SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di Ingegneria

Dottorato in Ingegneria Ambientale

XXXI Ciclo

STRUCTURAL JOINT INVERSION OF ELECTRICAL AND SEISMIC TOMOGRAPHY DATA

Candidate:

Ing. Lucia Palladini

Thesis advisor:

Prof. Michele Cercato

Coordinator:

Prof.ssa M. R. Boni

2017/2018

Contents

Introduction

This research project has been focused on the achievement of the structuraljoint inversion of two geophysical methods. The final target is to obtain a high resolution characterization of the shallow subsurface.

The aim of determining petrophysical properties, structural boundaries, etc, can be obtained through the integration of different information that derives from various geophysical methods. In fact, since each method is sensitive to a specific physical property, their integration can lead to an accurate final model. However, if such integration is conducted individually inverting the data sets, the final model will be affected by the resolution limitations of each method. For this reason, an important tool has been developedin geophysical applications: thejoint inversion. Two different approaches can be used to carry out the joint inversion: the petrophysical one, in which a petrophysical relationship is used, and the structural one, in which a structural similarity between models is imposed (Gallardo and Meju, 2004). Specifically, I decided to implement the algorithm for the structural joint inversion and specifically the structural approach developed by Gallardo and Meju (2003, 2004), since from literature it results to be the most robust method in the joint inversion (Moorkamp, 2017). In this process, an objective function that includes the objective function of each geophysical method is build and simultaneously minimized. In conclusion, the joint inversion may improve the resolution of each geophysical model and bring to models that are more accurate and easier to interpret.

Specifically, in this thesis, the electrical resistivity tomography (ERT) and the seismic refraction tomography (SRT) have been used to carry out the joint inversion. Both these high-resolution methods can be crucial in environmental and engineering applications, as for the geotechnical characterization of a site or for the detection of hydrological resources. Since the resistivity range overlaps for the different materials, resistivity measurements cannot be related to a specific soil or rock. Because of that, it would be better to obtain other information, for example from the seismic tomography. In fact, this method allows not only the reconstruction of the seismic wave velocities with depth, but also to obtain a good lateral resolution. Instead, the Ground Penetrating Radar (GPR) has not been considered since it presents some limits in the investigation depth, due to the high attenuation of electromagnetic energy in porous conductive media.

In addition to the integrated inversion, another goal has been obtained in this thesis: the implementation of the forward modeling for the seismic method and specifically, the Multistencils Fast Marching Method (MSFMM). This method can be seen as an extension of the FMM, that is considered from literature the fastest and the most efficient method for the solution of the eikonal

equation and accordingly for the computation of the first arrivals traveltimes. In particular, the MSFMM improves the accuracy and the efficiency of the FMM, since it considers also the information that derives from the diagonal directions.

Both the algorithms, the one of the joint inversion and the one of the forward modeling for the seismic method, have been implemented in Python language and integrated in the open-source software pyGIMLi.

In the first chapter, the definition of hydrogeophysics has been given with a little description of the different geophysical methods and their applications in the hydrogeophysics field.

In chapters 2 and 3, a little summary of the hydrogeologic and electric properties and of their mutual relationships has been made, in order to highlight the most used petrophysical relationships for the different geophysical methods, even if the joint inversion has been conducted following the structural approach and not the petrophysical one.

In chapter 4 different elements are described: the forward modelings for both electrical and seismic method; the general inversion theory and the joint inversion, with the description of the implemented algorithm for the joint inversion. Specifically, concerning the forward modelings, the finite elements method is explained for the electrical method. Instead, for the seismic method, a brief description of the FMM for both, structured and unstructured mesh, is given. In addition, the Multistencils Fast Marching Method is presented with the implementation of the method itself. Then, the general inversion theory is explained with the joint inversion and the cross-gradients function. In the end, the inversion methodology and the implementation of the algorithm used for the joint inversion are presented.

Chapter 5 presents a brief description of the open source program that has been used and specifically the pyGIMLi package.

At the end, in chapter 6 the results obtained by the implementation of the algorithm for the joint inversion and for the seismic forward modeling are shown, while, in chapter 7 the final comments and conclusion are discussed.

1. Definition of hydrogeophysics

Hydrogeophysics has been developed over the years to improve our understanding of the hydrological processes of the shallow subsurface through the application of geophysical techniques. More specifically, hydrogeophysics is capable of investigating the subsurface processes to estimate the flux parameters and transport models. This is especially important for the shallow subsurface, which contains not only the majority of water resources employed for agricultural support, but is also the recipient of most pollutants and contaminants (Rubin and Hubbard, 2005).

The importance of geophysical methods lies in the fact that traditional sampling techniques used for the characterization or monitoring of the shallow subsurfaceare point-like by definition and may not be capable to image the key hydrogeologic features of the subsoil because of lateral variations, scale effects and heterogeneity.On the other hand, geophysical methods are capable of imaging lateral and vertical variations in the physical parameters, providing high-resolution models, which are representative of the physical reality and are also useful to reduce and optimize the number of direct measurements needed for site characterization.

The choice of the geophysical technique to be used, depends on several factors; first of all, it depends on the desired resolution and depth of investigation. In addition to that, it also depends on the sensitivity of the geophysical technique to the investigated object, on the site conditions and obviously on time, funds, acquisition and interpretation costs. Generally, different geophysical techniques are used on the same site, since they are able to investigate different physical parameters (Rubin and Hubbard, 2005).

A better comprehension of the shallow subsurface is obtained by combining geophysical and hydrological data and this is the main goal of hydrogeophysics (Fig.1): sampling techniques provide point-likemeasurements, as pumping test or slug test, useful for obtaining information near a well; measurements made with tracers and variations in hydraulic load during pumping tests are suitable to obtain informations about the hydraulic properties between wells, while geophysical methods may give informations at small, medium and large scale. The data obtained from each metodology have different scales, but can be combined to obtain estimates that can be more precise and accurate when compared to the uncoupled results. However, the integration of geophysical data presents several difficulties, such as uncertainty and non-uniqueness of the inverted solutions, in addition to differences in scale between the methods and, not of secondary importance, the importance to establish reliable relationships between the investigated geophysical parameters and the hydrogeological properties. In fact, geophysical measures do not directly provide informations

about hydrogeological parameters. Most of the petrophysics relations, that link hydrogeological parameters (as porosity, type and amount of water between pores, etc.) to geophysical characteristics (as seismic velocity, electrical resistivity and dielectric constant) are not univocally known. The most used relationships in electromagnetics are Topp's and Archie's law (Lesmes and Friedman, 2005), in which electrical conductivity and porosity are linked, while in seismic methods Wyllie's law (Wyllie et al., 1956), linking seismic velocity to porosity. However, in many cases, it is necessary to obtain site-specific empirical relationships that link geophysical characteristics to the parameters of interest either in the field or in laboratory. Table 1 shows geophysical properties that each method investigates and the hydrological characteristics estimated by them.

Table 1 (Binley et al., 2015)

Geophysical Method	Geophysical Properties	Examples of Derived Properties and States
DC resistivity	Electrical conductivity	Water content, clay content, pore water conductivity
Induced polarization	Electrical conductivity, chargeability	Water content, clay content, pore water conductivity, surface area, permeability
Spectral induced polarization	As above but with frequency dependence	Water content, clay content, pore water conductivity, surface area, permeability, geochemical transformations
Self-potential	Electrical sources, electrical conductivity	Water flux, permeability
Electromagnetic induction	Electrical conductivity	Water content, clay content, salinity
Ground penetrating radar	Permittivity, electrical conductivity	Water content, porosity, stratigraphy
Seismic	Elastic moduli and bulk density	Lithology, ice content, cementation state, pore fluid substitution
Seismoelectrics	Electrical current density	Water content, permeability
Nuclear magnetic resonance	Proton density	Water content, permeability
Gravity	Bulk density	Water content, porosity

Summarizing, through geophysical methods, one or more geophysical properties are obtained, placing some sensors on the investigated area in response to a natural or artificial source. To estimate the subsoil characteristics, a relationship between the geophysical and hydrogeological properties is necessary. This is often unknown theoretically and it has to be obtained empirically.

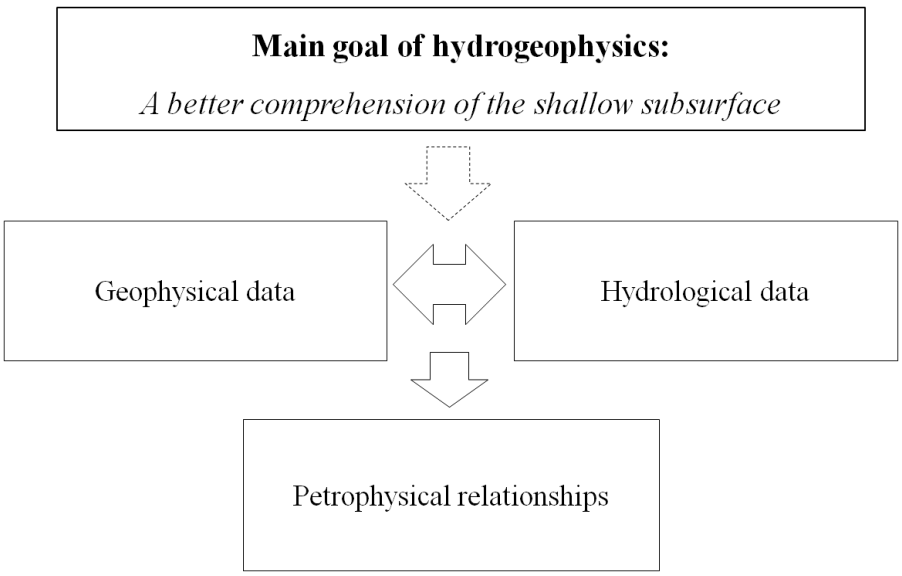

Figure 1: Scheme summarizing some concepts of hydrogeophysics. The goal of a better comprehension of the shallow subsurface can be obtained through the integration between geophysical and hydrological data. This integration requires the presence of petrophysical relationships

In the next lines, I report the applications to hydrogeophysics of the two geophysical methods that I will examine in the thesis, and specifically the resistivity and seismic methods, as explained in the Introduction.

Seismic methods provide informations about the mechanical properties of materials, and some properties, such as the compressibility module, are affected by the presence of water. Therefore, the seismic methods employing P waves can be used to derive the saturation of unconsolidated materials (Kirsch, 2006). Surface seismics methods use frequencies well below 1kHz, with maximum resolutions of about 1m or so. For example, Parra et al. (2006) aims to evaluate the applicability of reflection seismic (P waves) performed on the surface (using a frequency range 18-550Hz), as a possible technique for the hydrogeological characterization of the subsoil in a site in Florida. Some petrophysical informations, such as the porosity or permeability at the site, were known thanks to high-resolution seismic surveys carried out in aborehole (geophysical logs) with also samples taken and studied to reconstruct the site geology. The results show a good correlation between the porosity and permeability data obtained from the surface seismic with those obtained through the geophysical logs. Besides, there are many studies that propose an integrated inversion of seismic data together with hydrological data, in order to obtain more precise models, such as those of permeability in the study of Rubin et al. (1992).

The resistivity method allows to derive the spatial distribution of some soil characteristics, such as the low frequency resistivity, affected by lithology, chemical composition of the fluids and the water content of the examinated area. The method uses low frequencies and can be employed in different configurations, based on measurement target, variability of electrical properties, accessibility of the site and ability in analyzing data. Some case studies, carried out in Norway to study the flow and transport processes during the snow melting period (French et al., 2002), highlight the ability of borehole resistivity investigations (cross-borehole) to identify the areas of greatest infiltration in the subsoil, which is very useful for the development of appropriate flow and transport models. The study conducted by Binley et al (2002), focused on the unsaturated subsoil (2-10m) to obtain seasonal variations of water content in an English site and propose a correlation between electrical conductivity data and the net rain data. Specifically, the resistivity data were obtained through resistivity method performed in a single hole and subsequently converted into conductivity data through a site-specific petrophysical relationship. The study highlights the change in water content during the two years of monitoring, showing less infiltration in the winter months. In addition, it shows similar values of the water content obtained through two different methodologies: resistivity method and the GPR (100MHz), both with measurements in a borehole

up to 10m depth; the petrophysical relations used to derive the water content are the Crim's formula (Wharton et al., 1980) and the Archie's formula (Archie, 1942).

As regards the use of several geophysical techniques at a site, Paprocki (2000) highlights the good correlation between the results obtained using GPR together with the resistivity method and the neutron probe. Specifically, the conductivity map obtained through the ERT seems to be well correlated with the water content map obtained through the GPR, probably because the water content is a volume parameter and the ERT method provides a volume measurement. GPR proves to have a higher resolution than ERT, even ifit provides two-dimensional images, while ERT three-dimensional ones.

2. Hydrogeologic and electric properties of rocks and soils

One of the most important purposes of hydrogeophysics is to estimate with accuracy the hydraulic properties of the subsoil: in water saturated formations, the goal is to obtain an estimate of the porosity and hydraulic conductivity. On the other hand, in partially saturated environments, it is important to characterize the water content and the hydraulic conductivity as a function of saturation. Before examining the relationships between the hydrogeological properties and the geophysical parameters, the hydrogeological and electrical properties of rocks and soils will be briefly reviewed.

2.1. Hydrogeologic properties

Rocks and soils are composed of grains and pores; in particular, the porosity n can be defined as the ratio between the pore volume, V_p, and the total volume of the sample, V_t (Lesmes and Friedman, 2005):

$$n = \frac{V_p}{V_t} \qquad (2.1)$$

The porosity due to the space between the grains or to the rock fragments is called primary porosity, while the porosity due to tectonic fractures is called secondary porosity. Unconsolidated sediments that are made up of clayey matrix generally exhibits lower porosity than coarse unconsolidated deposits.

In the unsaturated area, the pore space is filled with air and water and the water is kept under tension (negative pressure). Water saturation, S, is defined as the ratio between the volume of water V_w and the total volume of the pores, V_p (Lesmes and Friedman, 2005):

$$S = \frac{V_w}{V_p} \qquad (2.2)$$

The water content θ is defined as the ratio of the volume of water to the total volume of the sample (Lesmes and Friedman, 2005):

$$\vartheta = \frac{V_w}{V_t} \qquad (2.3)$$

The flow of a fluid in a saturated porous medium is controlled by the saturated hydraulic conductivity K_s. For homogeneous and isotropic materials, it is given by the relation (Lesmes and Friedman, 2005):

$$K_s = \frac{k_s \cdot \rho_w \cdot g}{\mu} \qquad (2.4)$$

where k_s is the hydraulic permeability, μ is the fluid dynamic viscosity, ρ_w is the fluid density and g is the acceleration of gravity. The hydraulic conductivity characterizes the dynamic behavior of an aquifer and greatly influences the yield of a well or the dispersion of a contaminant; it is defined as the velocity of the fluid in passing through a material due to a certain pressure gradient. It has the unit of measurement of velocity (m/s).

The hydraulic conductivity is high for the sands (10^{-2}-10^{-7}m/s), while it is much lower for clayey materials (10^{-8}-10^{-11}m/s) (Domenico and Schwartz, 1990). It is mainly a function of the distribution and size of the pores, their connectivity and the tortuosity of the porous network. In granular materials, pore topology depends on grain distribution, packing and cementation (Lesmes and Friedman, 2005).

The permeability model of Kozeny and Carman (Lesmes and Friedman, 2005) assumes that the flow through a porous medium can be represented by a flow through a set of capillaries, believing that each capillary represents an independent flow path through the sample and that the effective length of this capillary, L_a, is greater than the macroscopic length of the sample, L. We can therefore define the tortuosity:

$$T = \left(\frac{L_a}{L}\right)^2 \qquad (2.5)$$

From considerations of laminar flow through the tubes, we can write the hydraulic permeability as (Lesmes and Friedman, 2005):

$$k_s = \frac{n \cdot r_h^2}{a \cdot T} \qquad (2.6)$$

where a is a tube shape factor, while r_h is the hydraulic radius. This expression can be written in a different way considering that the hydraulic radius is the reciprocal of the specific surface, S_p, which is the relationship between the area and the pore volume. Furthermore, the tortuosity can be linked to the formation factor F, through the relationship:

$$T = \left(\frac{L_a}{L}\right)^2 = n \cdot F \qquad (2.7)$$

and from that:

$$k_s = \frac{1}{a \cdot F \cdot S_p^2} \qquad (2.8)$$

that is a more practical relationship since F and S_p are physical characteristics that can be measured, unlike T and r_h. In the hazen model, instead, the hydraulic conductivity is given by (Lesmes and Friedman, 2005):

$$K_s = C \cdot d_{10}^2 \qquad (2.9)$$

where the d_{10} corresponds to the size of the grains for which 10% of the sample is finer, while the C coefficient depends on the grain sorting. The Hazen model can be applied to sediments whose d_{10} varies between 0.1 and 3.0mm. The hydraulic conductivity of partially saturated rocks and soils is a function of the saturated hydraulic conductivity, K_s, of the saturation level and of the water retention properties of the medium.

2.2. Electric properties: resistivity, conductivity and dielectric constant

Electric resistivity, ρ, is an index of the resistance encountered by the electric current passing through a rock formation. In particular, considering a conductor of length l and section S, we can say that:

$$R = \frac{\Delta V}{I} = \rho \cdot \frac{l}{S} \qquad (2.10)$$

where R is the resistance of the material, ΔV is the difference in potential and I is the current. The resistivity, ρ, is measured in Ω·m and is the opposite of the electrical conductivity, which is measured in S/m. The resistivity of sediments saturated with water is lower than the one of rocks or unsaturated sediments, but greater than the one of clayey materials.

We can define materials based on their conductivity, such as conductors (metals and metal alloys), semiconductors and dielectrics. In particular, the resistivity value measured by the resistivity method is an apparent resistivity because the current lines that allow evaluating this parameter have crossed more rock formations.

This value is not simply the average value of resistivity of the layers encountered, but it is defined by a more complex function since it depends on both the resistivities of the various lithotypes crossed and on their thicknesses.

The conductive properties of a material can be represented through the complex conductivity, σ^*, the complex resistivity, ρ^*, or the complex permittivity, ε^*, with (Lesmes and Friedman, 2005):

$$\sigma^* = \frac{1}{\rho^*} = i \cdot \omega \cdot \varepsilon^* \qquad (2.11)$$

where ω is the angular frequency, $\omega = 2\pi f$. The complex conductivity can be expressed in terms of magnitude, $|\sigma|$ and phase, φ, or by real, σ', and imaginary, σ'', components (Lesmes and Friedman, 2005):

$$\sigma^* = |\sigma| e^{i\varphi} = \sigma' + i\sigma''. \qquad (2.12)$$

The dielectric constant, *k'*, is given by the ratio between the permittivity of the sample, ε', and the permittivity of vacuum, ε_0, and it is proportional to the imaginary component of the complex conductivity:

$$k' = \frac{\varepsilon'}{\varepsilon_0} = \frac{\sigma''}{\omega \cdot \varepsilon_0} \qquad (2.13)$$

It represents the ability of rocks or soil to polarize when subjected to an alternating electric field. The electrochemical phenomena concerning the induced polarization are essentially those that originate with the contact between an electrolyte (water present in the pores of the soil) and particles with electron conductivity (granules of oxides or metal sulfides).

Moreover, *k'* varies according to the frequency of the applied field: for low frequencies all types of polarization are triggered, while for high frequencies, interfacial and electrochemical polarizations can not be triggered because they are the ones that need more time to trigger. The dielectric constant can then be estimated either through low-frequency methods, such as induced polarization, or through high-frequency methods, such as GPR. The most widely measured electrical parameters in the field are high-frequency permittivity (or dielectric constant), and low-frequency conductivity or induced polarization.

2.3. Seismic velocities

Seismic velocities for compressional, V_p, and shear waves, V_s, as defined for an elastic, isotropic medium are linked to bulk modulus, *k*, which represents the increase of density due to compression, Young's modulus, *E*, defined as the ratio between the applied uni-axial strain and the deformation that derives from it, and shear modulus, μ, which express the ratio between strain and tangential deformations, through the relations (Kirsch, 2006; Bachrach and Nur, 1998):

$$V_p = \sqrt{\frac{3k + 4\mu}{3\rho}} = \sqrt{\frac{E \cdot (1 - \nu)}{\rho \cdot (1 + \nu) \cdot (1 - 2\nu)}} \qquad (2.14)$$

and

$$V_s = \sqrt{\frac{\mu}{\rho}} \qquad (2.15)$$

where ρ is the density and v is the Poisson's coefficient (defined as the ratio between the transverse and longitudinal deformation of a medium). Specifically, the bulk modulus, k, is defined as the resistance of a material to a change in volume due to pressure, while shear modulus, μ, is defined as the ratio between the shear strength and the shear deformation. If the velocities of P and S waves are known, the elastic parameters which we have considered before, can be derived through the formulas:

$$K = \rho \cdot \left(V_p^2 - \frac{4}{3}V_s^2\right) \qquad (2.16)$$

$$E = \rho \cdot V_s^2 \cdot \frac{3V_p^2 - 4V_s^2}{V_p^2 - V_s^2} \qquad (2.17)$$

$$\mu = \rho \cdot V_s^2 \qquad (2.18)$$

$$\nu = \frac{V_p^2 - 2 \cdot V_s^2}{2 \cdot \left(V_p^2 - V_s^2\right)} \qquad (2.19)$$

In most cases a high seismic velocity corresponds to a high density. In fact, the increase in density is usually accompanied by an increase in the ability of the rock to resist to compression and shear stresses. As a result, the increase in density implies a decreaseof compressibility and an increased stiffness module (Yilmaz, 2001).

Seismic velocities are influenced by many factors, such as pressure, temperature, saturation, type of fluid, porosity, etc. P waves are influenced by the porosity of the rock and by the saturation, unlike the S waves, which depend mainly on the characteristics of the matrix and not on saturation. In fact, since fluids have a zero cutting module, S waves cannot propagate through them, and consequently the velocity will not be affected by the presence of fluids in unconsolidated materials, unlike P waves. In addition, both velocities decrease with increasing porosity and clay content. The effect of clay on the seismic properties depends on its type and position inside the rock: for example, if the clay is part of the rocky matrix and is more compressible than quartz, the speed and the impedance will decrease as the clay content increases (Kirsch, 2006). Finally, the effect of the porosity on this seismic velocity ratio is influenced by the shape of the pores, particularly for limestone, which shows micro-fractures. The ratio between velocities increases with increasing porosity, whereas for sandstones with rounded pores, this ratio does not increase as much as the porosity increases. Generally, thanks to the P wave velocity it is possible to distinguish saturated and unsaturated

materials, while it is not so easy for clayey sediments. Because of the different porosity and cementation of the grains, near surface seismic prospecting, signals may have a peak frequency in the range of 20-200Hz, with a wavelength in the range of a few meters up to some tens of meters (Rabbel, 2006). With regard to the effect of pressure on the speed of the waves, with the increase in pressure there is an increase in the speed of the seismic waves, due to the reduction in porosity. However, this increase occurs up to a certain pressure value, ie up to the closure of most of the pores. The speed will then tend to an asymptotic value in this last phase.

The typical seismic velocities for P waves are:

- 300-600m/s forunsaturated sands;
- 1500-2000m/s forsaturated sands;
- 1500-2500m/s forpeat or saturated clays;
- >3000m/s for compact rocks.

The V_p/V_s ratio can be used as a lithology indicator (Wang, 2001). In Wang et al. (2009) V_p and V_s velocities are measured in laboratory on 12 different types of rocks, such as basalt, granite, etc., and on 4 types of sulfide minerals.

3. Relations between hydrogeologic and electrical properties of rocks and soils

3.1. Permittivity models

Permittivity models are applicable to permittivity measurements made in the 100MHz -10GHz frequency range. At lower frequency the polarization mechanisms cause an increase in permittivity of rocks and soils as the frequency decreases, while for very high frequencies the permittivity decreases when the relaxation frequency of the water molecules is reached (Lesmes and Friedman, 2005). The response of the high frequency permittivity of rocks and soils is mainly influenced by the water content, since the permittivity of water, $k_w \approx 80$ is much greater than the one of dry soils, k_s=4-8 and air, k_a= 1 (Lesmes and Friedman, 2005; Kirsch, 2006). Other factors that influence the permittivity are the shape of pores and grains, the temperature and the salinity level of the aqueous solutions. The permittivity model used is different depending on the degree of saturation of a medium. The most used model to derive the porosity, *n*, for a water saturated medium, consisting of two phases, water and grains, is given by the CRIM formula (Lesmes and Friedman, 2005):

$$\sqrt{k_{eff}} = n\sqrt{k_w} + (1-n)\sqrt{k_s} \qquad (3.1)$$

making the hypothesis that the time taken by an electromagnetic wave to pass through the porous medium is given by the sum of the times used to pass through the separated phases of the material, as if the two phases were in series.

Other models can be used to derive the properties of a heterogeneous material in terms of permittivity, volume fractions and microgeometric configurations, such as the "universal formula" (Lesmes and Friedman, 2005):

$$k_{eff} = k_w + \frac{\left[\frac{(1-n)(k_s-k_w)\left[k_w+a\left(k_{eff}-k_w\right)\right]}{\left[k_w+a\left(k_{eff}-k_w\right)+\frac{1}{3}(k_s-k_w)\right]}\right]}{\left[1-\frac{\frac{1}{3}(1-n)(k_s-k_w)}{\left[k_w+a\left(k_{eff}-k_w\right)+\frac{1}{3}(k_s-k_w)\right]}\right]} \qquad (3.2)$$

in which the parameter *a*, which can range between 0 and 1, takes into account the effect of the neighboring particles on the internal electric field of a reference particle. If there are no interactions between the particles, the parameter *a=0* and then the universal formula is reduced to the Maxwell-Garnett formula (Lesmes and Friedman, 2005):

$$k_{eff} = k_w + 3(1-n)k_w \frac{k_s-k_w}{k_s+2k_w-(1-n)(k_s-k_w)} \qquad (3.3)$$

These relationships can be further modified if the shape and orientation of the grains of the medium are taken into account.

In the case of unsaturated medium, the CRIM formula is written as (Lesmes and Friedman, 2005):

$$\sqrt{k_{eff}} = \vartheta\sqrt{k_w} + (n-\vartheta)\sqrt{k_a} + (1-n)\sqrt{k_s} \qquad (3.4)$$

where θ represents the water content.

In equation (3.4) the porosity of the medium must be known to derive the water content, but the porosity is in many cases highly variable and often unknown. For this reason, some empirical equations have been obtained that use polynomial functions, obtaining the water content directly from the permittivity measures. The most used relationship is the Topp formula (Topp et al., 1980):

$$k_{eff} = 3.03 + 9.3\vartheta + 146\vartheta^2 - 76.7\vartheta^3 \qquad (3.5)$$

which has been obtained by careful measurements on five different types of soils to establish the link between the water content and the dielectric constant in the 1MHz-1GHz frequency range using the TDR probe; it provides excellent results especially for coarse and medium-textured soils, but not for fine-textured soils. When permittivity measurements are performed at frequencies smaller than 100MHz, these relationships may overestimate the water content for fine-textured soils, also because the clay response can vary greatly depending on the type of clay, the conductivity of water and of temperature.

3.2. Electrical conductivity models

The electrical response of soils and rocks can be derived using frequency-independent conductivity models. These models are applicable to low frequency conductivity measurements, in which the real part of the conductivity can be approximated with the low frequency conductivity. The conductivity of rocks and soils is influenced by several factors, such as the water content, the conductivity of the solution and the lithology of the sample, but also the shape of the grains and pores; it generally increases with concentration, mobility and charge of the ions in the solution, as well as with the temperature of the solution itself (Lesmes and Friedman, 2005). If the electrical conductivity of the solution is known, the electrical conductivity measurements can be used to derive the effective porosity of the saturated formations or the water content of the partially saturated formations (Lesmes and Friedman, 2005).

In the case of water saturated media, the most used relationship to derive the electrical conductivity of a medium is Archie's formula (Archie, 1942):

$$\sigma_{eff}=\frac{\sigma_w}{F}=\sigma_w n^m \qquad (3.6)$$

where σ_w is the water electrical conductivity, F the formation factor, n the sample porosity and m the cementation index.The formation factor F can be an indicator of the hydraulic tortuosity. The cementation index, m, assumes the value *1.3* for unconsolidated sands, while $m=2$ for consolidated formations; it increases when the grains become less spherical, while it is not affected by the granulometric assortment. The Archie'srelation assumes that the effective porosity is equal to the total one of the sample and that the electrical conductivity in soils or rocks saturated with water is due to the migration of ions into solutions through the entire porous matrix. If there are isolated pores, through which the ions cannot migrate, the actual porosity will be less than the total one and as a consequence the Archie's law leads to an overestimation of the sample conductivity (Lesmes and Friedman, 2005).

In the case of unsaturated media, Archie noted that the conductivity of the formation increases with saturation (S) according to the power law:

$$\sigma_{eff}(S)=\sigma_{sat}S^d \qquad (3.7)$$

where σ_{sat} is the conductivity of the completely saturated sample, S the saturation and d the saturation index. The saturation index, d, is *2*for consolidated rocks and varies between 1.3 and 2 for unconsolided sands. It does not provide reliable results at low saturations and especially in fine-textured soils, where the effects of the surface are not negligible. If we neglect the surface conductivity, the electrical conductivity of a partially saturated medium can be obtained through the formula:

$$\sigma_{eff}=\sigma_w n^m S^d \qquad (3.8)$$

considering that the saturation index is greater than the cementation factor, i.e.$d>m$, because as the saturation decreases, the water film around the particles becomes thinner and increases the tortuosity. Since in many cases the electrical conductivity is expressed in terms of water content, θ, assuming that $m=d$, then:

$$\sigma_{eff}=\sigma_w \theta^m \qquad (3.9)$$

If the water content is determined independently through measurements using the TDR probe, this relation can be used to derive the conductivity of water in the medium (Lesmes and Friedman, 2005).

Since the formation factor can be estimated thanks to electrical conductivity measurements, many authors have looked for an empirical relationship to derive an in situ estimate of the value of hydraulic permeability, but since*S* is much more variable than *F* for different soils and rocks, unfortunately it was not possible to determine a relationship that has general validity, but all the relationships obtained are strongly site-specific.

3.3. Relations between seismic and hydrogeological properties

The elastic properties of the rocks are highly influenced by the porosity and consequently by the seismic velocity too. For consolidated rocks, the relationship between seismic velocity and porosity is Wyllie's law (Kirsch, 2006),

$$\frac{1}{V} = \frac{1-\varphi}{V_{MATRIX}} + \frac{\varphi}{V_{PORE}} \qquad (3.10)$$

where the V_{MATRIX} is the seismic velocity of the rock matrix or of the grains, V_{PORE} is the seismic velocity of the fluid and φ is the porosity. This relationship was then modified by Raymer et al. (1980), in:

$$V=(1-\varphi)^2 \cdot V_{MATRIX}+\varphi \cdot V_{PORE} \qquad (3.11)$$

Many speed-porosity relationships obtained in situ or in the laboratory show the decrease in the speed of the compressional waves as the porosity increases.

For unconsolidated rocks, such as sand or gravel, seismic velocities are strongly influenced by both porosity and water saturation. The degree of saturation does not affect the velocity of the seismic waves below 90% of saturation, since exceeding this limit, there is a sudden increase in the speed of P-waves, and a slight decrease in the S-waves (Kirsch, 2006).

Furthermore, we can write the density of a saturated rock through the relation:

$$\rho = \rho_f \cdot \varphi - (1-\varphi) \cdot \rho_s \qquad (3.12)$$

in which φ represents the porosity, ρ_f the density of the fluid, while ρ_s the density of the solid matrix. The link between seismic velocities and porosity can also be expressed considering the compressibility module, through the relation (Kirsch, 2006):

$$\frac{K_{sat}}{K_m - K_{sat}} = \frac{K_{us}}{K_m - K_{us}} + \frac{K_{fl}}{\varphi \cdot (K_m - K_{fl})} \qquad (3.13)$$

where:

- K_{sat} is the compressibility module of saturated material;

- K_{us} is the compressibility module of unsaturated material;
- K_m is the compressibility module of rock matrix;
- K_{fl} is the compressibility module of fluid phase.

The modules of saturated and unsaturated material can be derived from the relation (Kirsch, 2006):

$$K_{sat,usat}=\rho\cdot\left(v^2_{psat,usat}-\frac{4}{3}v_s^2\right) \qquad (3.14)$$

Summarizing, in Fig.2 the most used petrophysical relationships are represented. Specifically, they are Archie's law, Wyllie's law and the CRIM's method. As specified in the text, the petrophysical relationships link the measured geophysical characteristics, as electrical conductivity, seismic velocity and dielectric permittivity, to the petrophysical parameters, as porosity, saturation and water content. Both geophysical characteristics and petrophysical parameters are highlighted in the scheme.

