inversion in the Ecuadorian box begins to be poor (figure 6.2) and where the Peruvian data allow a better identification of the slab.

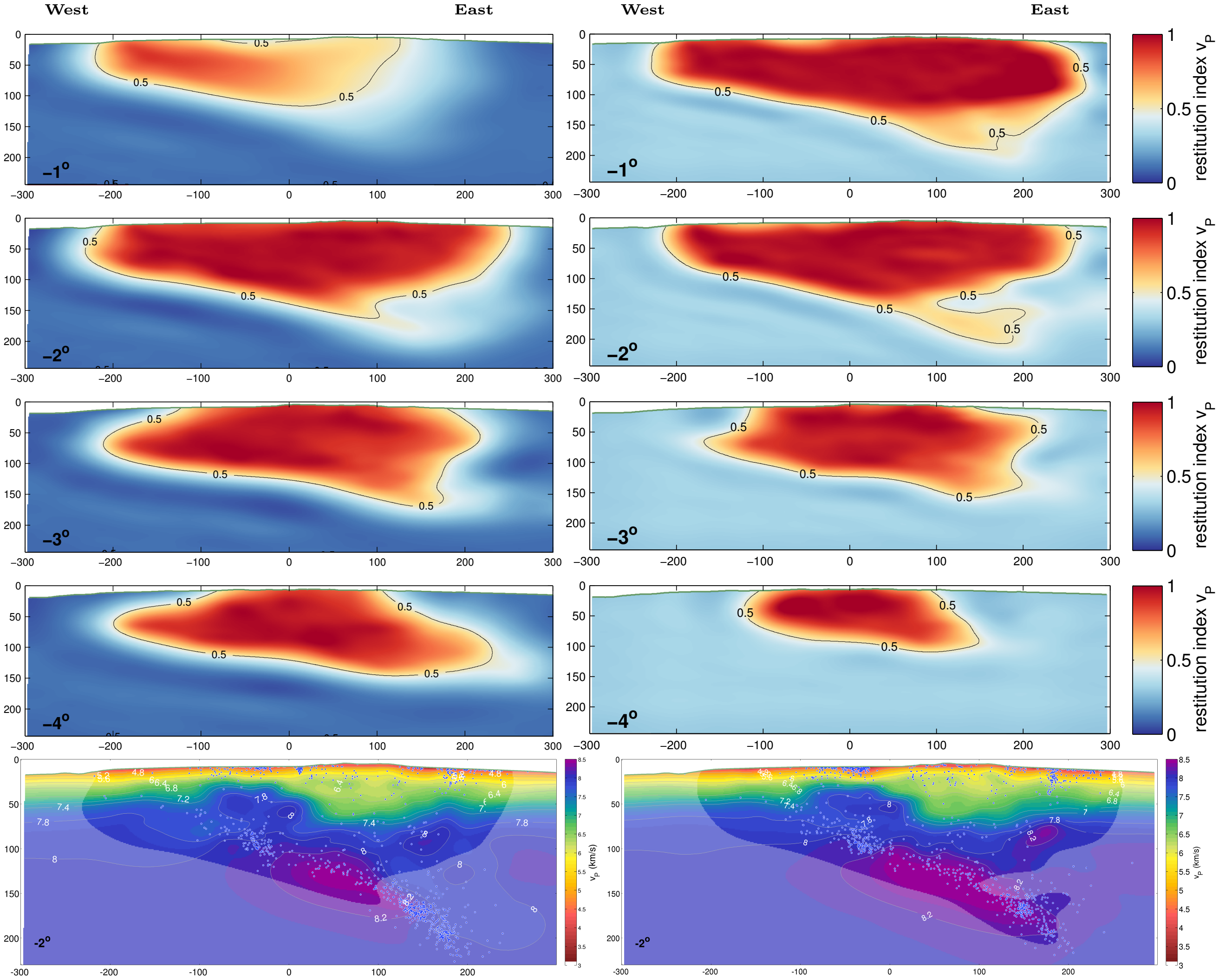

Figure 6.2: *Cross-sections of the restitution index (4 figures at the top of each column) and of P velocities (at the bottom). The results for the largest box including Northern Peru are represented in the left column, and that of the smallest one in the right column. The cross-sections are centered at 79.4 W and at the latitude indicated at the left bottom of each figure, with an azimuth of N 83° E. The confidence region, with an index greater than 0.5, are denoted by a greater transparency in the figures at the bottom*

For all these reasons, we will mainly focus on the results over the Ecuadorian box for interpretation.

6.2 Results for the models v_P and v_P/v_s

The cross sections of the P velocities, in absolute values and in relative values with respect to the prior model, and the ratio v_P/v_S are displayed between 0.5° N and 3.5° S in latitude (for the centre of the section) every 0.5° in figures 6.3-6.9. The seismicity indicated in each of these cross sections corresponds to the hypocenters located at less than 18 km from the section.

We can observe that the highest P velocities (greater than 8.2) are clearly localized within the slab south to 1.5° S in agreement with the usual conception of a cold-rigid plate downgoing into hotter and less dense surrounding materials (*e.g.* Stein and Wyssesion, 2003). Theoretical studies, taking into account a great variability of thermodynamical parameters of minerals in the Earth mantle, predict velocities v_P between 7.8 km/s to 8.5 km/s from 100 km to 300 km depth (e.g. Cammanaro et al., 2003). North to 1.5° S, a second domain arises in complement of this one at the East of the Wadati-Benioff zone, a little higher than the virtual continuation of this zone. We can also observe very high velocities in the mantle wedge between about the latitudes of 0° and 2° S, with a striking contrast with the low velocities beneath the Andean range. In the coastal region, the basins are clearly associated to low velocities in the upper crust.

The analysis of the v_P/v_S values is crucial to characterize partial melting and consequently the feeding process of the volcanic range. The generation of magma by the subduction of an oceanic plate follows a standard schema known as the genetic model (Daniel et al., 2006). Melting of hydrated mantle begins between 1000°C and 1100°C at a depth of about 100 km. This depth is not invariable but lies in the window between 65 and 130 km, depending on the structure of the volcanic arc (England et al., 2003). Actually, among all the parameters considered, the descent velocity of the slab is the one that shows the highest correlation rate with the volcano positions. Thus, considering this velocity as the leading phenomenological parameter amounts to considering that the partial melting of the mantle is produced at a critical temperature rather than a critical pressure (England et al., 2003). But this critical temperature depends on the quantity of water within the mineral. The water responsible of mantle hydration are fixed in specific minerals and transported down by the subduction process. These two minerals are the magnesian amphibole type pargasite and the magnesian mica type phlogopite. Once arrived to a depth of about 100 - 110 km, pargasite is destabilized and releases its incompatible elements and its water, which in turn causes the melting of the overlaying mantle wedge (Daniel et al., 2006). The partially melted materials produced by this magma genesis accumulate within large primary reservoir, elongate in the direction of the volcanic range at a depth of few tenth of kilometers beneath the volcano chains, as in Central Volcanic Andean Zone or in Java (Koulakov et al., 2009, 2011, Nur Abdullah, 2016). The partially melted materials are expected to present a low rigidity with respect to their bulk modulus, and thus can be characterized by great values of v_P/v_S. It allows the localization of these deep reservoirs along the range insuring the feeding of volcanoes.

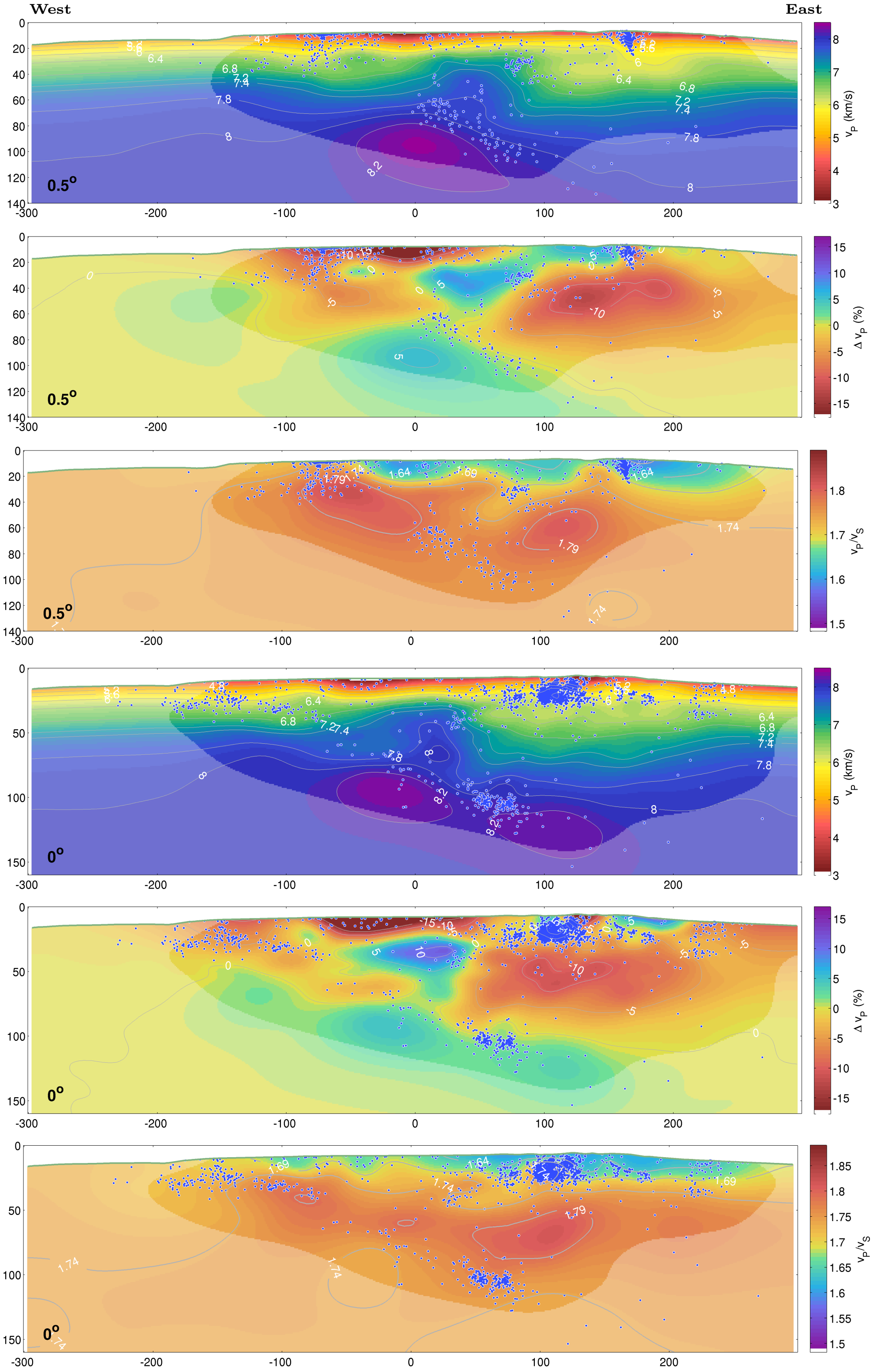

Figure 6.3: *Cross-sections of the P velocities in absolute values (km/s) or in relative values (%) with respect to the prior model, and of the v_P/v_S ratio. The cross sections are centered at 79.4° W and a latitude of 0.5° N (top) or 0° (bottom) with an azimuth of N 83° E. The two connected components of the Maldonado seismic nest can be easily identified at the latitude of 0° and at a depth of about 100 km. The scales are in kilometers and the confidence regions are denoted by a greater transparency*

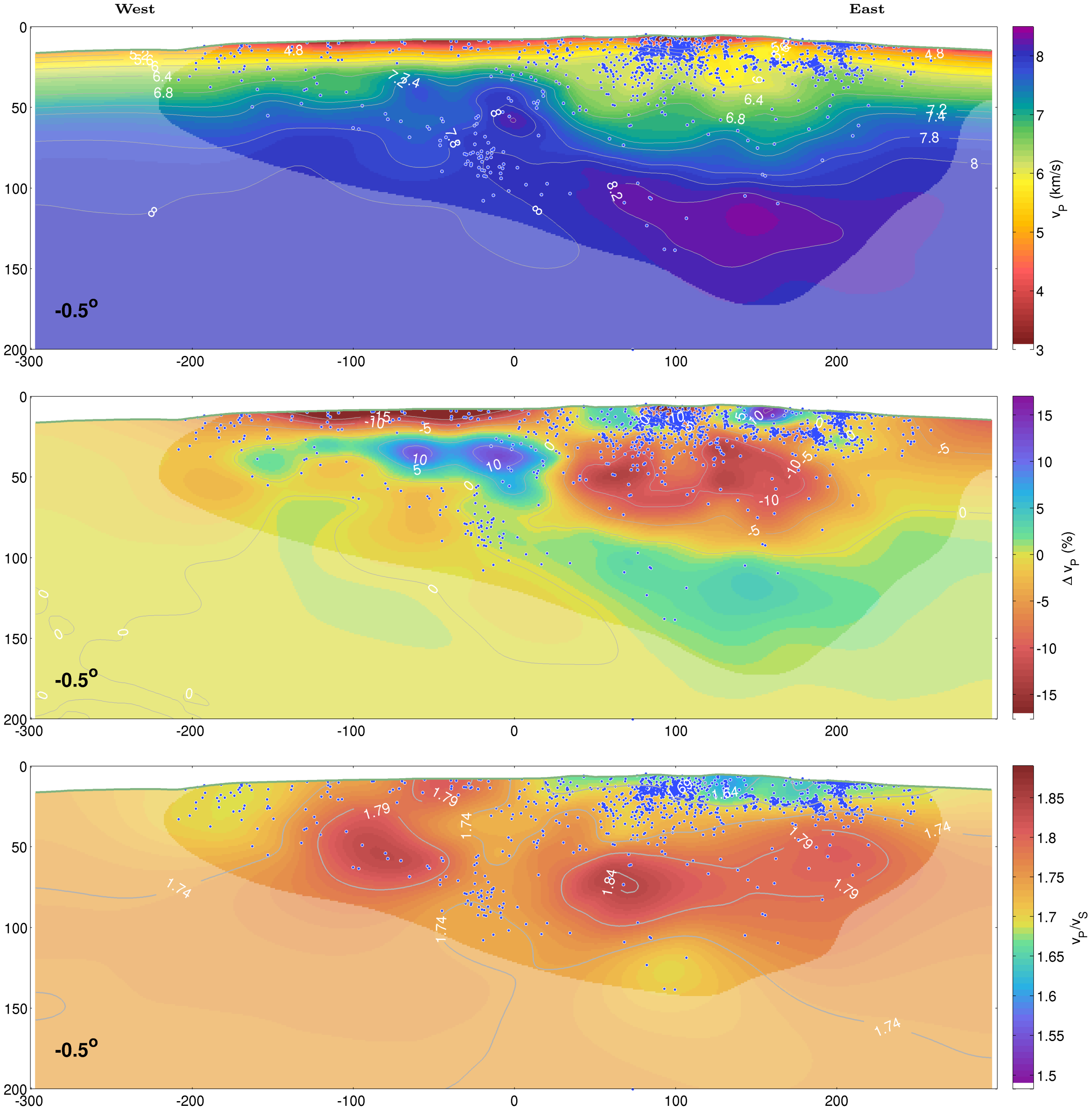

Figure 6.4: *The same as figure 6.3, except for the latitude of 0.5° S. Based on the Wadati-Benioff zone we could dream to a slab broken in 3 segments at this latitude.*

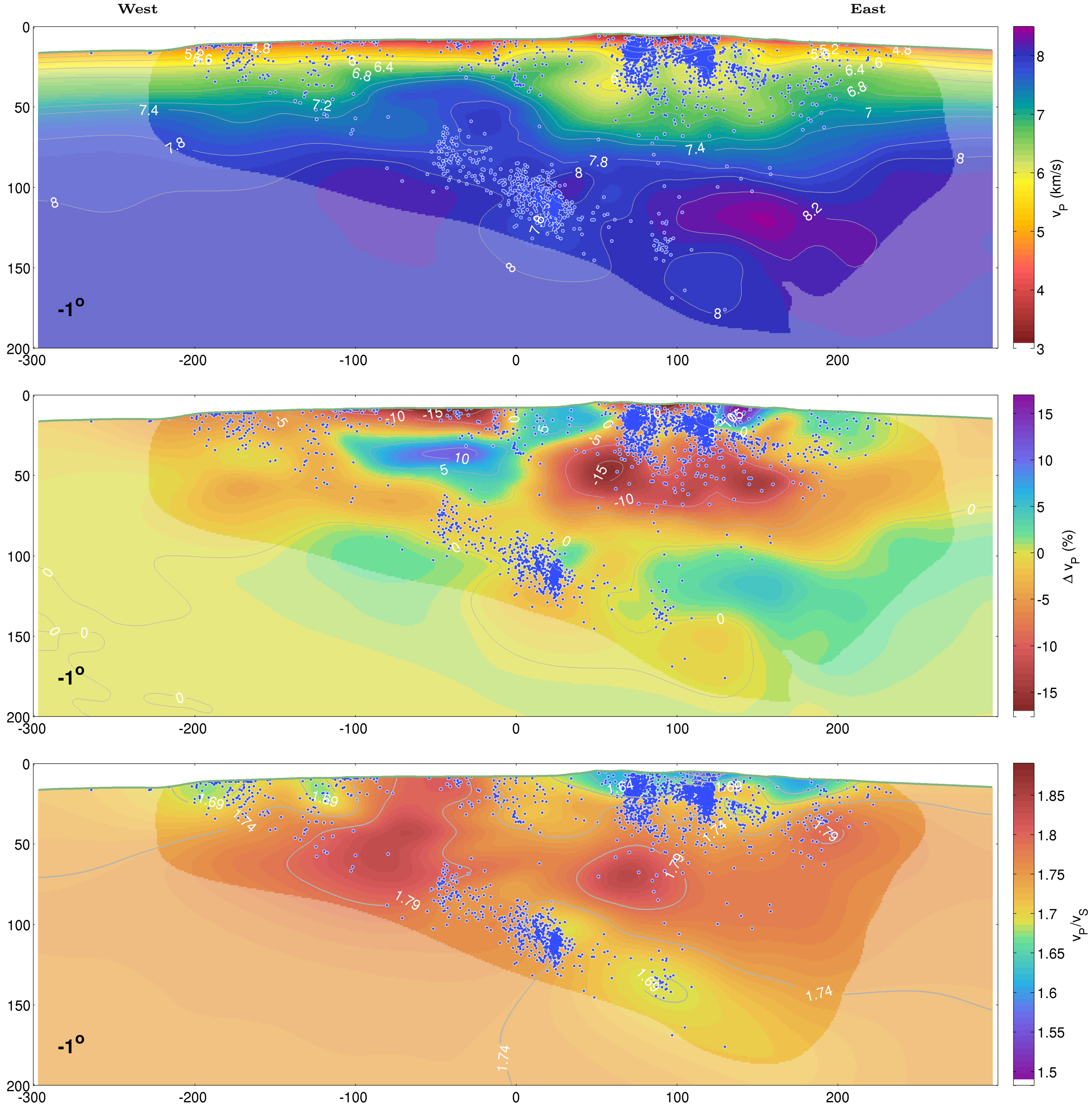

Figure 6.5: *. The same as figure 6.3, except for the latitude of 1° S.The La Maná seismic nest is clearly identifiable between about 110 km and 120 km depth. Note also the low values of the v_P/v_S ratio (about 1.69) within the slab at 135 km depth.*

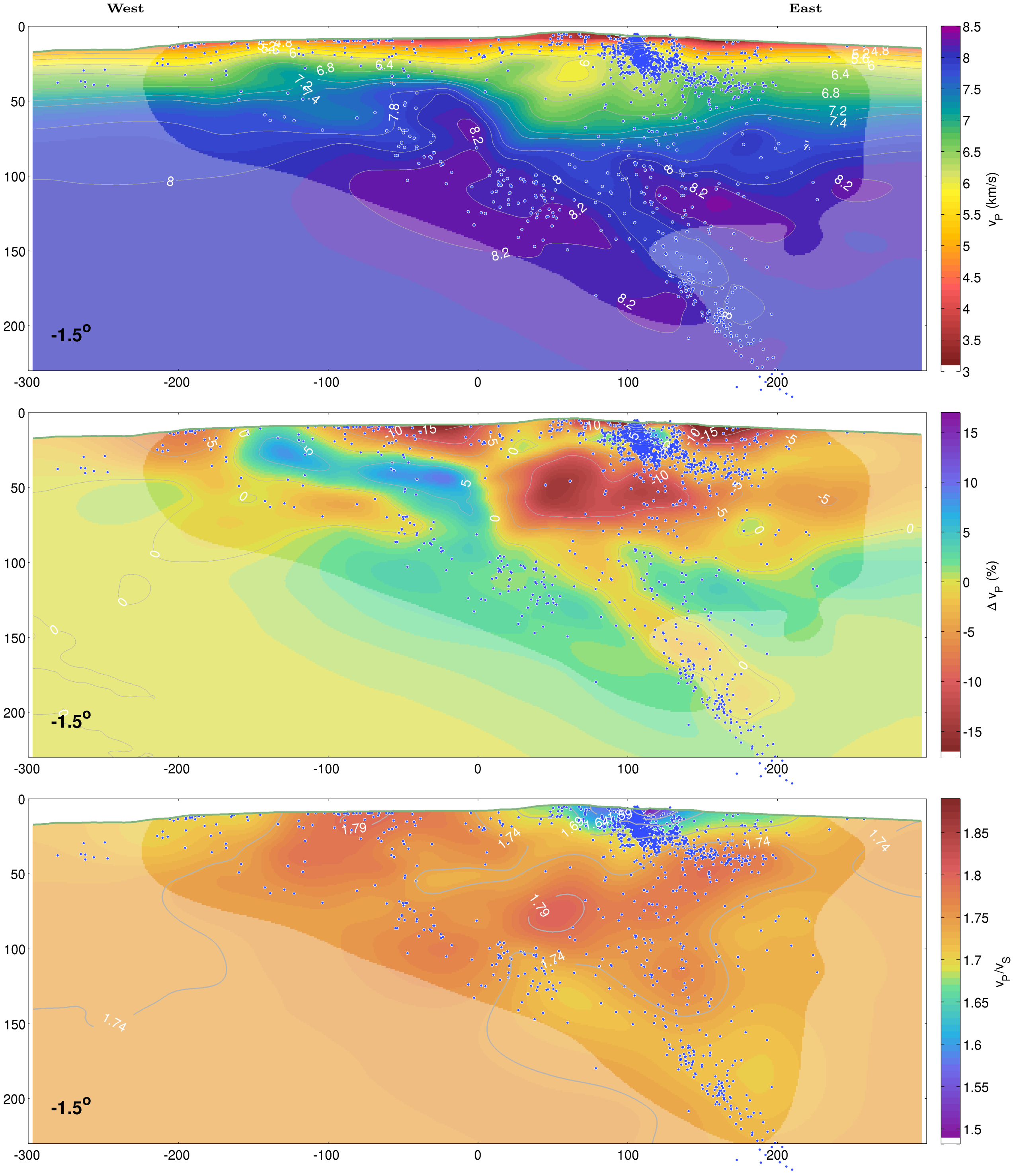

Figure 6.6: *The same as figure 6.3, except for the latitude of 1.5° S.The Puyo seismic nest is clearly identifiable between 165 km and 240 km depth, with a greater dip of about 40°. Low values of* v_P/v_S *are still discernible within the slab at about 135 km depth.*

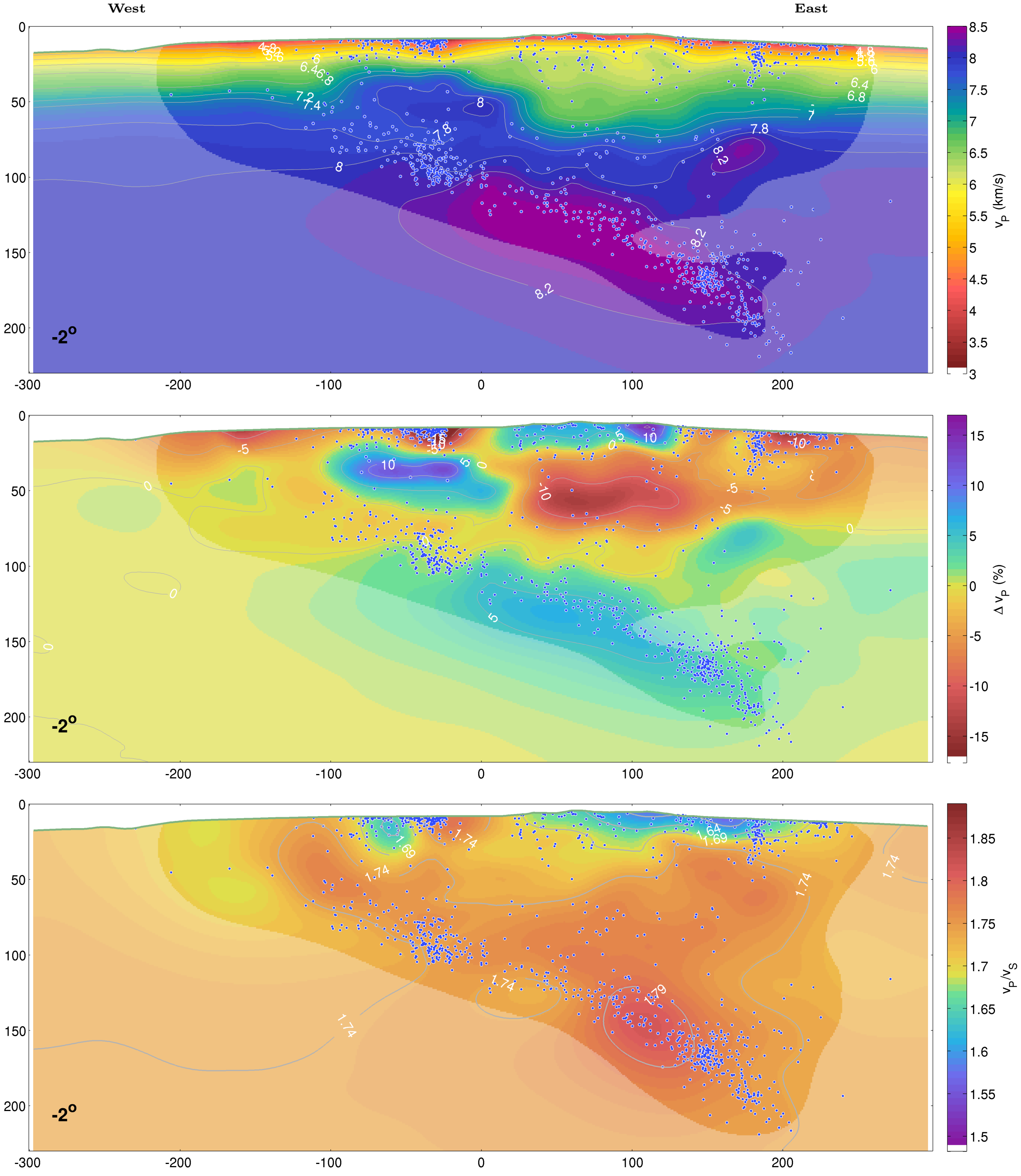

Figure 6.7: *The same as figure 6.3, except for the latitude of 2° S.The Puyo seismic nest is identifiable between 150 km and 210 km depth, with a lower dip than at 1.5° S. The Guyaquil nest is clearly identified at about 90 km depth. Low values of* v_P/v_S *are still discernible within the slab at about 135 km depth.*

