Università degli Studi di Cagliari

PHD DEGREE

SCIENZE E TECNOLOGIE DELLA TERRA E DELL'AMBIENTE

Cycle XXXI

TITLE OF THE PHD THESIS

Potential aquifer characterization through the integration of multivariate

geophysical datasets

Scientific Disciplinary Sector

Area 04 - Applied geophysics GEO/11

PhD Student: Stefano Bernardinetti

Coordinator of the PhD Programme   Prof. Aldo Muntoni

Tutor   Prof. Antonio Funedda
Co-tutor   Prof. Pier Paolo Bruno
                              Prof. Luigi Carmignani
                              Dr. Tommaso Colonna

To my wife Angela and my daughter Sofia,

for their support and strength during my absences.

Acknowledgements

Table of contents

6

represent low values and warm colors represent high values. In the left side of the map are classified observations with high dip (toward right reflectors), low to medium similarity, low energy and medium to low entropy. Maximum values of energy and similarity with low entropy and sub-horizontal reflectors are classified in the upper right part of the map.128

1 Abstract

The work developed in this thesis has the aim to provide a methodology for geophysical exploration and characterization of an aquifer system. Two of the most common geophysical techniques, electrical resistivity tomography and seismic P-waves refraction/reflection methods were applied in order to fully explore the subsurface.

Thanks to their sensitivity to factors as mineralogy, porosity, lithology, texture and pore content, electrical resistivity and acoustic waves can depict main subsurface changes connected with an aquifer system. Anyway, the use of a single geophysical parameter leads to an incomplete imaging with a not negligible level of uncertainty. The integrated use of both geophysical parameters (i.e. electrical resistivity and P-waves velocity), allowed to obtain complementary informations about the subsurface conditions, due to their different sensitivity. More in specific, were proposed two different approaches to explore and characterize the shallow part and the deep of the subsurface, because of to study deeper systems the loss in resolution is one of the drawbacks. In both cases were proposed two post-processing integration methods, where the shallower is based on the integration of electrical resistivity and P-waves seismic velocity from tomographic techniques; the deeper based on the multivariate analysis of a dataset of seismic attributes obtained by high resolution seismic reflection technique.

The results are: 1) a shallow unique model characterizing the main features of the subsurface with homogeneous geophysical facies corresponding to different hydrogeological conditions; 2) a high-resolution large scale unique model depicting main structural and stratigraphic changes meaningful in aquifer characterization.

The two approaches are based on machine learning clustering techniques, where the k-means with Silhouette index were applied to find out the optimal number of geophysical facies from tomographic datasets (i.e. electrical resistivity and P-waves velocity) and Self-Organizing maps (or Kohonen maps) were used to highlight similar pattern in seismic attributes multivariate datasets.

The integration methods were applied in two very different geological and hydrogeological environments, for the characterization of an aquifer in an intra-Apennine basin in Tuscany (Italy) and for the characterization of a complex hydrothermal system into the Solfatara volcano (Naples, Italy). Also the seismic data acquisition were carried out through two different sources, an 8-caliper seismic gun and the higher resolution vibrational source IVI Minivib. This choice was taken with the aim to demonstrate the consistency of the integration methods proposed, varying geological and hydrogeological settings.

In the case study of Laterina, high-resolution electrical resistivity tomography and seismic reflection/refraction surveys were carried out with the aim to characterize shallow unconfined aquifers in alluvial deposits and a deeper confined aquifer in rock. The tomographic integration revealed six geophysical facies consistent with stratigraphic information provided by a borehole

placed near the survey line. Moreover, allowed to infer the setting of main aquifers and aquitards and estimate the saturated from unsaturated zone for the shallower aquifer. The integration of seismic attributes provided a large scale image of the subsurface highlighting shallow alluvial recent and Villafranchian fluvio-lacustrine deposits above the Monte Cervarola Unit, mainly made up of sandstone/marls with a medium to high degree of fracturation.

The integration processes were applied at the case of Solfatara volcano with the purpose of: 1) highlight the shallower fluids pattern (i.e. gas and liquid phase) connected with the Fangaia mud pool and CO_2 flux, and 2) the main structural paths of uprising for fluids and deep reservoirs within the caldera. The tomographic images provided three different geophysical facies, one with low velocity low resistivity consistent with poorly consolidated sediments, and two high velocity with low and high resistivity, respectively associable with liquid dominated and gas dominated porous media. The integrated model obtained by Self-Organizing maps provided a meaningful image of the caldera highlighting main sub vertical discontinuities interpreted as fracture or ring fault, main paths for subsoil fluids circulation. Moreover, several sub-horizontal bright spot within the caldera were well depicted, consistent with vapor saturated media.

2 Introduction

The exploration of a hydrogeological complex system passes through a multiscale multitemporal study, from the basin to the site of interest scale, with geological, hydrogeological and geophysical surveys. The latter has the important role of providing an image of the subsoil and spatializing sparse informations from geological and hydrogeological fieldworks.

The need to obtain a detailed hydrogeophysical characterization of the subsurface and its interpretation, for fluids and groundwater resources management, is hard to tackle. As known, different deposits and rocks show different geophysical responses, due to mineral type, grain packing arrangement, porosity, permeability, and pore content (i.e. gas or fluid type). The use of several geophysical methods allows in theory to obtain complementary informations about the subsurface conditions, due to the different sensitivity of the geophysical methods to different parameters, such as P- and and S-wave velocities, bulk electrical conductivity, density, etcetera. A knowledge of multiple parameters co-located can reveal hydrogeological features in well constrained way thanks to the sensitivity of different geophysical techniques in subsurface variations.

The goal of this thesis is to provide a unique geological model of the aquifer by synthesizing and optimizing the information provided by several geophysical methods. This approach greatly reduces the degree of uncertainty and subjectivity of the interpretation by exploiting the different physical and mechanic characteristics of the aquifer. In general, to study deep systems the loss in resolution with the increasing of surface space, lead to different strategies to explore the shallow and deep part of a potential aquifer.

In this thesis I developed two approaches: the first approach has the aim to characterize the shallow part of an aquifer, through the integration of electrical resistivty and P-waves seismic refraction tomograms; the second approach has the aim to provide geological structural informations on the deeper part of the survey area, thanks to the integration of some key-attributes provided by high resolution seismic reflection surveys.

The geophysical parameters used in the tomographic post-processing integration i.e. electrical resistivity and P-wave velocity) are sensitive to key hydrological factors and can be helpful to discriminate aquifers from aquitards. Moreover, the integration of seismic attributes can provide a large-scale model that highlights main geological/hydrogeological features.

One of the strength of this work is the application of integration methods in two very different geological and hydrogeological environments, an alluvial valley in an intra-appennine basin in Tuscany (Italy) and a volcano caldera in the Campi Flegrei Area (Napels, Italy). Also the seismic data acquisition were carried out through two different sources, an 8-caliper seismic gun and the higher resolution vibrational source IVI Minivib. This choice

was taken with the aim to demonstrate the consistency of the integration methods proposed, varying geological and hydrogeological settings.

In Chapter 3 the geophysical techniques and attributes used for aquifer characterization are reviewed. A specific review of parameters which influence resistivity and velocity, in order to point out the capability of different geophysical parameters to reveal subsurface features, mainly as porosity, fracturation and saturation. Furthermore, basic theories about electrical resistivity and seismic refraction methods are explaned. Main steps of processing flows to obtain the electrical resistvity and seismic refraction tomography are exposed, as an introduction to the use of seismic attributes for reservoir charcterization.

Chapter 4 contains a review of geophysical integration methods, mainly based on geostatistical approaches. Then the two proposed integration method developed in this thesis are described: the first based on k-means clustering method in order to reveal geophysical homogeneous facies, the second based upon the Kohonen maps. For both approaches data processing and restitution are explained, with an interpretation key of the results. In particular for the tomographic integration the k-means and Silhouette index and their use on Matlab are exposed. The Self-Organizing maps or Kohonen maps, applied as integration method for the higher dimension multivariate dataset integration, provided by seismic reflection (i.e. seismic attributes), are exposed in this chapter. Specifically is described their use on the programming code R and procedures developed to produce graphical results.

Chapter 5 contains the application of the two methods developed in this thesis for the characterization of a shallow aquifer at the country of Laterina. The geological setting is characterized by shallow alluvial deposits which constitute the unconfined aquifer. Beneath an aquitard level is present made by clayey Villafranchian fluvial-lacustrine deposits. At the bottom of those deposits, an irregular bedrock of marl/sandstone with different degree of fracturation is present and constitutes the deeper confinated aquifer. The post-processing tomographic integration is performed on the bivariate dataset composed by electrical resistivity and P-waves seismic velocity along a line of about 420 m, giving several geophysical facies characteristics of different geological/hydrogeological settings. Then, the seismic attribute dataset provided by the high resolution seismic acquisition are integrated as explained in the chapter 4 in order to highlight at deeper depths main geological features intended as lithology variations and fractured/faulted zones.

Chapter 6 contains the the application of integration processes at the complex hydrothermal system of Solfatara Volcano. The study area is a dynamic and evolutive active volcano interested mainly by fluids circulation in liquid and gas phases. The post-processing tomographic integration is performed on 2 orthogonal co-located tomograms

provided by electrical resistivity and P-waves seismic refraction tomography, with the aim to distinguish liquid dominated from gas dominated zones. The deeper high-resolution dataset composed by seismic attributes is integrated in order to image a unique model of the subsurface a greater depth to the previous analyses with main structural pattern meaningfull in a hydrogeological optic.

3 Geophysical datasets

Aquifer exploration requires a multidisciplinary, time-lapse and multi-scale study because of water resources are influenced by many parameters, often spatially located far from the target area. Moreover seasons have their influence on groundwater resource and this has a great impact on small hydrogeological basin and shallow aquifers. In this field, geophysical/hydrogeophysical techniques can be applied to reveal geological settings at several scales and main fluid characteristics intended as quality level of the water.

As known, different materials show different geophysical response such as their resistivity, acoustic velocity, magnetic permeability and density. These parameters are influenced by the mineral type, grain packing arrangement, porosity, permeability, pore content (i.e. gas and/or fluid type) and temperature. In most geophysical surveys, therefore, to highlight geological features it is important that the variations in geophysical parameters show large differences with surrounding material. In general, any material is described by ranges of each property, many times overlaid and this increase the difficult to interpretate geophysical results with low uncertainty.

Here the main idea is to characterize the subsoil for shallow and deep target using tomographic data and reflection data, respectively ERT, SRT and seismic attributes from P-waves reflection sections.

The approach to use tomographic and reflection data can give a huge improvement in understanding the subsoil because of the different range of depth. Geoelectrical resistivity method has as main targets the discrimination of fluid content, its quality in terms of salinity, grain size and mineral content (i.e. clay minerals) and fluid circulation in fractured rocks. Seismic refraction tomography has the role of highlighting boundaries between deposits and rocks, dry or saturated conditions in alluvial deposits and weathering of rocks such as fractured or altered zones.

High resolution seismic reflection provides a reliable dataset for seismic attributes calculation. Their use is powerful in discriminating stratigraphical variations, discontinuities, geological structures, which are essential for understanding in fluid circulation and storage on subsurface.

3.1 Overview of geophysical methods for groundwater research

Main efforts in applied geophysics have been made from the oil and gas industries, where thanks to the high value of the resources, have been spent many researches to improve the reliability of exploration techniques. Hydrogeophysics takes advantage of these efforts in a sense, but however, subsoil conditions in terms of pressure, temperature and lithology are far from shallow geological/hydrogeological settings. Moreover, in the shallow part of the subsoil physical variations are more accentuated, geophysical responses can differ and this doesn't allow to apply many petrophysical relationship with geophysical

parameters (Rubin and Hubbard, 2006). Hydrogeophysics is a recent field started in the 1990s and with a constant rate of growth, as showed in Binley et al., 2015, which tries to integrate geophysical methods in hydrogeological studies.

One can use geophysics in two way, the first is use it in a not well-known area, to explore the subsoil with the aim to plan direct surveys (boreholes, etc.), and the second is to project, across geophysical surveys, local and sparse geologic and hydrogeological interpretation from borehole (e.g., Lochbühler et al., 2013, Ruggeri et al., 2013) and outcroppings .

Also the scale of geophysical measurements and their resolutions plays an important role in groundwater research. Another interesting point of view is to differentiate the survey plan in three subsequent temporal phases:

- Large scale surveys to delineate mains geological/hydrogeological features;

- Site scale surveys for a detailed imaging of the subsurface;

- Borehole logs for a punctual hydrogeophysical characterization.

Survey designing is a crucial point to characterize the aquifer and to do this one must know the targets and choose the appropriate geophysical methods to achieve them. For this, we have to take into consideration that the geophysical target has physical properties for the aquifer, which can be distinguished from geological and hydrogeological background. As seen in MacDonald et al., 2001, an incorrect choice of technique has been cited as one of the primary reason for the failure of many geophysical surveys.

One of the strength of the geophysical surveys is to provide spatially extensive information for the subsurface in a minimal invasive way. The target for mapping groundwater can be considered not directly the water but the geological setting in which the water exists, water included. A multitude of geophysical methods have been used effectively in hydrological investigations. Details of such methods can be found in many texts (e.g., Rubin and Hubbard, 2006; Vereecken et al., 2006). The widely applied methods for hydrological purposes are: direct current resistivity, induced polarization, self-potential, ground penetrating radar, electromagnetic induction, seismic refraction, nuclear magnetic resonance, and microgravity. Moreover, in the last years is increasing the application of seismic reflection technique for aquifer exploration and characterization.

One can summarize geophysical techniques for their effectiveness to reveal physical parameters of the subsoil:

- Electrical and electromagnetic methods reliable to highlight porosity, permeability, salinity and water flux.
- Ground penetrating radar for the capability to reveal water content, porosity and stratigraphy.

- Seismic methods have been used to delineate stratigraphy, pore fluid saturation, bedrock and fractured rock systems;
- Seismoelectrics to highlight water content and permeability
- Nuclear magnetic resonance to highlight water content and permeability
- Gravity and magnetics, have been used to map regional aquifers and large scale basin features;

The electrical resistivity tomography in DC can delineate well, vertical and lateral changes in subsoil, due to different characteristics as mineralogy, porosity, fractures and fluid circulation for soil and rocks. Due to the quadripoles size, this method find its best application in shallow aquifers (100-150m depth). DC resistivity tomography has been widely applied for groundwater research. Rai el al, 2013 demonstrate the use of this method in a granite formation to find water-bearing zones. Abdulkadir and Eritrp, 2017, show an application of the DC resistivity imaging over thermal springs. Electrical resistivity can be deployed in surface or in borehole (e.g., Daily et al., 1992; Singha and Gorelick, 2005) with galvanic or capacitive coupling. This allows to have a high flexibility in terms of application and to explore the subsoil at different scales and different resolution: e.g. the use of resistivity in borehole logging, better if coupled with gamma ray and caliper, is a powerful tool to distinguish aquifers from aquitard with extremely high resolution (<1 m independently from the depth).

Self-potential is a passive method which measures electrical fields in surface and borehole. One of the most interesting source of signal is groundwater flow. The relation between SP and water flow has been used in the past only in a qualitative way. Recently have been developed relationship between volumetric fluid flux and volume current density for quantitative use of SP (e.g., Revil et al., 2003, Revil et al., 2006).

Electromagnetic methods allow to explore subsurface in a huge range of scale, from borehole to the airborne passing through surface application. Common applications are characterization of aquifer and saline intrusion (e.g., Yechieli et al., 2001). Differently from ERT, these methods reach greater depth of investigation and are based on measuring electromagnetic fields in frequency or time domain. In a theoretical point of view, the variation of primary electromagnetic field (induced by an alternate current) produces a secondary electromagnetic filed and its decay gives information about subsoil conductivity. Results are in terms of resistivity and are influenced by subsoil characteristics as mineralogy, porosity, fractures and fluid circulation for soil and rocks. Controlled Source Audio Magneto Telluric is a method based on measuring electrical and magnetic field by using a controlled-source to produce 2D and 3D resistivity models. This technique can be applied to the study of aquifers in large scale and for deeper results than ERT (e.g., Blake et al., 2016, Šumanovac et al., 2018).

GPR is a method based on propagation of electromagnetic signal with frequency ranging from about 10 MHz to 1 GHz, from a transmitter to a receiver. Application can be done in surface as reflection survey and in cross-borehole to measure transmitted waves. The most applied configuration for surface surveys is based on a couple of transmitter-receiver with fixed offset: when an electromagnetic wave reaches a contrast in dielectric constant, part of the energy is reflected and measured from by the receiver. However, the use of multi-offset configuration is increasing, in a similar way of seismic reflection survey. Due to the relationship between radar velocity and moisture content, (Topp et al., 1980) GPR is widely applied especially in cross-borehole configuration for groundwater research (e.g., Irving et al., 2009). One of the weakness of this method is the strong attenuation of electromagnetic signal in conductive media, reducing the depth of investigation to few meters in unconsolidated subsoils.

The use of seismic surveys in groundwater exploration have traditionally relied on seismic refraction to map the depth and geometry of bedrock surface underlying unconsolidated and to delineate the P-waves velocity increasing from unsaturated to saturated soil (Wallace, 1970). More complex theory-based approaches exists which derive from the principles of the elastic wave propagation within saturated and unsaturated porous media (Foti et al., 2002). In other cases the hypothetical aquifer is identified by manipulating the ratio Vp/Vs (e.g. Stümpel et al., 1984; Castagna et al., 1985; Nicholson and Simpson, 1985). The high costs for reflection seismic restricted in the past the use for groundwater purpose, but in the last years, there is an increment about the use in shallow subsurface characterization and water research (e.g., Whiteley et al., 1998; Gruber et al., 2003; Francese et al., 2005; Martin et al., 2013). Moreover, new techniques have been developed to exploit the link between seismic and electromagnetic energy (Revil et al., 2014a; Sava et al., 2014). Seismoelectric phenomena in sediments arise from acoustic wave–induced fluid motion in the pore space, which perturbs the electrostatic equilibrium of the electric double layer on the grain surfaces (Block et al., 2006). Interesting results are in Dupuis et al., (2006) where the amplitude of co-seismic seismoelectric signals showed a possible linkage with porosity/permeability.

Differently from other methods, nuclear magnetic resonance permits the direct quantification of hydrogeological properties, as mobile water content (Yaramanci et al., 2002). The proton Magnetic Resonance Sounding (MRS) is a geophysical technique specially designed for hydrogeological applications. It is based on the principle of Nuclear Magnetic Resonance (NMR) and allows the non-invasive detection of free water in the subsurface (Legchenko et al., 2002). Also in this case depth of investigation and resolution depend on its application, either in borehole or in surface, and site conditions (Legchenko et al., 2002), with in general best performances in borehole.

Gravimetry is a method originally developed for exploration and geodesy purposes. Gravimeters are sensitive to differences in gravity caused by local variations in density of

the subsoil. The measurement of temporal variability is the main technique used in gravimetry, which is well-related to the groundwater level variation. The temporal variability of gravimetric results is the key for monitoring groundwater storage changes (e.g., Gehman et al., 2009) and constraining hydraulic parameter estimates (e.g., Blainey et al., 2007). Gravimetric studies for this purpose have been done from local scale (hundreds of meters) (e.g., Gehman et al., 2009, Christiansen et al., 2011) to global scale (Jin et al., 2013).

An important group of geophysical techniques are based on integrated method, but here the focus is only about effectiveness of single method. Any combination appears clearly as an improvement for groundwater research.

3.2 ERT in DC current

ERT in DC current method has been proved as one of the most reliable and effective for groundwater exploration, because of the resistivity is influenced by lithology, pore fluid chemistry and water content. The greater the electrical contrast between the soil matrix and heterogeneity, the easier is the detection.

Electrical methods may be applied in a range of survey configurations. The approach adopted will depend on: 1) the objectives of the survey, 2) the expected spatial variability of electrical properties, 3) access to suitable electrode sites, 4) equipment availability, and 5) data processing capabilities.

In surface applications, ERT can be used for 2D, 3D and multitemporal (e.g. Michot et al. 2003), depending on the spatial resistivity variation, the depth and the resistivity contrast between target and background values. Usually 2D profile are performed by using mutielectrode configuration with automatic measurements of potential difference across couples of electrodes. In this manner one can obtain a profile of resistivity values along vertical and horizontal directions. This is one of the main strength for the ERT method and allows to reveal subsoil changes also in complex environment. 3D applications are often used where targets, as mainly plumes or filtration path, are well defined in a portion of subsoil. Moreover, multitemporal ERT surveys have been proved useful to highlight the changing of water content in the vadose zone.

3.2.1 Theory and basic principles

As explained above, the purpose of electrical resistivity survey is to individuate the spatial resistivity distribution into the soil.

The resistivity of a medium is defined as follows:

$$\rho = R\left(\frac{S}{L}\right) \qquad \text{Eq. 1}$$

21

where R is the electrical resistance, L is the length of a cylinder and S is the area of its cross-section.

The electrical resistance R is defined by the Ohm's law as follows:

$$R = \frac{V}{I} \qquad \text{Eq. 2}$$

where V is potential and I is the current.

In a homogeneous half-space, with point source of current, the current flow is radial from this point and electrical equipotential surfaces are hemispherical.

In this configuration the potential can be expressed through the resistivity and the current:

$$V = \frac{\rho\, I}{2\, \pi\, R} \qquad \text{Eq. 3}$$

In an elementary quadripolar configuration, on homogeneous subsoil and with four electrodes called C1, C2, P1, P2 (Figure 1), the injection of current from C1-C2 dipole, produces a current flow and a correspondent potential field with equipotential surfaces, which can be measured through P1-P2 dipole. The potential difference measurement between P1-P2 dipole can be expressed in analytical way by the equation:

$$\Delta V = \frac{\rho\, I}{2\, \pi} \left[\frac{1}{C1P1} - \frac{1}{C2P1} - \frac{1}{C1P2} + \frac{1}{C2P2} \right] \qquad \text{Eq. 4}$$

where C1P1, C2P1, C1P2 and C2P2 represent geometrical distances among every couple of electrodes C1, C2, P1, P2.

The electrical resistivity can ben then calculated as follows:

$$\rho = \left[\frac{2\pi}{(1/C1P1)-(1/C2P1)-(1/C1P2)+(1/C2P2)} \right] \frac{\Delta V}{I} = K \frac{\Delta V}{I} \qquad \text{Eq. 5}$$

Where K is a geometrical factor that depends on the relative disposition of the four electrodes (quadripolar configuration).
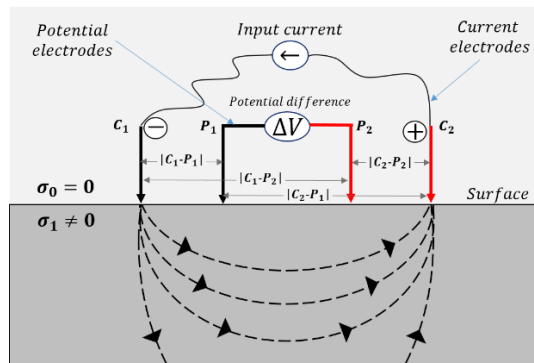
Figure 1 – Elementary quadripolar configuration in homogeneous half-space.

In non-homogeneous subsoil, the current flow distribution depends on the resistivity of the medium, and the current flow is concentrated into the more conductive volume.

Potential difference patterns provide information on the form of subsurface heterogeneities and of their electrical properties (Kearey et al., 2002).

3.2.2 *Factors affecting resistivity*

Subsurface electrical resistivity depends on textural and structural characteristics and is sensitive to the water content (Sheets and Hendrickx, 1995), therefore can be considered as a proxy for the variability of soil physical properties (Banton et al., 1997). More in specific, the electrical resistivity is a function of solid part, mainly grain size and mineralogy, voids intended as pore size distribution and connectivity, fluids content on pore and their resistivity properties and temperature. Also gases have a contribution on resistivity: the air produces a high resistivity because of its insulator property; another gas often presents in subsoil is CO_2 that shows an increasing of electrical conductivity when is dissolved in pore-water or a decrease when is in bubbles (Yang et al., 2015).

Electrical resistivity values for soil, range from few Ωm to several tens thousands Ωm. In terms of electrical conductivity, most of the soils are composed by insulating grain minerals, clayey conductive minerals and in minimum cases of metallic conductive minerals. Electrical conductivity mechanism differs for the above cases: in a prevalence of insulator grain soils, the current flow is mainly governed by electrolytic conductivity and take place through the moisture-filled pores; in case of clayey soils, resistivity depends to electrical charge density at the surface and is related to capacity of cation exchange (Table 2); when massive mineral are present without discontinuities the current flow is due to the free electrons displacement.
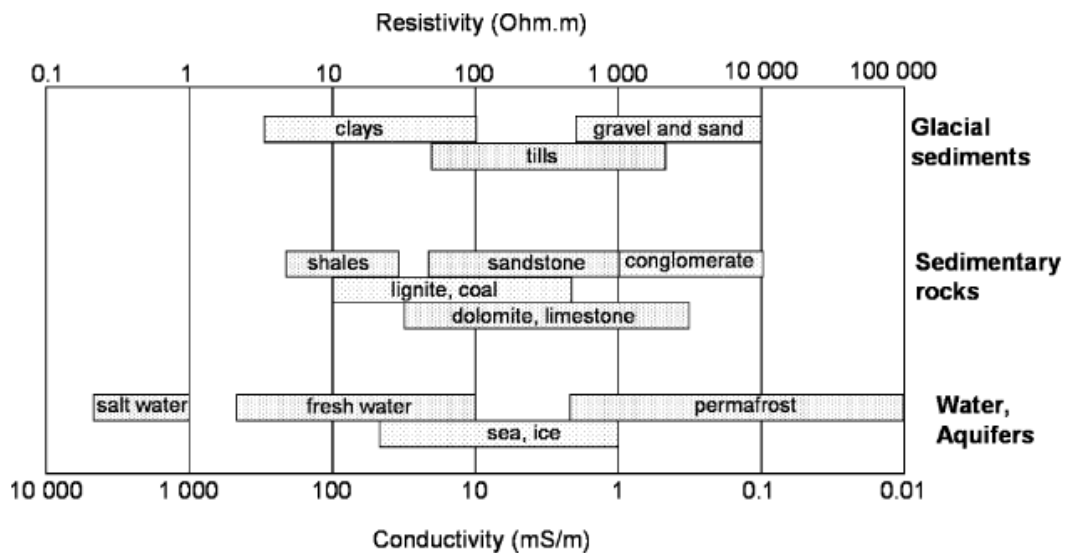
Figure 2 - Typical ranges of electrical resistivities of rocks and water (after modified Palacky, 1987)

The voids distribution and their connection, depending on depositional environment (i.e. grain size and textural structure) and its subsequent alteration, establishes both air and water content and consequently resistivity variation for the subsoil.

The Archie's law puts in relation resistivity with porosity for water saturated geological materials in absence of surface conductivity (i.e. without clay):

$$F = \frac{\rho}{\rho_w} = a\Phi^{-m} \qquad \text{Eq. 6}$$

where F is the formation factor often assumed to be an indicator of hydraulic tortuosity, ρ and ρ_w are resistivity of formation and pore water, *a* and *m* are constants related respectively to the coefficient of saturation and the cementation index, and Φ is the porosity. Archie's law implicitly assumes that the effective porosity (Φ_e) is equal to the total porosity (Φ) of the sample, and that all electrical conduction in a water-saturated rock or soil results from the migration of ions in the bulk pore-solution. If there are isolated pores through which ions cannot migrate, then ne<n, and Archie's law will over predict sample conductivity (Rubin and Hubbard, 2006).

Archie showed that the m ranged from 1.3 for unconsolidated sands to approximately 2.0 for consolidated sandstones (Table 1). Jackson et al. (1978) made electrical conductivity measurements on natural and artificial sand samples, determining that m increased as the grains became less spherical while variations in grain size and sorting had little effect.

Table 1 – Archie's law exponents (m) of different consolidated and non-consolidated media (from Rubin and Hubbard, 2006)

I. MEDIUM	Porosity Range	m, Archie's Exponent	Reference
clean sand	0.12-0.40	1.3	Archie (1942)
consolidated sandstones	0.12-0.35	1.8-2.0	
glass spheres	0.37-0.40	1.38	Wyllie and Gregory (1955)
binary sphere mixtures	0.147-0.29	1.31	
cylinders	0.33-0.43	1.47	
disks	0.34-0.45	1.46	
cubes	0.19-0.43	1.47	
prisms	0.36-0.52	1.63	
8 marine sands	0.35-0.50	1.39-1.58	Jackson et al. (1978)
glass beads (spheres)	0.33-0.37	1.20	
quartz sand	0.32-0.44	1.43	
rounded quartz sand	0.36-0.44	1.40	
shaley sand	0.41-0.48	1.52	
shell fragments	0.62-0.72	1.85	
fused glass beads	0.02-0.38	1.50	Sen et al. (1981)
fused glass beads	0.10-0.40	1.7	Schwartz and Kimminau (1987)
sandstone	0.05-0.22	1.9-3.7	Doyen (1988)
polydisperse glass beads	0.13-0.40	1.28-1.40	de Kuijper et al. (1996)
fused glass beads	0.10-0.30	1.6-1.8	Pengra and Wong (1999)
sandstones	0.07-0.22	1.6-2.0	
limestones	0.15-0.29	1.9-2.3	
Syporex®	0.80	3.8	Revil and Cathles III (1999)
Bulgarian altered tuff	0.15-0.39*	2.4-3.3	Revil et al. (2002)
Mexican altered tuff	0.50*	4.4	
glass beads	0.38-0.40	1.35	Friedman and Robinson (2002)
quartz sand	0.40-0.44	1.45	
tuff particles	0.60-0.64	1.66	

*connected (inter-granular) porosity
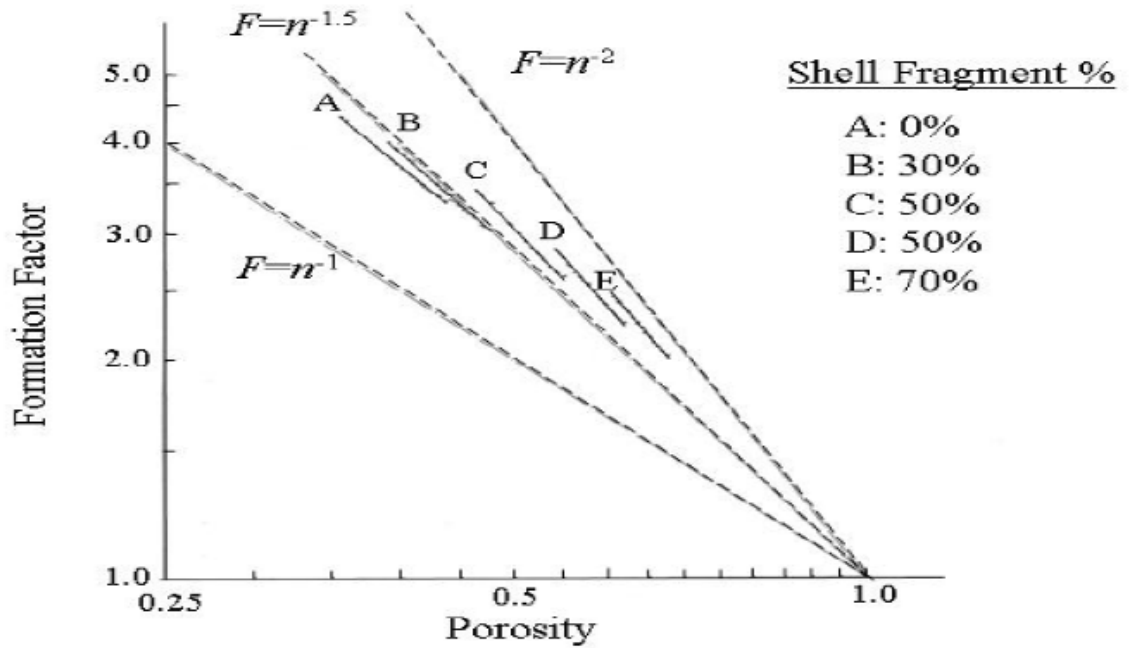
Figure 3 – A log-plot of formation factor versus porosity for granular mixtures consisting of rounded quartz grains and plately shell fragments (after Jackson et al., 1978). The dashed lines are Archie's law predictions for cementation indices of 1.0, 1.5, and 2.0 (from Rubin and Hubbard, 2006)

In presence of conductivity due only to the water ionic content, known the resistivity of pore water, the resistivity of formation and a, m empirical parameters, is possible to calculate the porosity. The Archie's law appears a powerful law for porosity estimation but is limited mainly when surface conductivity is not negligible and in part also in vadose zone.

When the amount of clay or clayey lithologies is not negligible, the Archie's law become inappropriate to describe conductivity and one of the most commonly model used to take into consideration surface conduction is Waxman and Smits law.

$$\sigma = \frac{\sigma_w}{F} + \sigma_{surface} = \frac{1}{F}\left(\sigma_w + BQ_v\right) \quad \text{Eq. 7}$$

where B is the equivalent ionic conductance of the clay exchange cations and Q_v is a measure of effective clay content.

When clay is immersed in a liquid, adsorbed ions at the surface can partially dissociate themselves from the clay particles (Table 2) and become available for ionic conductivity. The large specific surface in clay may supply many ions, leading to a much lower resistivity than in coarse grained soils (Fukue et al., 1999). The amount of clay, in a soil with high content of granular solid, and mainly its uniform distribution into the soil structure can have also a high impact on resistivity, giving high conductivity values.

26

Table 2 –Exchange capacity of common clays (after Keller and Frischknecht, 1966; from McNeill 1980).

Clay	Exchange Capacity
Kaolinite	3 to 15 m-equiv/100 g
Halloysite . $2H_2O$	5 to 10
Halloysite . $4H_2O$	40 to 50
Montmorillonite	80 to 150
Illite	10 to 40
Vermiculite	100 to 150
Chlorite	10 to 40
Attapulgite	20 to 30

Electrical conductivity in subsoil, excluding surface conduction in clay and metallic conduction in presence of massive metals, is principally electrolytic, related to the displacement of ions in pore-water. Under this consideration, water content plays a fundamental role on electrical conductivity. As shown for laboratory tests in Figure 4, for values lower than 15%, a minimal increment of water content decreases exponentially the resistivity.
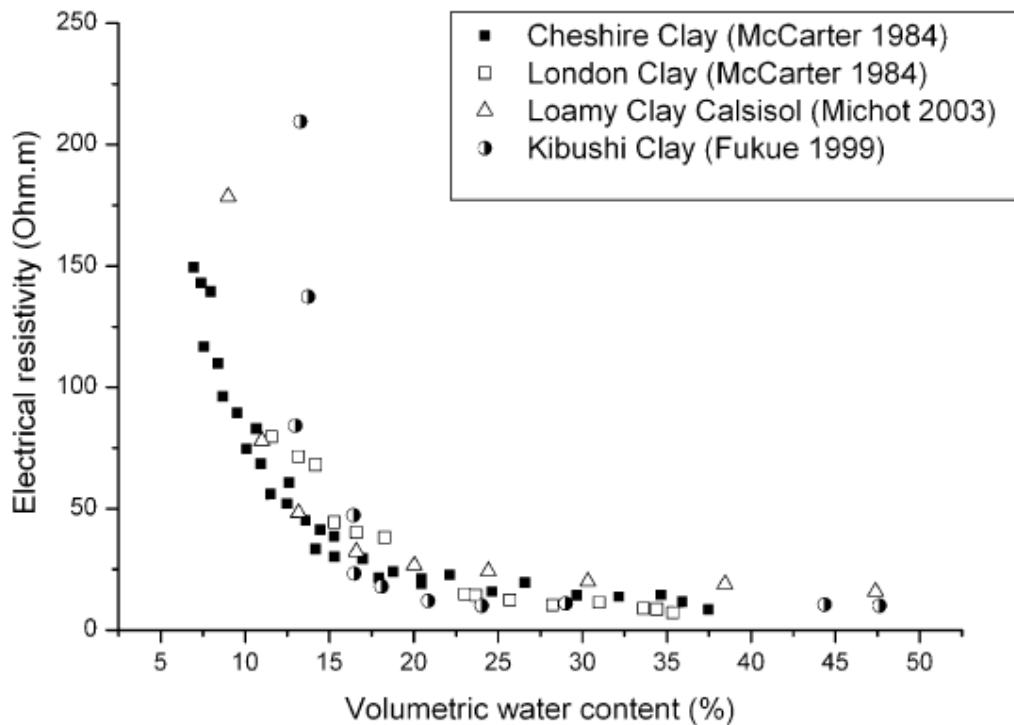
Figure 4 – Relationship between the volumetric water content and the electrical resistivity for different soil types (from Samouëlian et al., 2005).

Archie (1942) proposed an empirical equation for coarse-grained soils between electrical resistivity and water content, based upon laboratory measurements of clean sandstone samples. The relationship was a modified form of the previous equation, taking into account the porosity can be filled by another medium, as gas or oil:

$$S^n = \frac{F\rho w}{\rho} = \frac{a\rho_W}{\Phi^m \rho} \qquad \text{Eq. 8}$$

27

where S is the saturation degree and n a related parameter.

Conductivity of an electrolyte is proportional both to the total number and kind of ions in the solution and the viscosity of the water (McNeill, 1980). Many studies were conducted about the estimation of water content and an early issue was salinity variation. An interesting study was showed from Kalinski et al., (1993), where the volumetric water content is in relation with resistivity for different pore-water conductivity.
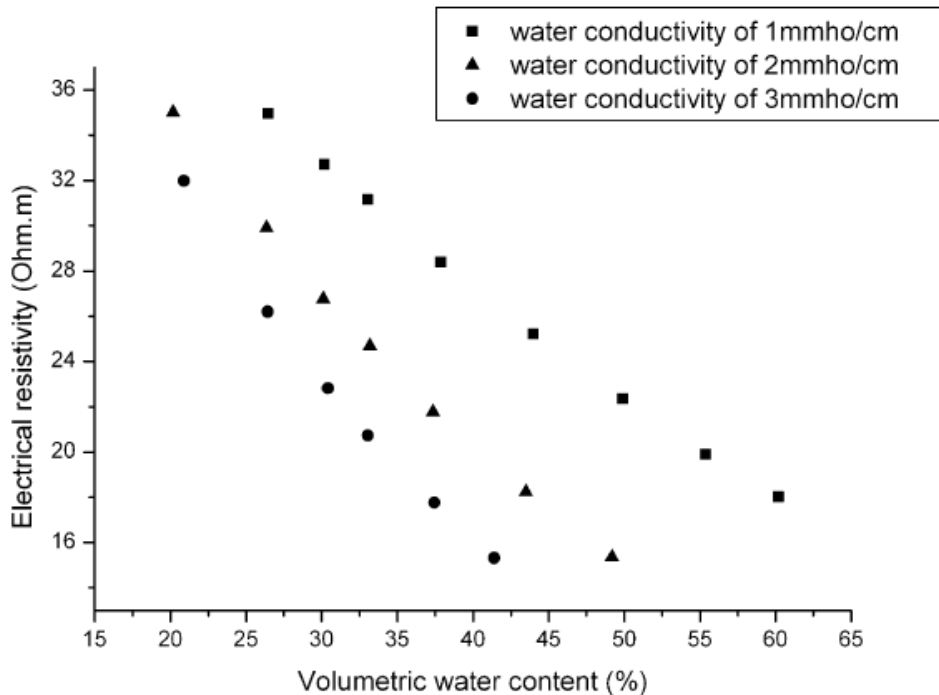
Figure 5 – Relationship between the volumetric water content and resistivity for different values of pore-water conductivity (from Samouëlian et al., 2005).

Thus, electrolytic conductivity can have a large range of variation because of its dependences on ionic composition and concentration, which may change into the soil. On this fact are based several ERT works with the aim to delineate salt water intrusion into coastal areas (e.g. Maillet et al., 2005; De Franco et al., 2009; Zarroca et al., 2011).

Another factor affecting electrical resistivity is temperature: ions agitation increases with the temperature when the viscosity of fluid decreases and consequently decreases resistivity. Campbell et al., (1948) showed a linear increasing of conductivity with temperature in the range from 15° to 35°C with a constant of 2.02% every °C:

$$\sigma_t = \sigma_{25°C} \left[1 + \alpha(T - 25°C) \right] \qquad \text{Eq. 9}$$

where $\sigma_{25°C}$ is the conductivity reference at 25°, σ_t is the conductivity at the temperature experiment and α is the correction constant.

3.2.3 Array configuration

ERT has been applied in a wide range of survey configuration into the space, from 1D to 3D, for azimuthal surveys and multitemporal acquisition, depending on target geometry and its spatial and temporal variability.

One-dimensional configuration with four-electrodes cells A, B, M, N are commonly applied in the field for vertical electrical sounding (VES). VES consists of carry out measures of apparent resistivity increasing subsequently electrodic step, to reveal changes of resistivity with the depth. The assumption is that resistivity varies only with the depth, and the subsoil consists in horizontal layers. This is a limit when the subsoil has lateral variations and in this case 2D or 3D give more correct results.