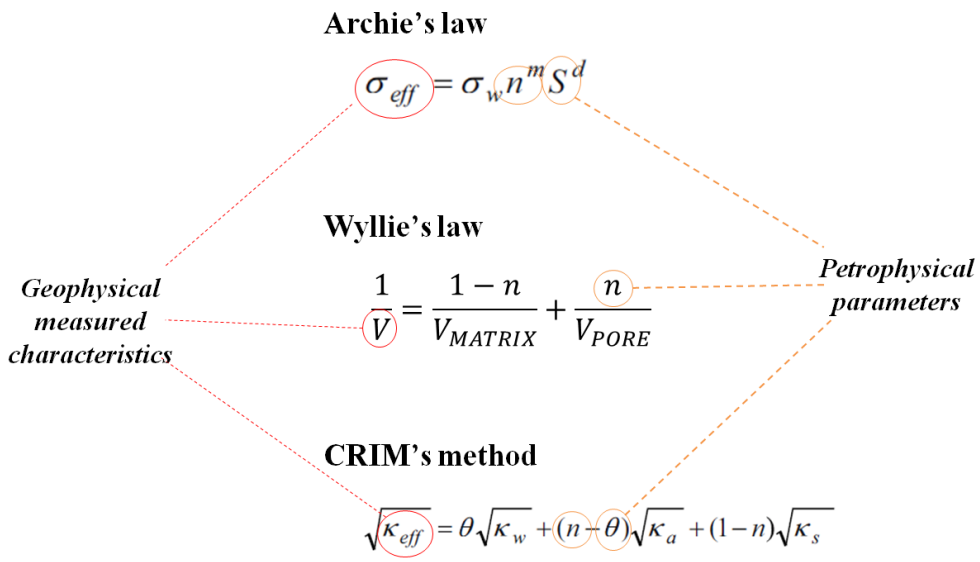

Figure2: Scheme in which the most used petrophysical relationships are represented: Archie's law, Wyllies's law and CRIM's method. They link the geophysical measured characteristics to the petrophysical parameters.

4. Inversion methodology

The main goal of geophysical methods is the derivation of the properties of the subsurface from the measurements of physical phenomena using surface or boreholes techniques: apparent resistivity data are obtained from electrical and electromagnetic methods, while seismograms from seismic methods using man-made explosions (and similar) or recording natural earthquakes (Sen, 2013).The problem of deriving the properties or a model of the earth from the geophysical observations is called *inverse problem* and is a very difficult problem because of the complex structure of the subsurfaceand of its general ill-posedness (Sen, 2013; Zhdanov, 2015). Generally, the real geology is approximated by a model, trying to determine the model parameters from the geophysical data (Zhdanov, 2015). The final interpretation will depend not only on the ability of the geophysicist in approximating the real geology using approriate models, but also in solving the inverse problem (Zhdanov, 2015). In order to choose the best model and to understand how the data are affected from it, it is necessary to calculate the theoretical data for an assumed earth model. The numerical modeling of geophysical data for some given parameters is called *forward problem* and it implies the derivation of a mathematical relationship between data and model (Menke, 1989; Sen, 2013; Zhdanov, 2015). However, since analytic solutions are impossibile to obtain for complex earth models, numerical methods based on finite difference or finite element formulations are used. In this way, predicted geophysical data are obtained for specific geological structures and compared with real data (Zhdanov, 2015). In the next paragraphs, first, I will examine the forward modellings for the two geophysical methods considered for the joint inversion (as described in the Introduction), and specifically the DC resistivity and the seismic refraction tomography, and, then, the inverse problem.

4.1. Forward Modelling

For both DC resistivity and refraction tomography, the forward problem is nonlinear: in DC resistivity, the nonlinearity is due to the interdependence of the distribution of resistivities and current flow paths (Stummer et al., 2002) and specifically, to fields of the electrical charge, that, accumulated in areas where electrical conductivity is changing, influence the magnitude of charge in other areas and specifically (Oldenburg et Li, 2005); in refraction tomography, the nonlinearity is due to the fact that the ray path depends on the velocity structure (Oldenburg et Li, 2005). In the next paragraph I will describe the theory of the two geophysical methods considered for the joint inversion: the electrical and the seismic methods.

4.1.1. Electrical methods

Introduction to the finite elements method

The forward problem, which, in the electrical methods, refers to the behavior of a medium subject to a stationary electric current and, consequently, the integration of the Poisson equation, can be solved with different methods. On way to solve the problem is the finite difference scheme, whose point of reference is the work of Dey and Morrison (1979), that is generally used in commercial software, as Res2Dinv (deGroot-Hedlin and Contestable, 1990); another approach is the finite elements method (Jin, 2002).

In particular, the finite element method is a numerical technique used in order to obtain approximate solutions to a boundary-value problem of mathematical physics. The Poisson's equation, that describes the electric filed behavior, is a differential equation that needs boundary conditions to be solved, such as Dirichlet and Neumann conditions (Jin, 2002). It is desirable to solve boundary-value problems analitically, but since an analytical solution can be obtained only for a few special problems, various approximate methods have been developed, as the Ritz and Galerkin methods (Jin, 2002), which are the most used.

The *Ritz method* is a variational method in which the boundary value problem is formulated in terms of a variational expression, that is called functional (Jin, 2002). The approximate solution is obtained by minimizing the functional, rather than the differential equation, using derivatives with respect to variables that define a certain approximation to the solution, using some expansion functions defined over the domain.

The *Galerkin's method* is based on weighted residuals: this method starts directly from differential equations that define the physical problem; instead of finding a variational system mathematically equivalent, the method uses an approximate solution, replacing this solution in the differential equation. However, since the solution is approximated, this operation causes an error, called residual error; even if it is not possible to delete this error, the usual approach is to force the integral of the error, or more generally, the product of the error and some weight function, to be zero (Pridmore et al., 1981; Jin, 2002).

In both methods it is very important to choose the trial function, defined on the entire domain, that can represent the true solution of the equation. However, since in many situations it is very difficult to find such function, especially for 2D or 3D problems, it is possible to divide the entire domain in small subdomains and use trial functions defined over each subdomain. Since the subdomains are small, such trial functions are usually simpler than entire-domain functions. The basic principle for

the finite elements method is replacing the entire domain with a certain number of subdomains, in which the unknown function is represented by simple interpolation functions with unknown coefficients (Jin, 2002; Pridmore, 1981).

Therefore, the finite element method can be outlined in the following points:

- discretization of the domain: the solution of the problem is approximated through a finite number of unknown coefficients;
- selection of the interpolation functions;
- formulation of the system of equations (e.g. Ritz or Galerkin methods can be used);
- solution of the system of equations.

The discretization of the domain is the most important step in the finite element method, since the manner in which the domain is discretized will affect the computer storage requirements, the computation time and the accuracy of the numerical results. The entire domain is subdivided into a number of subdomains, called elements. When the domain is a one-dimensional domain, the elements are short line segments interconnected to form the original line (Fig.3a); for a bidimensional domain, the elements are small triangles and rectangles. The triangular elements are used for irregular regions, instead the rectangular ones are used in the discretization of rectangular regions (Fig.3b). In a three-dimensional domain, the elements are represented by tetrahedral, triangular prisms or rectangular prisms (Fig.3c) (Jin, 2002).

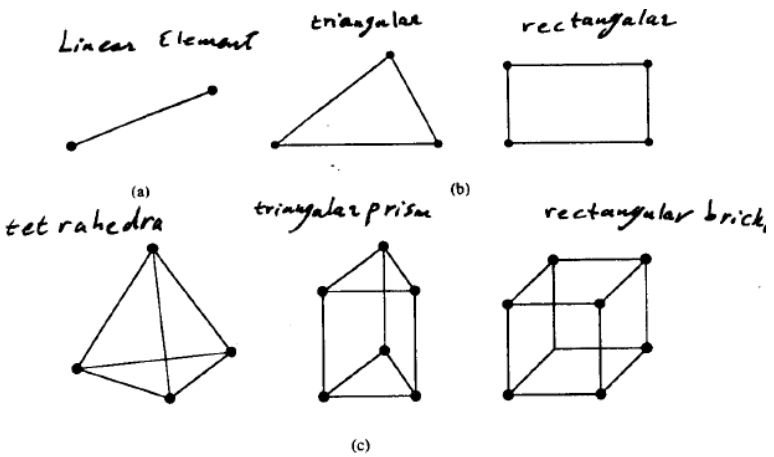

Figure 3: Finite elements: a) one-dimensional, b) two-dimensional, c) three-dimensional (Jin, 2002)

The problem is formulated considering the unknown function at the nodes of each element. More specifically, the one-dimensional element, the line, has two nodes, one at each endpoint; the triangular element has three nodes, the vertices and the tetrahedron four nodes. In the computer code, it is necessary to give a complete description of a node, consisting of the coordinate values of the nodes, the local number that indicates the node position within the element, and the global

number, which is used to assign the position within the entire domain. In the discretization we cannot have overlapping or empty places between elements and it is not possible for the vertex of an element to be inside another element. In addition to that, in the definition of the triangular elements, one has to avoid narrow angles, to avoid large computation errors (Jin, 2002). Even if increasing the number and reducing the size of the elements give in principle a more accurate solution, it produces more unknowns and consequently, much computer storage and computation time. Therefore, we can use smaller elements where the solution can have intense variations, and the larger ones where the variation is less accentuated (Jin, 2002).

The second step is the selection of an interpolation function that provides an approximation of the unknown solution within an element. The interpolation can be a polynomial of first, second or higher order; even if higher-order polynomials are very accurate, they result in a more complicated formulation than the lower-order polynomials; for this reason, the simple linear interpolation is widely used. Once the order of the polynomial has been chosen, we can obtain an expression for the unknown solution in an element, *e*, in the form:

$$\widetilde{\varphi}^e = \sum_{j=1}^{n} N_j^e \varphi_j^e = \{N^e\}^T \{\varphi^e\} = \{\varphi^e\}^T \{N^e\} \qquad (4.1)$$

where n is the number of nodes in the element, φ_j^e is the value of the unknown function φ at node j of the element, and N_j^e is the interpolation function for node j. Moreover, the $\{\cdot\}$ indicates a column vector and the superscript T denotes the transpose of the vector. An important characteristic of the interpolation functions is that they are nonzero only within the element, instead they are zero outside this element (Jin, 2002).

The third step is the formulation of the system of equation to use, for instance the Ritz variational or the Galerkin methods. Then, we set up the equations over all elements, to form the system of equations and we apply the required boundary conditions: the Dirichlet boundary condition, which assign a certain value to the function at boundary, or the Neumann condition, which requires the normal derivative of the function to vanish at the boundary (Jin, 2002). The last step is solving the system of equation, as:

$$[K]\{\varphi\} = \{b\} \qquad (4.2)$$

Application of the finite element method to the resistivity problem

Resistivity data are generally interpreted by 1D or 2D models, but considering the complex responses observed in many field environments, it is necessary to obtain 3D interpretation

(Pridmore et al., 1981; Sasaki, 1994). Numerical solutions to the bi-dimensional or three-dimensional direct current resistivity problem have been obtained in literature using different numerical techniques, as finite-difference or finite-elements techniques. The finite-difference method has been used by Dey and Morrison (1979) in order to model 2D or 3D structures, while the finite elements method has been applied to 2D resistivity problems by Coggon (1971) and to 3D resistivity problems by Pridmore et al. (1981).

First of all, we have to define the boundary value problem. Following Dey and Morrison (1979), we can write Ohm's law, that relates the current density J to electric field intensity E, and isotropic conductivity σ:

$$J=\sigma E \qquad (4.3)$$

Since the electric field E is conservative, i.e.it is the gradient of the potential u, it is possible to write:

$$E = -\nabla u \qquad (4.4)$$

If the principle of conservation of charge and the equation of continuity are considered, it can be obtained:

$$\nabla J=\frac{\partial \rho}{\partial t}\delta(x_s)\delta(y_s)\delta(z_s) \qquad (4.5)$$

where ρ is the charge density for a point whose coordinates (x,y,z) are specified in the Cartesian domain by the Dirac delta function.

In this way, we can write eq.4.3 as:

$$-\nabla\cdot\sigma\nabla u=\frac{\partial \rho}{\partial t}\delta(x_s)\delta(y_s)\delta(z_s) \qquad (4.6)$$

where (x_s,y_s,z_s) are the coordinates of the source point of the injected charge. For a point electrode at the origin (Pridmore et al. ,1981), we can write eq.4.6 as:

$$-\nabla\cdot\sigma\nabla u=I\delta(x_s)\delta(y_s)\delta(z_s) \qquad (4.7)$$

where I is the current in the electrode. This equation is known as the Poisson's equation (Binley and Kemna, 2005).

In order to obtain numerical solutions for the Poisson's equation, it is necessary to give some boundary conditions, such as the continuity of the potential φ and the continuity of the component of current normal to the boundary:

$$\frac{\partial u}{\partial n}+\alpha u=0 \qquad on \partial\Omega=\partial\Omega_s \cup \partial\Omega_I \qquad (4.8)$$

At the air-earth interface, the homogeneous Neumann condition *(α=0)* is applied to avoid current flow through the boundary along the outward normal *n* and so $\frac{\partial u}{\partial n}$ must be zero, whileσ is negligible in the air (Pridmore et al., 1981).

The finite element method (FEM) is employed to solve the Poisson's equation with given boundary conditions, following the approach of Rucker (2011). From a general point of view, two approaches can be used for the FEM method: the variational formulation and the weighted residuals. However, it has been demonstrated that for the DC method, the two approaches give the same results (Rücker, 2011) and we choose to work with the method of the weighted residual.

Summarizing shortly the Galerkin method, for a generalized partial differential equation, it is possible to obtain an approximated solution, which solves the equation with a residuum, *R*, that has to be multiplied by a weighting function, *w*. In particular, it is possible to write the weighted residual (Rücker, 2011):

$$\int_{\Omega} wR\, d\Omega=0 \qquad (4.9)$$

This expression (Eq.4.9) is then applied to Eq.4.7, and after some passages (Rücker, 2011), considering the boundary conditions too, the weak formulation is obtained (Eq.4.10):

$$\int_{\Omega} \sigma \nabla w \nabla u d\Omega+\int_{\partial\Omega_I} \sigma w \alpha u d\partial\Omega=\int_{\Omega} w \nabla \cdot j d\Omega \qquad (4.10)$$

Dividing the entire domain in a finite number of small subdomains and selecting an interpolation function that provides an approximation of the unknown solution within an element, according to the finite elements method, we can derive an approximated solution for the function *u* (Eq. 4.11):

$$u \approx u_h = \sum_{i=1}^{n} N_i \cdot u_{h_i} \qquad (4.11)$$

where *n* is the number of nodes that depends on the element that we choose and *N* are the interpolation rules for the potential u_h in each node.

Applying this approximation (Eq.4.11) to the weak formulation (Eq.4.10), it is possible to write the approximation for the DC problem, as:

$$\sum_{i=1}^{N} u_{h_i}\left(\int_{\Omega} \sigma \nabla N_j \nabla N_i d\Omega + \int_{\partial\Omega_I} \sigma\alpha N_j N_i d\partial\Omega\right) = \int_{\Omega} N_j \nabla \cdot j d\Omega for j = 1 \dots N \qquad (4.12)$$

Eq.4.12 can be written as a symmetrical system of equations **Au=b**, dividing the entire domain into C elements, that have a constant conductivity $\sigma^{(e)}$ and B_1 inner boundaries (see Rücker (2011) for more details).

Since the Dirac delta function in the source term of DC equation (Eq.4.7) leads to infinite potential gradients at the source position, this singularity can be removed splitting the total potential *u*, into a primary, u_p, and a secondary, u_s, part, as described in Coggon (1971):

$$u = u_p + u_s \qquad (4.13)$$

The primary potential, u_p, is caused by the current source in a uniform half-space with the conductivity σ_p and since it satisfies the source term, it is possible to write:

$$\nabla \cdot \left(\sigma_p \nabla u_p\right) = -\nabla \cdot j \qquad (4.14)$$

The secondary potential u_s is caused by the inhomogeneity, with the anomalous conductivity σ_s. The final formulation is written as:

$$\nabla \cdot (\sigma \nabla u_s) = \nabla \cdot \left(\left(\sigma_p - \sigma\right)\nabla u_p\right) in\Omega \qquad (4.15)$$

$$\frac{\partial u_s}{\partial n} + \alpha u_s = \frac{\partial u_p}{\partial n} + \alpha u_p = 0 \quad on\partial\Omega_I \qquad (4.16)$$

Since the system of equation has to be solved for each current source (number of electrodes used) and that the matrix A is symmetrical, sparse, positive definite and square, different numerical methods can be used and generally the conjugate gradient methods are very efficient (Rücker, 2011). I will discuss this method in depth in paragraph 4.5.

4.1.2. Seismic methods: The Fast Marching Method

A wide range of applications requires the development of numerical techniques for tracking moving interfaces or advancing fronts. In particular, in seismic applications, it is essential to compute the first traveltimes. The problem of computing first arrival traveltimes is equivalent to tracking an interface normal to seismic rays (Sethian and Popovici, 1999). Advances in numerical analysis have led to computationally efficient tools for tracking moving interfaces, such as level set methods (Sethian, 1999b). This method is intended to track interfaces whose motion is governed by the Hamilton-Jacobi equation, which is a particular class of first-order hyperbolic partial differential

equation (PDE). One important Hamilton-Jacobi equation is the Eikonal equation, which is a first-order nonlinear PDE whose solution tracks the motion of monotonically advancing fronts (Sethian and Popovici, 1999). Several methods have been proposed to solve this equation, among which the most stable and consistent is the Fast Marching Method (FMM), which can be used both for Cartesian domains and for triangulated surfaces (Fomel, 1997; Sethian, 1999a). The FMM is a grid based numerical algorithm that tracks the evolving front along a narrow band of nodes whose values are updated solving the Eikonal equation, using upwind finite-difference approximations of the gradient term (Sethian, 1999a, b; Rawlinson and Sambridge, 2004). The FMM exploits the same idea of the Dijkstra's method, which is an algorithm for computing smallest cost paths on a network (Sethian and Vladimirsky, 2000) and can be considered a continuous version of Dijkstra's algorithm (Sethian, 1999a; Bronstein et al., 2008). However, the Dijkstra's technique, that has the same computational complexity of the FMM, is numerically inconsistent, since, given two points on the graph, it produces the network with minimum length, which may not be optimal (Kimmel and Sethian, 1998, Sethian, 1999b).

We can write a brief description of the FMM. First of all, we can describe the boundary value problem (e.g. the description of an interface that is strictly expanding or contracting) (Sethian, 1999a, b). Consider a closed interface Γ, a boundary, that can be a curve in 2D or a surface in 3D and that separates one region from another. If Γ moves only in one direction (outward or inward) with speed F and has a monotonically increasing or decreasing propagation according to the sign of *F*, we can compute the arrival time *T(x)* of Γ across each point *x* to track the position of the front (Fig.4).

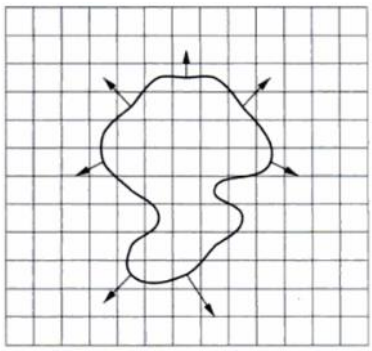

Figure 4: Crossing time calculation for an expanding front (Sethian, 1999b)

The equation of the motion of this front can be derived starting from the one-dimensional case, considering that the distance *x,* is the product between the speed *F* and the time *T* (Eq.4.17):

$$F = \frac{dx}{dT} \qquad (4.17)$$

In multiple dimensions, the equation of motion is given by (Eq.4.18):

$$|\nabla T|F = 1 \qquad (4.18)$$

in which the arrival time *T* of the initial position of the front is set to zero. If the speed depends only on the position, *x*, such equation reduces to a nonlinear first-order PDE, which is known as the Eikonal equation (that is a particular case of the more general Hamilton-Jacobi equation when F depends only on the position *x*). This equation describes a front that is always expanding or contracting and can be written also as (Eq.4.19):

$$|\nabla_x T| = s(x) \qquad (4.19)$$

where *s(x)* is the slowness, which is the reciprocal of velocity. This equation shows that the gradient of the arrival time surface is inversely proportional to the speed of the front (Rawlinson and Sambridge, 2004).

Since the evolving curve or surface can develop corners (where the evolving interface focuses), rarefaction fans (where it expands) or regions where it changes topology, it is important to deal with these behaviors accurately. The Eikonal equation cannot be easily solved in the presence of high velocity gradients, because the equation requires that gradient to be defined. One way of dealing with this problem is to seek a "weak solution", that results in a continuous T(x) but not necessarily a continuous gradient (Rawlinson and Sambridge, 2004) and specifically, to construct entropy-satisfying approximations of the gradient term. In order to find this weak solution, a viscous term can be added to the Eikonal equation:

$$|\nabla_x T| = s(x) + \varepsilon \nabla_x^2 T \qquad (4.20)$$

where ε is a smoothing parameter. The viscous term added to the Eikonal equation guarantees the solution to be smooth in the entire domain. As ε goes to zero, we find the weak solution that corresponds to the first-arriving wave front. (Sethian, 1999a; Rawlinson and Sambridge, 2004). The same solution can be obtained using the entropy condition (Sethian, 1999a, b): if the boundary curve is imaged as a source for a propagating flame, it will be satisfied the requirement that 'once a pointburns, it stays burnt' (Sethian, 1999b). In particular, this entropy condition in this case means that "once a corner has developed, the solution is no longer reversible" (Sethian, 1999b). Consequently, the propagation of a first-arriving wave frontsatisfies an entropy condition because it can only pass through a point once.

The numerical approximation of ∇T that selects the physically correct viscosity weak solution, is an upwind finite difference scheme, that uses upwind gradient operators, which take into account the

direction of information flow. In fact, in order to approximate the Eikonal equation, we can discretize the domain into a collection of grid points. In this way, we can define the gradient operators (using the Taylor expansion) and use the forward, D^+, and backward, D^-, operators (Sethian, 1999a; Rawlinson and Sambridge, 2004):

$$\left[max(D_a^{-x}T,-D_b^{+x}T,0)^2+max\left(D_c^{-y}T,-D_d^{+y}T,0\right)^2+max\left(D_e^{-z}T,-D_f^{+z}T,0\right)^2\right]_{ijk}^{1/2}=s_{i,j,k} \qquad (4.21)$$

where T is the traveltime, (i,j,k) are grid index variables in (x,y,z), s_{ijk} is the slowness and the integer variables a, b, c ,d, e, f define the order of accuracy of the upwind finite-difference operator used in each of the six cases (Rawlinson and Sambridge, 2004). In particular:

$$D_l^{-x}T_i=\frac{T_i-T_{i-1}}{\partial x} \qquad (4.22)$$

where δx is the grid spacing in x. The structure of Eq. 4.21 suggests that the information propagates from smaller values to larger values of traveltimes. To achieve this result, the FMM constructs traveltime T in a downwind fashion using the upwind known values and employing a narrow band approach (Rawlinson and Sambridge, 2004). In particular, the points that belong to the existing front are firstly freezed and the points around the existing front are put in a narrow band that marches forward. The essential characteristic of the algorithm is the selection of which gridpoint in the narrow band will be updated. The algorithm proceeds as follows: first, we tag points that have their values correctly calculated as *alive;* points that are one grid point away, within the narrow band and that have trial values, as *close*, and points that have no value calculated as *far* (Fig.5).

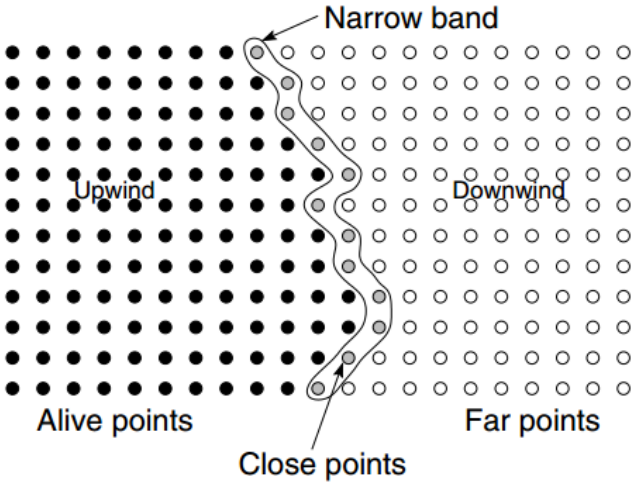

Figure 5: The principle of the method (Rawlinson and Sambridge, 2004)

The trial values are calculated using the upwind scheme with alive points only, that lie upwind of the close points. We can describe the loop with the following operations:

- identify the *close* point with minimum traveltime and tag it as an *alive* point;
- tag all neighbours of this point as *close* and compute their traveltimes;
- repeat this process to calculate the traveltimes for all gridpoints.

A heap sort algorithm (Cormen et al., 2001) is used to rapidly locate the gridpoint in the narrow band with minimum traveltime. The algorithm works because the process of recomputing the T values at upwind neighboring points cannot yield a value smaller than any of the accepted points. The scheme is very fast and solves the equation in O (N logN) steps, where N is the number of total points in the grid. The Eikonal equation is solved at each grid point and the accuracy of the method is dependent on the mesh size and on the order of the finite-difference approximation that is used (Sethian and Popovici, 1999; Rickett and Fomel, 2000).

The Fast Marching Method on triangulated surfaces

The FMM, in the attempt to solve the Eikonal equation and to determine the propagation of first-arrival seismic wave fronts given a slowness distribution, was focused in the first applications (Sethian and Popovici, 1999; Rickett and Fomel, 2000) on regular rectilinear 2D grids. Then, it was extended to unstructured 2D triangular grids thanks to Kimmel and Sethian (1998), Sethian (1999a, b) and Rawlinson and Sambridge (2004). Fomel (1997) obtained an equivalent result using a different formulation of the problem and in particular in terms of the Fermat's principle in the case of eikonal solvers. This approach was used also in Lelièvre et al. (2011).

Following the procedure of Fomel (1997) updated in Lelièvre (2011), we can consider the triangle ABC (Fig.6). The traveltimes in A and B are t_A and t_B respectively and are known quantities since they have been calculated in the previous step of the algorithm and the problem is to calculate the traveltime in node C. The assumption is that the wave fronts that travel through the cell are linear and that any ray travelling through the cell is also linear (Lelièvre et al., 2011). In the figure, ξ is a normalized distance along the line from A to B and ρ is the distance from node C to a point ξ on line AB (as in Fig.6). In the end, ξ_0 is the normalized projection of node C on line AB and ρ_0 is the length of the normal from C to ξ_0.

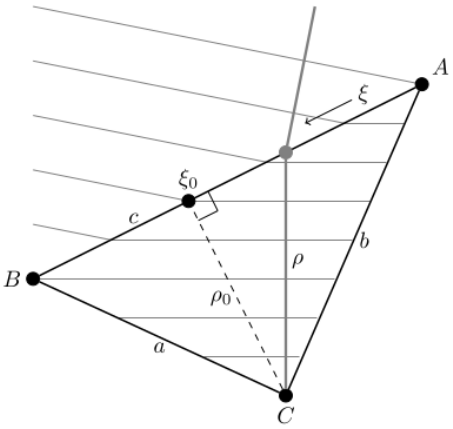

Figure 6: Triangular cell ABC. Wave fronts are represented as thin grey lines, instead the ray throught node C is represented as a thick grey line that is perpendicular to wave fronts

The traveltime in C can be written as:

$$t_C(\xi)=t(\xi)+s\rho=t_A+u\xi+s\sqrt{c^2(\xi-\xi_0)+\rho_0^2} \qquad (4.23)$$

where $u=t_B-t_A$ with $t_B>t_A$ (the front passes firstly through A and then through B), c is the length of line AB (since ξ is a normalized distance) and s is the slowness of the triangular cell.

The term $t_A+u\xi$ is the linearly interpolated traveltime at point ξ on line AB and $s\rho$ is the traveltime for the path from that point to node C, calculated as slowness multiplied by distance, following Pythagora's theorem considering the triangle $C\xi_0\xi$. Since Fermat's principle states that light traveling between two points always chooses the quickest path, in this case, the ray to C should correspond to a local traveltime minimum. Therefore, we search accordingly for the value of ξ that minimizes the traveltime through the cell and therefore, equation 4.23.

After some algebra, Lelièvre et al. (2011) found the simple expression for t_C:

$$t_C=t_A+u\xi_0+wc^{-1}\rho_0 \qquad (4.24)$$

that is equivalent to the one derived by Fomel (1997), where w is the square root of:

$$w^2=s^2c^2-u^2\geq 0 \qquad (4.25)$$

The traveltime calculated in this way corresponds to a wave transmitted through the cell, but considering also the traveltimes associated to head waves along sides AC and BC, as in Podvin and Lecomte (1991):

$$t_C=min\left(t_A+u\xi_0+wc^{-1}\rho_0,\ t_A+sb,\ t_B+sa\right) \qquad (4.26)$$

As discussed in Kimmel and Sethian (1998), the traveltime of the transmitted wave is accepted only if the solution comes from within the triangle, i.e. 'the gradient of the solution at a grid point should point into the triangle from which it is updated'. In fact, the ray that passes to node C should travel through the cell and intersect AB face. Moreover, Kimmel and Sethian (1998) states that in this procedure acute triangulation are required because a front that enters the side of the triangle, have to arrive to two points before the third is reached (monotonicity condition). This restriction ensures that the wave front marching is performed in a monotonic manner, so that each new traveltime is computed later than those already determined. As in Rawlinson and Sambridge (2004), we can consider fig.7:

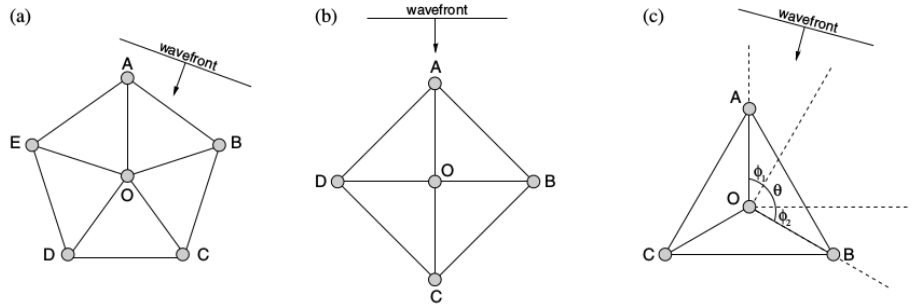

Figure 7: The figure represents a wavefront impinging on different types of triangulated domain. a) Acute triangles, b) Right angle triangles and c) obtuse triangles.

If we want to calculate the traveltime in point O, in the case of acute triangles (Fig.7a) the front passes through two adjacent nodes before node O and consequently the solution is accurate. In the domain with right angles, the wave front passes through two adjacent points before point O, except when the front is perpendicular to AO, BO, CO or DO. In these cases, the solution is anyhow accurate because the traveltime is given for example for the edge AO by

$$T_O = T_A + s_0 AO \qquad (4.27)$$

In case of obtuse triangles (Fig.7c), wavefronts that have normals that lie in the zone defined by θ will pass through A and B before O. Instead wavefronts with normals that lie in φ zones, will pass through A or B before O. Consequently, the solution is not accurate.

The assumption that has been made for the local traveltimes update, is to assume linear wavefronts at the scale of the grid cells, leading to the first-order modeling accuracy. However, this assumption can cause errors close to a point source, where the wave fronts are almost circular (in 2D). In reality, even in a homogeneous medium the wave fronts are spherical and they can be considered as planar only at large distances from the source (Lelièvre et al., 2011). Several authors have dealt with this problem. Lelièvre et al. (2011) proposed to make the initialization radius as large as possible, but this holds only in cases where there is a homogeneous near-source region. If a

heterogeneous area exists near the source, a finer discretization can be used in its vicinity, and a coarser grid far enough from the source.

4.1.3. The Multistencils Fast Marching Method

The main advantage of FMM compared to the other techniques is that it maintains a narrow band that separates the grid points of known solutions from the unknown ones. In addition to that, the solution at every grid point is computed according to an order that is consistent with the way the wave fronts propagate. However, this method is not accurate along diagonal directions, because it doesn't consider the information along diagonal directions and consequently, large numerical errors can occur along these directions (Hassouna and Farag, 2007). For this reason, different approaches have been introduced to improve the accuracy and the efficiency of the FMM as originally proposed. In particular, Hassouna and Farag (2007) implemented a highly accurate method in order to solve the Eikonal equation on Cartesian domains: the Multistencils Fast Marching Method (MSFMM). This method comes from the FMM, but solves the Eikonal equation at each grid point using different stencils that cover the entire neighbor points. The resulting solution have to satisfy the upwind condition, as in the FMM, which means that, in order to construct the solution downwind, we have to use the information upwind, and consequently, the known times calculated in the previous steps, using the difference schemes. Stencils that are not aligned with the natural coordinate system can be used and, in these cases, the Eikonal equation is approximated using directional derivatives, where higher order finite difference schemes are used.FollowingHassouna and Farag (2007), we consider the stencil S_v (Fig.8), that intersects the Cartesian domain at the grid points p_1, p_2, q_1 and q_2, the unit vectors $\overrightarrow{r_1} = [r_{11} r_{12}]^T$ and $\overrightarrow{r_2} = [r_{21} r_{22}]^T$ along $\overline{p_2 p_1}$ and $\overline{q_2 q_1}$ respectively (Fig.8).