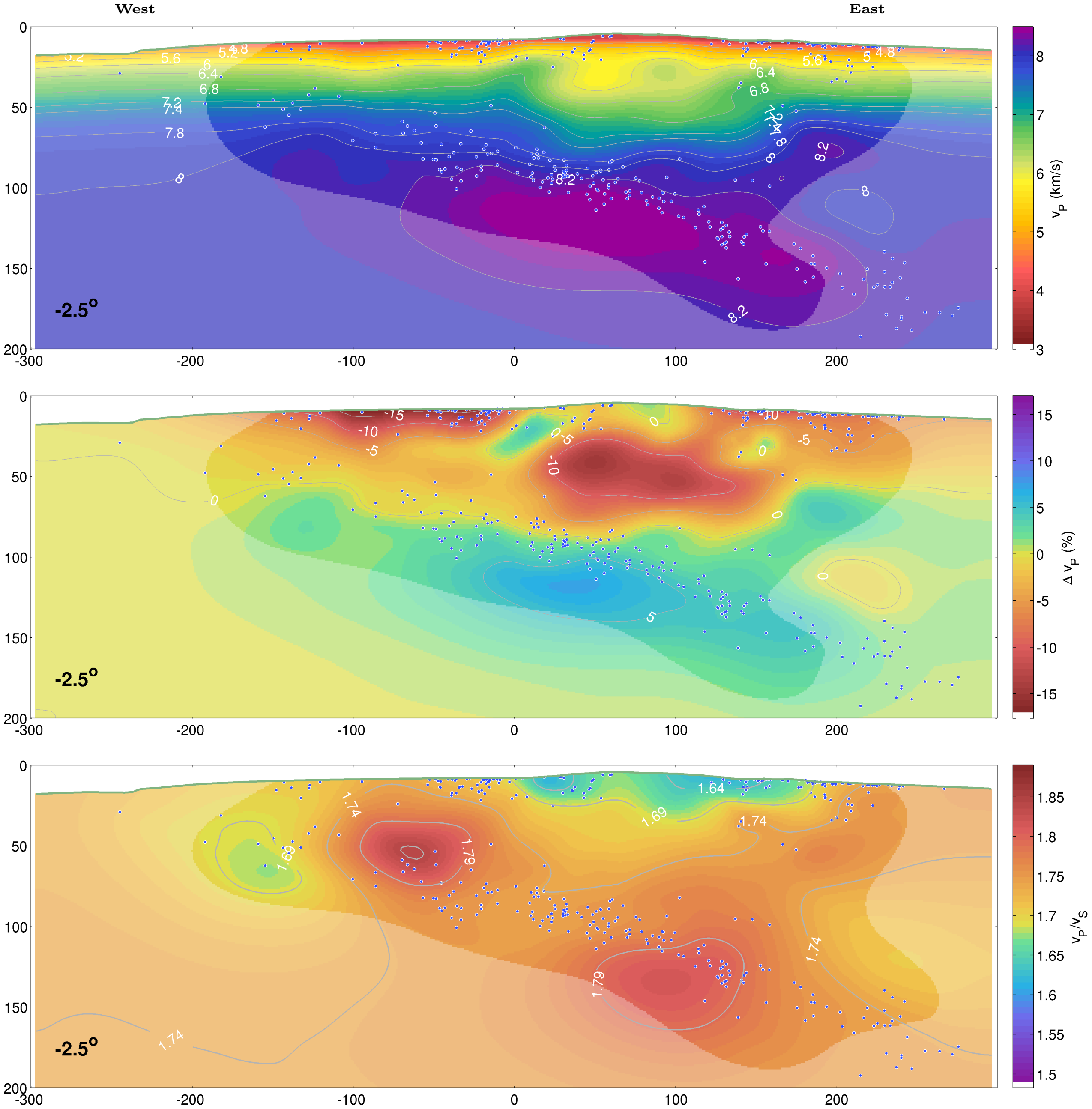

Figure 6.8: *The same as figure 6.3, except for the latitude of 2.5° S.*

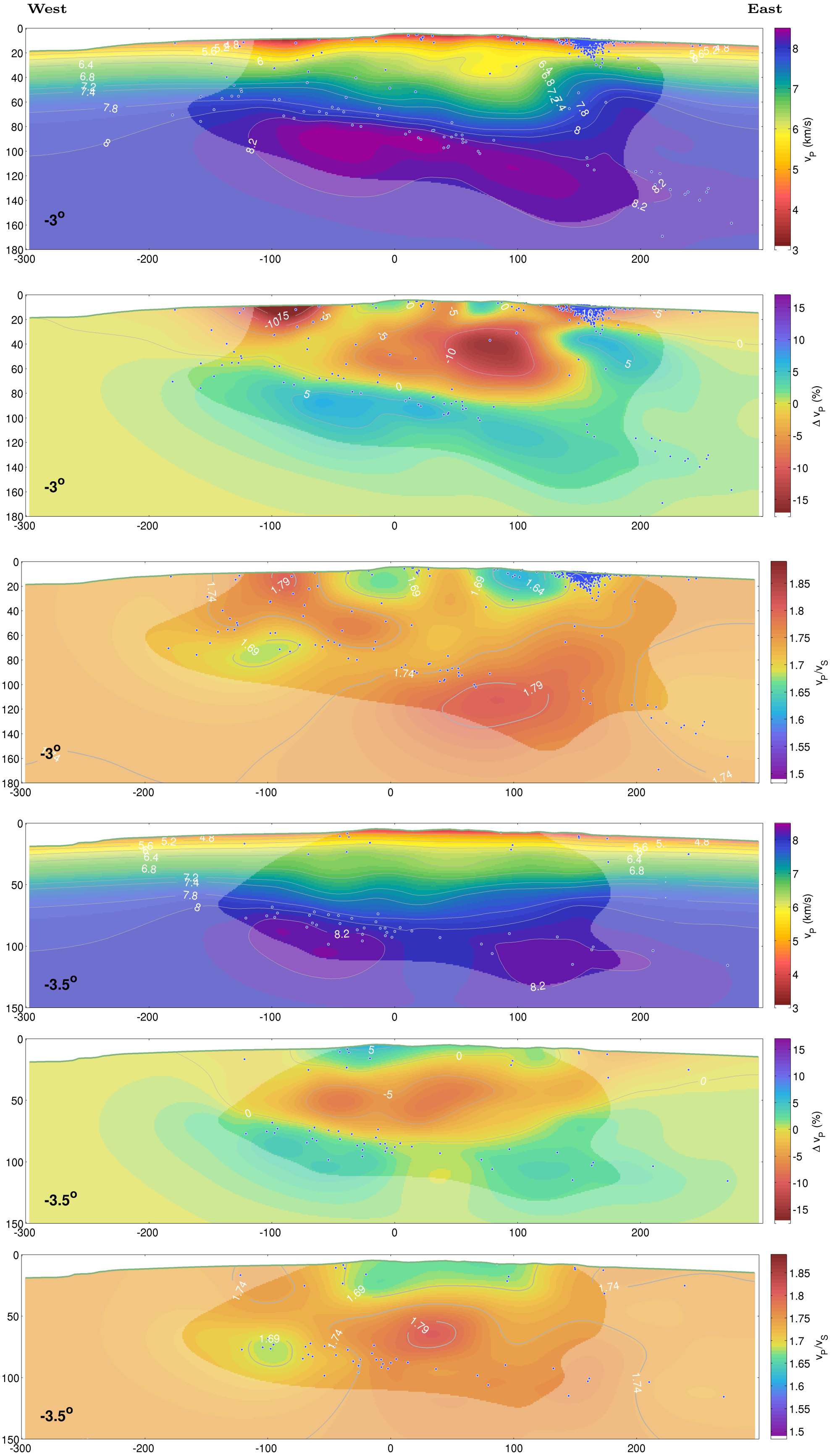

Figure 6.9: *The same as figure 6.3, except for the latitude of 3° S at the top and of 3.5° S at the bottom.*

The cross sections along the volcanic range that are displayed in figures 6.10, 6.11, 6.12, 6.13 clearly show great values of the v_P/v_S ratio within an elongate zone between 2° S and 0.5° N at depths between 35 and 80 km. We can observe that the south limit of this zone, which is lateraly centered on the western cordillera, coincides with the latitude (2.005° S) of the Sangay that is the southest volcano of the arc, as can be seen in figures 6.11, 6.12. To precise the spatial distribution of the high values of v_P/v_S, the domains corresponding to values greater than 1.8 are displayed in figure 6.14, and those to values greater than 1.82 in figure 6.15.

It confirms that the greastest values of the ratio correspond to the reservoir beneath the western cordiera. We can also observe that apart from two small domains relativelly shallow, the one localized at about 3.5° S and 79.2 W and the other at the East of the Antisana volcano, the other large connex domain where the ratio is greater than 1.8 is located in the portion of the slab at about 150 km depth south to 2° S that is continuously connected to the flat Peruvian slab (Figure 6.14).

We can also observe low values for the velocity ratio in the uppermost part of the crust along the volcanic range that characterize cold volcanic materials. A relativelly low value of this ratio is also discernible within the slab at a depth of about 135 km between 0.5° S and 2°S.

Finally the figures 6.11, 6.12, 6.13 show that the part of the Andean range which corresponds to the volcanic arcs clearly coincide with a thickening of the crust suggesting an involvement of the magmatism in the orogenesis of this part of the Andes.

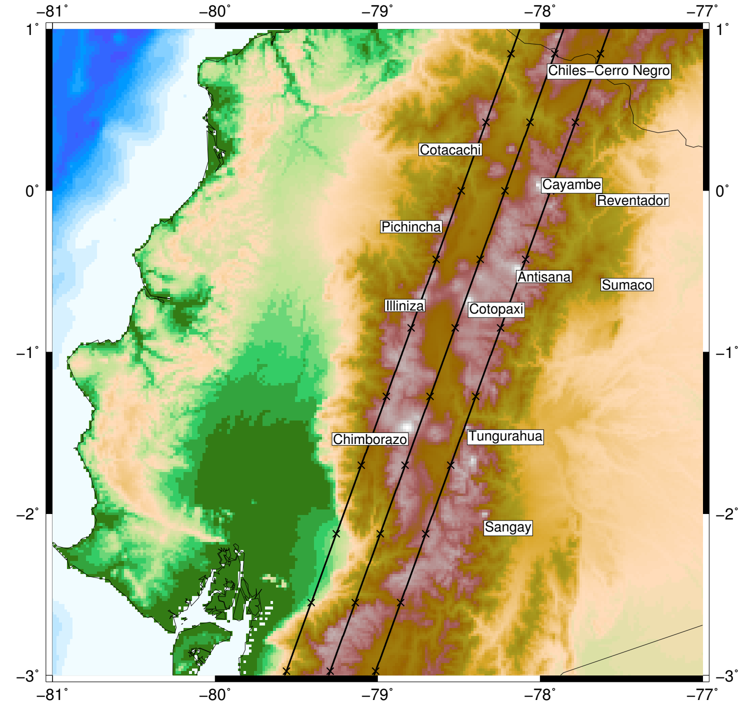

Figure 6.10: *Position of the traces of the 3 cross sections along the Andean range centered at the latitude of 1.7° S with an azimuth of N 20° E and with a cross every 100 km. The longitudes of the 3 centres are respectively 78.55° W, 78.83° W and 79.1° W.*

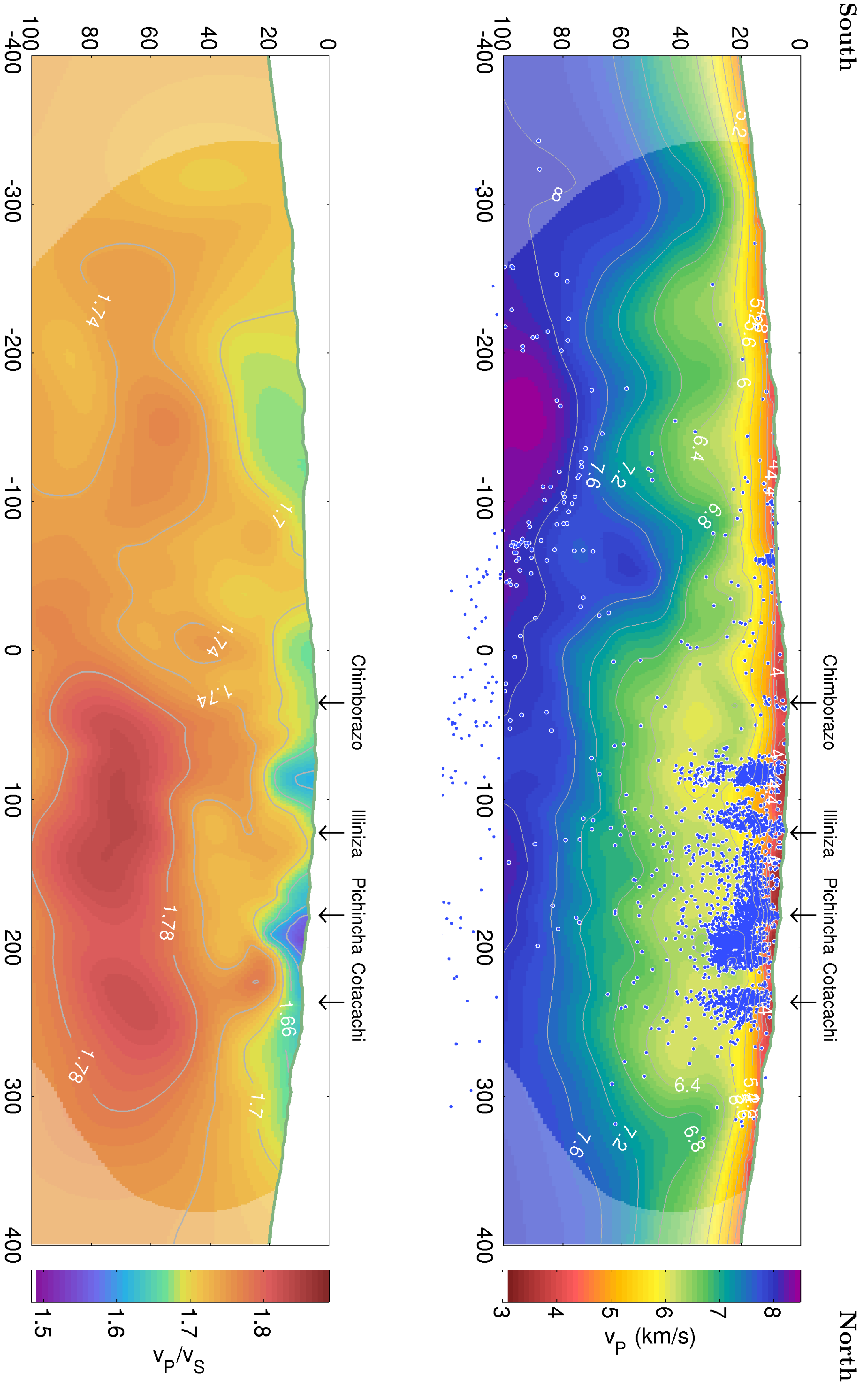

Figure 6.11: Cross section along the Western Cordillera centered at 1.7° S and 79.1° W with an azimuth of N 20° E. The scale is in kilometers

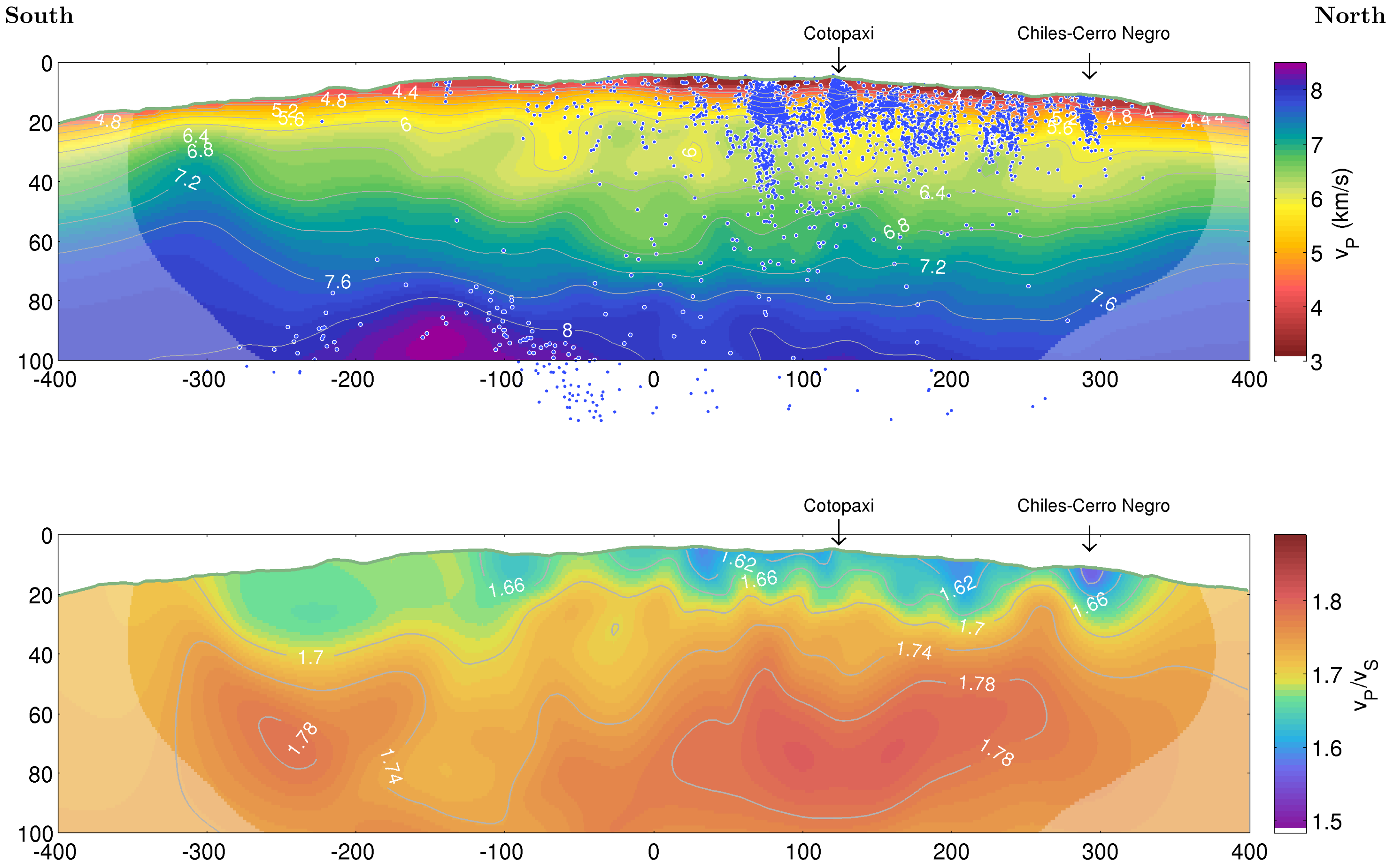

Figure 6.12: *Cross section along the Interandean zone centered at 1.7° S and 78.83° W with an azimuth of N 20° E. The scale is in kilometers*

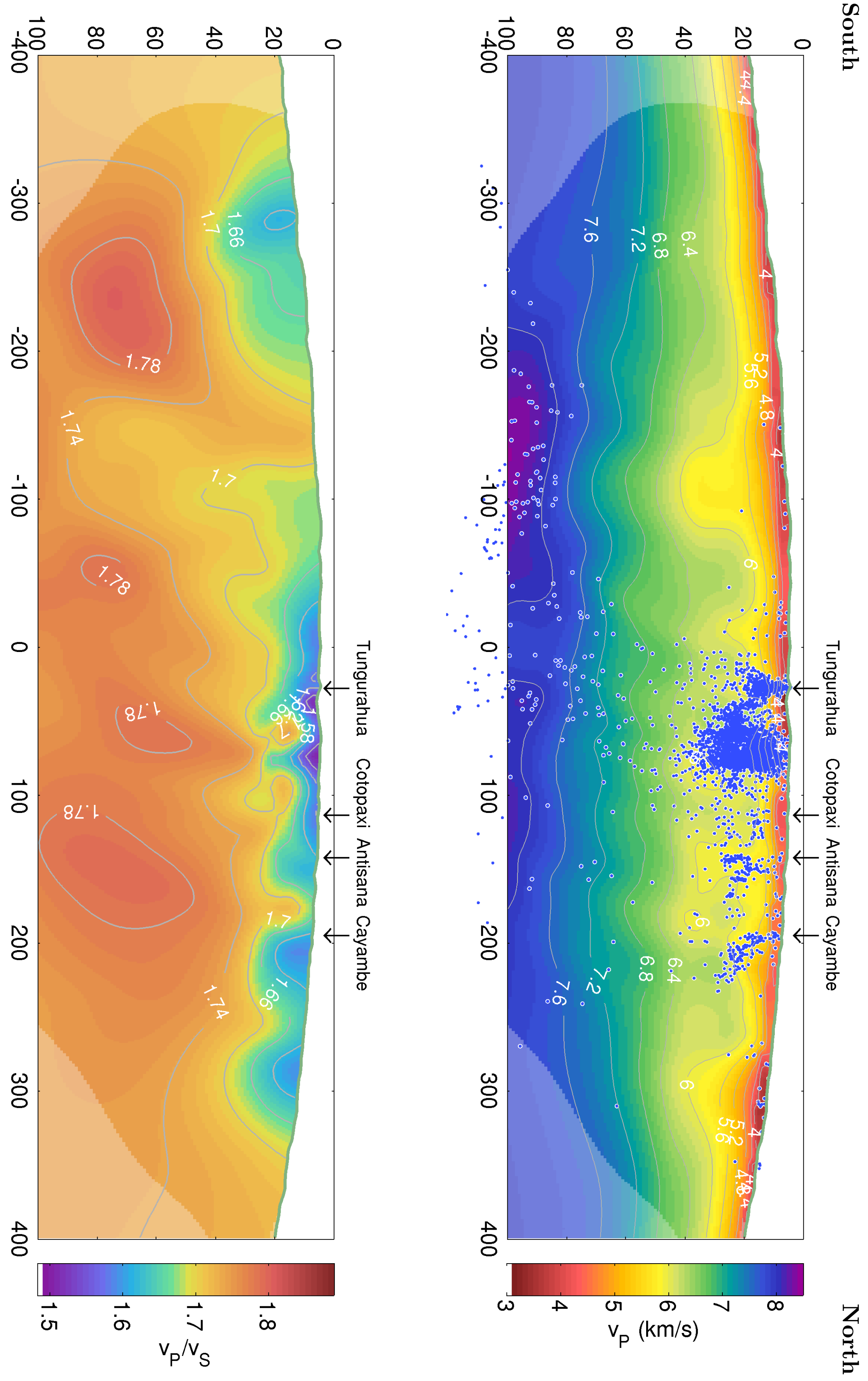

Figure 6.13: *Cross section along the Easttern Cordillera centered at 1.7° S and 78.55° W with an azimuth of N 20° E. The scale is in kilometers*

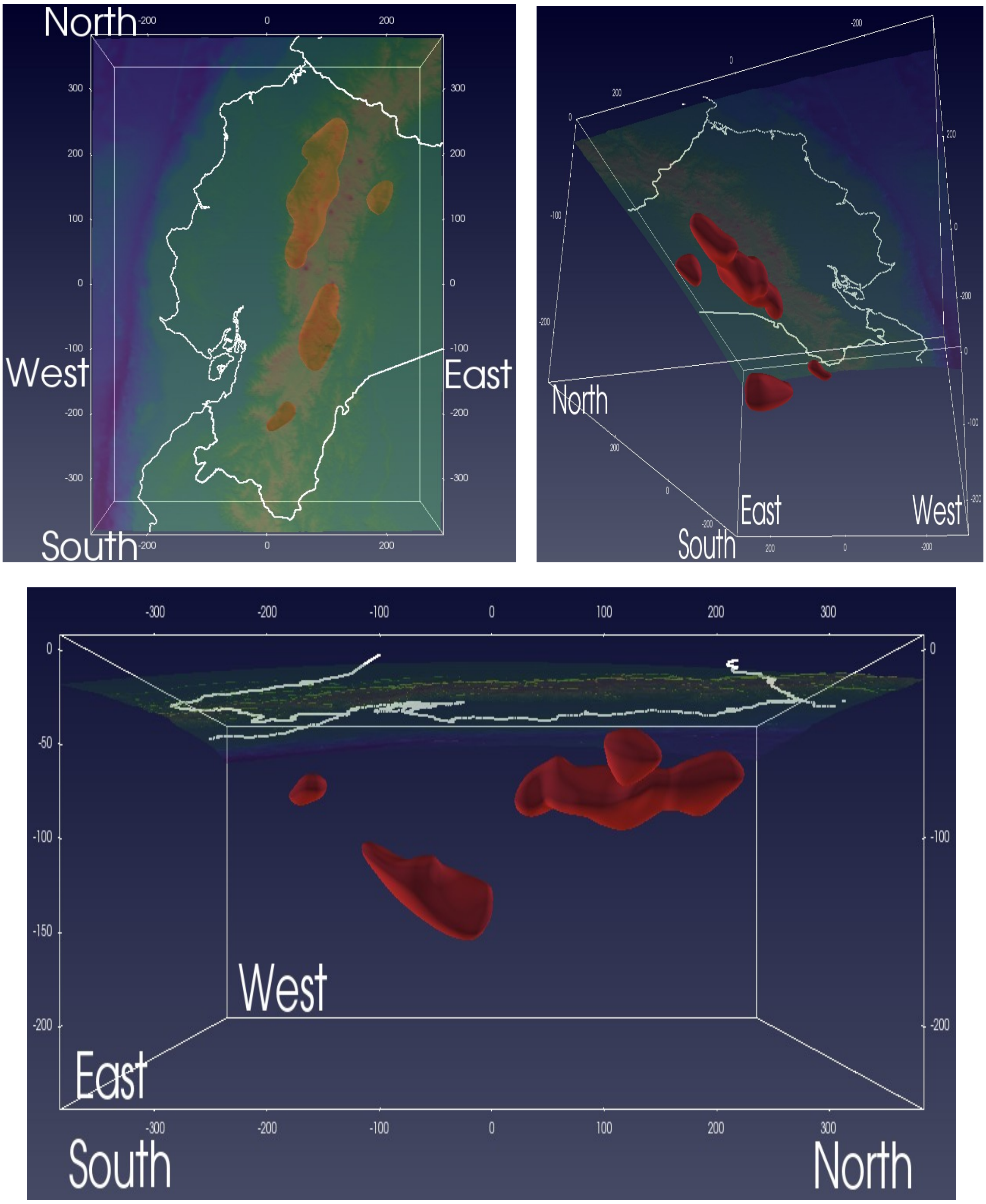

Figure 6.14: *The 4 domains corresponding to* $v_P/v_S \geq 1.8$

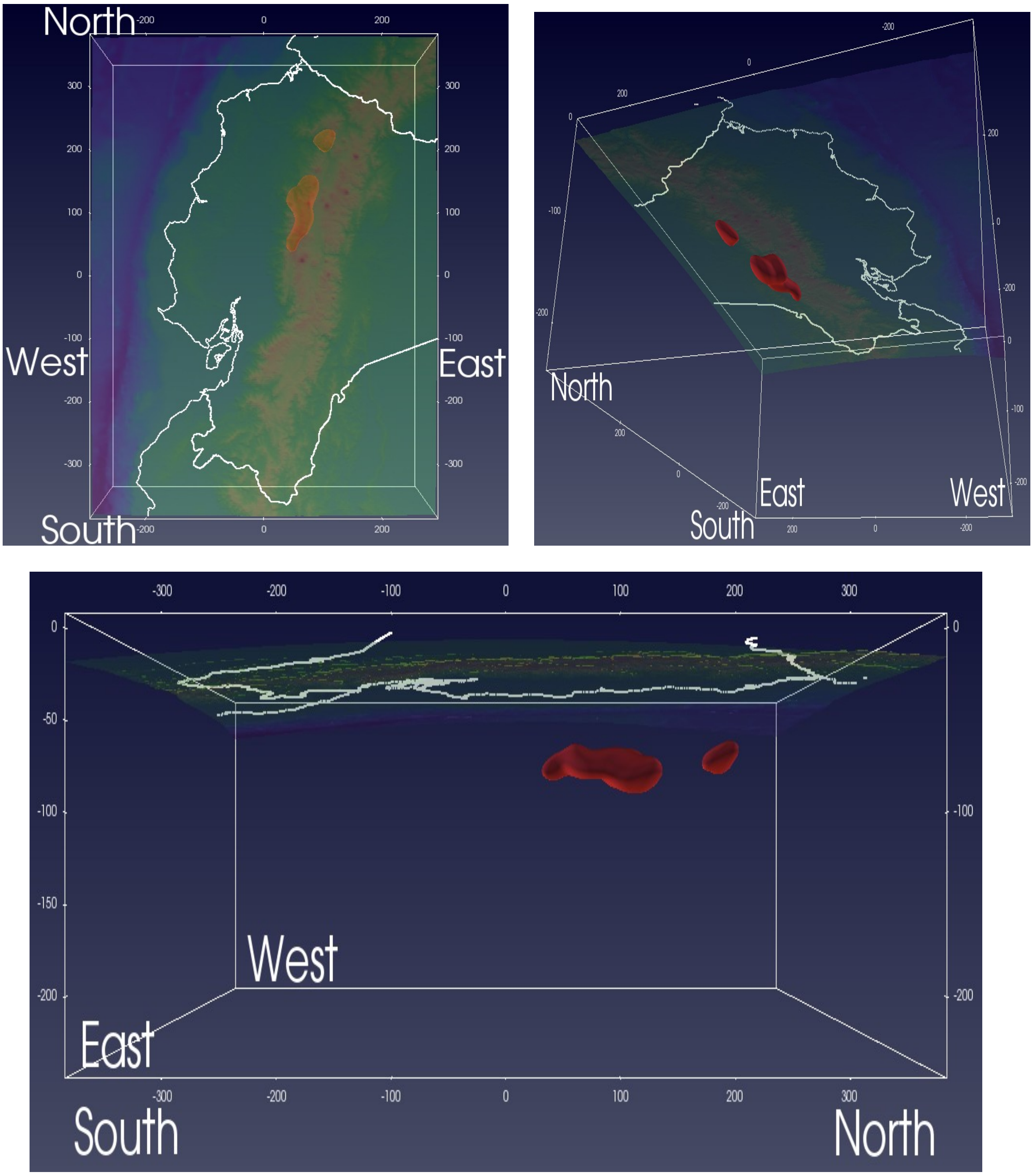

Figure 6.15: *The 2 domains corresponding to* $v_P/v_S \geq 1.82$

6.3 The localization of the seismicity

Since the arrival times are inverted for both the velocity fields and the event locations, we obtain, as a result, new locations of the seismicity. The figures 6.16 and 6.17 respectively display the spatial distribution of the seismicity shallower or deeper than 75 km depth. The general pattern of these distribution is evidently similar to the one obtained a priori in chapter one and three, but nevertheless, the inversion enhances the accuracy of the locations and allows a better resolution of this pattern.