Two-dimensional multi-electrodic arrays provide 2D resistivity image of the subsoil. The current and potential electrodes are maintained at a regular fixed distance from each other and are progressively moved along a line at the soil surface. At each step, one measurement is recorded. The set of all these measurements at this first interelectrodic spacing gives a horizontal profile of resistivity values. The inter-electrode spacing is increased then by a factor n = 2, and a second measurement level is done. This process (increasing the factor n) is repeated until the maximum spacing between electrodes is reached. One can notice that the larger the n-values, the greater the depths of investigation (Figure 6). As the distribution of the current also depends on the resistivity contrasts of the medium, the depth of investigation deduced conventionally from the spacing is called the "pseudo-depth".
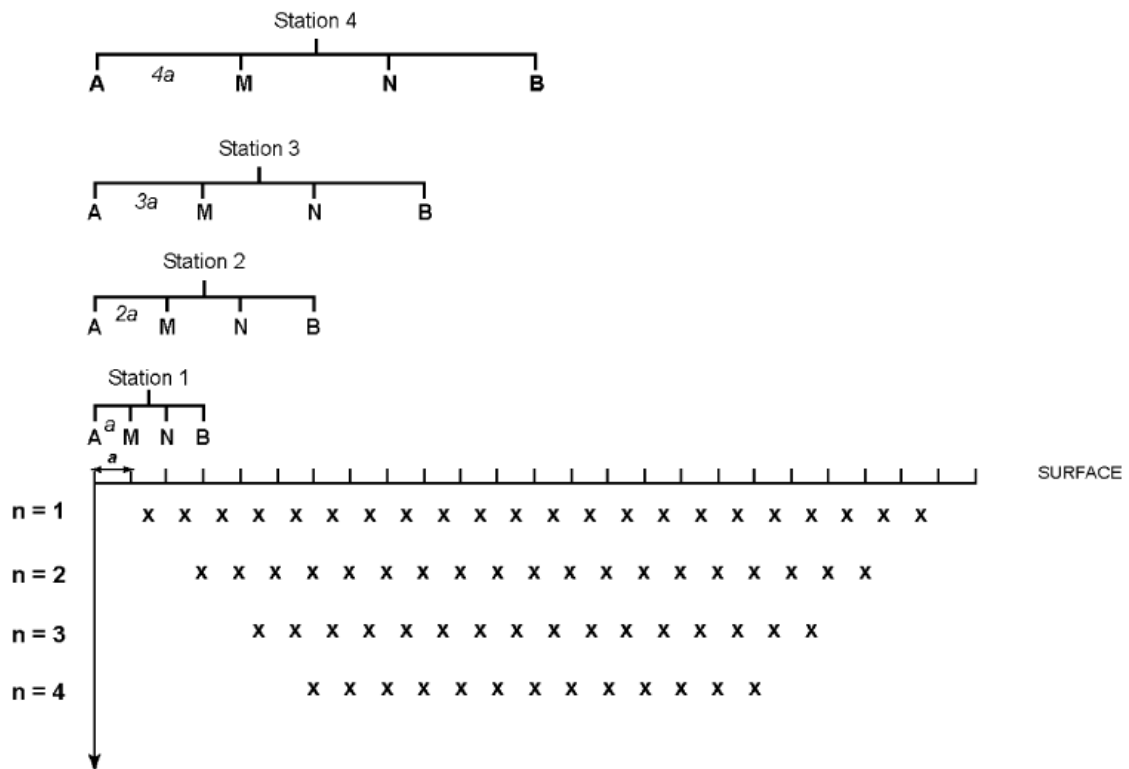
Figure 6 – Distribution of resistivity measurements on a 2D electrical resistivity pseudo section with n varying from 1 to 4.

The reciprocal position of each four electrodes for measurements, two for current and two for potential, leads to different configurations, which will be well explained later.

Table 3 (Samouëlian et al., 2005) summarizes for different 2D array configurations and compares the following characteristics for all the arrays: the sensitivity of the array to horizontal and vertical heterogeneities, the depth of investigation, the horizontal data coverage and the signal strength.

Table 3 – Technical characteristics of different quadripolar configurations; the labels are classified from (+) to (++++), equivalent at poor sensitivity to high sensitivity (from Samouëlian et al., 2005).

	Wenner	Wenner–Schlumberger	Dipole–dipole	Pole–pole	Pole–dipole
Sensitivity of the array horizontal structures	++++	++	+	++	++
Sensitivity of the array vertical structures	+	++	++++	++	+
Depth of investigation	+	++	+++	++++	+++
Horizontal data coverage	+	++	+++	++++	+++
Signal Strength	++++	+++	+	++++	++

Three dimensional analyses are more effective and realistic in case of anisotropic target and can be in two ways: by doing just in-line measures or in-line and cross-line measures. The first method consists of producing a three dimensional electrical dataset by assembling a network of two-dimensional parallel pseudo sections. In this case, an accurate three-dimensional result is thus recorded if electrical anomalies are preferentially oriented and if the in-line measurement electrodes are perpendicular to the orientation of the anomalies. In case of heterogeneous subsoil two-dimensional pseud sections should be acquired at least in two perpendicular directions (Zhou et al., 2001; Chambers et al., 2002). The second method is based on the use of in-line and cross-line quadripoles measurements conferring a huge survey fold on the investigated volume and several directions of measurements by adopting cross-line quadripoles. The resulting electrical resistivity imaging is less influenced by the direction of quadripoles as in two-dimensional case, thus recommended over anisotropic target.

Strong anisotropic targets can be revealed also with azimuthal resistivity survey (Taylor et al., 1988; Boadu et al., 2005): resistivity values are carried out rotating electrodic array to obtain resistivity values as function of the azimuth where the resultant ellipse of resistivity highlights the main direction of anisotropy.

ERT borehole logging has an important role in fine reservoir characterization, its high resolution lead to a punctual description of subsoil changes if compared to the surface methods.

In order to plot the resistivity raw data from survey, a consolidated method is the pseudo section. The resistivity value is placed horizontally at the center of the set of electrodes used to make that measurement. The vertical location is placed at a distance proportional to the gap between the electrodes. For convention, in surveys using the dipole-dipole array, one common method is to place the plotting point at the intersection

of two lines starting from the mid-point of the C1-C2 and P1-P2 dipole pairs with a 45° angle to the horizontal. Another manner is to place the vertical position of the plotting point at the median depth of investigation (Edwards, 1977), or pseudo depth, of the electrode array used. This pseudo depth value is based on the sensitivity values or Frechet derivative for a homogeneous half space. The pseudo section plot obtained by contouring the apparent resistivity values is a convenient means to display the data (Loke, 2004).

The pseudo section obtained gives a distorted picture of resistivity because of the distribution of values depends on the array configuration used for survey and on the true resistivity of subsurface (Figure 7). The pseudo section can be used as a graph for quantitative evaluation of the dataset as well as picking out bad apparent resistivity measurements.
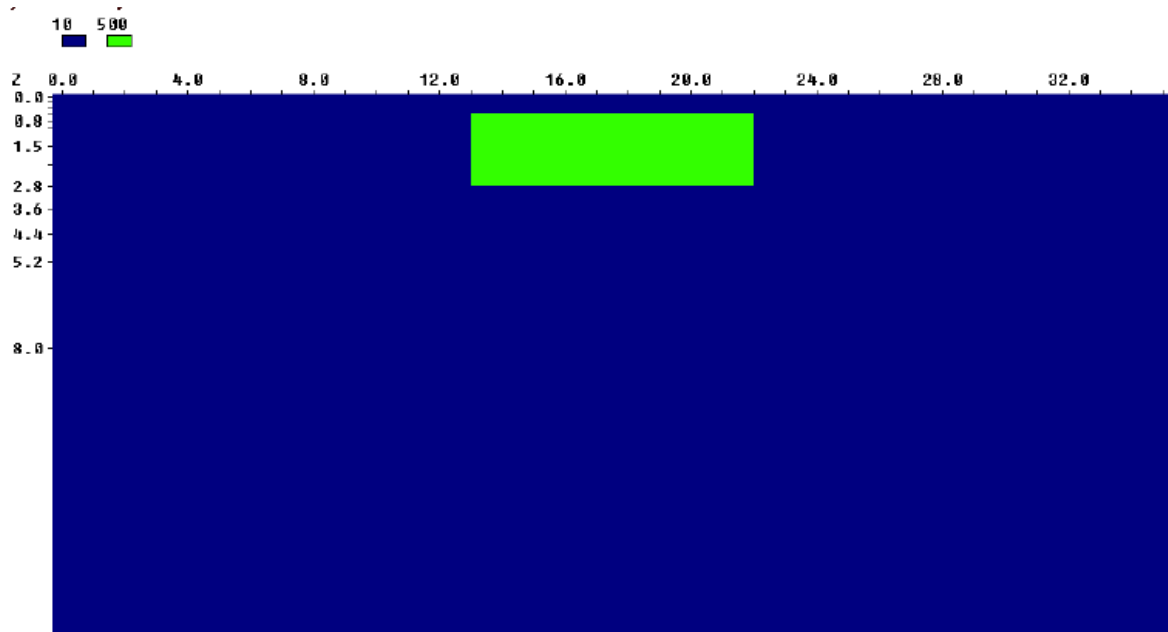
Figure 7 – Resistivity model with a rectangle of 500 ohm*m in a uniform 10 ohm*m space.
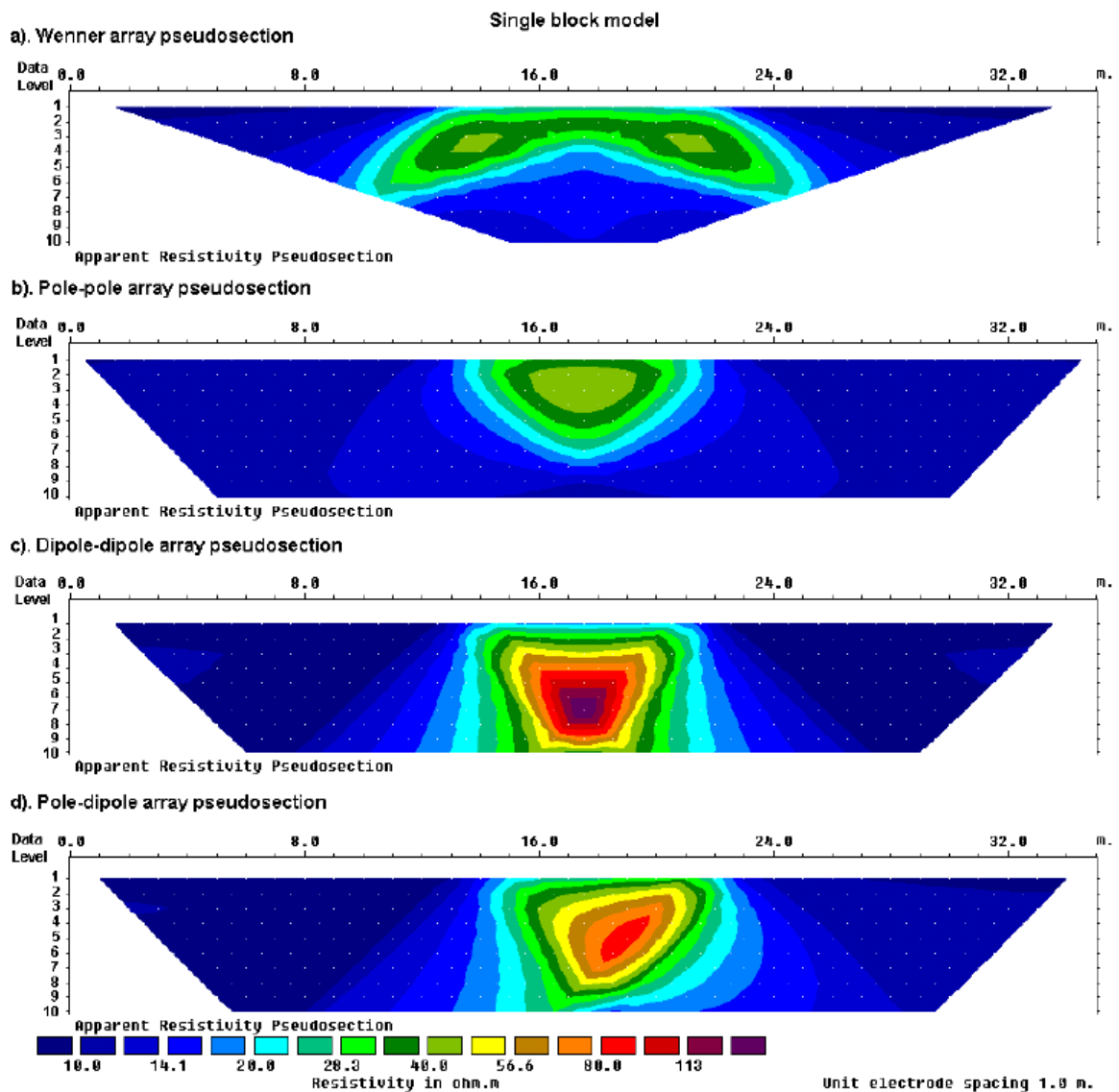
Figure 8 – Pseudo sections from four different array configuration of the same resistivity model in figure 6 (from Loke, 2004).

The choice of the best quadripolar configuration is crucial, because it affects results and is strictly depending on: geometry of the target, sensitivity of the resistivity meter and background noise level (Loke, 2004). Main features to take in account in the choice of the array configuration are depth of investigation, sensitivity to the vertical or horizontal variation of resistivity, signal strength and horizontal coverage. The most common arrays (Figure 9) used in two-dimensional survey are Wenner, Dipole-Dipole, Wenner-Schlumberger, Pole-Pole, Pole-Dipole and less common Multiple Gradient. The implementation of the quadripolar measurements sequence should be realized to avoid polarization effects (Dahlin, 2000). This happens in the case of one electrode of current is used subsequently within a short time as potential electrode (Loke, 2004).

32

Table 4 – The median depth of investigation (ze) for the different arrays (Edwards, 1977). L is the total length of the array. Note identical values of ze/a for the Wenner-Schlumberger and pole-dipole arrays. The geometric factor is for an "a" value of 1 m. For the pole-dipole array, the array length "L" takes into account the active electrodes C1, P1 and P2

		z_e/a	z_e/L	Geometric Factor	Inverse Geometric Factor (Ratio)
Wenner Alpha		0.519	0.173	6.2832	0.15915 (1.0000)
Wenner Beta		0.416	0.139	18.850	0.05305 (0.3333)
Wenner Gamma		0.594	0.198	9.4248	0.10610 (0.6667)
Dipole-dipole	n = 1	0.416	0.139	18.850	0.05305 (0.3333)
	n = 2	0.697	0.174	75.398	0.01326 (0.0833)
	n = 3	0.962	0.192	188.50	0.00531 (0.0333)
	n = 4	1.220	0.203	376.99	0.00265 (0.0166)
	n = 5	1.476	0.211	659.73	0.00152 (0.0096)
	n = 6	1.730	0.216	1055.6	0.00095 (0.0060)
	n = 7	1.983	0.220	1583.4	0.00063 (0.0040)
	n = 8	2.236	0.224	2261.9	0.00044 (0.0028)
Equatorial dipole-dipole					
	n = 1	0.451	0.319	21.452	0.04662 (0.2929)
	n = 2	0.809	0.362	119.03	0.00840 (0.0528)
	n = 3	1.180	0.373	367.31	0.00272 (0.0171)
	n = 4	1.556	0.377	841.75	0.00119 (0.0075)
Wenner - Schlumberger					
	n = 1	0.519	0.173	6.2832	0.15915 (1.0000)
	n = 2	0.925	0.186	18.850	0.05305 (0.3333)
	n = 3	1.318	0.189	37.699	0.02653 (0.1667)
	n = 4	1.706	0.190	62.832	0.01592 (0.1000)
	n = 5	2.093	0.190	94.248	0.01061 (0.0667)
	n = 6	2.478	0.191	131.95	0.00758 (0.0476)
	n = 7	2.863	0.191	175.93	0.00568 (0.0357)
	n = 8	3.247	0.191	226.19	0.00442 (0.0278)
	n = 9	3.632	0.191	282.74	0.00354 (0.0222)
	n = 10	4.015	0.191	345.58	0.00289 (0.0182)
Pole-dipole	n = 1	0.519	0.260	12.566	0.07958 (0.5000)
	n = 2	0.925	0.308	37.699	0.02653 (0.1667)
	n = 3	1.318	0.330	75.398	0.01326 (0.0833)
	n = 4	1.706	0.341	125.66	0.00796 (0.0500)
	n = 5	2.093	0.349	188.50	0.00531 (0.0334)
	n = 6	2.478	0.354	263.89	0.00379 (0.0238)
	n = 7	2.863	0.358	351.86	0.00284 (0.0178)
	n = 8	3.247	0.361	452.39	0.00221 (0.0139)
Pole-Pole		0.867		6.28319	0.15915 (1.0000)

a). Wenner Alpha

```
     C1          P1          P2          C2
    •←——a——→•←——a——→•←——a——→•
            k = 2 π a
```

b). Wenner Beta

```
     C2          C1          P1          P2
    •←——a——→•←——a——→•←——a——→•
            k = 6 π a
```

c). Wenner Gamma

```
     C1          P1          C2          P2
    •←——a——→•←——a——→•←——a——→•
            k = 3 π a
```

d). Pole – Pole

```
     C1          P1
    •←——a——→•
        k = 2 π a
```

e). Dipole – Dipole

```
   C2     C1              P1    P2
  •←—a—→•←———na———→•←—a—→•
       k = π n(n+1)(n+2) a
```

f). Pole – Dipole

```
   C1                P1    P2
  •←————na————→•←—a—→•
      k = 2 π n(n+1) a
```

g). Wenner - Schlumberger

```
   C1          P1  P2          C2
  •←———na———→•←—a—→•←———na———→•
       k = π n(n+1) a
```

h). Equatorial Dipole – Dipole

```
  C2                            P2
  •                             •
  ↑                    b        ↑
  a ←—————————na—————————→ a
  ↓                             ↓
  •                             •
  C1                            P1

       b = n a
       k = 2 π b L / (L - b)
       L = (a * a + b * b)^{0.5}
```
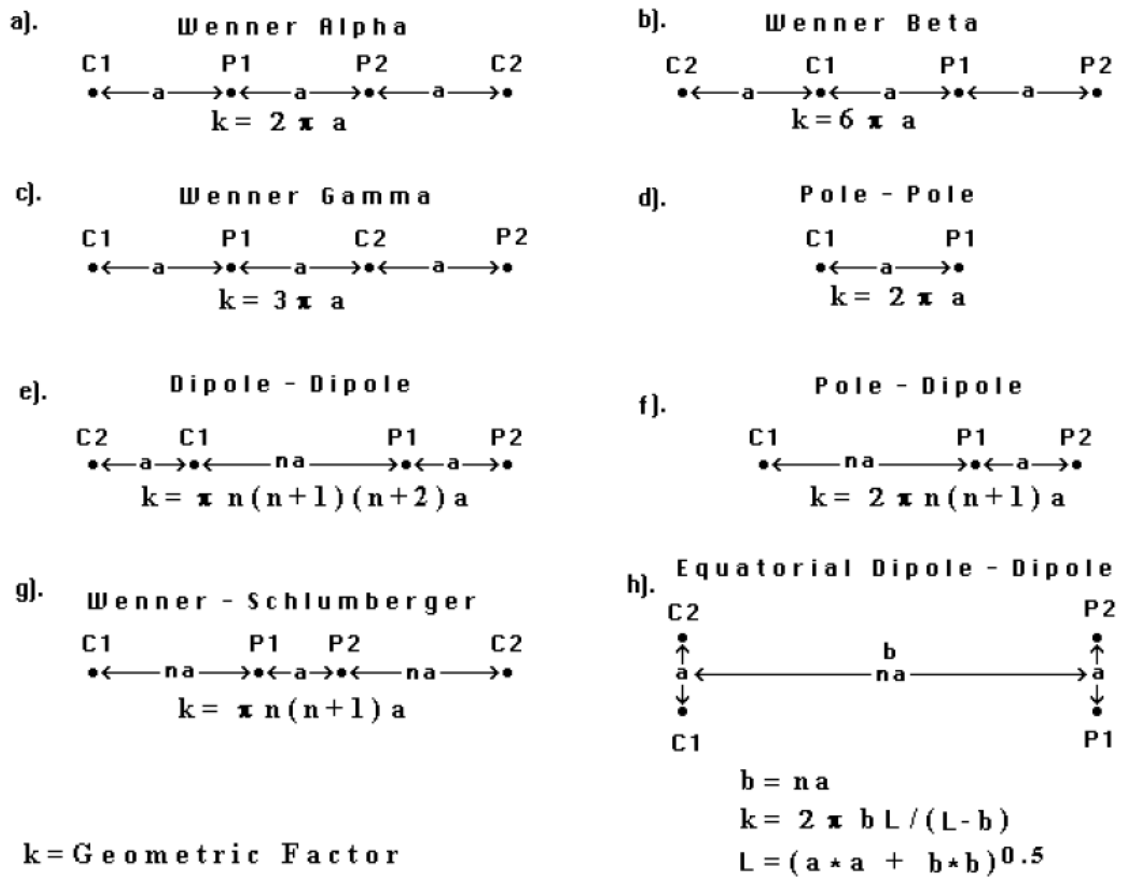
$$k = Geometric\ Factor$$

Figure 9 – Common array used in resistivity surveys and their geometric factors where "a" is the dipole length and n the dipole separation factor (from Loke, 2004).

The Wenner configuration was introduced by the works carried by the University of Birmingham research group (Griffiths and Turnbull, 1985; Griffiths, Turnbull and Olayinka, 1990). For a four-electrode array, there are three possible permutations of the positions of the electrodes (Carpenter and Habberjam, 1956). The most common array is the Wenner Alpha. In Figure 23a, the sensitivity plot for the Wenner Alpha array shows almost horizontal contours above the center of the quadripol. Due to this, the Wenner array has a good resolution for vertical changes but relatively poor horizontal resistivity changes, i.e., is vocated to resolve horizontal structures. Moreover, the Wenner array has the strongest signal among the common arrays, which can be useful in case of high background noise. A drawback of this configuration is the poor lateral coverage, especially when applied with a small number of electrodes.
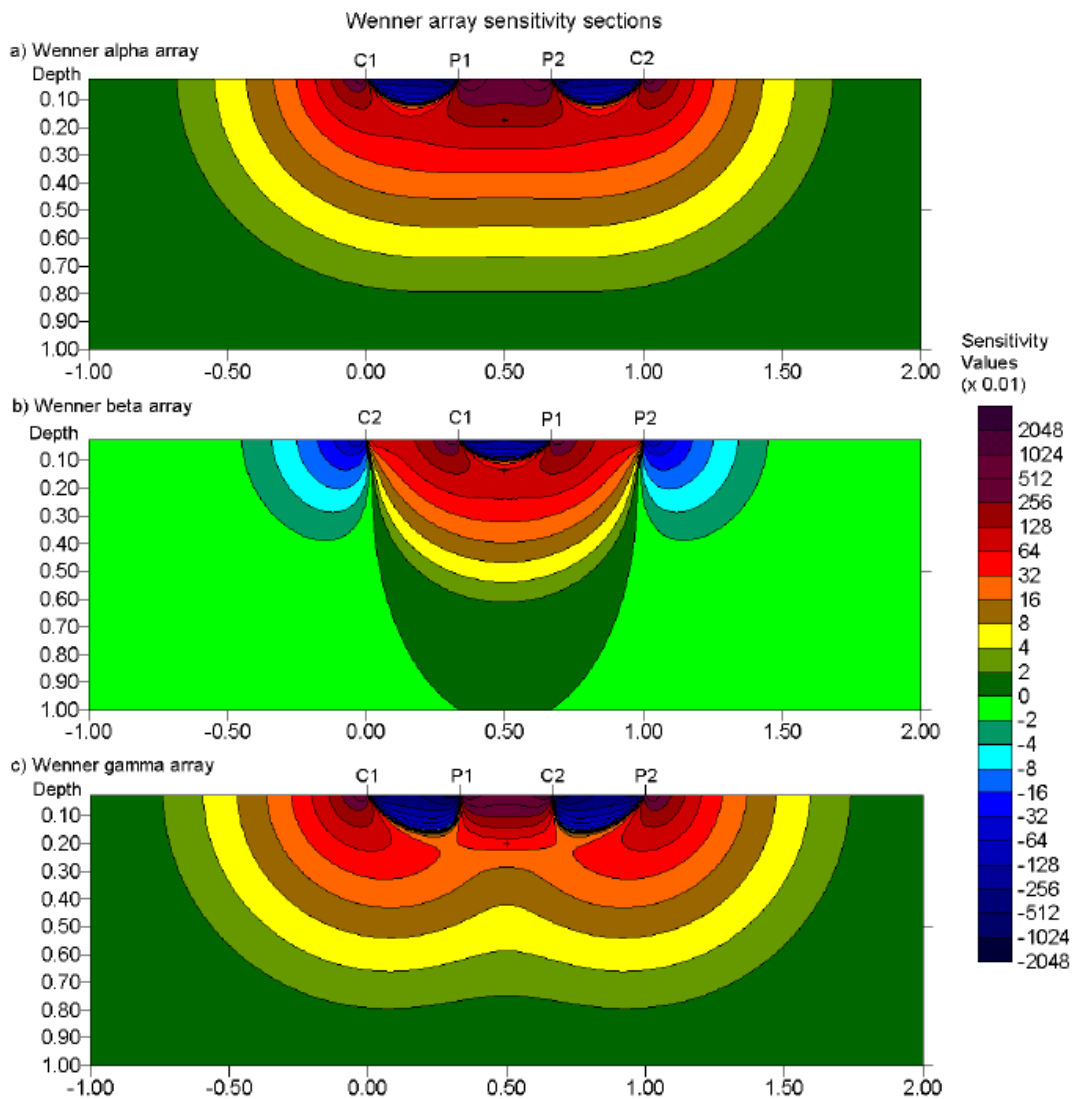
Figure 10 – 2D sensitivity sections for the Wenner array in a) alfa, b) beta and c) gamma configurations (from Loke, 2004).

The spacing between the current electrodes (C1-C2) pair, "a", is the same as the distance between the potential electrodes pair (P1-P2). The ratio of the distance between the C1 and P1 electrodes to the C2-C1 (or P1-P2) is known as "n". The investigation depth increases with a and n.

Figure 11 shows the sensitivity sections for this array for "n" values ranging from 1 to 6. The highest sensitivity values are generally located between the current dipole pair, as well as between the potential pair, that means the array is most sensitive to resistivity changes below the two dipoles. The almost vertical contour pattern for the sensitivity at "n" values greater than 2, leads the dipole-dipole array to be very sensitive to horizontal changes in resistivity and relatively insensitive to vertical changes in the resistivity. Thus it is good for revealing vertical structures, such as dykes, vertical discontinuities and cavities, but relatively poor in revealing horizontal structures such as horizontal stratigraphic layers. Dipole Dipole array has been, and is still, widely used in I.P. surveys because of the low EM coupling between the current and potential circuits (Loke, 2004).
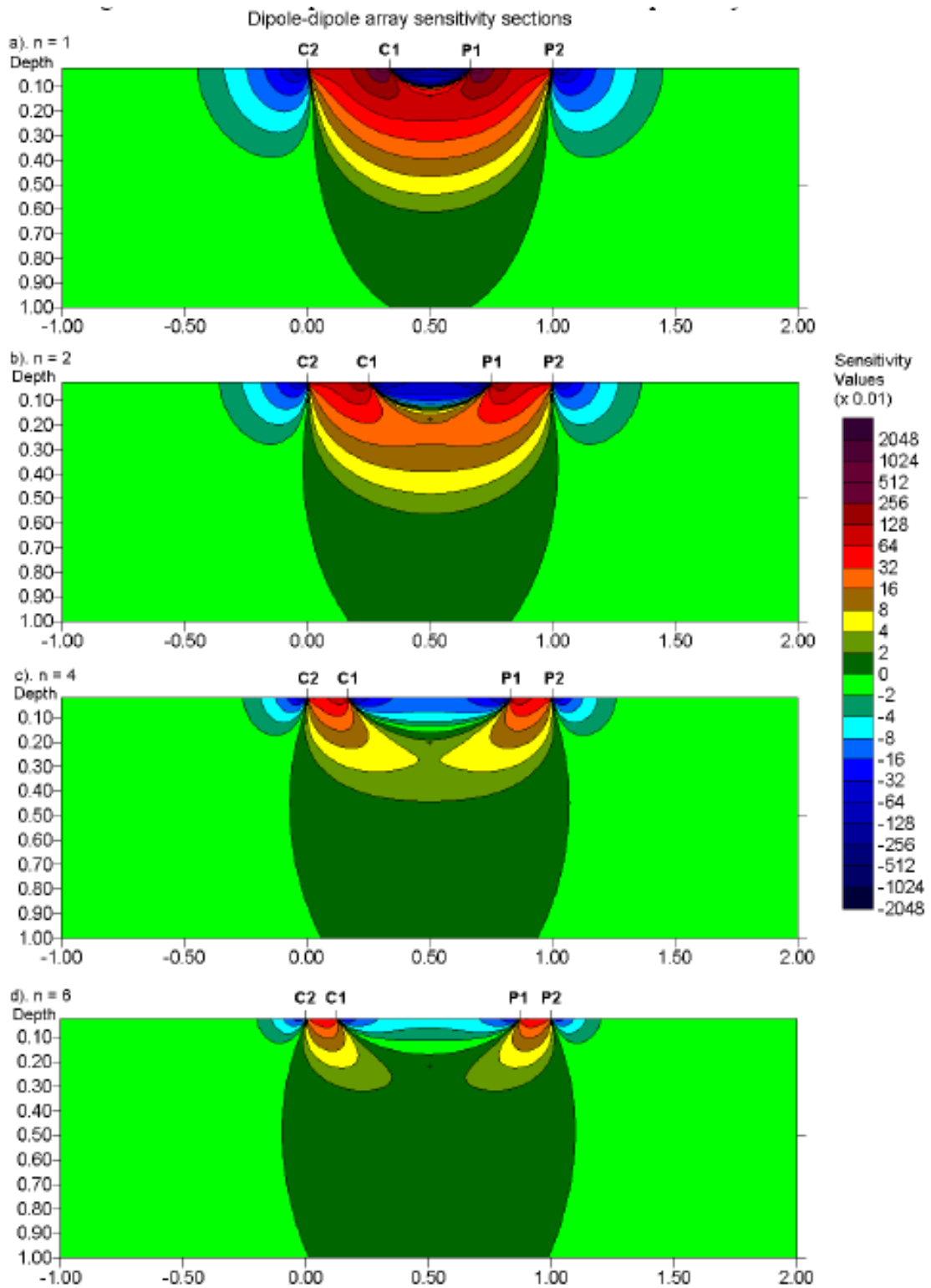
Figure 11 – 2D sensitivity sections for the dipole –dipole array: a) n=1, b) n=2, c) n=4, d) n=6 (from Loke, 2004).

This array has better lateral data coverage than the Wenner (Figure 8) that could be an important strength to extend laterally the survey. The voltage of this array is inversely proportional to n^3, thus the strength of the signal decreases rapidly with the increasing of n. For the same current, the voltage measured by the resistivity meter drops by about 56 times when "n" is increased from 1 to 6 (Loke, 2004, Table 2).

Wenner Schlumberger is a hybrid configuration between the Wenner and Schlumberger arrays (Pazdirek and Blaha 1996). The classical Schlumberger array is one of the most commonly used arrays for resistivity sounding surveys, where current dipoles are outer and spaced n times the distance between the potential dipole (Figure 9). A special case of this array with n=1 is the Wenner configuration.

Varying the n value from 1 (Wenner array) to 6 (the classical Schlumberger array) its sensitivity pattern changes in a way that confers a moderate sensitivity to both, horizontal (for low n values) and vertical (for high n values) structures. In areas geologically complexes, where both types of resistivity changes are expected, this array might be a good compromise between the Wenner and the Dipole-Dipole (Loke, 2004). Compared with other classical configurations, the Wenner-Schlumberger has a signal strength lower than the Wenner array, but higher than the dipole-dipole and the pole-dipole. This configuration produces slightly wider 2D sections than the Wenner array (Figure 13), but narrower than that obtained with the dipole-dipole array.

The Schlumberger reciprocal is a manipulation of classical Wenner-Schlumberger, where the current dipole is in the inner part of the quadripol. The larger distance of potential electrodes instead of the classical Wenner-Sclumberger makes it more sensible the telluric noise. One advantage of this configuration, in addition to be moderately sensitive to both horizontal and vertical variations, is the possibility to use multi-channel optimization for simultaneous measurements. In presence of high number of electrodes this features lead to a high-resoluted very low time-consuming configuration.
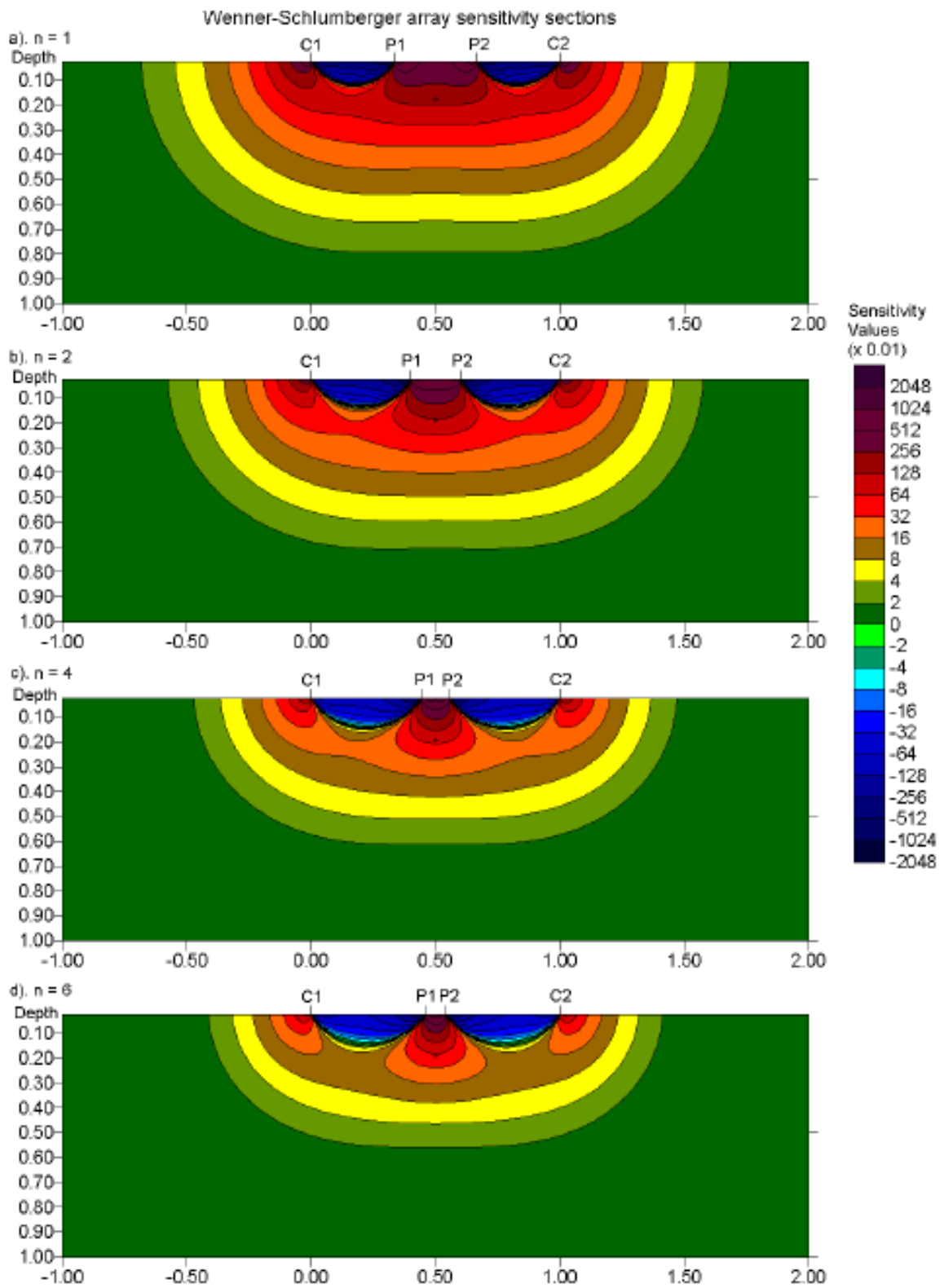
Figure 12 – 2D sensitivity sections for the Wenner-Schlumberger array: a) n=1, b) n=2, c) n=4, d) n=6 (from Loke, 2004)

i).
a). Wenner   b). Schlumberger

ii).
a). Wenner

b). Schlumberger

⊥ Electrode   + Datum point in pseudosection
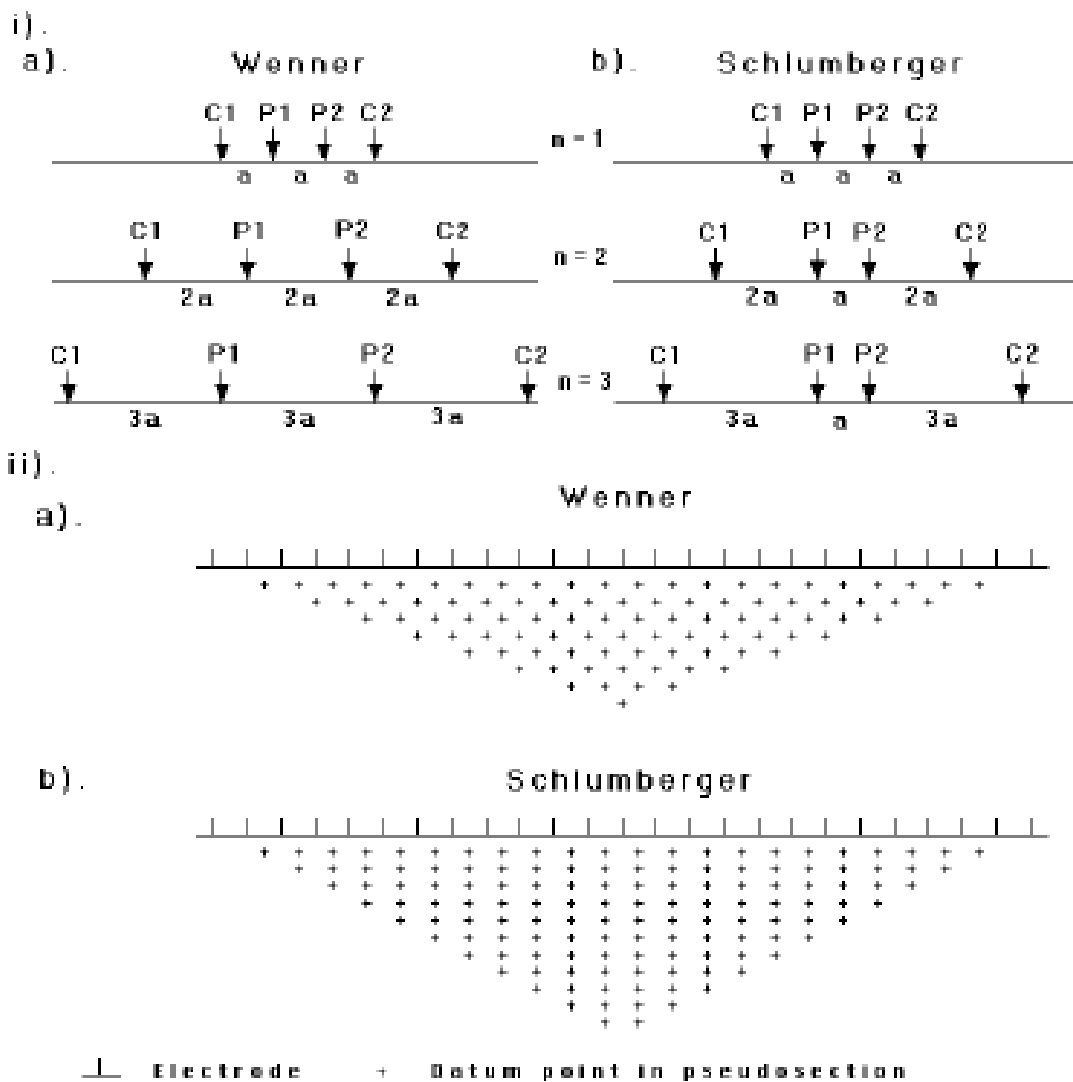
Figure 13 – A comparison of the (i) electrode arrangement and (ii) pseudo section data pattern for the Wenner and Wenner-Schlumberger arrays (from Loke, 2004).

The theoretical Pole-Pole array has only one current and one potential electrode (Figure 4d), but in real applications it does not exist. The errors introduced in measurements from a non-infinite distance of C2 and P2 electrodes from C1 and P1, are proportional to their distance compared to the C1-P1 spacing. To reduce at minimum above effects, the second current and potential electrodes (C2 and P2) must be placed at a distance that is more than 20 times the maximum separation between C1 and P1 electrodes used in the survey (Loke, 2004). In practice this is suitable for survey where the electrodic spacing is not more than few meters. However, there are inversion softwares that can take into account the real distance of C2-P2 poles. Another drawback is due to the large distance between the P1 and P2 electrodes that can pick up a large amount of telluric noise and severely degrade the quality of the measurements. The strengths for this array are the widest lateral coverage and the deepest depth of investigation. However, its resolution is the poorest among classical configurations.
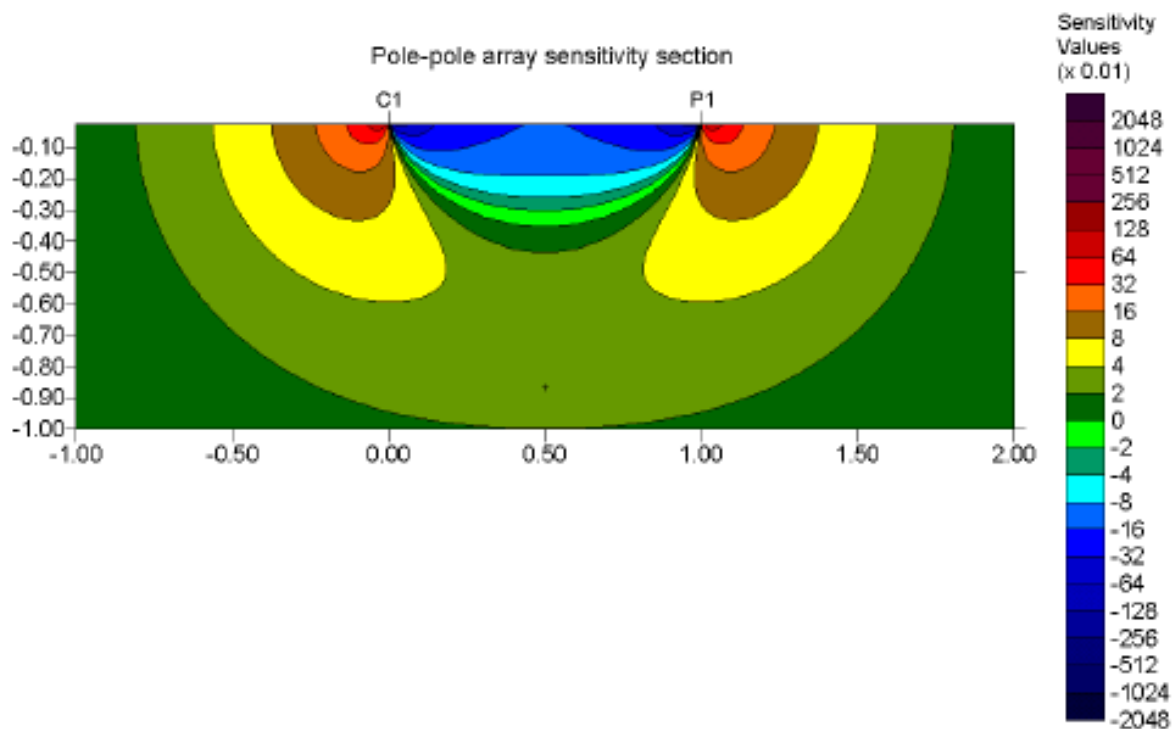
Figure 14 – The pole-pole array 2D sensitivity section (from Loke, 2004).