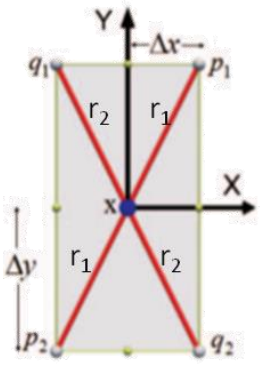

Figure 8: Stencil in which we consider the diagonal grid points

Therefore, we can consider the directional derivatives U_1 along $\overrightarrow{r_1}$ and U_2 along $\overrightarrow{r_2}$:

$$U_1=\overrightarrow{r_1}\cdot\nabla T(x)=r_{11}T_x+r_{12}T_y \qquad (4.28)$$

$$U_2=\overrightarrow{r_2}\cdot\nabla T(x)=r_{21}T_x+r_{22}T_y \qquad (4.29)$$

Thus, $U = R\nabla T(x)$ and, after some algebraic manipulation (see Hassouna and Farag (2007) for more details), we derive:

$$|\nabla T(x)|^2=U^T(RR^T)^{-1}U=s^2(x) \qquad (4.30)$$

where:

$$(RR^T)^{-1}=\frac{1}{\sin^2\varphi}\begin{pmatrix}-1 & \cos\varphi\\ \cos\varphi & -1\end{pmatrix} \qquad (4.31)$$

Substituting (4.31) into (4.30), the authors obtained a closed-form expression for the gradient of the arrival time along an arbitrary stencil:

$$U_1^2-2U_1U_2\cdot\cos(\varphi)+U_2^2=\sin^2\varphi\times s^2(x) \qquad (4.32)$$

If we consider the specific case in which $\Delta x = \Delta y = h$ and $\varphi = 90°$, (Fig.9) equation (4.32) becomes:

$$U_1^2+U_2^2=s^2(x) \qquad (4.33)$$

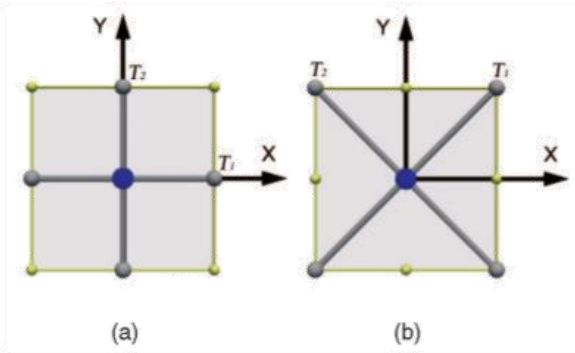

Figure 9: a) S1 stencil; b) S2 stencil. T1 and T2 are the minimum arrival times along the arms of the stencils.

Considering the S1 stencil (Fig.9a), using the first order approximation of the directional derivative, we have to solve equation (4.34):

$$\sum_{v=1}^{2}\max\left(\frac{T(x)-T_v}{h},0\right)^2=s^2(x) \qquad (4.34)$$

Conversely, for the second-order approximation of the directional derivative, the equation is (4.35):

$$\sum_{v=1}^{2}\max\left(\frac{3}{2h}[T(x)-T_v],0\right)^2=s^2(x) \qquad (4.35)$$

If we consider the stencil S2, that is aligned with the diagonal neighbors, for the first order approximation of the directional derivative, we have to solve the equation 4.36:

$$\sum_{v=1}^{2} max\left(\frac{T(x)-T_v}{\sqrt{2}h},0\right)^2 = s^2(x) \qquad (4.36)$$

and for the second-order approximation (Eq.4.37):

$$\sum_{v=1}^{2} max\left(\frac{3}{2\sqrt{2}h}[T(x)-T_v],0\right)^2 = s^2(x) \qquad (4.37)$$

These equations reduce to second order equations, whose solutions have to satisfy the causality relationship, which means that the information propagates from smaller to larger values of T. Specifically, the minimum solution is rejected because it does not maximize the equation, and the other solution is accepted if it is higher than the adjacent neighbors, T_1 and T_2. If this check is not true, the accepted solution is:

$$min\left(T_v + \|x-x_v\| \times s(x)\right), \qquad v=1,2 \qquad (4.38)$$

4.1.4. Implementation of the Multistencils Fast Marching Method

The proposed implementation of the Multistencils Fast Marching Methodwas carried out consideringa regular grid with cells with the same dimensions in x and y directions. The code has been written in Python programming language and integrated in my version of the pyGimli package (Rücker et al., 2017). The structure of the code was inspired by the Matlab code written by Kroon (2011).

First of all, I will describe the general structure of the code, and then I will examine in depth the most significant passages.

After the definition of the grid and of the sources for the seismic investigation, I assigned a value of slowness to each cell of the grid. Then, I individuated the cell in which the source was positioned and calculated the traveltimes analitically for the four nodes of the cell itself. The nodes whose value of time was calculated in this way were considered as *upwind* or *alive nodes*, while the remaining nodes were initialized with infinite value and are marked as *faraway or downwind nodes*. As in the structure of the FMM, the next step was the detection of the narrow-band nodes, which are close to the upwind ones calculated before. Each node of the narrow-band was examined and updated considering the known traveltimes and solving the Eikonal equation as in equations 4.34,

4.35, 4.36 and 4.37 (I will describe in detail the followed method in the next lines). Among all the narrow-band nodes, I extracted the point with the minimum traveltime and marked this point as upwind. Then, I individuated the nodes close to the upwind one, adding them in the narrow band. The traveltimes were again calculated as before and the process was repeated for all the nodes of the grid.

The update procedure is based on the upwind approximation to the eikonal equation. Specifically, as I said before, in the MSFMM we can consider two types of stencils: the S_1 stencil that is aligned with the natural coordinate systemand the S_2 stencil, which is aligned with the diagonal neighbors (Fig.9). Furthermore, for each stencil, we can consider the first or the second-order approximation of the directional derivative. In the end, we can have four cases:

- case A: the MSFMM considering only the S_1 stencil with a first order approximation of the directional derivative (Eq.4.34), which is the classic FMM;
- case B: the MSFMM considering not only the first, but also the second order approximation of the directional derivative for the S_1 stencil (Eq.4.34-Eq.4.35), which is the classic FMM with higher accuracy;
- case C: the MSFMM, considering also the stencil S_2 with the first order approximation of the directional derivative (Eq.4.34 and Eq.4.36);
- case D: the MSFMM, considering both the stencilsand using not only the first, but also the second order approximation of the directional derivative (Eq.4.34, Eq.4.35, Eq.4.36 and Eq.4.37).

The element which I examined in depth is linked to the value of slowness used in the update procedure and specifically in equations 4.34, 4.35, 4.36 and 4.37. Infact, I considered the slowness associated to each cell, as in Fomel (1997) and Lelièvre et al. (2011) in the traveltime computation, whilein literature (Sethian, 1999a; Hassouna and Farag, 2007) and in the Kroon's code, the value of slowness considered is the one associated to the node that has to be updated.

Specifically, as I said before, in the implementation of the classic FMM (case A), I considered the S_1 stencil (Fig.10) and Eq.4.34 that I write again for clarity:

$$\sum_{v=1}^{2} max\left(\frac{T(x)-T_v}{h},0\right)^2 = s^2(x) \qquad (4.34)$$

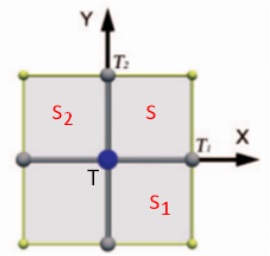

Figure 10: S1 stencil. T1 and T2 are the minimum arrival times along the arms of the stencils. s, s1 and s2 are the value of slowness for the three cells.

If only T_1 or T_2 are the known traveltimes in order to update point T in Fig.10, the head waves only will be considered and consequently, the traveltime for such point will be calculated using Eq.4.39 or 4.40 respectively:

$$max\left(\frac{T(x)-T_1}{h},0\right)^2=\left(min(s,s_1)\right)^2 \qquad (4.39)$$

$$max\left(\frac{T(x)-T_2}{h},0\right)^2=\left(min(s,s_2)\right)^2 \qquad (4.40)$$

The value of the slowness has been chosen as the minimum between the values of the adjacent cells, since, for the Fermat's principle, the ray follows the path that requires the shortest time.

If both T_1 and T_2 are the known traveltimes, not only the head waves, but also the transmitted wave has to be considered. Specifically, the best accuracy for such implementation is obtained computing the minimum traveltime t_T (Eq.4.44) between the ones calculated through equations 4.41, 4.42 and 4.43:

$$max\left(\frac{T_A(x)-T_1}{h},0\right)^2=\left(min(s,s_1)\right)^2 \qquad (4.41)$$

$$max\left(\frac{T_B(x)-T_2}{h},0\right)^2=\left(min(s,s_2)\right)^2 \qquad (4.42)$$

$$\sum_{v=1}^{2} max\left(\frac{T_C(x)-T_v}{h},0\right)^2=s^2 \qquad (4.43)$$

$$t_T=min(T_A,T_B,T_C) \qquad (4.44)$$

Specifically, for the head waves (Eq.4.41 and 4.42), I will consider for the slowness the minimum value between two adjacent cells (Fig.10), while for the transmitted wave (Eq.4.43) I will consider the slowness of the cell individuated by the two nodes with T_1 and T_2 traveltimes (Fig.10).

The same scheme has been followed for the classic FMM but considering the second-order approximation of the directional derivative (case B). Specifically, considering again Fig.10, if only T_1 or T_2 are the known traveltimes, t_T has been calculated using equations 4.45 or 4.46 respectively:

$$max\left(\frac{3}{2h}[T(x)-T_1],0\right)^2=\left(min(s,s_1)\right)^2 \qquad (4.45)$$

$$max\left(\frac{3}{2h}[T(x)-T_2],0\right)^2=\left(min(s,s_2)\right)^2 \qquad (4.46)$$

While, if both T_1 and T_2 are the known traveltimes, I will consider transmitted and head waves, calculating t_T as:

$$t_T=min(T_D,T_E,T_F) \qquad (4.47)$$

where T_D, T_E and T_F are calculated as:

$$max\left(\frac{3}{2h}[T_D(x)-T_1],0\right)^2=\left(min(s,s_1)\right)^2 \qquad (4.48)$$

$$max\left(\frac{3}{2h}[T_E(x)-T_2],0\right)^2=\left(min(s,s_2)\right)^2 \qquad (4.49)$$

$$\sum_{v=1}^{2} max\left(\frac{3}{2h}[T_F(x)-T_v],0\right)^2=s^2(x) \qquad (4.50)$$

Considering the S_2 stencil (Fig.11),

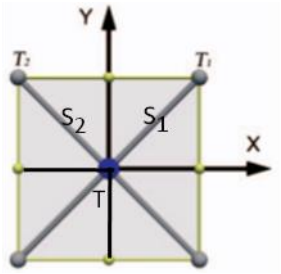

Figure 11: S2 stencil. T1 and T2 are the minimum arrival times along the arms of the stencils. s1 and s2 are the value of slowness for the two cells.

for the first order approximation of the directional derivative (case C), I considered Eq.4.36 that I write again for clarity:

$$\sum_{v=1}^{2} max\left(\frac{T(x)-T_v}{\sqrt{2}h},0\right)^2=s^2(x) \qquad (4.36)$$

Referring to Fig.11, if T_1 or T_2 are the known traveltimes, t_T will be calculated through equations 4.51 or 4.52 respectively:

$$max\left(\frac{T(x)-T_1}{\sqrt{2}h},0\right)^2=s_1^2(x) \qquad (4.51)$$

$$max\left(\frac{T(x)-T_2}{\sqrt{2}h},0\right)^2=s_2^2(x) \qquad (4.52)$$

In both cases the considered slowness is the one individuated by the nodes with traveltimes t_T and T_1 or T_2 respectively.

If both T_1 and T_2 are known, the traveltime in point T will be calculated with relation 4.53, considering for the slowness, the minimum value between the one of the two adjacent nodes:

$$\sum_{v=1}^{2} max\left(\frac{T(x)-T_v}{\sqrt{2}h},0\right)^2=\left(min(s_1,s_2)\right)^2 \qquad (4.53)$$

In the end, the same reasoning has been used for the S_2 stencil considering the second order approximation of the directional derivatives (case D). For T_1 or T_2 known traveltimes, equations 4.54 or 4.55 are used:

$$max\left(\frac{3}{2\sqrt{2}h}[T(x)-T_1],0\right)^2=s_1^2(x) \qquad (4.54)$$

$$max\left(\frac{3}{2\sqrt{2}h}[T(x)-T_2],0\right)^2=s_2^2(x) \qquad (4.55)$$

while for both T_1 and T_2 known traveltimes, Eq.4.56 was taken into account:

$$\sum_{v=1}^{2} max\left(\frac{3}{2\sqrt{2}h}[T(x)-T_v],0\right)^2=\left(min(s_1,s_2)\right)^2 \qquad (4.56)$$

In the end, the accuracy of this method improves by summing the coefficients of the two stencils.

4.2. Inversion Theory

Geophysical methods, like electric, electromagnetic or seismic ones, study different physical fields that are propagated in the subsurface in order to image its geological structures and estimate the associated physical parameters (Zhdanov, 2015). A key step in the geophysical data interpretationconsists in the comparison between the measured data and the synthetic or predicited data computed for the selected earth models. More specifically, the synthetic data are generally computed byanalytical or numerical modeling for the assigned earth model: this is called the forward modeling.

When interpreting geophysical data, one has to infer the subsurface geometry and physical properties from the observed data, which are by definition inaccurate due to experimental errors. The resulting inverse problem, provides an approximation of the real geology with a parametric model, whose theoretical response (predicted data) reproduce the field data within the experimental accuracy (Zhdanov, 2015). Due to the inherent ill-posedness in geophysical inversion, one has to consider:

- if the solution exists, since it might happen that the solution physically exists but from the mathematically point of view, there might be no adequate numerical model that fits the observed data;
- if the solution is unique, since different models can generate the same data;
- if the solution is stable, since the data are always contaminated with noise and one has to evaluate if the difference in the responses of different models is larger than the noise level (Zhdanov, 2015).

Practically, the subsurface domain is divided into many cells, using a structured or an unstructured mesh. The structured meshesare characterized by a regular connectivity and specifically, they can be composed of quadrilateral elements in 2D and hexahedral elements in 3D. The unstructured meshes are identified by irregular connectivity generally composed of triangles and tetrahedral in 2D and in 3D respectively. If the inverse problem is an even-determined problem, there are as many data as the number of model parameters, and consequently, there is only one solution. The inverse problem can be underdetermined, i.e. there are fewer data than the number of parameters, when we try to obtain estimates of the earth model parameters that are continuous from a finite set of measurements (Sen, 2013). In this case, many solutions exist and the model parameters estimates are non-unique. When there are more data than the number of model parameters, the problem is overdetermined; the “best” estimate (in some sense) can be obtained in a least squares sense.

However, in most situations the problem is mix-determined, as in seismic tomography problems, in which some cells are crossed by several rays, while others by none.

In our inverse solution, each cell has an assigned constant value of the physical property of interest. Since there are more cells than measured data, the problem is underdetermined and it can be solved as an optimization problem, in which the minimization of an objective function is formulated (Lelievre and Oldenburg, 2009). In the objective function Φ, the data misfit Φ_d is combined with the regularization term Φ_m as in eq.4.57:

$$\Phi = \Phi_d + \lambda \Phi_m \rightarrow min \qquad (4.57)$$

where λ is the regularization parameter that weights the regularization term for the resulting model and is a trade-off parameter between the two terms (Gunther et al., 2006; Lelievre and Oldenburg, 2009). If there is no regularization, the result can be an unstable and the solution non-realistic, since there is only the minimization of the misfit functional. Because of that, the regularization techniques are used. Specifically, the data misfit term is defined as the difference between the field data, that contain noise, d_i, and the predicted (synthetic) data, $f_i(\boldsymbol{m})$ that are calculated through the forward modeling f_i supposing a candidate model, m. In particular, it is defined as the sum-of-squares (Eq.4.58):

$$\Phi_d(\boldsymbol{m}) = \sum_{i=1}^{N} \left| \frac{d_i - f_i(\boldsymbol{m})}{\varepsilon_i} \right|^2 = \left\| D\big(d - f(\boldsymbol{m})\big) \right\|_2^2 \qquad (4.58)$$

in which N is the number of data and ε_i is the error associated with the field data and D=diag($1/\varepsilon_i$) (Gunther et al., 2006; Lelievre and Oldenburg, 2009) if the noise in the data is assumed to be uncorrelated. If the uncertainty in the observed data is large, its contribution to the misfit measure is small (Lelievre and Oldenburg, 2009). The Φ_m term is instead defined as the squared norm of the product between a constraint matrix $\boldsymbol{C}$ and the difference between the model $\boldsymbol{m}$ and a reference model, $\boldsymbol{m_0}$ as in eq.4.59 (Gunther et al., 2006):

$$\Phi_m(\boldsymbol{m}) = \|C(\boldsymbol{m} - \boldsymbol{m_0})\|_2^2 \qquad (4.59)$$

The $\boldsymbol{C}$ matrix is a sparse matrix, as defined in Gunther et al. (2006): $C \in R^{BxM}$, with 2M entries and B, the number of boundaries. Specifically, for regular grids, it is calculated as a discrete approximation of a partial differential operator of first or second order (Gunther et al., 2006).

As regards the regularization parameter λ, it has to be chosen appropriately. In fact, if it is too small, the minimization of the objective function is equivalent to the minimization of the misfit term alone, and consequently, since there is not regularization, the solution can be unstable; while if it is too large, the minimization of the objective function is equivalent to the minimization of the regularization term, and it can cause the solution to be closer to the a priori model, so that the field data are ignored in the inversion (Zhdanov, 2015). A method that can be used in order to choose the optimal value for the regularization parameter λ is the L-curve, a graphical method in which the misfit term Φ_d is plotted against the regularization term, Φ_m, for all possible values of λ. The result, in a log-log scale has an L shape appearance. The point that separates the vertical from the horizontal branches of the curve corresponds to the value chosen as the optimal value for the regularization parameter (Fig.12) (Zhdanov, 2015).

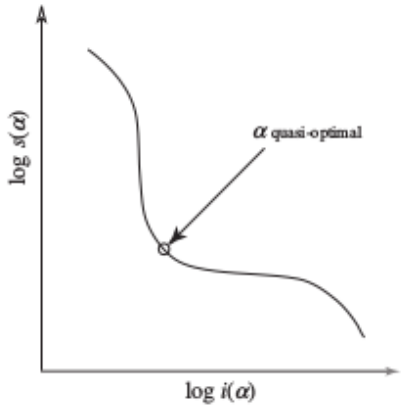

Figure 12: The L-curve, in which i(α) represents the misfit functional, while the s(α) the regularization functional (Zhdanov, 2015)

The solution of an inverse problem, formulated in Eq.4.57, is reduced to an optimization problem, since we try to find a set of model parameters for which the objective function has a minimum value (Sen, 2013). These optimization methods clearly depend on the shape of the objective function, that has peaks and troughs, for the non-linear problems. Specifically, the objective function can have local minima and a global minimum, which is the minimum of all the local minima. However, with the optimization methods we ideally try to find the global minimum of the objective function (Sen, 2013). Several optimization methods exist and a popular class of them goes by the name of "local" optimization methods, whose convergence depends on the choice of the starting model. In fact, these methods compute a search direction and a step length, from which a new model is calculated, after the interpreter has chosen a starting model. Evaluating the model, the process is repeated iteratively until there is no change in the new model or if the objective function reaches a minimum value. If the error function does not decrease during the process, the algorithm stops (Sen, 2013). Specifically, some of the methods that are commonly used consist of combinations/variations of: the Newton's method; the steepest descent and the conjugate gradient methods (Zhdanov, 2015).

In the Newton's method the error function (the objective function) is locally approximated by a quadratic function which is minimized exactly (Sen, 2013). Specifically, the error $E(m_{k+1})$ at point m_{k+1} can be approximated by the truncated Taylor's series (Eq.4.60) (Sen, 2013):

$$E(\boldsymbol{m_{k+1}}) \cong E(\boldsymbol{m_k}) + (\boldsymbol{m_{k+1}} - \boldsymbol{m_k})^T \nabla E(\boldsymbol{m_k}) + \frac{\mathbf{1}}{\mathbf{2}} (\boldsymbol{m_{k+1}} - \boldsymbol{m_k})^{\boldsymbol{T}} \boldsymbol{\nabla\nabla E}(\boldsymbol{m_k})(\boldsymbol{m_{k+1}} - \boldsymbol{m_k}) \qquad (4.60)$$

considering the gradient and the Hessian of the function. In particular, the point $\boldsymbol{m_{k+1}}$that minimizes this truncated series have to satisfy the condition (Sen, 2013):

$$\nabla E(\boldsymbol{m_k}) + \nabla\nabla E(m_k)(\boldsymbol{m_{k+1}} - \boldsymbol{m_k}) = 0 \qquad (4.61)$$

From that, we obtain (Sen, 2013):

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k - [\boldsymbol{\nabla\nabla E}(\boldsymbol{m_k})]^{-1} \nabla E(\boldsymbol{m_k}) = \boldsymbol{m}_k - \alpha_k \boldsymbol{H}^{-1}(\boldsymbol{m_k}) \nabla E(\boldsymbol{m_k}) \qquad (4.62)$$

in which $\boldsymbol{H}$ is the Hessian matrix and α_k is a step lenght chosen using for example a line search method.

Following the Newton's method and the approach in Gunther (2004), starting from the objective function (4.57):

$$\Phi = \Phi_d + \lambda \Phi_m \rightarrow min$$

we can write the gradient and the Hessian for the misfit function Φ_d (4.58):

$$\Phi_d(\boldsymbol{m}) = \sum_{i=1}^{N} \left| \frac{d_i - f_i(\boldsymbol{m})}{\varepsilon_i} \right|^2 = \left\| D\big(d - f(\boldsymbol{m})\big) \right\|_2^2$$

as, for the gradient:

$$\nabla_m \Phi_d(\boldsymbol{m}^k) = \boldsymbol{S}^T \boldsymbol{D}^T D(\boldsymbol{f}(\boldsymbol{m}^k) - \boldsymbol{d}) \qquad (4.63)$$

where **S** is the Jacobian or sensitivity matrix, given by:

$$S_{ij}(\boldsymbol{m}^k) = \frac{\partial f_i(m^k)}{\partial m_j} \qquad (4.64)$$

and for the Hessian:

$$\boldsymbol{H}_d^k = 2\boldsymbol{S}^T \boldsymbol{D}^T \boldsymbol{D} \boldsymbol{S} + 2(\nabla_m^T \boldsymbol{S}^T) \boldsymbol{D}^T \boldsymbol{D}(\boldsymbol{f}(\boldsymbol{m}^k) - \boldsymbol{d}) \qquad (4.65)$$

As in Gunther (2004), if we neglect the second term at right hand side of (4.65), that is generally difficult to compute, the Hessian will be:

$$\boldsymbol{H}_d^k = 2\boldsymbol{S}^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{S} \qquad (4.66)$$

and the method will be called the Gauss-Newton method.

Following the same method for the regularization term (4.59):

$$\Phi_m(\boldsymbol{m}) = \|C(\boldsymbol{m}-\boldsymbol{m_0})\|_2^2$$

We can write the gradient as:

$$\nabla_m \Phi_m = 2\boldsymbol{C}^T\boldsymbol{C}(\boldsymbol{m}-\boldsymbol{m_0}) \qquad (4.67)$$

and the Hessian:

$$\nabla_m^2 \Phi_m = 2\boldsymbol{C}^T\boldsymbol{C} \qquad (4.68)$$

In conclusion, using the Gauss-Newton method (Gunther, 2004), we can write the model update for the objective function, as:

$$(S^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{S} + \lambda\boldsymbol{C}^T\boldsymbol{C})\Delta\boldsymbol{m}^k = S^T\boldsymbol{D}^T\boldsymbol{D}\Big(\boldsymbol{d}-\boldsymbol{f}(\boldsymbol{m}^k)\Big) - \lambda\boldsymbol{C}^T\boldsymbol{C}\big(\boldsymbol{m^k}-\boldsymbol{m_0}\big) \qquad (4.69)$$

In Gunther (2004) the expression (4.69) was not solved using the direct inversion since the high number of model parameters and data could be a problem for computer time and memory usage. For this reason, iterative methods were used in order to obtain an approximate solution. Specifically, the conjugate gradient method was used (Gunther, 2004). I will examine this method in depth in paragraph 4.5.

4.3. Joint inversion

Groundwater modeling involves accurate information about aquifer characterization. These properties generally derive from point-like samplings, like geologic and hydrogeologic investigations in a restricted number of boreholes. However, these investigations are able to sample a small volume, mostly restricted around the boreholes. In addition, the laboratory experiments that determine some physical properties cannot represent the on-site conditions, because of the different scaling effects, confining pressure, degree of fracturation. An element that helps the reconstruction of hydrologic models at large scale is the use of crosshole or/and surface geophysical methods, that:

- investigate a much greater volume than drillings;
- are able to image the complexity and heterogeneity of the subsurface;
- provide a continuous although discretized model of the subsurface, not a punctual sampling (Doetsch et al., 2010).

Even if the results obtained from one geophysical method can help the improvement of the hydrologic model, ambiguities are yet present, because of the nonuniqueness in the inversion process (Doetsch et al., 2010). In fact, measurements are made on/or near the surface with distances that are not of the same length of geological variations and are contaminated by noise. This implies that geophysical inversion is an ill-posed problem, so that aclass of several, if not infinite number of models can fit the data to the same level (Moorkamp, 2017). In order to reduce the nonuniqueness, some external constraints can be introduced, as a priori information: via the regularizing term (e.g., smooth/sparse regularization) and/or by enforcing a particular model parametrization (e.g., by means of an under-parametrization of the domain and/or by incorporating the geometry of a desired feature in the discretization).

An approach that can help to achieve a more accurate and reliable reconstruction of the subsoil, is the joint inversion of two or several geophysical data sets. The joint inversion is defined (Moorkamp, 2017; Moorkamp et al., 2016) as the approach in which different types of data are inverted "*within a single algorithm, with a single objective function, and where all model parameters describing the property fields are adjusted concurrently throughout the inversion process*" (Moorkamp, 2017).
The joint inversion is different from the cooperative inversion (for example Günther and Rücker, 2006), that is defined as "*approaches where single dataset inversions are performed, sequentially or in parallel, and information is shared between the different inversions*" (Moorkamp et al., 2016).
However, the different geophysical methods are sensitive to different physical properties and their integration can bring much more confidence in the estimated models. Because of these different sensitivities of the various geophysical data, the joint inversion can be conducted using some petrophysical relationships, as for example resistivity and dielectric permittivity, that both depend on porosity (Linde et al., 2006). However, the petrophysical relationships also depend on additional physical parameters varying in space and that cannot be known in advance. Another type of joint inversion is based on imposing structural similarity between models (Haber and Oldenburg, 1997), without the use of an explicit petrophysical relationship. In particular, Gallardo and Meju (2003, 2004) developed a structural approach in which the vector cross-product of the gradients of two different models is forced to be zero at each iteration of the inversion, so it implies similar

directions of the two model gradient vectors (Doetsch et al., 2010). This method has been widely used and slightly modified by various authors (Linde et al., 2006; Tryggvason et al., 2006; Gallardo and Meju, 2007; Fregoso and Gallardo, 2009; Linde and Doetsch, 2010; Doetsch et al., 2010; Moorkamp et al., 2011; Hu et al., 2009; Bennington et al., 2015; Demirci et al., 2017, 2018; Shi et al., 2017; Pak et al., 2017; Wang et al., 2017) and it has become one of the most popular coupling approach in the joint inversion (Moorkamp, 2017). It is also considered one of the most robust methods in the joint inversion of nearsurfacegeophysical field data (Linde and Doetsch, 2016). This approach will be explained in detail in the next paragraph.

4.3.1. Cross-gradientsfunction

In their work, Gallardo and Meju (2003,2004) proposed a procedure for the inversion of two different geophysical data sets, specifically DC resistivity ad seismic P-wave travel times, using a new function, called cross-gradients constraints, with the purpose of determining 2D resistivity and velocity models in which the changes in the parameters occur in the same direction, regardless of their magnitude (Fregoso and Gallardo, 2009). Infact, the approach of the authors is that, even if the petrophysical relationships are not known, the geophysical properties depend on the same geology of the subsoil in the near surface field (Linde et al., 2006) and consequently, they show a structural similarity. The authors noted that the variations of the properties in the subsurface might occur in any direction and that they can be characterized using two parameters, magnitude and direction. The geophysical properties are considered structurally similar if the changesin magnitude and direction are common to both of them (Gallardo and Meju 2003, 2004). These attributes can be represented mathematically by the product of geophysical properties' gradients:

$$\vec{t}(x,y,z) = \nabla m_1(x,y,z) \times \nabla m_2(x,y,z) \qquad (4.70)$$

The authors considered that the two models are structurally similar if the cross-gradients function is zero everywhere, since it implies that the two geophysical models are exactly parallel and consequently, t is minimized (Gallardo and Meju 2003,2004, Bennington et al., 2015). This function allows changes in one of the models, without forcing the other to change necessarily (Linde and Doetsch, 2010; Bennington et al., 2015). Furthermore, it considers the direction of the change rather than the magnitude (Linde and Doetsch, 2010). The cross-gradients function can then be written as, for the 2D case and considering the considered geophysical methods:

$$t(x,z) = \left(\frac{\partial m_r(x,z)}{\partial z}\right)\left(\frac{\partial m_s(x,z)}{\partial x}\right) - \left(\frac{\partial m_r(x,z)}{\partial x}\right)\left(\frac{\partial m_s(x,z)}{\partial z}\right) \qquad (4.71)$$

where the subscripts *r* and *s* represent resistivity and slowness, respectively. The derivatives have been written (Gallardo and Meju, 2003,2004) using the forward differences:

$$t \cong \frac{4}{\Delta x \Delta z}\left(m_{rc}(m_{sb}-m_{sr})+m_{rr}(m_{sc}-m_{sb})+m_{rb}(m_{sr}-m_{sc})\right) \qquad (4.72)$$

where the first subscript represents the geophysical parameter considered and specifically in this case *r* for resistivity and *s* for slowness, and the second subscript represents the considered cell, *c* for center, *b* for bottom and *r* for right, as in Fig.13.

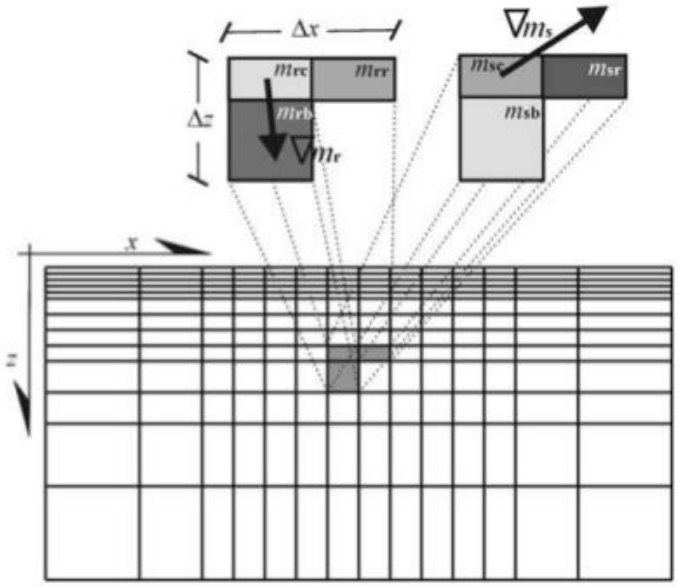

Figure 13: Mesh used to represent the models of the subsurface in Gallardo and Meju (2004). In detail, it is represented the scheme that has been used by the authors for the discretization of the cross-gradients function. (Gallardo and Meju, 2004)

Several authors have used this approach, like Linde et al. (2006), which were able to improve hydrogeologic studies with the joint inversion of crosshole electrical resistivity and ground penetrating radar.However, the authors used various inversion procedures: Gallardo and Meju (2003, 2004, 2007) and Gallardo (2007) included the cross-gradients function as an equality constraint and then solved the minimization problem with the method of Lagrange multipliers.Also Hamdan and Vafidis (2012) use the method of the Lagrange multipliers, slightly modifying the original equations of Gallardo and Meju (2004); Linde et al. (2006), Tryggvason and Linde (2006), Linde and Doetsch (2010), Doetsch et al. (2010) added the cross gradients in the objective function and then minimize them in a least square manner, solving the system at each iteration with the conjugate gradient method LSQR. Moorkamp et al. (2011) and Shi et al. (2017) too, added the cross-gradients function to the objective function, optimizing it employing a limited memory quasi-Newton approach and using the line search direction to update the model. Finally, Hu et al. (2009) use a Gauss-Newton method solving the joint-inversion problem in an iterative and sequential

manner applying the cross-gradients constrain to one fixed model and one updated model, but their approach is classified as a cooperative joint inversion.