The figure 6.17, which displays the seismicity at the so called intermediate depths and that is mainly located in the slab, allows the clear identification of 4 seismic nests that form an almost perfect geometric figure, with 3 of them, Maldonado, La Maná and Guayaquil at depths between 80 and 115 km, and the last one, Puyo, at deeper depths, between 160 and 240 km. We will come back to the detail of these seismic nests further on when describing the slab geometry and will now focus on the shallower seismicity.

The seismicity of the margin is concentrated in clusters relatively perpendicular to the coast line between about 1.5° N and 1.5° S and that define a clear segmentation of the margin in relation with the carnegie ridge which has been the subject of numerous experiments and studies (*e.g.* Manchuel et al., 2011; Font et al., 2013). The shallow seismicity in the Cordilleras region can be explained by three sources, the volcanic crisis, the current volcano-tectonic events and the tectonic deformations. The volcanic crisis that occured during the study period (table 6.1) concern the Chiles-Cerro Negro (November 2014 to present), the Guagua Pichincha (1991-2001), the Reventador (2002 to present), the Cotopaxi (August 2015 to present) and the Tungurahua (1999 to present). This kind of seismic activity, which, though spatially concentrated, is not permanent, must be catalogued as seismic swarms.

There are other shallow seismic clusters beneath volcano edifices that clearly correspond to volcano-tectonic activity (see also figure 6.11) as for the Cotacachi in the north, the Illiniza, and the Atacazo-Ninahuilca and the Corazón located between the Pichincha and the Illiniza.

Volcano	**Activity**	**Latitude**	**Longitude**
Chiles-Cerro Negro	Hydrothermal	$0.817^o N$	$-77.938^o W$
Cotacachi	None	$0.364^o N$	$-78.351^o W$
Reventador	Strombolian, vulcanian	$-0.077^o S$	$-77.656^o W$
Guagua Pichincha	Lava dome, ash explosions	$-0.171^o S$	$-78.598^o W$
Atacazo-Ninahuilca	None	$-0.353^o S$	$-78.617^o W$
Corazón	None	$-0.539^o S$	$-78.669^o W$
Cotopaxi	Phreatic, ash explosion	$-0.677^o S$	$-78.436^o W$
Illiniza	None	$-0.659^o S$	$-78.714^o W$
Tungurahua	Strombolian, vulcanian	$-1.467^o S$	$-78.442^o W$

Table 6.1: *Volcanoes related to seismic clusters during the study period.*

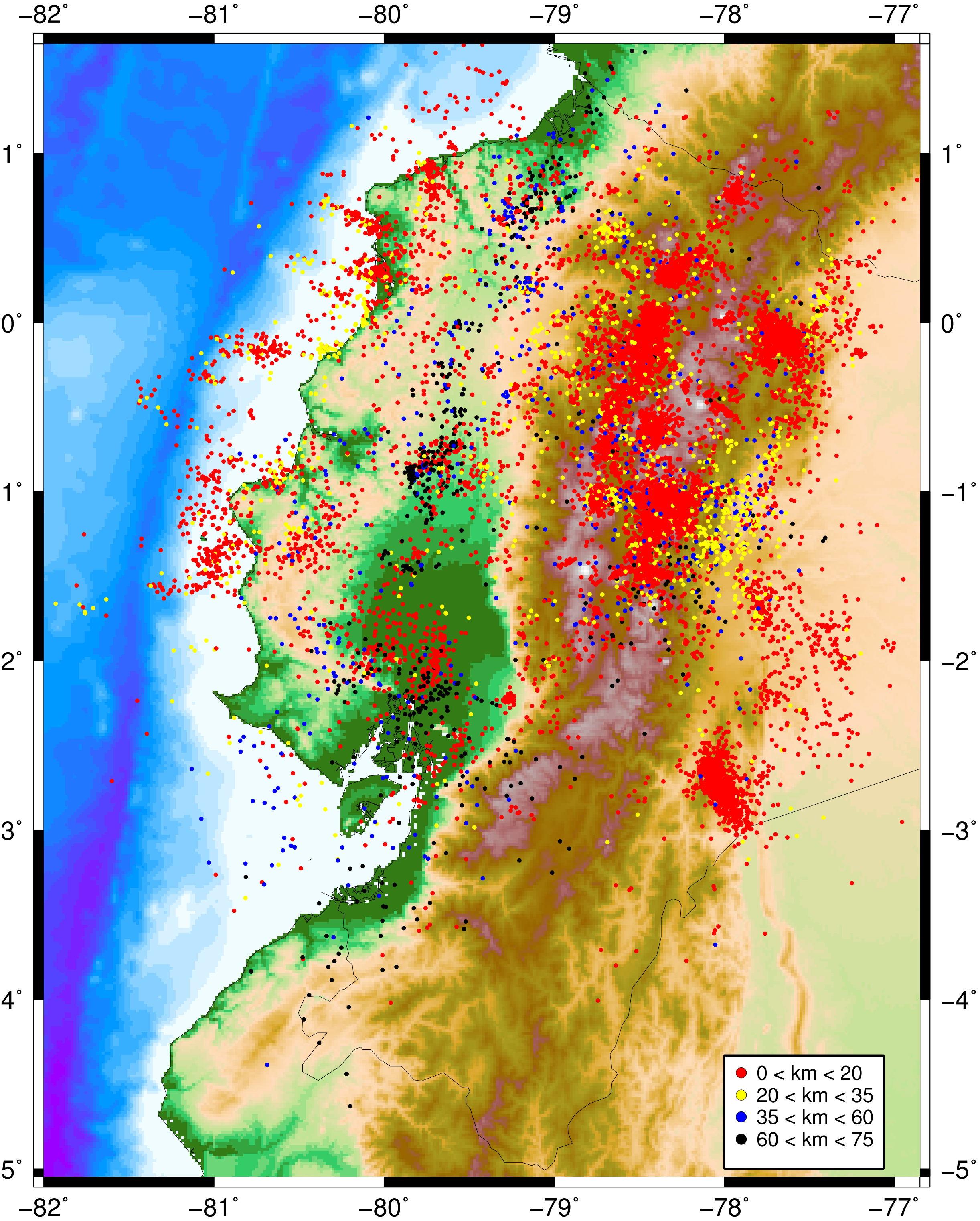

Figure 6.16: *Final localization of the seismicity shallower than 75 km depth.*

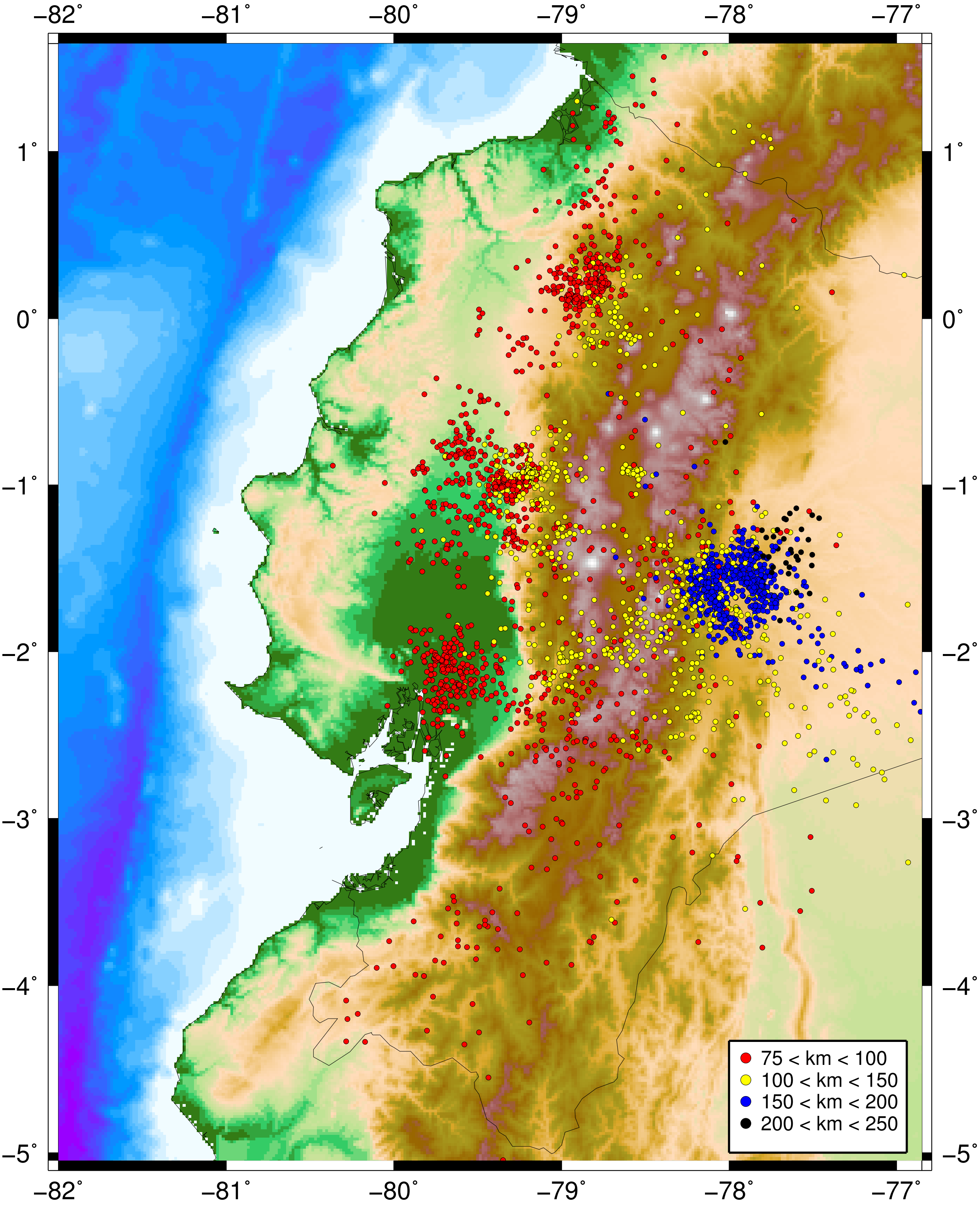

Figure 6.17: *Final distribution of the intermediate seismicity.*

But the intense shallow seismic activity around volcano edifices along the Western Cordillera must be related to the Quito-Latacunga fault-fold system that defines the western edge of the Quito-Latacunga microblock (Alvarado et al., 2016). In figure 6.16 we can clearly follow the seismicity of this Quito-Latacunga fault-fold system down to its connection, south of the Chimborazo, with the Chingual-Cosanga-Pallatanga-Puna fold system (CCPP), and observe that it contains not only this volcano-tectonic activity but also some purely tectonic clusters as the one located (around 1.08° S, 78.4° W) south of the city of Pujilí.

There are also several large shallow clusters clearly related to tectonic deformations. The Pisaymbo cluster (at about 1° S, 78.5° W) which is the principal source of released seismic energy in Ecuador. A preliminary relocalization of the event cluster by using a double difference approach allowed us to determine a mean N 70° E principal direction of the cluster (Araujo, 2009) that approximately corresponds to the local strike of the CCPP that crosses the cluster. Another important cluster is the one located South-east

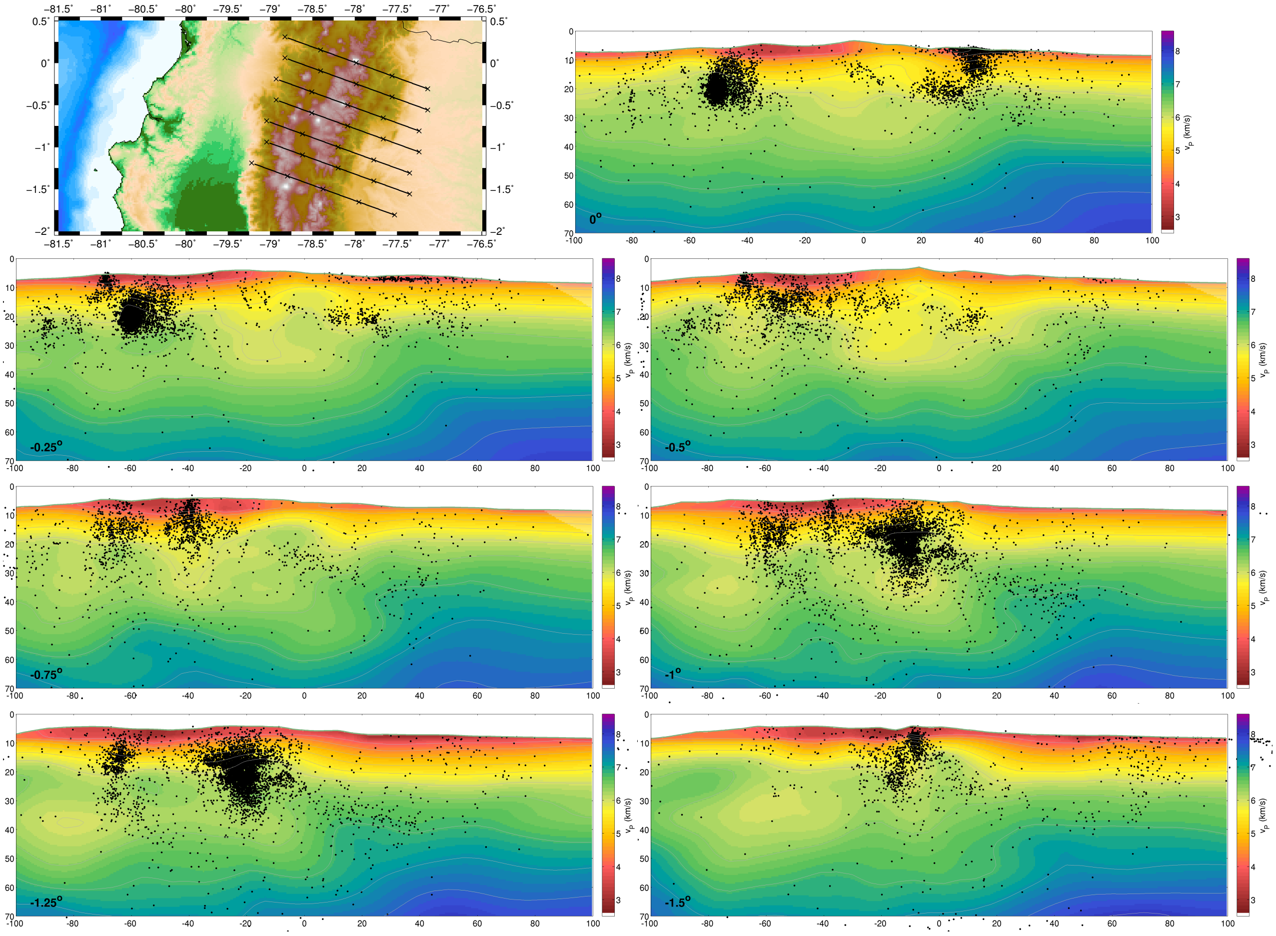

Figure 6.18: *Cross sections of the seismicity and P velocities in the crust, perpendicular to the cordilleras, between 0° and 1.5° S with an azimuth E 20° S. The latitude indicated at the bottom left of each figure corresponds to that of the centre of the cross-section. The traces of the sections are represented in the figure at the top left*

of the city of Macas and centered at approximately 2.7° S and 78° W. This cluster (figure 6.16) has a principal direction of N 10° W, orthogonal to the direction of subduction convergence, which corresponds to the one observed (Figure 6) by Legrand et al. (2005) during their 1999-2000 experiment involving 10 seismic stations. This cluster is related to Napo-Cutucú fault system and is basically generated by the thrust corresponding to the eastwards continental wedge propagation at the Eastern Subandean Belt (ESB) (Legrand et al., 2005; Alvarado et al., 2016). We can clearly follows in figure 6.14 the seismicity related to the ESB. We can also observe that the cluster located at about 0° and 77.7° W with a principal azimuth of N 40 ° W, and which contains the volcano-tectonic seismicity of the Reventador volcano, is also clearly related to the ESB. Another smaller tectonic shallow cluster is discernible in figure 6.14 north of Guayaquil, in the south part of Progresio basin.

Clear lineaments in the seismicity are also discernible in figure 6.16, as the CCPP system (see Alvarado et al, 2016, figures 1 and 6a) notably near the Guayaquil gulf and near Pallatanga (see also the cross section, figure 6.11 at the abcissa about -70 km). We can also observe a seismicity lineament starting from the CCPP system East-North of Pallatanga (small cluster) and bordering the Western cordillera at west up to Colombia.

In order to have a look at the sutures related to the accretion of the oceanic terranes (Guiller et al. 2001), we zoom into the crust model in figure 6.18 that displays cross-sections, perpendiculary to the cordilleras, of the seismicity and P velocities. We can observe a clustering of the seismicity around a plan that dips westward, with an increasing angle between 0° and 1.5° S. We can also observe that this plan is associated to a dipping westward of the lower crust as in a continental collision setting, notably around the latitudes of the Pisayambo cluster (between 0.75° S and 1.25° S°).

6.3.1 Evaluating the localization of some specific events

In order to evaluate the impact of the tomography on the localization of the earthquakes, and as we cannot consider 25 462 events one by one, we will focus on 5 sufficiently well observed events that occured during our study period and are of specific interest, either because of their large magnitude or tectonic setting, or with respect to seismic hazard. Notably, we will compare our prior and final localisations, respectively obtained with the LOCIN and the INSIGHT code, with those provided by the *Instituto Geofísico* (IG) and by the *United States Geological Survey* (USGS).

- The Pujilí earthquake ($M_w = 5.9$) occured the 28th March 1996 in the Interandean region, 100 km south of Quito. The earthquake, which produced important damage in adobe buidings, was located at 1.02° S, 78.78° W at a depth of 15 km by the IG (Guéguen et al., 1998), while the USGS localized it at 1.04 S, 78.74 W and at a depth of 33 km. Our localization after tomography at 1.03° S, 78.70° W and a depth of 28.5 km is very close to the prior one, at the western limit of the interandean depression (figure 6.19).

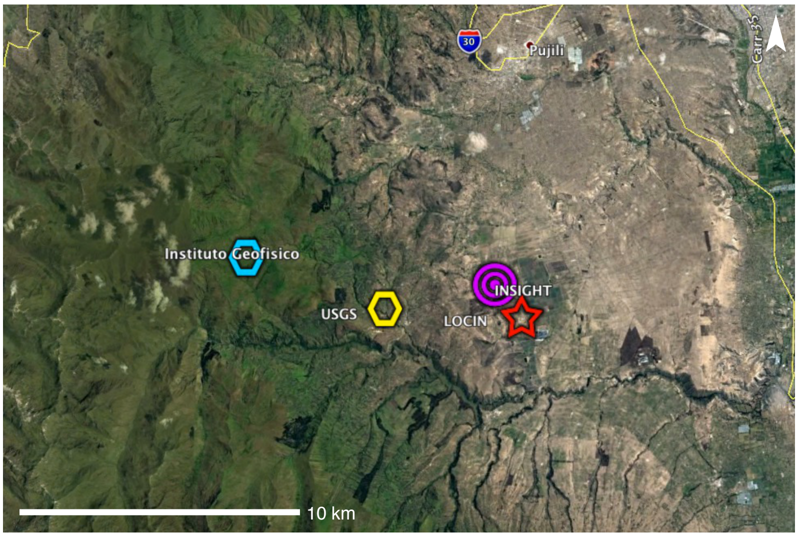

Figure 6.19: *The various Pujilí earthquake locations. Our localization after tomography (1.03° S, 78.70° W, depth of 28.5 km) put the hypocentre at the limit of the Interandean depression at an intermediate depth between those determined by the IG and by the USGS.*

- On the 4th August 1998 an earthquake ($M_W = 7.1$) occurred near the city of Bahía de Caráquez in an area of intense seismic activity which lies between Cabo Pasado and Manta (about 0.25°S - 0.75° S) in the Ecuadorian margin. This segment of the coastal margin produced two other large earthquakes over the last century, the 1896 ($Ml = 7.0$) and the 1956 ($M_S = 7.3$) events (Segovia et al., 1998). The IG localized the earthquake at 0.55° S, 80.53° N and at a depth of 20 km, while the USGS gave 0.59° S, 80.39° W and a depth of 33 km. Our localization after tomography is in between: 0.56° S, 80.47° W, at a depth of 28.7 km (figure 6.20).

Figure 6.20: *Bahía earthquake locations. Our final location (0.56° S, 80.47° W, depth of 28.7 km) lies in between those of the IG and USGS.*

- On the 12th August 2014 occured an eartquake ($M_W = 5.1$) north-west of Calderón city near Quito, with a maximum intensity of 6 *ems* corresponding to slight damages as thin cracks in walls and masonry (IG, 2014).

  The source of this seismic event can be related to the Quito Fault System (IG, 2014), over which the compressive deformation is localized in the Interandean Depression and is now considered as the western limit of the Quito-Latacunga microblock (Alvarado et al., 2016). The Quito Fault System is also connected with the Guayllabamba Fault System at the east in the north of the city (Alvarado et al., 2014). The survey of these fault systems close to Quito is obviously crucial in regard of seismic hazard for the city.

  Our location of the earthquake (0.06° S, 78.42° W, depth of 14.9 km) is about 4 km westwards from the localisation of IG (0.06°S, 78.38° W, depth 4 km) and 10 km deeper, while the USGS put the hypocenter 8 km north-east of the IG location (0.02 S, 78.32 W) at a depth of 11.9 km (figure 6.21).

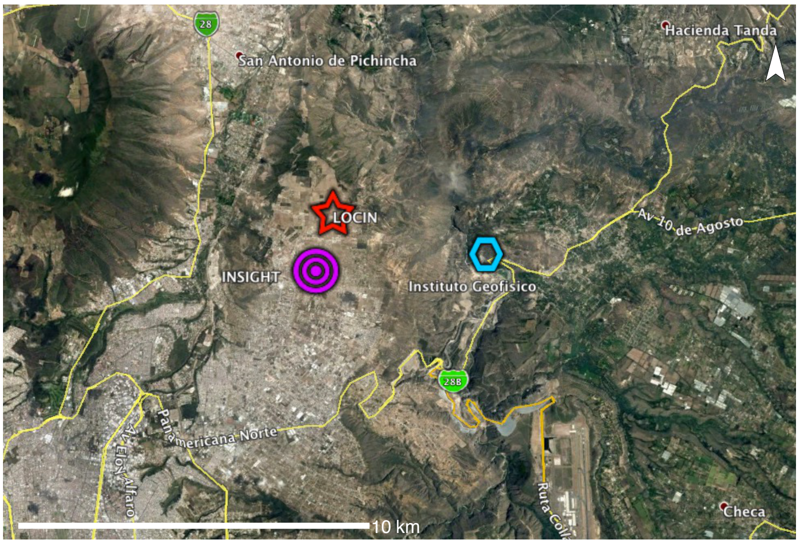

Figure 6.21: *Locations of the Quito (12th August 2014) earthquake. Our location after tomography is 0.06° S, 78.42° W, at a depth of 14.9 km.*

- The Pisayambo seismic cluster constitutes the greatest source of seismic energy released in Ecuador (IG, 2012; 2013) producing about 250 events with magnitudes greater than 2.8 each year (Araujo, 2012). The precise location of the event $M_w \simeq 5.0$ that occured on 26th March, 2010 in the area of the Pisayambo lake is very interesting to consider since the INSAR data evidenced a slip event, with a maximum displacement of 40cm, along the Bolichuco-Taurisacha fault (BTF) during a period compatible with this event (Champenois et al., 2014).

  Moreover, a field work (Baize et al. 2014) allowed the ground verification of cumulative displacements and of the 2010 faulting along the BTF on the one hand and, on the other hand, the results of the INSAR inversion (Champenois et al.

2014) are in agreement with the GCMT focal mechanism of the 26th March 2011. Consequently, Baize et al. considered that there is a large chance that this observed ground displacement corresponds to this seismic event, even though the IG location is about 12 km west of the fault and the USGS one 30 km east. More precisely, the IG localized the event at 1.14° S, 78.47° W and 2.6 km depth, and the USGS at 1.10° S, 78.13 W and 10 km depth. Our localization at 1.15° S, 78.36° W and 8.8 km depth (very close to the prior one) fits very well with the BTF (Figure 6.22) and does confirm the connection of the ground displacements with the seismic event.