As the Pole-Pole, the Pole-Dipole array has in theory one current pole C2. In practice the second current pole is placed far from the line at least 5 times the maximum C1-P1 spacing, to introduce negligible errors in measurements. The pole-dipole array also has relatively good horizontal coverage and a significantly higher signal strength compared with the dipole-dipole array and it is not as sensitive to telluric noise as the pole-pole array. Thus, can be used in situations where the background noise limitates the use of Dipole-Dipole. Its asymmetrical configuration can lead to asymmetries into the tomographic resistivity image; a manner to avoid this fact is to repeat measurement in forward and reverse configuration (Loke, 2004).

The sensitivity contour pattern in Figure 15 become more vertical with the increasing of n and due to this also the Pole-Dipole has a good sensitivity to the lateral changes (i.e. vertical structures).

The signal strength is lower when compared with the Wenner and Wenner-Schlumberger arrays but higher than the dipole-dipole array. For I.P. surveys, the higher signal strength (compared with the dipole-dipole array) combined with the lower EM coupling (compared with the Wenner and Wenner-Schlumberger arrays) due to the separation of the circuitry of the current and potential electrodes makes this array an attractive alternative.
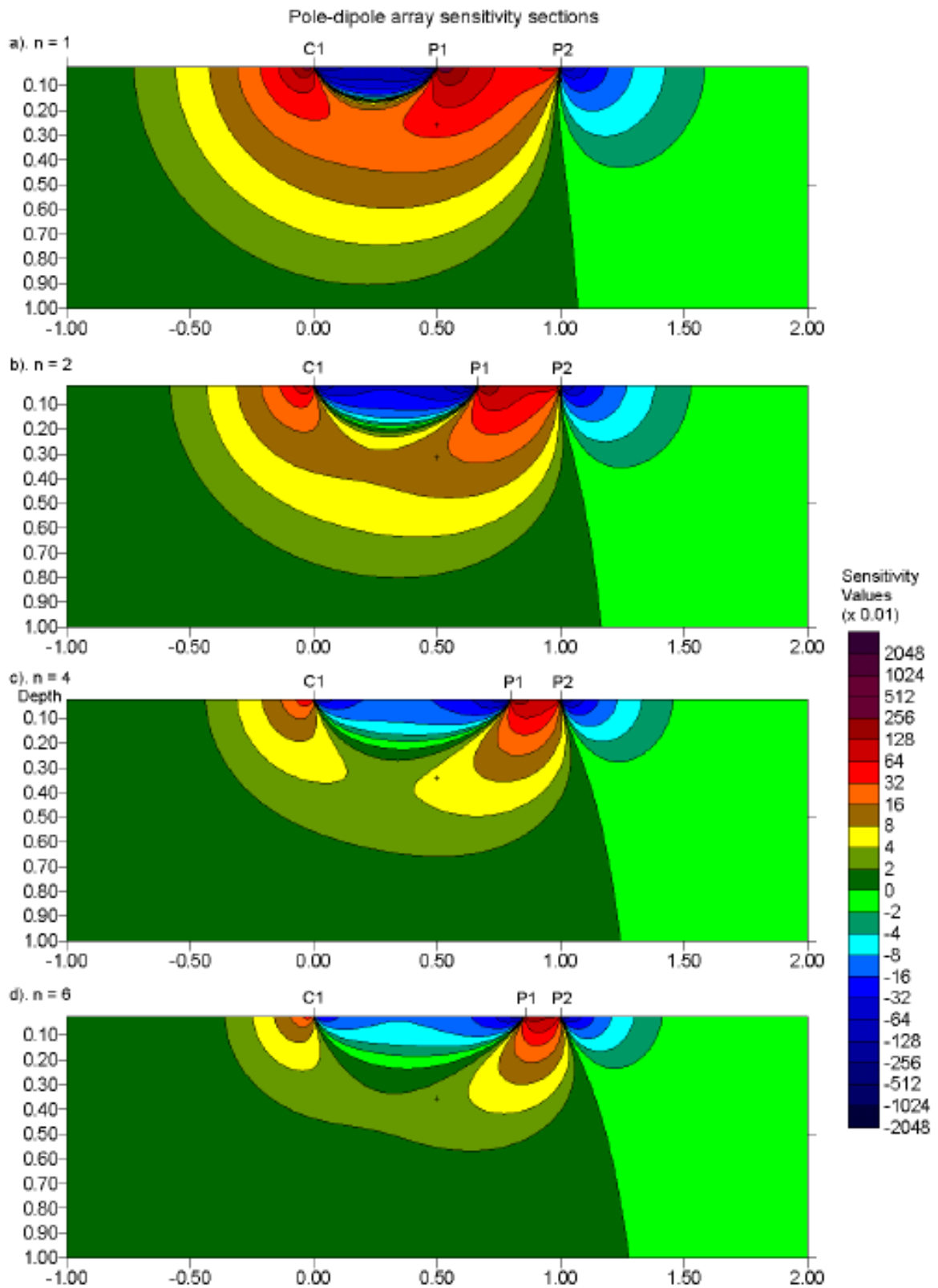
Figure 15 –The pole-dipole array 2D sensitivity section. The sensitivity sections with (a) N=1, (b) n=2, (c) n=4 and (d) n=6.

The Multiple Gradient is a relatively new array developed primary for multi-channel resistivity meter systems (Dahlin and Zhou, 2006). A multi-channel system has the advantage to acquire simultaneously several resistivity measurements with the potential

electrode pairs at different location, maintaining at the same locations the current electrodes. The current electrodes in the traditional gradient array are fixed at the end of the line, while measurements are made with the potential electrodes at different positions along the line. In the multiple gradient array, different sets of measurements are made with the current electrodes at different locations.

The results obtained by this array are comparable to those obtained by the Wenner-Schlumberger and Pole-Dipole arrays (Dahlin and Zhou, 2004) while it has more favorable noise characteristics (Zhou and Dahlin, 2003).

The Schlumberger reciprocal is an array for multichannel acquisitions, as the multi gradient array, where the electrodes disposition is similar to the Wenner-Schlumberger but the current electrodes C1 and C2 are in the inner part of the quadripol. A drawback of this array is suffering possible effects of telluric currents in large dipole configurations (Loke, 2004) with respect to the Wenner-Schlumberger because of the reciprocal configuration. Conversely, allows imaging data also in complex environments, in presence of lateral and vertical resistivity variations, with tomographic models similar to the classical Wenner-Schlumberger (Bernardinetti et al., 2018).

In practice when the array is fixed depending on the target and the maximum number of electrodes available, a usual technique to extend horizontally the survey, is the roll-along. This method consists in moving forward a part of the array after completed each sequence of measurements; the array is moved generally of a constant number of electrodes, based on the required coverage of the data at the bottom of the tomography. The roll-along allows to acquire continuous data for a high horizontal extension with the planned interelectrodic step (Figure 16).
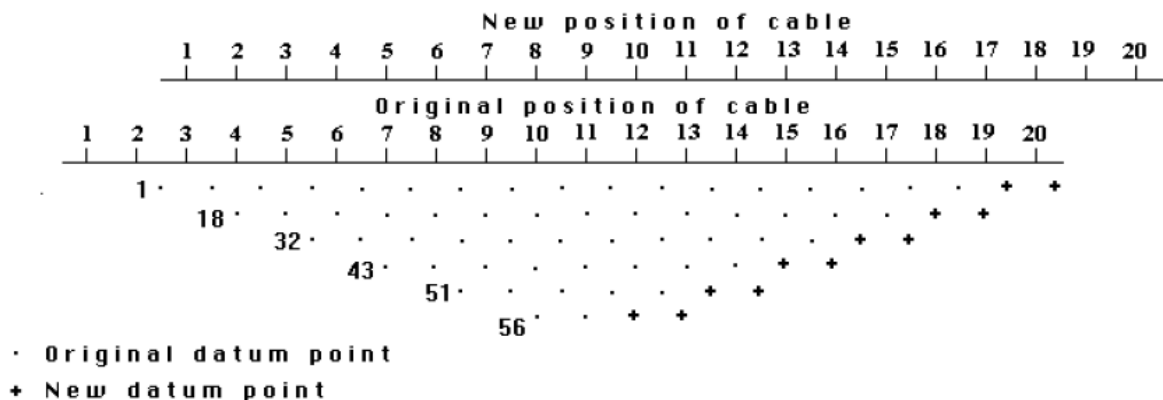
Figure 16 – An example of lateral extension of a 2D survey with the roll-along technique: the bigger points are carried out with the new position of electrodes.

3.2.4   *Inversion method*

As explained in the § 3.2.3 the immediate result of a 2D and 3D is respectively a resistivity pseudo section or a pseudo volume. Thus, to obtain an electrical resistivity image is necessary a conversion from apparent resistivity values to the resistivity model. In order to achieve this conversion one must apply the RES2DINV (Loke, 2002) inversion routine.

The 2-D model used by the inversion program, which consists of a number of rectangular blocks, is shown in Figure 2. The arrangement of the blocks is loosely tied to the distribution of the data points in the pseudo section. The distribution and size of the blocks is automatically generated by the program using the distribution of the data points as a rough guide. The depth of the bottom row of blocks is set to be approximately equal to the equivalent depth of investigation (Edwards 1977) of the data points with the largest electrode spacing. A forward modelling subroutine is used to calculate the apparent resistivity values, and a non-linear least-squares optimization technique is used for the inversion routine (deGroot-Hedlin and Constable 1990, Loke and Barker 1996). The program supports both the finite-difference and finite-element forward modelling techniques.
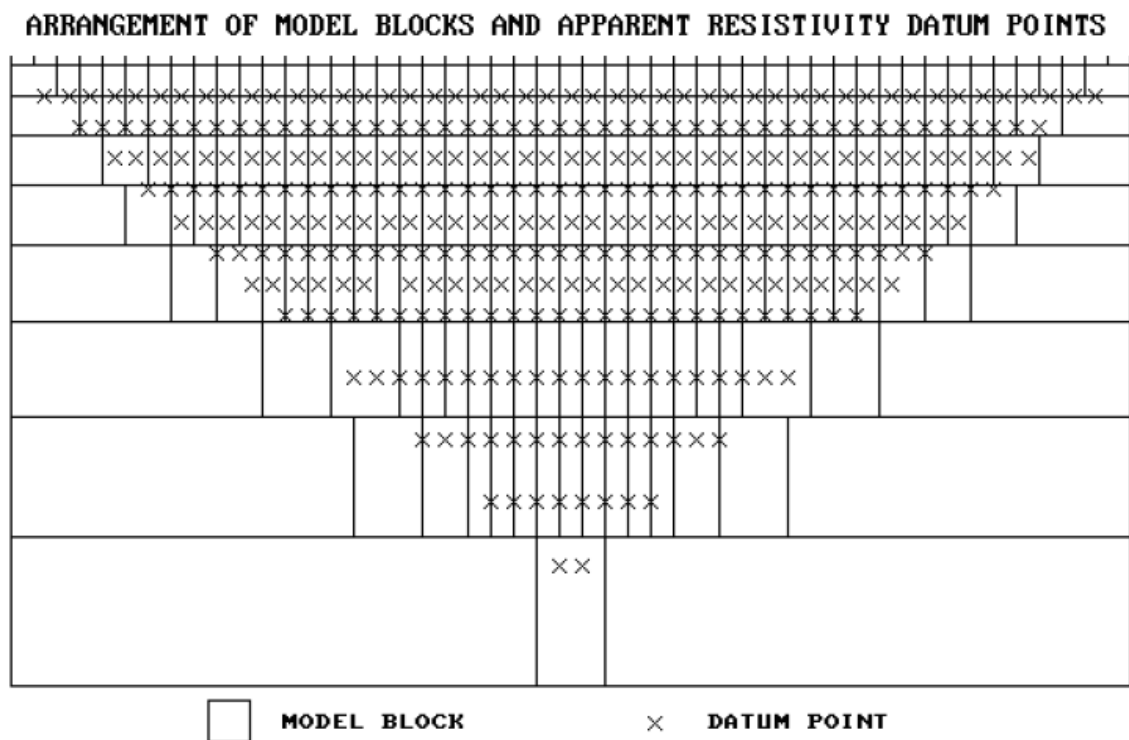
Figure 17 – Arrangement of the blocks used in a resistivity model with the data points in the pseudo section.

The smoothness-constrained least-squares method (L2–norm) is based on the following equation:

$$(\mathbf{J}^t\mathbf{J}+\lambda\mathbf{I})\Delta\mathbf{q}_k=\mathbf{J}^t\mathbf{g}$$

where

$\mathbf{I}$ = identity matrix

43

λ = Marquardt or damping factor

J = matrix of partial derivates (or Jacobian matrix)

Δq = model parameter change vector

g = discrepancy vector between observed data and model response data

One advantage of this method is that the damping factor and flatness filters can be adjusted to suit different types of data.

The purpose of this program is to determine the resistivity of the rectangular blocks that will produce an apparent resistivity pseudo section that agrees with the actual measurements. The layer discretization depends on the array configuration applied to carry out the survey and in general the thickness of each subsequent layer is increased by a value ranging from 10% to 25% with the depth. The depths of the layers can also be changed manually by the user.

The optimization method basically tries to reduce the difference between the calculated and measured apparent resistivity values by changing the resistivity of the model blocks. This difference can be assessed by the root-mean squared error (RMS). Iterations stop when reached a defined number or the variation in terms of RMS is lower than a threshold value; on other hand additional iterations don't give substantial improvement at the results. However the model with the lowest possible RMS error can sometimes show large and unrealistic variations in the model resistivity values and might not always be the "best" model from a geological perspective (Loke, 2004). In general the most prudent approach is to choose the model at the iteration after which the RMS error does not change significantly. This usually occurs between the 3rd and 5th iterations (Loke, 2010).

The conventional smoothness-constrained least squares method (L_2-norm) (deGroot-Hedlin and Constable 1990) attempts to minimize the square of the changes in the model resistivity values. This formulation is given by:

$$(\mathbf{J^t J}+\lambda \mathbf{W^t\, W})\Delta \mathbf{q_k}=\mathbf{J^t g}-\lambda \mathbf{W^t\, W\; q_k}$$

where q is the model resistivity vector.

The resulting resistivity image is a model with a smooth variation in the resistivity values and is suitable in the case of subsurface resistivity varies in smooth manner. In presence of sharp boundaries with high resistivity contrast the smoothness-constrained method produces images far from the real distribution of resistivity: the resistivity will change in a too smooth manner with respect to the real distribution.

When the robust model constrain inversion method (L_1-norm) is applied for the processing of the resistivity model, the program will attempt to minimize the absolute changes in the resistivity values. His formulation is:

$$(\mathbf{J^tR_dJ}+\lambda\mathbf{W^t\,R_w\,W})\Delta\mathbf{q_k}=\mathbf{J^t\,R_d\,g}-\lambda\mathbf{W^t\,R_m\,W\,q_k}$$

This constrain tends to produce models with sharp interfaces between different regions with high contrast of resistivity values, and within each region the resistivity value is almost constant. This might be more suitable for areas where such a geological situation exists, such as the soil-bedrock interface over a homogeneous bedrock.
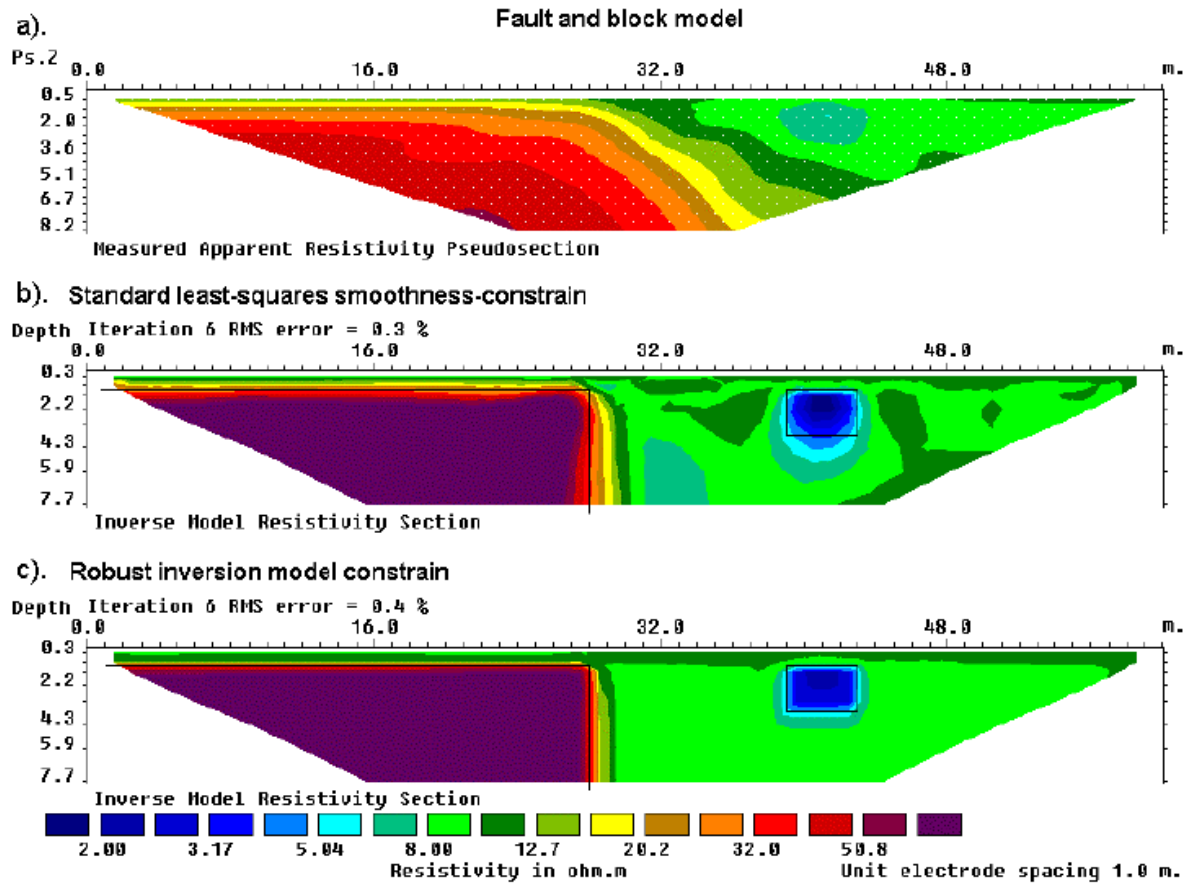
Figure 18 – Example of inversion results using the smoothness-constrain and robust inversion model constrain. (a) Apparent resistivity pseudo section (Wenner array) for a synthetic test model with a faulted block (100 ohm*m) in the bottom-left side and a small rectangular block (2 ohm*m) on the right side with surrounding medium of 1010 ohm*m. The inversion models produced by (b) the conventional least-squares smoothness-constrained and (c) the robust inversion method.

3.3 P – Waves seismic refraction tomography

In hydrogeological investigations, it is necessary to discriminate a number of essential properties of water bearing formations (i.e. aquifers) before confidence can be gained for long term supply, which preserves the integrity of the aquifer system for all present and future users. Seismic surveying can help mainly in the measurement and resolution of few essential aquifer characteristics as:

- Aquifer extensions individuated by boundary conditions such as faults and other important structures.
- Revealing unconsolidated sediments or rocks and their degree of fracturation.
- Saturated from unsaturated deposits discrimination.
- Inferring of sparse drill-hole information.

Over many decades the P-waves velocity has been used for an indirectly characterization of porous aquifers. Many attempts on the definition of the Vp value for the saturated level, demonstrated its non-univoqueness, with values ranging from 1200 to 1800 m/s (Hasselström, 1969; Zohdy et al., 1974; Sander, 1978; Haeni, 1986). Tomographic studies showed that the water table corresponds to a P-wave velocity of 1100–1200 m/s (Azaria et al., 2003; Zelt et al., 2006), with values as low as 1000 m/s (Watson et al., 2005). Although the range of velocity for saturated sediments might vary from about 1100 to 1800 m/s, Vp can be applied as a proxy to reveal changes in saturation because of the significant lower Vp values in unsaturated porous alluvial deposits. In addition, the 2D refraction tomography can be used to show the large variation of Vp between unconsolidated sediments and rocks.

3.3.1   Theory and basic principles

Seismic refraction is a method for the characterization of subsoil based on the refraction phenomena, due to the changing of the subsurface velocity, according with the Snell law. When an incident wave reaches an interface between two layers with different seismic waves velocity, there is the generation of reflected waves and refracted waves. In the case of two layers with velocity V_0 and V_1, and $V_0 < V_1$, where V1 is the velocity of the lower layer, the critical refracted wave travels at the interface, with the velocity of the lower layer. Under the hypothesis of $V_1 > V_0$, a seismic signal generated at the surface produces direct waves travelling in the upper layer with velocity V_0, refracted waves penetrating the lower layer and refracted waves travelling at the interface between the two layers with velocity V_1.

In Figure 19 are shown the subsoil model and the correspondent travel time, in the case of three layers with horizontal asset and increasing velocity with the depth. The travel time is represented in a Cartesian graph with offset source-receivers and time of arrival at the geophones. Depicted lines with different angular coefficient are representative of direct and refracted waves, where the reciprocal of angular coefficients are the velocities of the

three layers. The distance in which changes the angular coefficient of travel time depends on the thickness of the layers.
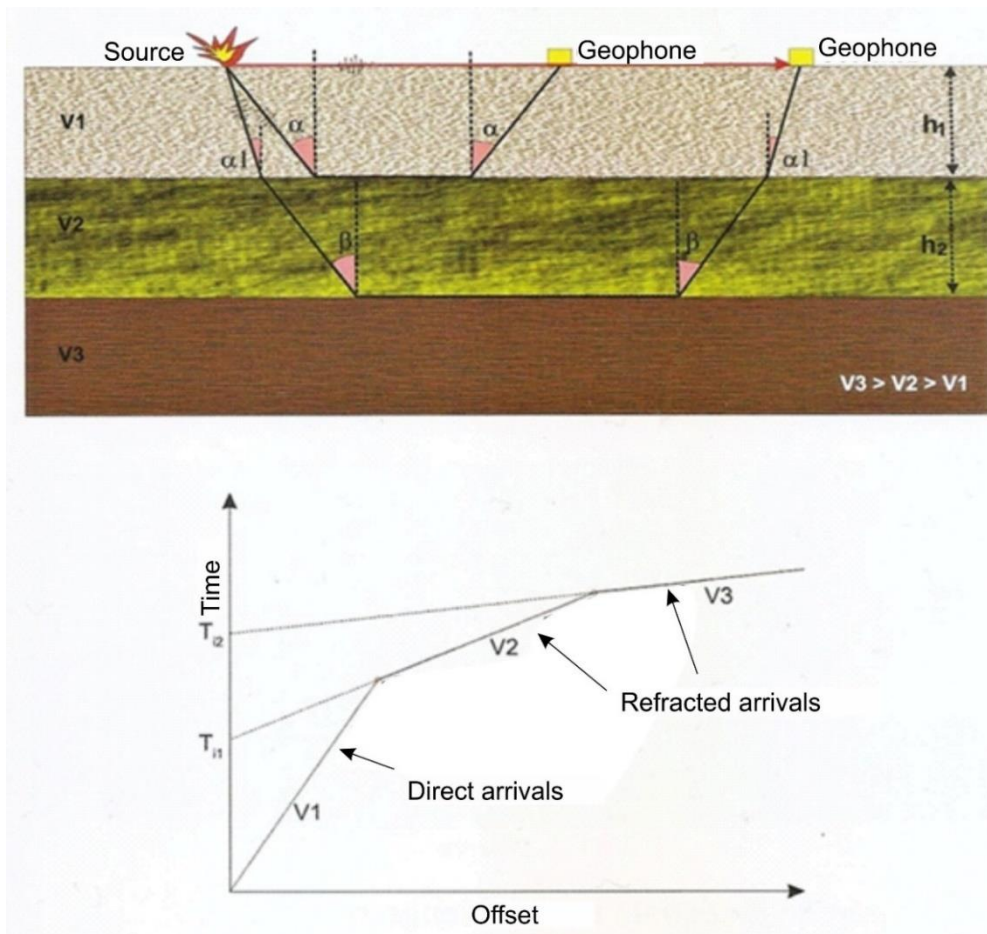
Figure 19 – Theoretical refraction in horizontal multi-layer space. In the upper part a schematic model of subsoil with 3 layers and source-receivers spreading on surface; in the lower part a measured travel time of first arrivals, with direct and refracted waves.

In a more general case with dipping interface between layers, appears a new unknown, β, the angle between a horizontal line and the interface (Figure 20). In this case, in order to resolve the problem, are required two symmetrical sources (i.e. conjugated sources). The reciprocal of angular coefficients allow calculating in this case apparent velocities, because the travel times are influenced also by the geometry of the interface. Anyway, the apparent velocities are still dependent on the real velocity of the layers, as follows:

$$V_2d = \frac{V1}{(\sin(\alpha - \beta))} \qquad \text{Eq. 10}$$

$$V_2c = \frac{V1}{(\sin(\alpha + \beta))} \qquad \text{Eq. 11}$$

With V_1 known from the dip of direct arrivals, the unknown α and β are obtained by the $V_2d = \frac{V1}{(\sin(\alpha - \beta))}$ Eq. 10 and $V_2c = \frac{V1}{(\sin(\alpha + \beta))}$ Eq. 11.
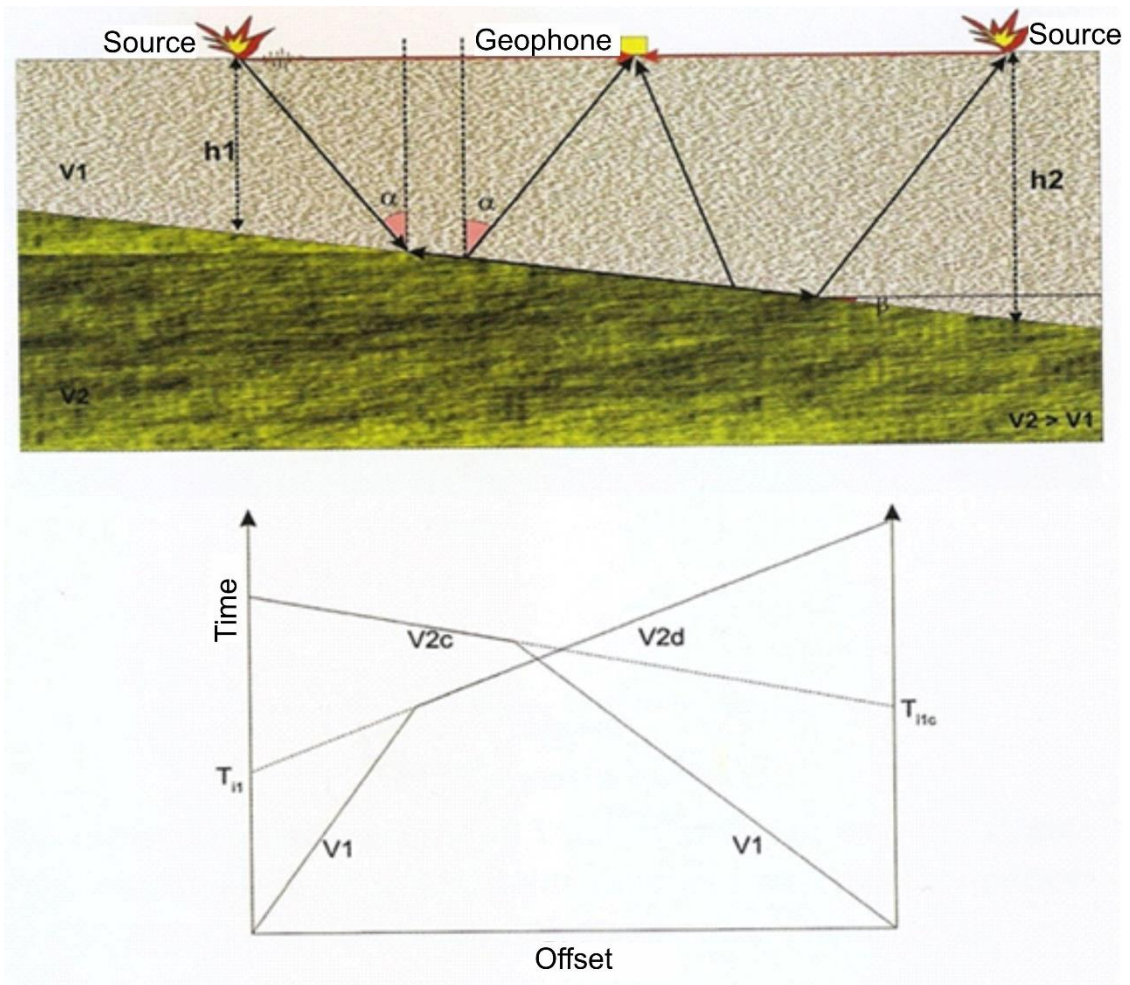
Figure 20 – An example of refraction in presence of a dipping interface between the two layers: in the upper part the subsurface model and in the lower the correspondent travel times.

Also more complex subsurface geometries can be revealed by the refraction method, as faulted zone. On travel times the effect of a fault, is produced by a different level of the top of the layer 3, which introduces a shifting of the refracted arrivals maintaining the same angular coefficient (Figure 21).
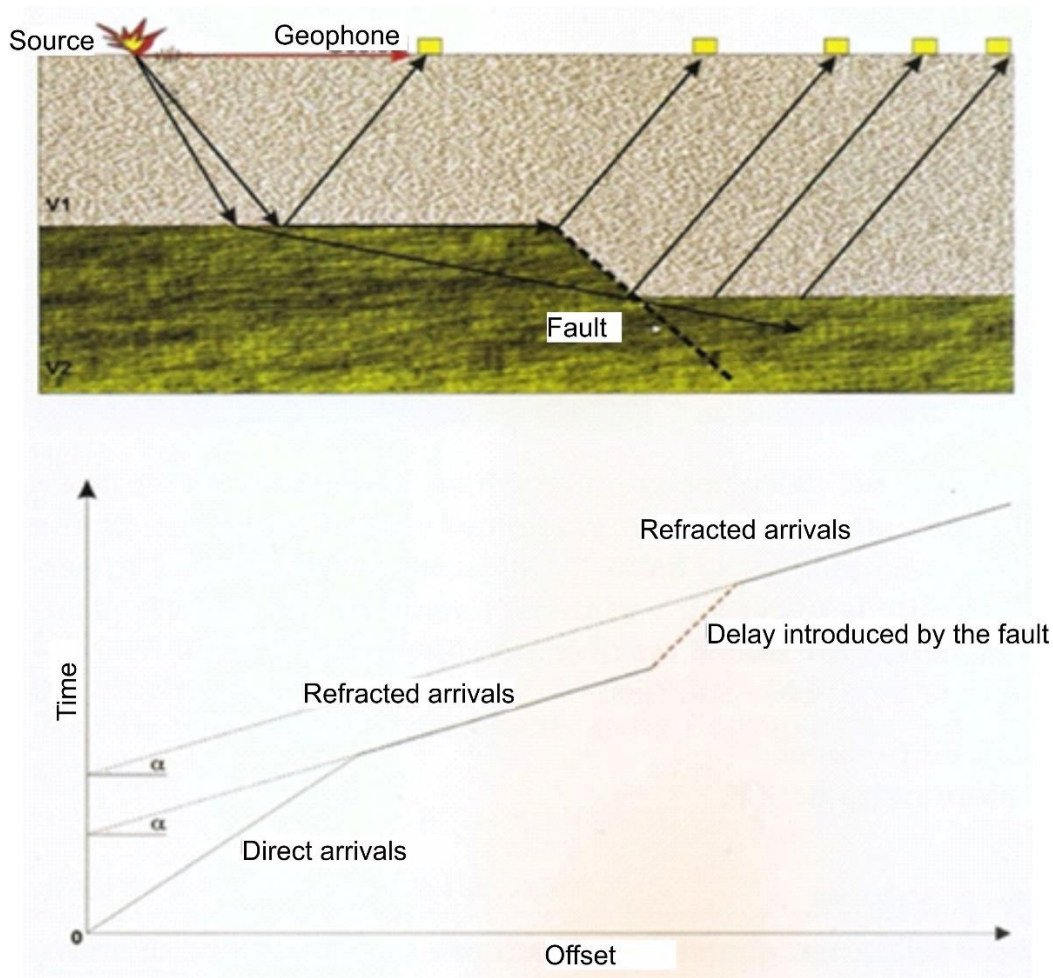
Figure 21 – An example of refracted arrivals in case of faulted zone: in the upper part the subsoil model with two layers and a faulted zone; in the lower part the travel time

3.3.2 Factors affecting P-waves velocity

P-waves velocity are mainly influenced by factors as porosity, pressure, mineralogy and saturation. Wyllie et al. (1956) showed the effect of porosity and saturation on synthetic media and sedimentary rocks and obtained an equation as follows:

$$\frac{1}{Vp} = \frac{\varphi}{Vw} + \frac{1-\varphi}{Vm} \qquad \text{Eq. 12}$$

where Vp is the p-waves velocity of the rock, Vw is the pore-fluid velocity, Vm is the matrix velocity and φ is the porosity.

The equation is obtained under assumptions of intergranular porosity, in absence of clay, normally consolidated that are far from the real conditions. Anyway, the Wyllie equation showed a good reliability in case of clean sandstones with porosity from 10% up to 25%, increasing then the error with the porosity for values higher than 25% (Han et al., 1986). An improvement was proposed by Raymer et al., (1980), that proposed different equations varying the porosity; in Figure 22 (Knackstedt et al., 2005) are compared results from Wyllie and Raymer equations for different samples of silty sands.
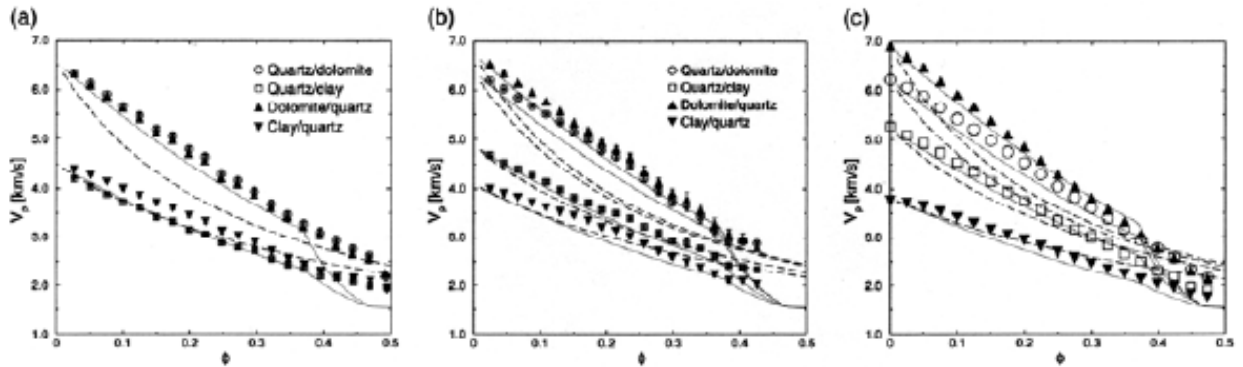
Figure 22 – A comparison of simulation results for water saturated silty sands with Wyllie equations (dashed line) and Raymer equation (dotted line). Minerals ratio are: (a) 1:1, (b) 2:1, (c) 4:1. The Raymer equation gives better prevision of the Wyllie equation, but in many cases underestimates Vp for a given Φ (from Knackstedt et al., 2005).
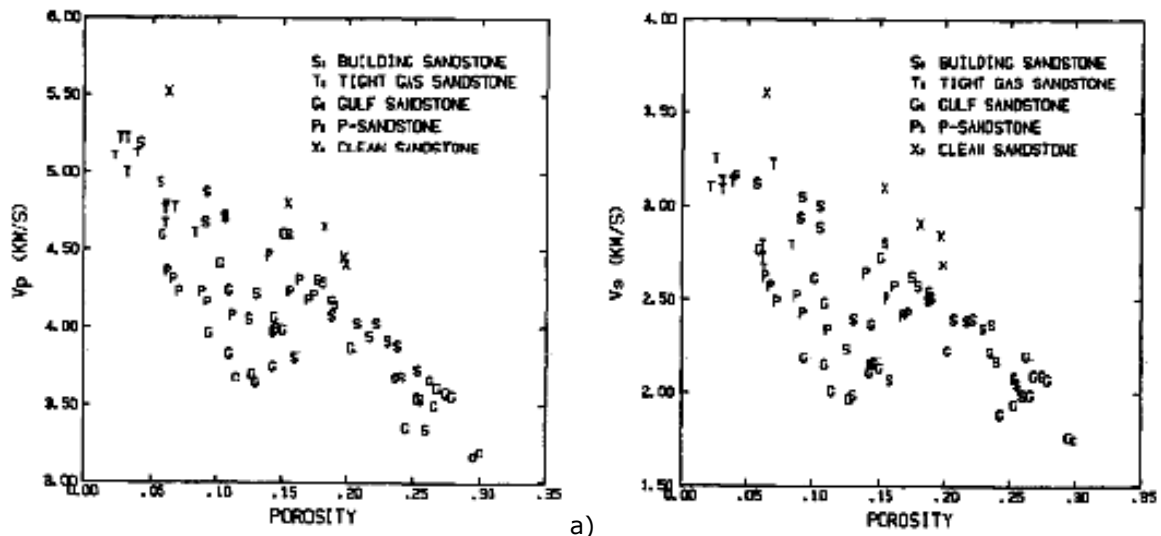
Figure 23 – P waves (a) and S waves velocity versus porosity for 75 samples of sandstone with Pc=40 MPa and Pp= 1 MPa (from Han et al., 1986).

Petrographic analyses showed that intragranular clay particles create an elevate micro porosity reducing pores dimensions (Scotellaro, 2007). Thus, the effect of clay content is not negligible for velocity prediction models. As seen in Kowallis et al. (1984), a large variation of velocity related to the clay content is due to the clay minerals micro porosity. Taking into consideration results from Tosaya and Nur (1982) with Kowallis et al. (1984), dry and saturated sandstones have similar velocity-porosity-clay content relationships. The velocity decreasing with an increased clay content is due to the micro porosity from clay, since in presence of clay an increase of porosity gives lower velocities than the expected in the case of clay free sandstone.