Another interesting aspect that emerges from the works of the different authors is that the cross-gradients function can be defined in different ways (Linde and Doetsch, 2010). It can be defined for the total model, as in Gallardo and Meju (2003) or for the model update with respect to a reference model (Linde et al., 2006). The discretized cross-gradients function can be obtained usingforward (Gallardo and Meju, 2004) or central (Linde and Doetsch, 2010) differences and the formulation is different if we have a uniform discretization (Linde and Doetsch, 2010) or a nonuniform cell spacing using rectangular cells (Gallardo and Meju, 2003, 2004).

4.4. Inversion Metodology

Two slightly different approaches have been developed for thejoint inversion, both based on gradient direction of resemblance. Specifically, as in Meju and Gallardo (2016): "*one approach seeks exact structural resemblance between the sought images*. *In the other approach, structural resemblance is encouraged, rather than imposed, by minimizing a weighted norm of the structural resemblance constraint*".
Specifically, Meju and Gallardo (2016) used the cross-gradient function (4.73):

$$\Delta\vec{\varphi}=\vec{t}=\nabla m^{(1)}\times\nabla m^{(2)} \qquad (4.73)$$

in multiphysics imaging and imposed two geophysical model to be structurally similar, in the inversion process, when:

$$\Delta\vec{\varphi}=\vec{t}=\nabla m^{(1)}(x_i,y_i,z_i)\times\nabla m^{(2)}(x_i,y_i,z_i)=\vec{0} \qquad (4.74)$$

where $m^{(1)}$ and $m^{(2)}$ are the electrical resistivity and the seismic velocity respectively. In their work, in particular, the data are assumed to have a random noise with an uncorrelated Gaussian distribution with zero mean and variance σ^2. In this inverse problem, the regularization is performed considering the smoothness assumptions and using all available a priori information through the a priori model m_0. The errors associated with the a priori model have a Gaussian distribution with the covariance matrix $C_{m_0 m_0}$. The final objective function is:

$$s\left(\hat{\mathbf{m}}^{(1)},\hat{\mathbf{m}}^{(2)}\right)=\left\|\begin{matrix}\mathbf{d}_0^{(1)}-f^{(1)}\left(\hat{\mathbf{m}}^{(1)}\right)\\ \mathbf{d}_0^{(2)}-f^{(2)}\left(\hat{\mathbf{m}}^{(2)}\right)\end{matrix}\right\|^2_{C_{dd}^{-1}}+\left\|\begin{matrix}\alpha^{(1)}\mathbf{D}\hat{\mathbf{m}}^{(1)}\\ \alpha^{(2)}\mathbf{D}\hat{\mathbf{m}}^{(2)}\end{matrix}\right\|^2+\left\|\begin{matrix}\hat{\mathbf{m}}^{(1)}-\hat{\mathbf{m}}_0^{(1)}\\ \hat{\mathbf{m}}^{(2)}-\hat{\mathbf{m}}_0^{(2)}\end{matrix}\right\|^2_{C_{m_0m_0}^{-1}} \quad (4.75)$$

subject to:

$$\{\vec{t}_i\}=\{\vec{0}\}$$

In the expression (4.75) **d** is the data vector, $f(\mathbf{m})$ is the forward modeling operator, **D** is the discrete version of the second-order Tikhonov regularization matrix, C_{dd} is the covariance matrix associated to each geophysical data set and α is a constant damping operator for each model. Because of the nonlinear behaviour of the forward modelling and the cross-gradients function, the solution for the objective function is found with the linearization around the starting model $\mathbf{m_o}$ in the first passage and with a two stage minimization process in an iterative way. Specifically, the solution at iteration k+1 is:

$$\hat{\boldsymbol{m}}_{k+1}^{(j)}=\hat{\boldsymbol{m}}_k^{(j)}+\boldsymbol{N}_k^{(j)^{-1}}\left(\boldsymbol{n}_k^{(j)}-\boldsymbol{B}_k^{(j)^T}\boldsymbol{q}_k\right) \qquad j=1,2 \quad (4.76)$$

in which:

$$\boldsymbol{N}_k^{(j)}=\boldsymbol{A}_k^{(j)^T}C_{dd}^{(j)^{-1}}\boldsymbol{A}_k^{(j)}+\alpha^{(j)}\boldsymbol{D}^T\boldsymbol{D}+\left(\boldsymbol{C}_{m_0m_0}^{(j)}\right)^{-1} \quad (4.77)$$

$$n_k^{(j)}=\boldsymbol{A}_k^{(j)^T}C_{dd}^{(j)^{-1}}\left(\boldsymbol{d_0^{(j)}}-\boldsymbol{f^{(j)}}\left(\hat{\boldsymbol{m}}_{\boldsymbol{k}}^{(\boldsymbol{j})}\right)+\boldsymbol{A}_{\boldsymbol{k}}^{(\boldsymbol{j})}\hat{\boldsymbol{m}}_{\boldsymbol{0}}^{(\boldsymbol{j})}\right)+\alpha^{(j)}\boldsymbol{D}^T\boldsymbol{D}\hat{\boldsymbol{m}}_{\boldsymbol{k}}^{(\boldsymbol{j})}+\left(\boldsymbol{C}_{m_0m_0}^{(j)}\right)^{-1}\left(\hat{m}_0^{(j)}-m_k^{(j)}\right) \quad (4.78)$$

where:

$$A=\frac{\partial f(m)}{\partial m_k} \quad (4.79)$$

are the partial derivatives of the model $\boldsymbol{m}=\boldsymbol{m_k}$ (i.e. Jacobian is called *S* in equations 4.63-4.66 and 4.69).

Then,

$$q_k=\left\{B_k^{(1)}N_k^{(1)^{-1}}B_k^{(1)^T}+B_k^{(2)}N_k^{(2)^{-1}}B_k^{(2)^T}\right\}^{-1}\left\{B_k^{(1)}N_k^{(1)^{-1}}n_k^{(1)}+B_k^{(2)}N_k^{(2)^{-1}}n_k^{(2)}+\Delta\varphi\left(\hat{m}_k^{(1)},\hat{m}_k^{(2)}\right)\right\} \quad (4.80)$$

and $\mathbf{B_k}$ is the jacobian matrix that is associated with the cross-gradients function. The term $\Delta\varphi$ is the cross-gradient function:

$$t(m^{(1)},m^{(2)})=\nabla m^{(1)}\times\nabla m^{(2)} \qquad (4.81)$$

that is linearized using a first order Taylor series expansion around $m_0^{(1)}$ and $m_0^{(2)}$. In particular, the updating of the two different models is made considering the cross-gradient function of the models in the previous iteration.

The second approach *"adopts inexact measures of structural resemblance but have the added flexibility that such measures may be treated as statistical realizations"* (Meju and Gallardo, 2016). Specifically, in this approach, that has been adopted by different authors, like Linde et al. (2006), Doetsch et al. (2010), Hu et al. (2009), the cross-gradients function is added to the objective function and the objective function can be written as:

$$s(\hat{\mathbf{m}}^{(1)},\hat{\mathbf{m}}^{(2)})=\left\|\begin{matrix}\mathbf{d}_0^{(1)}-f^{(1)}(\hat{\mathbf{m}}^{(1)})\\ \mathbf{d}_0^{(2)}-f^{(2)}(\hat{\mathbf{m}}^{(2)})\end{matrix}\right\|^2_{C_{dd}^{-1}}+\left\|\Delta\varphi(\hat{\mathbf{m}}^{(1)},\hat{\mathbf{m}}^{(2)})-\Delta\varphi_0\right\|^2_{C_{\varphi\varphi}^{-1}}+\left\|\begin{matrix}\hat{\mathbf{m}}^{(1)}-\hat{\mathbf{m}}_0^{(1)}\\ \hat{\mathbf{m}}^{(2)}-\hat{\mathbf{m}}_0^{(2)}\end{matrix}\right\|^2_{C_{m_0m_0}^{-1}} \qquad (4.82)$$

in which $C_{\varphi\varphi}^{-1}=\alpha I$ and α is a constant trade-off parameter (Meju and Gallardo, 2016). Following the Gauss-Newton approach, relation 4.82 can be written as:

$$\hat{\boldsymbol{m}}_{k+1}^{(j)}=\hat{\boldsymbol{m}}_k^{(j)}+\boldsymbol{H}_k^{-1}\boldsymbol{r}_k^{(j)}\quad j=1,2 \qquad (4.83)$$

in which the term H_k is the Hessian matrix, that can be written as:

$$\boldsymbol{H}_k=\begin{bmatrix}\boldsymbol{A}_k^{(1)^T}\boldsymbol{C}_{dd}^{(1)^{-1}}\boldsymbol{A}_k^{(1)}+\boldsymbol{B}_k^{(1)^T}C_{\varphi\varphi}^{-1}\boldsymbol{B}_k^{(1)}+\left(C_{m_0m_0}^{(1)}\right)^{-1} & \boldsymbol{B}_k^{(1)^T}\boldsymbol{C}_{\varphi\varphi}^{-1}\boldsymbol{B}_k^{(2)}\\ \boldsymbol{B}_k^{(2)^T}\boldsymbol{C}_{\varphi\varphi}^{-1}\boldsymbol{B}_k^{(1)} & \boldsymbol{A}_k^{(2)^T}\boldsymbol{C}_{dd}^{(2)^{-1}}\boldsymbol{A}_k^{(2)}+\boldsymbol{B}_k^{(2)^T}\boldsymbol{C}_{\varphi\varphi}^{-1}\boldsymbol{B}_k^{(2)}+\left(\boldsymbol{C}_{m_0m_0}^{(2)}\right)^{-1}\end{bmatrix}$$

and the expression for r_k is:

$$\boldsymbol{r}_k^{(j)}=\boldsymbol{A}_k^{(j)^T}\boldsymbol{C}_{dd}^{(j)^{-1}}\left(\boldsymbol{d}_0^{(j)}-\boldsymbol{f}^{(j)}\left(\hat{\boldsymbol{m}}_k^{(j)}\right)\right)+\boldsymbol{B}_k^{(j)^T}\boldsymbol{C}_{\varphi\varphi}^{-1}\left(\Delta\varphi_0-\Delta\varphi\left(\hat{\boldsymbol{m}}_k^{(1)},\hat{\boldsymbol{m}}_k^{(2)}\right)\right)+\left(\boldsymbol{C}_{m_0m_0}^{(j)}\right)^{-1}\left(\hat{\boldsymbol{m}}_0^{(j)}-\boldsymbol{m}_k^{(j)}\right) \qquad (4.84)$$

As underlined in Meju and Gallardo (2016), this approach is conceptually identical to the other approach when $C_{\varphi\varphi}^{-1}=\alpha I$ and $\alpha\to\infty$ and *"the first term on the right-hand side containing the observables f(m) are augmented with the regularization terms (Dm)"* as in eq.4.75. Therefore, this minimization process can be solved with different techniques, like for example the nonlinear conjugate gradients method for large-scale 3D multiphysics imaging problems, as suggested by the authors (Meju and Gallardo, 2016), or the conjugate gradient method as suggested in Zhadanov (2015).

4.5. Algorithm for the joint inversion

The objective function for the proposed joint inversion can be written using the complete formulation, as (4.85):

$$\Phi(m_1,m_2)=\left\|\begin{matrix}D_1(f(m_1)-d_{obs1})\\D_2(f(m_2)-d_{obs2})\end{matrix}\right\|_2^2+\begin{matrix}\lambda_1\\\lambda_2\end{matrix}\left\|\begin{matrix}C(m_1-m_{01})\\C(m_2-m_{02})\end{matrix}\right\|_2^2+\lambda_{CG}\|t(m_1,m_2)-t_0\|_2^2\rightarrow min \qquad (4.85)$$

in which: D_1 and D_2 are the matrices that take into account the errors of data fields for both the considered geophysical methods; $f(m_1)$ and $f(m_2)$ are the two forward modelings and d_{obs1} and d_{obs1} are the field data acquired for the two methods; C is the constraint matrix; m_1 and m_2 are the two models and m_{01} and m_{02} the a priori models; λ_1 and λ_2 are the two regularization parameters for the two methods; t and t_0 are the cross-gradients function and its a priori value respectively and finally λ_{CG} is the regularization parameter that weights the cross-gradients function. Obviously, if no a priori information is available, it is eliminated from the complete formulation (4.85).

Using the Gauss-Newton method (as in Günther, 2004) the misfit and regularization terms can be written as in Chapter 4.2 (equations 4.58 and 4.59), while, for the term of the cross-gradients function, $\Phi_{CG}=\|t(m_1,m_2)-t_0\|_2^2$, the gradient can be written as:

$$\frac{\partial\Phi_{CG}}{\partial m_1}\cong B_1^T(t(m_1,m_2)-t_0) \qquad (4.86)$$

$$\frac{\partial\Phi_{CG}}{\partial m_2}\cong B_2^T(t(m_1,m_2)-t_0) \qquad (4.87)$$

where B_1 and B_2 are the Jacobian matrices associated with the cross-gradient function. Then, the Hessian can be written as:

$$H\cong\begin{bmatrix}B_1^TB_1 & B_1^TB_2\\B_2^TB_1 & B_2^TB_2\end{bmatrix} \qquad (4.88)$$

Finally, the model update for the objective function can be written as:

$$\left\{\begin{bmatrix}S_1^TD_1^TD_1S_1 & 0\\0 & S_2^TD_2^TD_2S_2\end{bmatrix}+\begin{bmatrix}\lambda_1C^TC & 0\\0 & \lambda_2C^TC\end{bmatrix}+\lambda_{CG}\begin{bmatrix}B_1^TB_1 & B_1^TB_2\\B_2^TB_1 & B_2^TB_2\end{bmatrix}\right\}\cdot\begin{bmatrix}\Delta m_1\\\Delta m_2\end{bmatrix}=$$

$$\begin{bmatrix}S_1^TD_1^TD_1(d_{obs1}-f_1(m_1)\\S_2^TD_2^TD_2(d_{obs2}-f_2(m_2)\end{bmatrix}-\begin{bmatrix}\lambda_1C^TC(m_1-m_{01})\\\lambda_2C^TC(m_2-m_{02})\end{bmatrix}-\lambda_{CG}\begin{bmatrix}B_1^T(t(m_1,m_2)-t_0)\\B_2^T(t(m_1,m_2)-t_0)\end{bmatrix} \qquad (4.89)$$

Since even in this case the high number of model parameters and data could be a problem for computer time and memory usage, I decide to use the conjugate gradient method (Hestenes and

Stiefel, 1952) in order to solve the system of equations 4.89 approximately. The conjugate gradient method has been widely used in solving large scale system of equations and it has been successfully applied to geophysical problems (Scales, 1987). It is an iterative method for solving linear or linearized system of equations (Pilkington, 1997) and it has been proved that this method is fast, simple and accurate and it can be adapted not only to sparse systems, but also to normal equations (VanDecar and Snieder, 1994).

Starting from the system of equations 4.89, we can write the algorithm for the conjugate gradient method (following the approach of Gunther (2004), Gunther et al. (2006) and VanDecar and Snieder (1994)). In particular, following the approach of VanDecar and Snieder (1994), I wrote the system of equations (4.89) as the least squares system:

$$\begin{pmatrix} DS \\ \sqrt{\lambda}C \\ \sqrt{\lambda_{CG}}B \end{pmatrix} \Delta m = \begin{pmatrix} D(d_{obs}-f(m) \\ -\sqrt{\lambda}C(m-m_0) \\ -\sqrt{\lambda_{CG}}(t-t_0) \end{pmatrix} \qquad (4.90)$$

in which the terms take into account the two models like in the system (4.89). Then, I wrote the different terms of the conjugate gradients method as in VanDecar and Snieder (1994) and after some algebra, I obtained the following conjugate gradients algorithm in which, only for convenience in writing, I called:

$$\tilde{t}=t-t_0 \qquad (4.91)$$

$$\tilde{m}=m-m_0 \qquad (4.92)$$

The conjugate gradients algorithm will be (following the notations in Gunther, 2004):

$k=0$

$z_0=D(d_{obs}-f(m)-S\Delta m_0)$

$y_0=(\tilde{t}+B\Delta m_0)$

$p_0=r_0=D^TS^Tz_0-\lambda C^TC(\tilde{m}+\Delta m_0)-\lambda_{CG}B^Ty_0$

while $r_k \neq 0$:

$q_{k+1}=D\cdot S\cdot p_k$

$w_{k+1}=B\cdot p_k$

$$\alpha_{k+1}=\frac{\|r_k\|^2}{q_{k+1}^T\cdot q_{k+1}+\lambda p_k^TC^TCp_k+\lambda_{CG}w_{k+1}^Tw_{k+1}}$$

$\Delta m_{k+1}=\Delta m_k+\alpha_{k+1}\cdot p_k$

$z_{k+1}=z_k-\alpha_{k+1}q_{k+1}$

$y_{k+1}=y_k+\alpha_{k+1}\cdot w_{k+1}$

$r_{k+1}=D^T \cdot S^T \cdot z_{k+1}-\lambda \cdot C^T C(\Delta m_{k+1}+\tilde{m})-\lambda_{CG} B^T y_{k+1}$

$\beta_{k+1}=\frac{\|r_{k+1}\|^2}{\|r_k\|^2}$

$p_{k+1}=r_{k+1}+\beta_{k+1} \cdot p_k$

$k=k+1$

After obtaining the Δm, in the iterative scheme, the model is updated using a line search parameter (Gunther, 2004):

$m^{k+1}=m^k+\tau^k \Delta m^k$ (4.93)

in which k denotes the iteration number.

In fig.14 the flowchart of the joint inversion is outlined. Specifically, the inversion process starts from the definition of the field or synthetic data and of the initial model, from which the predicted data are calculated for both geophysical methodologies. Then, the objective function, composed of the two data misfits, regularization terms and the cross-gradients function, is minimized. After the minimization of the objective function, if the stopping criteria are satisfied, the final earth model is found, otherwise the model is updated and the predicted data are again calculated, starting a new minimization process.

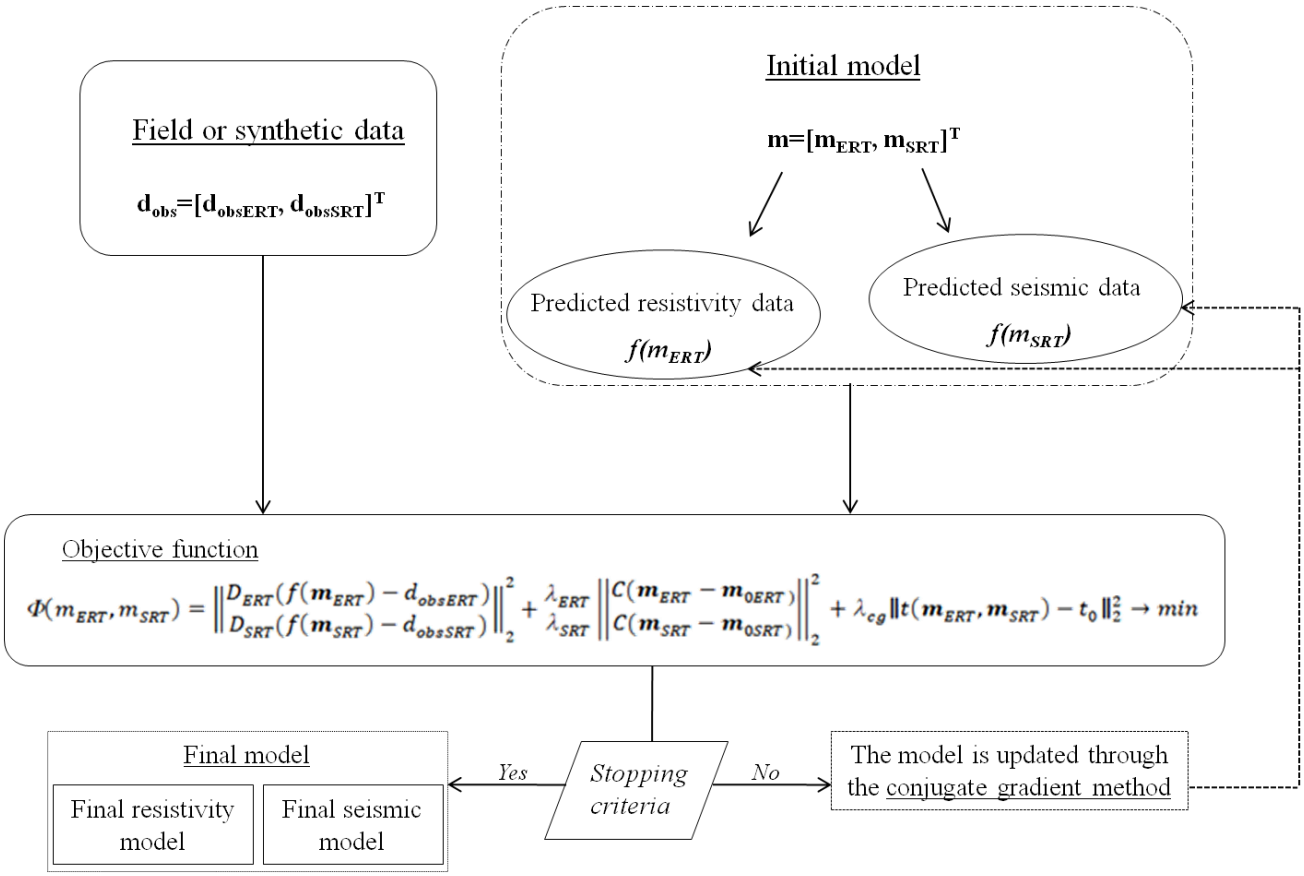

Figure 14: Flowchart of the implemented joint inversion for resistivity and seismic methods

5. pyGIMLi

The pyGIMLi package is an open-source library for geophysical modeling and inversion written in C++ and Python (Rücker et al., 2017). Specifically, as in Rücker et al. (2017), the runtime sensitive parts are implemented in C++, in order to have the maximum computational efficiency, while the connections with this "core library" are implemented in pure Python language, in order to obtain the maximum flexibility, since Python language is easier and more intuitive than C++. Because of this particular configuration, it is possible to add to the pyGIMLi package all the codes are needed for specific applications.

The package allows conducting a geophysical inversion for both field and synthetic data, since it provides a flexible generation of the parameter mesh, that can be regular or unstructured. In the case of unstructured mesh, an external mesh generator is used, which is called *Triangle.* Furthermore, different forward modeling codes for various geophysical methods are present. The used inversion routine employs the conjugate gradient method in order to solve the inverse problem (Fig.15).

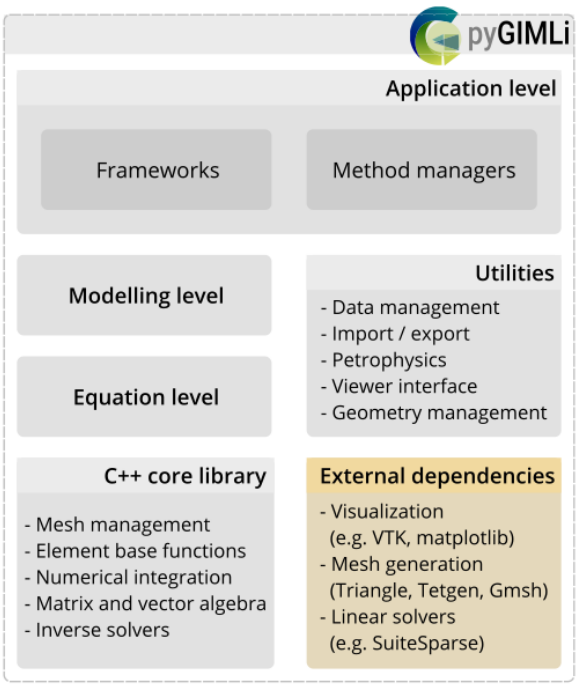

Figure 15: Software architecture of pyGIMLi (from Rücker et al., 2017)

As the connections and several codes are written in pure Python, it is possible to use also the Python libraries, as the plotting library *Matplotlib*, or alternatively one can export the final result using visualization programs such as *Paraview*.

6. Results

The results of joint inversions of both synthetic and real-data will be presented in this chapter. In a first step the FMM implemented in the pyGIMLi package will be used for seismic forward modeling. Then, I will examine the results obtained from the implementation of the MSFMM as the seismic forward modeling. For the synthetic inversions, the inverted results will be compared with the analytic solutions for simple cases. In addition, the inverted sections obtained from the separated and joint inversion using the MSFMM as forward modeling will be compared to point out the improved reconstruction obtained by the joint inversion of electrical and seismic data.

6.1.1. Joint inversion results

6.1.2. Synthetic model without noise

Separated Inversion

The first synthetic model consists of two layers with resistivity and seismic velocity that increase with depth (Fig.16).More specifically, the resistivity increases with depth from 500 to 1000Ω·m, while the seismic velocity is 1000m/s and 3000m/s, respectively for the surface layer and the deeper one (Fig.16). The model is intended to represent near surface discontinuities and a step 2m high is introduced in the central part of the section.

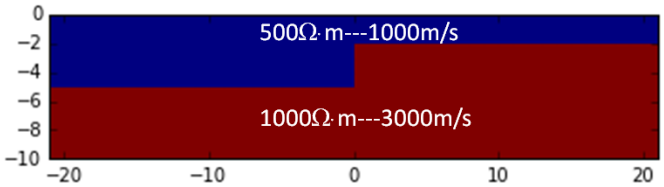

Figure 16: Synthetic model with values of resistivity and seismic velocity that increase with depth.

The electrical tomography measurements were simulated using the dipole-dipole configuration, to provide adequate resolution in the identification of the objects. For the measurements, 41 electrodes, 1m spaced, were used. The synthetic data were obtained using the forward modeling algorithm of the pyBert package (Rücker, 2011; Günther et al., 2006) previously described (Chapter 4) and no noise was added.

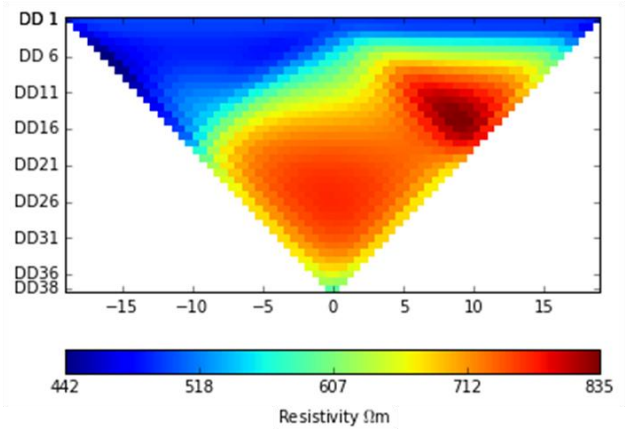

Figure 17: Synthetic electrical resistivity response obtained from the model without noise

For the seismic refraction tomography measurements, 41 geophones, 1m spaced apart, and 9 shots, with a shot spacing of 5m, were used. The choice of using such number of shots and geophones derives from the need to individuate appropriately the step at the center of the section. The synthetic traveltimes were obtained employing the Fast Marching Method (FMM) (described in Chapter 4) implemented in the pyGIMLi package (Rücker et al., 2017) at first for the triangular mesh and then adapted to the structural mesh. Also in this case, no noise was added to the synthetic data. The layer geometry and the seismic velocity values are well identified by the FMM (Fig.18).

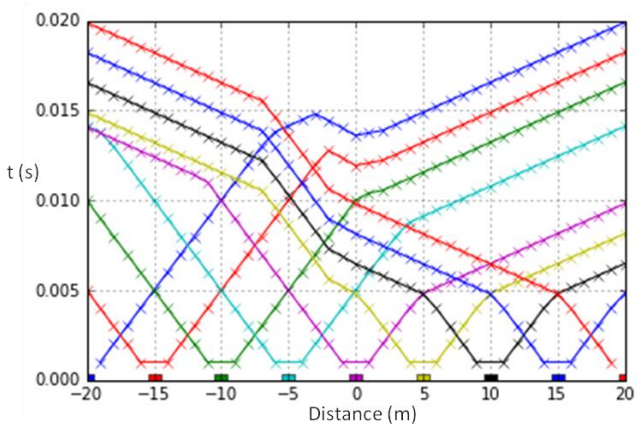

Figure 18: Computed travel times responses obtained from the model without noise

The synthetic data were then separately inverted using the pyGIMLi package. The lambda values used in the uncoupled inversion were chosen arbitrarily (since the data were error free and accordingly it has not been possible to obtain the L-curve) and set at a reference value of 30, for electrical resistivity and 50 for seismic travel times. The inversion was conducted without using a reference model. The convergence is reached after only 2 iterations for the ERT section and after 3 iterations for the SRT one,with the abort criterion of $\chi^2 = \frac{\Phi_d}{N} = 1$ and the inverted models show a good reconstruction of the model (Fig.19).

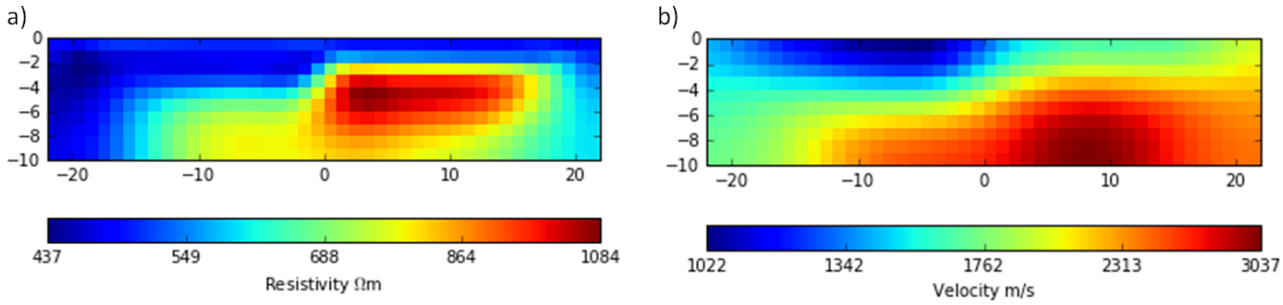

Figure 19: Separated inversion results: a) ERT; b) SRT

In fact, the result of the uncoupled inversion for ERT (Fig.19a) indicates the presence of two layers with values ranging from about 400 to 1000Ω·m; the step is clearly imaged in the middle of the section. A similar result has been obtained for the SRT separated inversion (Fig.19b) in which the seismic velocitiesrange fromabout 1000m/s to about 3000m/s. For the SRT inverted model, the step is individuated in a less clear way than the ERT section, whereas the layer at higher speed is wider than the one individuated by the ERT inverted model. Using the modelsobtained by the uncoupled inversion, it is possible to calculate the cross-gradients function (Gallardo and Meju, 2003, 2004) as described in Chapter 4, using the forward differencesand a uniform discretization (Fig.20):

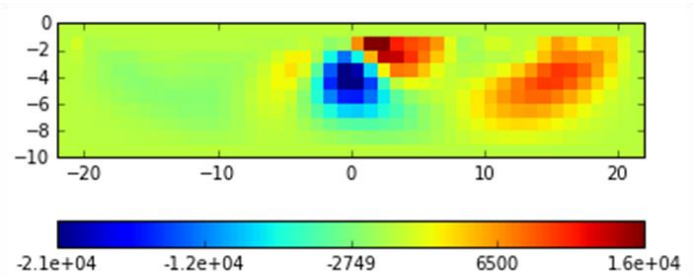

Figure 20: Cross-gradients function calculated for the separated inversion

In the section of the cross-gradients function high and low cross-gradients values represent areas of structural difference of the separated inverted models.

Another chart that is widely used in the literature on the joint inversion (Linde et al., 2006; Linde and Doetsch, 2010) is the scatter plot, that is a type of graph in which, for example, two variables for a set of data are represented using Cartesian coordinates. This type of graph can be used to individuate some correlation between variables. In our case, using the results of the separated inversion, the scatter plot (Fig.21) does not show clear different zones.

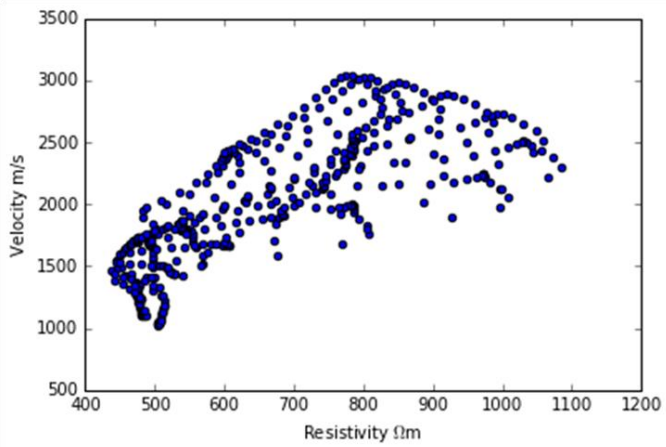

Figure 21: Scatter plot of seismic velocity versus electrical resistivity for the separated inversions of ERT and SRT.