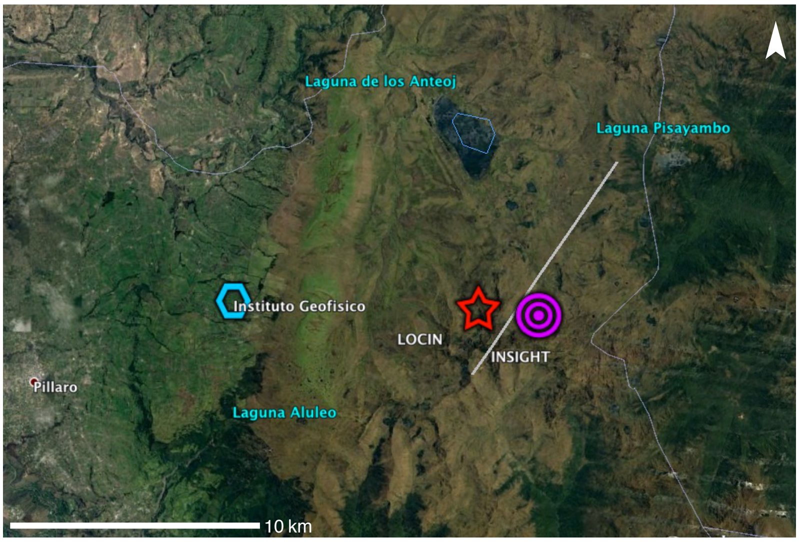

Figure 6.22: *Localizations of 26th March 2010 Pisayambo earthquake. Our final location (1.15° S, 78.36° W and 8.8 km depth) fits very well with the trace of the BFT, the approximative position of which is represented by a straight line (after Baize et al. 2014).*

- The Macas cluster also represents a large shallow seismic activity in the eastern subandean belt. In this area, located a few tenth kilometers southeast from Macas near the cordillera de Cutucu, a $M_w \simeq 7.0$ shallow earthquake occured the 3rd October 1995. Due to its magnitude, the event was widely recorded, even though there was no RENSIG station nearby. It allows us to measure the impact of the tomography on the prior localization since 20 P phases have still been recorded by the RENSIG. The event was localized at 77.95 W, 2.81 S and 16 km depth by the IG, at 77.88 W, 2.75 S and 24.4 km by the USGS, at 77.82 W, 2.77 S and 23.6 km depth by Engdahl et al. (1998). Our prior localization is close to that of the IG, 77.98 W, 2.83 S and at an altitude of 0.5 km, while the location after tomography is 77.86 W, 2.91 S and at an altitude of 0.3 km, about 15 km away from the prior one, and about 20 km south from that of the USGS. Our final location is only a few kilometers west of the trace of the Morona frontal thrust (Figure 6.23 and Legrand et al., 2005, figure 3) to which is related the earthquake according to Legrand et al. (2005).

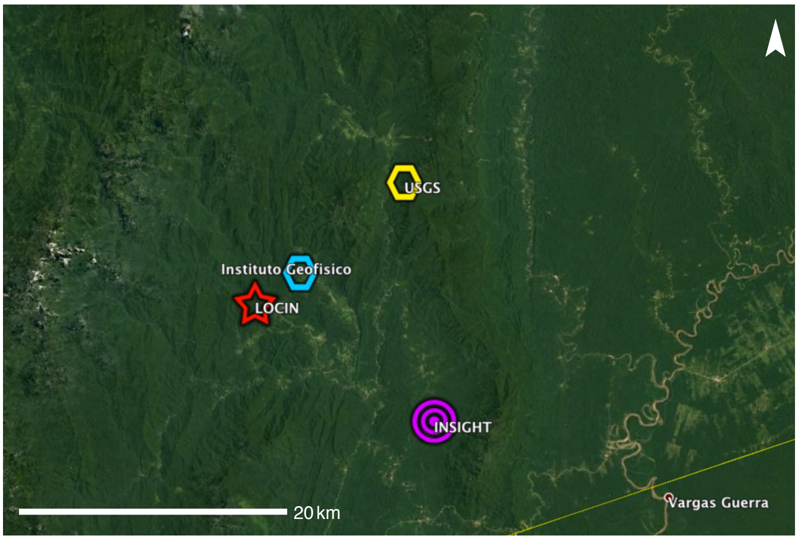

Figure 6.23: *Localizations of 3rd October 1995 Macas earthquake. Our final location (77.86 W, 2.91 S, at an altitude of 0.3 km) is close to the Morona frontal thrust.*

6.3.2 Investigating the slab

By considering the Wadati-Benioff zone, the figures 6-3,...,6-9 and 6.17 allow an investigation of the geometry of the slab between 0.5° N and 3.5° S with a good lateral resolution as only the earthquakes located less than 18 km of each cross-section are represented. It defines for the positive latitudes a continuously dipping slab down to a depth of about 125 km with an angle of about 23° that corresponds to the results of Guillier et al., 2001 and Manchuel et al., 2011. Between 0.25° N and 0.25° S, a first intermediate nest is discernible between depths of 90 and 120km that we will name Maldonado because it is located beneath the city of Pedro Vicente Maldonado, 70 km west of Quito (see also figure 6.17). This nest cannot be detected from the international data due to the relatively low magnitude ($M_w \leq 4$).

Around latitude of 0.5° S the Wadati-Benioff zone is not very clearly defined everywhere. The segment of the Wasdati-Benioff zone between the depths of 60 km and 110 km has a dip angle of 31°. Moreover the velocity anomaly ($v_P \geq 8.2$) could suggest a flatter slab below.

At about 1° S, the Wadadati zone can be clearly defined with the second intermediate nest (la Maná) at depths between 90 and 120 km. The corresponding dip angle is again about 25°. This angle is almost the same at 1.5° S down to 120 km, but the third nest (Puyo), which begins to be discernible at this latitude, presents a dip angle of about 40 ° at depths ranging between 160 and 240 km. That represents a sudden change in dip angle, unless considering this plan as inside the slab and related to tearing, in which case the slab may be considered as continuous from the trench. Alternativelly one may consider that the slab is broken in two pieces. One from the trench, down to 130 km depth, the other one with a dip angle of 40 ° up to about 90 km and appriximately associated to the eastern high velocity anomaly. We will try to analyze this scenario afterwards with the tomography of seismic clusters.

At 2° S the situation is much better with a clear continuity of the slab and the emergence of the fourth intermediate nest (Guayaquil) that ranges between 75 and 110 km. South to 2° S, the Puyo nest is no longer visible, and the dip angle decreases down to 10° at 3° S.

6.3.3 Mapping the slab

Based on the definition of the Wadati-Benioff zone, we have constructed a net of points corresponding to the surface of the slab in the large inversion box (between about 1.5° N and 5° S) by using the seismicity obtained either in this box or in the small one depending on the latitude. We have adopted a conservative point of view in regard to the Puyo nest in stopping the net at the change of dipping. Practically, every 25 km in the direction corresponding to the geographic meridian we have pointed the top of the Wadati-Benioff zone every 20 km in the direction corresponding to the parallel of latitude as far as possible.

To interpolate the data and to obtain a smooth surface, we use the Gaussian formalism of the inverse problem by using the algorithm in the data space (equation 2.29), which in this case is equivalent to krigging. We use the following correlation kernel :

$$Cor(\mathbf{x},\mathbf{x}') = \frac{1}{1+c}\left\{\varphi\left(\frac{||\mathbf{x}-\mathbf{x}'||}{L}\right) + c\varphi\left(\frac{||\mathbf{x}-\mathbf{x}'||}{l}\right) + \varphi\left(\sqrt{\frac{(x-x')^2}{l_{NS}^2} + \frac{(y-y')^2}{l_{EW}^2}}\right)\right\}$$

for any $\mathbf{x} = (x, y)$ and $\mathbf{x}' = (x', y')$ in the reference plan of the box, and where:

- φ is the real function defined by $\varphi(r) = 1/cosh(r)$,
- $0 \leq c \leq 1$ is a positive constant,
- L, l_{EW}, l_{NS} are three correlation lengths.

L is the greater length and insures the smoothing at long wave-length; l_{EW} and l_{NS} are two small lengths corresponding respectively to the East-West and the North-South direction that insure the fitting of the small wave-lengths. We use two different lengths at this scale due to the greater variability of the slab in the E-W direction than in the N-S direction.

To obtain a sufficiently smooth model while keeping the main feature of the raw data whe chose: $l_{EW} = 45km$, $l_{NS} = 70km$, $L = 300km$, $c = 0.1$, and a damping parameter $\sigma_d/\sigma_m = 1/15$ The resulting slab surface is displayed in figure (fig. 6.24).

The general pattern of this slab model is broadly similar to that of the global model Slab 1.0 (Hayes et al., 2012) in our study zone. The main common feature is the kind of

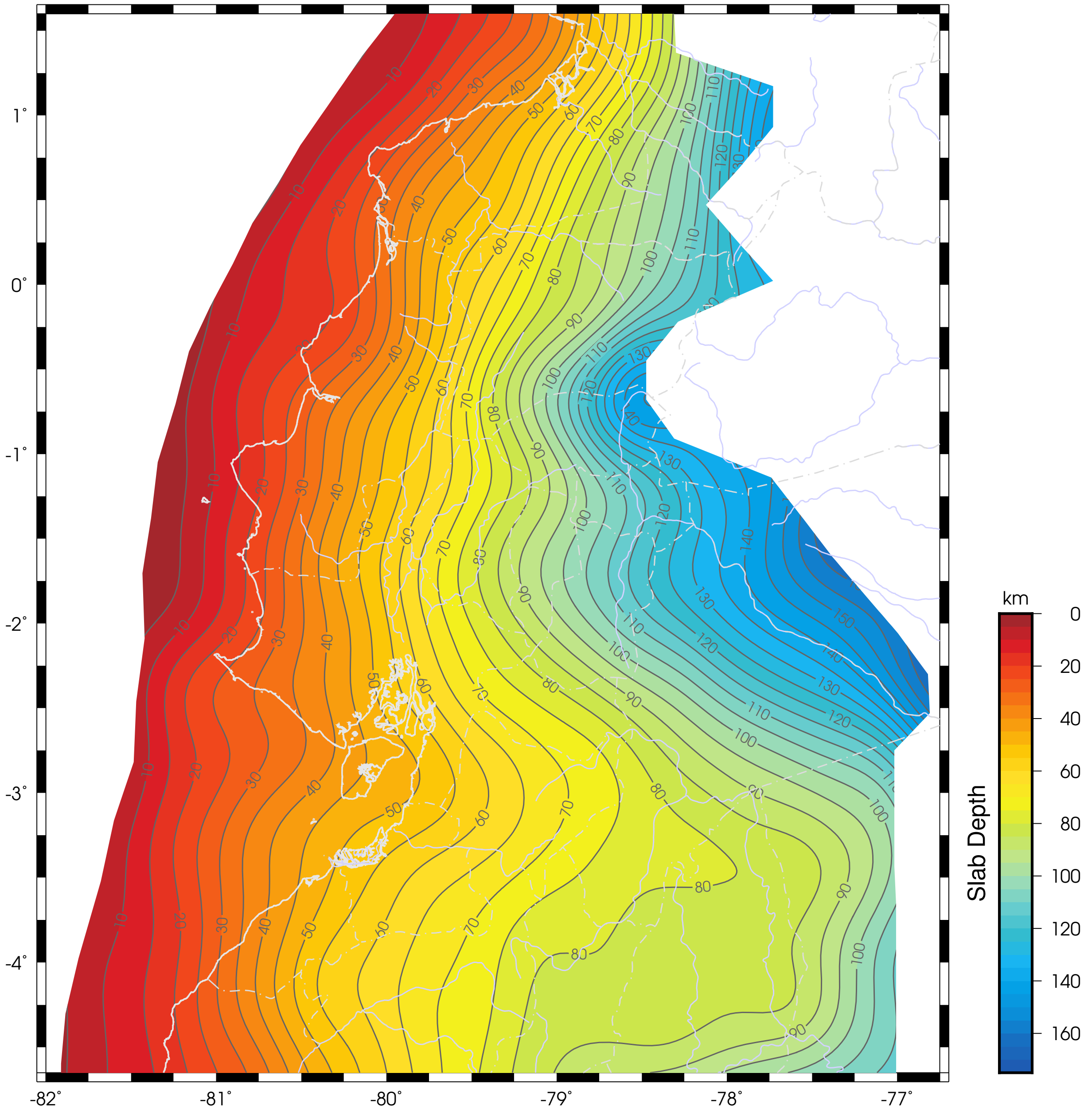

Figure 6.24: *Depth of the slab derived of our tomography results. The western boundary corresponds to the oceanic trench and the eastern boundary to the deepest part of the Wadati-Benioff zone detected by our tomography seismicity, aside from the Puyo nest.*

promontory that forms the surface at about 3.5° S-4° S at the latitude of the main concavity of the trench line, and which can be interpretated as an ondulation of the slab to accomodate the bending while keeping the same surface. Actually our model presents two other smaller ondulations with two clear trough lines and two ridge lines that respectively passes through the Maldonado and the La Mana nests. The line linking the Guayaquil and the Puyo nests, which approximately coincides with the extrapolation of the Grijalva fracture zone, *i.e.* to the initial spreading center between the Cocos and

Nazca plates, and could corresponds to a weakening of the plate (Yepes et al. 2016), leads to the southern trough line.

We can also observe that the starting slope near the trench of tle Slab 1.0 model is steeper than that of our model and much steeper than those observed by the SAL2, SAL6, and SIS4 transects (Graindorge et al., 2004, Gailler et al., 2007; Salares et al., 2005)

6.4 The intermediate-depth seismic nests

The intermediate seismicity consists of those earthquakes with depths ranging between 50 km and 300 km (Frohlich, 2006; Houston, 2007). In the study area this intermediate depth seismicity concentrates into 4 clusters (see figure 6.17), Puyo, La Maná, Maldonado and Guayaquil, which may be catalogued as nests since their activity is permanent and isolated from the nearby one (Prieto et al., 2012; Beauval et al., 2013); which distinguished them from aftershock sequences or swarms. The intermediate-depth seismicity occurs at temperature and pressure conditions that should not allow ordinary fracturation, and although it represents approximately 25 % of the global seismicity, the underlying physical mechanism remains unclear (Frohlich, 1989, 2006; Houston, 2007; Prieto et al. 2012) A first proposed mechanism is the dehydration embrittlement in which the dehydration of the slab generates high fluid pore pressure that counteracts the high ambient normal stress that inhibits the microcrack propagation (Houston, 2007). The viability of this mechanism was showed by experiments in serpentinite (Rayleigh and Patterson, 1965). The fluid could come from hydrated sediments embedded in the subduction or from phase transition in the oceanic plate at depth (Kirby et al., 1996; Davies, 1999). The popularity of this mechanism is due to thermo-mechanical model that theoretically explain the double seismic zone in various conditions (Seno and Yamanaka, 1996; Yamasaki and Seno, 2002). Despite of the success of dehydration embrittlement hypothesis, the difficulty to reproduce the experimental conditions in laboratory makes the cause of intermediate-depth seismicity a relatively open question. Other phenomenological models exist, as, for instance, the model of anticrack faulting in a metastable phase in which it is the volume change and heat release due to a phase transformation, rather than a fluid release, that trigger a shear instability (Houston, 2007). A third possible mechanism is a thermal shear instability produced when a feedback between the temperature-dependent rheology and the shear deformation generates viscous heating giving raise to an apparently abrupt failure on a very thin shear zone (Houston, 2007).

As the primary mechanism of this seismicity is not well understood, the empirical study of nest seismicity is crucial to bring a sufficient amount of data (Prieto et al., 2012). Deep nests constitute a relatively rare phenomenon, which should indicate exceptional conditions. For instance, the deep nest closest from the Puyo one is the well documented Bucaramanga nest (Vargas and Mann, 2013) 1100 km away. This unusual phenomenon could be the tear or breaking of the slab as it had been proposed in the case of Gribraltar Arc, the Lesser Antilles and the Tonga Trench (see Meighan et al., 2013) or in case of

Bucaramanga nest (Vargas and Mann, 2013). However, this latter nest lies in a very particular geodynamical setting, which has also been interpreted as the collision of the Nazca and Caribbean slabs (Zafiri et al., 2007). In any case, seismic nests indicate a discontinuity in the slab structure that can be analyzed by seismic tomography. For example, in the case of slab delamination, they could be related to a low P-velocity anomaly that could results of tearing or of the strengthening of slab dehydration due to the upwelling of hot asthenosphere materials (Bernal-Olaya et al., 2015).

We will now focus on the 4 Ecuadorian nests by using the double difference variant of the INSIGHTcode, described in section 2.4.2, that allows the relocalization of the earthquakes The code inverses differences of arrivall times at each station for both the event location and the velocity field inside a small box containing the nest (figure 6.25), with the results of regional tomography as prior model and prior locations.

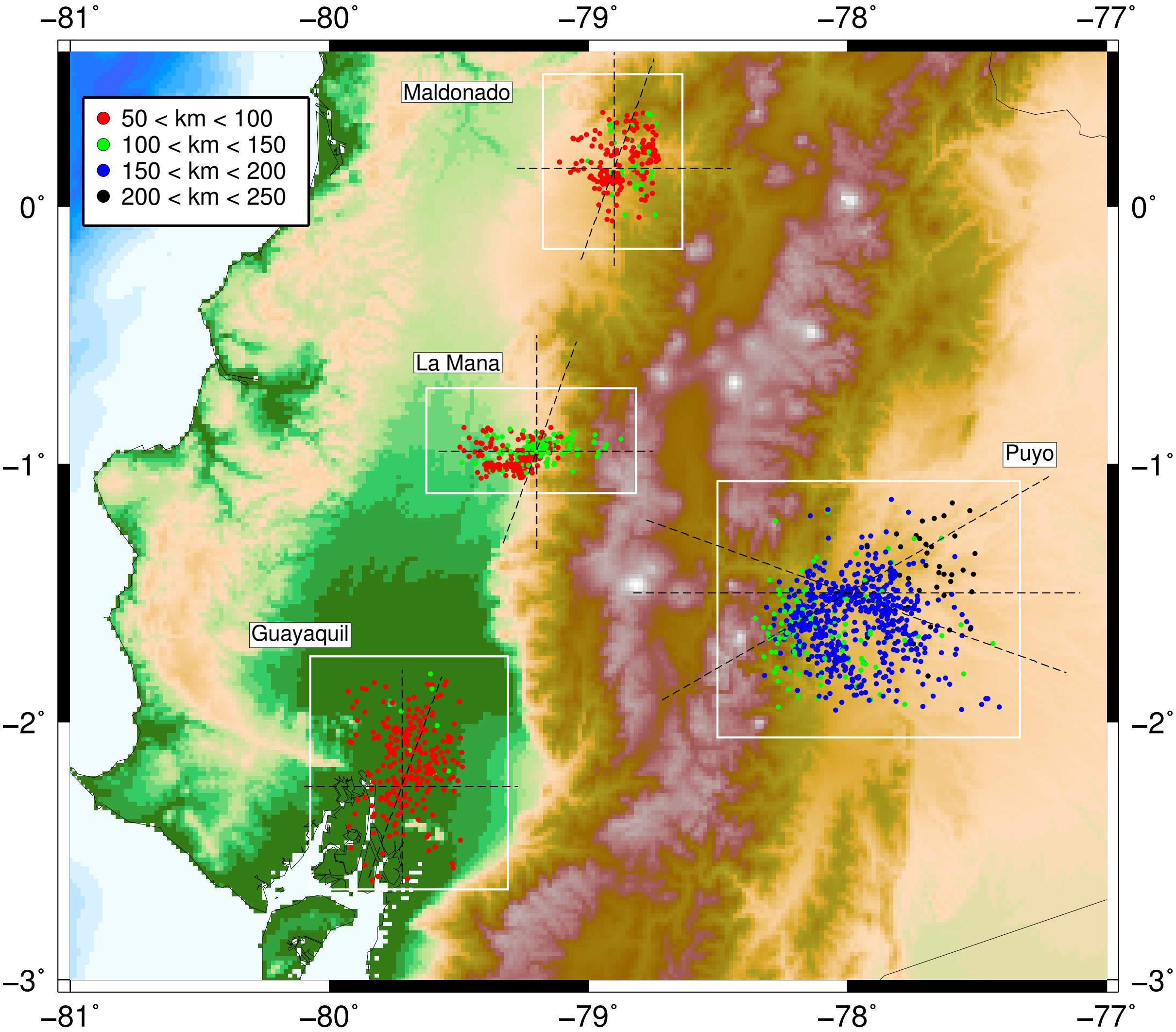

Figure 6.25: Positions of the four seismic nests analyzed by differencial tomography. The traces of the inversion boxes are denoted by white thin line. The traces of the considered cross sections are represented by dashed black lines.

- The La Maná nest

La Maná is a small city located at the limit between coastal plains and Western Cordillera 110 km southwest of Quito.The inversion box is centered at 0.95° S, 79.2° W and at a depth of 104 km. The dimensions of the inversion box are 91 km in the E-W direction, 46 km in the N-S direction and 67 km in the vertical direction (figure 6.25). The nest inside the box consists of 509 earthquakes. We have assigned the values of 10 km to the smoothing lengths ξ_H, ξ_V, and respectively 250 m/s and 0.05 to the physical standard deviations of v_P and v_P/v_S. To determine the tuning parameter for the regularization we have computed L-curve diagrams for various values of the damping parameters ξ_0 and λ related to the velocity model and locations (figure 6.26)

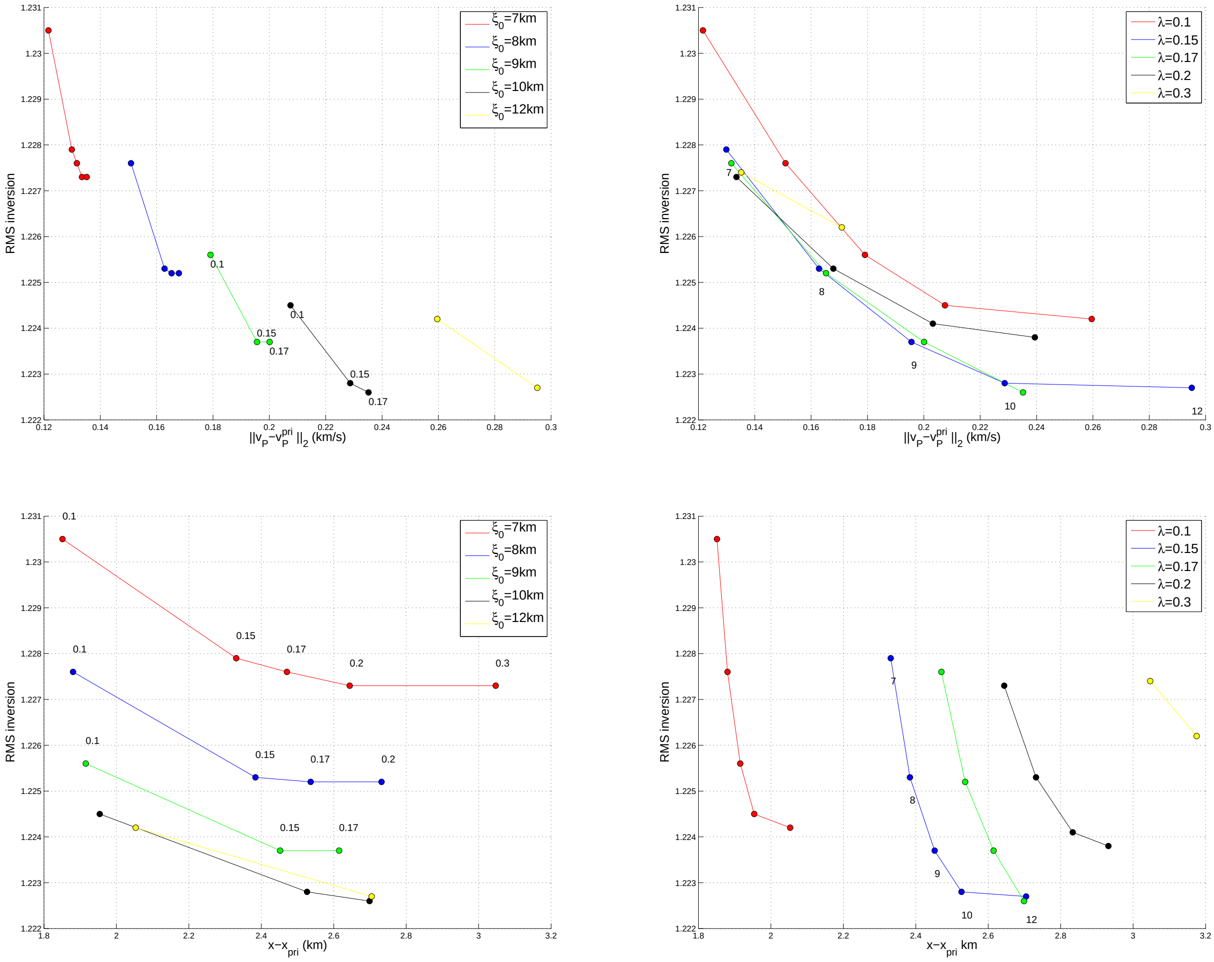

Figure 6.26: *RMS of the fitting variable versus the norm of the increment of the P velocity field with respect to the prior one (top), and versus the mean distance between the hypocenters and the prior ones (bottom). The values near each point correspond to λ on the left panel and to ξ_0 (in km) on the right one.*

The analysis of these L-curves leads to the values of $\xi_0 = 10km$ and $\lambda = 0.17$ for the tuning parameters. The set of inversion parameters is thus:

σ_{V_P}	σ_{V_P/V_S}	ξ_H	ξ_V	ξ_0	λ
$250m/s$	0.05	$10km$	$10km$	$10km$	0.17

The figure 6.27 displays 3 cross sections oriented WE, SN and N 20° E (see figure 6.25) of the nest. It shows that the nest is clustering on a mean almost vertical plan, oriented W-E. It also confirms the low values of the P velocity in the mantle above the slab, even though we can observe a numerical oscillation in the velocity.

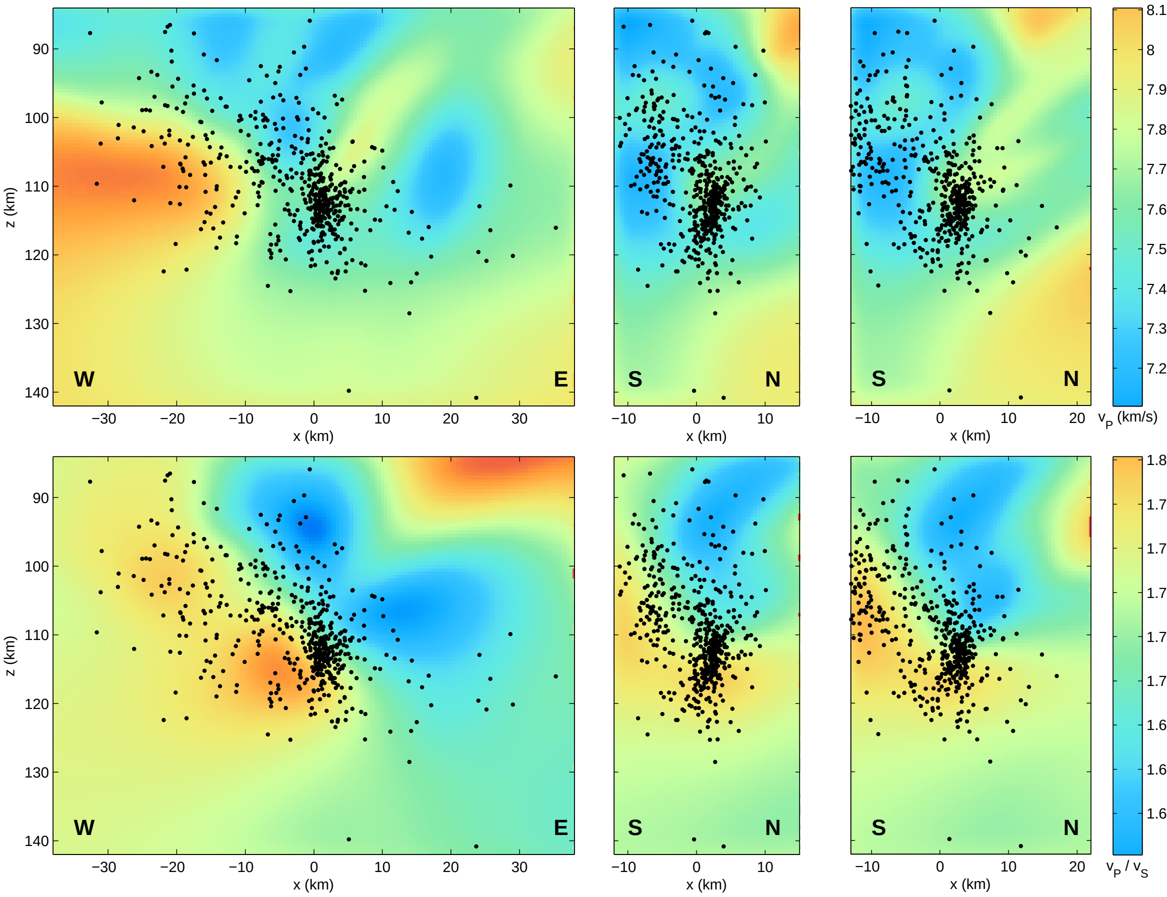

Figure 6.27: *Vertical cross sections of the La Maná nest in the W-E direction (left), S-N (middle), N-20° E (right). All the seismicity is represented and z is the vertical coordinate in the reference box.*

- The Maldonado nest

  Pedro Vicente Maldonado is a small city located 70 km west of Quito. The corresponding inversion box is centered at 0.15° N, 78.9° W and at a depth of 95 km, with dimensions of 61km E-W, 76 km N-S and 41km in the vertical direction. The nest contains 224 events inside the box. The values of the inversion parameters are set as follows:

σ_{V_P}	σ_{V_P/V_S}	ξ_H	ξ_V	ξ_0	λ
$250m/s$	0.05	$10km$	$10km$	$12km$	0.2

The relocalization displayed in figure 6.28 shows the normal dipping of the Benioff zone in the W-E cross section. It also shows that the seismicity is spread between to components lying in vertical plans.