Parra et al., (2003), conducted several velocity experiments on rocks samples at the core scale to assess the influence of porosity on P-waves and S-waves velocities, differentiating pore structure. In general, P and S-wave velocities decrease with increasing porosity in all facies. This velocity variation occurs because P- and S-wave velocities depend on the shear bulk moduli, or rigidity, of the rock matrix. In effect, the pores reduce the structural integrity and hence the rigidity of the rock. Thus, flow units associated with connected moldic/vuggy-porosity are low-velocity zones.
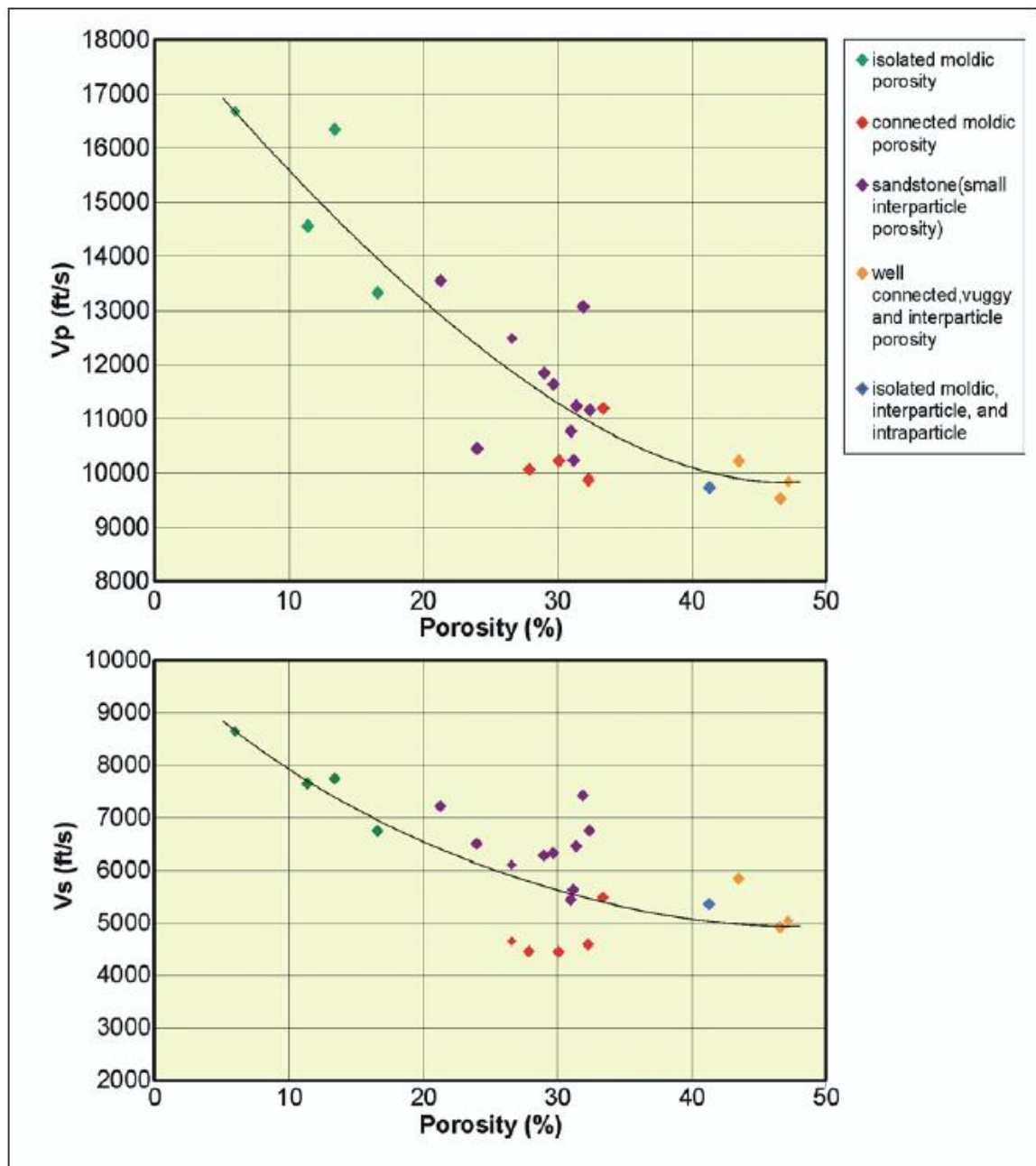
Figure 24 – Vp and Vs plotted against porosity for five lithofacies (from Parra et al., 2003)

Marion et al., (1992) demonstrated that clay affects velocity and porosity based on its volume content on samples of sand and clay mixtures. In case of clay volume fraction is lower than sand porosity (c < φ), clay particles are present in intergranular spaces reducing the porosity. The effect is an increasing of the velocity with the increasing of clay content. In case of clay volume fraction is greater than sand porosity(c > φ), the granules are suspended into the clayey matrix. The rising of clay content leads to an increasing of porosity and a reduction of velocity.
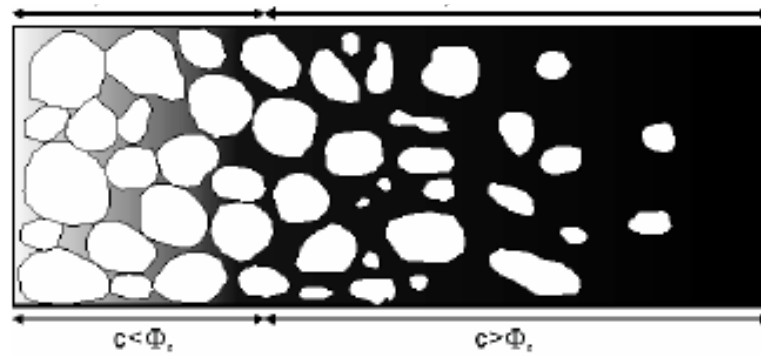
Figure 25 – Sand grains disposition with the increasing of clay content in the sand-clay mix, where "c" is the volume fraction of the clay compared with the porosity of the sand Φs (from Marion et al., 1992). When c< Φs the soil is mainly composed by sands and silty-sands and for c>Φs silts (from Marion et al., 1992).
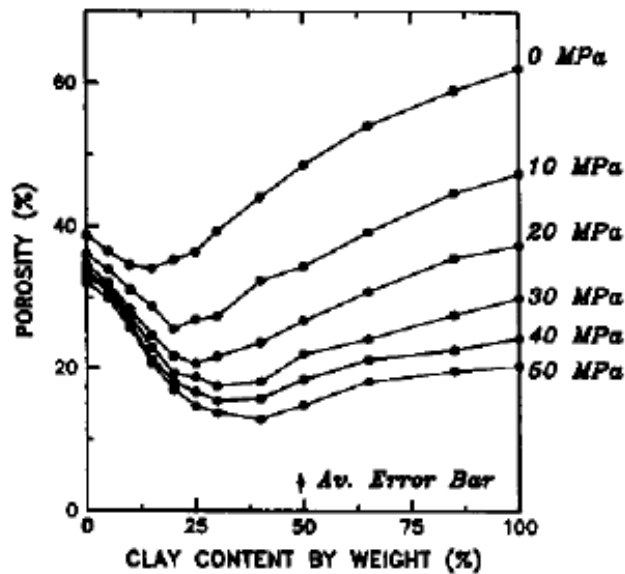
Figure 26 – Porosity variation with the clay content in mix of saturated sand-clay: a minimum value of porosity for a clay content ranging from 20% to 40%; the pressure reduces the porosity and move the minimum porosity towards higher clay content (from Marion et al., 1992)
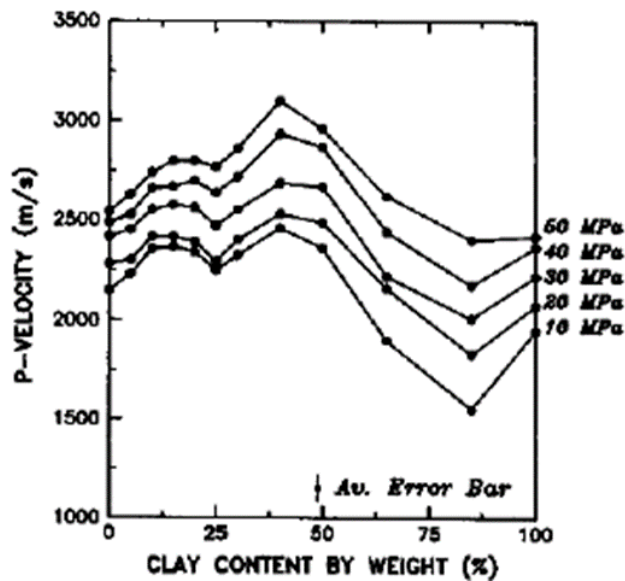
Figure 27 – P waves velocity variation with the clay content in samples of saturated sand-clay. The maximum values are in correspondence of a clay content of 40%. The increasing of velocity is in the order of 20-30% higher than pure sand or clay. Minimum values at 85% clay content are due to errors in the measure length of the samples (from Marion et al., 1992).

Theoretical approaches demonstrate that in a sandy deposit, the maximum increase of VP, occurs at a water content corresponding to a saturated state, Sr=1.00, where Sr is the degree of saturation (Biot, 1956a,b; 1962). Moreover, Yoshimi et al., (1989) demonstrated that "low saturation" in sand, is related to VP significantly smaller than 1500 m/s and saturated sand have velocity greater than 1500 m/s.

3.3.3   Inversion method

Starting from an acquired seismic dataset, it is necessary a tomographic inversion technique to obtain a continuous two dimensional P waves velocity model (Iyer and Hirahara, 1993).

First arrivals with an adequate signal to noise ratio are manually picked on common shot gathers and times are verified for consistency using the reciprocity rules of Ackermann et al. (1986). Then, they are processed by using travel-time tomography (Iyer and Hirahara, 1993). Theoretical travel times are carried out through the ray-tracing method, based on Huygens principle (Hayashi and Takahashi, 2001), starting from an initial P-wave velocity model. This method was developed to applicate the seismic refraction tomography, where surface topography and subsurface geological structure are complex and are present strong lateral velocity variations.

The discrete form of i-th travel time computation is expressed as follows:

$$t_i = s_1 l_{i1} + s_2 l_{i2} + s_3 l_{i3} + \quad + s_N l_{iN} \quad \text{Eq. 13}$$

where

s_i is the slowness (i.e. the inverse of velocity) for the i-th cell

l_{ij} is the raypath for each pair of source receiver

For M pairs of source – receiver and in matrix notation is:

$$LS = \begin{bmatrix} l_{11} & \cdots & l_{1N} \\ \vdots & \ddots & \vdots \\ l_{M1} & \cdots & l_{MN} \end{bmatrix} \begin{pmatrix} s_1 \\ s_N \end{pmatrix} = \begin{pmatrix} t_1 \\ t_M \end{pmatrix} = T \qquad \text{Eq. 14}$$

The initial model consists of a layered and quadrangular cell with the velocity is increasing with the depth. As in Figure 28 at the boundary of cells are placed nodes and rays are lines connecting these nodes and the path is the shortest possible (Moser, 1991) for each couple of source-receiver.
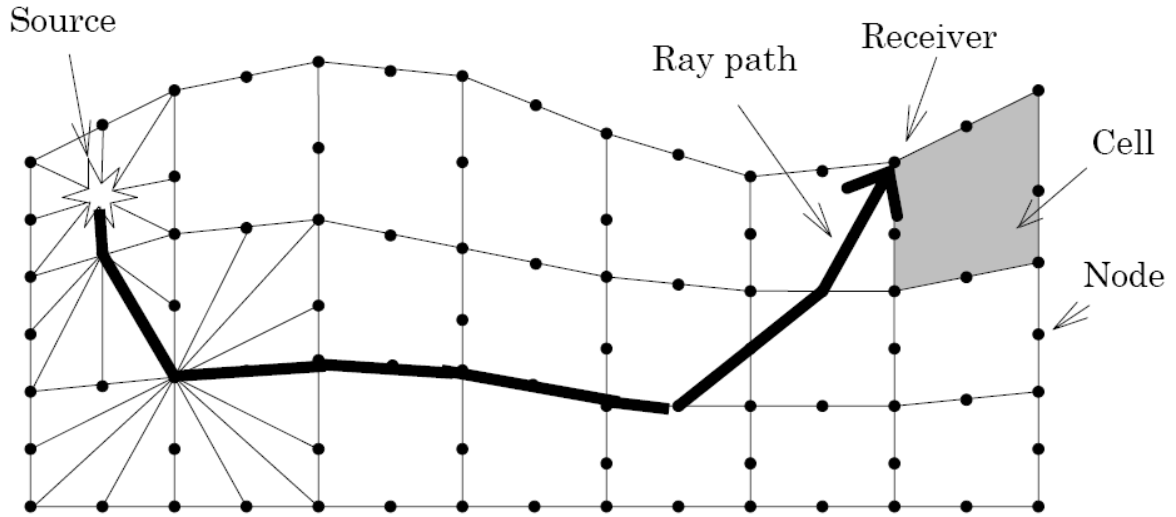
Figure 28 – Finite element model with nodes and ray paths of the seismic signal from source to receiver.

Based on such a velocity model, the aim of the inversion method is find the minimum travel time between source and receiver for each source-receiver pair. Thus, travel time residuals (i.e. observed minus theoretical) are minimized by upgrading the velocity model with an iterative nonlinear least-square method based on the simultaneous iterative reconstruction technique (Gilbert, 1972).

The nonlinear least square inversion works as in the following steps:

1. Calculate theoretical value Y0 for the initial value X0

$$Y_0 (Z) = Y (Z, X_0) \qquad \text{Eq. 15}$$

2. Calculate residuals ΔY between theoretical value Y_0 and observed Y

$$\Delta Y = Y - Y_0 \qquad \text{Eq. 16}$$

3. Calculate the correction value for $X(\Delta Y)$ by the least square method

$$RMSE_0 = \sqrt{\frac{\Delta Y^T_0 \, \Delta Y_0}{M}} \qquad \text{Eq. 17}$$

$$(A^T A) \, \Delta X = A^T \, \Delta Y \qquad \text{Eq. 18}$$

4. Calculate the new estimated value for X1

$$X_1 = X_0 - \Delta X \qquad \text{Eq. 19}$$

5. Iteration restarts from the step 1

The tomographic inversion stops when the residual error is close to an acceptable value, i.e. 1/8 of the dominant period of the P-pulse (Maraio et al., 2018).
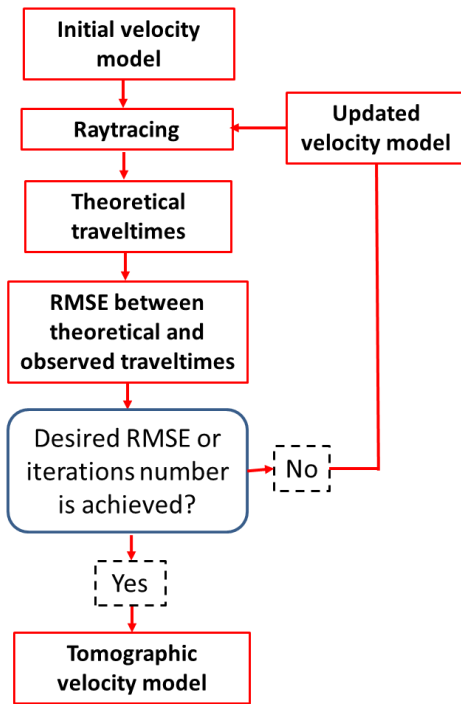
Figure 29 – Flow diagram of the seismic refraction tomography inversion.

3.4 Seismic attributes from P – waves seismic reflection

Seismic reflection surveying may offer a high level of resolution at depth and the ability to map detailed structures, which cannot be matched with any other method. Seismic reflection is able to recover information that may contribute to revealing aquifer geometry and system fluxes paths. As for the other geophysical method applied in the exploration phase, seismic reflection can be used to infer borehole informations. The use of high resolution seismic reflection by the groundwater research is improving, thanks to the high capacity to provide clear images of the main features related to the aquifers (i.e. mainly distinguish lithologies and fractured/faulted zones). As seen in Martin et al., (2013) the high resolution near surface seismic images are also able to clearly discern sedimentary packages within the major aquifer providing a higher level of understanding of groundwater resource.

Seismic facies analysis is a powerful qualitative technique applied for stratigraphic analysis on seismic data and reservoir characterization (West et al., 2002). Seismic facies can be distinguished through several characteristics of the seismic signal such as amplitude, continuity, geometry and frequency, which are strictly dependent from the subsoil.

Seismic attributes, powerful to reveal facies, are a quantitative measure of seismic characteristic of interest (Chopra et al. 2005), useful for interpretation of geological structures, stratigraphy and rock/pore fluid properties. The above geological parameters are strictly correlated with seismic data, because of the reflectivity of the earth depends on geometrical and physical features of subsoil. Dipping strata, impedance contrast due to lithology variation or textural variations, intended as coarse grains or fine laminated depositions are well imaged by seismic reflection.

Attributes evolved from 1930s to nowadays thanks to the increase of computational powerful and highlight different characteristics of the seismic data, depending on the calculation and the parameter took into consideration. In a historical review, Taner et al. (1994) classify attributes in two groups, geometrical and physical. Under geometrical group, they include dip, azimuth, and continuity, who reveal geometric features of seismic section and into the physical group parameters as amplitude, phase and frequency. In literature one can find many others way of classification as prestack or poststack, in time or depth, or as in Liner et al. (2004) in general and specific groups.

Seismic attributes are largely applied in geological reservoir characterization. Their use increased also in ultra-shallow high-resolution surveys such as GPR, prevalently for archeological and geological characterization (e.g. Forte et al., 2012; Zhao et al., 2013; dos Reis et al., 2014). In an analogous manner, an aquifer can be considered a porous or fractured media with textural and structural characteristics different from the background.

In this optic the use of attributes, from high resolution and shallow seismic dataset, can be applied to the imaging of aquifers.

In this work are used several attributes with the aim to reveal geological features as lithological variations, fractures and stratigraphy. To this end, were taken into consideration texture attributes, similarity, energy, instantaneous frequency and dip, because of their attitude to reveal the expected geological and structural pattern.

3.4.1   Texture attributes

Texture analysis of seismic data was introduced by Love and Simaan (1984) to reveal patterns of common seismic signal character. This inspiration came from the suggestion that zones of common signal character are related to the geologic environment in which their constituents were deposited (Chopra et al., 2006).

Seismic texture analysis is a quantitative way to describe and classify the seismic image, based upon a multitrace analysis, in which is effectively quantified the spatial distribution of seismic reflectors.

The best way to produce texture attributes was find recently, by using statistical measures to classify seismic textures using grey-level co-occurrence matrices (GLCMs) (West et al, 2002; Gao, 2003). The GLCMs essentially represent the joint probability of occurrence of grey-levels for pixels with a given spatial relationship in a defined region. Then the normalized GLCM matrix is a kind of a probability matrix that tells us how probable it is to find pairs of neighboring amplitudes in the area (volume) around the evaluation point. The GLCMs are then used to generate statistical measures of properties like coarseness, contrast and homogeneity of seismic textures, which are useful in the interpretation of oil and gas anomalies (Chopra et al., 2006).

In OpendTect the GLCM is computed on re-scaled data. The input data is re-scaled linearly to 4-bits (GLCM 16 x 16), or to 5-bits (GLCM matrix 32 x 32) by giving the clipping range of the input data. Neighbors then are compared in the inline and cross-line directions.
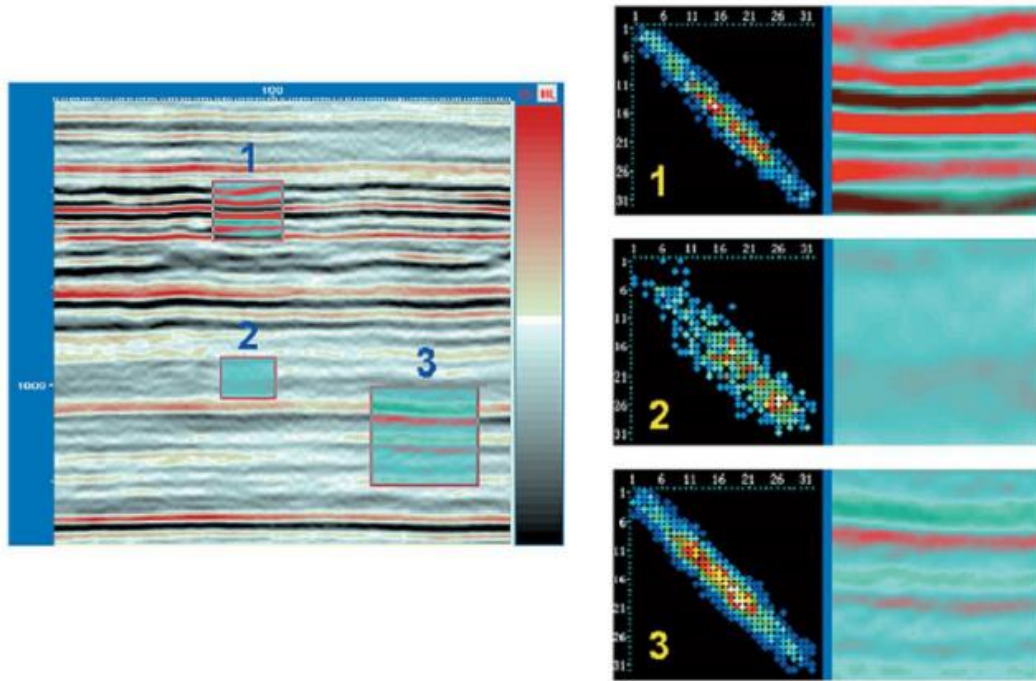
Figure 30 – Computation of GLCMs for different reflection characteristics. The matrix size chosen for computation is 32. Into the three different windows the results of GLCMs computation (from Chopra et al., 2006).

In Figure 30 one can see the GLCM computation for 3 different region of the image where GLCMs are shown to the right. In general, for strong continuous reflections, the GLCM exhibits a tight distribution along the diagonal. Low amplitude regions (or transparent) exhibit values near the center. Discontinuous or incoherent reflections have more occurrences farther away from the diagonal. Window 2 has lower amplitudes as well as incoherent reflections and so the GLCM shows a scatter about the diagonal.

In OpendTect there are 3 groups of texture attributes: contrast group, orderliness group and the last using descriptive statistics of the GLCM texture measures.

In all equations given below N denotes the size of the GLCM matrix; i refers to the column and j to the row. P is the GLCM Probability matrix.

Contrast

$$\text{Contrast} = \sum_{i,j=0}^{N} P_{i,j}(i-j)^2$$

Eq. 20

When i and j are equal, the cell is on the diagonal and (i-j) = 0. These values represent amplitudes entirely similar to their neighbor, so they are given a weight of 0. If i and j differ by 1, there is a small contrast, and the weight is 1. If i and j differ by 2, contrast is increasing and the weight is 4. The weights continue to increase exponentially as (i-j) increases. Contrast is a measure of local variations within neighbor pixel and high values are related with low homogeneity of the image.

58

Dissimilarity

$$\text{Dissimilarity} = \sum_{i,j=0}^{N-1} P_{i,j} |i-j|$$

Eq. 21

In Dissimilarity the weights with which GLCM probabilities are multiplied, increase linearly away from the diagonal (along which neighboring values are equal).

Homogeneity

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$$

Eq. 22

Dissimilarity and Contrast result in larger numbers for more contrasting windows. If weights decrease away from the diagonal, the result will be larger for input areas (volumes) with little contrast. Homogeneity weights values by the inverse of the Contrast weight, with weights decreasing exponentially away from the diagonal. Homogeneity is a measure of similarity of pixels and can be used to quantify continuity of reflectors.

Angular Second Moment and Energy

$$\text{ASM} = \sum_{i,j=0}^{N-1} P_{i,j}^2 \qquad \text{Energy} = \sqrt{\text{ASM}}$$

Eq. 23

ASM and Energy use the GLCM probability as a weight for itself. The name for ASM comes from Physics, and reflects the similar form of Physics equations used to calculate the angular second moment, a measure of rotational acceleration. High values of ASM or Energy occur when the input area (volume) is very orderly, as index of textural uniformity and can be useful to highlight geometry and continuity (Chopra et al., 2006).

Entropy

$$\text{Entropy} = \sum_{i,j=0}^{N-1} P_{i,j} \left(-lnP_{i,j}\right)$$

Eq. 24

Entropy is the opposite of energy; it is a measure of chaos. In physics (thermodynamics) entropy refers to the quantity of energy that is permanently loss to heat ("chaos") every time a reaction or a physical transformation occurs. Entropy cannot be recovered to do useful work. Because of this, the term is used in non-technical speech to mean irremediable chaos or disorder. High values of entropy occur for images texturally not uniform.

GLCM Mean

$$\text{GLCM mean } \mu_i = \sum_{i,j=0}^{N-1} i\left(P_{i,j}\right) = \mu_j = \sum_{i,j=0}^{N-1} j\left(P_{i,j}\right)$$

<div align="right">Eq. 25</div>

The left hand equation calculates the mean based on the reference pixels, i. The right-hand equation calculates the mean over the neighbor pixels, j. These two values are identical because OpendTect computes a symmetrical GLCM, where each amplitude is counted once as a reference and once as a neighbor.

GLCM Variance

$$\text{Variance } \sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j}\left(i - \mu_i\right)^2 = \sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j}\left(j - \mu_j\right)^2$$

<div align="right">Eq. 26</div>

Variance is a measure of the dispersion of the values around the mean. It is similar to entropy. It answers the question "What is the dispersion of the difference between the reference and the neighbor pixels in this input area (volume)?" GLCM Variance in texture measures performs the same task as does the common descriptive statistic called variance. It relies on the mean, and the dispersion around the mean, of cell values within the GLCM. However, GLCM variance uses the GLCM, therefore it deals specifically with the dispersion around the mean of combinations of reference and neighbor amplitudes, so it is not the same as variance of input amplitudes that can be computed with the "Volume Statistics" attribute.

GLCM Standard Deviation

$$\text{Standard Deviation } \sigma_i = \sqrt{\sigma_i^2} = \sigma_j = \sqrt{\sigma_j^2}$$

<div align="right">Eq. 27</div>

There is no particular advantage to using Standard Deviation over Variance, other than a different range of values.

GLCM Correlation

$$\text{GLCM Correlation} = \sum_{i,j=0}^{N-1} \left[\frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$$

<div align="right">Eq. 28</div>

The Correlation texture measures the linear dependency of input amplitudes on those of neighboring amplitudes. GLCM Correlation is quite a different calculation from the other texture measures described above. As a result, it is independent of them (gives different information) and can often be used profitably in combination with another texture measure. It also has a more intuitive meaning to the actual calculated values: 0 is

uncorrelated, 1 is perfectly correlated. GLCM Correlation can be calculated for successively larger window sizes. The window size at which the GLCM Correlation value declines suddenly may be taken as one definition of the size of definable objects within an image. If the input is completely uniform the GLCM variance is 0 and the correlation function is undefined. OpendTect will in that case return the value 1.

3.4.2 Geometrical and physical attributes

Similarity, dip, energy are attributes useful to delineate structural features and variations both vertical and lateral connected to the seismic amplitude.

Similarity is an attribute that returns trace-to-trace similarity properties as continuity of reflector and can highlight well interruption of reflectors. Similarity is a measure of coherency and shows how much waveforms within traces in the gate are similar. A value of 1 means that segments within the gate are identical, on the other hand 0 means they are completely dissimilar. In OpendTect, the calculation works in this way: first, try to find the direction of best match at every position, known as Dip, and then using the dip to calculate the best similarity between adjacent traces.

Dip is an attribute that indicate the dip of maximum coherency: its use reveals bedding of strata and discontinuities.

Energy is calculated as the squared sum of the sample values in the specified time-gate divided by the number of samples in the gate. The energy is a measure of reflectivity in the specified time-gate, higher the energy, higher the amplitude. This attribute enhances lateral and vertical variations related to the amplitude of the seismic signal and can reveal continuity (or discontinuity) of geological features. This attribute may be a proxy of major lithology changes, unconformities, gas and fluid accumulation.

Instantaneous attributes represents instantaneous variation of various parameters for the seismic complex signal computed sample-by-sample and returns such values for each sample location. These attributes can be collectively termed instantaneous attributes since they concisely and quantitatively describe the seismic waveform (or character) at any sample point (White, 1991). In particular, instantaneous frequency can be applied because of its effective discrimination powerful related to both, wave propagation effects and depositional characteristics. Higher frequencies can be indicator of sharp interfaces originated by thin laminated soils and lower frequencies are indicative of more massive bedding geometries. Thus, this attribute can be used as discriminator of lithology variation due to its capacity of distinguish sand rich regions (low frequencies) from shale rich (high frequency) (Chopra et al., 2005). As seen in Moon et al. (1993) low values of frequency are produced also by intense fracturing zones.

4 Multivariate analyses as integration methods

The need to obtain a detailed hydrogeological characterization of the subsurface and its interpretation for the groundwater resources management, often requires to apply several and complementary geophysical methods (Bedrosian et al., 2007). As seen in chapter 3, different materials show different geophysical responses, due to mineral type content, grain packing arrangement, porosity, permeability, and pore content (i.e. gas or liquid type). The use of several geophysical methods allows in theory to obtain complementary information about the subsurface conditions due to the different sensitivity of the geophysical methods to different parameters such as P- and S-wave velocities, bulk electrical conductivity, density, and etcetera. A co-located knowledge of different parameters can reveal geological features well constrained by the use of different geophysical parameters. The goal of the approach in this thesis is to provide a unique geological model of the aquifer by synthesizing and optimizing the information provided by several geophysical methods. This approach greatly reduces the degree of uncertainty and subjectivity of the interpretation by exploiting the different physical and mechanic characteristics of the aquifer. Due to the above reasons, the integrated use of several geophysical methods, can lead to a well-constrained and unique conceptual model that shows various geo-facies conversely the univariate way. In such a way, the effectiveness of the individual geophysical methods can be improved to jointly determine the Earth structures from their multiple physical properties (Bosch, 1999; Bosch et al., 2002).

Multivariate approaches, cluster analysis techniques included, are powerful data-mining tools to reveal relationships among various petrophysical parameters (e.g., Gill et al., 1993; Barrash and Morin, 1997; Bosch et al., 2002). One of the strength of these approaches is the capability to classify data in similar groups and to find correlation without any prior relationship among variables. Due to this, cluster techniques are widely used for the earth sciences (e.g., Dumay and Fournier, 1988; Gill et al., 1993; Fechner and Dietrich, 1997; Dietrich et al., 1998; Hammah and Curran, 1998; Bosch et al., 2002; Güler et al., 2002; de Matos et al., 2006; Paasche et al. 2006; Roden et al., 2015; Di Giuseppe et al., 2017).

Several attempts for a quantitative joint interpretation methodology were based on a 2-D cluster recognition of the seismic and resistivity from Magnetotelluric models (e.g. Bauer et al., 2003; Haberland et al. 2003; Bedrosian et al. 2004; Maercklin et al. 2005). Furthermore, Bedrosian et al. (2007) showed a quantitative approach based on probability density function, to highlight lithological variations, independently of any prior relation linking electrical and seismic parameters.

The main concept of cluster analysis is to group multivariate dataset in a multidimensional space on the basis of distances (i.e., statistical measure of similarity) among observations. In this way, starting from different geophysical parameters, one can

obtain clusters of data points where variables are similar on the joint parameter space (i.e. the space composed by multivariate dataset). As clustering method I applied k-means algorithm and Silhouette index (SI) to assess the validity of the clustering. Moreover, the dissimilarity (or similarity) calculated through SI has been used also as uncertainty to image integrated tomography in the original space (i.e. 2D section).

If the clustering methods find their best application in low dimensionality multivariate datasets, conversely don't work well in high dimensionality datasets. On the other hand, there are many application of geophysical data integration with large number of variables, developed by using techniques of dimensionality-reduction widely applied as Principal component Analysis and Artificial Neural Network.

Exploratory analysis, where visualization plays a very important role, has become more and more difficult given the increasing dimensionality of the data. There is a real need for methods that provide meaningful mappings into two dimensions, so that we can fully utilize the pattern recognition capabilities of our own brains. There are many approaches to mapping a high dimensional data set into two dimensions, of which principal component analysis (PCA, Jackson 1991) is probably the most used followed from the Self-Organizing Maps (SOM) (Kohonen 2001). Furthermore, SOMs have seen many diverse applications in a broad range of fields, such as medicine, biology, chemistry, image analysis, speech recognition, engineering, computer science and geophysics.

Seismic facies analysis can be accomplished through the use of pattern recognition techniques. Given the appropriate combination of seismic attributes, one can identify lateral and vertical changes in the reservoir. When a priori geological information is missing, seismic facies analysis is called unsupervised and is performed through unsupervised learning or clustering algorithms (Duda and Hart, 2001).

In this thesis, data analysis was performed using Matlab for integration of tomographic dataset and R for integration of seismic attributes. In following paragraphs will be explained parameters strictly relative to the two programming languages above mentioned, as applied into the developed codes.

4.1   Integration of tomographic data set through k-means algorithm

The proposed integration method is based on the recognition of similarities on a geophysical multivariate dataset, where 2D co-located tomograms are the different variables. In this way, one can identify discrete zones of the subsurface with similar geophysical response, which can be called "geophysical facies".

The multivariate dataset can be divided in groups, using clustering algorithms, producing partitions in which dissimilarities among observations are minimal and are maximal among clusters. 2D independent and complementary tomograms constitute the multivariate dataset (i.e. bulk resistivity, induced polarization, P and S velocity waves, etc.).

This method allows integrating dataset during post-processing phase, without any prior constrain among geophysical parameters, often site-dependent and not easy to find out, with a high impact on results. In this manner, the proposed integration approach can be applied in different geological settings, starting from the most suitable geophysical parameters, useful to highlight the desired target.

For a successful approach to this work, basic steps have to be take into consideration (Fayyad et al., 1996):

- Feature selection: to choose geophysical variables useful to reveal attended results in cluster analysis;
- Clustering algorithm: to select a clustering algorithm which lead to a partitioning that fits well the dataset;
- Validation of the results: as the clustering number is not known a priori, is necessary assess the optimal solution using an adequate criteria;
- Interpretation of the results: to reach the right conclusion based on clustering results and other meaningful evidence.

Following the clustering approach from Di Giuseppe et al., (2014), in this thesis I applied k-means method to reveal similarities in bivariate geophysical datasets composed of electrical resistivity and P-waves seismic refraction tomography.

A distinction can be made between hierarchical and partitioning clustering approaches (i. e. k-means). Hierarchical techniques produce a tree of nestled clusters starting to group similar single observations, whereas partitioning techniques produce homogeneous groups for a predetermined number of clusters, with the possibility for an instance to change clusters at any re-computation of centroids. Partitioning techniques are known to be less susceptible to outliers and to be computationally more efficient than hierarchical methods (Tronicke et al., 2004). Because of its conceptual simplicity and

algorithmic robustness, the k-means approach is one of the most popular and widely used clustering techniques.

In literature, one can find several integration of tomographic dataset, based on k-means algorithms. Tronicke et al. (2004) integrated cross hole georadar velocities and attenuation tomography to characterize heterogeneous alluvial aquifers using k-means cluster analysis, with effectiveness in imaging hydrostratigraphic zonation. Di Giuseppe et al. showed the effectiveness of k-means algorithm to reveal different geostructural zones over a faulted zone for a bivariate shallow dataset (2014) and for a trivariate deep dataset (2017).

Electrical resistivity tomography (ERT) and seismic refraction tomography (SRT) have been largely applied for geological exploration, more in specific for groundwater research and their reliability has been proved across last decades in many different situations. Tomographic methods are powerful techniques to highlight lithological and indirectly hydrogeological (intended as grain size basin deposits, fracturation of rocks and saturation) variations, both in vertical and horizontal directions. The co-located use of ERT and SRT increases the knowledge of physical condition of subsoil as grain size and saturation: if is true that saturation degree can vary the conductivity for a fixed lithology, it is also true that reference values for saturated-unsaturated subsoil have a wide range. Fractured-saturated rocks generally have low resistivity values and this can lead to a dubious interpretation: clayey soil or fractured rocks or clayey rocks? By knowing also the velocity value, the subsurface interpretation becomes more constrained in a reduced range of lithologies (Figure 31).
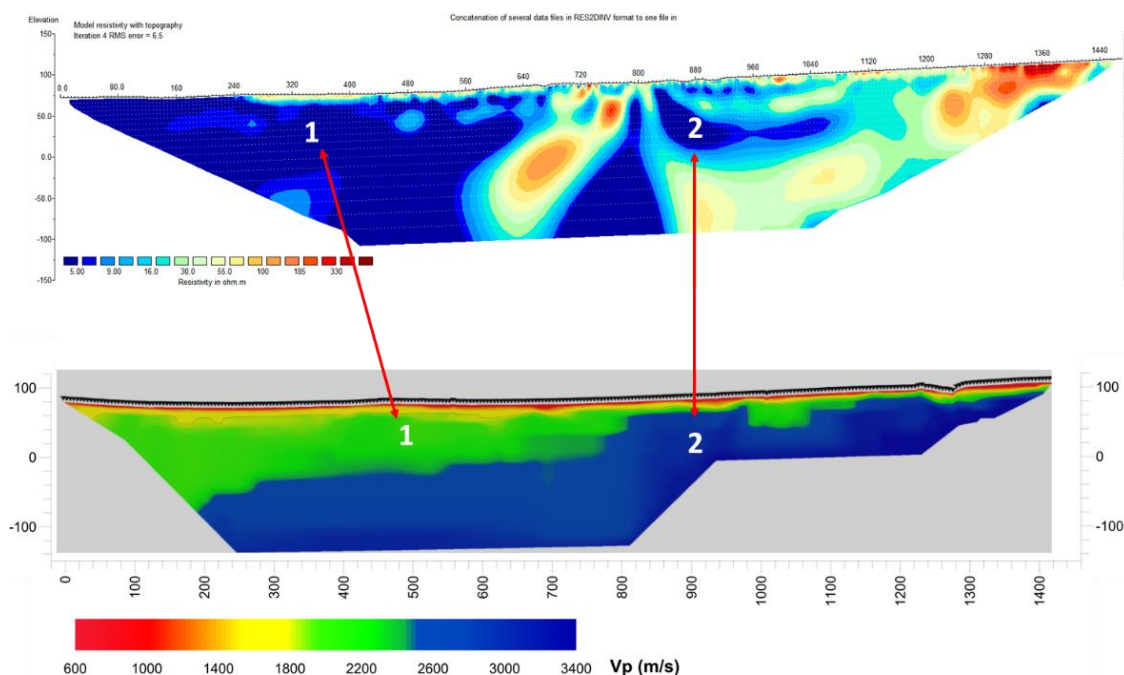
Figure 31 – A comparison between co-located ERT (above) and SRT (below), where zones 1 and 2 have the same resistivity but different velocity. Zone 1 is representative of clay and zone 2 of clayey marls.

Thus, the coupled use of these two geophysical parameters reduces the degree of uncertainty in interpretation and give a clearer description of the subsoil condition. The resolution of the integrated tomography is directly related to the resolution of input dataset.

In order to analyze similarities within the bivariate dataset it occurs that grid data point be co-located. Starting from the geoelectrical resistivity and seismic refraction models acquired along the same line, with their respective spatial discretization and generally different, to obtain co-located grids an initial procedure of homogenization is required. To achieve this, data must be interpolated over the same grid spacing.

Usually, electrical tomographies have a non-uniform distribution, which progressively coarseness with the depth, instead of seismic tomographies which have a regular grid with the depth (Figure 32). Bedrosian et al., (2007), integrating a dataset constituted by magnetotelluric and seismic 2D images, interpolated both models with grid dimensions from the more resoluted, in order to avoid a loss of data and preserving the information content of both models. Here with the same aim to preserve informations and evaluate possible noise effects from the interpolation, results of the integration are compared method as follows: in the first attempt both 2D models are interpolated with the grid spacing from the lowest resoluted; in the second attempt 2D models are interpolate with a more resoluted common grid spacing (Figure 32). The two interpolated datasets are then integrated and compared by using the Silhouette results: one can consider as acceptable the grid that doesn't vary the cluster distribution with respect to the lower grid spacing. In this manner is possible to preserve the best resolution into the integrated image.
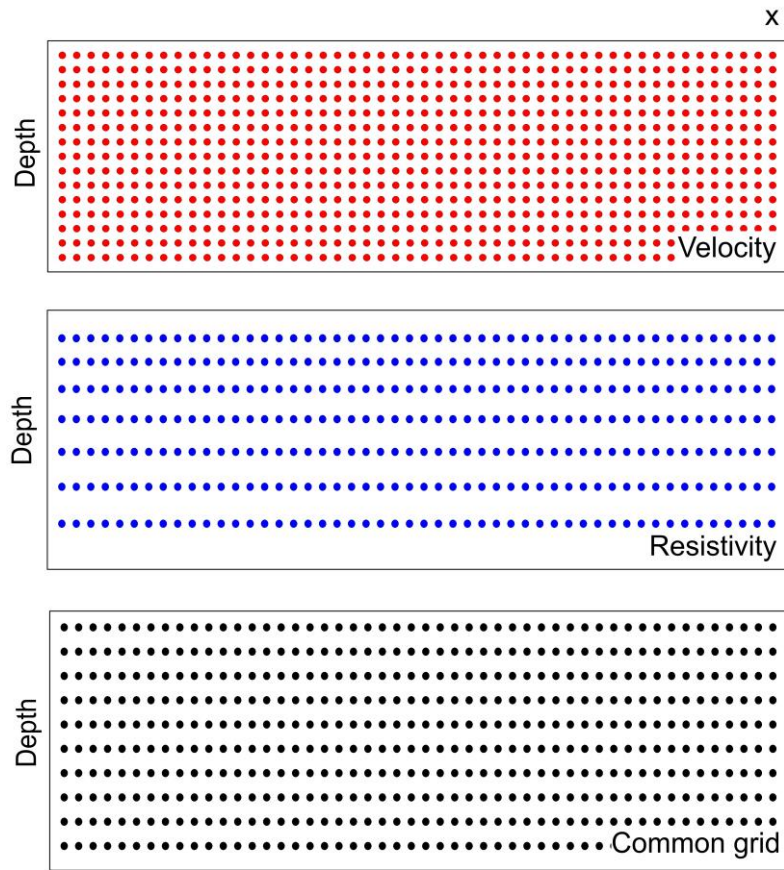
Figure 32 – Different grid spacing for velocity model, electrical resistivity model and common grid for the co-locating interpolation process.

The integration code to produce clusters from the bivariate dataset, is based upon k-means algorithm (Lloyd, 1982), a partitive and unsupervised method to find similar observations in data exploratory analysis.

K-means is a partitional clustering algorithm, non-hierarchical and unsupervised, which allows separating a dataset in k cluster based on distances among points. The objective function of the algorithm is to minimize within cluster variance and to maximize it among clusters.
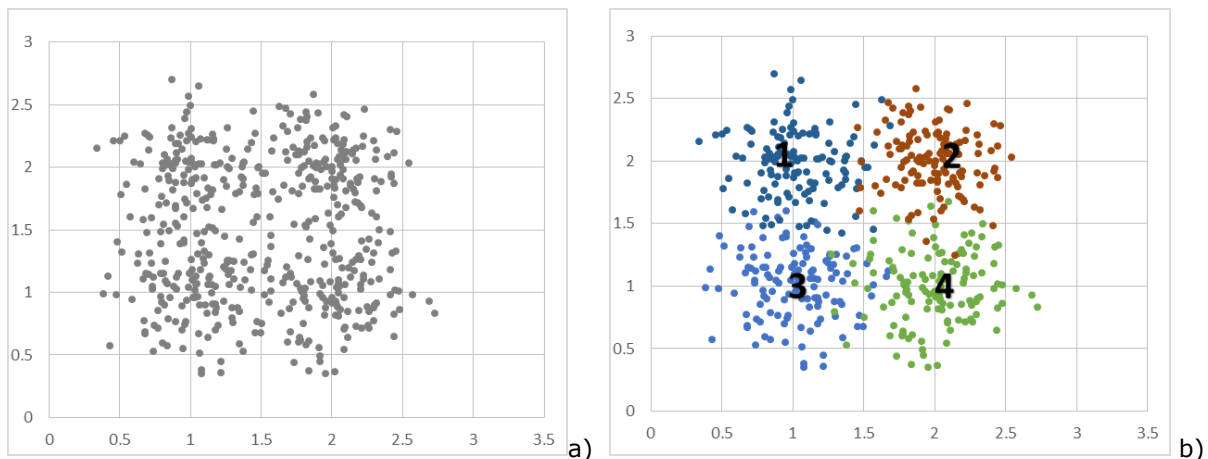
Figure 33 – In a) a synthetic dataset with four clusters and in b) a clustering analysis performed by k-means.

The algorithm is initialized by defining in random way (not ever) k centroids and assigning to the j-th centroid, samples closer (assignment step). The second step (update step) consists in recalculate mean values intended as new k – centroids. The process to re-assign observations to the k centroids is reiterated with centroids, which the mean values are previously calculated. Iterations stop when the within cluster variance reaches minim values and maximum intra-cluster values. Distances in p-dimensional space between observations, used for minimization, can be calculated with several method as below:

- *Squared Euclidean distance* (default). Each centroid is the mean of the points in that cluster

$$d(x,c)=(x-c)(x-c)' \quad \text{Eq. 29}$$

- *Sum of absolute differences*, i.e., the $L1$ distance. Each centroid is the component-wise median of the points in that cluster.

$$d(x,c)= \sum_{j=1}^{p} |xj - cj| \quad \text{Eq. 30}$$

- One minus the *cosine* of the included angle between points (treated as vectors). Each centroid is the mean of the points in that cluster, after normalizing those points to unit Euclidean length.

$$d(x,c) = 1 - \frac{xc'}{\sqrt{(xx')(cc')}} \qquad \text{Eq. 31}$$

- One minus the sample *correlation* between points (treated as sequences of values). Each centroid is the component-wise mean of the points in that cluster, after centering and normalizing those points to zero mean and unit standard deviation.

$$d(x,c) = 1 - \frac{(x - \vec{x})(c - \vec{c})'}{\sqrt{(x - \vec{x})(x - \vec{x})'}\sqrt{(c - \vec{c})(c - \vec{c})'}} \qquad \text{Eq. 32}$$

where:

x is an observation

c is a centroid

p is the number of variables

$$\vec{x} = \frac{1}{p}\left(\sum_{j=1}^{p} xj\right) \vec{1p}$$

$$\vec{c} = \frac{1}{p}\left(\sum_{j=1}^{p} cj\right) \vec{1p}$$

$\vec{1p}$ is a row vector of p ones.