Joint inversion

The joint inversion is performed using the same conditions used for the separated inversion, as well as the lambda values for ERT and SRT. As regards the value of the lambda, that will be indicated as lambdaGC, that weights the cross-gradients function, a few trials are necessary to obtain the appropriate weight. Specifically, following Linde et al. (2008), I decided to map the mean value of the vector obtained for the cross-gradients function as a function of the different lambdaGC values. Using lambda values from 100 to 1700 and plotting the mean values of the cross-gradients vector as a function of lambda, the lambda value in correspondence of which the smallest value of the cross-gradients function can be selected (Fig.22).

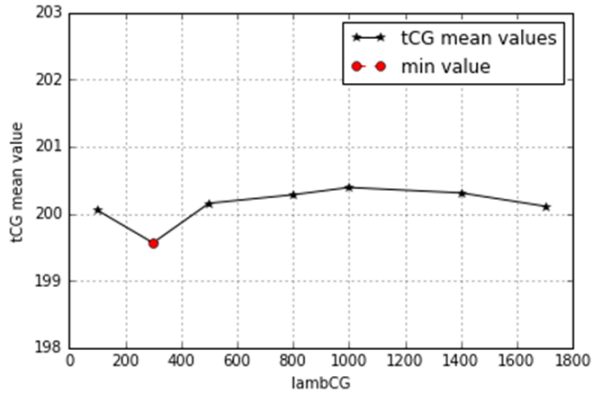

Figure 22: Mean value of the cross-gradients function as a function of lambda. The red point represents the optimum lambda for the joint inversion

In this case, the appropriate weigh for the cross-gradients function in the inversion seems to be the 300 value (red point in Fig.22).

The convergence of the joint inversion was obtained after 4 iterations using the stop criterion of $\chi^2=1$. The final inverted models (Fig.23) show results that are similar to the results obtained for the separated inversion, but that better characterize the investigated area for example for the better approximation of real values.

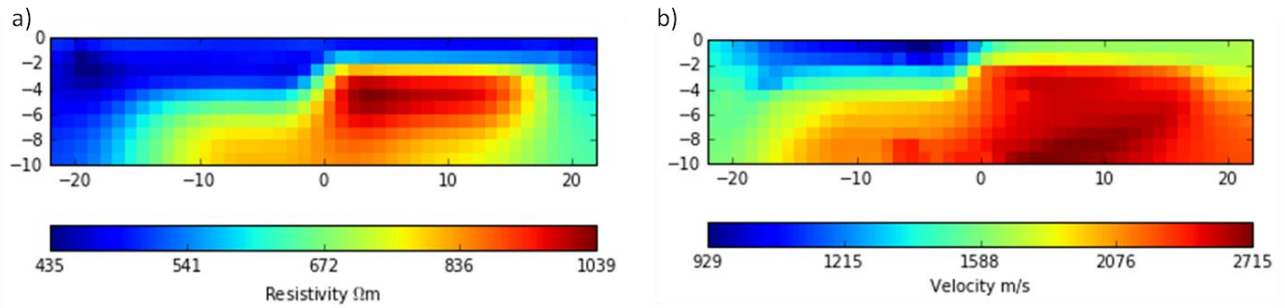

Figure 23: Joint inversion results: a) ERT; b) SRT

In the ERT section, the resistivity maximum value is lower than the one obtained for the separated inversion and closer to the theoretical model. The SRT section individuates the two layers and the step in the central part of the inverted model more clearly and the maximum velocity value is slightly lower than the one of the separated inversion. The associated cross-gradients function (Fig.24) shows values that are smaller than the ones obtained for the separated inversion, and this was the target of the joint inversion. This is in agreement to the fact that the models obtained with the joint inversion have enhanced structural attributes than the separated inversion.

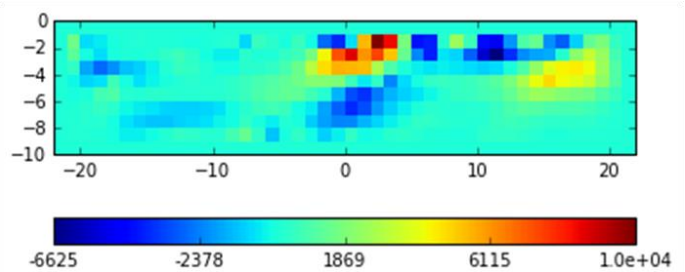

Figure 24: Cross-gradients function calculated for the joint inversion

The importance of the joint inversion approach can also be appreciated observing the scatter plot (Fig.25), as it shows less dispersion of data that are more grouped and easier to interpret than the separated inversion.

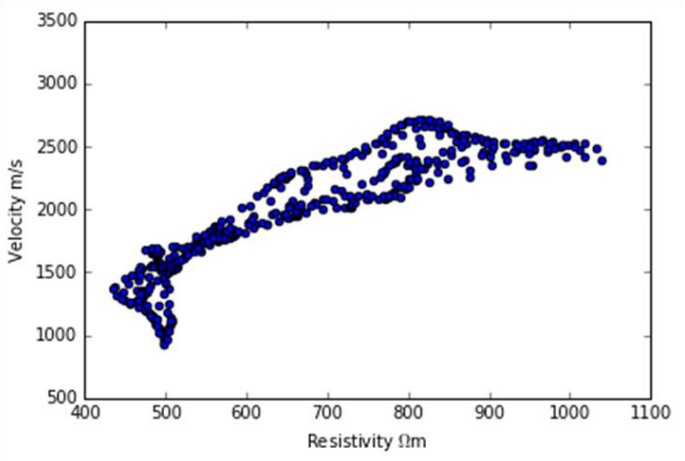

Figure 25: Scatter plot of seismic velocity versus electrical resistivity for the joint inversion of ERT and SRT.

6.1.3. *Synthetic model with noise*

Separated Inversion

In this paragraph I will examine the results of the joint inversion obtained adding some noise to the synthetic data using the same model as in the previous paragraph. Specifically, electrical resistivity data were contaminated with 1% of Gaussian noise, while seismic data with Gaussian noise of 1 ms standard deviation. The synthetic data were obtained using the same configurations described for the synthetic model without noise: 41 electrodes, 1m spaced using the dipole-dipole configuration for the electrical tomography measurements, while 41 geophones, 1m spaced, and 9 shots, 5m spaced for seismic refraction tomography measurements. The forward modeling algorithms that have been used were the one of the pyBert package (Rücker, 2011; Günther et al., 2006) for the electrical tomography measurements, while the FMM was used for the seismic refraction tomography measurements. The responses obtained in both cases are slightly different than the ones obtained without noise, but correctly locate the presence of the two layers and of the step in the central part of the section (Fig.26).

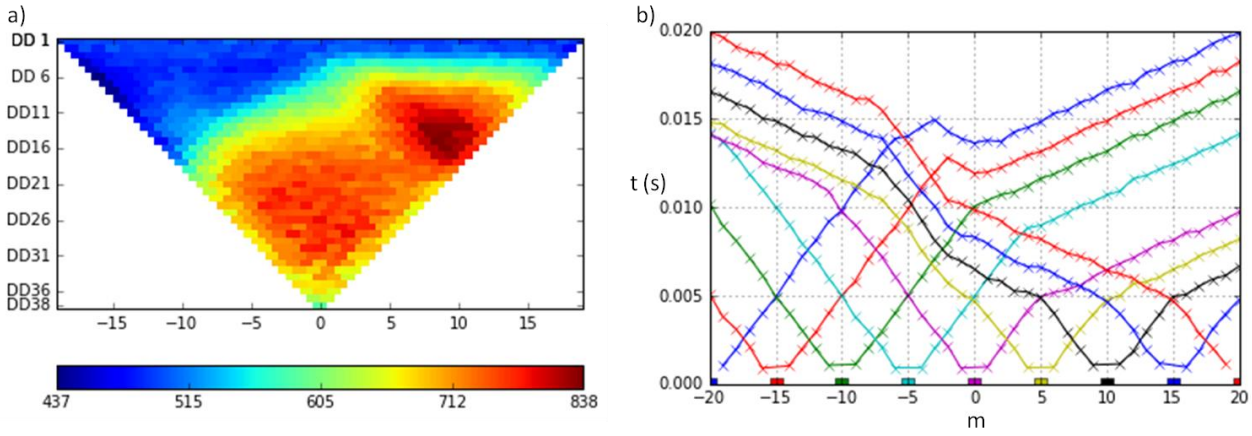

Figure 26: Computed responses for a)ERT and b) SRT

The separated inversion was conducted using the pyGIMLi package as previously described and the lambda value was chosen with the L-curve, described in Chapter 4.2, for both ERT and SRT. Specifically, a series of inversions were conducted using different values of lambda: for ERT, the lambda values start from 10 reaching the value 1258 (Fig.27), while for SRT, from 153 to 1258, since the values lower than 153 did not give satisfying results (Fig.28).

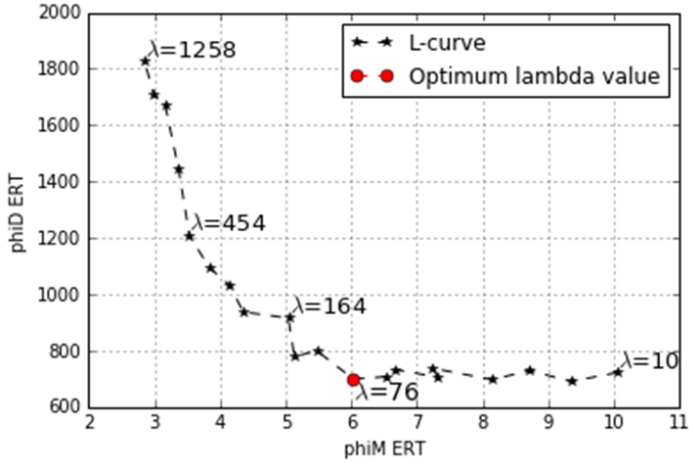

Figure 27: L-curve for ERT synthetic model; the red point is the optimal value for the regularization parameter in the separated inversion

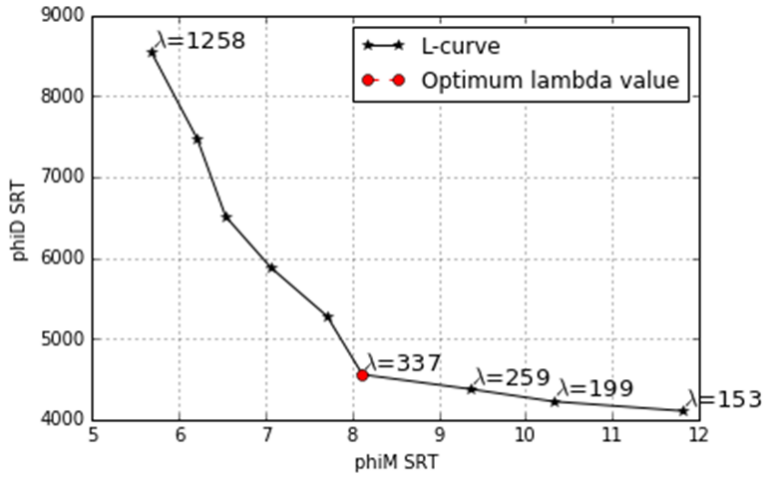

Figure 28: L-curve for SRT synthetic model; the red point is the optimal value for the regularization parameter in the separated inversion

As previously described, the optimal value on such curve was individuated and highlighted as a red point; specifically, the optimal values were 76 for ERT and 337 for SRT.

The convergence was reached after 5 iterations for ERT and 8 iterations for SRT. In both cases, the stop criterion is $\chi^2=1$. The ERT section obtained after the inversion indicates two distinct layers and of the step in the center of the section (Fig.29a). Furthermore, the minimum and maximum resistivity values were close to those of the model, ~500Ω·m and ~1000Ω·m respectively (Fig.29a). The results of the SRT inversion are similar: the two layers are clearly individuated and the step geometry too (Fig.29b). With regard to the velocity values, the maximum value (~3800m/s) is much higher than the velocity value of the true model (Fig.29b).

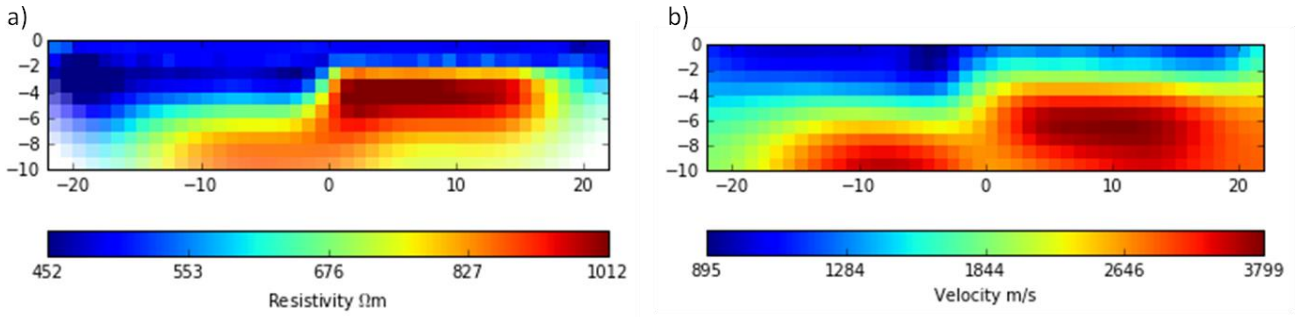

Figure 29: Separated inversion sections for a) ERT and b) SRT.

The cross-gradients function calculated with the results of the separated inversion, as for the synthetic model error free, shows the areas of structural difference between the two models mostly in the right part of the section (Fig.30).

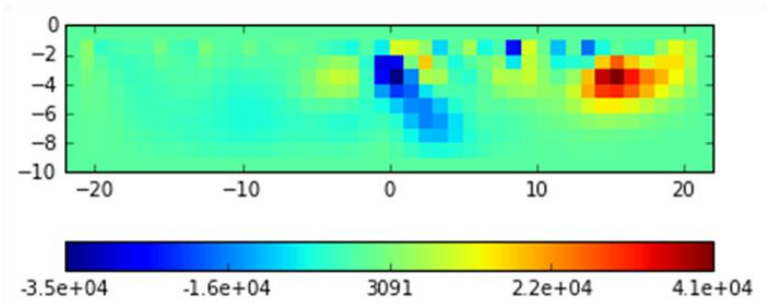

Figure 30: Cross-gradients function for the separated inversion of the noisy data.

In the end, the scatter plot for the separated data is shown (Fig.31). As in the error free model, also in this case, it is not possible to distinguish the different layers.

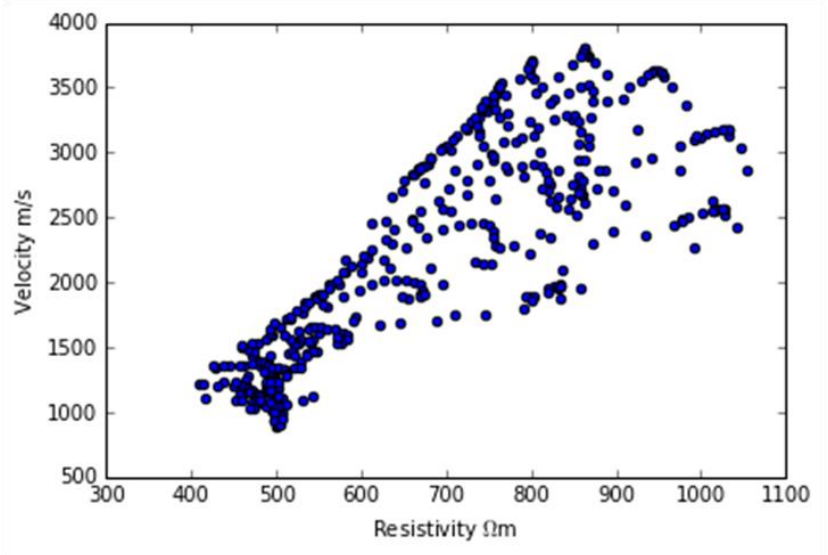

Figure 31: Scatter plot for the separated inversion of data.

Joint inversion

The lambda values used for the separated inversion were also used for the joint inversion, which was performed using the same conditions of the separated inversion. However, since in this case the

number of iterations needed to reach convergence for ERT was different from that for SRT, it was more difficult to obtain good results from the joint inversion. In fact, if the joint inversion is conducted as for the error free model, the obtained values for the seismic velocity data are too high and too different compared to the model and the two layers are not well identified. Accordingly, I examined the results obtained from each iteration of the SRT separated inversion and I noticed that the section obtained after the first iteration was very different from the one obtained after two iterations. The same thing did not happen for the ERT inversion that showed stable results comparing the sections obtained after the first and the second iteration. In conclusion, probably the joint inversion could not give satisfactory results because of the instability of the SRT inversion. Because of that, I decided to conduct the first 5 iterations of the SRT separated inversion, to achieve a stable result, and I used the obtained model to start the joint inversion. In summary, when the number of iterations needed to reach convergence for one method is very different from the one of the other method, the joint inversion must be conducted carefully and started with a stable model.

As regards the lambdaCG value used in the joint inversion, as described in the previous paragraph, I mapped the mean value of the vector obtained for the cross-gradients function for different lambdaCG values in the range of 100-5000 (Fig.32). The minimum value was the one obtained for a lambdaCG value of 500, as highlighted in Fig.32 with a red point.

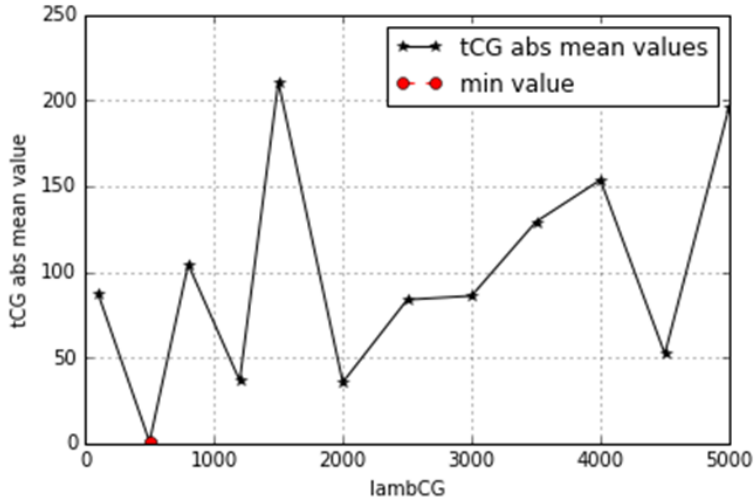

Figure 32: Mean value of the cross-gradients function for the noisy synthetic model. The red point highlights the optimum lambda for the joint inversion.

Using the obtained lambda values, the convergence of the joint inversion was reached in 5 iterations, using also in this case the stopping criterion of $\chi^2=1$, and the final results are satisfactory both for ERT and for SRT, even if the maximum value of SRT inverted model is too high (3916m/s) and too different from the model (Fig.33).

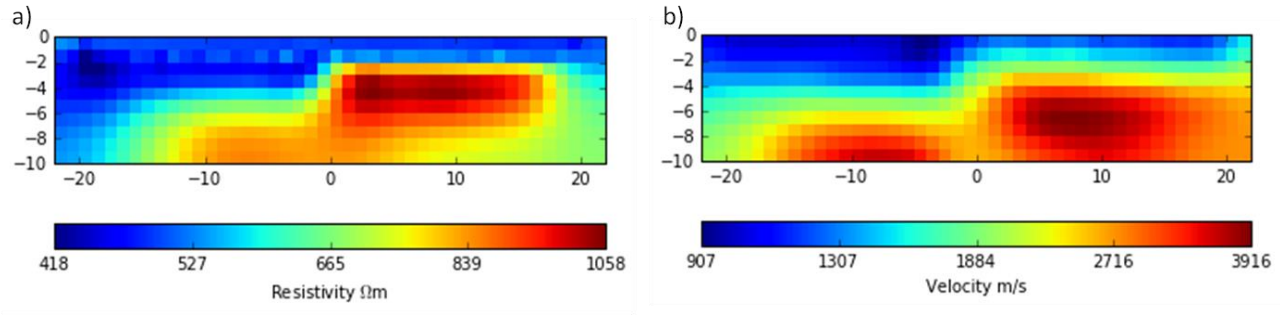

Figure 33: Joint inversion sections for a) ERT and b) SRT.

As regards the cross-gradients function (Fig.34), from the comparison between the one obtained from the joint inversion and the one obtained from the separated inversion, it emerges that it has been reduced thanks to the joint inversion, and consequently it means that after the joint inversion, the ERT and SRT model have an enhanced structural similarity compared to the results of the separated inversion.

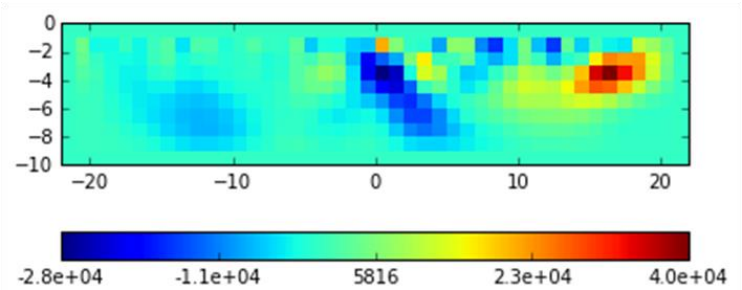

Figure 34: Cross-gradients function for the joint inversion of the noisy data.

The scatter plot obtained with the results of the joint inversion (Fig.35) show a smaller data dispersion when compared to the uncoupled inversion, even if minimal, but does not allow an easier zonation, as for the synthetic model without noise, first of all because of the presence of the noise in the data and probably also for the instability of the SRT separated inversion.

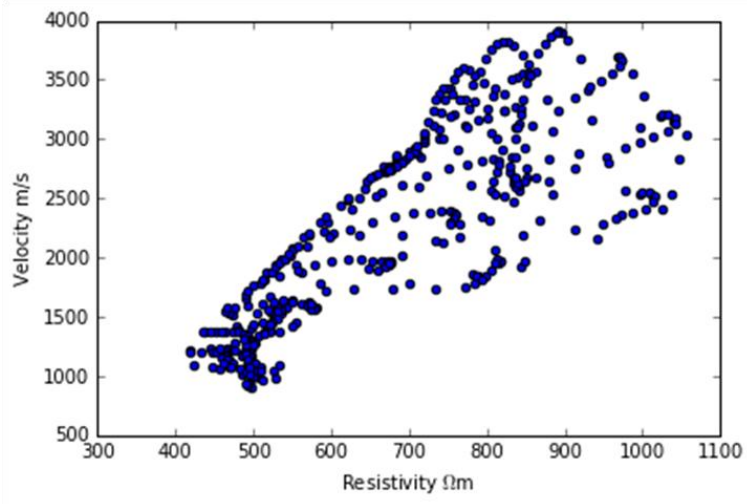

Figure 35: Scatter plot for the separated inversion of data.

6.1.4. Field data: Rieti

Separated Inversion

The field data hereby described were acquired to image and characterizethe shallow subsurface around and below a historic building situated near Rieti (central Italy town), around 80km north of Rome. The construction, which has been built in 1910 as a two-floor masonry building, is used as a National research centre for agricultural studies (Fig.36).

Figure 36: Photograph of the historical building (Cercato and De Donno, 2018).

This building has several structural problems, as fractures on the load-bearing walls and possible differential settlements phenomena in the soil foundations, probably triggered by the 2009 L'Aquila earthquake, and worsen by the seismic sequence of Central Italy (2016-2017). Specifically, the foundations of the building, as deduced from historical sources, are probably 1m wide and extended to a depth that ranges between 1.5 and 2.5m (Cercato and De Donno, 2018). From a geological point of view, the area is located "within the travertine outcropping area, with variable soil thickness above the travertine bedrock" (Cercato and De Donno, 2018). Different types of geophysical measurements carried out at this site (Fig.37): 2D tomographic investigations, specifically electrical resistivity, P-and SH-wave seismic tomography, located along three lines (L1, L2 and L3 in Fig.37); three-dimensional ERT data, acquired near the building using C and L-shaped arrays. Lines L2 and L3 were acquired to obtain a geophysical model of the undisturbed subsoil near the building, while the L1 line was a high-resolution survey that investigatesthe soil conditions that influence the building. Three dimensional ERT measurements have been necessary for the reconstruction of the shape of the foundations. In 2017, the excavation of two trenches and the execution of a Standard Penetration Test (Fig.37) were eventually carried out to collect some geotechnical data and samples, to validate the reconstructed model, as a support for the geological interpretation of the geophysical results.

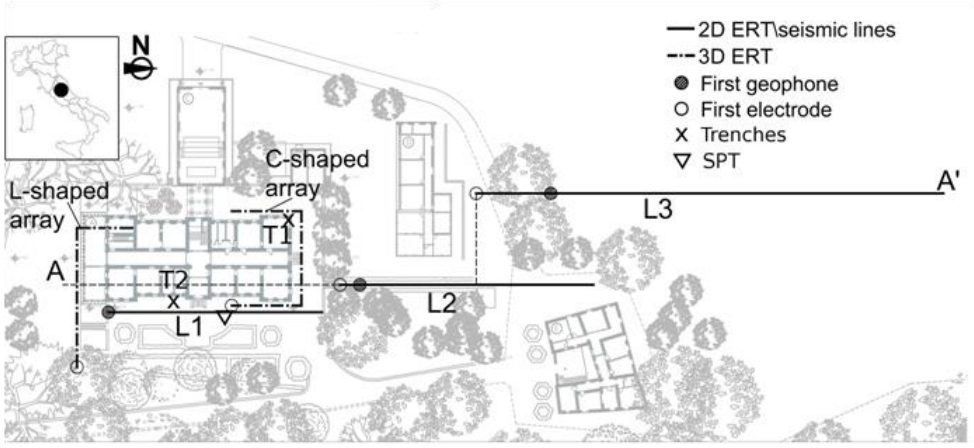

Figure 37: View of the examined site with the location of geophysical and geotechnical measurements (Cercato and De Donno, 2018).

Here the focus is on the L3 line, which presents a complete dataset for joint inversion and allows the reconstruction of the shallow subsoil in the area near the building. The ERT measurements were acquired with 48-electrodes IRIS Instruments SyscalPro48 and using a combination of dipole-dipole and Wenner-Schlumberger configurations with stainless steel electrodes 2m spaced apart. The seismic refraction tomography (SRT) data were acquired using a 48-channel system of 4.5 Hz vertical geophones 1m spaced (I have examined the P-wave data) and a 7 kg hammer on a steel plate as source, with a Geometrics Geode seismograph at a sampling rate of 0.125 ms. The acquisition was performed "with a geophone streamer, made up by abrasion-resistant steel tripod plates connected and towed by a high tensile-strength band" (Cercato and De Donno, 2018).

In particular, as it is possible to see in Fig.37, the ERT survey is longer than the SRT's, but since for the joint inversion the same extension of data is necessary, I eliminated the data of the ERT survey that were over the SRT one.

The ERT and SRT data were firstly examined separately, performing the separated inversion with the pyBert and pyGimli packages (Rücker, 2011; Günther et al., 2006), using the finite element algorithm of the pyBert package as the ERT forward modeling, while the FMM of the pyGimli package as the SRT forward modeling, as in the synthetic model. Specifically, in order to choose the optimal value of lambda for the inversion, the L- curve was computed using for ERT a range of lambda values of 10-1258 (Fig.38), while a smaller range, 100-1258 for SRT (Fig.39), since lambda values <100 gave not satisfactory results. The optimal lambda values obtained with the L-curve are 164 for ERT data and 316 for SRT data (Figs.38-39).

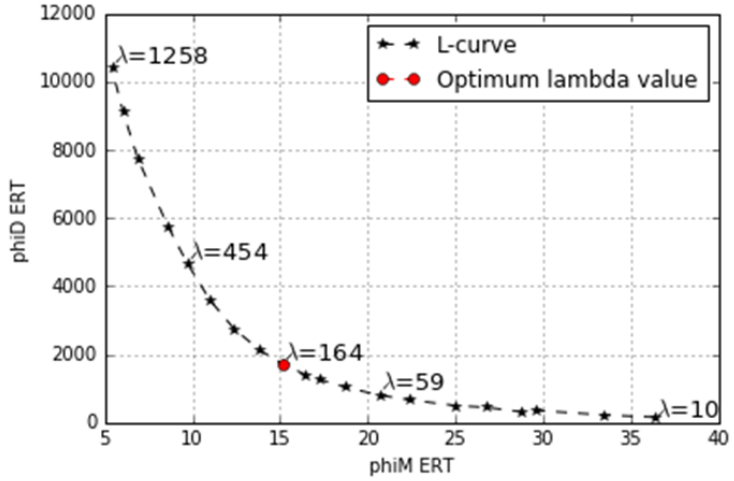

Figure 38: L-curve for ERT field data. The red point represents the optimum lambda value.

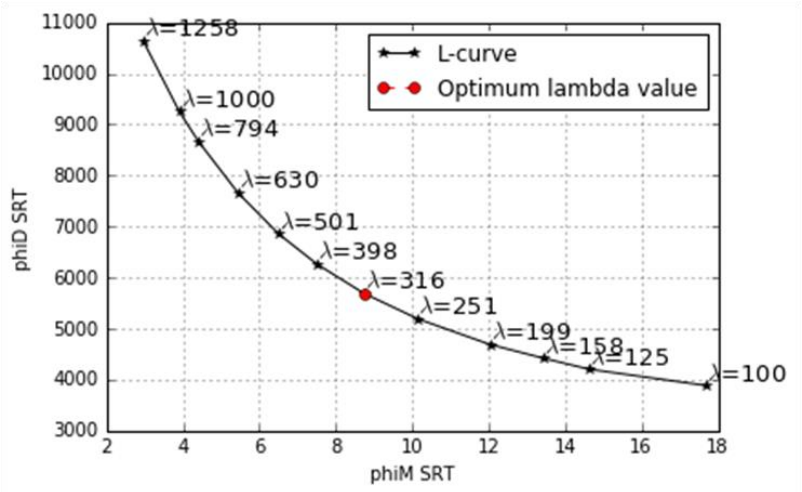

Figure 39: L-curve for SRT field data. The red point represents the optimum lambda value.

The results obtained from the separated inversion confirmed the results obtained by Cercato and De Donno (2018) with different algorithms and programs. The depth of the investigation was higher for the ERT when compared to the SRT inversion; in order to perform the joint inversion, I limited the depth of the model to 10m to achieve a comparable coverage with both methods.

The ERT convergence was reached with 4 iterations, while the SRT one with 5 iterations. In both cases, the stop criterion is given by $\chi^2=1$. The ERT section showed three different layers: the conductive ones with values of resistivity <20Ω·m for the shallower and in the range of 30-60Ω·m for the middle one, and a deeper resistive layer, that is visible in the left part of the section, with a resistivity >100Ω·m, that probably represents a travertinous formation (Fig.40).

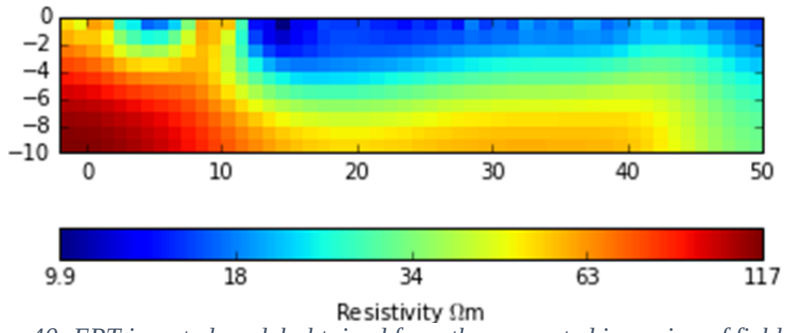

Figure 40: ERT inverted model obtained from the separated inversion of field data

The SRT inverted model confirms the ERT inverted results, since a surface layer with low P-wave, with velocities in the range of 300-600m/s is well imaged. These values are typical of a weathered layer. Then, velocities increase with depth, reaching velocities >800m/s for the middle layer and >1000m/s (Fig.41) in the left part of the section in a position that overlaps the one of the ERT section. These high velocities (>1000m/s) suggest the presence of a travertineous layer, that is visible only in the left part of the ERT section, because of the maximum depth of 10m that is reached by the SRT inversion.

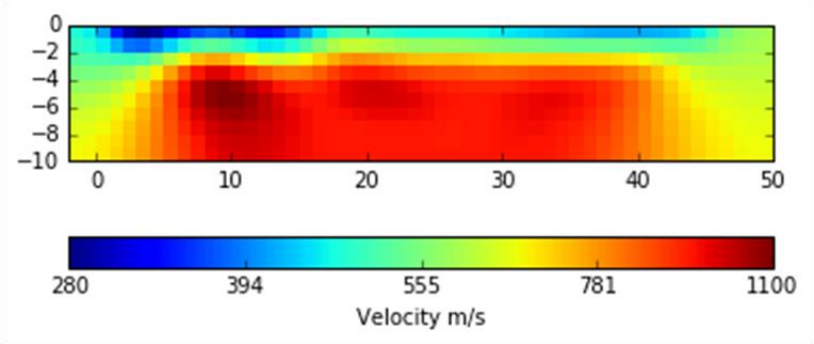

Figure 41: SRT inverted model obtained from the separated inversion of field data

The section of the cross-gradients function (Fig.42) obtained from the ERT and SRT inverted models and calculated using the forward difference scheme and a uniform discretization, exhibits some areas of structural difference in the left part of the section, with values around zero in most of the section.