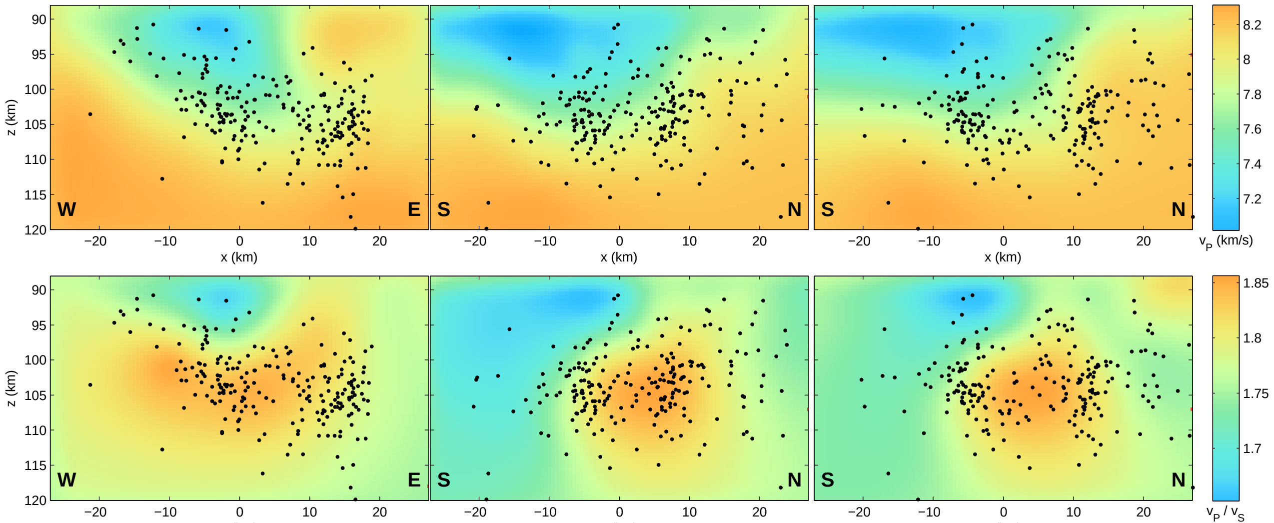

Figure 6.28: *Vertical cross sections of the Maldonado nest in the W-E direction (left), S-N (middle), N-20° E (right). All the seismicity is represented and z is the vertical coordinate in the reference box.*

- The Guayaquil nest

This cluster, which a priori presents the greatest spatial dispersion and is the shallowest of the four nests, is located beneath the city of Guayaquil. We identified 126 events within the inversion box of 86 km N-S, 101 km E-W and 45 km heigth, centered at 2.25° S, 79.72° W and at a depth of 84 km. The parameters obtained for the regularization are as follows:

σ_{V_P}	σ_{V_P/V_S}	ξ_H	ξ_V	ξ_0	λ
$250m/s$	0.05	$10km$	$10km$	$9km$	0.17

The cross sections displayed in figure 6.29 show a P velocity structure corresponding to the Wadati-Benioff directionality in agreement with the standards of subduction zones. Despite relocalization the cluster remains widespread.

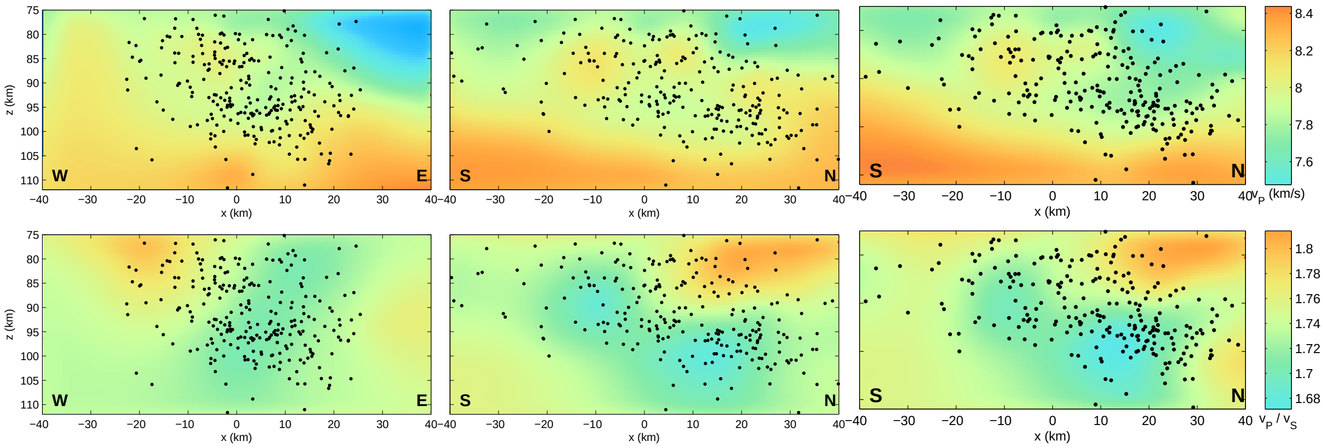

Figure 6.29: *Vertical cross sections of the Guayaquil nest in the W-E direction (left), S-N (middle), N-20° E (right). All the seismicity is represented and z is the vertical coordinate in the reference box.*

- The Puyo nest

  The Puyo cluster is the unique source of deep seismicity in Ecuador and is, properly speaking, the only deep nest in Ecuador. Its name is related to the Puyo city located 160 km southeast of Quito. The inversion box is centered at 1.5° S, 78.0° W and at a depth of 171 km, with dimensions of 131 km E-W, 111 km S-N and a heigth of 143 km .

  721 seismic events are a priori located within the box. Based on a L-curve analysis (figure 6.30) we set the damping parameters at ξ_0 =7 km and $\lambda = 0.2$.

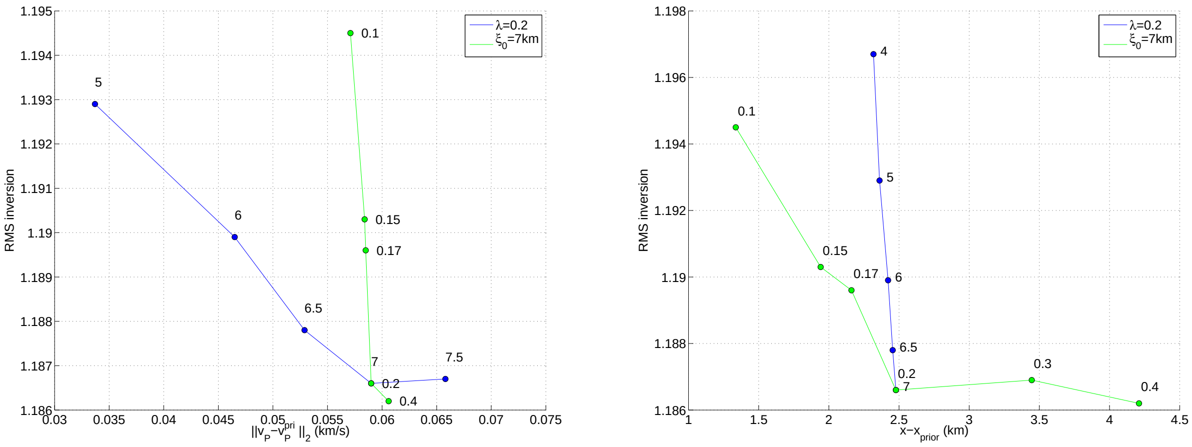

Figure 6.30: *RMS of the fitting variable versus the norm of the increment of the P velocity field with respect to the prior one (left), and versus the mean distance between the hypocenters and the prior ones (right). The values near the points correspond either to* ξ_0 *(km) along the blue lines or to* λ *along the green lines*

The complete list of inversion parameters is as follows :

σ_{v_P}	σ_{v_P/v_S}	ξ_H	ξ_V	ξ_0	λ
$250m/s$	0.05	$10km$	$10km$	$7km$	0.2

The figure 6.31 displays the 3 cross sections oriented W-E, E 20° S and E 30° N (see figure 6.25) of the nest. This latter orientation points out a preferential direction of the relocalization of the nest (right of figure 6.31) which is associated with a normal structure of the velocities, and approximately coincides with the direction of the Grijalva Fracture Zone. (figure 6.24). The orientation E 20° S corresponds to the line joining the La Mana and the Puyo nests.

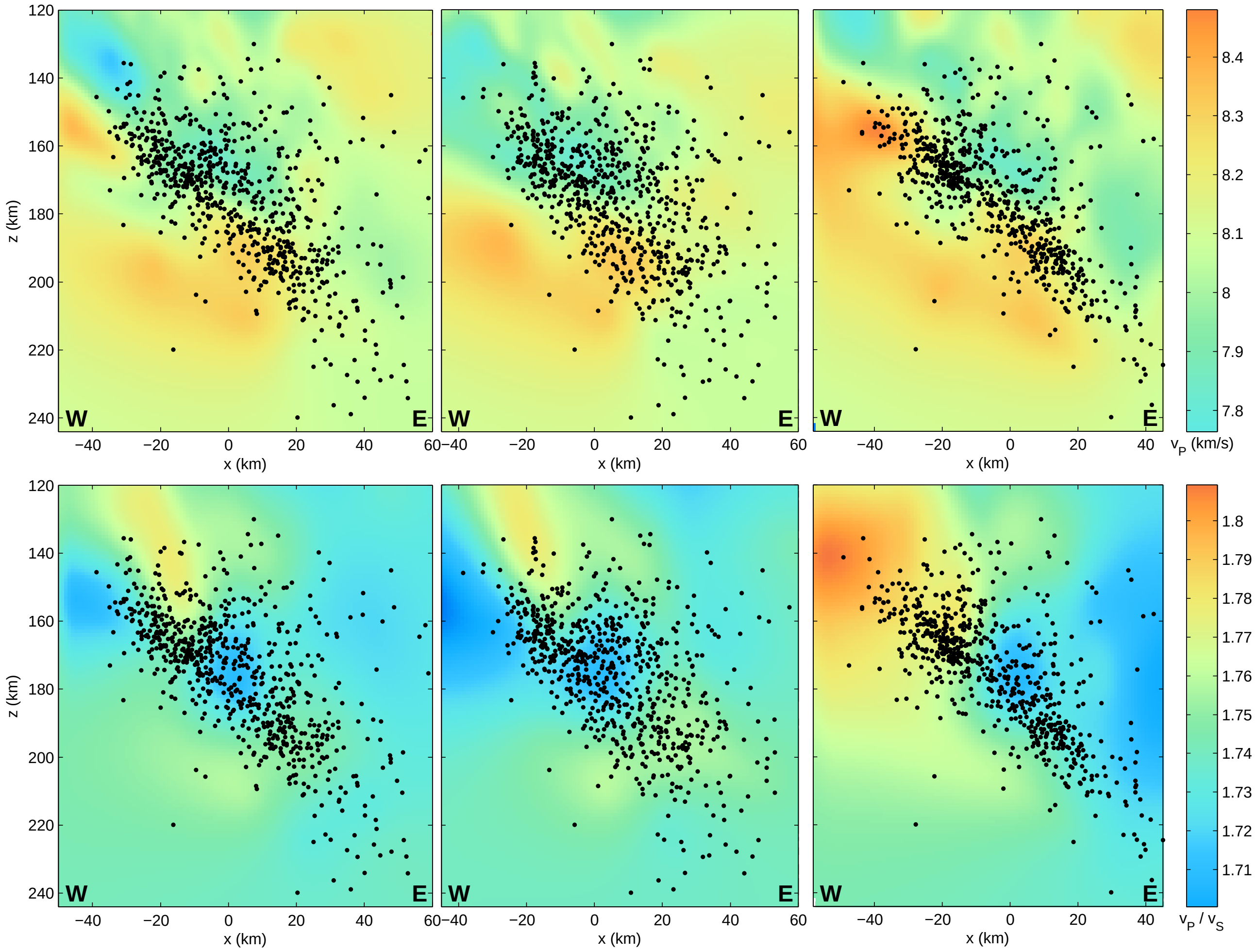

Figure 6.31: *Vertical cross sections of the Puyol nest respectively oriented W-E (left), E 20° S (middle), E 30° N (right). All the seismicity is represented and z is the vertical coordinate in the reference box.*

The relocalization shows that the nest is actually well concentrated over a plan, and allows a precise determination of this clustering plan with a strike angle of about N 30° W and a dip angle of about 40°. Furthermore, the Harvard CMT catalogue shows that the focal normal mecanisms of the 24 $M_w \geq 5$ events available on the website are very homogeneous with a strike of the nodal plans compatible with that of the nest.

6.5 A discontinuous slab ?

Let us now come back to the topology of the slab in trying to adopt the opposite point of view of the conservative and mainstream paragraphs 6.3.2-6.3.3. Let us start by considering the cross section oriented N 20° E that passes through the 3 nests of Guyaquil, La Maná and Maldonado and the perpendicular one that joins the La Maná and the Puyo nests (figure 6.32).

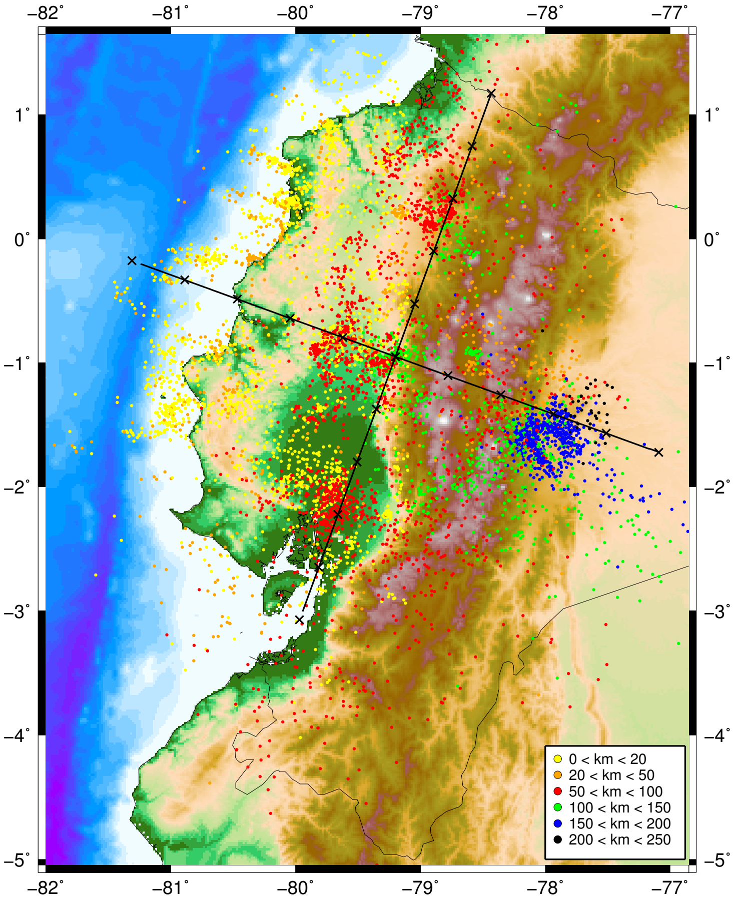

Figure 6.32: *Positions of the two orthogonal cross sections centered near La Maná with orientations N 20° E and E 20° S. The shallow seismicity beneath and at the east of the Andean range has been removed.*

The figure 6.33 shows that the La Maná nest corresponds to a discontinuity of the domaine $v_P > 8.2$ km/s of the P wave velocity model that would suggest a discontinuity between a northern slab and a southern slab or at lest a change in the constitution of the slab.

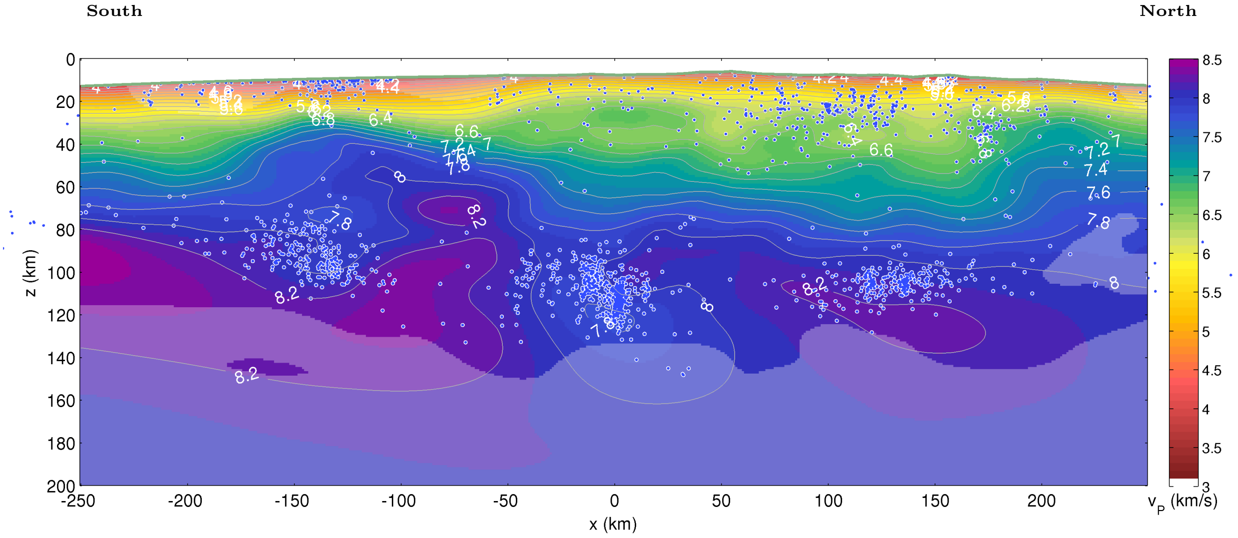

Figure 6.33: *Cross section of the P velocity model centered at the La Maná cluster and oriented N20° E.*

To follow this anomaly let us have a look at the perpendicular cross section oriented E 20° S that connects the La Maná with the Puyo nest (figure6.32). The discontinuity extends down to 170 km depth (figure6.34) and a new high velocity domain arises above the Puo cluster. The Wadati-Benioff zone is well defined down to the La Maná nest, but between this cluster and the Puyo nest the seismicity is sparse and not in the continuity of the slab (6.34).

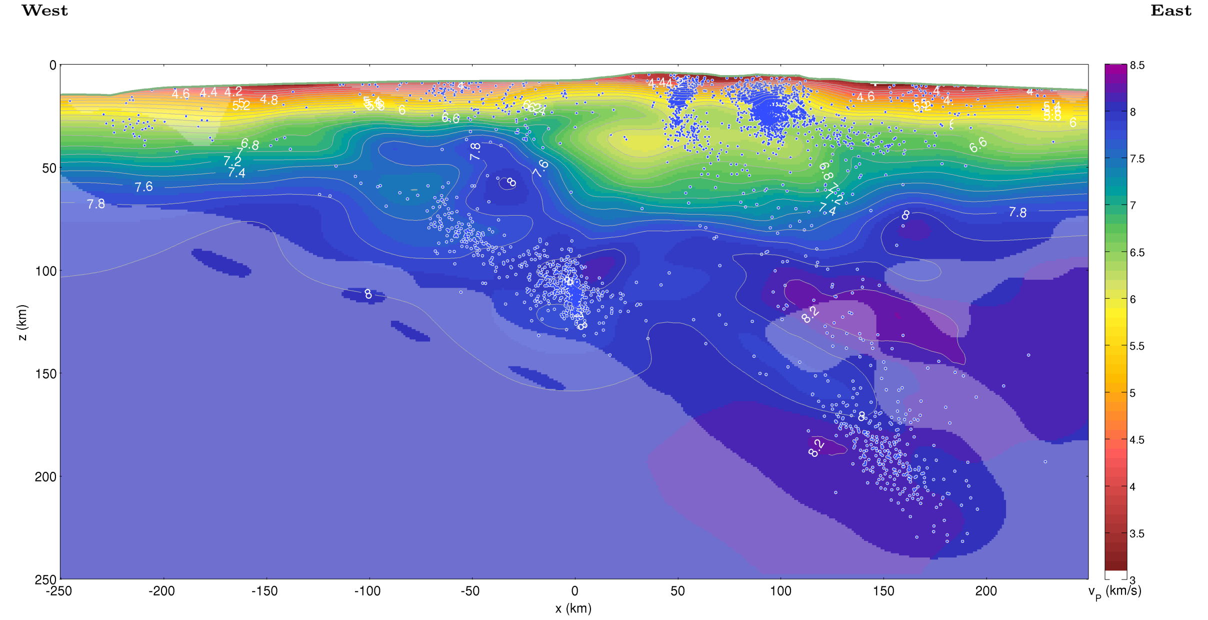

Figure 6.34: *Cross section of the P velocity model centered at the La Maná cluster and oriented E20°S. The La Maná and Puyo nests are clearly identifiable.*

To go further on in the exploration of this high velocity discontinuity the figure 6.35 displays in 3-D the domain defined by $v_P \geq 8.1$ km/s that gives us a rough representation of the subducted Nazca plate. The figure at the top does replace in the 3-D context the two cross sections 6.33 and (6.34).

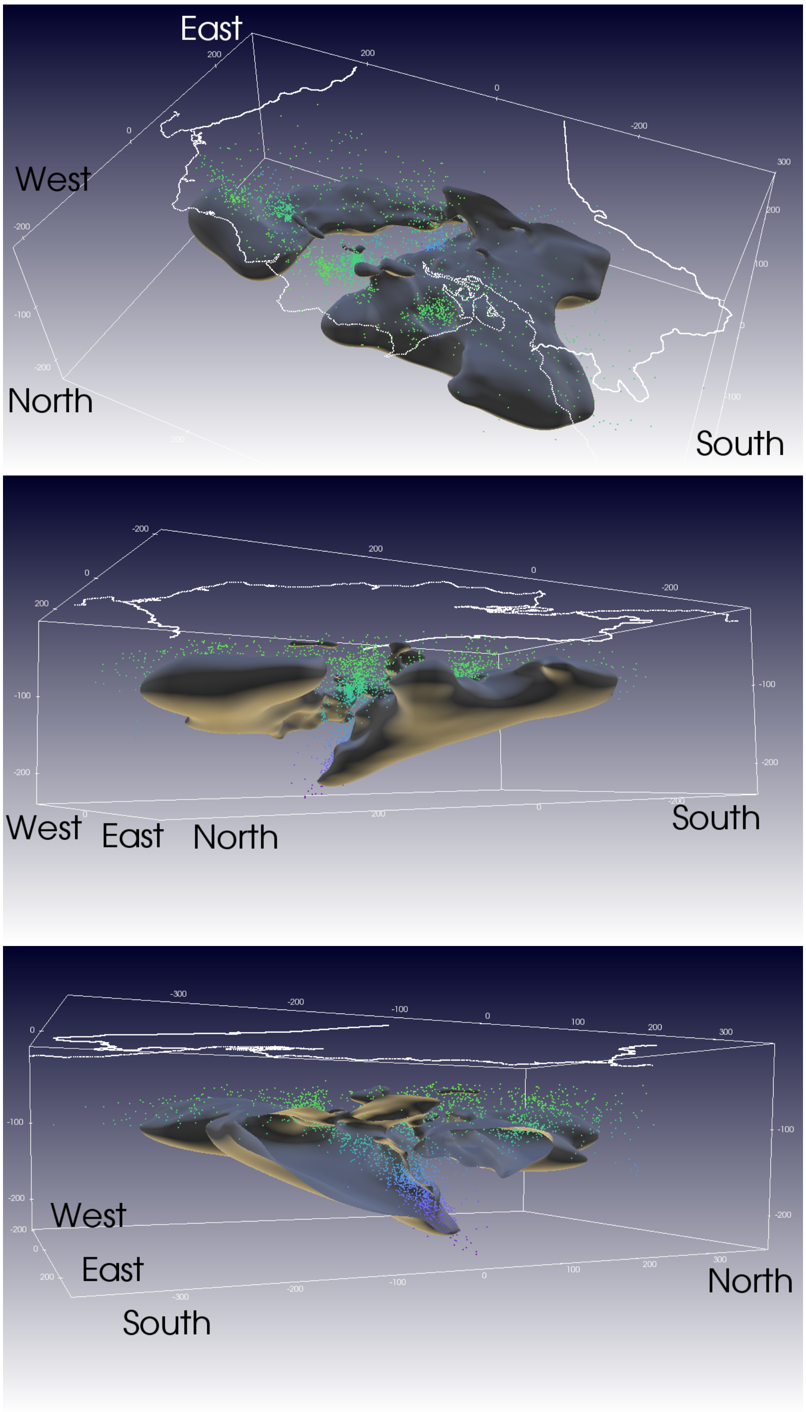

Figure 6.35: *Velocity isovolumes $v_P \geq 8.1 km/s$ in the result model. Seismicity represented as dots.*

Furthermore, the intermediate-depth seismicity, superimposed in 6.35 shows that the Guayaquil and Maldonado clusters are directly related with changes in the dip of the subduction (bottom figure 6.35). Instead, the La Maná cluster is directly situated in the zone of low velocity which connects with the Puyo cluster (figure 6.35).

We can also observe that the Puyo nest is continuously connected to the southern slab without any discontinuity in the slope, but rather by a continuous seismicity from the Peruvian flat slab to an increasingly sloping southern slab passing under the northern structure. (midle and bottom, figure 6.35),

To make this latter point completely clear, the figure displays in 3-D the seismicity as seen from the south-east and from the north-west. It shows the continuity of the seismicity of the Peruvian slab down to the Puyo nest, and from the Puyo nest up to the Guayaquil cluster.