In Matlab ambient, above distances are called respectively: *sqeuclidean*, *cityblock*, *cosine* and *correlation*. In this thesis, the distance is calculated as the square of Euclidean distance because is the most widely employed for this kind of clustering.

One of the major weakness of the k-means is the possibility to converge in local minima instead of global and leading to an apparent optimal clustering. The solution is influenced from the initial centroids. Instead to define in random way initial centroids, with the option "Start" and "Cluster" one can launch a preliminary clustering analysis on a 10 % subsample of the dataset. The individuated centroids will be the starting values for the complete clustering. Coupled with this, it is possible to repeat several times the algorithm to avoid local minima with the option "Replicates" and the relative number. In the integration code the value of 8 was individuated after several trials as the minimum number to obtain the same results.

The number of clusters, k, is an input value for this algorithm and is unknown; the solution is to individuate a quite wide range (i.e. bigger than the presumed number of cluster) of k values for the dataset and to produce corresponding clustering analyses. Thanks to several validation methods, based generally on distances among observation

within cluster and intra-cluster, one can assess the right assignment of the observation into the clusters.

K-means is sensitive to the data scale and to avoid this it's important to have the starting dataset with similar ranges for the variables, because this can affect results and leading to not well partitioned clusters. In order to do this, a crucial step is the "data sphering" (Martinez & Martinez, 2005), to remove dominance effects in the clustering due to the different variances of the data.

4.1.2 Data sphering

A preliminary data processing is necessary before to apply any process of data mining because of different variances of geophysical variables. The multivariate dataset is usually normalized because one of the standard deviation dominate in the clustering procedure. This step called "data sphering" is based upon the calculation of covariance matrix.

Starting from the dataset one can calculate average values for the multivariate dataset, the covariance matrix and its eigenvalues and eigenvectors. The dataset transformation is achieved by using this formula (Martinez & Martinez, 2005):

$$Z_i = \Delta^{-0.5} Q^T (x_i - \bar{x}), \; i = 1, \ldots, N \qquad \text{Eq. 33}$$

where Q are eigenvectors for the covariance matrix, Δ is the diagonal matrix of eigenvalues, x_i is the i-th sample for any geophysical variable and $\bar{x}$ is its average value.

Z_i is the new scaled multivariate dataset suitable as input for data mining processing.

4.1.3 Validation of k-means clustering

Since clustering algorithms discover clusters, which are not known a-priori, the final partitioning of a dataset requires some sort of evaluation in most applications.

Any assessing tool in clustering validation has a series of issues to resolve, which can be resumed in: 1) quality of clusters, 2) the degree with which a clustering scheme fits a specific data set, 3) the optimal number of clusters in a partitioning (Halkidi et al., 2001).

As seen in the previous paragraph, the optimal number of clusters, k, is unknown at the beginning and is an input value to start the analysis. Due to this, one proceed doing analysis for several K values, from two to a number sufficiently greater than attended geofacies and then evaluate results for every k analyzed. There are several methods to validate clustering results based on different considerations that can individuate the optimal number of clusters. In Figure 34 appears so clear that the optimal clustering for the example is three; however choosing the wrong parameters for the clustering algorithm leads to a non-optimal clustering, i.e. wrong solution.
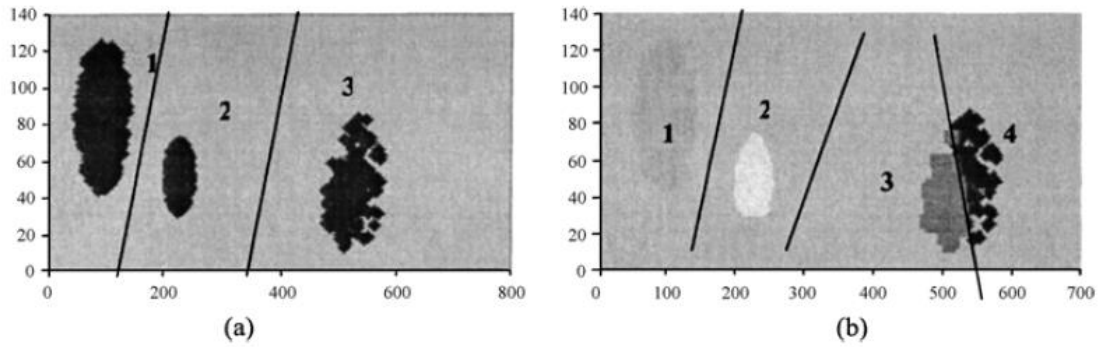
Figure 34 – (a) A dataset that consists of 3 clusters and (b) the result from the application of k-means when we ask four clusters (from Halkidi et al., 2001).

In this work the Silhouette index (SI) (Rousseeuw, 1987), was applied. SI compares distances of every i-th observation within cluster with average extra-cluster distances (Figure 35). In this way, every observation assigned in a cluster has a value of Silhouette index, that assess how well is placed any observation within the cluster, compared to the samples into the other clusters. On other words, the silhouette value measures the degree of confidence in the clustering assignment of a particular observation (Brock et al., March 2008).
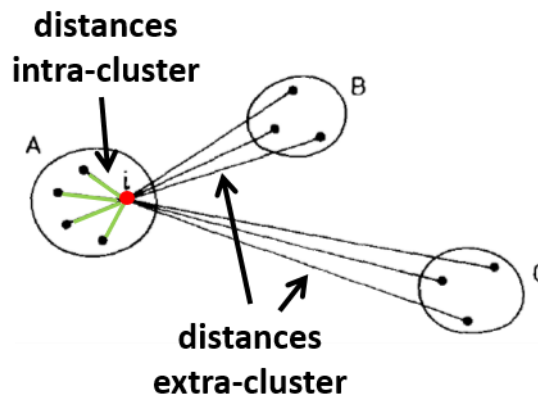
Figure 35 – Intra-cluster and extra-cluster distances for the i-th observation in cluster A.

$$s(i) = \begin{cases} 1 - a(i)/b(i) & \text{if } a(i) < b(i), \\ 0 & \text{if } a(i) = b(i), \\ b(i)/a(i) - 1 & \text{if } a(i) > b(i). \end{cases}$$

Eq. 34

The SI ranges between -1 and 1: value greater than 0 and close to 1 means observation well-matched to its own cluster, at values near to 0 corresponds observation with not clear assignment and SI values <0 are typical of misclustered observations. One obtains the optimal clustering for the k value, which gives the maximum value of SI.

72

Its assessment of compactness and separation conforms literally to the stated goal of clustering; a relatively small a_k compared to b_k means that x_k has been successfully grouped with its similar points in the same cluster in a way that separates from its dissimilar points (Rawashdeh and Ralescu, 2012).
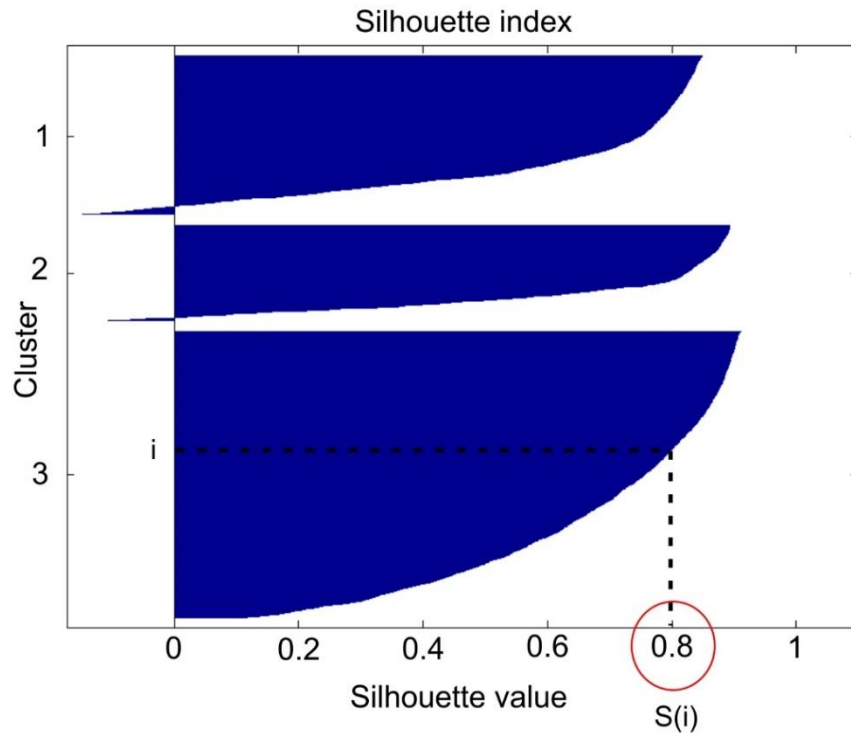
Figure 36 – Graphical Silhouette representation for 3 clusters: in the red circle the Silhouette value for the i-th observation within the cluster 3.

Compared with R-Squared index (Sharma S., 1996) or others (Halkidi et al., 2001), based on distances within and intra cluster, which reveal the optimal clustering at the "knee", SI gives less chances to choose the wrong number of cluster: the optimal k is where SI reaches the maximum value. Find the "knee" is not so obvious and easy as it appears (Figure 37) and to choose the wrong number of clusters can lead to an unsuccessful analysis. The difficult increases with not well-separated clusters, as in the case of tomographic multivariate dataset, where images have variation in a smoothed way.
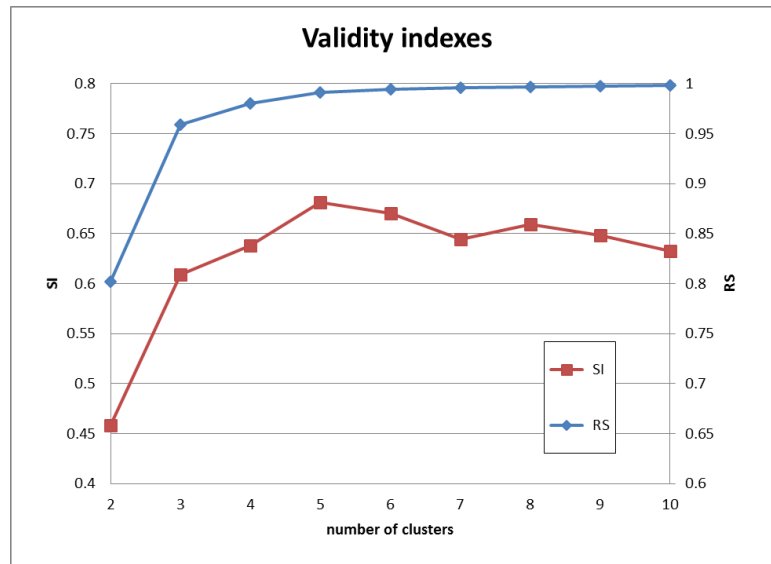
Figure 37 – Comparison between R squared (in blue) and Silhouette index (in red): the optimal number of clusters for the Silhouette index is univocally individuated with the maximum value in k=5 instead of the R squared index.

When lithofacies information is available, the number of clusters (i.e. geophysical facies) should be close to the number of different lithologies present, which can be compared with the attended optimal cluster number. In this way this information can help to individuate the optimal partitioning and to avoid wrong solutions.

4.1.4   Results and graphical restitution

The SI can be also interpreted as a measure of membership for a particular observation assigned in a cluster. Observations far from to the centroid and close to the borders of any cluster have a high degree of uncertainty about their positioning, with a SI close to 0 or lower. A hard clustering representation in the original 2D space (i.e., tomographic space) shows clusters with sharp border, intrinsic characteristic of k-means method. In a similar manner that in Paasche et al., (2010), where the saturation of integrated images was based on membership function, here is applied SI as saturation because of its capability to exhibit the membership of each observation.

In a well-separated original data in multivariate space, with visible natural clusters, the attended degree of uncertainty in clustering for observations is lower than in a contiguous dataset. However, tomographic imaging have as intrinsic characteristic a continuous variation of the values, changing in a smoothed way. The increasing or decreasing of any geophysical parameter depends mainly on physical variation into the subsoil, geophysical technique, inversion method and its resolution. Thus, tomographic data vary without discontinuity and this leads to multivariate dataset with clusters not well separated as expected. In order to limitate this drawback, SI can be used as saturation value to reveal zones with uncertainty about clustering, in a similar way to the soft clustering (Paasche et al., 2010); this approach can reduce known limits of k-means about ambiguity of clustered data at the borders and can highlight zones where geophysical

74

parameters are changing. This improvement can give more complete and realistic results instead of classical hard clustering imaging, showing geofacies where observations are well clustered, local dissimilarities and transition zones.

4.2   Integration of seismic attributes through Self Organizing Maps

The object of seismic interpretation is to extract all the geologic information possible from the data as it relates to structure, stratigraphy, rock properties, and perhaps reservoir fluid changes in space and time (Liner, 1999). More in specific, seismic attributes reveal features, relationships, and patterns in the seismic data that otherwise might not be noticed (Chopra and Marfurt, 2007).

Over the last decades artificial neural networks (ANNs) have been subject to an increasing interest in seismic attributes integration to discover patterns. These groups and patterns in the data, identified by the Self-Organizing Maps (SOM) analysis routinely, reveal geologic features important in the interpretation process (Roden et al., 2015).

Self-organizing maps, or Kohonen maps (Kohonen, 2001), belong to the category of competitive learning based clustering. The SOM is known to be one of the most important tools for the unsupervised seismic facies analysis (Coléou et al., 2003).

The SOM is an excellent tool in exploratory phase of data mining. It projects input space on prototypes of a low-dimensional regular grid that can be effectively utilized to visualize and explore properties of the data, i. e., a SOM groups together similar data in a manner that provides effective visualization of multivariate dataset. The mapping produced by SOM is usually many to one, the projection images on the SOM are local averages of the input data, comparable to the k-means averages in classical vector quantization (Gersho, 1982; Gray, 1984). As known from its implementation, SOM networks learn to cluster groups of similar input patterns from a high dimensional input space in a non-linear fashion onto a low dimensional (most commonly two-dimensional) discrete lattice of neurons in an output layer (Kohonen, 2001). This is achieved in manner that similarity relations into the output map represent similarity relations of input patterns, i.e., similar input seismic attributes are classified in neurons physically located close to each other (preserving topology). Thus, the output map can be used as a tool to reveal different groups, identificating various distinctive features (Taner et al., 2001).

SOM analyses are applied in many cases to facilitate the interpretation of seismic surveys and to predict seismic facies in terms of lithology (Taner et al., 2001) thanks their capacity to learn and recognize similar features. In this thesis I applied SOM analysis to individuate homogeneous geophysical facies based on seismic attributes, with the aim to highlight zones strictly related to geological features. More in specific, I used geometric, physics, textural and instantaneous seismic attributes to characterize 2D sections in terms

of lithology variations and discontinuities for hydrogeological purpose. The joint use of these attributes can be powerful to reveal both primary and secondary permeability thanks to the effectiveness to reveal lithology differences, as coarse or laminated deposits and main discontinuities as fractured/faulted zones in rocks. Because the SOM can be interpreted as mapping of the input n-dimensional space into a two dimensional grid that preserve the topological structure of original data space and because seismic attributes measure the changes in geology, SOM preserve the topological relation of the underlying geology (de Matos et al., 2006).

Before performing attribute integration a preliminary analysis is needed, for assessing the most representative variables into the multivariate dataset. His was done by using the Principal Component Analysis technique to compute the cross-correlation in a multi-dimensional dataset. Principal Component Analysis represents a rotation of the multi-dimensional point cloud so that the maximum variability is projected onto the pair-wise combination of axes (Prasad et al. 2005). The use of these types of multi-attribute classifications makes forecasting more reliable and much easier to interpret (Linari 2003).

4.2.1   Kohonen algorithm in R

The SOM is a type of neural network suitable for unsupervised learning (Kohonen, 1997). This is done in such a way that neurons physically located close to each other in the output layer of the SOM have similar input patterns (Kalteh et al., 2008).

Typically, SOM networks learn to cluster groups of similar input patterns from a high dimensional input space in a non-linear fashion onto a low dimensional (most commonly two-dimensional) discrete lattice of neurons in an output layer (Kohonen, 2001). SOMs combine competitive learning with dimensionality reduction by smoothing the clusters with respect to an a-priori grid. One of the main characteristics of these networks is the topological ordering property of the clusters generated. Clusters objects are mapped in neighboring regions of the grid, delivering an intuitive visual representation of the clustering. SOMs are reported to be robust and accurate with noisy data (Mangiameli et al., 1996). On the other hand, SOMs suffers from the same problems such as those of dynamical clustering: sensibility to the initial parameters settings and the possibility of getting trapped in local minimum solutions (Jain et al., 1999).

The process starts by assuming that input variables (i.e. seismic attributes), can be represented as vector in the space R^n, x = [x_1, x_2,…, x_n]. The objective of the algorithm is to organize the dataset of input seismic attributes in a geometric structure known as SOM. Each SOM unit, defined by a prototype vector, is connected with its neighborhood, which in 2D usually produces hexagonal or rectangular structure maps. One can assume that the map has Q elements and there are Q prototype vectors as w_i = [w_{i1}, w_{i2},…., w_{in}], with i =1,2,….,Q, where n is the number of seismic attributes (Figure 38). The algorithm starts

by assigning a so-called codebook vector to every unit, which will play the role of a typical pattern, a prototype, associated with that unit. Usually, one randomly assigns a subset of the data to the units. During training for each learning step, objects are repeatedly presented in random order to the map. The distances between x and all the prototype vectors are computed. The map unit with the smallest distance m_b to the input training vector is called best matching unit or "winning unit; a weighted average is used, where the weight of the new object is one of the training parameters of the SOM. The learning rate α, it is typically a small value in the order of 0.05. During training, this value decreases so that the map converges.
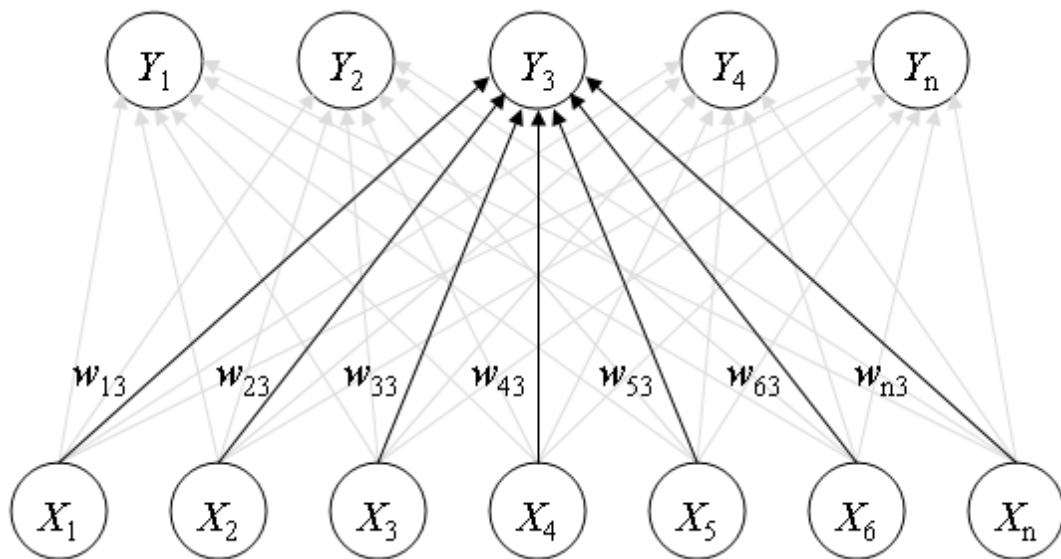
Figure 38 – Input and output layers with the indication of weights.

The spatial constraint lies in the fact that SOMs require neighboring units to have similar codebook vectors. This is achieved by not only updating the winning unit, but also the units in the immediate neighborhood of the winning unit, in the same way. The updating equation for the i-th unit is expressed as follows:

$$w_i(t+1)=w_i(t)+\alpha(t)h_{bi}(t)[x-w_i(t)]$$

where

t   is the iteration

α is the learning rate

h_{bi} is the neighborhood size centerd at the winner unit

The winning unit, i.e. the one most similar to the current training object, will be updated to become even more similar. The spatial constraint lies in the fact that SOMs require neighboring units to have similar codebook vectors. This is achieved by not only updating the winning unit, but also the units in the immediate neighborhood of the winning

77

unit in the same way. The size of the neighborhood decreases during training as well, so that eventually (in R implementation of The Kohonen package after one-third of the iterations) only the winning units are adapted (Wehrens and Buydens, 2007). (Figure 39). Thus, SOM prototype vectors are updated at each iteration. The algorithm terminates after a predefined number of iterations
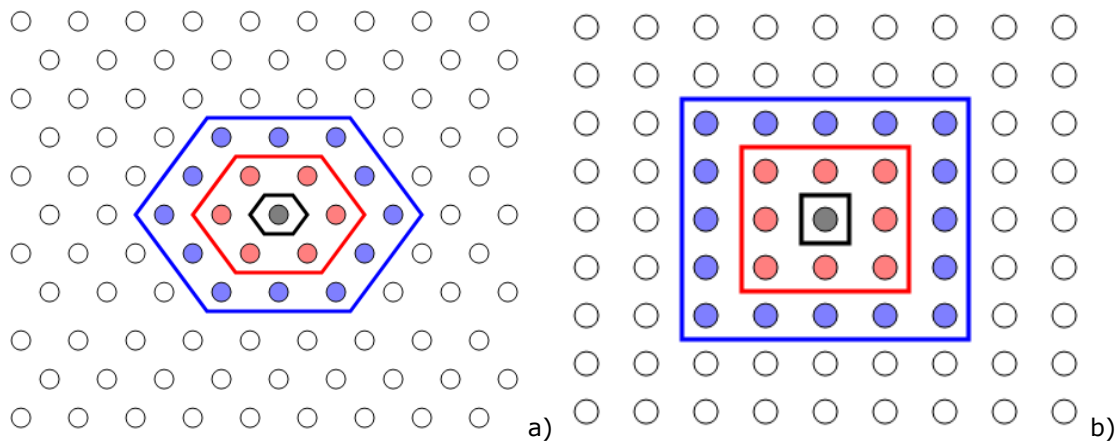
Figure 39 – Two different disposition of neurons with their respectively neighborhood highlighted with blue, red and black (winner unit) colors: in a) hexagonal configuration and in b) rectangular configuration.

Initialization parameters have their importance, because as seen above, SOM can get trapped in a local minimum solution. Even with identical settings, repeated training of a SOM will lead to sometimes even quite different mappings, because of the random initialization. However, the conclusions drawn from the map remain remarkably consistent, which makes it a very useful tool in many different circumstances (Wehrens et al., 2007).

Initialization parameters such as net (or grid) sizing and number of iterations are the first step for a successful SOM analysis. The main idea about sizing the SOM is: find it by trial–and-error. One may have to test several sizes of the SOM to check that the cluster structures are shown with a sufficient resolution and statistical accuracy (Kohonen, 2014). A common practice for defining the shape of the map is calculating the first two principal dimensions in which the variances of the dataset input are; the ratio between the number of neurons in the two directions of the grid is proportionated to the ratio between the two principal components (Kohonen, 2014). The principal component analysis is performed in R (Abdi and Williams, 2010; Kassambara, 2017). The correlation between a variable and a principal component (PC) is used as the coordinates of the variable on the PC. The representation of variables differs from the plot of the observations: the observations are represented by their projections, but the variables are represented by their correlations (Abdi and Williams, 2010). Positively correlated variables are grouped together. Negatively correlated variables are positioned on opposite sides of the plot origin (opposed quadrants). The distance between variables and the origin measures the quality of the variables on the factor map. Variables that are away from the origin are well represented

78

on the factor map. The quality of representation of the variables on factor map is called "cos2" (i.e. square cosine, squared coordinates). A high cos2 indicates a good representation of the variable on the principal component and the variable is close to the circumference of the correlation circle. A low cos2 indicates that the variable is not perfectly represented by the PCs and the variable is close to the center of the circle (Abdi and Williams, 2010). As recommended by Vesanto and Alhoniemi, 2000, the grid should have a number of nodes well above the number of real cluster in the dataset.

The number of iterations should be elevate enough to permit at the algorithm finding an absolute minimum solution. In R thanks to the Kohonen package, training progress can be inspected by using "changes" output, which describes the average distance of an object with the closest codebook vector (Wehrens et al., 2007).

The "som" function has several input parameters as follows with the indication of default values:

- grid: the rectangular or hexagonal grid of units. The format is the one returned by the function somgrid from the class package.
- rlen: the numer of iterations, i.e. the number of times the data set will be presented to the map. The default is 100;
- alpha: the learning rate, determining the size of the adjustments during training. The decrease is linear, and default values are to start from 0.05 and to stop at 0.01;
- radius: the initial size of the neighborhood, by default chosen in such a way that two-thirds of all distances of the map units fall inside this number. The size of the neighborhood decreases linearly during training; after one-third of the iterations only the winning unit is being adapted and the algorithm corresponds to k-means.
- init: optional matrix of codebook vectors. If it is not given, randomly selected objects from the data are used. This feature can be useful when re-training a map with new data.
- toroidal: by default, FALSE. If TRUE, the edges of the map are not real edges, and data are actually mapped to a torus. Put differently: opposite map edges are joined together.
- KeepData: default value equals TRUE. However, for large data sets it may be too expensive to keep the data in the som object, and one may set this parameter to FALSE.

4.2.2   *Results and mapping*

As stated above, SOM analysis preserves topological disposition of the input dataset. The main result of a SOM analysis is the classification of similar observations within the neurons.

A trained SOM can be seen as an elastic net that folds over the input data provided. Similar input samples are classified in neurons physically located close to each other (preserving topology), thus, closer neurons represent similar values of the original multivariate dataset. Consequently, in order to display the multivariate dataset into the original space (seismic section), the best solution is a multidimensional color map based on the neuron position. In this thesis is proposed a 3D color map based on the two

dimensions of the net and Umatrix as values for RGB index. The first and second dimensions of the net, scaled into the range from 0 to 255, are respectively R and G and the Umatrix value is the B value. Thus, the RGB color map take into consideration the real distribution of neurons and shows similar neurons with similar colors. Moreover, the Umatrix value is useful to highlight groups of observations, into the multidimensional input space, far from the others (Figure 40).
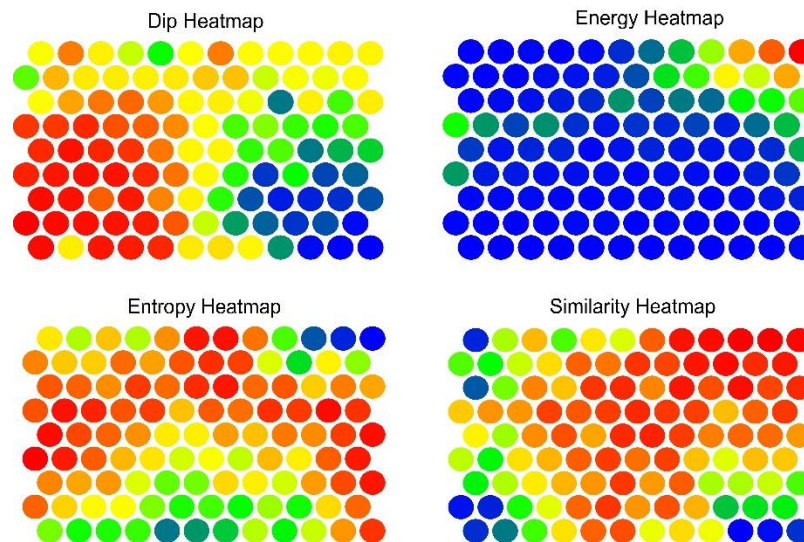
Figure 40 – Distribution of average attribute values on a trained SOM for four different attributes: cold colors represent low values and warm colors represent high values. The distribution of classified observations on neurons appears clear, e.g. in the right upper corner are classified observations with high energy and similarity, low entropy and medium dip.

Heatmaps are a way to represent the average values of input observations classified into each neuron. These are powerful to show 2D distribution of different variables across the map, facilitating a visual interpretation of SOM classification. Usually the variation range of the variables is showed by a gradational color scale, with cold colors for the lowest value (i.e. blue) and hot colors for the highest value (i.e. red) (Figure 40). The heatmaps are produced by averaging variables classified into each neuron.

The most used method to visualize the cluster structure of a SOM are based on distance matrix techniques (Ultsch, 1993), in particular the unified distances matrix (Umatrix) (Figure 43). This item is a visualization technique that allows to represent structural properties of the high dimensional input space. The Umatrix shows distances between prototype vectors of neighboring map units and the winner unit (Figure 41); resulting values can be interpreted as a "digital elevation model": there are valleys where input vectors are quite similar and hills where input vectors are dissimilar (Figure 42).
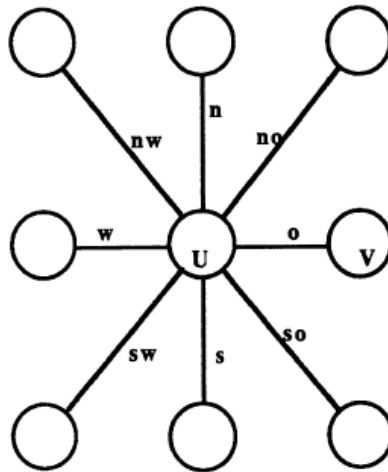
Figure 41 – The winner neuron "U" and its and its immediate neighbors (from Ultsch, 1993).
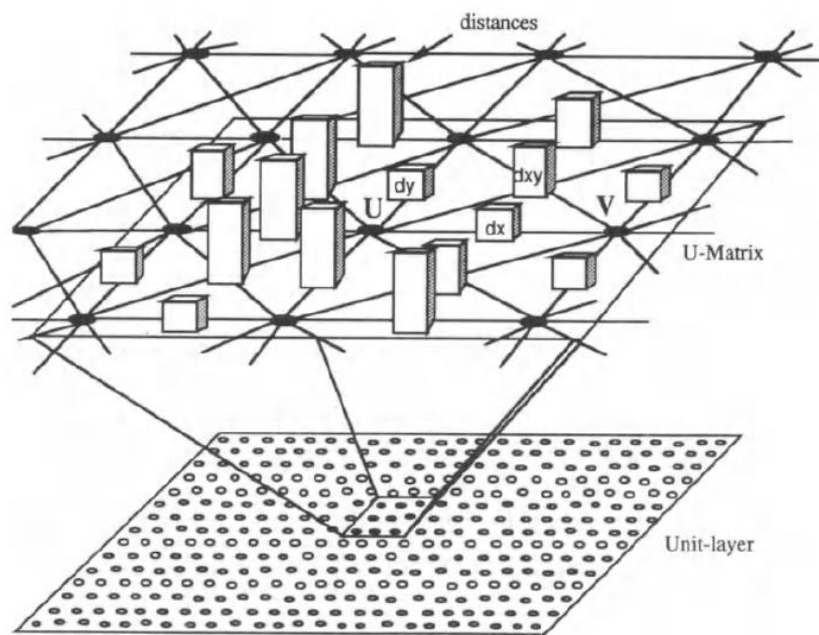
Figure 42 – A representation of neighborhood distances between the winning neuron and its immediate neighbors where distances can be interpreted as a digital elevation model.
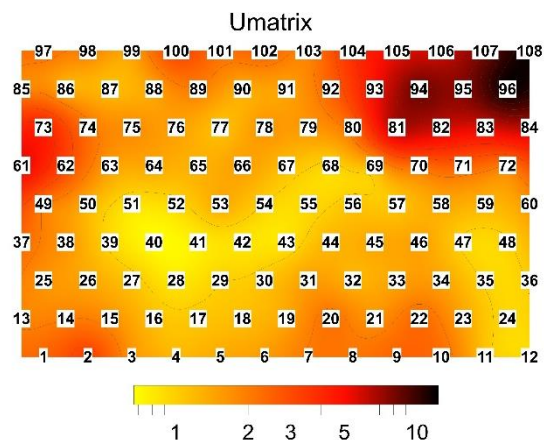
Figure 43 – U-matrix for a trained SOM: yellow color represents closest neurons and black furthest neurons that reflect distances among observations in 2D projected data space.

5   Shallow aquifer characterization: Laterina

The study area, into the municipality of Laterina (Arezzo, Italy), is a shallow basin filled by lacustrine and alluvial deposits (Pleistocene and Holocene epochs, Quaternary period), with alternated silt, sand with variable content of pebbles and clay where the bottom is represented by arenaceous-pelitic rocks (Monte Cervarola Unit, Tuscan Domain, Miocene epoch). This shallow basin constitutes the unconfined superficial aquifer to be exploited and the fractured substrate can constitute the confined aquifer. To improve the geological model obtained from a detailed geological survey were performed electrical resistivity and P wave refraction/reflection tomographies along the same line in order to obtain different, independent and complementary datasets.

5.1   Geological setting

The study area is located at NE of Arezzo (Tuscany, Italy) in the southern part of the upper Valdarno Valley between the Prataomagno massif and Chianti Mountains. The Valdarno valley is an asymmetrical half graben, formed during an extensional phase of the Northern Apennines (Carmignani and Kligfield, 1990; Boccaletti et al., 1990; Sagri et al., 2004) which produced low- and high-angle normal faults, oriented about NW/SE. The basin is filled up by Plio-Pleistocene sediments (Figure 44; Boccaletti et al., 1996).
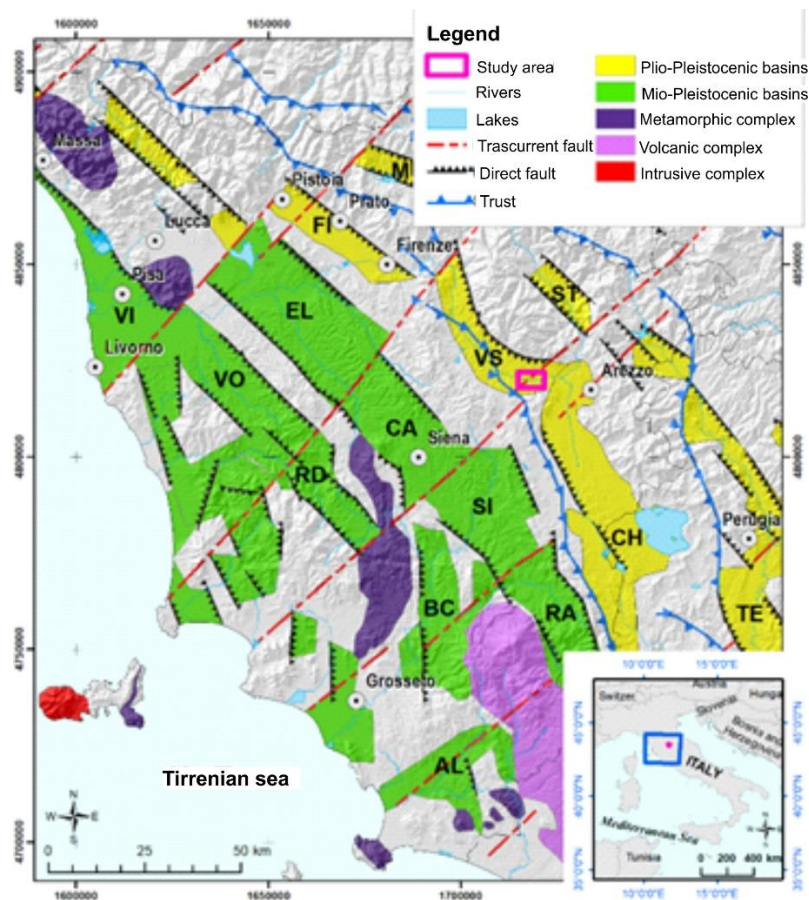
Figure 44 - Simplified tectonic map of Tuscany where located the Neogenic and Quaternary basins of the Northern Apennine: AL Albegna, BC Baccinello, CA Casino, CH Chiana, CT Casentino; EL Elsa, FI Firenze, MU Mugello, RA Radicofani, RD Radicondoli, SI Siena, TE Tiberino, VI Viareggio, VO Volterra, VS Valdarno Superiore (modified from Martini and Sagri,1993).

The hills of the study area are characterized by fluvial-lacustrine deposits of Villafranchian age, belonging to the "Ciuffenna River Sub-synthem" associated to the third depositional phase of upper Valdarno basin. This Sub-synthem is made up of "Pian di Tegna silts", "Levane sands" and "Laterina conglomerate". The "Laterina conglomerate" is mainly formed by rounded pebbles of limestone and lenses of sands and are in unconformity above the "Montevarchi Sub-synthem" or on the "Monte Falterona" Fm (Figure 45).
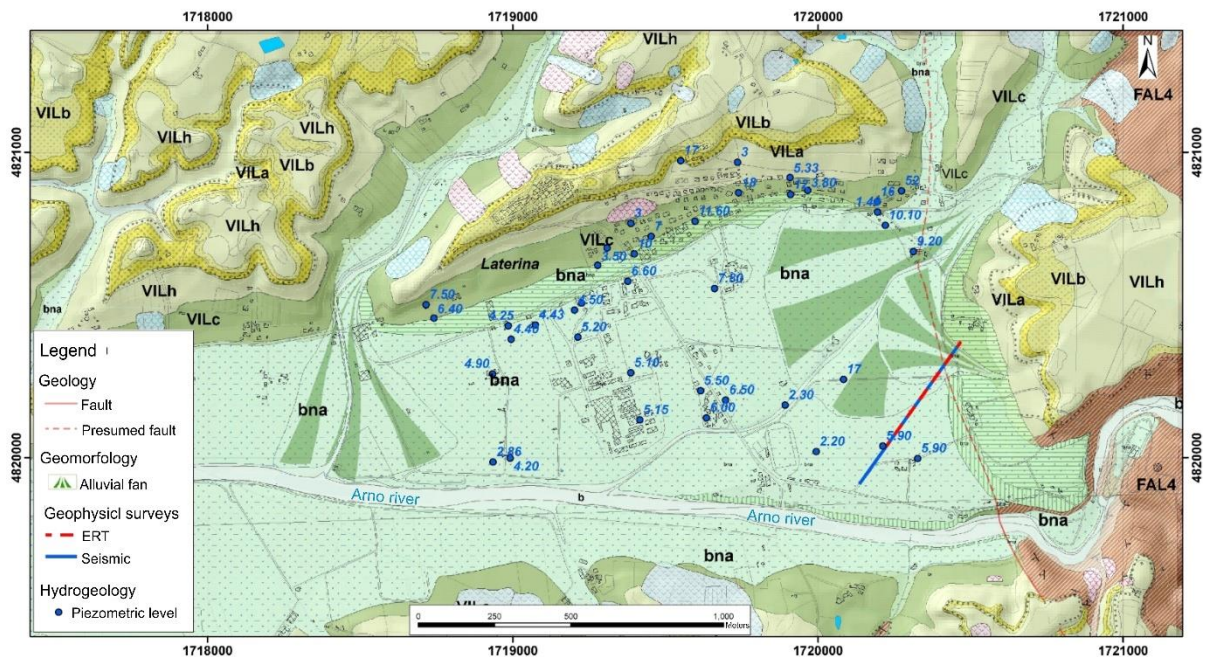
Figure 45 - Geological map of the investigated area with the location of geophysical surveys. b= actual alluvial deposits; bna=recent deposits; VILa, b, c, h= Villafranchian fluvial-lacustrine deposits; FAL4=Tectonic Unit Cervarola –Falterona. Dashed in red the integrated line and in blue the seismic line. Points in blue are representative of the groundwater depth levels.

The "Montevarchi Sub-synthem", associated to the second depositional phase of upper Valdarno valley, is formed by "Oreno River silt" of the Lower Pleistocene age, interbedded with layers of sand/silty-sand or gray/blue-clays and locally with pebbles lenses. Beneath the "Oreno River silt" are present the "Ascione River Clays", of Lower Pleistocene age, characterized by clays, silty-clays and clayey-sands with plant residues. Locally are present peaty clays and layers of lignite. There are frequently intercalations of sands and sands with pebbles.

The bedrock is represented by the "Lonanno Member" (Aquitanian-Burdigalian age) belonging to the Monte Cervarola Unit, formed by fine-grained sandstone and decimetric levels of siltstone, marly-siltstone and shale (Figure 46 and Figure 47). Locally are presents layers of coarser sandstone with a thickness of 1.5 m.

Figure 46 - Alternation of decimetric layers of sandstone, siltstone and shale belonging to the "Loanno member" of Monte Cervarola Unit.

Figure 47 - Sandstone and siltstone, belonging to the "Loanno member" of Monte Cervarola Unit, cropping out along the Bregine River.

In the Northeast and South part of the study area the Monte Cervarola Unit is covered, with tectonic boundary, by the "Monte Senario Breccia" belonging to the Canetolo Unit. This formation is mainly formed by Eocenic – Oligocenic calcarenite and shale. In the eastern part of the study area the geological fieldworks highlights an high-angle-normal fault with strike N/NW-S/SE and dip direction toward W, connected with the extensional phase of the Northern Apennines, which formed the upper Valdarno basin. Evidences of this structure, cataclasites and fault breccias (Figure 48) are present in the Southeast part of the Laterina basin.

Figure 48 - Cataclastic layer in sandstone and siltstone, belonging to the "Loanno member" of Monte Cervarola Unit.

Into the valley are present alluvial continental Quaternary deposits, mainly formed by sand and silt with lenses of pebbles. In the East part of the valley, are present two alluvial fans, formed by the Bregine River (Figure 45) and other minor rivers. Sediments of these alluvial fans are mainly formed by silty-sands with lenses of pebbles. (Figure 49, Figure 50 and Figure 51).

Figure 49 - Polygenic gravel with rounded pebbles in sandy matrix. Are well evident embricate structures.

Figure 50 - Pluridecimetric level of gravel in coarser sand deposit.

Figure 51 - Panoramic view on the alluvial fan deposit of Bregine River. Are evident lenses of gravel alternated with coarser and fine sand.

Figure 52 – Clayey silts of Villafranchian age outcropping at the base of alluvial fan deposits.

In summary, the geological model purposed for the study area can be depicted in three stratigraphic units (Figure 45):

- Holocene alluvial deposits made by sand and gravel for a thickness ranging from 5 to 20 m (label "bna" in Figure 45). Into the studied area are present two main alluvial fans, where the principal was formed by Bregine River and other minor rivers. This deposit host a probable shallower and unconfined aquifer.