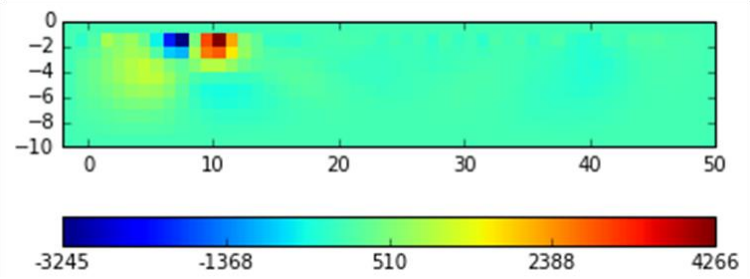

Figure 42: Cross-gradients function obtained from the data of the separate inversion

In the end, the scatter-plot (Fig.43) shows a certain dispersion of data, typical of field data, which are affected by measurement and modeling error.

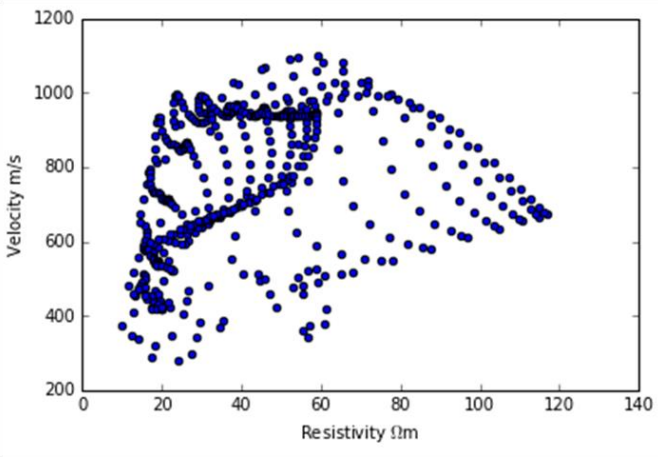

Figure 43: Scatter plot of the separated inverted data

Joint inversion

As for the joint inversion of synthetic data, also for the field data the joint inversion was conducted using the lambda values of the separated inversion and the same conditions of the separated inversion. The lamdaCG value was chosen as the minimum mean value of the cross-gradients vector, as for the synthetic model. Specifically, from the graph in fig.44 the lambdaCG value which corresponds to the minimum mean value is 4000.

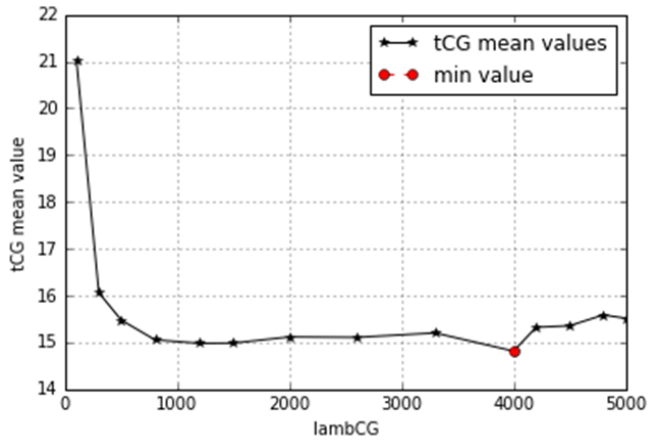

Figure 44: Graph of the mean value of the cross-gradients function for the field data. The red point represents the optimum lambda for the joint inversion

The joint inversion conducted with these parameters reached the convergence after 5 iterations. As in the previous cases, the stop criterion is given by the $\chi^2=1$. The ERT inverted model is very similar to the one obtained by the separated inversion (Fig.45), while the SRT section shows some differences with respect to the corresponding separated inversion. In fact, it individuates the presence of the first layer, with velocities in the range of 300-600m/s, and of a second layer, with higher velocity, between 700 and 900m/s, as in the separated inversion, but since the maximum value is 949m/s, the travertinous area seems not to be individuated (Fig.46).

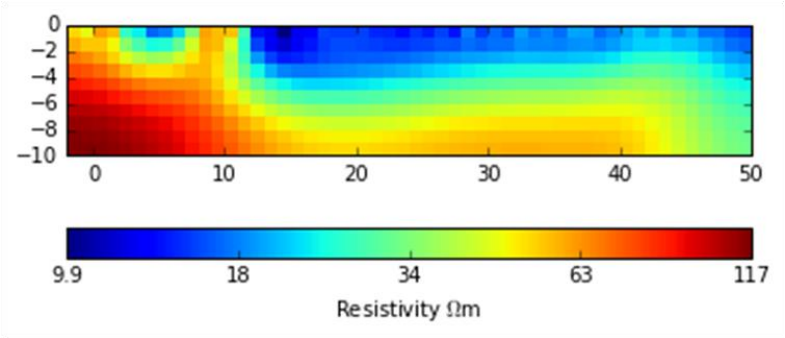

Figure 45: ERT inverted model obtained from the joint inversion of field data.

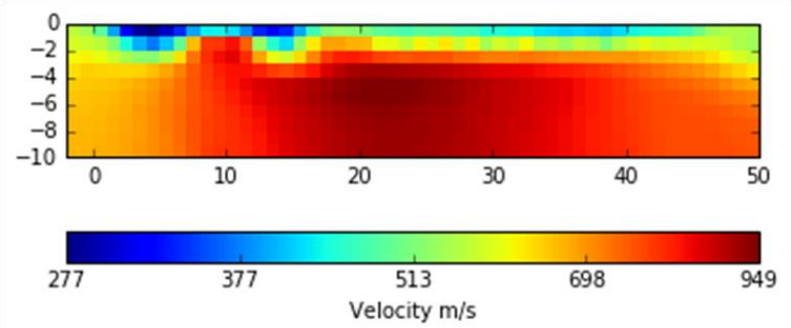

Figure 46: SRT inverted model obtained from the joint inversion of field data.

However, the significant sections for the joint inversion, as the cross-gradients function and the scatter plot, show the enhancement of the subsoil reconstruction obtained through the joint inversion. In fact, the cross-gradients function shows a remarkable reduction of the cross-gradient term, meaning that the two models have a better structural similarity than the separated inversion (Fig.47) and the scatter-plot highlights a less dispersion of data than the separated inversion, even if it is not possible to define different zonations of the subsurface (Fig.48).

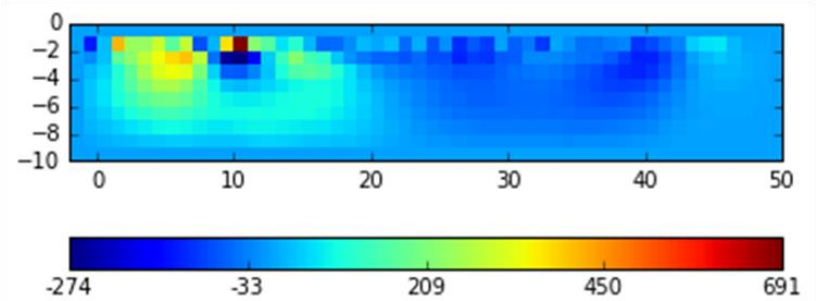

Figure 47: Cross-gradients function for the joint inversion.

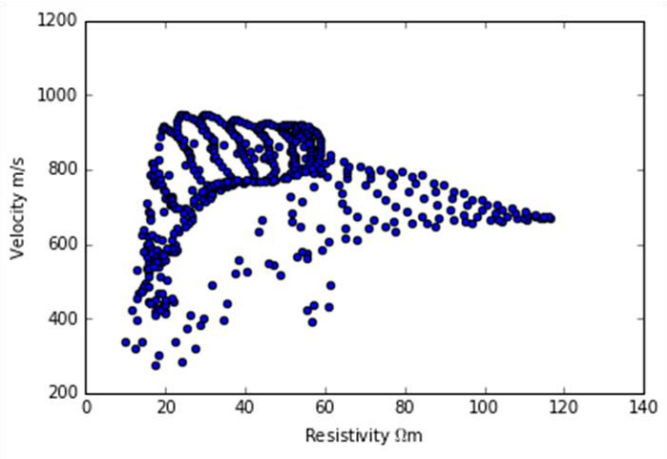

Figure 48: Scatter plot of the joint inversion data.

6.2. Seismic forward modeling results

In this paragraph I will describe the results obtained from the implementation of the seismic forward modeling, the Multistencils Fast Marching Method (MSFMM), reported in paragraph 4.1.4. At first, I will compare the traveltimes obtained with the MSFMM algorithm with the known analytic solution for a two-layer model. The results will be described for two regular grids with different size of the cells, in order to understand how and if the solution becomes more accurate with a finer grid. Then, I will introduce the sections obtained from the separated and joint inversion using the MSFMM as forward modeling for the synthetic model described in paragraph 6.1.1 and 6.1.2 with and without noise in the data and for the field data of Rieti, described in paragraph 6.1.3. The algorithm of the MSFMM has been included in my version of the pyGimli package and inversion routine (Rücker et al., 2017).

6.2.1. Comparison between computed and analyc traveltimes for a two-layer model

In order to understand the accuracy of the MSFMM's implementation, a two-layer synthetic model whose solution is known, has been examined. Specifically, the model is composed of two horizontal layers with increasing velocities with depth, 1000m/s and 3000m/s for the first and second layer respectively (following an example of the pyGimli package (Rücker et al., 2017)) (Fig.49).

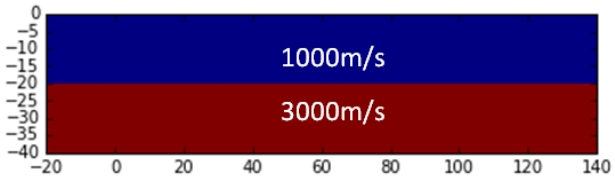

Figure 49: Synthetic two-layer model.

The analytic traveltimes have been computed (Rabbel, 2006) and compared to the traveltimes computed through the MSFMM algorithm.

Specifically, recalling the four cases examined for the MSFMM, as in paragraph 4.1.4:

- case A: the MFMM considering only the S_1 stencil with a first order approximation of the directional derivative (Eq.4.34), which is the classic FMM;
- case B: the MFMM considering not only the first, but also the second order approximation of the directional derivative for the S_1 stencil (Eq.4.34-Eq.4.35), which is the classic FMM with higher accuracy;
- case C: the MSFMM, considering also the stencil S_2 with the first order approximation of the directional derivative (Eq.4.34 and Eq.4.36);

- case D: the MSFMM, considering both the stencils and using not only the first, but also the second order approximation of the directional derivative (Eq.4.34, Eq.4.35, Eq.4.36 and Eq.4.37).

I will examine the traveltimes obtained through the implementation of the MSFMM for each case and considering two regular grids: a grid with dimension of the cells of 4mx4m (Fig.50a) and another finer grid with dimension of the cells of 2mx2m (Fig.50b). The different size of the cells for the two grids is important to assess the discretization impact and the effects in the accuracy.

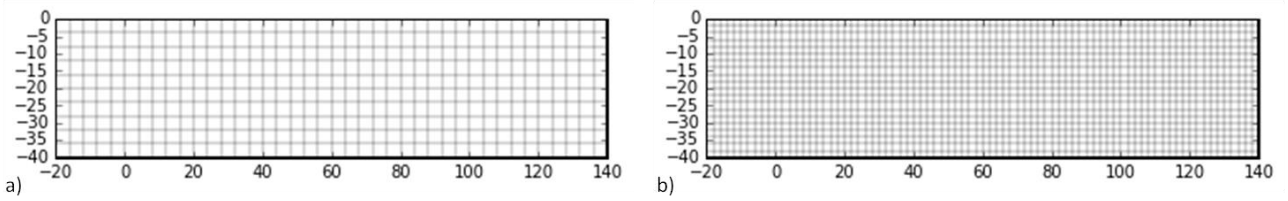

Figure 50: Regular grids involved in computing the traveltimes for the MSFMM. a) grid with dimension of the cells of 4mx4m; finer grid with dimension of the cells of 2mx2m

In the next lines, I will report in the graphs not only the comparison between calculated and analytical traveltimes, but also the error, i.e. the normalized difference between the two time vectors, and the wavefronts. In particular, the image of the wavefrontsclearly depicts the physics of the phenomenon.

Case A

In the traveltime graphs of Fig.51a, d, it is clearly visible that the difference in the cell dimension of the grids strongly affects the final result. In fact, the straight line that represents the critically refracted arrivals for the 2x2m grid results to be closer to the analytical solution than the one obtained for the 4x4m grid.

This result is also confirmed by the graph that represents the error (Fig.51b, e), calculated as the difference between the two vectors, analytical and computed traveltimes, normalized by the analytical solutions. Clearly, the absolute error occurs at the points where the traveltimes start to sense the velocity change (Fig.51b, e). For larger offset, the error decreases asymptotically. The minimum value for the error is around -1.5% for the 4x4m grid (Fig.51b), while it decreases to around -0.7% for the 2x2 grid (Fig.51e). A similar behavior has the maximum error, which is greater than 1.5% for the 4x4m grid decreasing to around 1.3% at the maximum offset, while for the 2x2m grid it is greater than 1%, but decreases to a value that is less than 1% at 140m distance.

Regarding the wavefronts (Fig.51c, f), for both grids the second layer is well identified, while the wavefronts seem to be more accurate for the 2x2m grid.

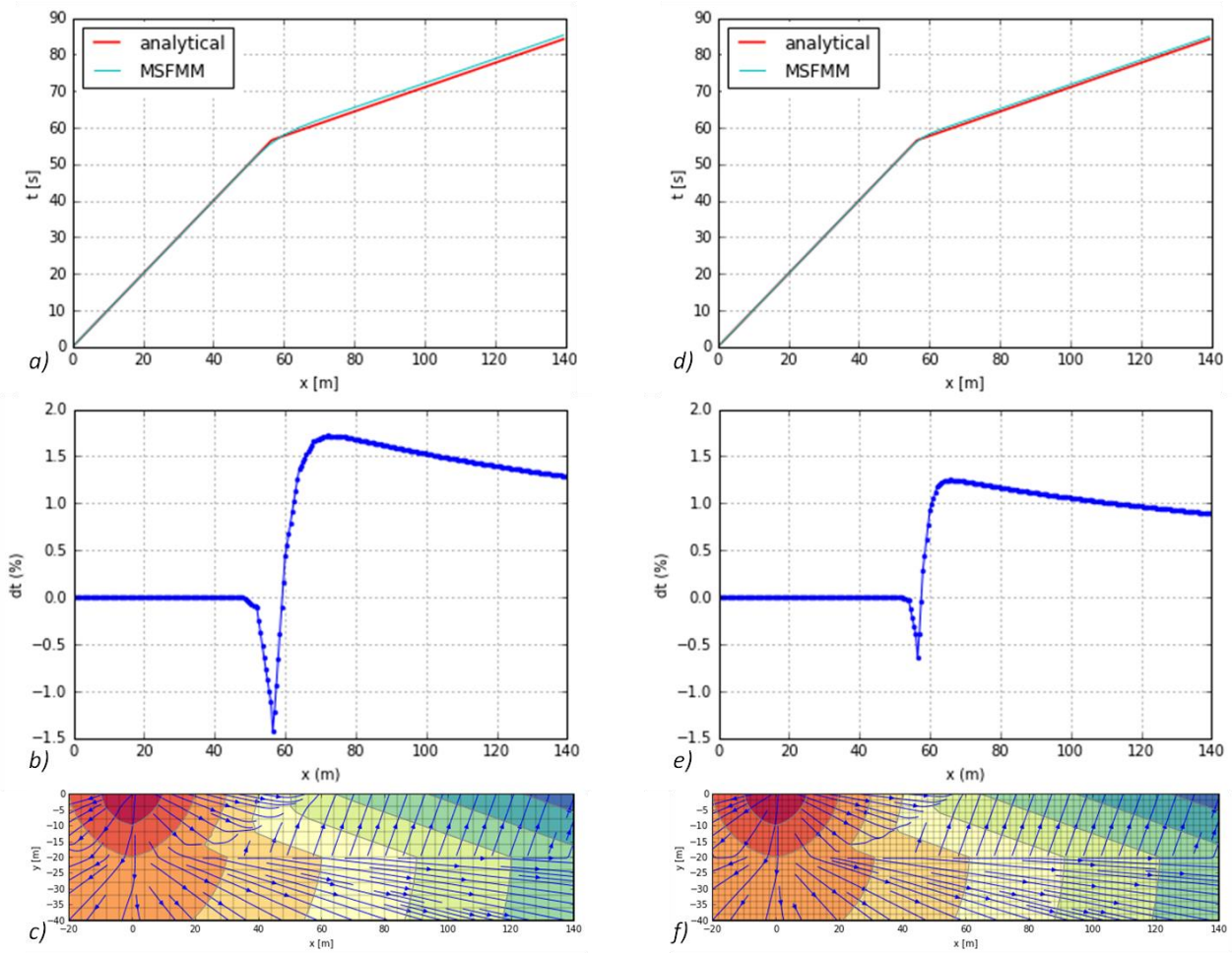

Figure 51: Graphs obtained for case A, the classic FMM: a), b) and c) obtained for the grid with dimensions 4x4m, while d), e) and f) for the 2x2m grid. a) and d) represent the traveltime diagrams; b) and e) represent the errors, calculated as the difference between the analytical and the computed traveltimes vectors; c) and f) represent the wavefronts for the two grids.

Case B

As for the classic FMM (case A), also for the FMM with the second order approximation of the directional derivative, the dimension of the cells influences the accuracy of the final result. This is clearly visible both in the traveltime graph (Fig.52a, d), in which the critically refraction arrivals for the 2x2m grid are closer to the analytical solution than the one for the 4x4m, and in the error graph (Fig.52b, e), where the minimum error decreases from around 1% (4x4m grid) to around 0.5% (2x2m grid), while the maximum error decreases from 1.5% (4x4m grid) to 0.8% for the 2x2m grid. The error that is reached at 140m is around 0.9% for the 4x4m, while it is around 0.5% for the 2x2m grid. Furthermore, the results obtained seem to be more accurate than the classic FMM, since the error further decreases. In fact, the maximum error that is around 1.8% for the classic FMM, reaches the value that is less than 1.5% for the 4x4m grid using the second order approximation, while for the 2x2m grid it decreases from around 1.3% to around 0.8%. The decrease is similar for the minimum error. Moreover, the wavefronts (Fig.52c, f) are more accurate in this case than for the classic FMM.

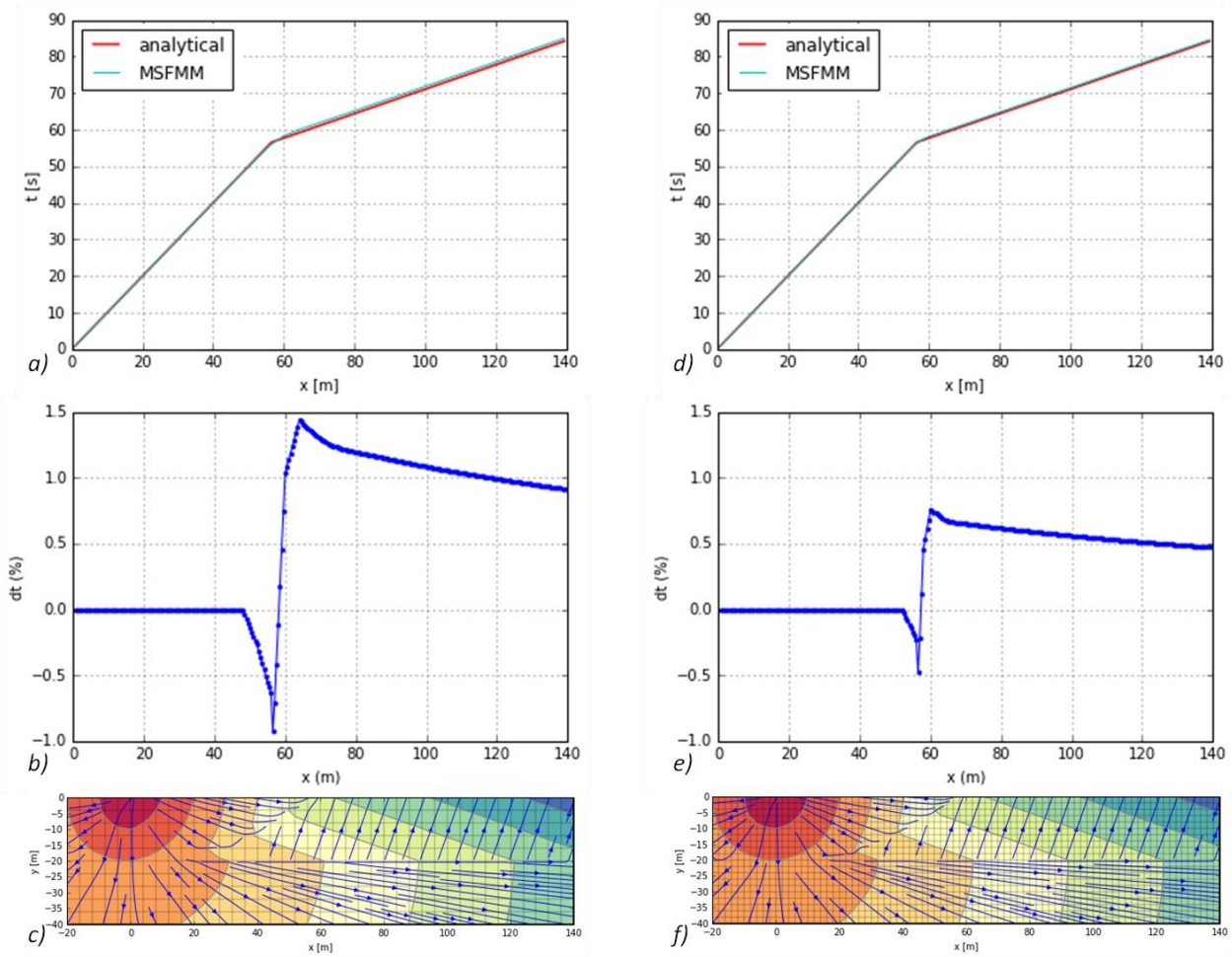

Figure 52: Graphs obtained for case B: a), b) and c) obtained for the grid with dimensions 4x4m, while d), e) and f) for the 2x2m grid. a) and d) represent the traveltime diagrams; b) and e) represent the errors, calculated as the difference between the analytical and the computed traveltimes vectors; c) and f) represent the wavefronts for the two grids.

Case C

In this case the cross-neighbors are used with the first order approximation of the directional derivative. The results are similar to the ones of case A and B, as regards the influence of the cell dimension of the grids. In fact, the critically refracted traveltimes (Fig.53a, d) obtained for the 2x2m grid are closer to the analytical ones than for the 4x4m grid and consequently, the error is reduced for the 2x2m grid. The minimum error is reduced from around 1.7% for the 4x4m grid to 0.8% for the 2x2m grid, almost 1%, while the maximum error is less reduced, since from little more than 1% for the 4x4m grid it reaches the value around 0.9% for the 2x2m grid (Fig.53b, e). Besides, the error is further reduced compared to the classic FMM considering both the first and the second order approximation of the directional derivative. In fact, examining the maximum error for the 4x4m grid, it is around 1.7% in case A and is reduced to around 1.4% for case B, while in this case it is further reduced reaching the value of around 1.1%. Instead, for the 2x2m grid the maximum error is reduced from around 1.3% for case A, to 0.8% for case B, while it reaches a value around 0.9% for case C.

With regard to the wavefronts (Fig.53c, f), it seems that they are more accurate than the classic FMM (case A and B).

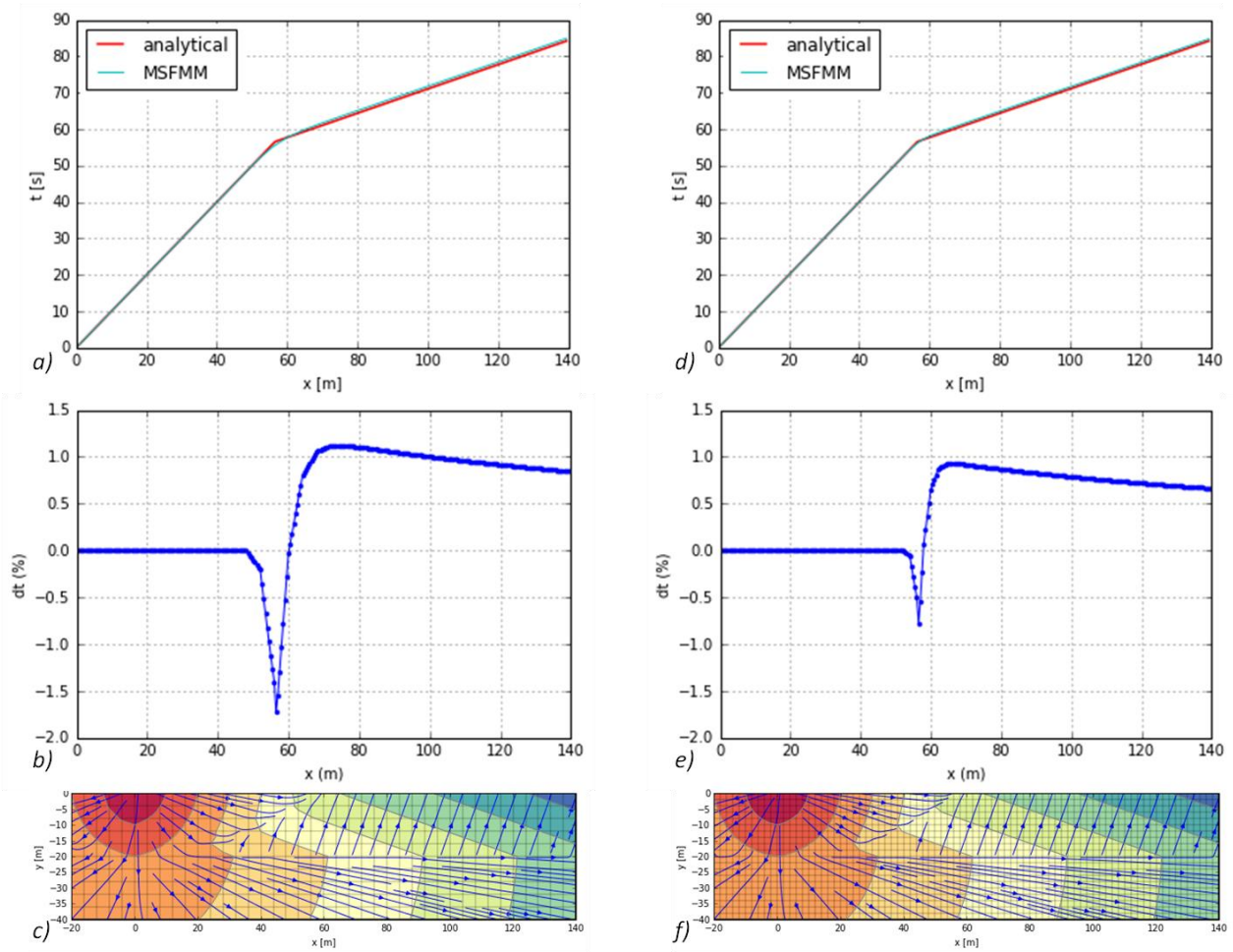

Figure 53: Graphs obtained for case C: a), b) and c) obtained for the grid with dimensions 4x4m, while d), e) and f) for the 2x2m grid. a) and d) represent the traveltime diagrams; b) and e) represent the errors, calculated as the difference between the analytical and the computed traveltimes vectors; c) and f) represent the wavefronts for the two grids.

Case D

In this last case, the two stencils are used and also the first and second approximations of the directional derivative. As in the previous cases, the final result is influenced by the cell size of the grid (Fig.54a, d). However, the error is only negative and reaches the asymptotically value of 1.4% for the 4x4m grid, while the 0.8% for the 2x2m grid (Fig.54b, e). Furthermore, the figure that shows the wavefronts (Fig.54c, f) highlights a shape that is more accurate than the other cases.

However, in all the four cases, the computation error is in the range of the accuracy of the first arrival picking.

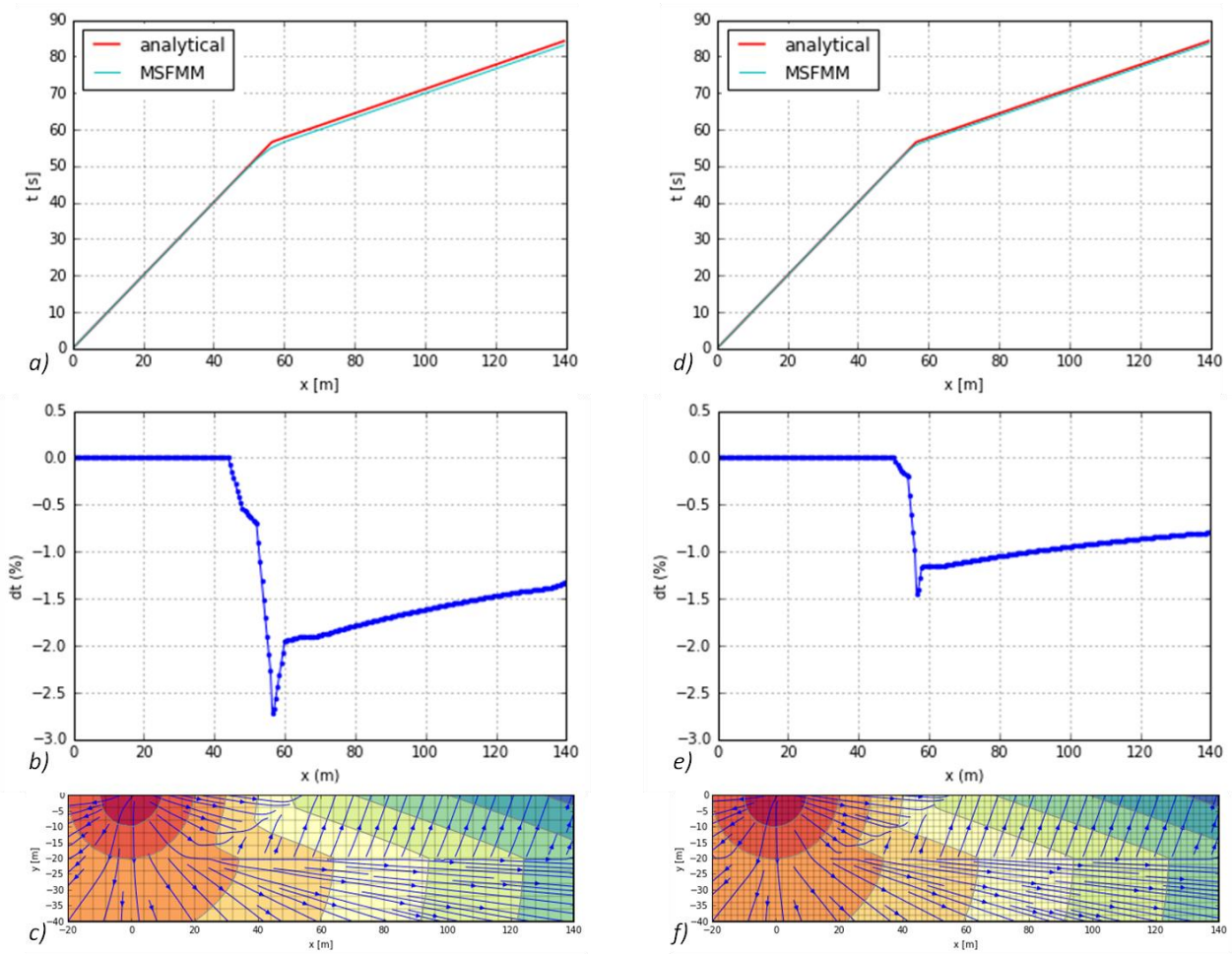

Figure 54: Graphs obtained for case D: a), b) and c) obtained for the grid with dimensions 4x4m, while d), e) and f) for the 2x2m grid. a) and d) represent the traveltime diagrams; b) and e) represent the errors, calculated as the difference between the analytical and the computed traveltimes vectors; c) and f) represent the wavefronts for the two grids.

6.2.2. *Separated SRT inversion for the synthetic model without noise*

The synthetic model examined, is the same described in paragraph 6.1.1. It consists of two layers with the seismic velocity that increases with depth, specifically from 1000 to 3000m/s with a step 2m high in the central part of the section (Fig.55).

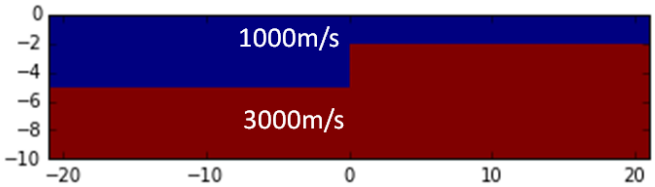

Figure 55: Synthetic model with values of resistivity and seismic velocity that increase with depth.