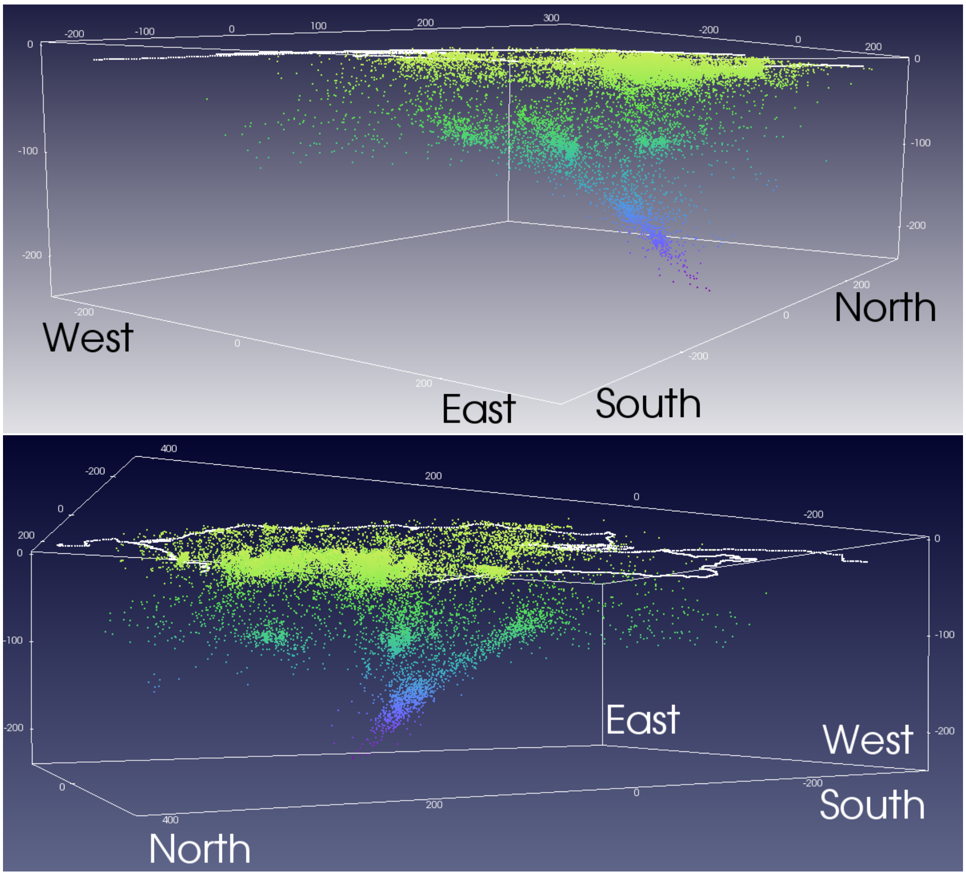

Figure 6.36: Seismicity of the Puyo seismic nest showing a continuity of the southern slab

In order to try to explain the velocity difference between 8 km/s and 8.2 km/s that is related to this apparent gap of the slab at the level of the La Maná nest, let us consider the relation between the P wave velocity and the pressure (depth) and temperature in the mantle (Cammanaro, Goes, Vacher and Giardin, 2003 and figure 6.37). The left graphic shows that a decrease in velocity from 8.2 to 8 km/s corresponds to an increase in temperature of about 500°C at a depth of 100 km and for the same rheology.

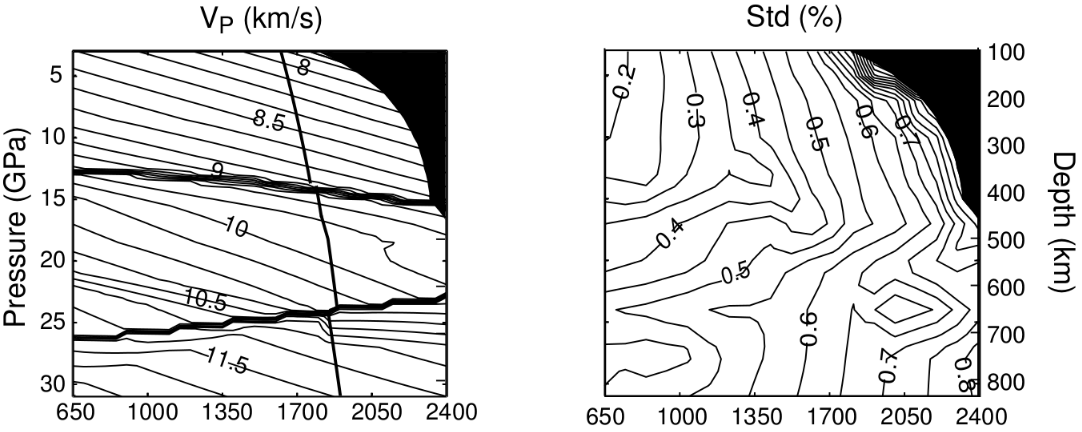

Figure 6.37: *Theoretical values of P wave velocity (left) and its standard deviations (right) for the Earth mantle. Results are reported in function of pressure and depth in vertical axes and temperature (°C) in horizontal axis. Figure from Cammanaro, Goes, Vacher and Giardin, 2003.*

This low velocity anomaly in the slab does not allow us to claim that there exist a tearing of the slab up to the trench. It could rather exhibit a zone where the subducted Nazca plate is hotter. Moreover, our present dataset does not allow a very accurate resolution of the slab near the trench. The incorporation of the data of the ADN experiments and of the last sequence following the April 2016 earthquake would surely substancially improve the resolving power of the whole data set. The Puyo cluster, however, clearly represents a zone where the slab has broken since in the three dimensional representation we can see the southern slab passing beneath the northern slab (figure 6.35). This could confirm and clarify the high velocity zone above the Puyo cluster as showed in the tomographic P velocity images of this cluster (figure 6.31).

The orientation about E 20° S of this gap cannot be related to the Grijalva Fracture Zone, which corresponds to the initial spreading center between the Cocos and Nazca plates, and the orientation of which is approximately that of a line joining the Guayaquil with the Puyo nest. But, although the opening of this ridge has been stopped between the coastal subduction zone and the Panama fault to the West for about 10 million years (Londsdale, 2005), the opening ridge, when subducting, swept before all the shallower mantle beneath central and northern Ecuador.

6.6 Mohography

The velocity model, displayed in figures (6.3,...,6.9) allows us to build a model of Moho depth. Since this depth usually corresponds to a jump in P velocity between about 6.9-7.0 km/s and 7.9-8 km/s, we search at each latitude and longitude the depth corresponding to the maximum of the norm of the velocity gradient in the range of velocity between 7.2

km/s and 7.4 km/s. This approach is thus a trade off between the strategy consisting in identifying the Moho with an iso-velocity surface and the one consisting in taking the depth of maximum norm of the gradient.

We have tested the approach with a lot of models corresponding to a wide range of tuning parameters, varying the damping length ξ_0, the vertical and horizontal smoothing lengths ξ_V and ξ_H. We observed a relative stability for the resulting Moho depths. Our final choice corresponds to the following set of tuning parameters for the model (resulting of the data inversion): ξ_0=5 km, ξ_H =35 km, ξ_V= 15km that yields the raw model displayed in Figure 6.38. You can observe that we prefered to use a stronger damping (a lower ξ_0) and a shorter vertical smooting length ($\xi_V = 15km$) than for the tomographic model diplayed in figures (6.3,...,6.9) to obtain a better vertical resolution near the base of the crust.

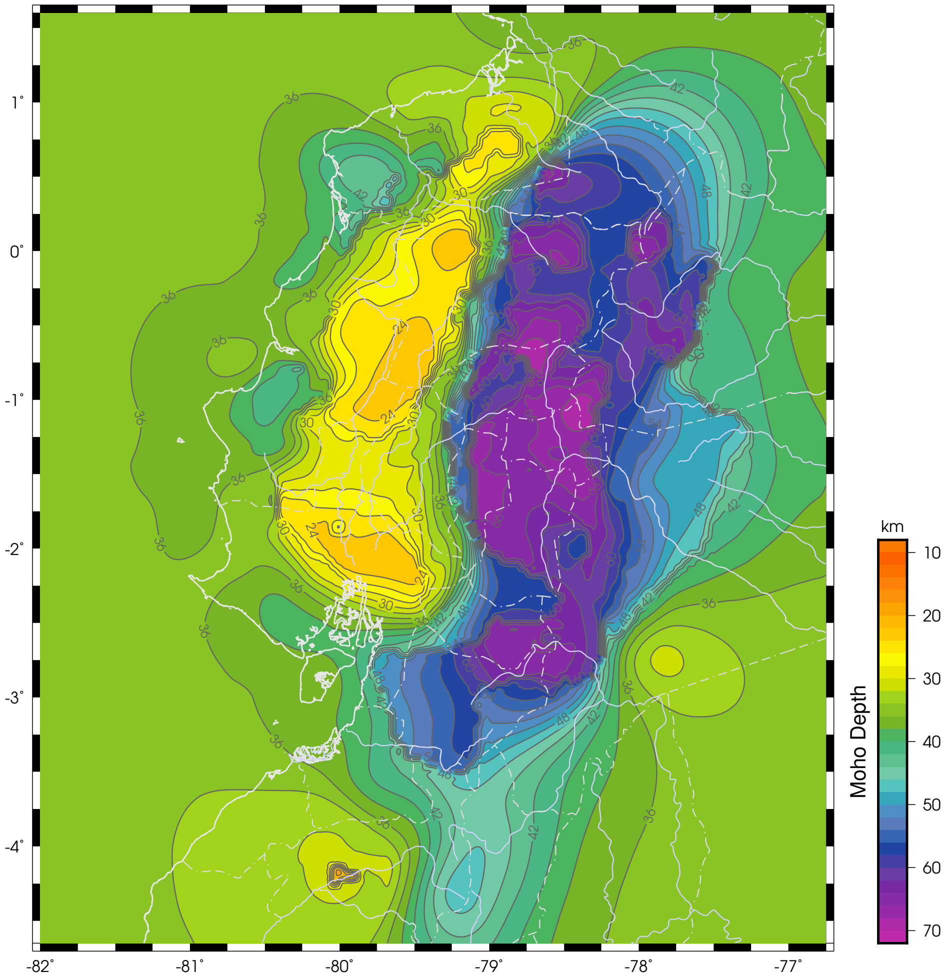

Figure 6.38: *Maximum of gradient velocity norm between isovelocities $7.2km/s$ and $7.4km/s$ for the resulting tomography model of v_P.*

This raw Moho model (figure 6.38) must be clearly smoothed and complement with information in the North East and the West to compensate the lack of resolution of our tomography in these areas.

6.6.1 Complementing information on Moho depth

For the west area over the Pacific ocean and the margin, we took advantage of the seismic profiles determined by travel time inversion of wide angle seismic data acquired during the marine seismic experiments SISTEUR (Collot et al. 2002) and SALIERI (Flueh et al., 2001) off the Ecuadorian coastine. More precisely we used (Figure 6.39):

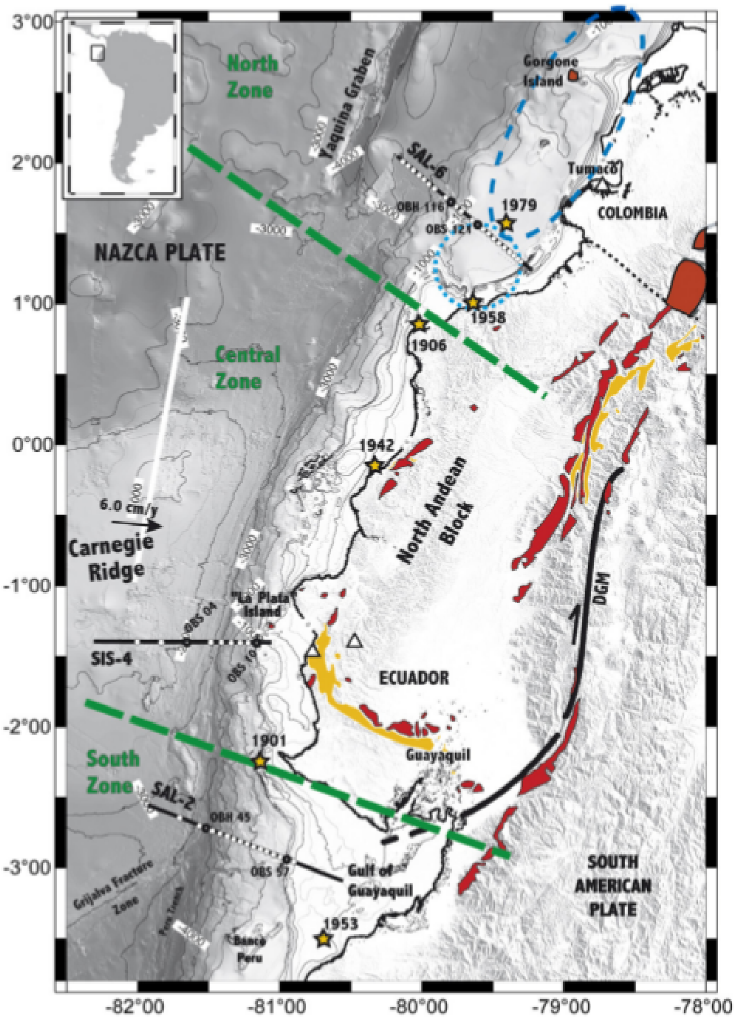

Figure 6.39: *Positions of the three seismic profiles used for the mohography profile 2 SALIERI (white line, Sallares et al., 2005), SIS-4 and SAL-2 (Gailler et al., 2004; Graindorge et al. 2004). The figure is from Gailler et al., 2004.*

- the results of the profile2 (SALIERI) across the Carnegie ridge, running alongside the Ecuadorian coast at the west of the trench (Sallares et al., 2005, Figure 3 and Gailler et al., 2007, figure 5).

- The SIS(TEUR)4 transect at 1.4°S along the southern flank of the Carnegie ridge (Gailler et al., 2007, figure 5).
- The SAL(IERI)2 transect in the Guayaquil gulf between about 2.5° S, 82° W and 3° S, 80.5 W (Gailler et al., 2007, figure 5, Graindorge et al., 2004, figure 4).

From the figures of these three profiles indicating the Moho discontinuity, we have built a N-S profile of the oceanic crust thickness at 81.875° W between 1.5° N to 3.5° S. (figure 6.40). This profile approximately corresponds to the crust thickness of the profile 2 (SALIERI) given by Sallares et al. (2004) between 1° N and 1° S. It rougthly matches the thickness of Graindorge et al. (2004) and Gaillet et al. (2007) at about 2.6° S and reaches a thickness of 7.5 km at the northern side (3° S) of the Grijalva Fracture Zone (GFZ). At the south of the GFZ the thickness is set at 5.5 km taking the jump of 500 m across the GFZ of the bathymetry (Londsdale, 2005) into account through a simple isostasy evaluation. At the northern limit the extrapolated value is 7 km.

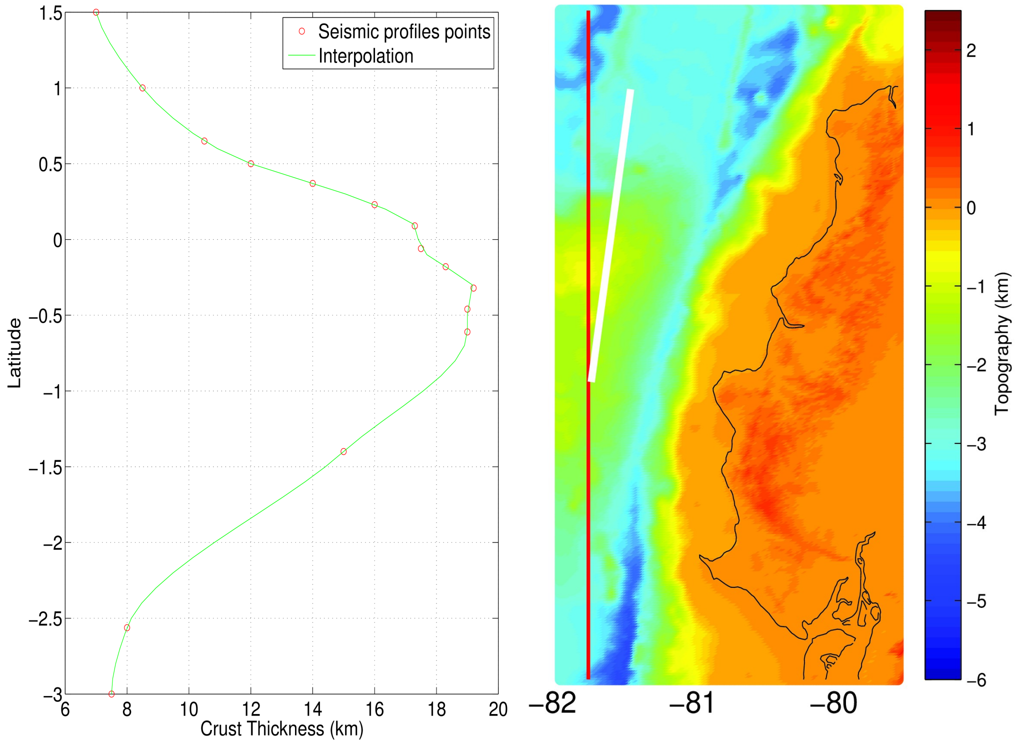

__Figure 6.40:__ Left: the crust thickness profile at 81.875° W deduced from the three wide-angle seismic profiles (SALIERI and SISTEUR). Right: map of the bathymetry (the trace of the left cross section is indicated in red).

To obtain the Moho depth down to a depth of 35 km, the general strategy was to propagate this reference profile of crust thickness following the azimuth of convergence (N 83° E) and to add either the local bathymetry or the depth of the slab as determined in paragraph (6.3.3), depending on the position with respect to the trench. We also virtually propagate the GFZ across Ecuador to take the corresponding jump in crust thickness into account. The precise algorithm is as follows:

1. In the North of the GFZ and to the West of the trench, the 81.875° W crust thickness (fig. 6.40) is propagated in the direction of the azimuth of N 83° E throughout the area, and the local bathymetry is added at each point to obtain the local Moho depth.

2. In the South of the GFZ and to the West of the trench, the thickness of the crust is set at the constant value of 5.5 km, and as in the previous item the local bathymetry is added to obtain the Moho depth.

3. In the north of the GFZ and to the East of the trench, the value of the crust thickness is, as in the first item, the reference thickness (at 81.875° W) propagated following the azymuth N 83° E, but, this time, multiplied by the factor $1 - |\lambda - \lambda(trench)|/2$, where λ and $\lambda(trench)$ stands respectively for the longitude of the point and the longitude of the trench at the latitude of the point in degree. This correcting factor encounts for the thinning of the crust when subducting. The depth of Moho is obtained by adding this resulting thickness to the depth of our slab model instead of the bathymetry.

4. In the south of the GFZ and to the East of the trench, the evaluation is the same as in the previous item but in assigning the value of 5.5 km to the crust thickness before correction.

In the North-East of the box, we use the Isostasy approach to complement the information. Taking as reference a crust at the sea level with a thickness of 35 km, the simple isostasy yields :

$$\Delta = kh \quad \text{with} \quad k = \frac{\rho_s}{\rho_m - \rho_c} \quad \text{and} \quad \Delta = z - 35$$

where z is the depth of Moho (in km), h is the altitude, ρ_s, ρ_c, ρ_m are respectively the density at the top, the bottom of the crust and at the top of the mantle. Taking respectively for these latter densities the values of about 2.7, 3.0, 3.3 yields k of the order of 9. Based on our model of the depth of Moho and on the model Etopo1 we have sampled this coefficient k in a domain between 0.5° S and 0.5° N, 77.5° W and 79.2° W adjacent to the zone to infer, and obtained a median value of 8.7133 that confirms our Moho model. Based on this value of k we have extrapolated the depth of Moho over the N-E of the box by using the above formula and a locally averaged altitude.

The resulting models of the Moho depth, West and North-East of the box, are displayed in figure 6.41. We have also incorporate 11 values of the Moho depth derived from the profile SAL2 in the Guayaquil gulf (Gailler et al., 2007) because our model does not fit well these data, as can be seen in the figure 6.41.

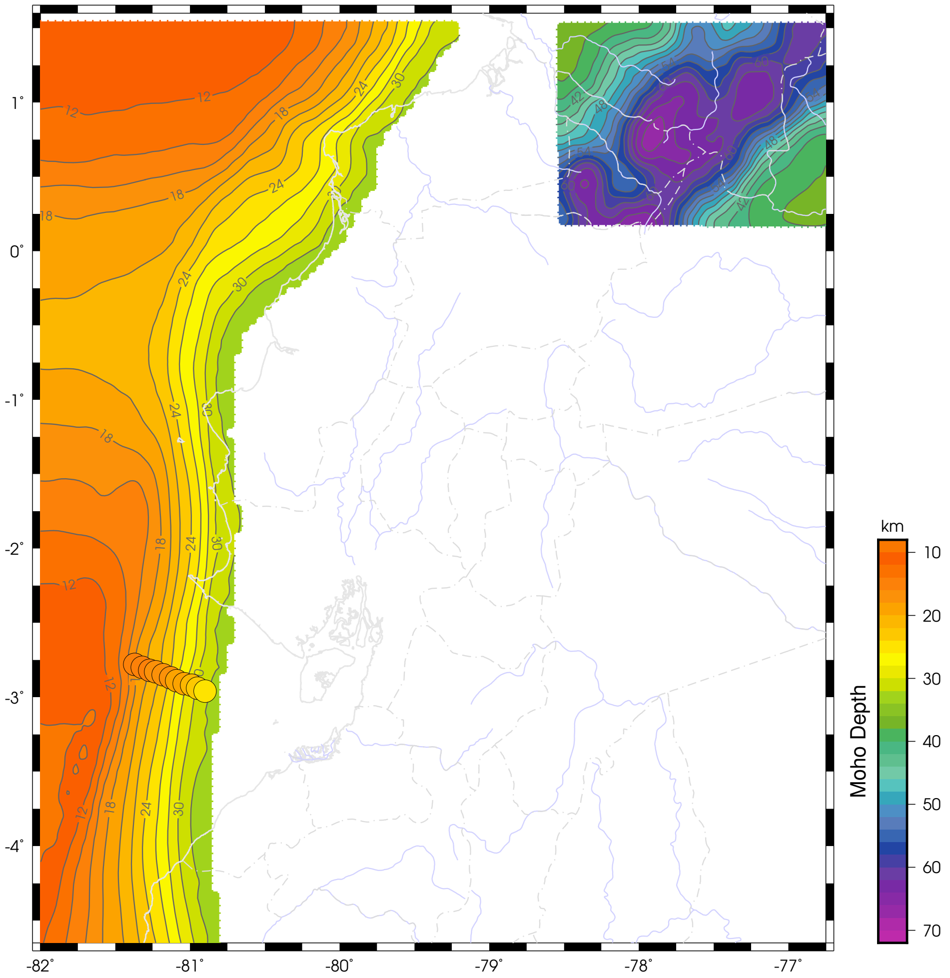

Figure 6.41: *Moho depth model derived from seismic profiles SALIERI and SISTEUR (West of the box), and by isostasy (Nort-East of the box). The 11 depths values derived from the profile SAL2 (Gailler et al., 2007) are represented by disc, the color of which corresponds to the scale on the right.*

6.6.2 Matching Moho information

To match all this information we adopt the same stochastic formalism as for the slab depth in (6.3.3). The data are the various depth values coming from either our raw model (figure 6.38) or the complementary models (figure 6.41). The correlation kernel is the kernel :

$$Cor(\mathbf{x},\mathbf{x}') = \frac{1}{cosh\left(\frac{||\mathbf{x}-\mathbf{x}'||}{L}\right)}$$

for any $\mathbf{x} = (x, y)$ and $\mathbf{x}' = (x', y')$ in the reference plan of the box, and where L is the smoothing length.

To set the standard deviation of the data, we have analytically defined a weight function $w(x, y)$ (figure 6.42) ranging between 0 and 1 and that takes the value 1 in the area in which the influence of the raw model must be enhanced and 0 where it must be removed.

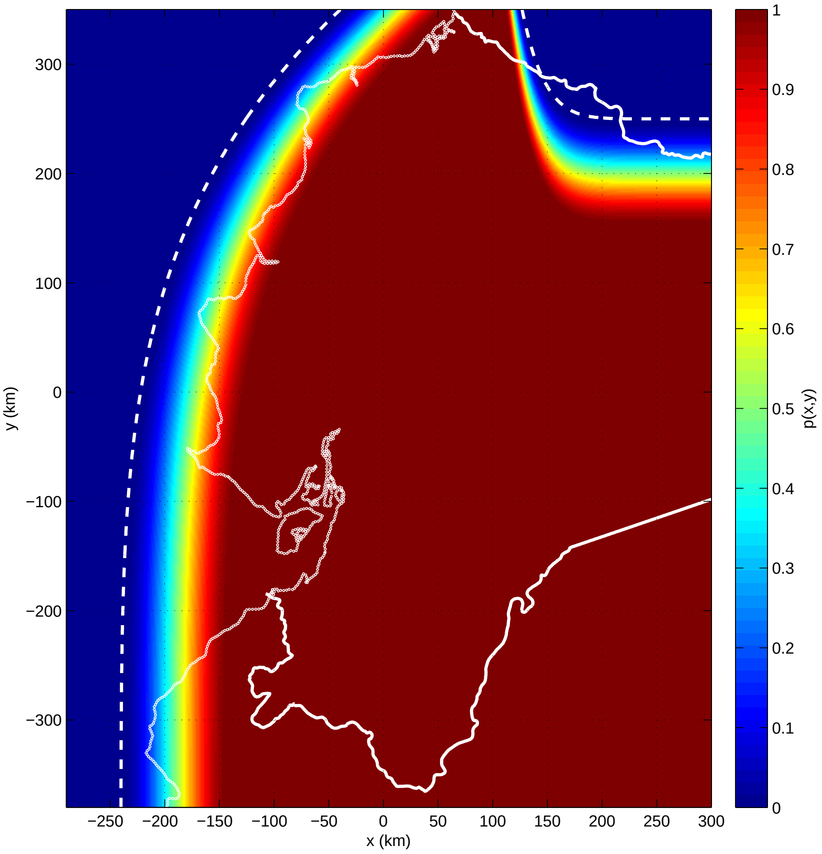

Figure 6.42: *The weight function $w(x, y)$ for standard deviations*

By using this function $w(x, y)$ the standard deviation $\sigma_d(x, y)$ of a datum d located at (x,y) are given by:

$$\sigma_d(x,y) = \sigma_d^{min} + \begin{cases} w(x,y)(\sigma_d^{max} - \sigma_d^{min}) & \text{for data of the raw model} \\ \\ (1 - w(x,y))(\sigma_d^{max} - \sigma_d^{min}) & \text{for data of a complementary model} \end{cases}$$

except for the data corresponding to the SAL2 profile for which σ_d increases linearly from 1.5 km to 4 km when going eastward. We set $\sigma^{min} = 1$ km and assigned the value 20 km to σ^{max} for the western region and 30 km for the North-eastern area; so that the standard deviations of the raw data ranges from 1km in the red area to 30 km in the northern blue area and to 20 km in the western; those of the northern complementary data from 1 km in the northern region to 30 km in the other ones, and those of the western complementary data from 1km in the western region to 20 km in the other ones.

We have tested a wide range of values of L between ten and a few hundred of kilometers for various values of σ_m and finally set: $L = 35$ km and $\sigma_m = 10$ km to obtain the model displayed in figure 6.43.