- Villafranchian silt and gray-blue clay with plant rests belonging to the "Ciuffenna River Sub-synthem" (Lower- Middle Pleistocene age), show a thickness variable from 10 to 50 m (VILa, VILb, VILc, VILh, in Figure 45). From a hydrogeological point of view this is the main aquitard between the shallow and the deep aquifer.
- Miocene sandstone, belonging to the "Monte Cervarola" Unit, represent the bedrock of the studied area (FAL4, Figure 45). The bedrock is formed by fine-grained sandstone and siltite alternated with marly-pelitic layers. Locally are present coarse sandstone layers with metric thickness. Fractured portions of "Monte Cervarola" Unit represents a conspicuous deep aquifer.

In the study area is presumed the presence of a shallow aquifer system constituted by the alluvial deposits, mainly sand and pebbles, with a thickness ranging from few meters to 20 meters and are locally present suspended and secondary aquifers. The recharge of unconfined aquifer is due to prevalently rainfalls and runoff and is influenced by seasons and human activities in volume and quality. A more stable and disposable groundwater resource can be represented by the deep and confined aquifer in fractured "Monte Cervarola" Unit. The spatial extension and tectonic structures presents on this formation lead to a high productive confined aquifer. The groundwater monitoring showed variable values of the water depth from surface across the valley, with values ranging from 2.2 m to 17 m indicating the presence of several aquifers in agreement with the above mentioned hypothesis (Figure 45).
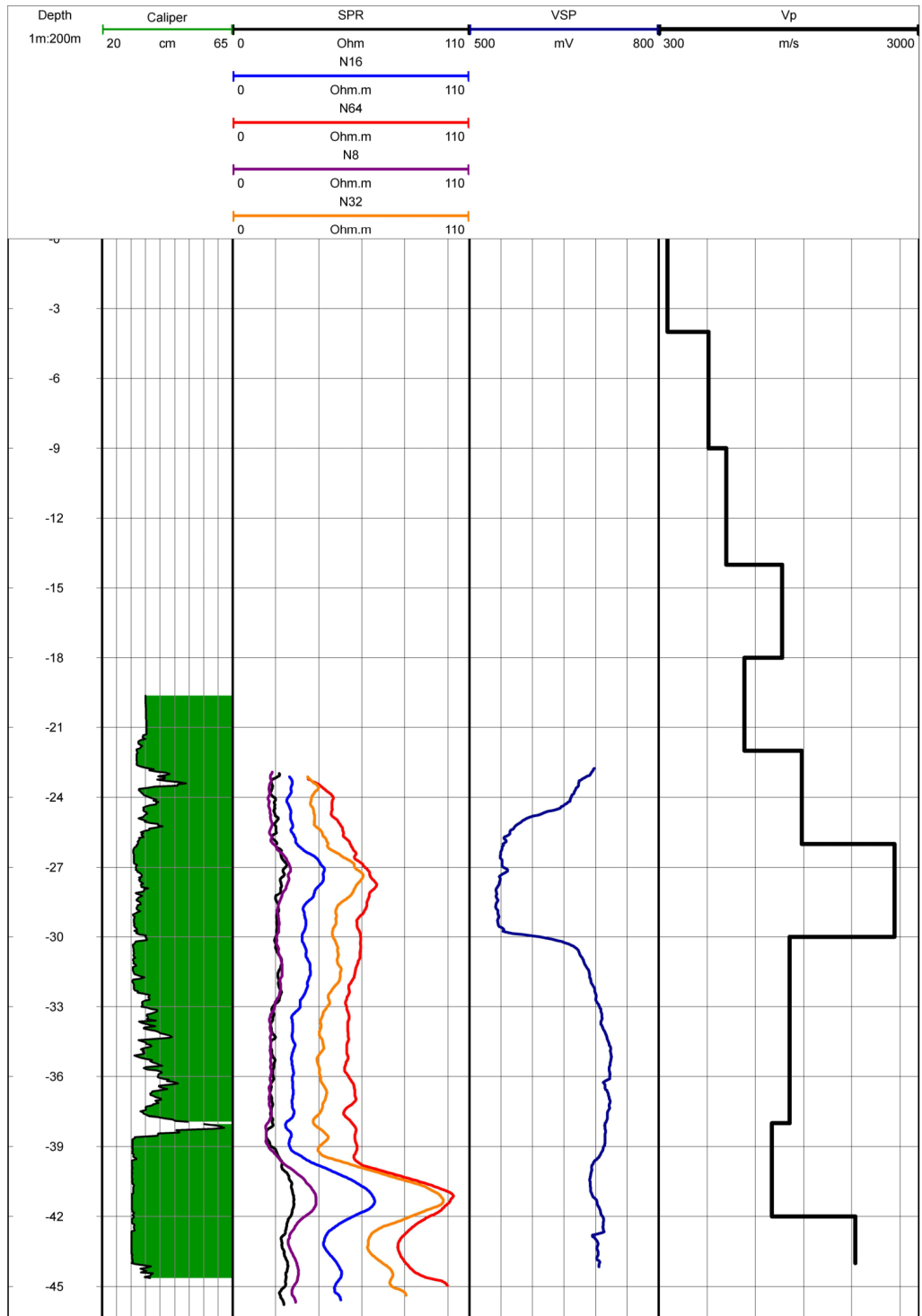
Figure 53 – Geophysical logs of the borehole "Laterina 1" located as in Figure 45. From the depth of 9 m to 18 m the well is cased.

5.2 Geophysical data set

The aim of the geophysical surveys was revealing main geological features, in which to discriminate aquifers from the aquicludes on alluvial deposits (i.e. bna) and the shallow part of the bedrock (i.e. Monte Cervarola unit). As stated into the § 3, in surface survey planning, the first objects to take into account are the dimensions of the target (thickness, lateral extension and depth) and the contrast of its physical properties compared to the background. In particular, the preliminary hydrogeological surveys (i.e. groundwater monitoring) coupled with previous geophysical information across the basin, showed the presence of a multi-level groundwater. These evidences, together with the geological/geomorphological knowledge from fieldworks highlighted the chance of a complex environment, with lateral variations of coarse sediments (i.e. gravel and pebble) locally trapped in silty/clayey soil and an articulated bedrock at an estimated depth, ranging from 20 to more than 60 m. Thus, the right approach was to take into consideration both vertical and lateral variations with an adequate resolution for the target.

Co-located ERT in dc current and SRT were acquired along a profile, marked in Figure 54, for a length of 425 m across the more vocated area for the shallow and the deeper aquifer. Seismic acquisition was extended up to about 560 m for a better exploration of the basin (Figure 54).
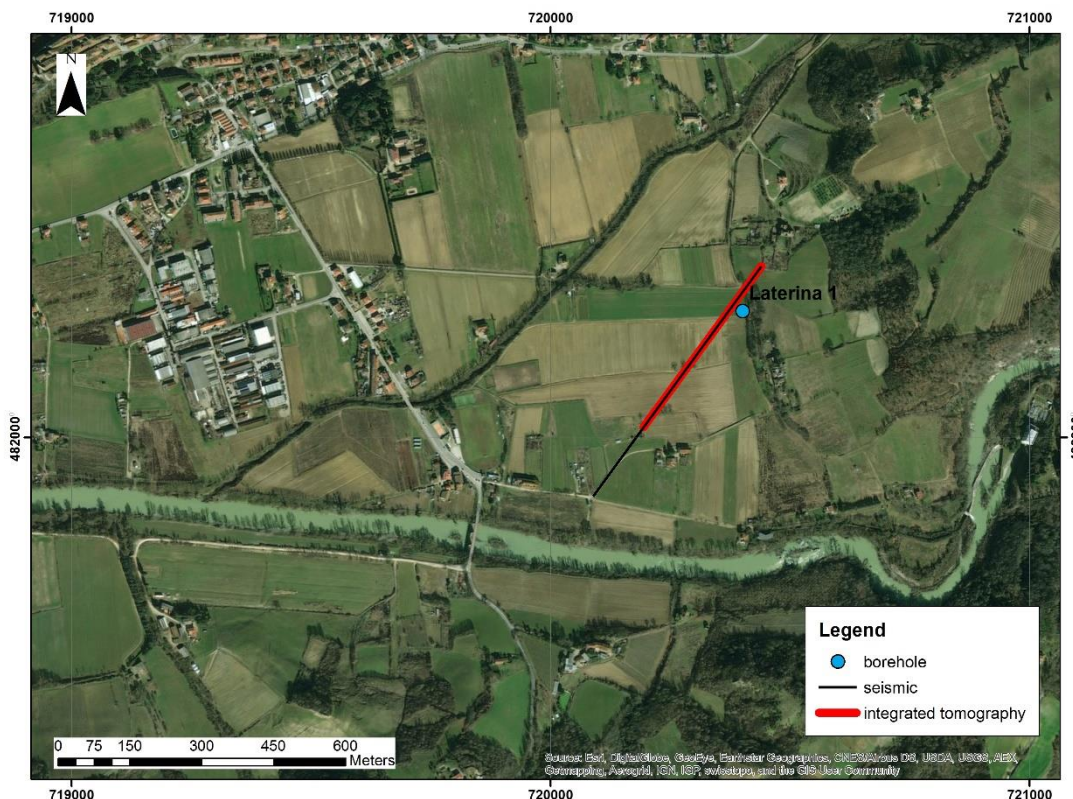
Figure 54 – Location of geophysical surveys and the borehole "Laterina 1" on satellite imagery: in red the integrated line, in black seismic survey.

5.2.1 ERT

DC current resistivity dataset was acquired with a georesistivimeter IRIS Syscal Pro Switch, with 96 electrodes and an electrodic step of 3 m, for a total length of 285 m. The choice of the electrodic step was a compromise to have an adequate solution for the shallow part of the target (i.e. shallower aquifer) (Villafranchian fluvial-lacustrine deposits) and for exploring enough in depth the upper part of the sandstone/marl bedrock (Monte Cervarola Unit). In order to fully explore horizontally the total length of about 420 m, was applied the roll-along technique (Loke, 2004): the base array was moved forward of 24 electrodes for two times and data was acquired just for the remaining part (Figure 55) at each movement.
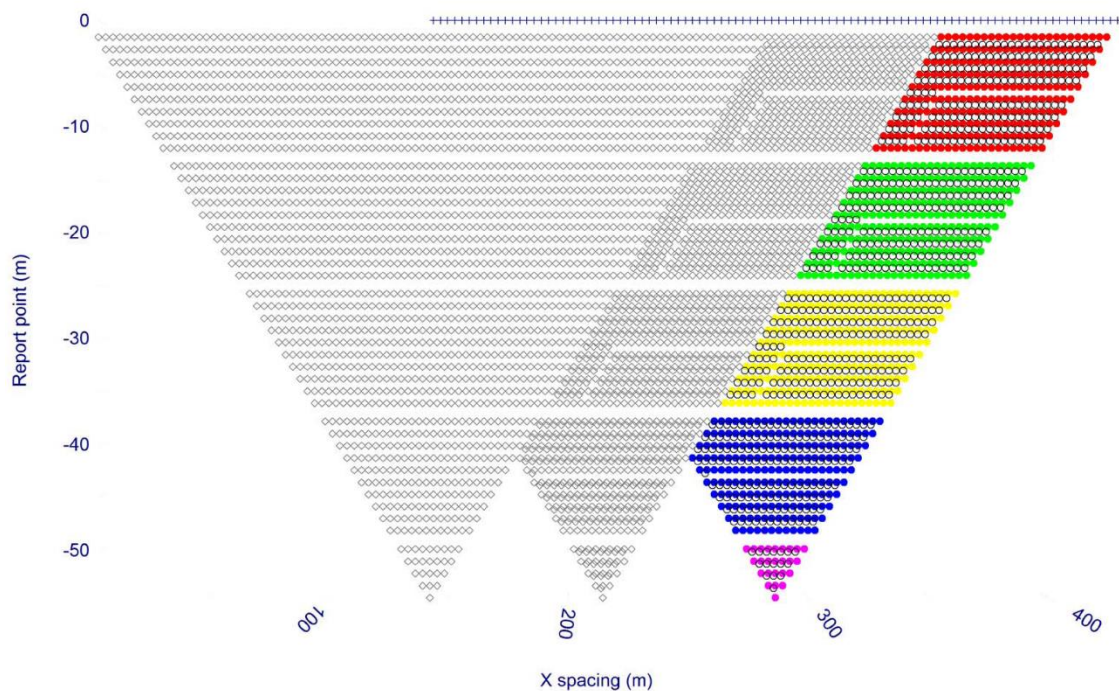
Figure 55 – Planned pattern of quadripoles with base and two roll-along movements for the ERT survey.

The quadripolar configuration applied was a manipulation of classical Wenner-Schlumberger, the Schlumberger reciprocal, where the current dipole is in the inner part of the quadripol. The Wenner Schlumberger quadripol allows to have a moderate sensitivity to the vertical and lateral variations, a compromise between Wenner and Dipole-Dipole quadripoles (Loke, 2004). In this way, as expected, complex variations of resistivity can be detected (Bernardinetti et al., 2018). The site far from the urban area yields the background noise negligible, also if the quadripolar configuration has a better signal to noise ratio compared to the Dipole – Dipole. From the point of view of the productivity, the reciprocal configuration of dipoles allows to acquire with multichannel optimization up to 9 quadripoles simultaneously, reducing the acquisition time considerably if compared with classical Wenner-Schlumberger. Technical parameters for the acquisition are summarized in Table 5.

Table 5 – Acquisition parameters for the ERT survey.

Injection duration	Injection Voltage	Stack number		Q factor	Number of Electrodes	Interelectrodic spacing
		min	Max			
500 ms	400 Volts	3	6	5%	96	3 m

Data processing can be divided in 2 steps. In the first step were removed bad data using threshold values represented by measurements of st-dev>5%, abs (ΔV) <1 mV, apparent resistivity <0 Ω*m and resistivity values that appears as outliers. In the second step was carried out the tomographic inversion starting from apparent resistivity values trough Res2Dinv as described at the § 3.2.4. At the purpose was applied the inversion routine based on the smoothness-constrained least-squares method.

The 2D model used was discretized by using the model refinement option that allows to enhance the model discretization and take into consideration large horizontal resistivity variations of the ground (Loke, 2004). At last, was applied the tool "Edit data - RMS statistic error" to remove noisy data, where the percentage difference between the observed and calculated apparent resistivity values was far and well separated from the average values.

The final electrical resistivity model was obtained at the fourth iteration with an RMS of 4.1%. The choice to contain the number of iteration is due to the possibility of creating artefacts by increasing the number of iterations.
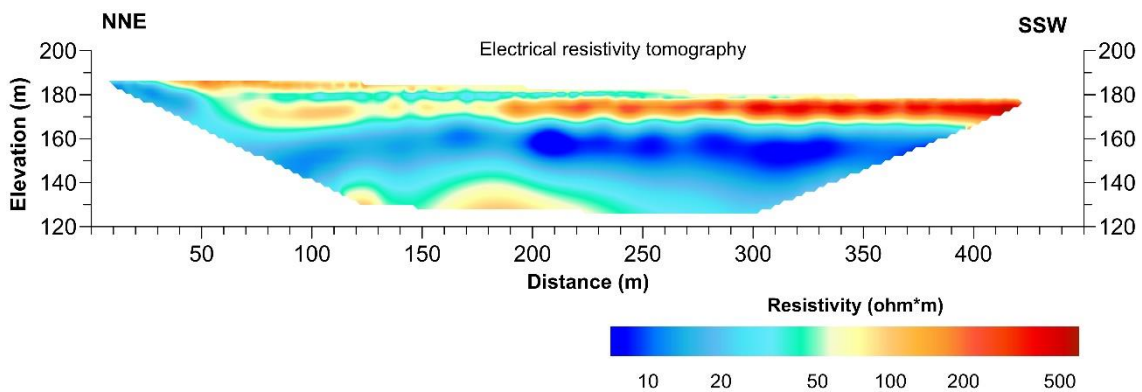
Figure 56 – Electrical resistivity tomography carried out with Schlumberger reciprocal quadripolar configuration.

The resistivity model in Figure 56 showed a range from 5 to 500 ohm*m, with maximum depth of 50/60 m. Into the tomography are visible vertical alternances of resistivity, more accentuated from the distance of 200 m; from 0 to 200 m there is a series of alternated resistive-conductive layers from the surface to the bottom, where is delineated an increasing of resistivity. A 60 m from the beginning of the tomography is well visible a lateral variation of resistivity, ranging from values of 20 up to 100 ohm*m. The layer from the elevation of about 170 m to the 140 m showed a lateral variation, in particular a decreasing of resistivity, from the progressive 200 m to the end of the ERT.

5.2.2 SRT

The seismic dataset was acquired along the same profile of the geoelectrical measurements, but with the major length of 576 m. The reason for the longer profile was for fully exploring the geological setting of the Monte Cervarola unit FAL4 into the basin, towards the Arno River.

Seismic surveys were planned with the aim to produce high resolution and redundant datasets, useful to carry out reliable tomograms and seismic section also in complex geological conditions.

In order to build an accurate tomographic velocity model with deep refractions and to record wide angle reflection, was applied the dense wide-aperture profiling. Several authors show the ability of seismic acquisition in complex geological environment by using densely spaced wide-aperture geometry (Improta et al., 2003; Bruno et al., 2010). Compared to the typical common midpoint narrow-aperture acquisition, the dense wide-aperture permits to acquire from near vertical reflections to large amplitude post-critical reflections and deep refracted waves, due to the large range of offset. The interpretation of redundant deep-penetrating refracted waves and reflection data acquired from a wide range of incidence angles, may provide accurate information on velocity distribution and interface geometries (Improta et al., 2002).

The array was deployed of two sub array of 192 active channels, equipped with 40 Hz natural frequency geophones, spaced of 2 m and sources were fired into the soil each 4 m with an 8-caliper seismic gun (Figure 57).
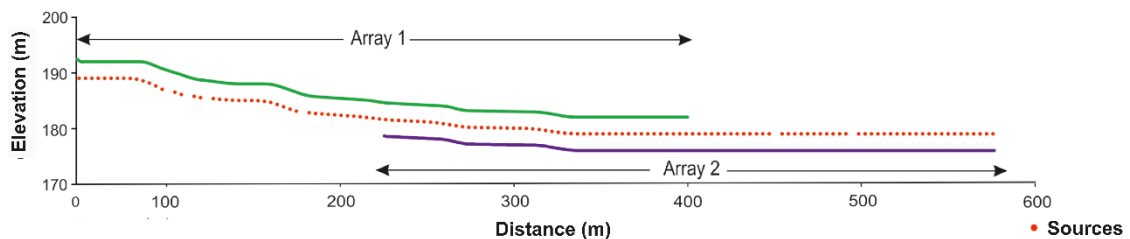
Figure 57 – Acquisition pattern configuration with an array of 192 active geophones moved forward with an overlap of half line: red dots are sources nominally fired each 4 m with an 8 mm caliper seismic gun.

This configuration allowed to acquire high resolution dataset with offset ranging from 1 up to about 380 m and a ratio maximum offset-depth of the target higher than 3 (Figure 58). Technical parameters for the acquisition are summarized in Table 6.
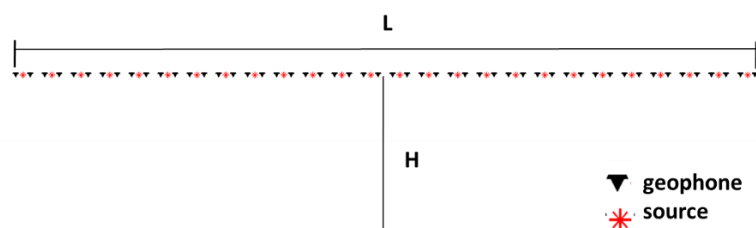
Figure 58 – Theoretical scheme of the survey length compared with the depth of the target.

Table 6 – Technical parameters for the acquisition of seismic refraction/reflection dataset.

Seismograph Geode Geometrics	8
Active channels	192
Natural frequency of geophones	40 Hz
Geophones interval	2 m
Source interval	4 m
Array number	2
Sampling interval	1 ms
Record length	1 s

First arrivals were manually picked on about 27000 seismic traces and were verified by using the reciprocity rules of Ackermann et al. (1986). Data were inverted as explained in § 3.3.3 through an iterative reconstruction technique (SIRT). The initial model was provided by traditional techniques of refraction (Burger, 1992). The final velocity model, i.e. tomographic model, was obtained after 20 iterations with a RMSE of 1.59.
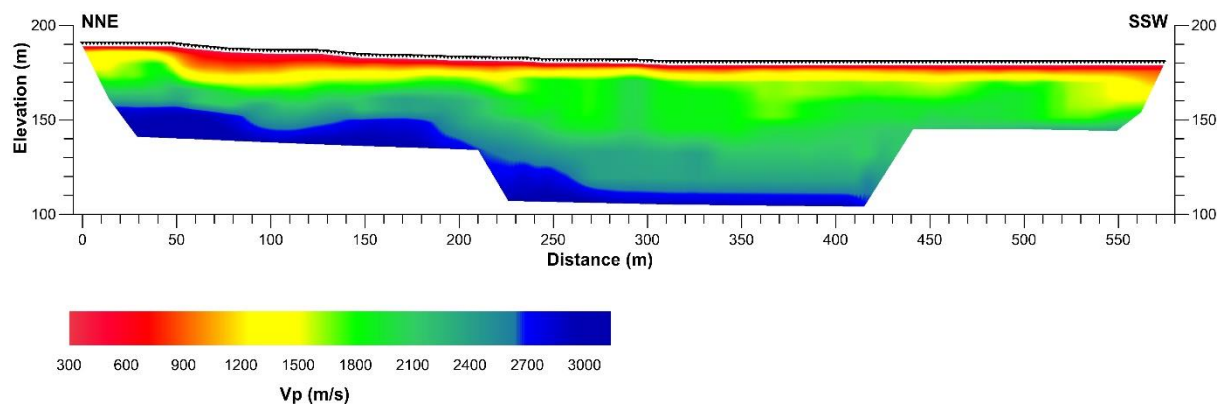
Figure 59 – Seismic refraction tomography carried out with Simultaneous iterative reconstruction technique.

P-waves velocity for the seismic tomographic model varies from 300 m/s up to more than 3000 m/s. The shallower part of the tomographic model (i.e. 10-15 m from the surface) showed a high vertical gradient and a strong lateral variation, at the distance of 50 m, which highlighted a deepening of about 10 m for the 1200 m/s contour level. The tomography showed another deepening into the middle part, showing in general a step structure.

5.2.3 Seismic attributes

Seismic attributes were carried out from the shallow depth converted Common Mid Point stack (Figure 60) for a total length of about 540 m, previously processed into the CGT Group research project "Sfruttamento sostenibile della risorsa idrica sotterranea dell'area di Laterina (Arezzo)". The main processes applied were: signal to noise enanchement, top and bottom muting, spiking and predictive deconvolution, refraction

94

statics, cdp sorting, semblance velocity analysis, normal move out correction, stacking and depth conversion. The acquisition parameters, well described in the previous paragraph, with nominal spacing of geophones and sources respectively of 2 m and 4 m allowed the high resoluted CDP spacing of 1 m.
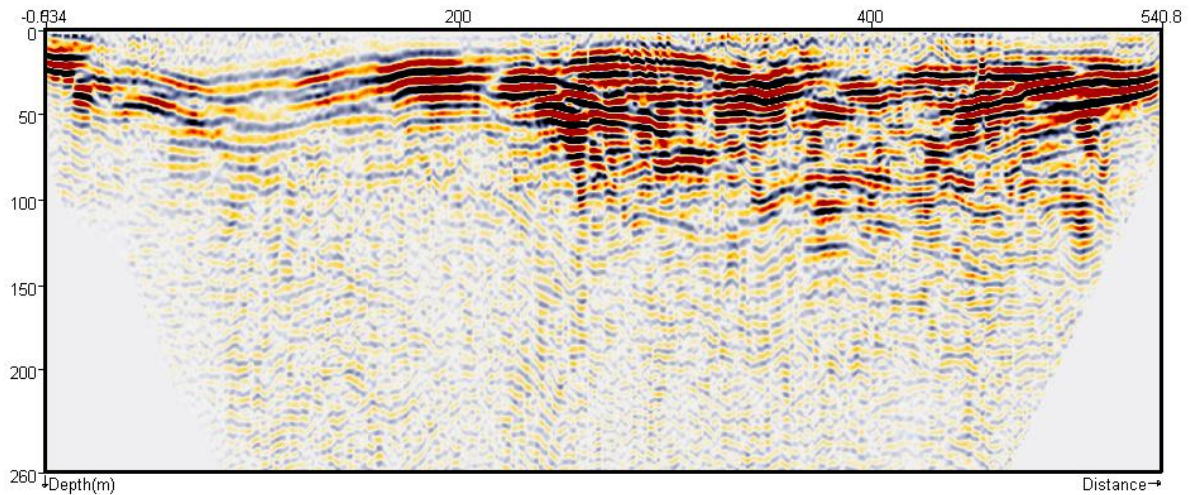
Figure 60 – CMP depth converted stack on shallow target of Laterina valley.

Seismic attributes were computed using OpendTect on the CMP stack for a depth up to 150 m, to analyze the part of the signal with higher signal to noise ratio and to avoid introducing sources of noise into the dataset. In order to enhance main stratigraphic changes useful for a larger scale study of the area, were produced energy, entropy, instantaneous frequency and similarity. These attributes were carried out with the depth gate from -3 to 3 m, and for texture attribute, step out nr=2 and GLCM size 32x32.

In general the seismic signal revealed an increasing of depth from the start to the end of the line with a maximum depth of about 120 m in correspondence of CDP 400.

The energy attribute showed zones with high energy (in red) from about the CDP 150 the end of the line mainly in a depth range included from 20 to 70 m with a main absence of high values from about the CDPs 350 to 400. Entropy attribute highlighted lowest values (in yellow) connected with the highest values of energy. The entire section can be separated in two main areas, before and after the CDP 200. From the start up to the CDP 200 entropy has values medium and locally low, then an increasing of values connected with the high discontinuous setting of reflectors.

Instantaneous frequency in the CDPs interval from 50 to 170 and in the depth range from 15 to 70 m revealed a zone with a homogeneous low frequency value and a smaller portion with similar values in the CDPs from 450 to 500 from the depth of 30 to 60 m. Higher values of instantaneous frequency dominate the section from about the CDP 200 to the end of the section. Highest values of instantaneous frequency are connected with local abrupt variation of phase of the seismic signal.

Similarity highlighted main continuous reflectors in white, locally interrupted with black sub-vertical paths, mainly present from about the CDP 200 to the end of the section.

All attributes highlighted a more complex image of the subsurface starting about from the CDP 200 to the end of the section, where medium to high energy and similarity zones are truncated by low similarity and high entropy patterns.
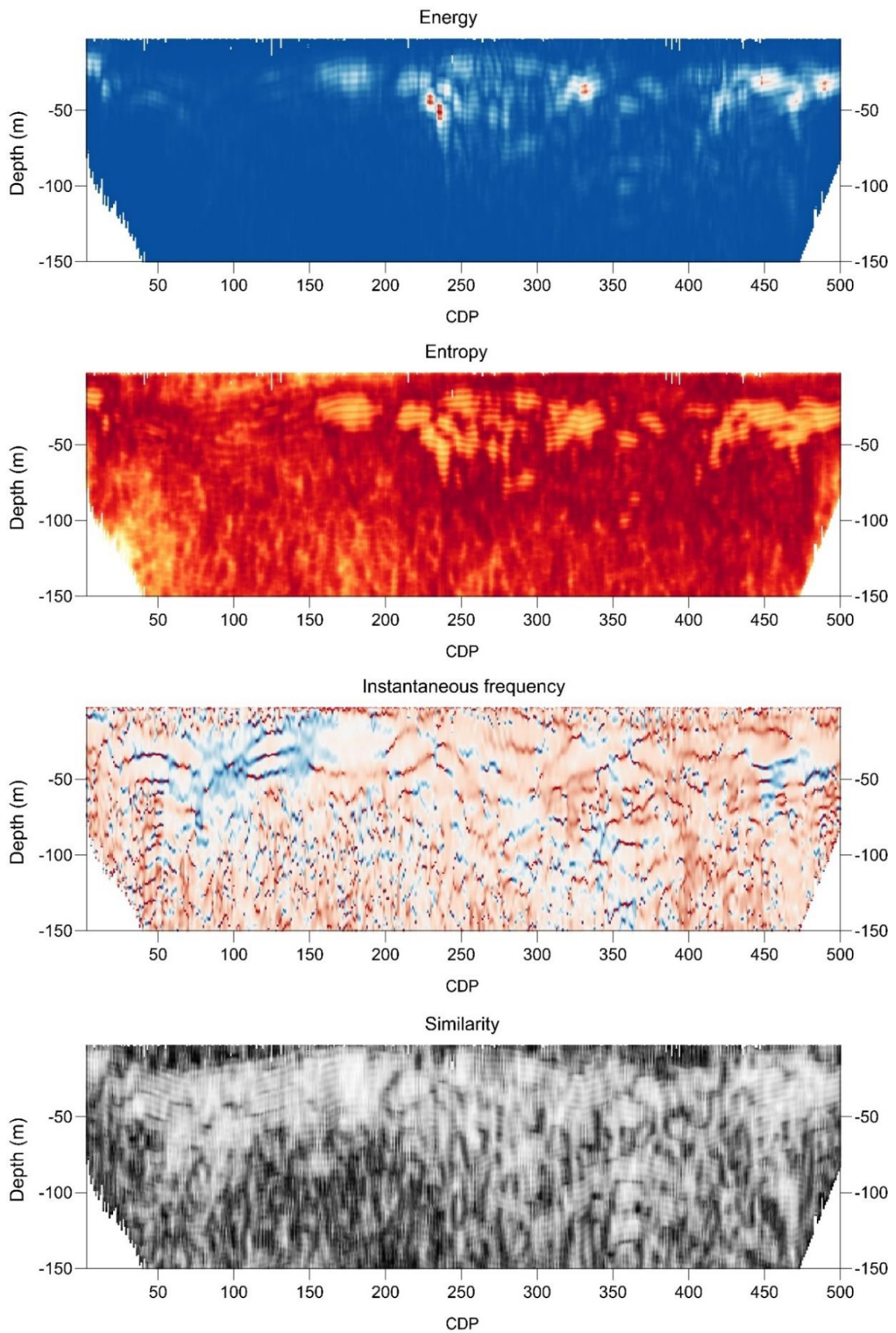
Figure 61 – Geometrical, physical and texture attributes produced from the depth converted CMP stack: energy is plotted with blue white red palette, where the red color is for the maximum values; entropy is plotted with a palette yellow to dark red, with lowest values for the yellow color; instantaneous frequency is plotted with the blue-white-red palette with lowest values for blue and highest for red; similarity is plotted with a grey level scale, where the white represents the maximum values and the black the minimum.

5.3 Results

5.3.1 Integrated tomography

Geoelectrical and seismic tomograms were sampled at co-located positions for a maximum length of 420 m by applying the triangulation with linear interpolation process (Surfer®) (Figure 62).
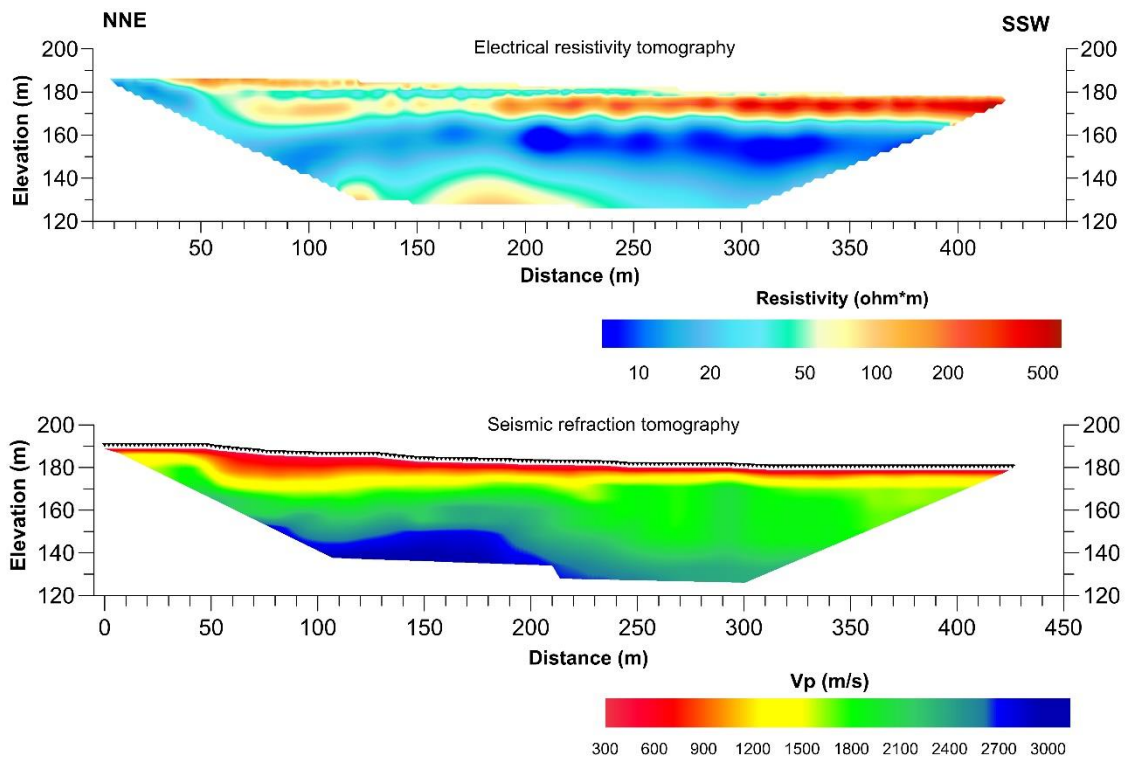
Figure 62 – Co-located ERT and SRT along the survey line.

The integration code was applied in a range of k varying from 2 up to 10, in agreement with the geophysical facies attended from the geological setting. Moreover, minimum values of Silhouette index obtained for the analyzed values from k equal 2 up to 10 are consistent with the assumed k range, i.e. the maximum value of the validity index is inner to the analyzed range.

The dataset was integrated by using the velocity in km/s and the $\log_{10}$ of the resistivity. In Figure 65 is showed the sphered dataset in which are well visible relationships between the two variables.

In order to preserve a high resolution dataset, tomographic models were sampled with a uniform grid spacing of 2x2 m, a compromise between the more resoluted ERT and the lower resoluted SRT (2x7 m grid spacing). A preliminary analysis to test the possible influence of the interpolation process in the integration technique is necessary: were compared integration results from 2x7 m, the lower resolution of the seismic model, and from 2x2 m common grid spacing. In Table 7 are showed the validity index results from the two analyses with 2x7 and 2x2 m grid spacing: in both cases the analyses showed the

optimal cluster number k=6. These results allowed using the more resoluted spacing, 2x2 m, as grid discretization for the sampling of co-located values.

Table 7 – Silhouette values for k analyzed from 2 to 10 for the 2x2 m and 2x7m grid spacing. In both cases the maximum is obtained for k=6.

	grid spacing	
K	2x7	2x2
2	0.60	0.60
3	0.61	0.63
4	0.62	0.65
5	0.64	0.65
6	0.68	0.66
7	0.66	0.64
8	0.65	0.63
9	0.66	0.62
10	0.62	0.64
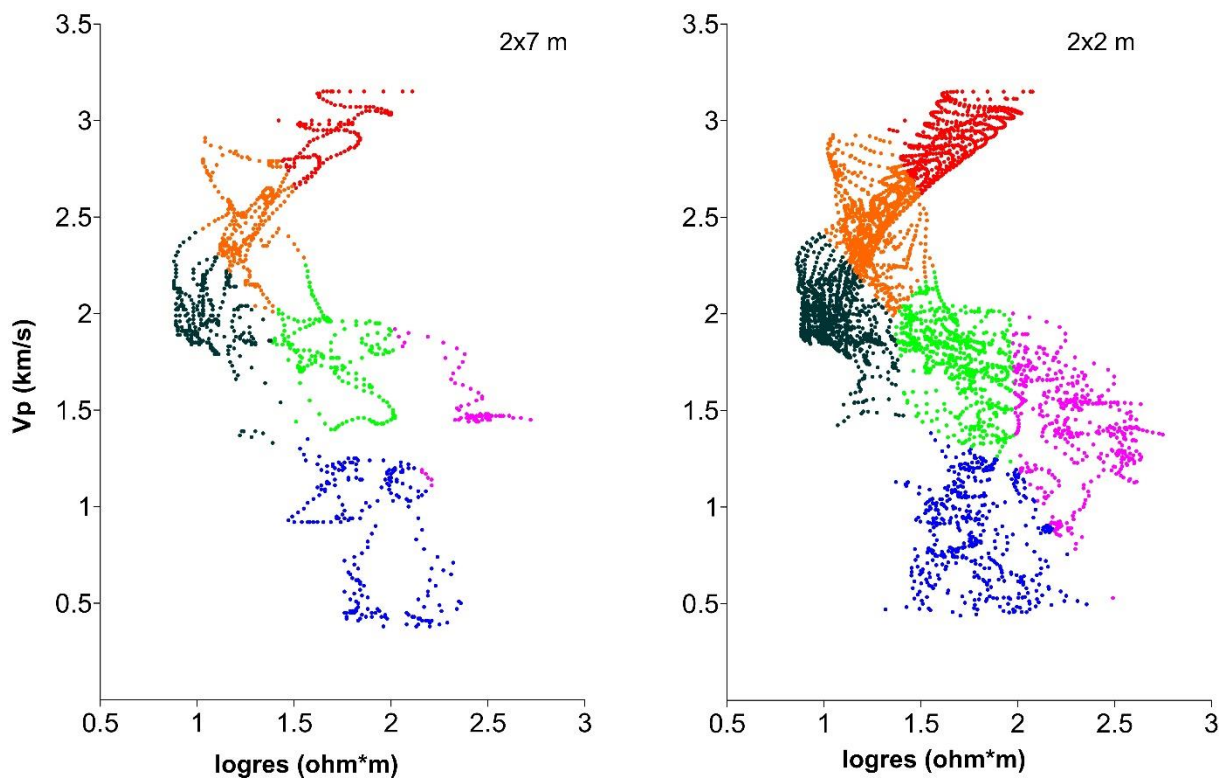
Figure 63 – Clustered bivariate dataset on the joint parameter space: on the left for 2x7 m grid spacing and on the right side for 2x2 m grid spacing.

The two analyzed datasets partitioned with cluster assignment colors (Figure 63) are showed in the joint parameter space in case of 2x7 and 2x2 m grid spacing. In Figure 64 are showed the two integrated images, respectively 2x7 and 2x2 m: the resulting images exhibited substantially the same macro geophysical facies and as attended an image with higher resolution in the case of 2x2 m grid spacing.
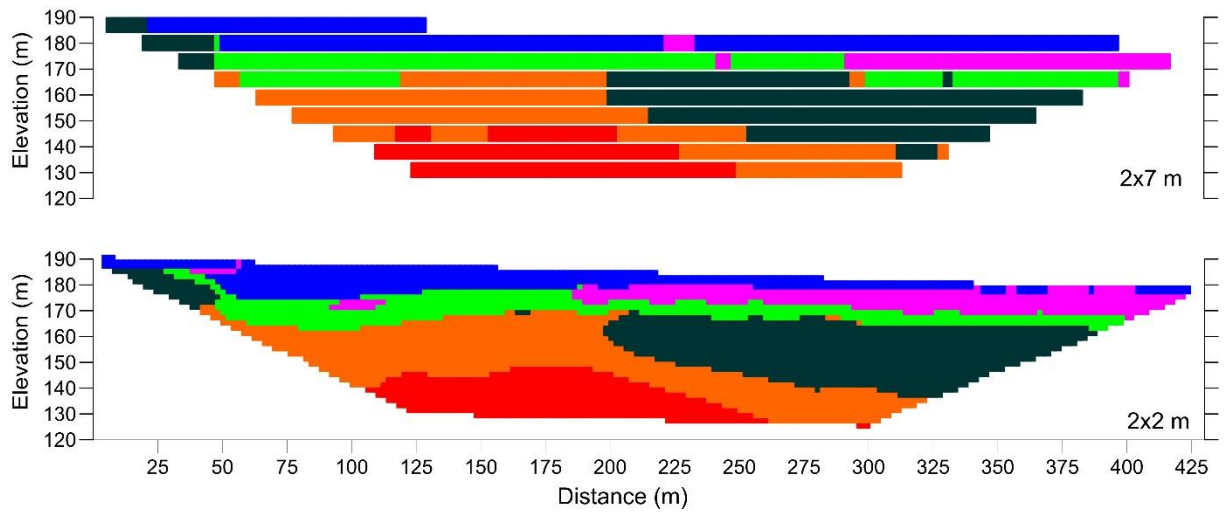
Figure 64 – Integrated tomograms, above with 2x7 m grid spacing and below with 2x2 m spacing: the tomography produced with 2x2 m common grid has a higher resolution with respect to the 2x7 m spaced dataset.
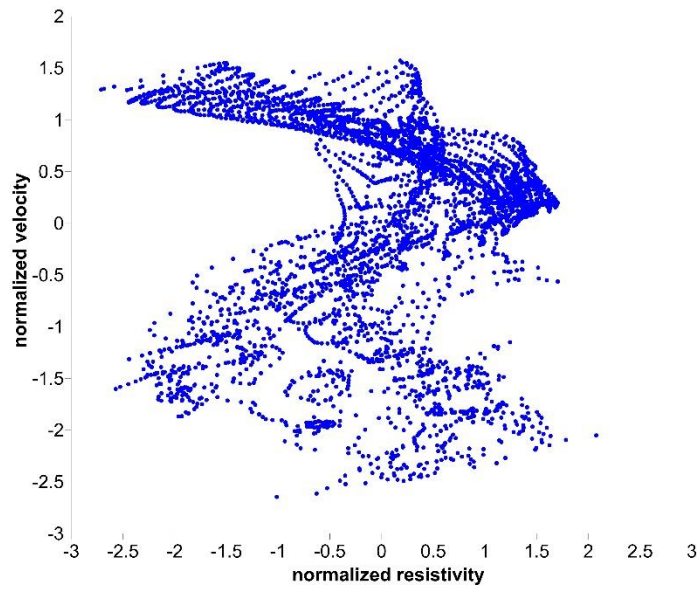
Figure 65 – Spherized dataset used as input of the clustering process.