The configuration used for the seismic refraction tomography measurements is the same described in paragraph 6.1.1 and specifically, 41 geophones, 1m spaced, and 9 shots, 5m spaced. The forward modeling used is the implemented MSFMM (described in Chapter 4). As previously, no noise has been added to the synthetic data and a regular mesh has been used. The separated inversion was carried out for each case of the MSFMM, using the same lambda value in order to compare the

results. Hence, I used 30 as lambda value and the same conditions as in the separated inversion carried out using the FMM implemented in pyGimli as forward modeling. In all the following cases, the convergence is obtained when the stop criterion of $\chi^2 = \frac{\Phi_d}{N} = 1$ is reached.

Case A

The image of the computed traveltimes (Fig.56a) obtained for the synthetic model shows that the used forward modeling detected the different velocity of the two layers, since the traveltime curves show different slopes, and the step, for the anomaly that is visible in the center part of the section. The section obtained from the separated inversion after 4 iterations (Fig.56b) clearly highlights the two different layers, with the same values of the model, since the minimum value is 1107m/swhile the maximum 3015m/s. The step is well individuated, since in the left part of the sectionthe presence of a higher velocity layer is detected around 5m in depth, while in the right part around 2m in depth, as in the model. The results obtained in Fig.56 for the classical FMM are slightly different from the ones of Fig.19 obtained using the FMM implemented in the pyGIMLi package since the implementations are slightly different.

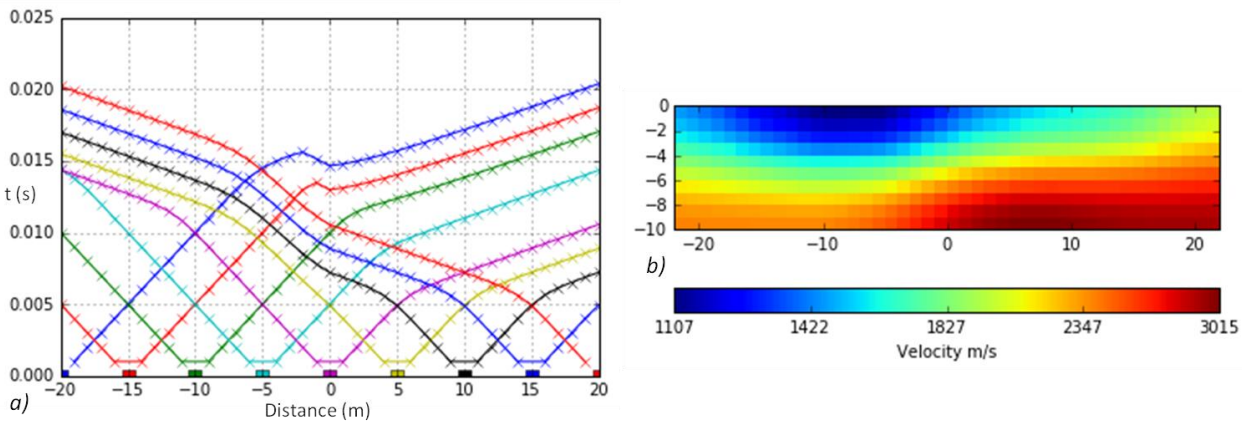

Figure 56: a) Computed travel times responses for the synthetic model without noise for case A; b) section obtained by the separated inversion of synthetic data.

Case B

As in case A, also in case B (classic FMM using the second-order approximation of the directional derivative) the graph of the computed traveltimes (Fig.57a) highlights not only the presence of two layers with different velocities, but also the step in the middle part of the section, even if the traveltimes result to be lower than the ones of case A. The section of the separated inversion (Fig.57b), obtained after 4 iterations, shows the presence of the two different layers and of the step in the central part of the section, but differently from the one of case A, it shows a smaller range of velocity values, even if very close to the ones of the model (1113-2950m/s), and it seems to

individuate in a better way the depth of the second layer. In fact, in the left part of the section, the second layer is better identified at the depth of 5m than the section of case A.

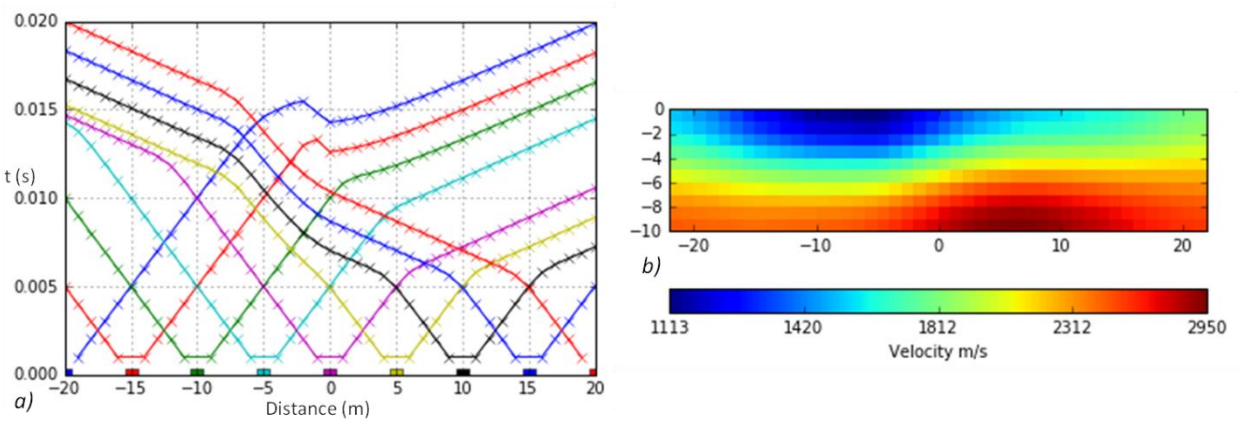

Figure 57: a) Computed travel times responses for the synthetic model without noise for case B; b) section obtained by the separated inversion of synthetic data.

Case C

In this case, in which the cross-neighbors are used, the calculated traveltimes result to be lower than the ones of case A (Fig.58a), while the section of the separated inversion (Fig.58b) shows a range of values that is closer to the one of the model than case A. Furthermore, the obtained model seems to be more accurate than case A and B, since the depth of the second layer is well identified not only in the left part of the section at the depth of 5m, but also in the right part at a depth of around 2-3m.

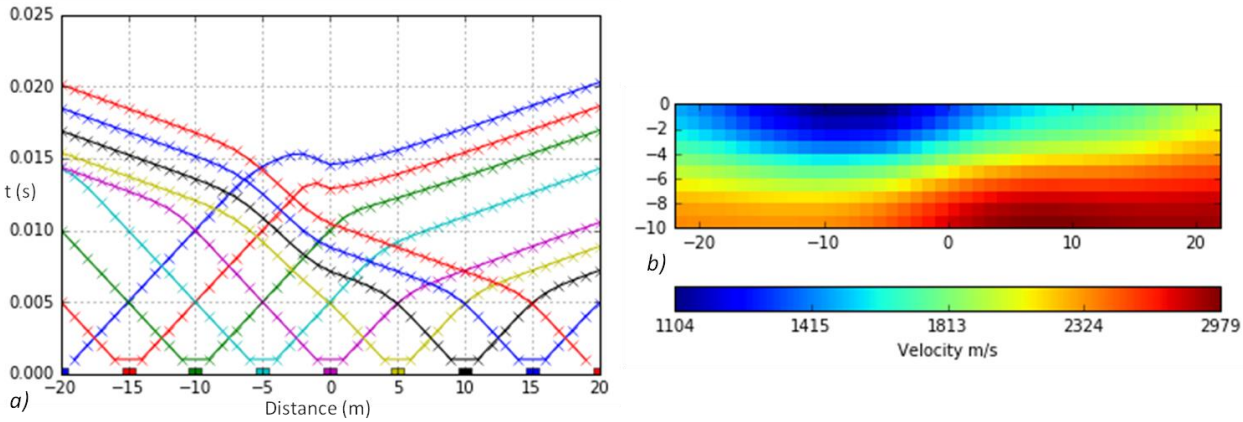

Figure 58: a) Computed travel times responses for the synthetic model without noise for case C; b) section obtained by the separated inversion of synthetic data.

Case D

In the end, using both S_1 and S_2 stencils and in both cases the second order accuracy, the traveltime graph (Fig.59a) shows computed traveltimes that are lower than the ones calculated for the other cases. Furthermore, the velocity section (Fig.59b) obtained from the inversion process after 4 iterations shows a range of values that is smaller than the one of case C, and as in case B, a better accuracy as regards the detection of the second layer in the left part of the section at a depth that is

around 5m, but not in the right part of the section, where the model of case C results to be more accurate.

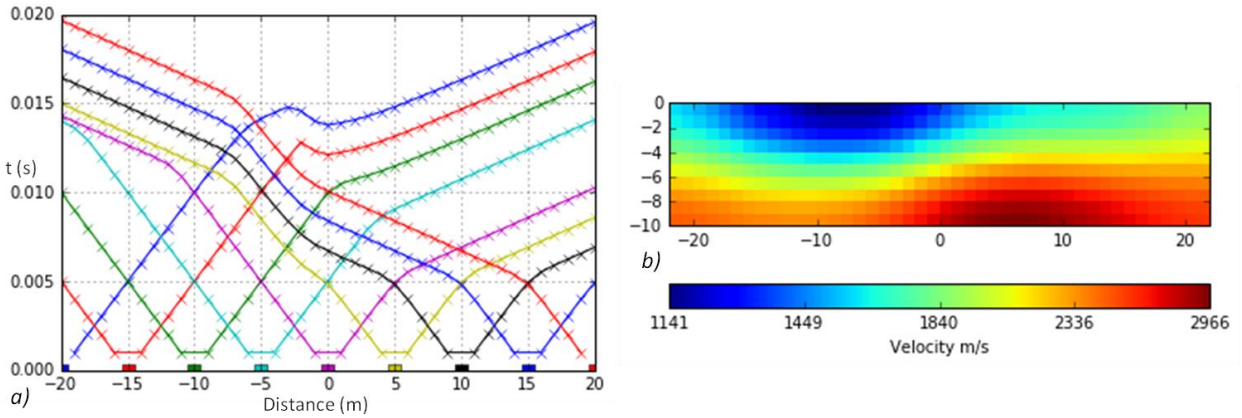

Figure 59: a) Computed travel times responses for the synthetic model without noise for case D; b) section obtained by the separated inversion of synthetic data.

In summary, in all the four cases, good results have been obtained since the range of velocity values is very close to the one of the model, the two layers are well identified and also the step in the middle part of the section. However, between all the cases, it seems that the best result is obtained for case C, in which the cross neighbors are used. In fact, the range of values is closer to the one of the model (1104-2979m/s) and the depth of the layer with a higher velocity is better identified not only at the 5m, but also at the 2m depth. For this reason, I decided to describe the next results only for case C.

6.2.3. *Joint inversion for the synthetic model without noise*

In this paragraph, I will examine the results obtained from the joint inversion for the synthetic model described in the previous paragraph (6.2.2). I will describe also the separated inversion results, since they are important for the comparison with the joint inversion ones.

Only for clarity, I present again the examined model, with resistivity values of 500-100Ω·m and velocity values of 1000-3000m/s (Fig.60).

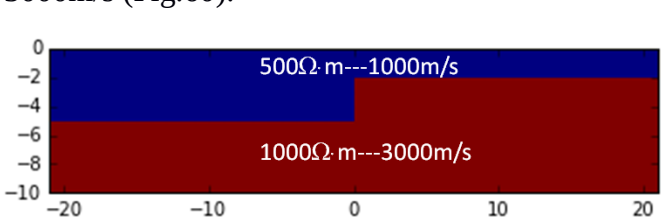

Figure 60: Synthetic model with values of resistivity and seismic velocity that increase with depth.

As described in paragraph 6.2.2, the electrical tomography measurements have been simulated using the dipole-dipole configuration, with 41 electrodes, 1m spaced, while the seismic refraction tomography measurements with 41 geophones, 1m spaced, and 9 shots, 5m spaced. The synthetic

responses have been obtained using the forward modeling of the pyBert package for the resistivity measurements, while the MSFMM (case C) for the seismic traveltimes computation (Fig.61).

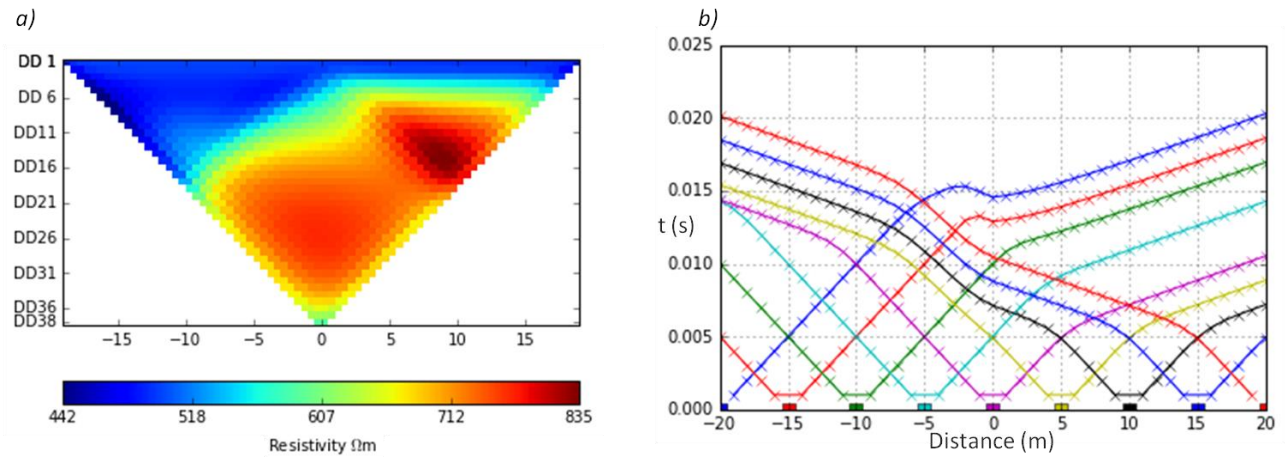

Figure 61: a) Synthetic electrical resistivity response obtained from the model without noise; b) Computed travel times responses obtained from the model without noise.

The synthetic data were firstly inverted separately, setting the lambda value to 30 for both the geophysical methods. The convergence is reached after 2 iterations for the resistivity model, while after 4 iterations for the seismic model (Fig.62). In both cases, the convergence is obtained following the stop criterion of $\chi^2=1$.

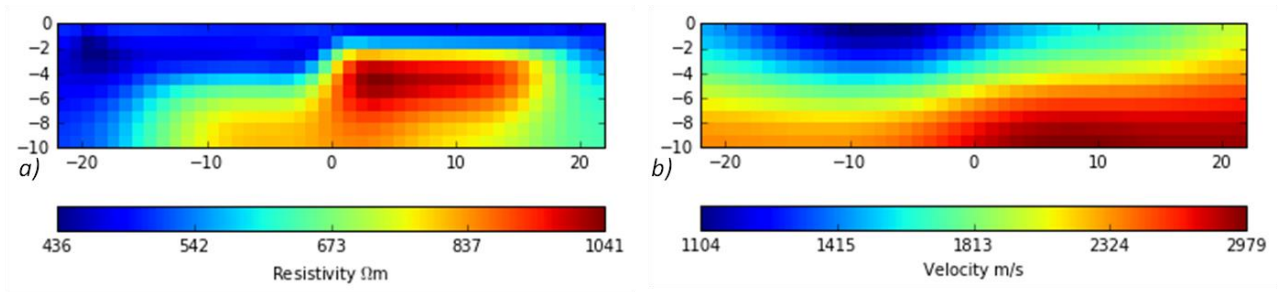

Figure 62: Separated inversion results: a) ERT; b) SRT.

Using the results obtained by the separated inversion, it is possible to calculate the cross-gradients function (Gallardo and Meju, 2003, 2004) as described in Chapter 4, using the forward differences and using a uniform discretization (Fig.63).

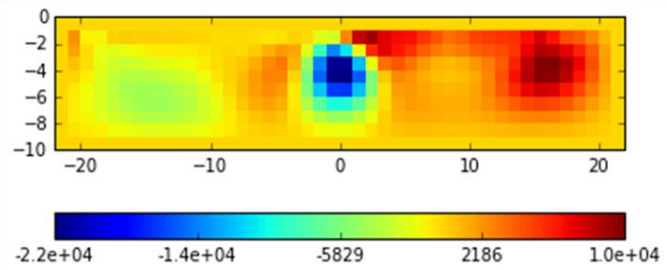

Figure 63: Cross-gradients function calculated for the separated inversion.

As described in the previous paragraph, the cross-gradients function shows areas of high and low cross-gradients values, which represent areas of structural difference of the separated inverted models.

As regards the scatter plot obtained from the results of the separated inversion (Fig.64), it shows a wide dispersion of data.

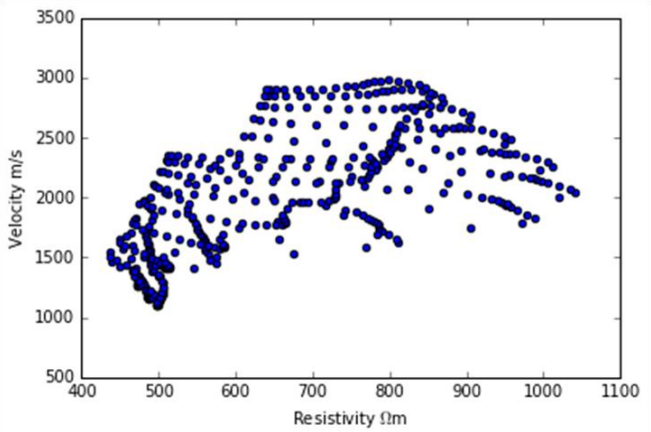

Figure 64: Scatter plot of seismic velocity versus electrical resistivity for the separated inversions of ERT and SRT.

Using the same conditions of the separated inversion, the joint inversion was conducted. As regards the lambdaCG value, it was chosen as the minimum mean value of the cross-gradients vector and specifically as 800 (Fig.65). The convergence was reached after 4 iterations, with the stopping criterion of $\chi^2=1$.

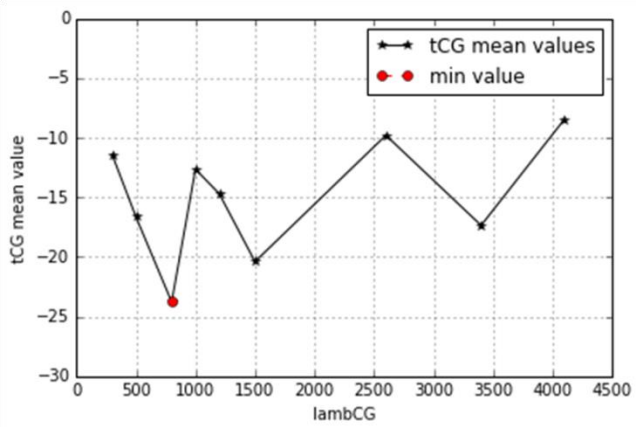

Figure 65: Graph of the mean value of the cross-gradients function for the synthetic modelwithout noise. The red point represents the optimum lambda value for the joint inversion

The sections obtained from the joint inversion (Fig.66) show for both the ERT and SRT sections the presence of two layers and the step in the middle of the section with values that are similar to the ones of the model. However, the result obtained for the seismic measurements (Fig.66b) shows a more accurate identification of the second layer in the right part of the section, at the depth of

around 2m, compared to the separated inversion, even if the maximum value obtained is lower than the one in the separated inversion.

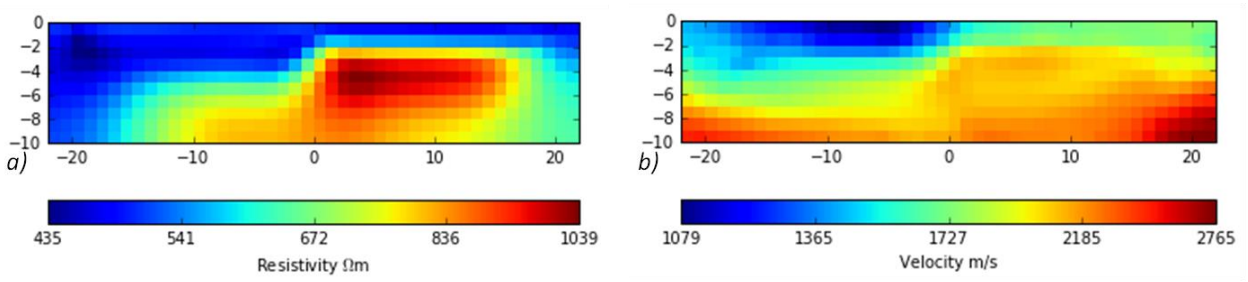

Figure 66: Joint inversion results: a) ERT; b) SRT.

The associated cross-gradients function (Fig.67) shows values that are lower than the ones of the separated inversion, even if not much.

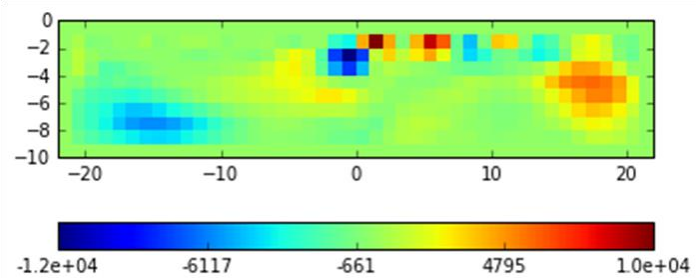

Figure 67: Cross-gradients function calculated for the joint inversion.

However, the cross plot (Fig.68) shows a less dispersion of data that are more grouped than the separated inversion.

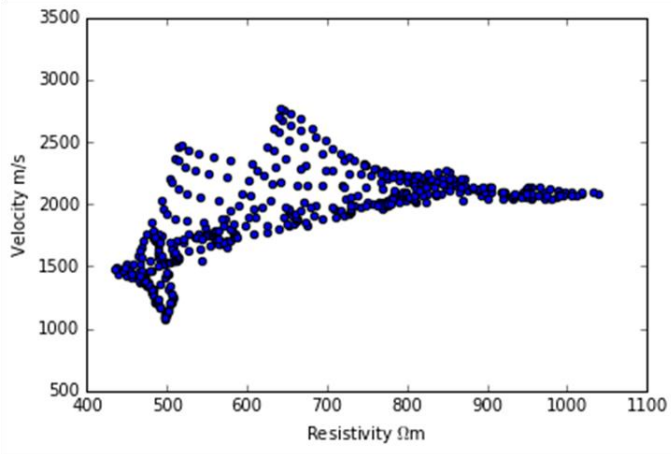

Figure 68: Scatter plot of seismic velocity versus electrical resistivity for the joint inversion of ERT and SRT.

6.2.4. Separated and joint inversion for the synthetic model with noise

The results obtained adding some noise to the synthetic model described in the previous paragraph will be examined. Specifically, as described in paragraph 6.1.2, synthetic data computed using the same configuration described for the synthetic model with no noise (41 geophones, 1m spaced, and 9 shots, 5m spaced), have been contaminated with Gaussian noise of 1ms standard deviation and 0

mean. The synthetic traveltimes have been obtained using the MSFMM with the cross-neighbors as forward modeling (Fig.70a). The response obtained is slightly different from the one obtained without noise (Fig.61b) due to the low noise level, but individuates the presence of the two layers and of the step in the central part of the section. The separated inversion reached the convergence after 6 iterations with the stopping criterion of $\chi^2=1$ and was conducted using the value 954 for lambda parameter, chosen through the L-curve (Fig.69).

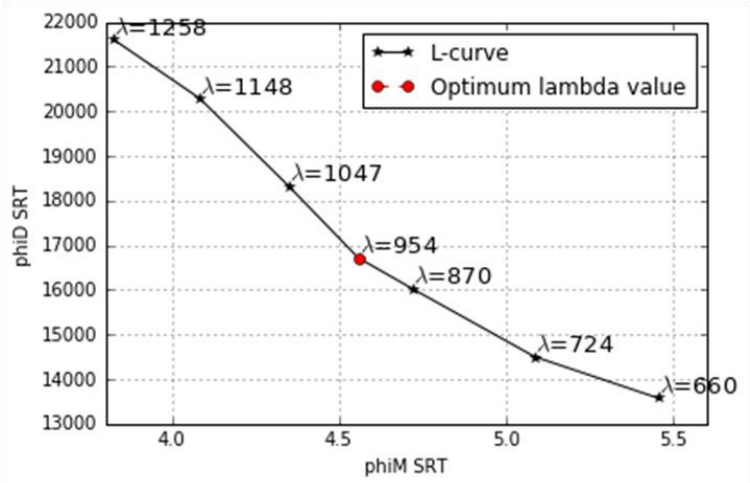

Figure 69: L-curve for SRT data. The red point indicates the optimum lambda value.

The section of the separated inversion (Fig.70b) shows that the two layers with different velocities are well individuated, even if the maximum values reached result to be too high than the ones of the model. The step in the middle part of the section is also well located both at the depth of 5 and 2m.

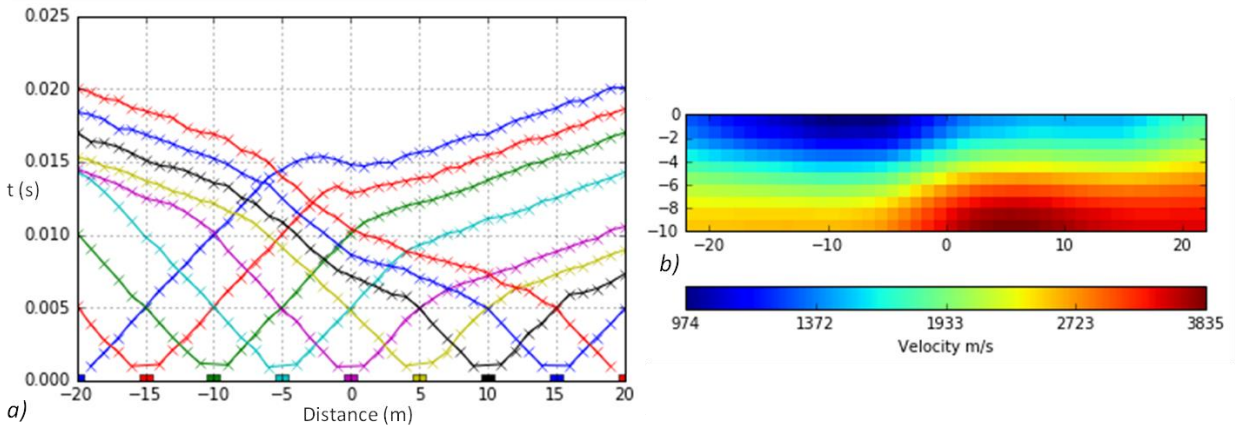

Figure 70: a) Computed travel times responses for the synthetic model with noise; b) section obtained by the separated inversion of synthetic data.

ERT synthetic data have been obtained using the same configuration used in the synthetic model without error and specifically 41 electrodes, 1m spaced using the dipole-dipole configuration. Furthermore, the data were contaminated with 1% of Gaussian noise. As in paragraph 6.1.2 (see Fig.29a), the forward modeling used was the one implemented in the pyBert package (Rücker, 2011; Günther et al., 2006). I show the obtained synthetic response (Fig.71a) and the inverted

model obtained by the separated inversion (Fig.71b) using for lambda value the one obtained with the L-curve criterion described in paragraph 6.1.2 and specifically, 76.

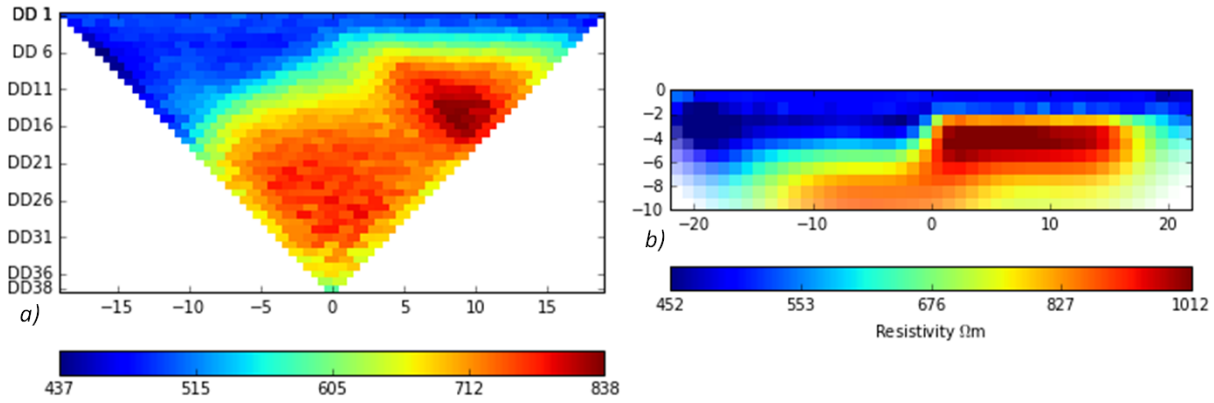

Figure 71: a) Computed resistivity response for the synthetic model with noise; b) section obtained by the separated inversion of synthetic data

The calculated cross-gradients function (Fig.72) shows high values and various areas of structural difference between the two models.

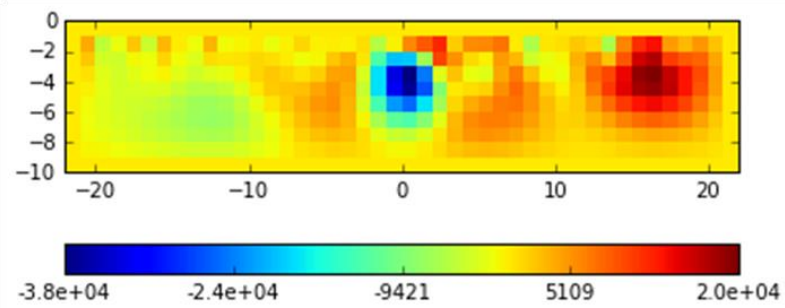

Figure 72: Cross-gradients function for the separated inversion of the noisy data.

In the end, the scatter plot (Fig.73) obtained from the results of the separated inversion shows a great dispersion of values.

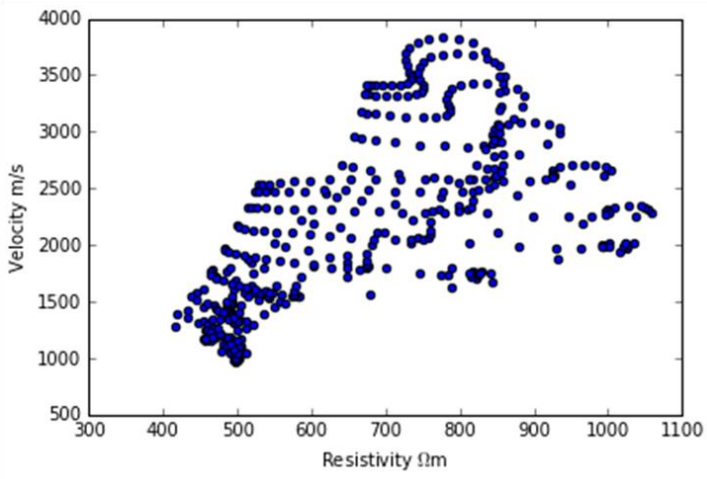

Figure 73: Scatter plot for the separated inversion of data.

Because of the too different results obtained from the first iterations of the SRT separated inversion, the first two iterations of the SRT separated inversion were carried out, until the result became

stable. Then, the obtained model was used to start the *joint inversion*. The value of the λ_{CG} parameter was chosen through the graph of the mean value of the cross-gradients function (Fig.74) and specifically as 1400.

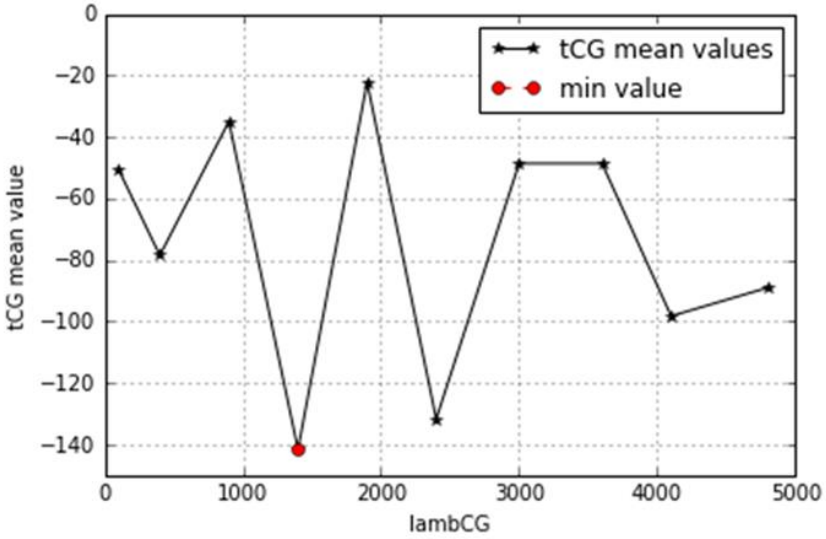

Figure 74: Graph of the mean value of the cross-gradients function for the synthetic model with noise. The red point represents the optimum lambda for the joint inversion

The inverted models obtained by the joint inversion (Fig.75) after 5 iterations (the convergence is obtained following the stop criterion of $\chi^2=1$) show that the range of values for the ERT section (Fig.75a) is almost the same of the separated inversion, while for the SRT inverted model (Fig.75b) it is smaller and closer to the starting model. Furthermore, the SRT section seems to individuate the step in the middle part of the section in a better way than the one of the separated inversion.