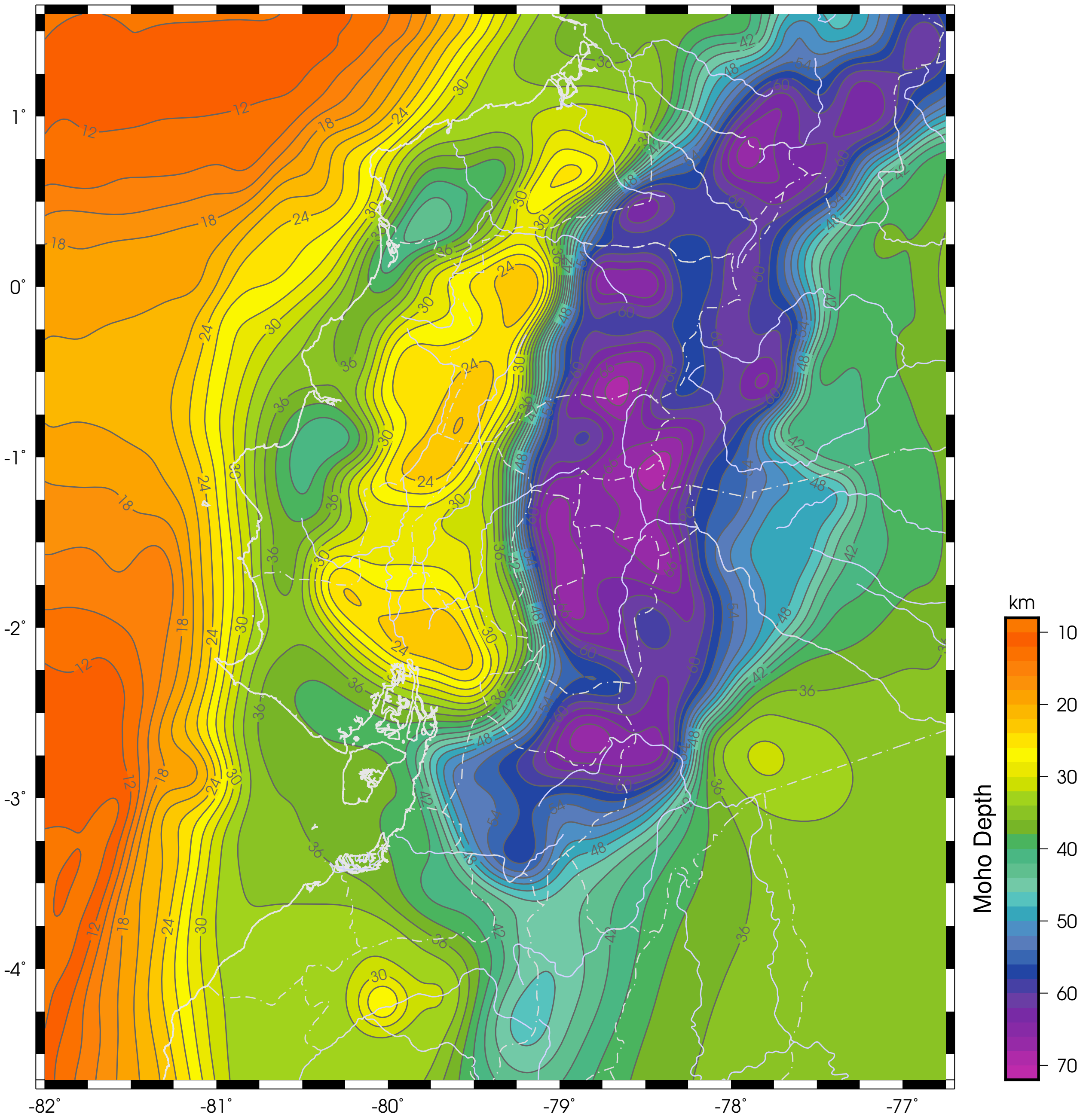

Figure 6.43: *Map of the Moho. The color scale for the depth is on the right side*

The figure allows the identification of the three sedimentary basins in the coastal region, namely the Borbón, Manabí and el Progreso basins, and further South the basin of Celica-Lancones. We can also observe the dramatic transition between the coastal plains and the Cordilleras, the InterAndean depression (outside the N-E zone), and the two holes in the continental margin North and South of the Carnegie ridge. The Eastern Andean Belt is also visible. We can also observe the impact of the points of the SAL2 profile on the model at the western limit of the Guayaquil gulf.

Finally, given the models of topography of the Moho and of the slab, we can compute the model of thickness of the mantle wedge over the region, which is displays in figure 6.44. The pattern of the figure is strongly correlated with the quasi-rigid part of the motions of the called Inca and North Andean slivers (Nocquet et al. 2014). It suggests that the promontory of the slab at the latitude of the Guayaquil gulf plays a key role in the Ecuadorian deformation at the surface.

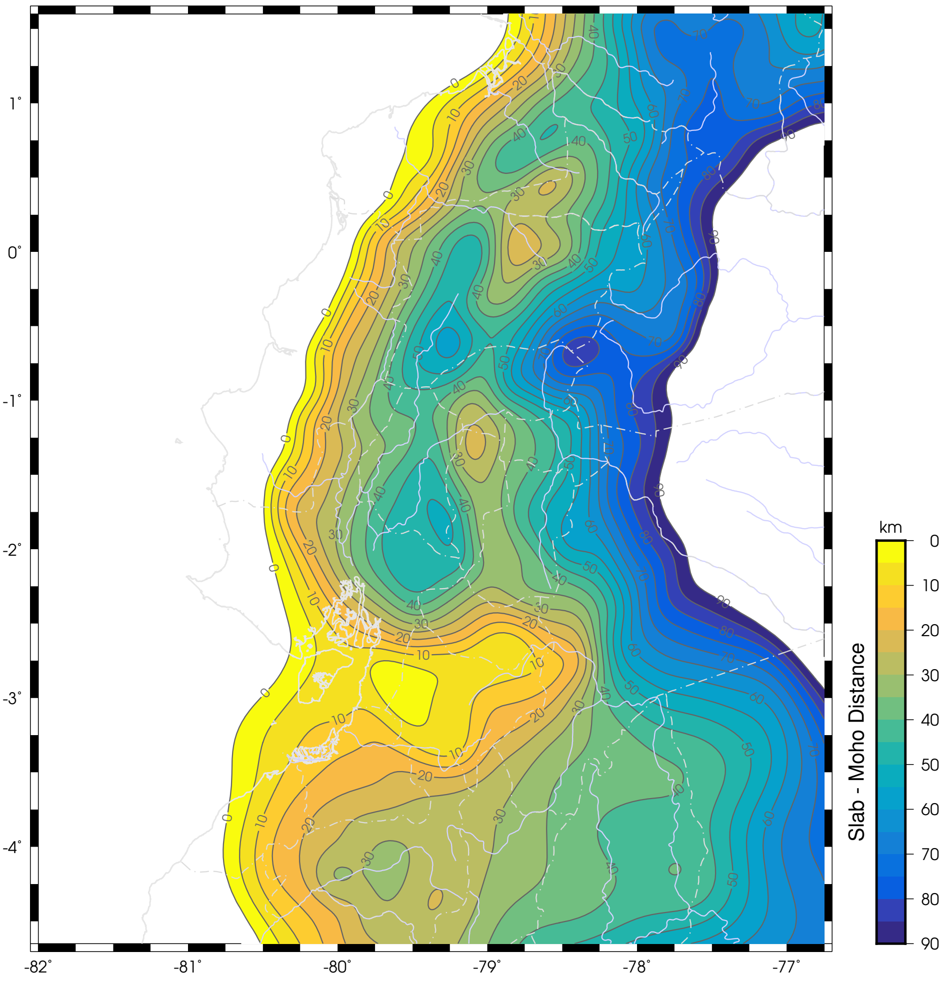

Figure 6.44: *Thickness of the mantle wedge as deduced from the slab and Moho models.*

6.7 Differential Tomography of the Cotopaxi volcano

To test the tomography by differences of arrival times between stations of a dense network described in paragraphs 2.4.2 and 4.3.4, we now focus on a relatively dense network settled on a geological object of interest: the Cotopaxi volcano.

Cotopaxi volcano (5897m) is situated 60 km south of Quito between the interandean valley and the rim of Chalupas caldera, which is part of Eastern Ecuadoran Cordillera. This active andesitic volcano, with a base diameter of 22 km and almost 3000 m of relief over the Limpiopungo plain, is covered by an icecap over the uppermost 1000 m of the cone. Rhyolitic and andesitic bimodal magmatism has occurred periodically during the past 0.5 Ma (Hall and Mothes, 2007; Mothes and Hall, 2008), and the volcanic edifice partly collapsed during the Colorado canyon episode (4,500 years BP) in the north-east sector of the volcano. The present andesitic episode that began following this event, is characterized by VEI=3-4 (Volcanic Explosivity Index) events approximately every 100-150 years (Hall and Mothes, 2008). Initial rhyolites of the 4,500 years BP Colorado canyon episode have a close chemical affinity with magmas emitted by Chalupas caldera. Interpreting U-Th isotopic data, Garrison et al. (2006) proposed that ryholites from Cotopaxi volcano are the result of partial melting of residual solidified Chalupas caldera magma. Based on the seismic monitoring of the volcano, Ruiz et al. (1998) pointed out a constant occurence of Long-Period (LP) events yielding a low activity in the volcano. This activity was located below the summit, from an altitude of 4 km down to a depth of 8 km under sea level. Métaxian et al. (2003) analyzed LP events by comparing events recorded on ice near the summit, and on rock near the snout of the north flank glacier. They pointed out that icequakes can be mixed up with volcanic LP events, when recorded far away from the glacier, because of similarities in waveform and spectral content. The seismicity significantly changed after 2000 with an increase of the number of LP events in January 2001, followed by a swarm of Volcano-tectonics (VT) earthquakes in November 2001 (Molina et al.,2008). Very Long-Period (VLP) events, presenting a broad spectral peak around 2s, were recorded by Molina et al.(2008) in 2002. By waveform inversion of these signals they determined a vertical crack extending north-south, at a depth ranging between 2 and 3 km below the north-east flank. They interpreted the VLP/LP seismicity as a magmatic intrusion that triggered a swarm of VT events. LP events were thought of as resonance of cracks located above the magma system. Since August 2015 Cotopaxi volcano has presented high phreatic activity with episode of steam explosions and ash emissions.

The strategy of tomography consists in defining a local inversion box that contains the roots of the Cotopaxi and the stations of a dense network and, based on the regional model and on the global event localisations, in inverting the differences of travel times between the stations for the structure and the locations of the events within the inversion box. The structure and the location of the events outside the inversion box remain unchanged through the inversion, but the P and S rays must be computed at each iteration over the whole reference box. For the events which are localized far away from the box, the difference of travel times between two stations of the box is mainly function of the local stucture as the rays almost coincide with each other near the source.

The inversion box is centered at 0.68° S and 78.43° W with a base of 36 km× 36 km and a vertical heigth of 30 km. Based on the local seismicity, the bottom of the box lies at a depth of 22 km below the reference ellipsoid at the center. The box contains 22 stations of the RENSIG network and 12 seismic stations of a temporary experiment conducted in 1996-1997 (Métaxian et al., 1999). Retaining only among the catalogue of 25 422 events, those of the events for which at least 5 phases were recorded by stations of the box, it remains 8 821 events for the tomography.

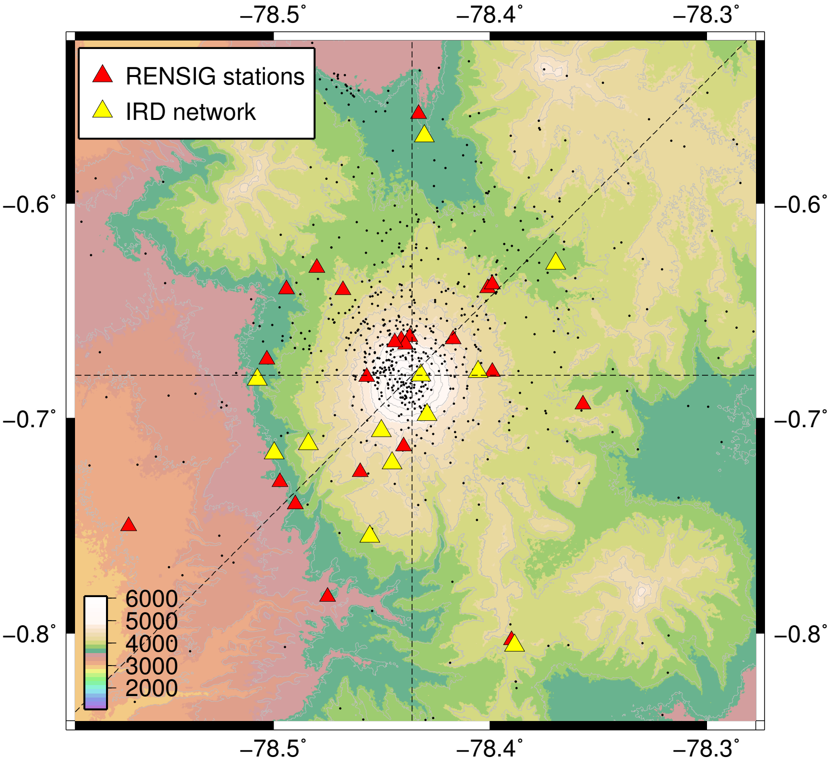

Figure 6.45: *Cotopaxi seismic stations used in study of differences tomography. Dashed lines gives direction of tomography cross sections.*

After several tests, we set the physical and tuning parameters as follows:

σ_{vP}	$\sigma_{vP/vS}$	ξ_H	ξ_V	ξ_0	λ
$750m/s$	0.10	$3km$	$3km$	$1.5km$	0.3

The resulting model is displayed in figure 6.46 along three cross sections, E-W, S-N and N 45° E.

The results for the v_P models (top figure 6.46) show a basement with velocity $v_P \simeq$ 6 km/s slopping westward toward the InterAndean valley, overlaid with lower velocity layers reaching $v_P \simeq 2.5$km/s near the surface in the south west flank of the edifice at the apex of the quasi vertical cluster of seismicity. This cluster of seismicity takes place in a region of low $v_P/v_S \simeq 1.6$ ratios that characterize gas-bearing rocks, *i.e.* high fluid compressibility, while high values of this ratio characterize liquid-bearing rocks, *i.e.* high fluid compressibility (Vanorio et al., 2005, and its bibliography). This suggests a seismic activity triggered by gaz activity rather than partial melted rocks and magma transport. The only zone of relatively high values of the ratio v_P/v_S is localized under the flank N-E of the edifice that coincides with the position of the seismic activity of long period and volcano-tectonic earthquakes detected during the unrest period of Cotopaxi in 2001-2002, and interpreted as triggered by a magmatic intrusion associated to a crack ranging at a depth of 2 or 3 km (Molina et al., 2008). We can also observe very low values (down to 1.54) of the ratio v_P/v_S beneath the S-W flank of the edifice and the Chalupas caldera situated to the east that are absolutely incompatible with the presence of partially melted rocks.

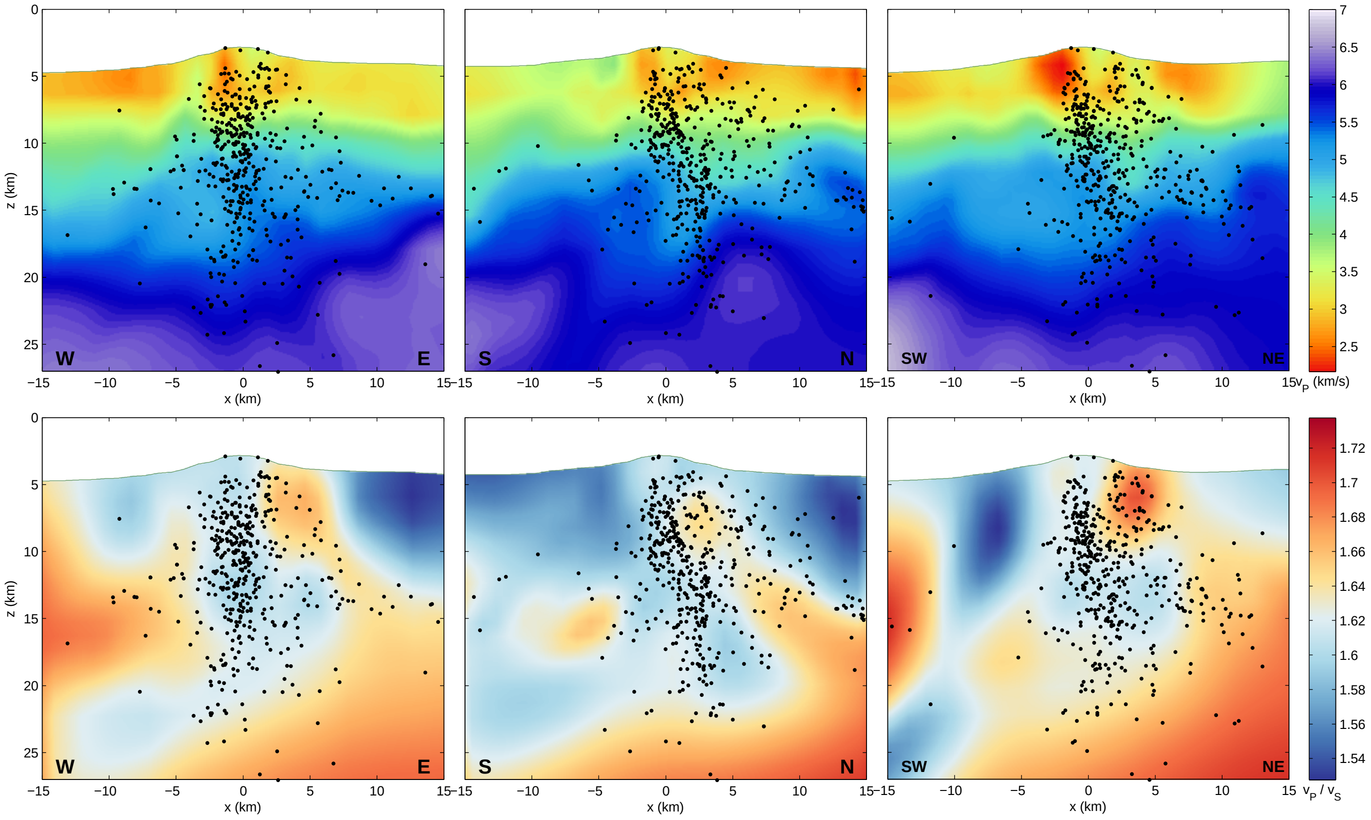

Figure 6.46: *Cross sections of v_P (top) and v_P/v_S (bottom) across the Cotopaxi volcano, oriented W-E (left), S-N (middle), N 45° E (right).*

Bibliography

[1] Abullah, N.-M., 2016, Multi-scale Seismic Tomography of the Substructures of the Merapi- Merbabu Volcanic Complex: A Preliminary Study, Master IPGP, Supervisor: B. Valette, Paris.

[2] Alvarado A., L. Audin, J.-M. Nocquet, E. Jaillard, P. Mothes, P. Jarrín, M. Segovia, F. Rolandone and D. Cisneros, 2016, Partitioning of oblique convergence in the northern Andes subduction zone: migration history and present-day boundary of the North Andean Sliver in Ecuador, Tectonics, doi: 10.1002/2016TC004117.

[3] Alvarado, A., L. Audin, J. M. Nocquet, S. Lagreulet,, M. Segovia, Y. Font, G. Lamarque, H. Yepes, P. Mothes, F. Rolandone, P. Jarrín, and X. Quidelleur, 2014, Active tectonics in Quito, Ecuador, assessed by geomorphological studies, GPS data, and crustal seismicity, Tectonics, 33, 67–83, doi:10.1002/ 2012TC003224.

[4] Alvarado A., 2012, Néotectonique et cinématique de la déformation continentale en Equateur, Thèse de doctorat, Université Grenoble Alpes.

[5] Amante C., and B.W. Eakins, 2009, ETOPO1 1arc-minute global relief model: procedures, data sources and analysis, NOAA Technical Memorandum NESDIS NGDC- 24.

[6] Araujo, S., 2012, Resolución del problema directo de tomografía sísmica: trazado de rayos sísmicos en el clúster de Pisayambo, La GRANJA, Revista de ciencias de la vida, 15(1) 2012: 48-54.

[7] Araujo, S., 2013, The Ecuadorian Moho, La Granja, Revista de Ciencias de la Vida , 18(2) 2013: 43-47, Universidad Politécnica Salesiana, Ecuador.

[8] Araujo, S., L. Troncoso, M. Ruiz, 2009, Relocalización por dobles diferencias del cluster sísmico de Pisayambo, La Granja. 10 (2). Pp. 27-34 ISSN: 1390-3799.

[9] Araujo S., B. Valette, V. Monteiller, M. Ruiz, 2014, Seismic Travel-Time Tomography of the Northern Andean Volcanic Zone in Ecuador, Geophysical Research Abstracts Vol. 16, EGU2014-12096-1, 2014 EGU General Assembly 2014.

[10] Baize S., L. Audin, T. Winter, A. Alvarado, L. Pilatasig, M. Taipe, P. Reyes, P. Kauffmann, H. Yepes, 2015, Paleoseismology and tectonic geomorphology of the

Pallatanga fault (Central Ecuador), a major structure of the south-american crust. Geomorphology, in press, 10.1016/j.geomorph.2014.02.030.

[11] Baize, S. H. Jomard, J. Champenois, P. Espin, A. Alvarado,2014, Preliminary report of the field trip to the laguna Pisayambo area (oct, 18-24,2014) 2010 Slip and cumulative faulting along Bolichuco-Taurisacha fault (BFT).

[12] Baldock, J.W., 1982, Geología del Ecuador: Boletín de la explicación del mapa geológico de la República del Ecuador, Ministerio de Recursos Naturales y Energéticos, Ecuador.

[13] Barragán, R. and P. Baby, 2002, Volcanogenic evidence of the north Andean tectonic segmentation: volcanoes Sumaco and Reventador, Ecuadorian Subandean zone, 5th International Symposium on Andean Geodynamics Universite Paul Sabatier, Toulouse, France, 16-18 Septembre 2002.

[14] Beauval, C., H. Yepes, P. Palacios, M. Segovia, A. Alvarado, Y. Font, J. Aguilar, L. Troncoso, S. Vaca, 2013, An Earthquake Catalog for Seismic Hazard Assessment in Ecuador, Bulletin of the Seismological Society of America, vol. 103, issue 2A, pp. 773-786.

[15] Belayouni, N., 2013, Nouveaux algorithmes efficaces de modélisation 2D et 3D : Temps des premières arrivées, angles à la source et amplitudes. Thèse de doctorat, École nationale supérieure des mines de Paris.

[16] Benz, H.M., B. A. Chouet, P. B. Dawson, J. C. Lahr, R. A. Page, and J. A. Hole, 1996, Three-dimensional and S wave velocity structure P of Redoubt Volcano, Alaska, J. Geophys. Res. 101, 8111–8128.

[17] Berger, P.. J-L. Got, C. Valdés, V., Monteiller, 2011, Seismic tomography at Popocatépetl volcano, Mexico, J. Vol. and Geo. Research, V. 200, Issue 3, p. 234-244.

[18] Bernal-Olaya R., Mann, P., C.A. Vargas, 2015, Earthquake, Tomographic, Seismic Reflection, and Gravity Evidence for a Shallowly Dipping Subduction Zone beneath the Caribbean Margin of Northwestern Colombia, Petroleum geology and potential of the Colombian Caribbean Margin: AAPG Memoir 108, p. 247–270.

[19] Billen, M., 2008, Modeling the dynamics of subducting slabs, Annu. Rev. Earth Planet. Sci. 2008. 36:325–56.

[20] Bonnardot M. A., Hassani R., Tric E., Ruellan and M., Régnier, 2008, Effect of margin curvature on plate deformation in a 3-D numerical model of subduction zones, Geophysical Journal International, 2008, 173 (3), p. 1084-1094.

[21] Borrero C., A. Pardo, C. M. Jaramillo, J. A. Osorio, A. Cardona, A. Flores, S. Echeverria, S. Rosero, J. García, H. Castillo, 2012, Tectonostratigraphy of the Cenozoic Tumaco forearc basin (Colombian Pacific) and its relationship with the

northern Andes orogenic build up, Journal of South American Earth Sciences 39 (2012) 75-92.

[22] Cammanaro, F., S. Goes, P. Vacher, D. Giardin, 2003, Inferring upper-mantle temperatures from seismic velocities, Physics of the Earth and Planetary Interiors 138 (2003) 197–222.

[23] Cediel, F., R. P. Shaw, and C. Caceres, 2003, Tectonic assembly of the Northern Andean Block, in C. Bartolini, R. T. Buffler, and J. Blickwede, eds., The Circum-Gulf of Mexico and the Caribbean: Hydrocarbon habitats, basin formation, and plate tectonics: AAPG Memoir 79, p. 815 – 848.

[24] Chambat, F, 1996, Figure de la Terre : Gravimétrie, régime de contraintes et vibrations. Thèse de Doctorat, Université Paris 7.

[25] Champenois, J., S. Baize, M. Vallée, H. Jomard, A. Alvarado, L. Audin, (in prep.). Evidences of surface rupture associated with a low magnitude (M5.0) shallow earthquake in the Ecuadorian Andes. To be submitted to J. Geophys. Research.

[26] Champenois, J., S. Baize, L., Audin, H. Jomard, A. Alvarado, V. Pinel, A. Hopper, H. Yepes, 2014, Applicability of InSAR techniques to study surface deformations in Ecuador, Presentation in :LMI seismes et volcanas dans les Andes du Nord, Quito,Septembre 2014.

[27] Charléty, J., S. Voronin, G., Nolet, I., Loris, F. J., Simons, K., Sigloch, and I. C. Daubechies, Global seismic tomography with sparsity constraints: Comparison with smoothing and damping regularization, Journal of Geophysical Reseach: Solid Earth, Vol. 118, 1-13, doi:10.1002/jgrb.50326, 2013.

[28] Chemenda, A., S. Lallemand and A. Bokun, 2000, Strain partitioning and interplate friction in oblique subduction zones: Constraints provided by experimental modeling, Journal of Geophysical Research, vol. 105, no. b3, pages 5567-5581, march 10, 2000.

[29] Chlieh, M., P.A. Mothes, J.-M. Nocquet, P. Jarrin, P. Charvis, D. Cisneros , Y. Font , J.-Y. Collot , J.-C. Villegas-Lanza , F. Rolandone , M. Vallée , M. Regnier , M. Segovia , X. Martin and H. Yepes, 2014, Distribution of discrete seismic asperities and aseismic slip along the Ecuadorian megathrust, Earth and Planetary Science Letters 400 (2014) 292–301.

[30] Collot, J.-Y., Charvis, P., Gutscher, M.-A., Operto, S., 2002. The Sisteur scientific party, exploring the Ecuador-Colombia active margin and inter-plate seismogenic zone. EOS 83 (185), 189-190.

[31] Collot, J.-Y., F. Michaud, A., Alvarado, B., Marcaillou, M., Sosson, G., Ratzov, S., Migeon, A., Calahorrano, A., Pazmino, 2009, Visión general de la morfología submarina del margen convergente de Ecuador- Sur de Colombia: implicaciones

sobre la transferencia de masa y la edad de la subducción de la Cordillera de Carnegie, in: Geología y geofísica marina y terrestre del Ecuador desde la costa continental hasta las islas Galápagos. Editores Jean-Yves Collot, Valenti Sallares, Nelson Pazmiño. INOCAR.

[32] Collot J-Y, Y. Legonidec, F. Michaud, B. Marcaillou, A. Alvarado, G. Ratzov, M. Sosson, E. Lopez, P. Silva, y el personal científico y técnico del INOCAR, 2006 (a), Mapas del margen continental del Norte de Ecuador y del Suroeste de Colombia : Batimetría, Releive, Reflectividad Acústica e Interpretación Geológica, publicación IOA-CVM-03-Post, INOCAR, Guayaquil, Ecuador, 2006.

[33] Collot J-Y, F. Michaud, Y. Legonidec, A. Calahorrano, F. Sage, A. Alvarado, y el personal científico y técnico del INOCAR, 2006 (b), Mapas del margen continental Centro y Sur de Ecuador : Batimetría, Relieve, Reflectividad Acústica e Interpretación Geológica, publicación IOA-CVM-04-Post, INOCAR, Guayaquil, Ecuador, 2006.

[34] Cook F.A., D.J. White, A.G. Jones, D.W.S. Eaton, J. Hall, and R.M. Clowes, 2010, How the crust meets the mantle: Lithoprobe perspectives on the Mohorovicic discontinuity and crust–mantle transition. Can. J. Earth Sci. 47: 315–351 (2010).