The optimal number of clusters was carried out with the Silhouette index as explained in § 4.1.3. In Figure 66 are showed Silhouette values for different number of clusters. The optimal clustering was obtained for the maximum value of Silhouette at k=6, anyway similar values of silhouette were obtained for k= 4, 5.
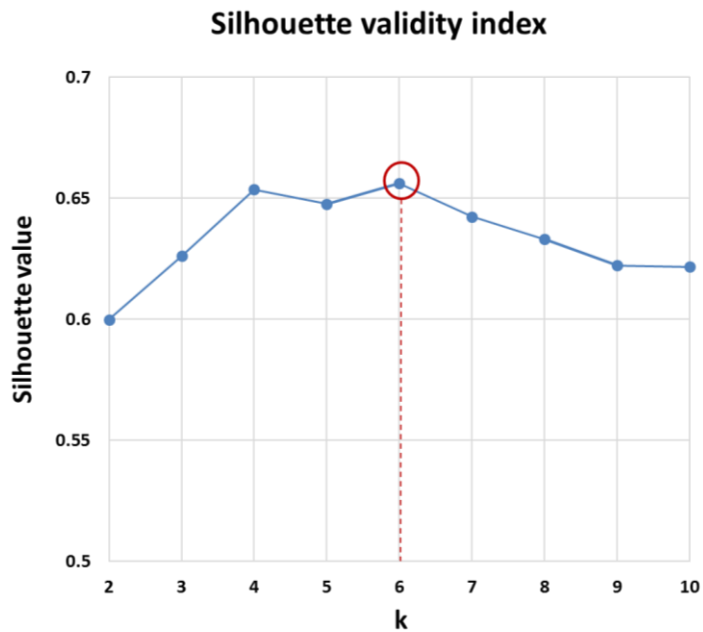
Figure 66 – Silhouette validity index for different k analyzed where the optimal clustering is individuated for the maximum value in k=6.

Table 8 – Average values and standard deviations for 6 clusters individuated for the optimal clustering. In the second column there are average Silhouette index values for each cluster.

Cluster	SI for the optimal k =6	Log_{10} Resistivity (ohm*m)	Velocity (km/s)
1	0.78	1.06±0.12	1.97±0.16
2	0.61	1.78±0.2	0.87±0.23
3	0.58	2.28±0.19	1.43±0.25
4	0.63	1.26±0.11	2.47±0.19
5	0.73	1.66±0.16	2.9±0.13
6	0.59	1.68±0.16	1.76±0.2

In Table 8 are summarized SI values for each cluster, with average values for the two variables and their standard deviation. In general and as expected, at an increment of SI corresponds a decreasing of the dispersion of both variables. The highest values of SI were reached by clusters 1 e 5, respectively with values of 0.78 e 0.73 (Table 8).

The integrated imaging of input ERT and SRT are showed in 3 different manners: 1) the tomography is showed by assigning a different color to each cluster, 2) the SI with a color scale ranging from the minimum to the maxim value and 3) the version saturated with SI. Low values of SI, ranging from -0.15 to 0.3 are indicative of the main uncertainties in the clustered observations, where the colors red and white represent the interfaces of geophysical facies. The version saturated with SI take into consideration the uncertainty in clustering analysis: the comparison with the hard clustering version shows few differences into the cluster 3 and 4 and the absence of the two lenses of cluster "3" in the left-shallow part of the tomography (Figure 68).
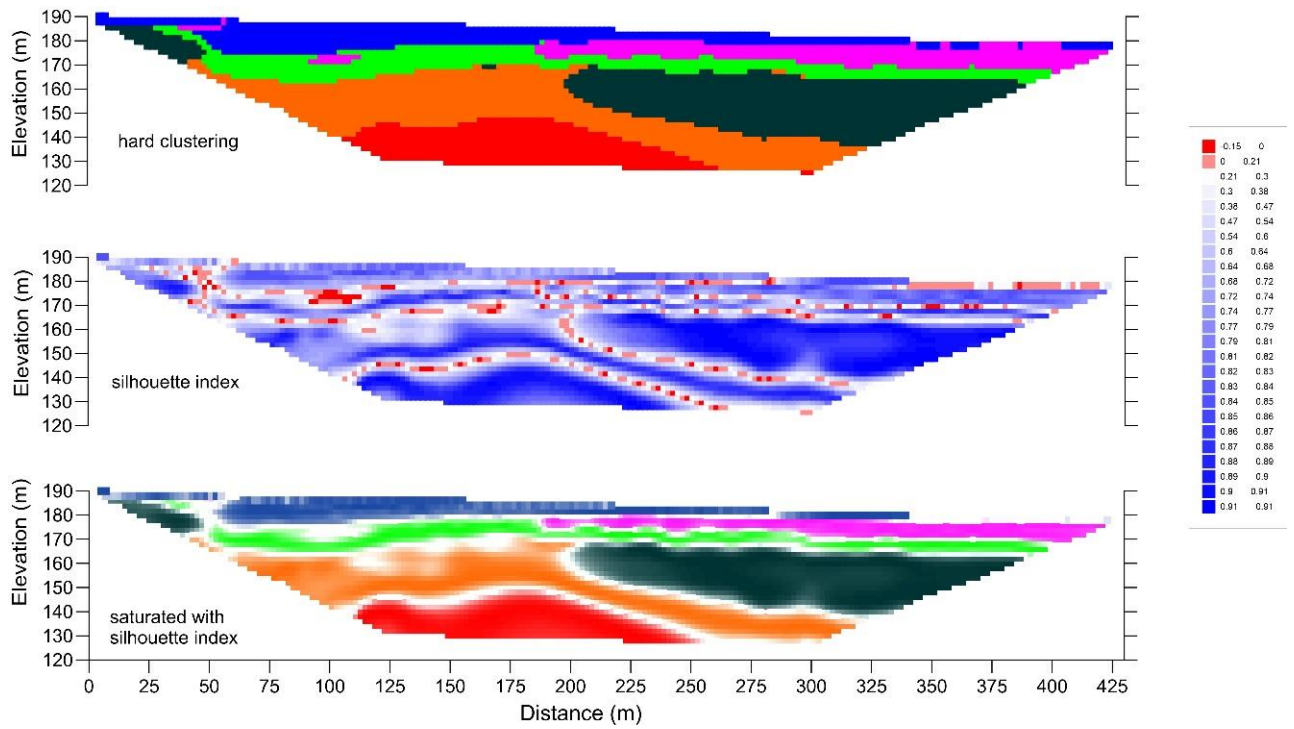
Figure 67 – Three different representation of the integration method: at the top the hard clustering version obtained with k-means algorithm, in the middle the tomography of Silhouette index where the red colors is representative of values lower than 0.21 and at the bottom the integrated tomography obtained with k-means algorithm saturated with Silhouette index.

In Figure 68 highlighted in black are observations with SI lower than 0 that are representative of the borders of clusters (i.e. geophysical homogeneous facies).
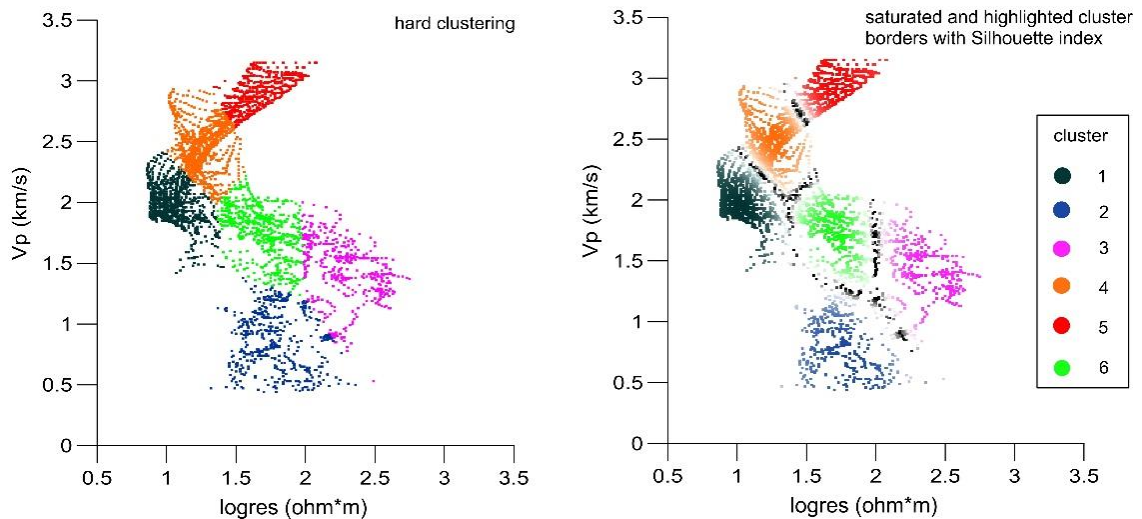
Figure 68 – The clustered dataset in the joint parameter space with clusters represented with different colors. In the right graph the dataset has colors saturated with Silhouette index and observations with values than 0 are highlighted in black.

The first three horizontal layer from the topographic surface are represented by the clusters 2, 3 and 6, with an increase of velocity with the depth and a maximum of resistivity for the cluster 3. Beneath are disposed the clusters 1 and 4, with a lateral marked decrease of velocity from the left part to the right of the tomography. Also the average resistivity

values showed a decrease from the cluster 4 to the 1. The deepest layer is individuated by the high velocity and high resistivity cluster 5 (Figure 67 and Figure 68).

5.3.2   *Integrated seismic attributes*

As described in the § 4.2, the first step in order to generate a neural network fitted on the dataset distribution, is estimate the two net dimensions (i.e. the number of the neurons along the X and Y directions). Through the principal component analysis one can estimate the ratio between the two sides of the net as the ratio between the variances of first two dimensions. Moreover, as suggested by Vesanto and Alhoniemi (2000), the number of the neurons should be bigger than the clusters into the dataset (i.e. expected homogeneous zones within the seismic image).

In order to produce the best integrated section, were carried out several SOM analyses with different grid dimensions on the same dataset. Thus, the visual inspection allowed to choose the most meaningful images, based on the subsurface features expected.

The dataset analyzed with the Self-Organizing maps was composed by energy, entropy, instantaneous frequency and similarity after sphering the data (i.e. scaling the dataset). The variables contribution and their quality were evaluated in the correlation circle (Abdi and Williams, 2010): in Figure 69, a), the three variables with their cos2 (i.e. square cosine, squared coordinates) are represented, where energy and entropy showed a better representation on the principal component compared with similarity and instantaneous frequency. Conversely a minimal difference in the contribution in principal component between the highest for energy and the lowest for instantaneous frequency. The variances for the first two dimensions are respectively 38.3% and 28.2% and their ratio is equal to 1.36.
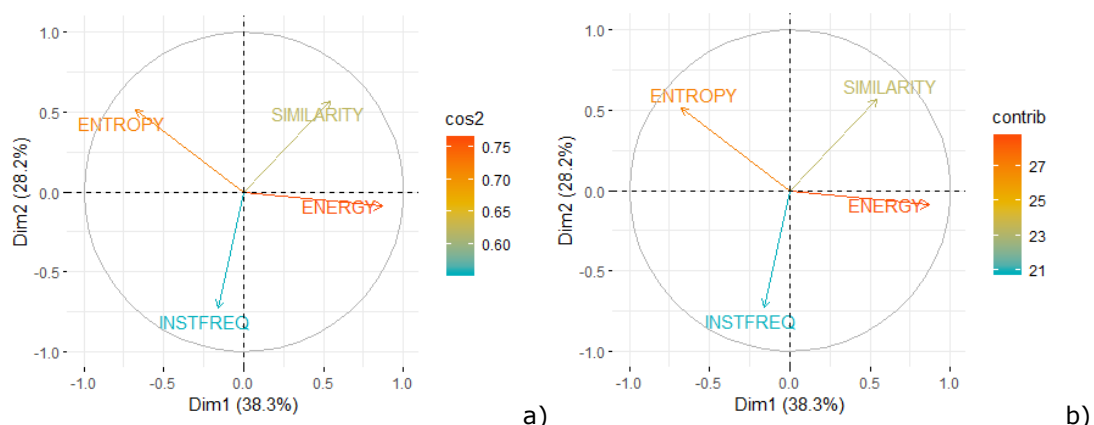
Figure 69 - Principal components analysis for the dataset composed by energy, entropy, instantaneous frequency and similarity seismic attributes: in a) are represented the attributes in the correlation circle with square cosine highlighted with the color scale, in b) are represented the attributes in the correlation circle with contribution highlighted with the color scale. In both analyses energy and entropy showed the best representation in first two dimensions.

The multivariate dataset composed by energy, entropy, instantaneous frequency and similarity was analyzed for the following neural grid with a ratio of the sides close to

1.36: 7x5, 10x7, 12x9 and 14x10. The best SOM analysis was obtained for the neural grid in the hexagonal disposition with size 7x5 (Figure 71), trained for 9000 iterations.

Analyzed dataset can be interpreted through the comparison of the RGB color map, the heatmaps and the Umatrix to reveal the main features. The observations classified into the neurons 5, 6, 7, 13 and 14 (Figure 71) are representative of zones with high similarity and energy from medium to maximum, with low values of entropy and instantaneous frequency. The high values of Umatrix for the same neurons confirms the large variation of energy and in minor part of entropy (Figure 70). On the integrated section are imaged as local discontinuous zones from the depth of about 20 m up to 60 m, for CDPs greater than 150. The neurons 4, 5, 12, 13, 19 and 20 show low to medium energy, high values of similarity and medium to low entropy and in the integrated section are representative of the zones included in the CDPs range from 70 to 200 and with a depth ranging from 20 to 70m. Neurons 24, 25, 30, 31, 32, 33 have high values of entropy with similarity variable from medium to high and are characteristics of the integrated section from the CDP 200 to the end of the section. The remaining neurons can be considered as background values, due to the low ratio signal to noise ratio of the seismic image.
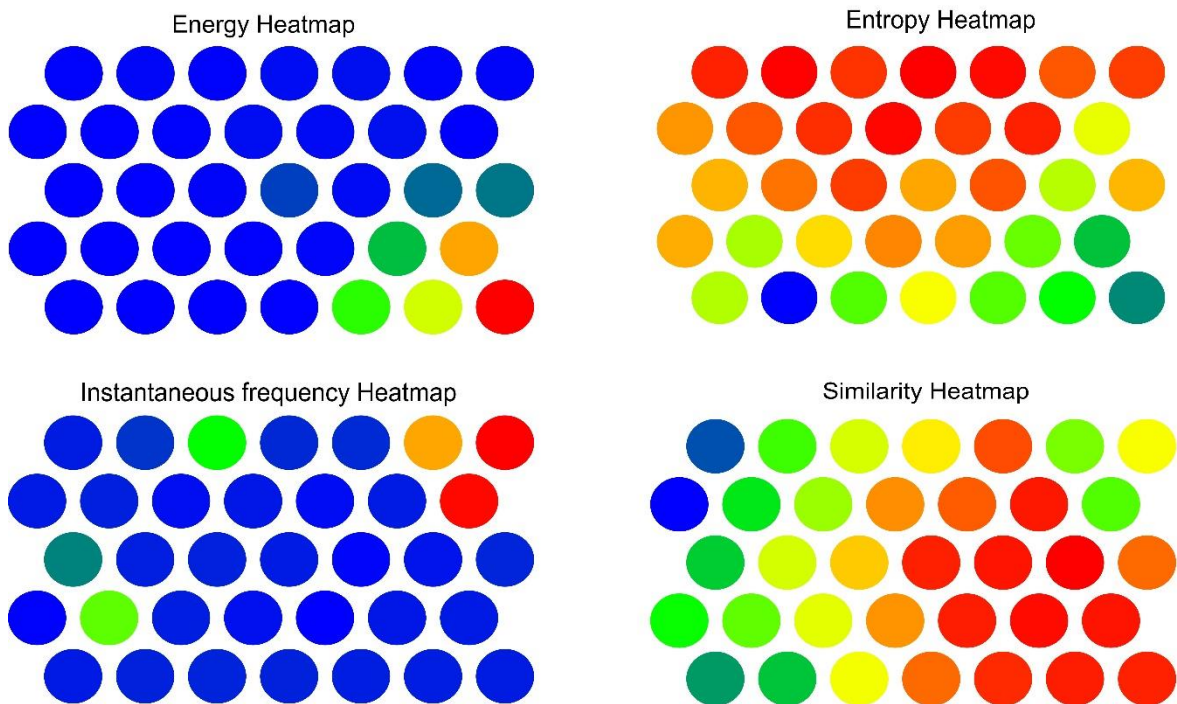
Figure 70 - Distribution of average attributes values on 7x5 trained SOM for the dataset composed by energy, entropy, instantaneous frequency and similarity: cold colors represent low values and warm colors represent high values. In the right lower corner of the map are classified observations with medium to high energy and similarity, low entropy and instantaneous frequency. In the upper right corner of the map are classified medium to high entropy and instantaneous frequency, low energy and medium to low similarity. In the half upper part of the map high entropy values with variable similarity.
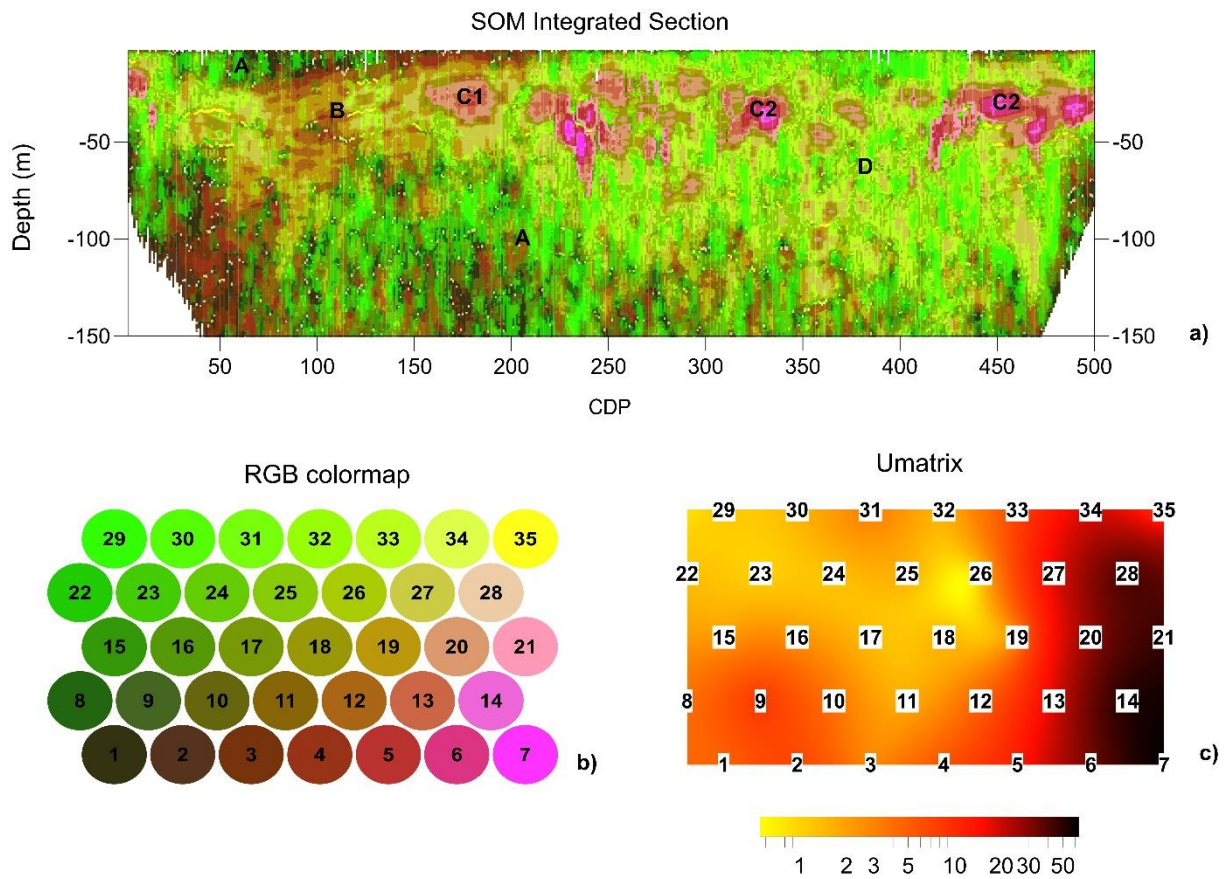
Figure 71 - In the upper part a) the integrated section with similar zones individuated with colors. In the lower left side b) the RGB color map based on the position of neurons and c) the U-matrix. With "A" are high entropy and low energy/similarity zones; with "B" high similarity, low energy, low entropy and frequency zones; with C1 and C2 high similarity and medium to high energy zones; with "D" high entropy with medium to high similarity. A main distinction in two different zones can be considered before and after the cdp 200, in agreement with the lateral discontinuity revealed by the tomographic method in Figure 67.

6 Complex hydrothermal aquifer characterization: Solfatara volcano

6.1 Geological Setting

Solfatara volcano is located at NE of Pozzuoli within the Campi Flegrei area (Figure 72), west of Naples (Sothern Italy), which is characterized by an active resurgent caldera system (Vitale and Isaia, 2014 and references therein) resulting from two large collapses related to the Campanian Ignimbrite (39 ka, De Vivo et al., 2001) and the Neapolitan Yellow Tuff (14 ka, Orsi et al., 1996) eruptions. The Campi Flegrei volcanic system is still active, since the last eruption occurred in 1538 A.D. at Monte Nuovo. The volcanic activity included both phreatic and phreatomagmatic events, as well as lava dome-forming eruptions in the area around the present crater rim (Isaia et al., 2009; Vitale and Isaia, 2014; Isaia et al., 2015). Geophysical (Bruno et al. 2007; Petrosino et al. 2012; Byrdina et al. 2014; Di Giuseppe et al. 2015; De Landro et al., 2016; Gresse et al 2017 and reference therein) and field data (Isaia et al. 2009; Isaia et al. 2015) allowed to reconstruct the stratigraphy of the shallow part of the crater, down to a depth of 150 m, and reconstruct the complex hydrothermal system located below the Solfatara crater which includes a mix of upwelling fluids, gases, and meteoric water. Isaia et al. (2015) suggested that the deeper part of the system (below 150 m) comprises fractured rocks of the pre-eruption basement, which are mostly composed of deposits from the Agnano-Monte Spina eruptive sequence (De Vita et al. 1999), and which form a stair-step-like structure that is cut by concentric, steep ring faults. These are covered by the Solfatara deposits, which range in composition from alkali-basalts to trachyte and phonolite, were mainly generated during explosive eruptions (e.g., De Vivo et al., 2001). These deposits are mainly represented by phreatic and phreato-magmatic ash, lapilli, as well as breccias and collapse breccias (Isaia et al. 2009; Isaia et al. 2015). The Solfatara unit is in turn overlain by the pyroclastic deposits from Astroni volcano (less than 2 km north of Solfatara) that cover most of the central sector of the caldera, and by more recent deposits. This collapsed sediment filled structure hosts a complex system of fluid circulation, which is controlled by both rock matrix permeability and fractures (Byrdina et al. 2014; Isaia et al. 2015).

The caldera has been characterized by periodic episodes of extended, low-rate ground subsidence and uplift, a phenomenon called bradyseism, accompanied by intense seismic and geochemical activity (e.g., De Natale et al., 1991). Sea level measurements, made on the ruins of a Roman market built near the seashore of Pozzuoli town, indicate a slow subsidence of the area since Roman times. This subsidence was first interrupted by a 7-m uplift, which took place west of Pozzuoli between 1530 and 1538 AD, ending with the Monte Nuovo eruption in 1538 AD (Di Vito et al., 1987). After the 1538 eruption, the subsidence continued until 1968. Then, rapid uplift started again, totaling about 170 cm

during 1969–1972 and 182 cm during 1982–1984 (Berrino et al., 1984). In contrast to the sinking phases, uplift episodes are accompanied by seismic activity. Earthquakes occur mostly along the coast around Pozzuoli, at Solfatara, and within the bay. Seismic activity does not extend outside the margin of the caldera and abruptly terminates at 3.5- to 4-km depth, suggesting a sharp transition from brittle to ductile behavior. The maximum recorded magnitude (4.0) was measured in 1984 (Orsi et al., 1999). Earthquakes are likely associated with the upward migration of a pressure front triggered by an excess of fluid pressure from a magmatic intrusion, and with the brittle readjustment of the inflated system occurring along some lubricated structures (Gaeta et al., 1998; Bianco et al., 2004; Saccorotti et al., 2007; Cusano et al., 2008 and reference therein).
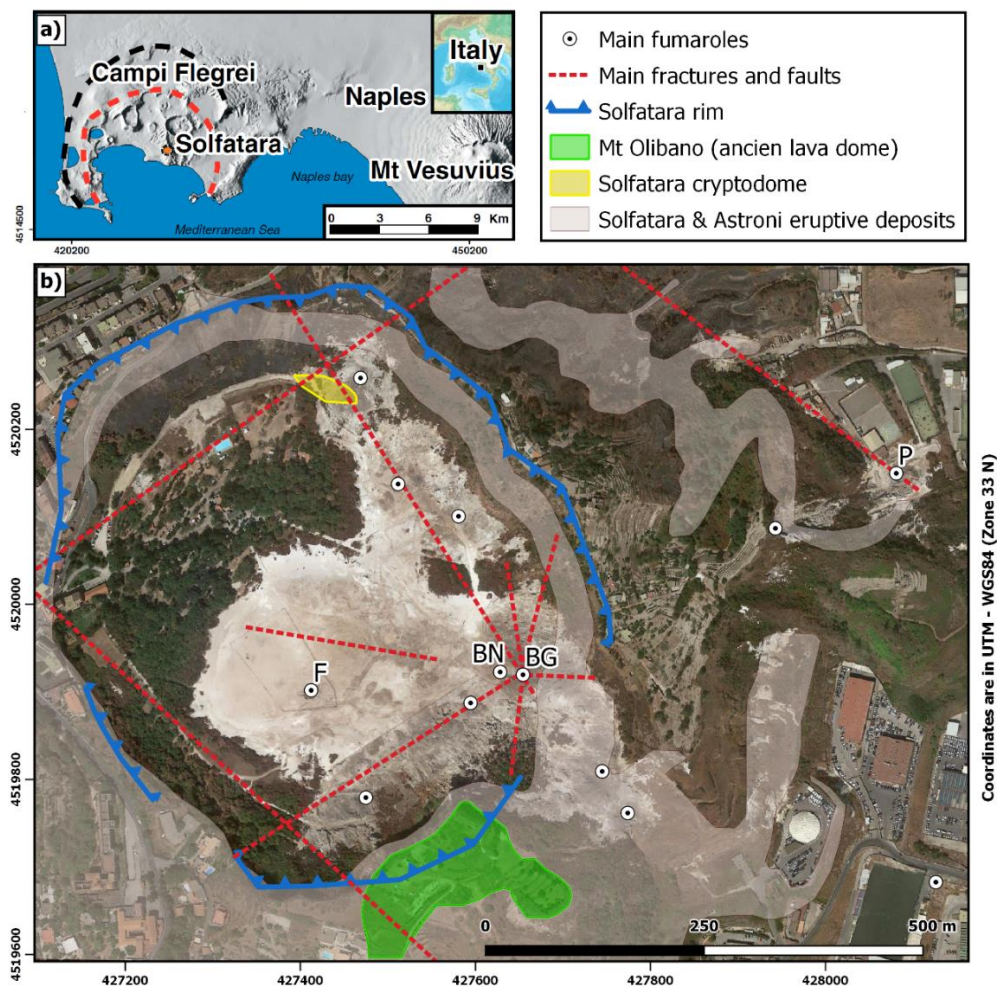
Figure 72 - Map of the studied area. a) Solfatara crater inside the Campi Flegrei, in the Neapolitan urban area (GIS database from Bechtold et al., 2005). Two major eruptions formed nested calderas, Campanian Ignimbrite and Neapolitan Yellow Tuff calderas (black and red dotted line). b) Solfatara satellite map (2014) with main geological units: Mount Olibano, Solfatara cryptodome, and eruptive deposits. Red dashed lines indicate the major fractures/faults. Main fumaroles are indicated with white circles and black dot: Bocca Grande (BG) and Bocca Nuova (BN) fumaroles and Fangaia mud pool (F) and the Pisciarelli fumarolic area (P) (modified from Gresse et al., 2017).

The Solfatara crater is characterized by intense diffuse degassing and fumarolic emissions (e.g., Todesco et al., 2010; Chiodini et al., 2005; Chiodini, 2009). Chiodini et al. 2001 showed that Solfatara releases about 1500 t/day of volcanic-hydrothermal CO_2, as a result of diffuse degassing through soil, during which about 3350 t/day of steam condense.

107

This hydrothermal water circulating in the shallow part of the hydrothermal system derive from a deeper zone where magmatic gases mix with meteoric water and rise as a plume to shallow depth (Caliro et al., 2007) and emerging from the mud pool at Fangaia (e.g., Bruno et al. 2007; Petrosino et al. 2012; Isaia et al. 2015). Above the mix zone, fluids flow essentially as gas phase from a high-temperature area (> 350 °C) to a shallow zone (< 300 m) where temperatures range from 190 to 230 °C (Caliro et al., 2007). Electric resistivity tomography (ERT) profiles of this shallow zone (100–300 m) revealed the existence of a two-phase system represented by liquid- and gas-dominated zones in the subsurface of Fangaia mud pool (Figure 73) and the fumaroles of Bocca Grande and Bocca Nuova, respectively (Byrdina et al., 2014). In this shallow steam-heated part of Solfatara crater, a predominant advanced argillitic alteration occurs, where sulfuric acid is created at, or above, the water table by the oxidation of H_2S (Rye,. 2005). This type of solfataric alteration causes the formation of mainly amorphous silica and alunite (Piochi et al. 2015; Mayer et al. 2016). The hydrothermal activity at the surface is characterized by an intense soil diffuse degassing, both inside and outside of the crater (Cardellini et al. 2017 and reference therein). Acidic (pH ~ 1.7) high-temperature (≥ 160 °C) fumaroles are mainly located in the eastern part of the crater, whereas hot springs, steam-heated pools (45–95 °C), and fumarolic vents are concentrated in its center (Glamoclija et al. 2004; Valentino and Stanzione 2004; Chiodini et al. 2011). The absence of vegetation represents the area of diffuse soil degassing as well as the solfataric alteration of the rocks.

Fumarole temperatures measured at the surface range up to 145 and 150 °C for Bocca Grande and Bocca Nuova, respectively (Valentino and Stanzione, 2004; Chiodini et al., 2015). The main components of the fumarolic emissions are H_2O and CO_2, followed by H_2S, H_2, N_2, and CH_4 (Caliro et al. 2007). Due to the buffering of magmatic fluids by the large hydrothermal system, components like SO_2, HCl, and HF are not detected at the surface (Chiodini et al. 2001; Moretti et al. 2013).
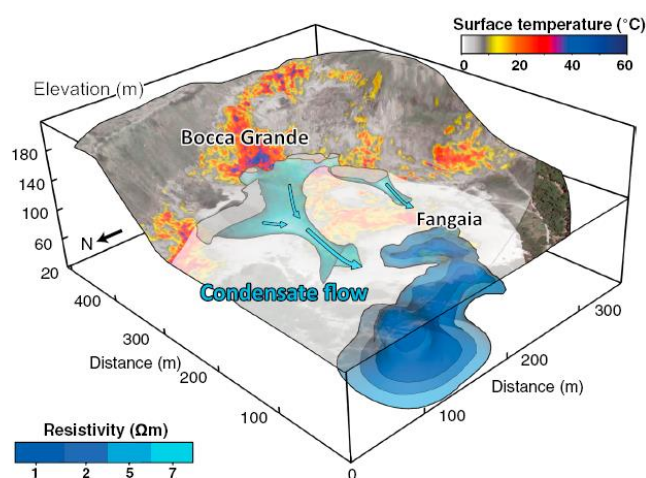
Figure 73 – Surface temperature (°C) map of the Solfatara crater with several 3D electrical resistivity isovalues. In light blue are indicated the condensate flow and the lower resistivity values the Fangaia liquid-dominated plume (From Gresse et al., 2017).

6.2 Geophysical data set

Two coincident high-resolution seismic reflection/refraction (Bruno et al., 2017) and electrical resistivity (Gresse et al., 2017) profiles were acquired in the Solfatara tuff cone in May and November 2014 (Figure 74), along with CO_2 flux and surface temperature measurements. The acquired data are a subset of the MedSuV - RICEN dataset, which also includes a wider series of time-lapse geophysical and geochemical experiments carried out within Solfatara volcano, with the aim of studying changes in the properties of the medium at small scales through repeated high-resolution multi-parameter observations over time. In order to fully explore and reconstruct in depth the crater, were acquired two orthogonal profiles (Figure 74).
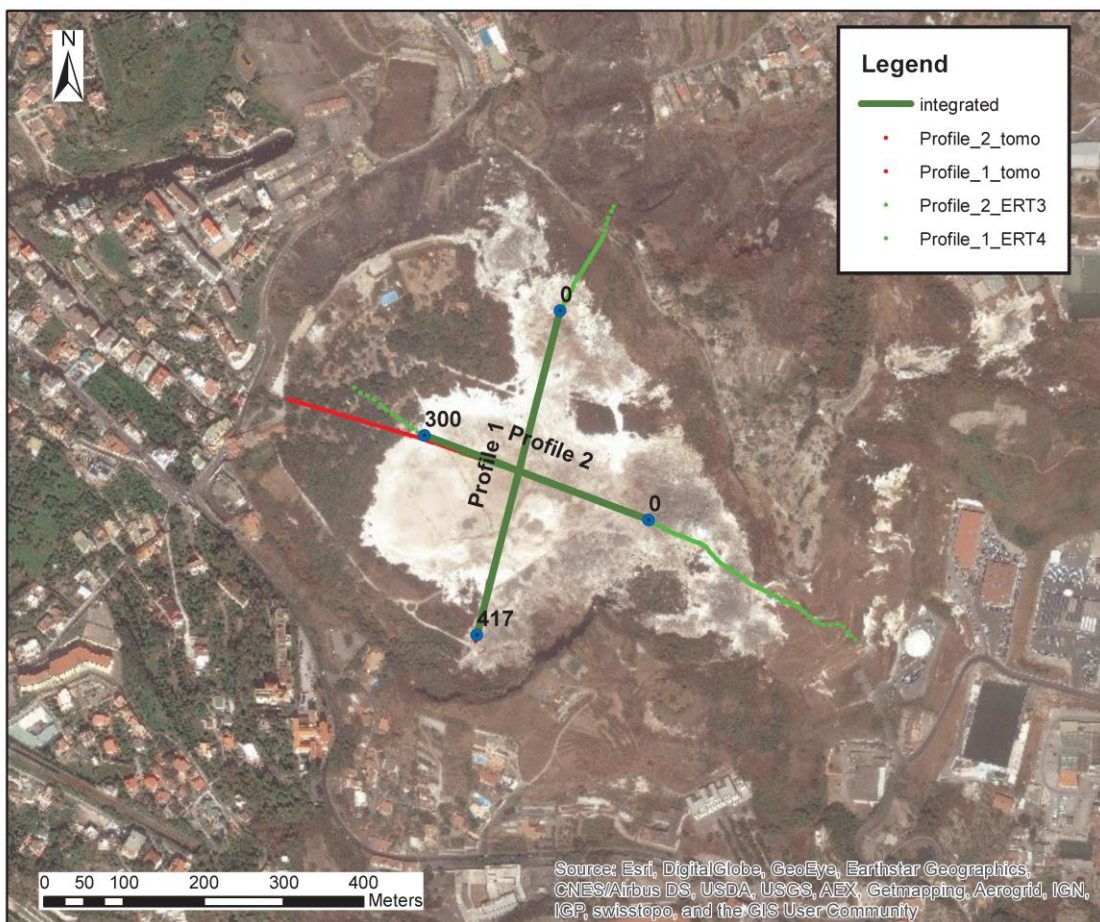
Figure 74 – 2D map with satellite imagery showing the location of seismic and geoelectrical surveys and the co-located datasets position used in the integrated tomography.

6.2.1 ERT

DC current resistivity datasets were acquired by stacking three to seven individual measures. Only measurements with a standard deviation below 5% were retained for the inversion. Furthermore, it appeared that in the globally conductive area of the Solfatara crater, an injection current below 50 mA was not high enough to ensure robust resistance measurements; consequently, were removed these data (Gresse et al., 2017).

NNE **Electrical resistivity tomography** SSW

ESE **Electrical resistivity tomography** WNW
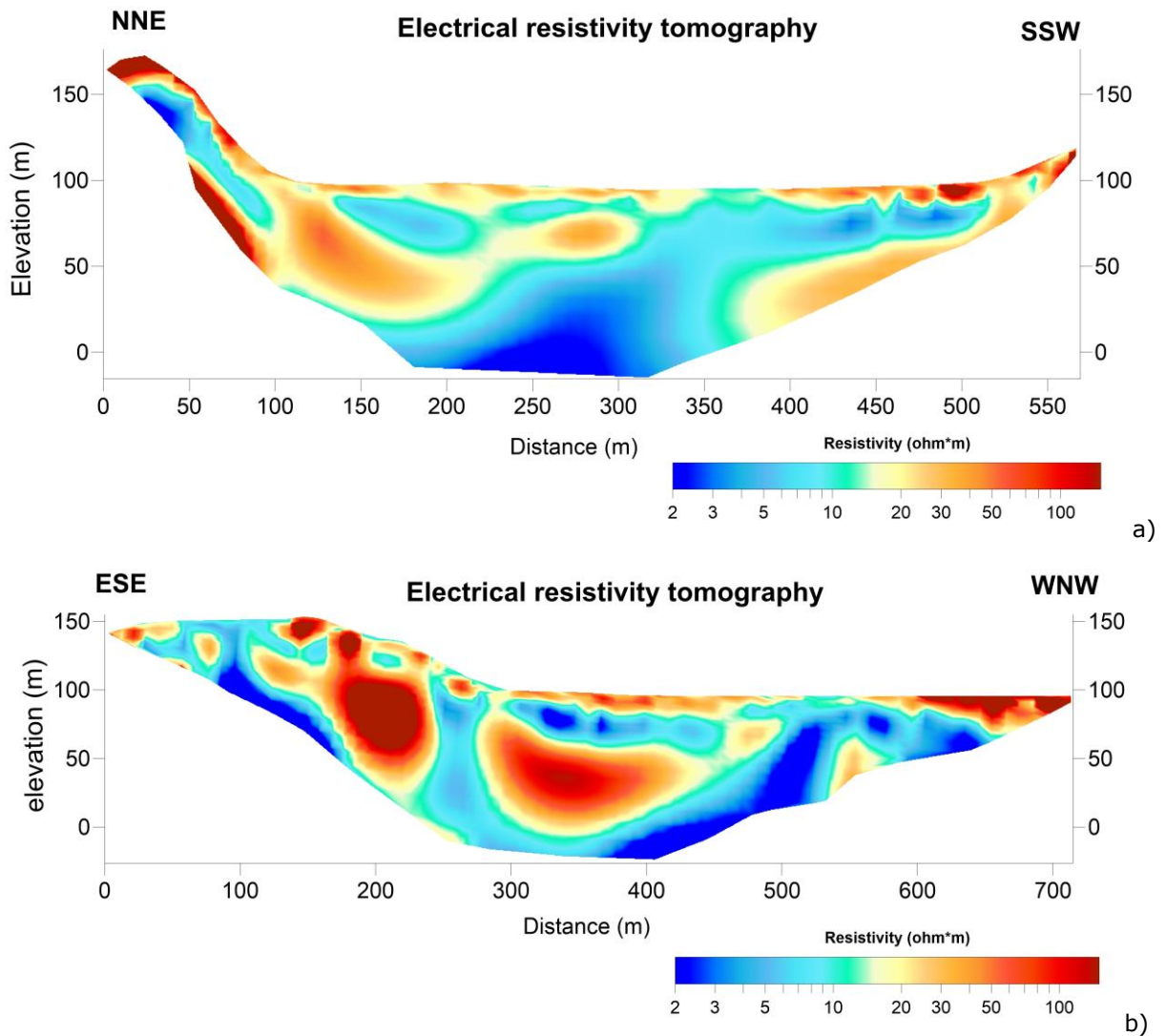
Figure 75 – Electrical resistivity tomography carried out on: a) "Profile 1" and b) "Profile 2"

The quadripolar configuration applied was the classical Wenner-Schlumberger, powerful to explore complex subsoils where are present lateral and vertical resistivity variations (Loke, 2004), with a high ratio signal-to-noise.

The tomographic inversion was carried out with Res2Dinv as described at the § 3.2.4. At the purpose was applied the inversion routine based on the smoothness-constrained least-squares method. The final electrical resistivity models were obtained at the fourth iteration with RMS 8.5 and 10.5%, respectively for "Profile 1" and "Profile 2".

The resistivity models ("Profile 1" and "Profile 2") showed resistivity values ranging from few ohm*m to more than 100 ohm*m. In both profiles is present a resistive (>30 ohm*m) layer with thickness of about 15 m, which is interrupted by a conductive zone at the progressive 350 m in "Profile 1" and at 530 m in "Profile 2". Beneath this layer, a more conductive one is present, with resistivity values lower than 10 ohm*m and a variable thickness from 5 to about 20 m. The resistivity models on both profiles imaged a large and sub-vertical conductive body with resistivity values lower than 5 ohm*m reaching the bottom of the tomograms.

6.2.2 SRT

In order to image the structure of the crater at larger depths (down to ~1 km), during the RICEN experiment (Bruno et al., 2017) also conducted two high-resolution seismic profiles along two orthogonal NNE (i.e. Profile 1) and WNW (i.e. Profile 2) directions (Figure 74). Their purpose is to locate fractures and faults, whose position and geometry are fundamental in understanding the mechanisms of massive degassing and hydrothermal fluid circulation occurring near the surface (Bruno et al., 2017). The first arrivals provided by these surveys were processed in order to obtain two high resolution refraction tomograms (Bruno et al., 2017).

Table 9 - Technical parameters for the acquisition of seismic refraction/reflection of Profile 1 and Profile 2.