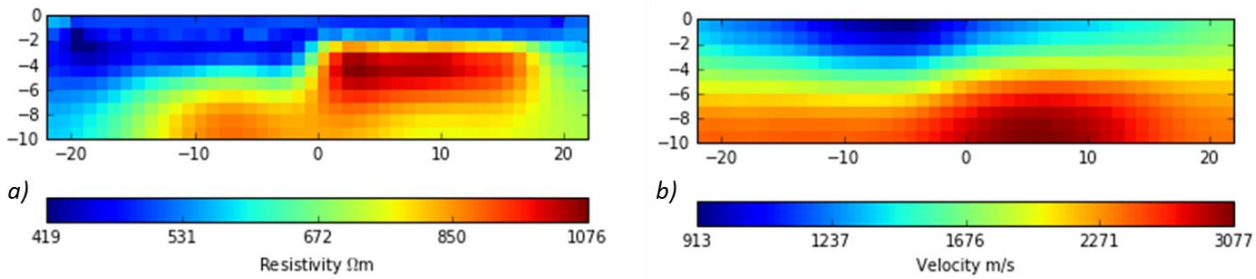

Figure 75: Joint inversion sections for a) ERT and b) SRT.

The cross-gradients function (Fig.76) shows values smaller than the ones of the separated inversion and consequently, areas with a better structural similarity.

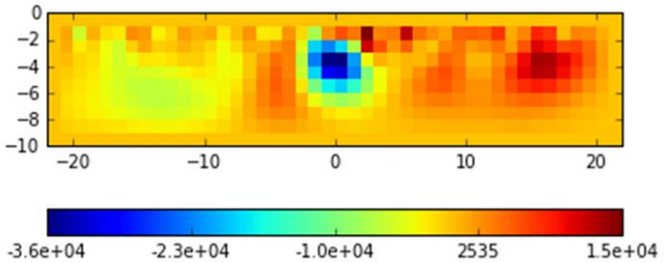

Figure 76: Cross-gradients function for the joint inversion of the noisy data.

The scatter plot (Fig.77) shows a less dispersion of data due to the joint inversion and consequently, a better individuation of the geology of the examinated subsoil.

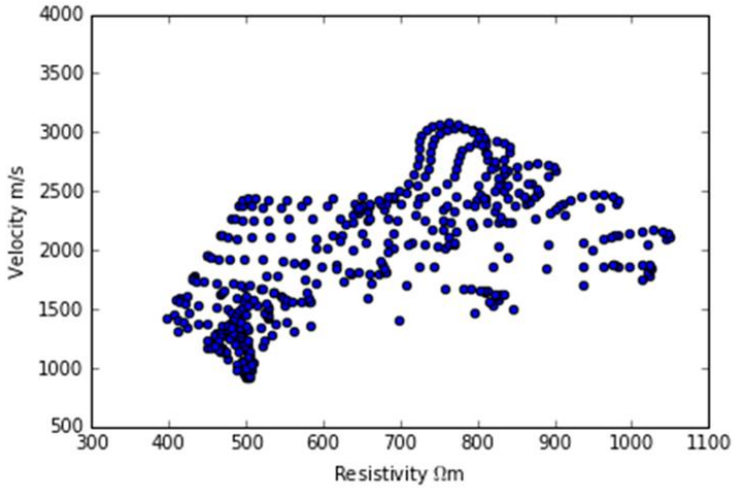

Figure 77: Scatter plot for the joint inversion of data.

6.2.5. Separated and joint inversion for field data

The MSFMM with cross-neighbors has been used as forward modeling not only for the synthetic data described in the previous paragraphs, but also for field data acquired at Rieti, described in paragraph 6.1.3 and labeled as Line 3.

The separated inversion for both geophysical data was carried out using the pyGIMLi package (Rücker et al., 2017). The optimum lambda values have been chosen for both ERT and SRT data through the L-curve. Specifically, for ERT data, the optimum lambda value results to be 164 (see paragraph 6.1.3), while for SRT data, the value 354 (Fig.78).

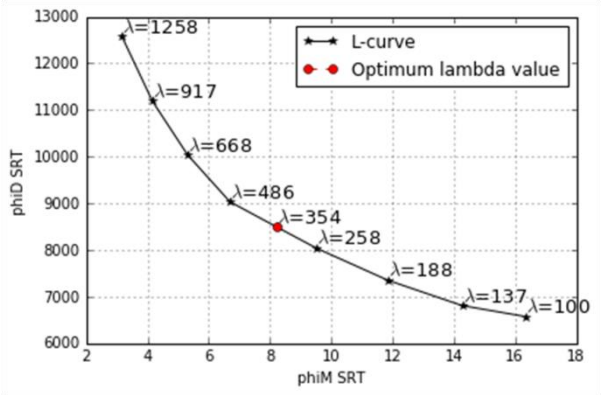

Figure 78: L-curve for SRT field data. The red point highlights the optimum lambda value.

In the following lines, I will discuss only the results for the seismic separated inversion, referring to paragraph 6.1.3 for the results obtained from the ERT one.

For the separated seismic inversion, the convergence was reached after 3 iterations using the stop criterion of $\chi^2=1$. The result (Fig.79) confirms the results obtained from the SRT inversion obtained using the FMM as forward modeling, even if with some differences.

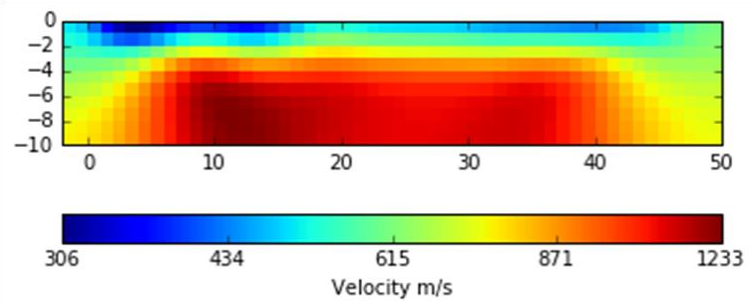

Figure 79: SRT inverted model obtained from the separated inversion of field data using the MSFMM as forward modeling.

In fact, from the comparison between the inversion sections obtained using the FMM and the MSFMM as forward modeling (Fig.80), it emerges that both locate a first layer with low P-wave, with velocities in the range of 300-600m/s. However, the presence of a layer with high velocities (>1000m/s) is detected in both the sections, even if in the one obtained using the MSFMM (Fig.80b) it results to be not only in the right part of section, but also slightly moved to the righ and wider than the one obtained with the FMM (Fig.80a).

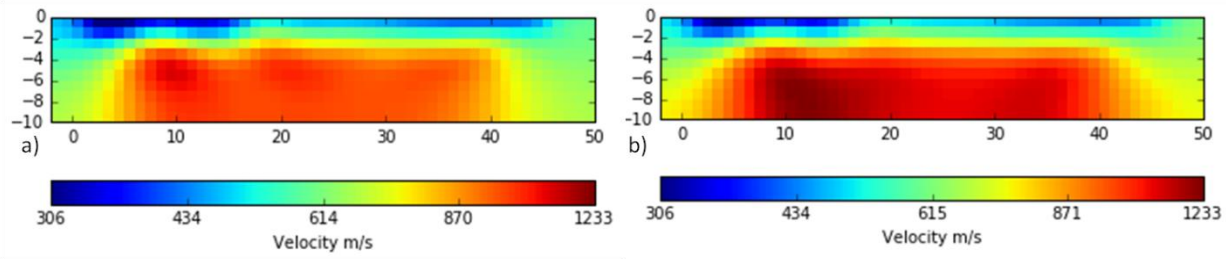

Figure 80: Comparison between the inversion results obtained using a) the FMM and b) the MSFMM as forward modelings.

The cross-gradients function (Fig.81), obtained using the results of the ERT and seismic separated inversions, shows values that are smaller than the ones obtained using the FMM as forward modeling for the seismic inversion and consequently a higher structural similarity between the two models already in the separated inversion. We can conclude that in this case, the different forward modeling influences the structural similarity between the two models.

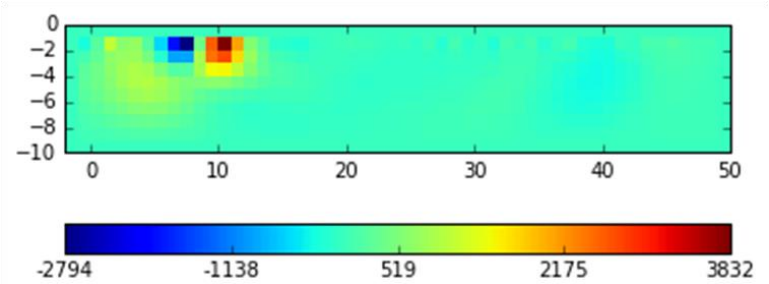

Figure 81: Cross-gradients function obtained from the data of the separate inversion

In the end, the scatter-plot (Fig.82) shows a dispersion of data that is similar to the one obtained using the FMM as forward modeling in the seismic inversion.

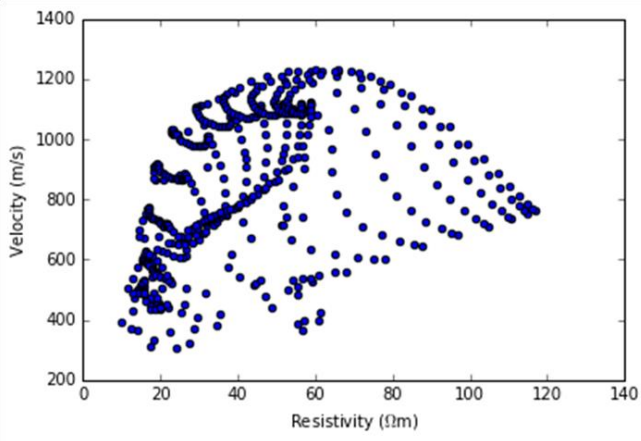

Figure 82: Scatter plot of the separated inverted data.

Regarding the ***joint inversion***, it has been performed using the same lambda parameters of the separated inversion, while λ_{CG}, the parameter weighting the cross-gradients stabilizer, has been chosen after few trials and specifically selecting 8700, the minimum mean value of the cross-gradients vector (as in Fig.83).

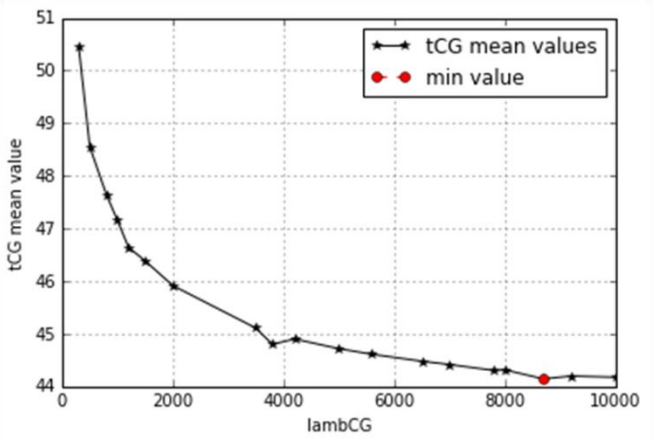

Figure 83: Graph of the mean value of the cross-gradients function for the synthetic modelwithout noise. The red point represents the optimum lambda for the joint inversion

The convergence was reached after 4 iterations, using also in this case the stop criterion of $\chi^2=1$. The inverted models of the joint inversion for ERT and SRT (Fig.84) show: for ERT the same result of the separated inversion, while a slightly different result for SRT.

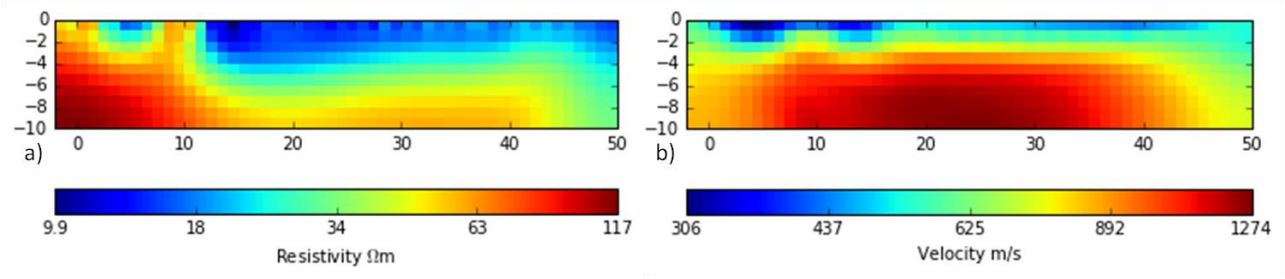

Figure 84: Joint inverted models for a) ERT and b) SRT.

In fact, in the SRT section (Fig.84b), the range of values is similar to the one of the separated inversion. Furthermore, the first layer, with low velocity values, is well identified, while the layer with high velocity values, which can represent a coarse middle layer, seems to have the higher values in the center of the section and specifically between 15 and 30m. This result is slightly different from the one of the separated inversion, in which the higher velocities were concentrated between 10 and 20m (Fig.79). However, the the cross-gradients function and the cross-plot show the goodness of the joint inversion. In fact, the obtained results for the cross-gradients function (Fig.85) are greatly reduced than the ones obtained from the separated inversion.

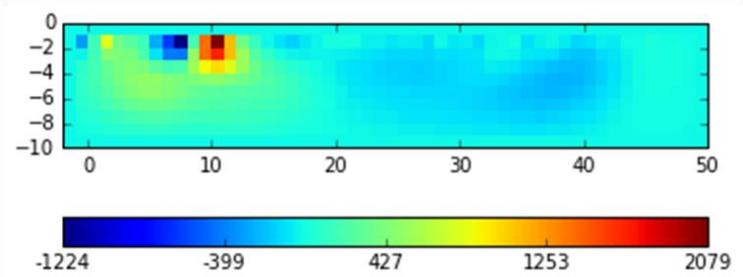

Figure 85: Cross-gradients function obtained from the data of the joint inversion

In the end, the dispersion of data is reduced in the scatter-plot for the joint inversion (Fig.86).

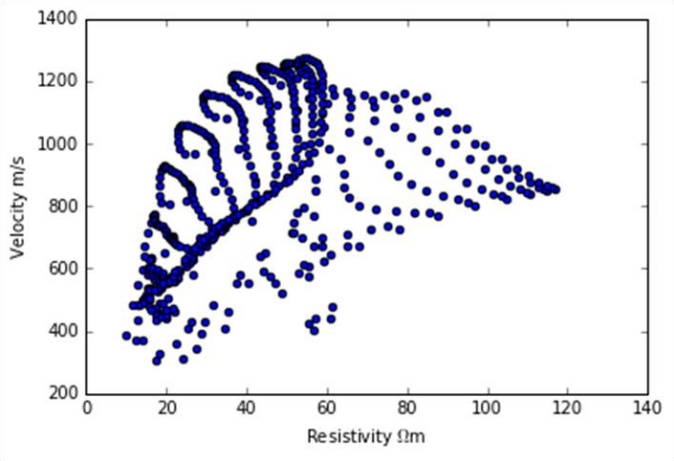

Figure 86: Scatter plot of the joint inverted data

Conclusions

In this thesis, the structural joint inversion of two geophysical methods has been performedon synthetic models and field data. The results support the use of the joint inversionto improve the resolution and reduce the interpretation ambiguities in the inverted models improving the structural similarity compared to the results of the separated inversion. The joint inversion with the use of the cross-gradients function provides a link between two different geophysical models without imposing a relationship between the measured parameters, which is helpful in the near surface, where these relationships are hard to find. In addition, the interesting element of the cross-gradients function is that the similarity is not forced where the changes are not present. It can happen that no variation in the first property is associated to a variation in the second one. In this case, the structural similarity is not artificially created in the section where the property's variation is not present. In conclusion, the approach of the cross-gradients function provides similar results only where it is possible to establish the similarity.

Furthermore, the MSFMM as forward modeling for the seismic method, and specifically the use of the information that derives from the diagonal directions, improves the inverted models in both the separated and joint inversions. In fact, especially in the synthetic models, the different layers of the subsoil are better individuated and the resolution is improved compared to the results obtained using the FMM as forward modeling.

However, some improvements are needed in the joint inversion especially when the number of iterations is different for the separated inversion of the two geophysical methods. In particular, the inversion process can be modified weighting the two methods in a different way and considering also the rate of convergence of each inversion process.

Another improvement can be applied to the MSFMM. Specifically, the algorithm can be implemented in another programming language, faster and more efficient, as C++ and integrated in the pyGIMLi library.

In the end, since through the joint inversion, a more reliable model can potentially be obtained, the next step might be the reconstruction of saturation or porosity sections.

References

Annan P., 2005. *GPR methods for hydrogeological studies*, in Hydrogeophysics, edited by Y. Rubin and S. S. Hubbard, chap. 7, 532 pp., Springer, New York.

Archie G. E., 1942. *The electrical resistivity log as an aid in determining some reservoir characteristic*, Pet. Trans. AIME, 142, 54–62.

Bachrach R. and A. Nur, 1998. *High-resolution shallow seismic experiments in sand. Part 1: Water table, fluid flow, and saturation*, Geophysics, 63(4), 1225 – 1233.

Bennington N. L., Zhang H., Thurber C. H. and Bedrosian P. A., 2015. *Joint inversion of seismic and magnetotelluric data in the Parkfiled region of California using the normalized cross-gradient constraint*, Pure appl. Geophys., 172(5), 1033–1052.

Binley A., P. Winship, L. J. West, M. Pokar, and R. Middleton, R., 2002. *Seasonal variation of moisture content in unsaturated sandstone inferred from borehole radar and resistivity profiles*, J. Hydrology, 267, 160–172.

Binley A. and A. Kemna, 2005. *Electrical methods*, in Hydrogeophysics, edited by Y. Rubin and S. S. Hubbard, pp. 129– 156, Springer, New York.

Binley A., S. S. Hubbard, J. A. Huisman, A. Revil, D. A. Robinson, K. Singha, and L. D. Slater, 2015. *The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales*, Water Resour. Res., 51, 3837–3866, doi:10.1002/2015WR017016.

Bronstein A. M., Bronstein M. M. and R. Kimmel, 2008. *Numerical Geometry of Non-Rigid Shapes*, Springer.

Cercato M. and G. De Donno, 2018. *Focusing on soil-foundation heterogeneity through high-resolution electrical and seismic tomography*, Near Surface Geophysics 16 (1), 1-12

Coggon J. H., 1971. *Electromagnetic and electrical modelling by the finite element method*, Geophysics, v. 36. p. 132-155.

Cormen H. T., C. E Leiserson., R. L. Rivest and C. Stein, 2001. *Introduction To Algorithms*, 10.2307/2583667.

Davis J. L. and Annan A. P., 1989. *Ground penetrating radar for high-resolution mapping of soil and rock stratigraphy*, Geophysical Prospecting 37, 531-551.

deGroot-Hedlin C. and Constable S. C., 1990. *Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data,* Geophysics, 55, 1613-1624.

Demirci I.˙, M. E. Candansayar, A. Vafidis, P. Soupios, 2017. *Two dimensional joint inversion of direct current resistivity, radio-magnetotelluric and seismic refraction data: an application from Bafra Plain, Turkey,* J Appl Geophys 139:316–330

Demirci I., Ü. Dikmen and M. E. Candansayar, 2018. *Two-Dimensional joint inversion of Magnetotelluric and local earthquake data: Discussion on the contribution to the solution of deep subsurface structures*. Physics of the Earth and Planetary Interiors. 275. 10.1016/j.pepi.2018.01.006.

Dey A. and Morrison H. F., 1979. *Resistivity modeling for arbitrarily shaped three-dimensional structures*, Geophysics, 44, 615–632.

Doetsch J., N. Linde, I. Coscia, S. A. Greenhalgh, and A. G. Green, 2010. *Zonation for 3D aquifer characterization based on joint inversions of multimethodcrosshole geophysical data*, Geophysics, 75, no. 6, G53–G64, doi: 10.1190/1.3496476.

Domenico P.A. and F. W. Schwartz, 1990. *Physical and Chemical Hydrogeology*, John Wiley & Sons, New York, 824 p.

Fomel S., 1997. *A variational formulation of the fast marching eikonal solver*, in SEP-95: Stanford Exploration Project, pp. 127–147.

Fregoso E. and L. A. Gallardo, 2009. *Crossgradients joint 3D inversion with applications to gravity and magnetic data*, Geophysics, 74, no. 4, L31–L42, doi: 10.1190/1.3119263.

French H. K., C. Hardbattle, A. Binley, P. Winship., L. Jakobsen, 2002. *Monitoring snowmelt induced unsaturated flow and transport using electrical resistivity tomography,* J. Hydrol. 267, 273–284.

Gallardo L. A., 2007. *Multiple cross-gradient joint inversion for geospectral imaging*, Geophysical Research Letters, 34, L19301, doi: 10.1029/ 2007GL030409.

Gallardo L. A. and M. A. Meju, 2003. *Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data*, Geophysical Research Letters, 30, 1658, doi: 10.1029/2003GL017370.

Gallardo L. A. and M. A. Meju, 2004. *Joint two-dimensional DC resistivity and seismic travel time inversion with cross gradients constraints*, J. Geophys. Res., 109, B03311, doi:10.1029/2003JB002716.

Gallardo L. A. and M. A. Meju, 2007. *Joint 2D cross-gradient imaging of magnetotelluric and seismic travel-time data for structural and lithological classification*, Geophysical Journal International, Volume 169, Issue 3, 1 June, Pages 1261–1272,https://doi.org/10.1111/j.1365-246X.2007.03366.x.

Grote K., S. S. Hubbard and Y. Rubin, 2003. *Field-scale estimation of volumetric water content using GPR groundwave techniques*, Wat. Resour. Res. 39(11), 1321, 10.1029/2003WR002045.

Günther T., 2004. *Inversion Methods and Resolution Analysis for the 2D/3D Reconstruction of Resistivity Structures from DC Measurements*, PhD thesis, Freiberg University of Mining and Technology, available from http://fridolin.tu-freiberg.de.

Günther T. and C. Rücker, 2006. *A new joint inversion approach applied to the combined tomography of dc resistivity and seismic refraction data.* In: 19th EEGS symposium on the application of geophysics to engineering and environmental problems.

Günther T., C. Rücker, K. Spitzer, 2006. *Three-dimensional modelling and inversion of DC resistivity data incorporating topography - II. Inversion*, Geophys. J. Int. 166 (2), 506–517.

Haber E. and D. W. Oldenburg, 1997. *Joint inversion: a structural approach*, Inverse Prob 13(1):63–77

Hamdan A. H. and A. Vafidis, 2012. *Joint inversion of 2D resistivity and seismic travel time data to image saltwater intrusion over karstic areas*. Environmental earth sciences. 68. 10.1007/s12665-012-1875-9.

Hassouna M. S. and A. A. Farag, 2007. *Multistencil fast marching methods: A highly accurate solution to the Eikonal equation on Cartesian domains,* IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(9) 1-12.

Hestenes M. R. and E. Stiefel, 1952. *Methods of Conjugate Gradients for Solving Linear Systems.* Journal of Research of the National Bureau of Standards, 49, 409-435.

Hu W., A. Abubakar, T. M. Habashy, 2009. *Joint electromagnetic and seismic inversion using structural constraints*, Geophysics, 74, R99–R109.

Kim C. R., 2001. *A physical model experiment on the hydrogeologic applications of GPR*, Ph.D. Dissertation, The Ohio State University.

Kimmel R. and J. A. Sethian, 1998. *Computing geodesic paths on manifolds*, Proc. Natl. Acad. Sci. USA, 95(15), 8431–8435.

Kirsch R., 2006. Ed. Groundwater Geophysics, A Tool for Hydrogeology. New York: Springer-Verlag.

Kroon D. J., 2011. *Accurate fast marching*, MATLAB Central File Exchange.

Jin J., 2002. *The Finite Element Method* in Electromagnetics (2nd ed.) Wiley–IEEE Press.

Lelièvre P. G., C. G. Farquharson and C. A. Hurich, 2011. *Computing first-arrival seismic traveltimes on unstructured 3d tetrahedral grids using the fast marching method*, Geophys. J. Int., 184, 885–896.

Lelièvre P. G. and D. W. Oldenburg, 2009. *A comprehensive study of including structural orientation information in geophysical inversions*, Geophys. J. Int., 178, 623–637, doi:10.1111/j.1365-246X.2009.04188.x.

Lesmes D. P. and S. P. Friedman, 2005. *Relationships between the electrical and hydrogeological properties of rocks and soils*, in Hydrogeophysics, edited by Y. Rubin and S. S. Hubbard, chap. 4, 532 pp., Springer, New York.

Linde N., A. Binley, A. Tryggvason, L. B. Pedersen, and A. Revil, 2006. *Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltime data*, Water Resour. Res., 42, W12404, doi:10.1029/2006WR005131.

Linde N., A. Tryggvason, J. E. Peterson and S. S. Hubbard, 2008. *Joint inversion of crosshole radar and seismic traveltimes acquired at the South Oyster Bacterial Transport Site.* Geophysics, 73, no. 4, G29–G37.

Linde N. and J. A. Doetsch, 2010. *Joint inversion of crosshole GPR and seismic traveltime data,* In R. D. Miller, J. H. Bradford, and K. Holliger, eds., Advances in nearsurface seismology and ground-penetrating Radar: SEG Geophysical Developments Series 15, 1–16.

Linde N. and J. Doetsch, 2016. *Joint inversion in hydrogeophysics and near-surface geophysics*. In: Integrated imaging of the Earth, M. Moorkamp, P. Lelievre, N. Linde, and A. Khan (Editors), ch.7, 119-135, John Wiley & Sons, Inc. Hoboken, New Jersey. Doi: 10.1002/9781118929063.ch7.

Meju M. A. and L. A. Gallardo, 2016. *Structural coupling approaches in integrated geophysical imaging.* In: Moorkamp M., Lelièvre P. G., Linde N., Khan A. (eds) Integrated imaging of the earth: theory and applications. Wiley, pp 49–67. ISBN 9781118929063. doi: 10.1002/9781118929063.ch4.

Menke W., 1989. *Geophysical Data Analysis: discrete Inverse Theory*, rev. edn, Academic Press, San Diego.

Moorkamp M., B. Heincke, M. Jegen, A. W. Roberts, R. W. Hobbs, 2011. *A framework for3-D joint inversion of MT, gravity and seismic refraction data*. Geophys. J. Int. 184,477–493

Moorkamp M., P. G. Lelivre, N. Linde, A. Khan, 2016. *Introduction.* In: Moorkamp M, Lelie`vre PG, Linde N, Khan A (eds) Integrated imaging of the earth: theory and applications. Wiley, pp 1–6. ISBN9781118929063. doi: 10.1002/9781118929063.ch1.

Moorkamp M., 2017. *Integrating electromagnetic data with other geophysical observations for enhanced imaging of the earth: a tutorial and review*. SurvGeophys. doi:10.1007/s10712-017-9413-7.

Oldenburg D. W. and Y. Li, 2005. *Inversion for applied geophysics: A tutorial*, Investigations in Geophysics, 13, 89–150.

Pak Y. C., T. Li, G. S. Kim, 2017. *2D data-space cross-gradient joint inversion of MT, gravity and magnetic data*, J. Appl. Geophys.,143, pp.212-222

Paprocki L., 2000. *Characterization of vadose zone in-situ moisture content and an advancing wetting front using crossborehole ground penetrating radar,* Masters Thesis, New Mexico Institute of Mining and Technology.

Parra J. O., C. L. Hackert and M. W. Bennett, 2006. *Permeability and porosity images based on P-wave surface seismic data: Application to a south Florida aquifer*, Water Resour. Res., 42, W02415, doi:10.1029/2005WR004114.

Pilkington M., 1997. *3-D magnetic imaging using conjugate gradients*, Geophysics, 62, 1132–1142

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(1), 271–284.

Pridmore D. F., G. W Hohmann., S. H. Ward, and W. R. Sill, 1981. *An investigation of finite-element modeling for electrical and electromagnetic data in three dimensions*, Geophysics, 46, 1009-1024.

Rabbel W., 2006. *Seismic methods*. In: Groundwater Geophysics. A Tool for Hydrogeology (ed. R. Kirsch), pp. 23–83. Springer.

Rawlinson N. and M. Sambridge, 2004. *Wave front evolution in strongly heterogeneous layered media using the fast marching method*, Geophys. J. Int., 156(3), 631–647.

Raymer L. L., Hunt E. R., Gardner J. S., 1980. *An improved sonic transmit time - porosity transform*. Trans. SPWLA, 21st Ann Log Symp:1-13

Rickett J. and S. Fomel, 1999. *A second order fast marching eikonal solver*, in Stanford Exploration Project, Report 100, pp 287-293.

Rubin Y., G. Mavko, and J. Harris, 1992. *Mapping permeability in heterogeneous aquifers using hydrologic and seismic data*, Water Resour. Res., 28(7), 1809–1816.

Rubin Y. and S. S. Hubbard, 2005. *Introduction to hydrogeophysics*. In: Rubin Y, Hubbard SS (eds) Hydrogeophysics, Chap. 1. Springer, Dordrecht, The Netherlands, pp 3–21

Rücker C., 2011. *Advanced Electrical Resistivity Modelling and Inversion using Unstructured Discretization*. PhD thesis, University of Leipzig.

Rücker C., T. Günther, F. M. Wagner, 2017. *pyGIMLi: An open-source library for modelling and inversion in geophysics*, Computers and Geosciences, 109, 106-123, doi: 10.1016/j.cageo.2017.07.011

Sasaki Y., 1994. *3-D resistivity inversion using the finite-element method*. Geophysics, **59**, 1839-1848.

Scales J., 1987. *Tomographic inversion via the Conjugate Gradient method*. Geophysics. 52. 179-185. 10.1190/1.1442293.

Sen M. and P. Stoffa, 2013. *Global Optimization Methods in Geophysical Inversion*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511997570.

Sethian J. A., 1999a. *Fast marching methods*, SIAMRev., Vol. 41, No. 2, pp. 199–235.

Sethian J. A., 1999b. *Level Set Methods and Fast Marching Methods*, second ed. Cambridge Univ. Press.

Sethian J. A. and A. M. Popovici, 1999. *3-D traveltime Computation using the fast marching method,* Geophysics, vol.64, no.2, pg.516-523.

Sethian J. A. and A. Vladimirsky, 2000. *Fast methods for the eikonal and related Hamilton–Jacobi equations on unstructured meshes*, Proc. Natl. Acad. Sci. USA, 97, pp. 5699–5703.

Shi Z., R. W. Hobbs, M. Moorkamp, G. Tian, L. Jiang, 2017. *3-D cross-gradient joint inversion of seismic refraction and DC resistivity data.* J. Appl. Geophys. 141:54–67. doi.org/10.1016/j.jappgeo.2017.04.008.

Stummer P., H. Maurer, H. Horstmeyer and A. G. Green, 2002. *Optimization of DC resistivity data acquisition: real-time experimental design and a new multielectrode system.* IEEE Transactions on Geoscience and Remote Sensing,40.

Topp G. C., J. L. Davis, and A. P. Annan, 1980. *Electromagnetic determination of soil water content: Measurements in coaxial transmission lines,* Water Resources Research, 16, 574-582.

Tryggvason A. and N. Linde, 2006. *Local earthquake (LE) tomography with joint inversion for P- and S-wave velocities using structural constraints*, Geophysical Research Letters, 33, L07303, doi: 10.1029/2005GL025485.

VanDecar J. C. and R. Snieder, 1994. *Obtaining smooth solutions to large, linear, inverse problems*, Geophysics, 59, 818- 829.

Wang Z., 2001. *Fundamentals of seismic rock physics*. Geophysics;66: 398–412.

Wang Q., S. Ji, S. Sun and D. Marcotte, 2009. *Correlations between compressional and shear wave velocities and corresponding Poisson's ratios for some common rocks and sulfide ores.*, Tectonophysics, 469, 61–72.

Wang K., H. Tan and T. Wang, 2017. *2D joint inversion of CSAMT and magnetic data based on cross-gradient theory.* Applied Geophysics, **14**(2), pp.279–290.

Wharton R. P., G. A. Hazen, R. N. Rau and D. L. Best, 1980. *Electromagnetic propagation logging: Advances in technique and interpretation*, Soc. Pet. Eng., Pap. No. 9267.

Wyllie M. R. J., A. R. Gregory and L. W. Gardner, 1956. *Elastic wave velocities in heterogenous and porous media*. Geophysics 21:41-70

Yilmaz Ö., 2001. *Seismic Data Analysis*. Soc. Explor. Geophys., Tulsa.

Zhdanov M.S., 2015. *Inverse Theory and Applications in Geoophysics*, Elsevier. Published.