[35] Corredor, F., 2003, Seismic strain rates and distributed continental deformation in the northern Andes and three-dimensional seismotectonics of northwestern South America, Tectonophysics 372 (2003) 147 – 166.

[36] Crase, E., Pica, A., Noble, M., A. Tarantola, 1990, Robust elastic nonlinear waveform inversion - application to real data, Geophysics, Volume: 55 Issue: 5 Pages: 527-538 Published: may 1990.

[37] Daniel J.-Y., A. Brahic, M., Hoffert, R., Maury, A., Schaaf and M., Tardy, 2006, Sciences de la Terre et de l'Univers, Deuxieme Edition, Libraire Vuibert.

[38] Deniaud, Y., P. Baby, C. Basile, M. Ordonez, G. Montenegro, G. Mascle, 1999, Overture et évolution tectono-sédimentaire du golfe de Guayaquil: bassin d'avant-arc néogène et quaternaire du Sud des Andes équatoriennes, C.R. Acad. Sci. Paris, Sciences de la terre et des planètes, 1999. 328, 181-187.

[39] Dumont, J-F., E. Santana, M-A. Bonnardot, N. Pazmiño, K. Pedoja and B. Scalabrino, 2014, Geometry of the coastline and morphology of the convergent continental margin of Ecuador. Geological Society, London, Memoirs 2014, v. 41, p. 327-338.

[40] Eaton D. 2006. Multi-generic origin of the continental Moho: insights from Lithoprobe. J. Terra Nova, vol. 18, no 1, pp. 34-36.

[41] Egbue, O. and J. Kellogg, 2010, Pleistocene to Present North Andean "escape" , Tectonophysics, Volume 489, Issues 1–4, 20 June 2010, Pages 248–257.

[42] Ehrhard, A., 1982, Fonctions aléatoires stables : densité du maximum d'une fonction aléatoire gaussienne, Thèse de 3eme cycle : math. : Strasbourg 1

[43] Engdahl, E.,R., van der Hilst, R. Buland., 1998. Global teleseismic earth- quake relocation with improved travel times and procedures for depth determination, Bull. seism. Soc. Am., 88, 722–743.

[44] England, P., R. Engdahl, W. Thatcher, 2004, Systematic variation in the depths of slabs beneath arc volcanoes, Geophys. J. Int. (2004) 156, 377–408.

[45] Evans, J.R. , D. Eberhart-Phillips , C. H. Thurber, User's manual for SIMULPS12 for imaging vp and vp/vs: a derivative of the "Thurber" tomographic inversion SIMUL3 for local earthquakes and explosions, Open-file Report 94-431, U.S. Department of Interior, U.S. Geological Survey, USGS-OFR-94-431.

[46] Flueh, E.R., Bialas, J., Charvis, P. and the Salieri Scientific Party, 2001. Cruise report SO159 SALIERI, GEOMAR, Kiel.

[47] Fluh, E., R. Milkereit, B. Meissner, R. Meyer, R.P. Ramirez, J.E. Quintero, J. Dei, C., A. Udias, 1981, Observaciones de refracción sísmica en el Noreste colombiano en la latitud 5.5N, 83-95, In: Investigaciones Geofisicas sobre las estructuras Oceano-Continentales del Occidente Colombiano, Bogota, Colombia.

[48] Font Y., Segovia M.,Vaca S. and Theunissen T, 2013, Seismicity patterns along the Ecuadorian subduction zone: new constraints from earthquake location in a 3-D a priori velocity model. Geophys. J. Int. 13.,

[49] Gailler, A, 2005, Structure de la marge d'Equateur-Colombie par modélisation des données de sismique grand angle marines. Thèse de Doctorat. Université de Nice-Sophia Antipolis.

[50] Gailler, A., P. Charvis , E. R. Flueh, 2007, Segmentation of the Nazca and South American plates along the Ecuador subduction zone from wide angle seismic profiles. Earth and Planetary Science Letters 260 (2007) 444 – 464.

[51] Garcia Aristizabal A., H. Kumagai, P. Samaniego, P. Mothes , H. Yepes, M. Monzier, 2007, Seismic, petrologic, and geodetic analyses of the 1999 dome-forming eruption of Guagua Pichincha volcano, Ecuador, J. Vol. Geo. Research, 2007, 161 (4), p. 333-351. ISSN 0377-0273.

[52] Garrison, J., J. Davidson, M. Reid, S. Turner, 2006, Source versus differentiation controls on U-series disequilibria: Insights from cotopaxi volcano, Ecuador, Earth Planet. Sc. Lett., 244, 548–565.

[53] Gaspari G., S. E. Cohn, 1999, Construction of corrélation functions in two or three dimensions, Q. J. R. Meteorol. Soc, 125, pp. 723-75.

[54] Gerya, T.V., D. Fossati, C. Cantieni, D. Seward, 2009, Dynamic effects of aseismic ridge subduction: numerical modelling, Eur. J. Mineral. 2009, 21, 649–661.

[55] Goyes, P., 2009, Fondos marinos de soberanía y jurisdicción del Ecuador, in: Geología y geofísica marina y terrestre del Ecuador desde la costa continental hasta las islas Galápagos. Editores Jean-Yves Collot, Valenti Sallares, Nelson Pazmiño. INOCAR.

[56] Graindorge, D., A., Calahorrano, P., Charvis, J.-Y. Collot, N., Bethoux, 2004, Geophy. Res. Letters, Vol. 31, doi:10.1029/2003GL018803, 2004.

[57] Guéguen P., J.-L. Chatelain, B. Guillier, H. Yepes, J. Egred, 1998, Site effect and damage distribution in Pujili (Ecuador) after the 28 March 1996 earthquake, Soil Dynamics and Earthquake Engineering 17 (1998) 329-334.

[58] Guillier, B., J.-L.Chatelain , È., Jaillard, H. Yepes, G. Poupinet, J.-F. Fels, 2001, Seismological evidence on the geometry of the orogenic system in central-northern Ecuador (South America), Geo. Res. Letters, Vol. 28, No. 19, P. 3749-3752, October 1, 2001.

[59] Gutscher, M.-A., , J. Malavieille , S. Lallemand , J.-Y. Collot, 1999, Tectonic segmentation of the North Andean margin: impact of the Carnegie Ridge collision, Earth and Planetary Science Letters 168 (1999) 255–270.

[60] Hall M. and P. Mothes, 2007, The rhyolitic–andesitic eruptive history of Cotopaxi volcano, Ecuador, Bull Volcanol DOI 10.1007/s00445-007-0161-2.

[61] Hansen, P.C., 1992, Analysis of discrete ill-possed problems by means of the L-curve, SIAM REVIEW Vol. 34, No. 4, pp. 561-580, December 1992.

[62] Hayes, G. P., D. J. Wald, R. L. Johnson, 2012, Slab1.0: A three-dimensional model of global subduction zone geometries, J. Geo. Res., Vol. 117, B01302, doi:10.1029/2011JB008524, 2012.

[63] Hickey, J., J. Gottsmann, P. Mothes, 2015, Estimating volcanic deformation source parameters with a finite element inversion: The 2001–2002 unrest at Cotopaxi volcano, Ecuador, J. Geophys. Res. Solid Earth, 120, 1473–1486, doi:10.1002/2014JB011731.

[64] Higley, D., 2004, The Talara Basin Province of Northwestern Peru: Cretaceous-Tertiary Total Petroleum System, U.S. Geological Survey Bulletin 2206-A, U.S. Department of the Interior, U.S. Geological Survey.

[65] Huber, P.J., 1964, Robust Estimation of a Location Parameter Ann. Math. Statist. Volume 35, Number 1 (1964), 73-101.

[66] Instituto Geofísico de la Escuela Politécnica Nacional, 2013, Informe sísmico para el Ecuador año 2013, `http://www.igepn.edu.ec/informes-sismicos/anuales/231--50`.

[67] Instituto Geofísico de la Escuela Politécnica Nacional, 2012, Informe sísmico especial N 1- sismo de Quito de agosto 2014, `http://www.igepn.edu.ec/noticias/897-informe-s%C3%ADsmico-especial-n%C2%B0-1-sismo-de-quito-12-de-agosto-2014`.

[68] Instituto Geofísico de la Escuela Politécnica Nacional, 2012, Informe sísmico para el Ecuador año 2012, `http://www.igepn.edu.ec/informes-sismicos/anuales/231--50`.

[69] Jaillard E., Herail G., Monfret T. and Worner G, 2002, Andean geodynamics: main issues and contributions from the 4th ISAG, Gottingen. Tectonophysics 345 (2002) 1–15.

[70] Jaillard, E., H., Lapierre, M., Ordonez, J., Toro Alava, A., Amortegui., J., Vanmelle 2009, Accreted oceanic terranes in Ecuador: Southern edge of the Caribbean plate?. Geological Society Special Publication, 2009, Geological Society, London, Special Publications (328), pp.469-485.

[71] Jolivet, L. and H.-C. Nataf, 2001, Geodynamics, Balkema Publishers, India.

[72] Katsumata A. and N. Kamaya, 2003, Low-frequency continuous tremor around the Moho discontinuity away from volcanoes in the southwest Japan Akio. GEOPHYSICAL RESEARCH LETTERS, VOL. 30, NO. 1, 1020.

[73] Kennett B.L.N., E.R. Engdahl , R. Buland, 1995, Constraints on seismic velocities in the earth from travel times Geophys. J. Int, 122, 108-124.

[74] Koulakov, I., E.I. Gordeev, N.L. Dobretsov, V.A. Vernikovsky, S. Senyukov, A. Jakovlev, 2011, Feeding volcanoes of the Kluchevskoy group from the results of local earthquake tomography, Geophys. Res. Letters, Vol., 38, L09305, doi:10.1029/2011GL046957, 2011.

[75] Koulakov, I., B. Zaharia, B. Enescu, M. Radulian, M. Popa, S. Parolai, and J. Zschau, 2010, Delamination or slab detachment beneath Vrancea? New arguments from local earthquake tomography, Geochem. Geophys. Geosyst., 10, Q03002, doi:10.1029/2009GC002811.

[76] Koulakov, I., T. Yudistira, B.-G. Luehr, Wandorno, 2009, P, S velocity and VP /VS ratio beneath the Toba caldera complex (Northern Sumatra) from local earthquake tomography, Geophys. J. Int. (2009) 177, 1121–1139.

[77] Koulakov I., 2009, LOTOS code for local earthquake tomographic inversion. Benchmarks for testing tomographic algorithms, Bulletin of the Seismological Society of America, Vol. 99, No. 1, pp. 194-214, doi: 10.1785/0120080013

[78] Latorre, D., J. Virieux, T. Monfret, V. Monteiller, T. Vanorio, J.-L. Got, H. Lyon-Caen, 2004, A new seismic tomography of Aigion area (Gulf of Corinth, Greece) from the 1991 data set, Geophys. J. Int. (2004) 159, 1013–1031.

[79] Leeds, A. R., 1977, Mantle velocities in the Colombia-Ecuador region. Bogota, Colombia, Colombia: Univ. Javeriana, Inst. Geofis, 237-242.

[80] Legrand, D., A. Calahorrano, B. Guillier, L. Rivera, M. Ruiz, D. Villagomez, H. Yepes, 2002, Stress tensor analysis of the 1998–1999 tectonic swarm of northern Quito related to the volcanic swarm of Guagua Pichincha volcano, Ecuador, Tectonophysics 344 (2002) 15 – 36.

[81] Legrand, D., Baby, P., Bondoux, F., Dorbath, C., De Berc, S. B., Rivadeneira, M, 2005, The 1999–2000 seismic experiment of Macas swarm (Ecuador) in relation with rift inversion in Subandean foothills. Tectonophysics, 395(1), 67-80.

[82] Leveque, J.J., L Rivera, G. Wittlinger,1993, On the use of the checker-board test to assess the resolution of tomographic inversions, Geophys. J. Int.1993, Volume 115, Issue 1.

[83] Loève M., 1965, Fonction aléatoires du second ordre In P. Lévy : Processus Stochastiques et Mouvement Brownien, 2e éd., Gauthier-Villars Rééd. Jacques Gabay.

[84] Lonsdale, P, 2005, Creation of the Cocos and Nazca plates by fission of the Farallon plate, Tectonophysics 404 (2005) 237–264.

[85] Machetel, P, 2008, Les pulsations de la Terre: Dynamique du Manteau. Vuibert.

[86] Manchuel, K., M. Régnier, N. Béthoux, Y. Font, V. Sallarès, J. Díaz, H. Yepes, 2011, New insights on the interseismic active deformation along the North Ecuadorian–South Colombian (NESC) margin,Tectonics, Vol. 30, TC4003, doi:10.1029/2010TC002757, 2011.

[87] MIT Open CourseWare, 2008, Introduction to Seismology, `http://ocw.mit.edu`.

[88] Martinod, J., L. Husson, P., Roperch, B,. Guillaume and N., Espurt, 2010, Horizontal subduction zones, convergence velocity and the building of the Andes, Earth and Planetary Science Letters ,Volume 299, Issues 3–4, 1 November 2010, Pages 299–309.

[89] Martinod J., B., Guillaume, N,. Espurt, C., Faccenna, F., Funiciello and V., Regard. 2012. Effect of aseismic ridge subduction on slab geometry and overriding plate deformation: Insights from analogue modeling, Tectonophysics 588 (2013) 39–55.

[90] Mégard, F., P., Roperch , M., Lebrat, C. Laj, T., Mourier and C., Noblet, L'Occident Équatorien: Un terrain océanique pacifique accolé au continent sud-américain, Bull. Inst. Fr. Et. And. 1987, XVI, N 1-2, pp. 39-54.

[91] Meighan, H. E., U. ten Brink, J. Pulliam, 2013, Slab tears and intermediate-depth seismicity, Geo. Res. Letters, Vol. 40, 4244–4248, doi:10.1002/grl.50830, 2013.

[92] Meschede, M., U. Barckhausen and H.U. Worm, 1998, Extinct spreading on the Cocos Ridge, Terra Nova, Volume 10, Issue 4, pages 211–216, July 1998.

[93] Meschede, M. and U. Barckhausen, 2000, Plate tectonic evolution of the Cocos-Nazca spreading center. In: Silver, E.A., Kimura, G., and Shipley, T.H. (Eds.), Proc. ODP, Sci. Results, Vol. 170.

[94] Métaxian, J., S. Araujo, M. Mora, P. Lesage, 2003, Seismicity related to the glacier of Cotopaxi Volcano, Ecuador, Geophys. Res. Lett., 30, 2003.

[95] Métaxian, J.-Ph., M. Ruiz, A. Nercessian, S. Bonvalot, P. Mothes, G. Gabalda, F. Bondoux, P. Briole, J. L. Froger, D. Rémy, 1999, Geophysical Studies of Cotopaxi volcano, Ecuador: Seismicity, Structure and Ground Deformation, ISAG 1999, Symposium on Andean Geodynamics, October 4–6 , 1999, Ed. IRD.

[96] Michaud, F., C. Witt and J-Y. Royer, 2009, Influence of the subduction of the Carnegie volcanic ridge on Ecuadorian geology: Reality and fiction, GSA Memoirs 2009, v. 204, p. 217-228.

[97] Molina, I., H. Kumagai, A. García-Aristizábal, M. Nakano, P. Mothes, 2008, Source process of very-long-period events accompanying long-period signals at Cotopaxi Volcano, Ecuador, Journal of Volcanology and Geothermal Research 176(1):119-133.

[98] Molina I., H. Kumagai, J-L Le Pennec, M. Hall, 2005, Three-dimensional P-wave velocity structure of Tungurahua Volcano, Ecuador, J. Vol. Geo. Research, 147 (2005) 144 – 156.

[99] Monteiller, V., J-P Métaxian, B. Valette, S. Araujo, 2010, Seismic tomography of Cotopaxi Volcano, Ecuador: a two-scale approach, Preprint submitted to Earth and Planetary Science Letters January 22, 2010.

[100] Monteiller, 2005, Tomographie à l'aide de décalages temporels d'ondes sismiques P : développements méthodologiques et applications, Thése Doctoral, Université de Savoie.

[101] Monteiller, V., J-L., Got, J., Virieux and P. Okubo, 2005, An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano, J. Geophys. Res., Vol. 110, B12306, doi:10.1029/2004JB003466.

[102] Mosegaard K., A. Tarantola, 2002, Probabilistic Approach to Inverse Problems, International Handbook of Earthquake & Engineering Seismology, Part A., p 237-265, Academic Press, 2002.

[103] Mothes, P. and M. Hall, 2008, Rhyolitic calderas and center clustered within the active andesitic belt of Ecuador's Eastern Cordillera, Collapse Calderas Workshop, IOP Conf. Series: Earth and Environmental Science 3 (2008) 012007, doi:10.1088/1755-1307/3/1/012007

[104] Nercessian, A., A. Hirn and A. Tarantola, 1984, 3-D seismic transmission prospecting of the Mont Dore volcano, France, Geophys. J. R. astr. Soc, 307-315

[105] Neveu, J., 1968, Espaces Gaussiens, Fonctions Aléatoires Gaussiennes, Presses de l'Université de Montréal

[106] Nocquet, J.-M., Villegas-Lanza, J.C., Chlieh, M., Mothes, P.A., Rolandone, F., Jarrin, P., Cisneros, D., Alvarado, A., Audin, L., Bondoux, F., Martin, X., Font, Y., Régnier, M., Vallée, M., Tran, T., Beauval, C., Maguiña Mendoza, J.M., Martinez, W., Tavera, H., Yepes, H., 2014. Motion of continental slivers and creeping subduction in the northern Andes. Nat. Geosci. 7 (4), 287–291.

[107] Nolet, G., 2008, A Breviary of Seismic Tomography, Cambridge University Press.

[108] Ocola, L. C., Aldrich, L. T., Gettrust, J. F., Meyer, R. P., & Ramirez, J. E., 1975, Project Nariño I: Crustal structure under southern Colombian-northern Ecuador Andes from seismic refraction data. Bulletin of the Seismological Society of America, 65(6), 1681-1695.

[109] Olver, F.W., D.W. Lozier, R.F. Boisvert, C.W. Clark, 2010, NIST Handbook of Mathematical Functions, Cambridge University Press.

[110] Paige, C. and M. Saunders, 1982, LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares, TOMS 8(1), 43-71 (1982).

[111] Pavlis, G.L. and J.R. Booker, 1980, The mixed discrète-continuous inverse problem : Application to the simultaneous détermination of earthquake hypocenters and velocity structure, J. Geophys. Res., 105, 4801-481.

[112] Pazmiño N., and F. Michaud, 2009, Descripción de los Sedimentos Marinos en la Cordillera Submarina de Carnegie, in: Geología y geofísica marina y terrestre del Ecuador desde la costa continental hasta las islas Galápagos. Editores Jean-Yves Collot, Valenti Sallares, Nelson Pazmiño. INOCAR.

[113] Pedoja, K., J.F. Dumont, M. Lamothe, L. Ortlieb, J.-Y. Collot, B. Ghaleb, M. Auclair, V. Alvarez, B. Labrousse, 2006 (a), Plio-Quaternary uplift of the Manta Peninsula and La Plata Island and the subduction of the Carnegie Ridge, central coast of Ecuador , Journal of South American Earth Sciences 22 (2006) 1–21.

[114] Pedoja, K., L. Ortlieb, J.F. Dumont, M. Lamothe, B. Ghaleb, M. Auclair, B. Labrousse , 2006 (b), Quaternary coastal uplift along the Talara Arc (Ecuador, Northern Peru) from new marine terrace data, Marine Geology 228 (2006) 73 – 91.

[115] Podvin, P. and Lecomte, I., 1991. Finite difference computation of traveltimes in very contrasted velocity models : a massively parallel approach and its associated tools. Geophys. J. Int., 105:271–284

[116] Ponce, G., 2014, Analysis of the seismicity in the central margin of Ecuador, Université de Nice Sophia Antipolis Unité de Formation et de Recherche (UFR), MASTER Pro2 AGI 2013 – 2014, Mémoire réalisé en Stage

[117] Potin, B., 2016, Tomographie, mohographie et localisation des séismes dans un modèle 3D : Application aux Alpes occidentales, Thése Doctoral, Université de Grenoble.

[118] Potin B., B. Valette, F. Thouvenot, and V. Monteiller, 2015, From local earthquakes tomography to Moho topography beneath the Western Alps, Geophysical Research Abstracts Vol. 17, EGU2015-2690, 2015 EGU General Assembly 2015.

[119] Prevot, R., J.-L. Chatelain, B. Guillier, H. Yepes, 1996, Tomographie des Andes equatoriennes; evidence d'une continuite des Andes Centrales, Comptes Rendus de l'Academie des Sciences, Serie II. Sciences de la Terre et des Planetes, 1996, Vol. 323, Issue 10, pp. 833-840.

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

[121] Reguzzoni M., Sampietro D, 2012, Moho estimation using GOCE data: a numerical simulation. In: International Association of Geodesy Symposia, Geodesy for Planet Earth, Kenyon SC, Pacino MC, Marti U (eds), 136:205-214.

[122] Reyes, P., O. Dauteuil, F. Michaud, 2011, Cuantificación relativa del levantamiento de la Cordillera Costera (Ecuador): resultados preliminares a partir de geomorfología cuantitativa, 7 Jornadas en Ciencias de la Tierra. Noviembre 2011, Quito.

[123] Rhea, S., G. Hayes, A. Villaseñor, K. P. Furlong, Tarr, A.C., and Harley Benz, 2010, Seismicity of the Nazca Plate and South America: U.S. Geological Survey Open-File Report XXXX, scale 1:12,000,000.

[124] Rodgers, C.D., 2000, Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing Co. Ltd.

[125] Ruiz, M., B. Guillier, J. Chatelain, H. Yepes, M. Hall, P. Ramon, 1998, Possible causes for the seismic activity observed in Cotopaxi volcano, Ecuador, Geophys. Res. Lett., 25, 2305–2308.

[126] Sallarès V., P. Charvis, E.R. Flueh, J. Bialas and the SALIERI Scientific Party, 2004, Seismic structure of the Carnegie ridge and the nature of the Galápagos hotspot. Geophys. J. Int. (2005) 161, 763–788.

[127] Szary, W., 2014, Introduction of global plate tectonics I, Earth2Energy Educational Publishing.

[128] Scholz, C.H, 2002, The Mechanics of Earthquakes and Faulting. Cambridge University Press.

[129] Segovia, M., J. Pacheco, N. Shapiro, H. Yepes, B. Guillier, M. Ruiz, A., Calahorrano, D. Andrade, J. Egred, 1999, The august 4, 1998, Bahia earthquake (Mw=7.1): rupture mecanism and comments on the potential seismic activity, Four/il ISA G. Goettingen (Gernuinv), O.J -06110119l)l).

[130] Shelly D.R., G.C.Beroza, S. Ide and S. Nakamula, 2006, Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. Nature Vol 442—13 July 2006.

[131] Taboada A., L.A. Rivera, A. Fuenzalida, A. Cisternas, H. Philip, H. Bijwaard, J. Olaya and C. Rivera, 2000, Geodynamics of the northern Andes Subductions and intracontinental deformation (Colombia). TECTONICS, VOL. 19, NO. 5, PAGES 787-813 OCTOBER 2000.

[132] Tarantola, A., 2005, Inverse Problem Theory, SIAM.

[133] Tarantola, A. et B. Valette, 1982 a, Generalized non-linear inverse problems solved using the least square criterion, Rev. Geophys. Space Phys., 20, 2, 219-232.

[134] Tarantola, A and Valette, B. 1982 b, Inverse problems=quest for information. Journal of Geophysics, (50):159-170.

[135] Tavera H., R. Vilca and G. Marín, 2006, Inferences on the Geometry of the Nazca Plate In Northwestern Perú based on data collected by a local seismograph network, Earth Sci. Res. J. Vol. 10, No. 1 (Jun. 2006): 15-24.

[136] Theunissen T., S. Chevrot, M. Sylvander, V. Monteiller, A. Villasenor, S. Benahmed, M. Calvet, 2013, 3D crustal and lithospheric structure of the Pyrenean orogenic wedge, Geophysical Research Abstracts Vol. 15, EGU2013-4000-3, 2013 EGU General Assembly 2013.

[137] Thouvenot, F. and J. Fréchet, 2006, Seismicity along the northwestern edge of the Adria microplate. In: Pinter, N., Grenerczy, G., Weber, J., Stein, S. & Medak, D. (eds), The Adria Microplate: GPS Geodesy, Tectonics and Hazards, NATO Sci. Ser. IV. Earth and Envir. Sci. 61. Springer, Dordrecht, 335–349.

[138] Turcotte, D. and G. Schubert, 2002, Geodynamics, Second Edition, Cambridge University Press.

[139] Trenkamp R., J. N. Kellogg , J. T. Freymueller, H. P. Mora, 2002, Wide plate margin deformation, southern Central America and northwestern South America, CASA GPS observations, Journal of South American Earth Sciences 15 (2002) 157-171.

[140] Udías, A., 2000, Principles of seismology, Cambridge University Press, New York.

[141] United States Department of the Interior Geological Survey National Center for Earthquake Research, 1975, Hypo71 (Revised): A computer program for determining hypocenter magnitude, and first motion pattern of local earthquakes, Open File Report 75-31, Menlo Park, California.

[142] Valette, B., 2011, Regularization of inverse problems, Extraction and Fusion of Information for Displacement measurement from SAR Imagery, Ecole de Printemps de Physique des Houches, mai 2011.

[143] Vanorio, T., J. Virieux, P. Capuano, G. Russo, 2005, Three dimensional seismic tomography from P wave and S wave microearthquake travel times and rock physics characterization of the Campi Flegrei Caldera. Journal of Geophysical Research: Solid Earth, 110(B3).

[144] Vargas, C. A., P. Mann, Tearing and Breaking Off of Subducted Slabs as the Result of Collision of the Panama Arc-Indenter with Northwestern South America, Bulletin of the Seismological Society of America, Vol. 103, No. 3, pp. 2025–2046, June 2013, doi: 10.1785/0120120328.

[145] Veloza, G., R., Styron, and M., Taylor, 2012, Open-source archive of active faults for northwest South America, GSA Today, v. 22, no. 10, doi: 10.1130/GSAT-G156A.1.

[146] Vergely, J. L., Valette, B., Lallement, R., Raimond, S.,2010, Spatial distribution of interstellar dust in the Sun's vicinity-Comparison with neutral sodium-bearing gas. Astronomy & Astrophysics, 518, A31.

[147] Waldhauser F. and W.L. Ellsworth, 2000, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, Bull. Seism. Soc. Am., 90, 1353-1368, 2000.

[148] Waldhauser, F., 2001, hypoDD: A computer program to compute double-difference earthquake locations, USGS Open File Rep., 01-113, 2001.

[149] Witt, C. and J., Bourgois, 2010, Forearc basin formation in the tectonic wake of a collision-driven, coastwise migrating crustal block: The example of the North Andean block and the extensional Gulf of Guayaquil-Tumbes Basin (Ecuador-Peru border area), Geological Society of America Bulletin 2010;122;89-108.

[150] Yamasaki, T. and T. Seno., 2003, Double seismic zone and dehydration embrittlement of the subducting slab, J. of Geophysical Res., Vol. 108, No., B4, 2212, doi:10.1029/2002JB001918, 2003.

[151] Yepes, H., L. Audin, A. Alvarado, C. Beauval, J. Aguilar, Y. Font, F. Cotton, 2015, A new view for the geodynamics of Ecuador: Implication in seismogenic source definition and seismic hazard assessment, Tectonics, 35, doi:10.1002/ 2015TC003941.

[152] Zarifi, Z., J. Havskov, A. Hanyga, 2007, An insight into the Bucaramanga nest, Tectonophysics 443 (2007) 93 – 105.