Seismograph Geode Geometrics	Up to 10
Active channels	Up to 240
Natural frequency of geophones	4.5 Hz
Geophones interval	2 m
Source interval	4 m
Seismic source	IVI MINIVIB
Sweep frequency range	5 – 150 Hz
Sweep length	15 s

As a seismic source, was used a single 6400 kg IVI-MINIVIB® vibroseis truck, which delivers a maximum theoretical peak force of ~27 kN at each sweep. The sources were spaced 4 m, located halfway between geophones, with some source gaps due to the unsafe surface conditions, especially along Profile 2. 117 vibration points were acquired along profile Profile 1 and 75 vibration points along Profile 2. At each vibration point, two sweeps were stacked and recorded respectively by a 216-channel (i.e. Profile 1) and 240-channel (i.e. Profile 2) array, with spacing between the sensors of 2 m. The source-receiver configuration allowed sampling a wide range of offsets, from a minimum of 0.5 m to a maximum of 453 m (Bruno et al., 2017).
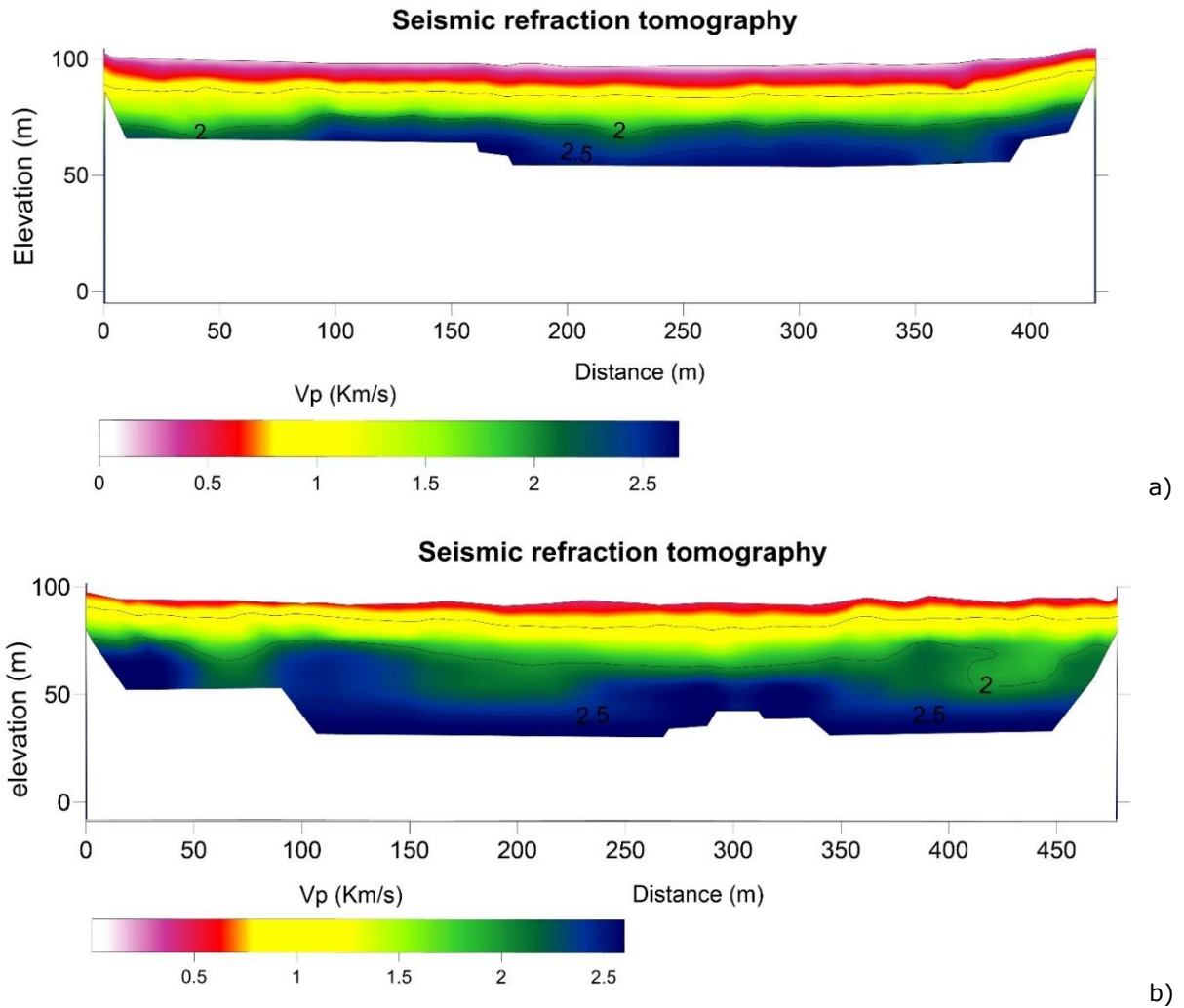
Figure 76 – Seismic refraction tomography carried out on: a) "Profile 1" and b) "Profile 2".

P-waves velocity for the seismic tomographic models varies from 300 m/s up to more than 2600 m/s. The shallower part of the tomographic model (i.e. 10-15 m from the surface) showed a limited vertical gradient, especially in the Profile 1 and not remarkable lateral variations, with velocity up to about 1000 m/s. In both profiles at the depth of about 20 m starts a vertical velocity increasing from about 1000 m/s to more than 2000 m/s. In the deeper part, the Profile 2 showed lateral velocity variations, in a range of about 2000 m/s to about 2600 m/s. In general the Profile 2 reached a maximum depth major than Profile 1, respectively 50 m and 70 m.

6.2.3   Seismic attributes

Seismic attributes were obtained from a depth converted Common-Reflection-Surface stack, provided by Bruno et al., (2017), a fast and cost-effective alternative to standard reflection processing which allows to greatly improve signal-to-noise ratio in settings where structural complexity and high levels of ambient noise make it challenging to obtain a reliable seismic image. This work highlighted the conspicuous improvement of the seismic imaging through a comparison with the classical CMP processing, where stack have more continuous reflectors and better imaged dipping reflectors. Moreover, as seen in the work of Pussak et al. (2014), CRS stack greatly improved the quality of seismic attributes instead of the conventional CDP stack. Above reasons made possible carry out reliable attributes from the Solfatara caldera complex environment.
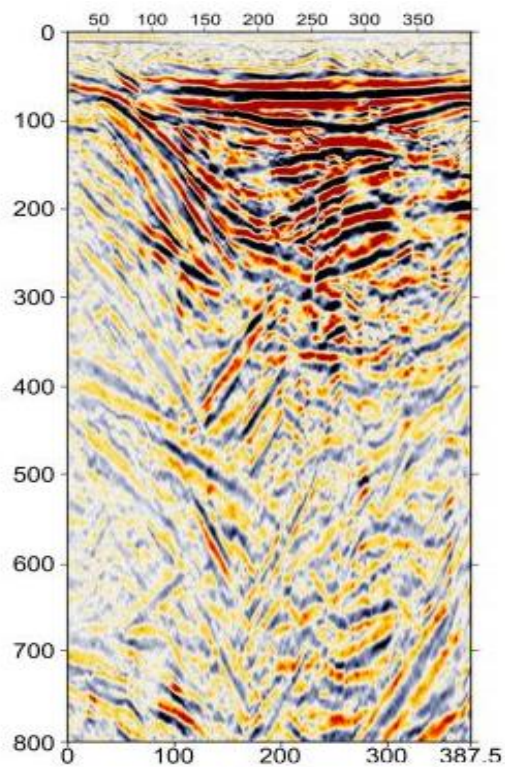
Figure 77 – CRS depth converted stack on Profile 1.

Seismic attributes were computed using OpendTect on the CRS stack obtained from the "Profile 1" for a maximum depth of 600m. In order to describe main structural features, connected with the hydrogeological activity of the caldera, were produced dip angle, energy, similarity and a set of texture attributes. These attributes were carried out with the depth gate from -5 to 5 m, and for texture attribute, step out nr=3 and GLCM size 32x32.
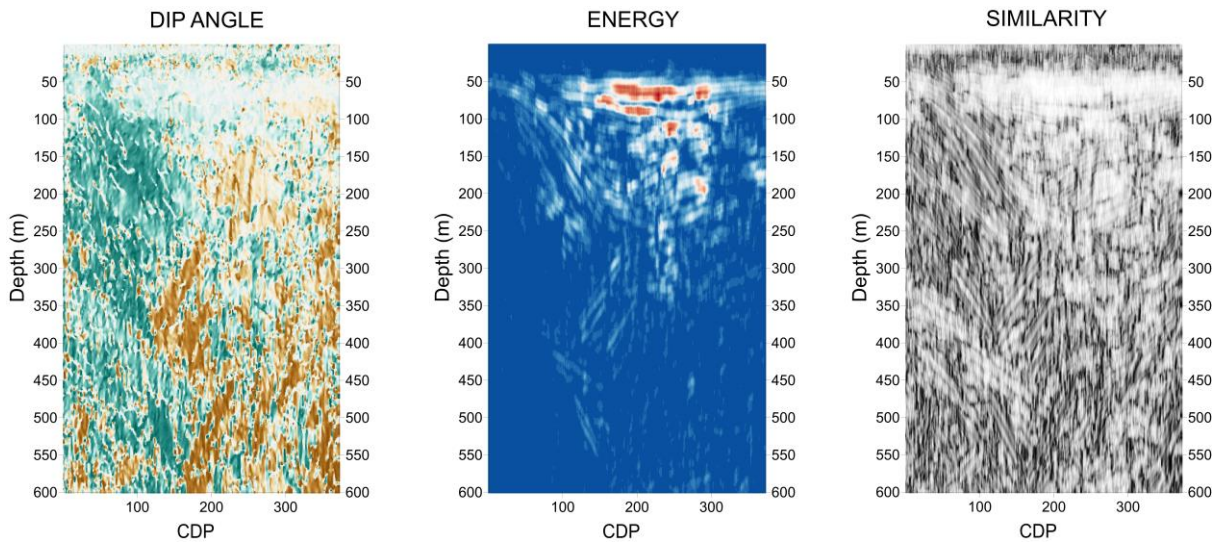
Figure 78 – Geometrical and physical attributes produced from the depth converted CRS stack on "Profile 1": in DIP ANGLE is plotted with brown with green gradational scale, where green colors are for reflections dipping toward the right side; in ENERGY blue white red palette, where the red color is for the maximum values of the attribute and in SIMILARITY a grey level scale, where the white represents the maximum values and the black the minimum.
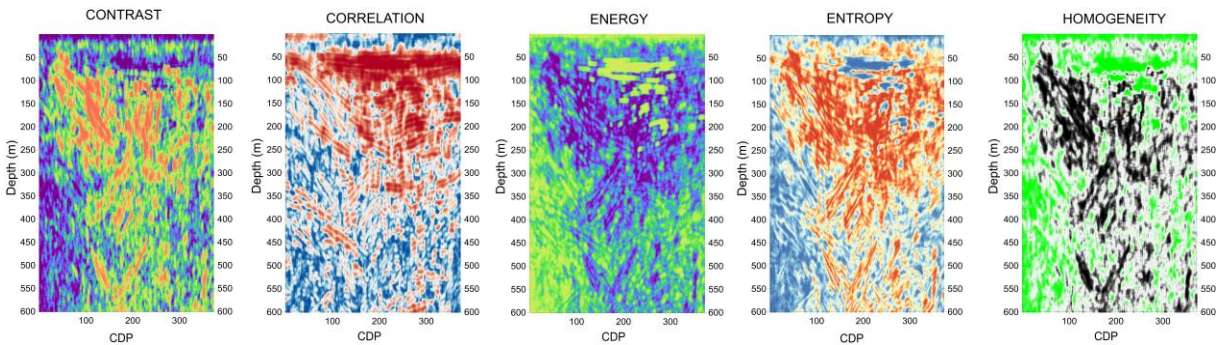
Figure 79 – Texture attributes produced from the depth converted CRS stack on "Profile 1": above attributes give complementary and redundant information about the texture of the seismic image.

The dip attribute highlighted a sub horizontal zone from 50 to about 100 m; below this zone are present narrow and large structures dipping toward the right side and the left with high angle to sub vertical.

The energy attribute showed a first horizontal high energy zone below the depth of 50 m (in red) cutted from a zone of medium energy at the CDP 250; a thinner high energy layer dipping towards the right side of the section reaching the minimum at the depth of about 100 m in correspondence to the CDP 250. Local high energy portions are presents from the CDP 230 to 300 in the depth range from 120 to 200 m. Medium to low energy zones with high dip are highlighted from the left border to the center of the section from the depth of 50 m to about 250 m. Local zones with medium to low energy are still present below the depth of 250 m in the CDPs range from 200 to 250.

Similarity highlights the main vertical interruptions of the seismic reflectors with black color and minors in grey; white color zones (i.e. high lateral continuity of reflectors) are generally computed in the same portions of high to medium energy.

With the aim to individuate similar portions of subsoil in terms of subsoil texture, were performed contrast, correlation, entropy, energy and homogeneity, from the texture attribute packet. Contrast highlighted major dissimilarities among neighbor pixels with the red color, which are disposed from high dip angle in the left part of the section to near vertical to the center of the section. The attribute correlation revealed continuity of reflectors (in red) with results close to the attribute similarity. Energy and entropy showed in opposite way a measure of order or disorder of the seismic image: yellow and blue colors, respectively for energy and entropy, individuated zones with high continuity and organization of reflectors.

6.3 Results

6.3.1 Integrated tomography

In Profile 1, geoelectrical and seismic tomographies were sampled at co-located positions for a maximum length of about 420 m applying the triangulation with linear interpolation (Surfer®) process (Figure 80). In Profile 2, geophysical models were sampled in co-located positions for a length of about 300 m (Figure 81).

Electrical resistivity tomography

Seismic refraction tomography
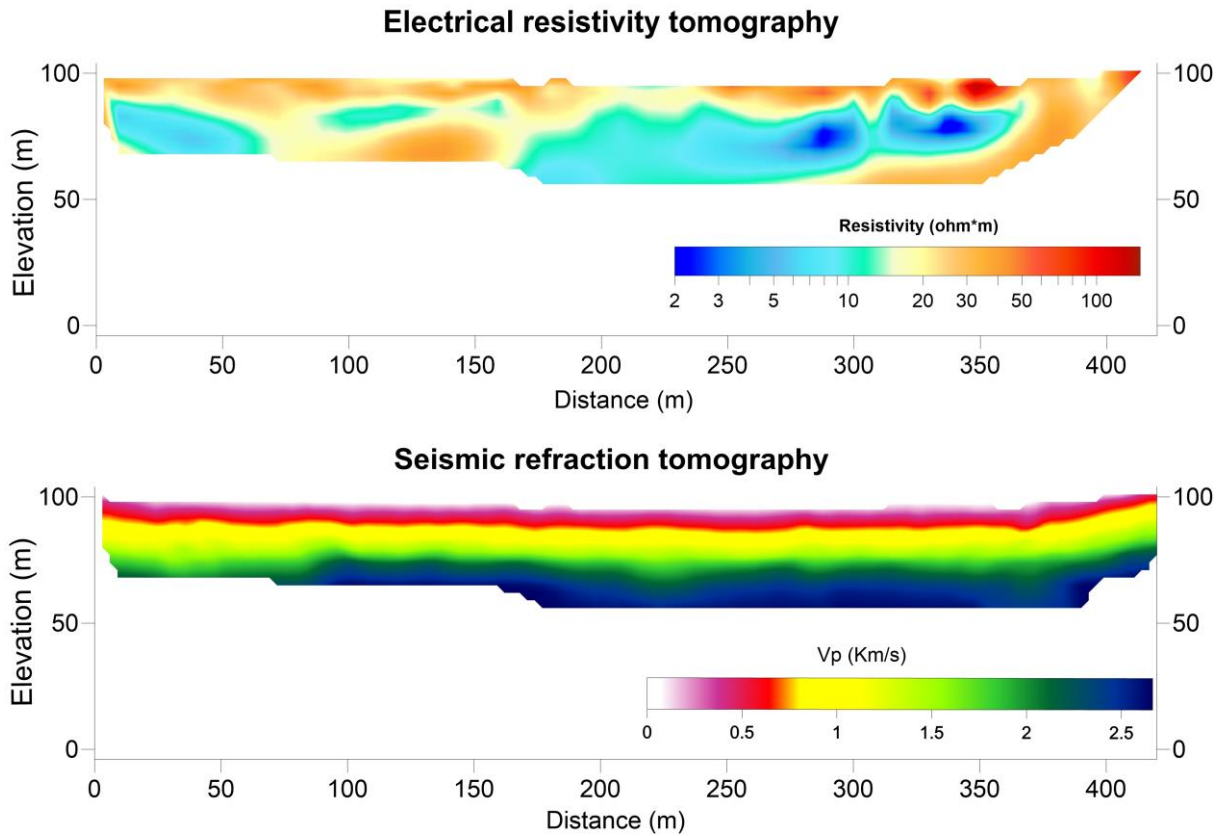
Figure 80 - Co-located ERT and SRT along the "Profile 1" for a length of 420 m.

Electrical resistivity tomography

Seismic refraction tomography
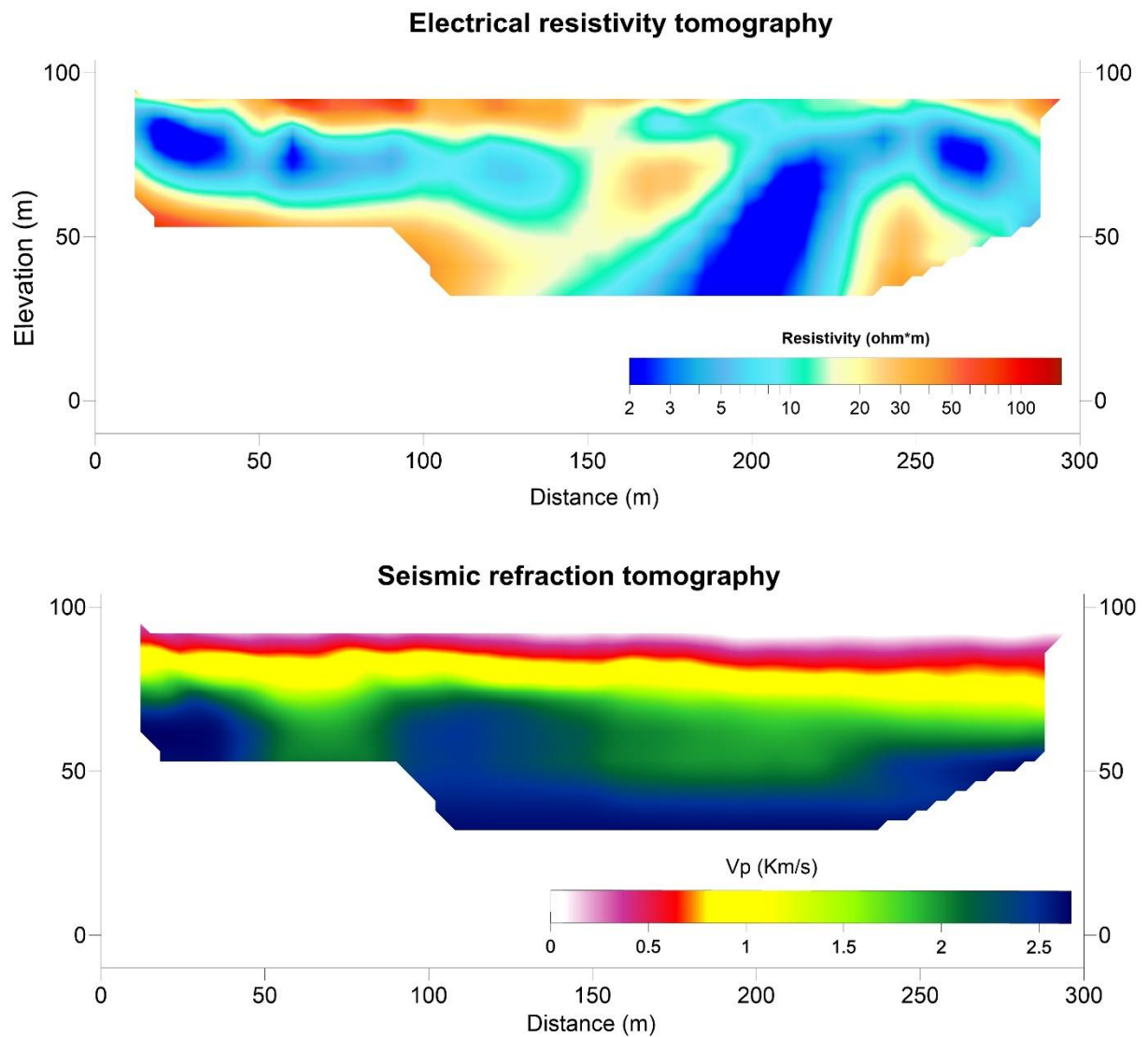
Figure 81 - Co-located ERT and SRT along the "Profile 2" for a length of about 300 m.

The integration code was applied in a range of k varying from 2 up to 10, in agreement with the geophysical facies attended from the geological setting. Moreover, minimum values of Silhouette index obtained for analyzed values of k from 2 to 10 are consistent with the assumed k range, i.e. the maximum value of the validity index is inner to the analyzed range.

The dataset was integrated by using the velocity in km/s and the $\log_{10}$ of the resistivity. The optimal clusters number was carried out with the Silhouette index as explained in § 4.1.3.

In order to preserve a high resolution from geophysical models, geoelectrical and seismic tomograms were sampled with a uniform grid spacing of 3x3 m, a compromise between the more resoluted SRT and the lower resoluted ERT. As explained in § 4.1, the ERT has a not uniform model spacing which coarses with the depth, with a maximum of vertical spacing of 6 m and lateral constant spacing of 10 m. The preliminary analysis to test the possible influence of the interpolation process in the integration technique is

developed as follows: were compared integration results from 10x6 m, the minimum resolution of the geoelectrical model, and from 3x3 m common grid spacing. In Table 10 are showed the validity index results from the two analyses with 10x6 and 3x3 m grid spacing for profile 1: in the first case the analysis showed the optimal cluster number k=6 and k=3 for the more resoluted dataset. The same test for the Profile 2 exhibited the same result for 3x3 and 10x6 grid spacing, with k=3. This result allowed using the more resoluted spacing, 3x3 m, as grid discretization for the sampling of co-located values in Profile 2.

Table 10 - Silhouette values for k analyzed from 2 to 10 for the 3x3 m and 10x6m grid spacing for the "Profile 1".
In the case of 3x3 m grid spacing the maximum is obtained for k=3 and for k=6 in the case of 10x6 m.

	grid spacing	
k	3x3	10x6
2	0.549	0.549
3	0.613	0.572
4	0.580	0.536
5	0.566	0.512
6	0.535	0.609
7	0.522	0.581
8	0.533	0.607
9	0.553	0.607
10	0.532	0.608

Table 11 - Silhouette values for k analyzed from 2 to 10 for the 3x3 m and 10x6m grid spacing for the "Profile 2".
In both cases the maximum is obtained for k=3.

	grid spacing	
k	3x3	10x6
2	0.583	0.577
3	0.654	0.644
4	0.596	0.598
5	0.579	0.575
6	0.593	0.539
7	0.574	0.541
8	0.582	0.555
9	0.579	0.573
10	0.564	0.566
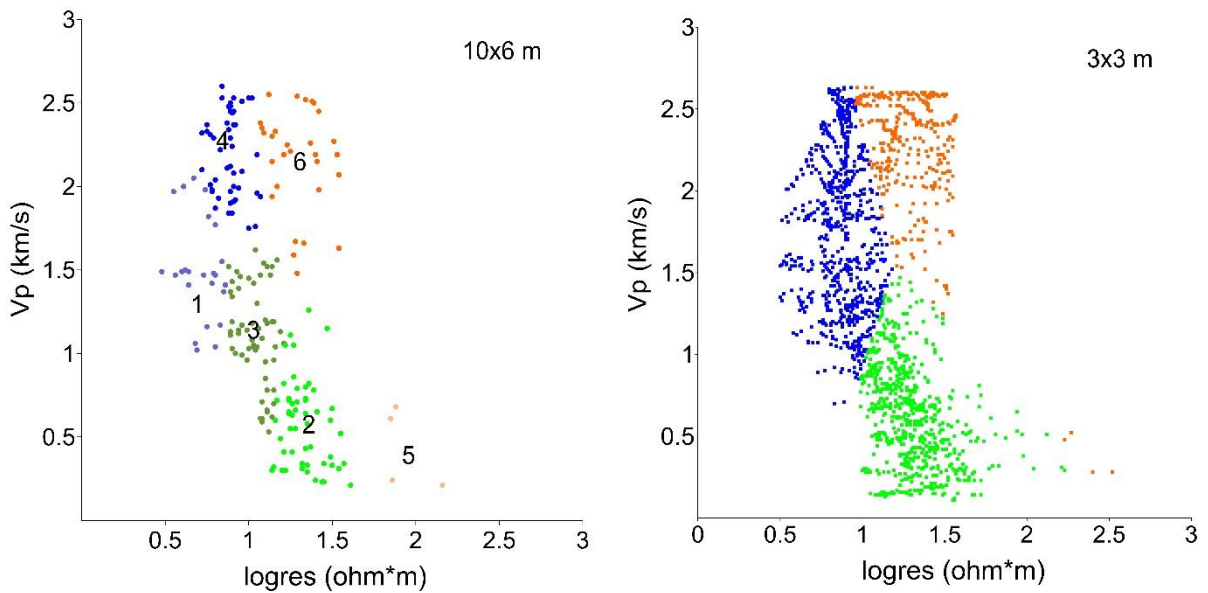
Figure 82 - Clustered bivariate dataset for "Profile 1" on the joint parameter space: on the left the analysis with grid spacing 10x6 m and on the right side with grid spacing 3x3 m. In the case of 10x6 m spaced dataset, the integration method individuated the optimum clustering for 6 clusters; in the case of 3x3 m spaced dataset, the integration method individuated the optimum clustering for 6 clusters.

In the analysis of Profile 1 dataset, clustered observation imaged with different colors (Figure 82) showed apparently different results for the low resoluted grid (10x6 m) with respect to the more one (3x3 m). In general the lower resoluted dataset has a minor density of observation, due to the loss of data in the shallower part of the ERT and in the horizontal spacing of the SRT. Moreover, clusters 1, 3, 4 are nested in the blue cluster of 3x3 m grid spacing analysis (Figure 82). Clusters 2 and part of 3 in the 10x6 m grid spacing analysis are nested in the green cluster in 3x3 m, and cluster 6 is directly comparable with the cluster orange in 3x3 m spaced grid dataset (Figure 82). Also at a visual inspection of observation on 2D tomographic space the nested distribution of cluster in 10x6 m spaced grid of Profile 1 leads to similar geophysical macro facies compared with the 3x3 m spaced grid (Figure 83). With the above considerations and in order to interpret similar geophysical facies on the two profiles, for the Profile 1 is take into consideration the analysis carried out on the 3x3 m common grid.
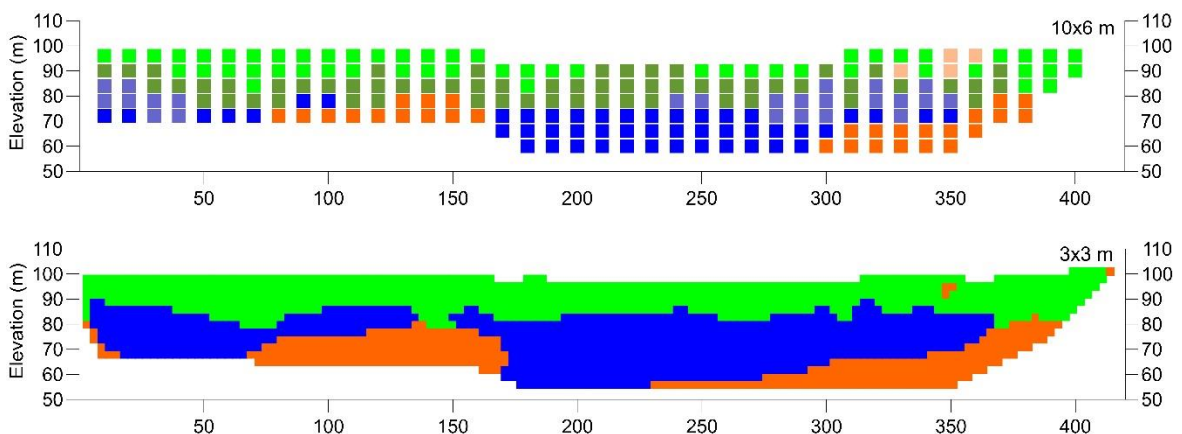
Figure 83 – Comparison between integrated tomograms in "Profile 1" with grid spacing 10x6 (above) m and 3x3 m (below). In the 10x6 spaced dataset nested clusters are visible with blue and light blue within the blue cluster in the 3x3 m spaced dataset.

119

For Profile 2, the two analyzed datasets partitioned for cluster assignment (Figure 84) are showed in the joint parameter space in case of 10x6 and 3x3 m grid spacing: the resulting images exhibited substantially the same macro geophysical facies and as attended, an image with higher resolution in the case of 3x3 m grid spacing.
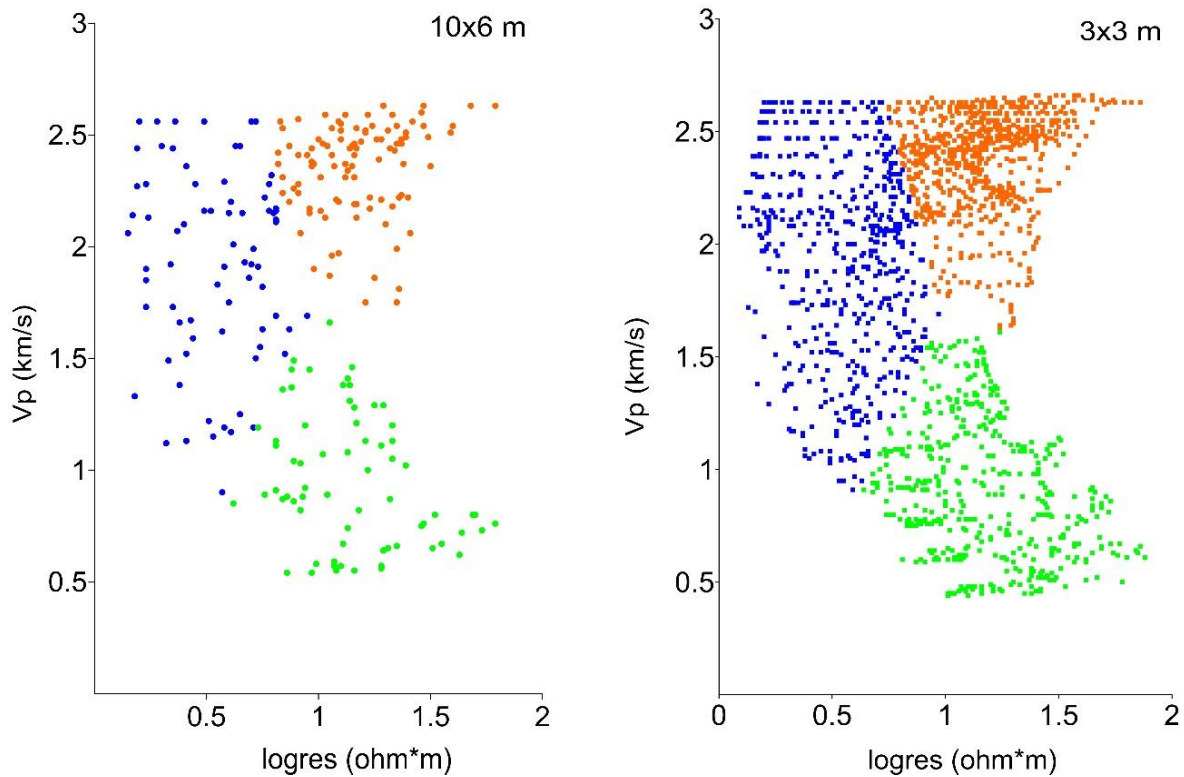
Figure 84 - Clustered bivariate dataset for the "Profile 2" on the joint parameter space: on the left with grid spacing 10x6 m and on the right side with grid spacing 3x3 m. In both cases the observations were separated in a similar way.

In Figure 85 are showed Silhouette values for different number of clusters for the dataset of Profile 1 on the 3x3 m common grid. The optimal clustering was obtained for the maximum value of Silhouette at k=3.
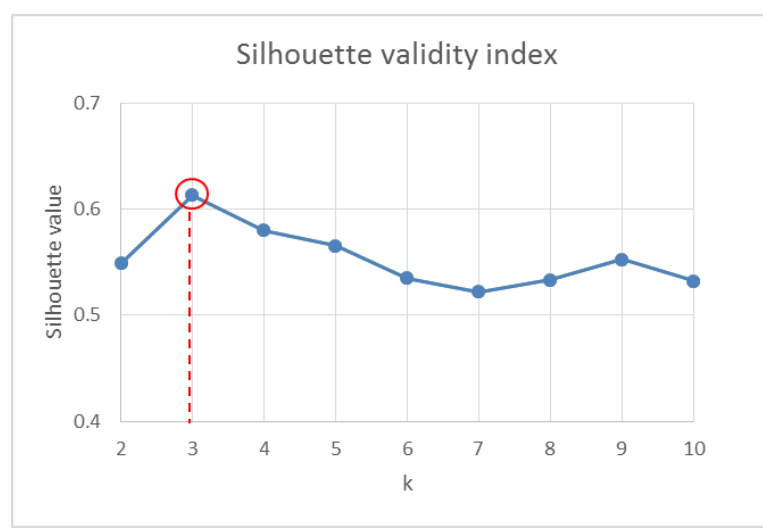
Figure 85 - Silhouette validity index of Profile 1 for different k analyzed (3x3 m grid spacing) where the optimal clustering is individuated for the maximum value in k=3.

Table 12 - Average values and standard deviations for 3 clusters individuated for the optimal clustering. In the second column there are values of Silhouette index for each cluster.

Cluster	SI for the optimal k =6	Log10 Resistivity (ohm*m)	Velocity (km/s)
1	0.63	1.29±0.19	0.63±0.31
2	0.64	1.29±0.2	2.26±0.37
3	0.56	0.88±0.13	1.76±0.48

In the Table 12 are summarized SI values for each cluster, with average values for the two variables and their standard deviation. In general and as expected at an increment of SI corresponds a decreasing of the dispersion of both variables. The highest values of SI were reached by clusters 1 and 2, respectively with values of 0.63 e 0.64 (Table 12).

The integrated imaging of input ERT and SRT are showed in 3 different manners: 1) the tomography is showed by assigning a different color to each cluster, 2) the SI with a color scale ranging from the minimum to the maxim value and 3) the version saturated with SI. Low values of SI, ranging from -0.15 to 0.3 are indicative of the main uncertainties in the clustered observations, where the colors red and white represent the interfaces of geophysical facies. The version saturated with SI take into consideration the uncertainty in clustering analysis: the comparison with the hard clustering version shows few differences into the cluster 1 and 3. The tomography of silhouette index highlighted main sub-horizontal contacts at the depth of about 20 m and a couple of vertical contact in the left side of the tomogram.
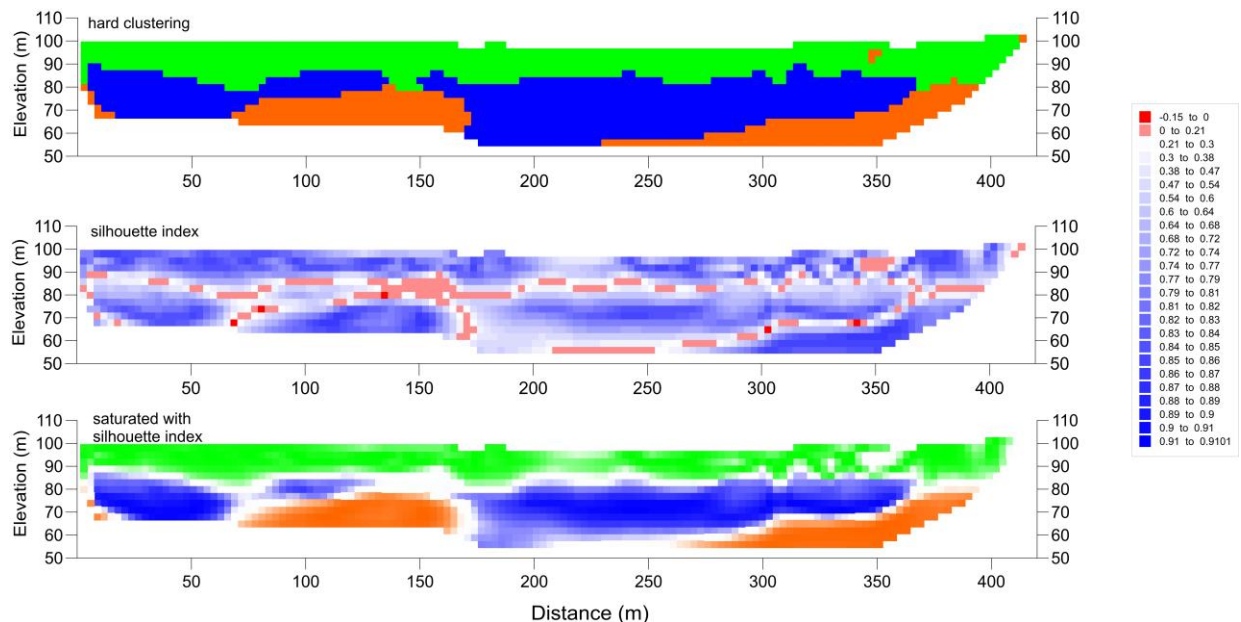
Figure 86 - Three different representation of the integration method on "Profile 1": at the top the hard clustering version obtained with k-means algorithm, in the middle the tomography of Silhouette index where the red color is representative of values lower than 0.21 and at the bottom the integrated tomography obtained with k-means algorithm saturated with Silhouette index.

In Figure 87 highlighted in black are observations of Profile 1 in the joint parameter space, with SI lower than 0.3, which are representative of the borders of 3 clusters. In black also few observations with high resistivity and low velocity ascribable as outlier.
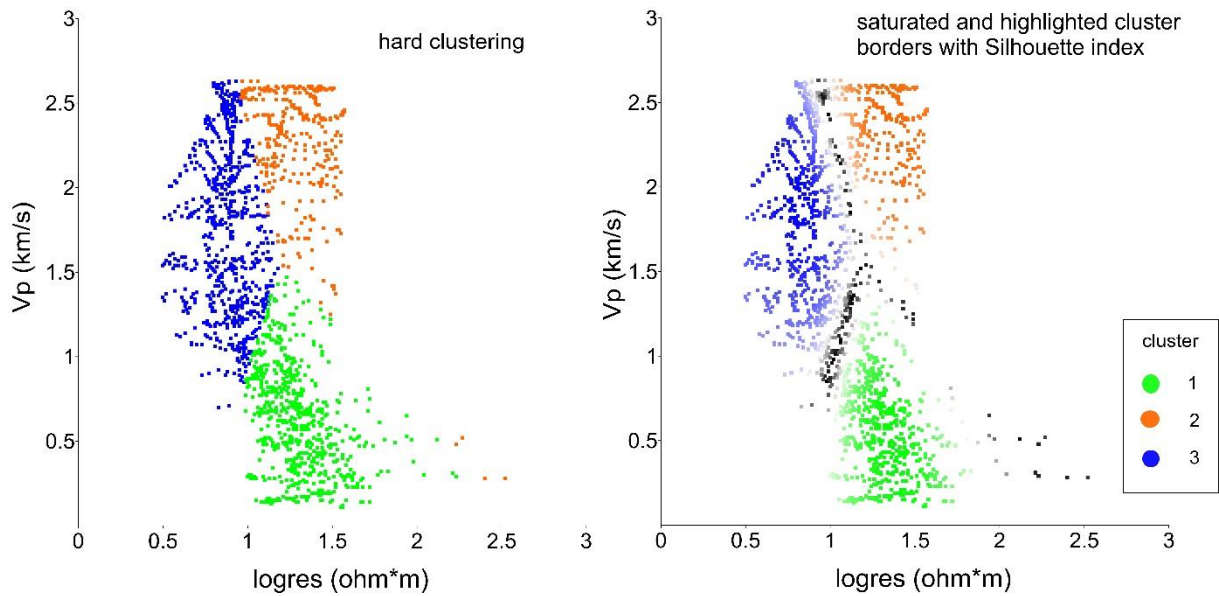
Figure 87 - The clustered dataset of Profile 1 (3x3 m grid spacing) in the joint parameter space with clusters represented with different colors. In the right graph the dataset has colors saturated with Silhouette index and observations with values than 0.21 are highlighted in black.

In Figure 88 Silhouette values for different number of clusters are showed. The optimal clustering was obtained for the maximum value of Silhouette at k=3.
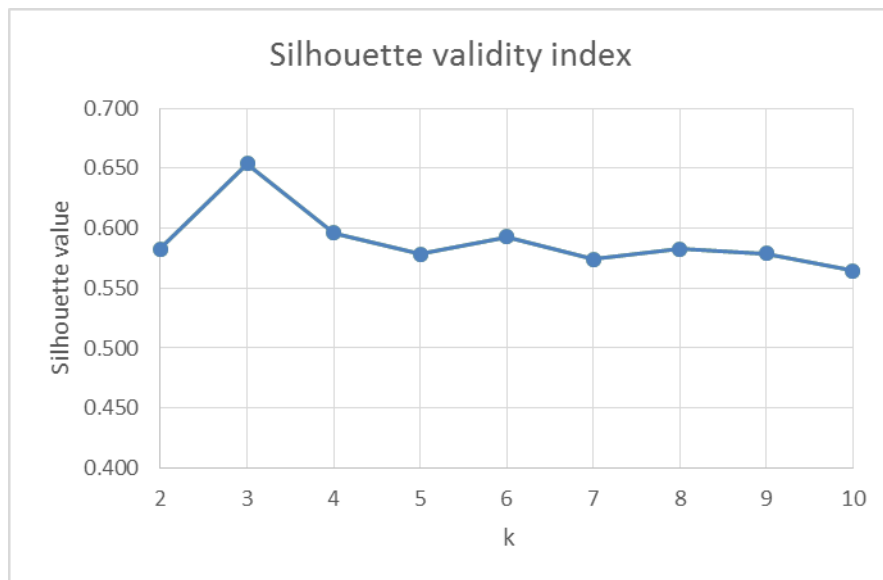
Figure 88 - Silhouette validity index of Profile 2 for the different k analyzed, where the optimal clustering is individuated for the maximum value in k=3.

Table 13 - Average values and standard deviations for 3 clusters individuated at the optimal clustering for Profile 2. In the second column there are values of Silhouette index for 3 clusters.

Cluster	SI for the optimal k =3	Log10 Resistivity (ohm*m)	Velocity (km/s)
1	0.66	1.18±0.26	0.91±0.3
2	0.76	1.17±0.20	2.37±0.21
3	0.53	0.55±0.21	1.94±0.44

In the Table 13 are summarized SI values for each cluster, with average values for the two variables and their standard deviation. In general and as expected at an increment of SI corresponds a decreasing of the dispersion of both variables. The highest value of SI were reached by clusters 2, with a value of 0.76.
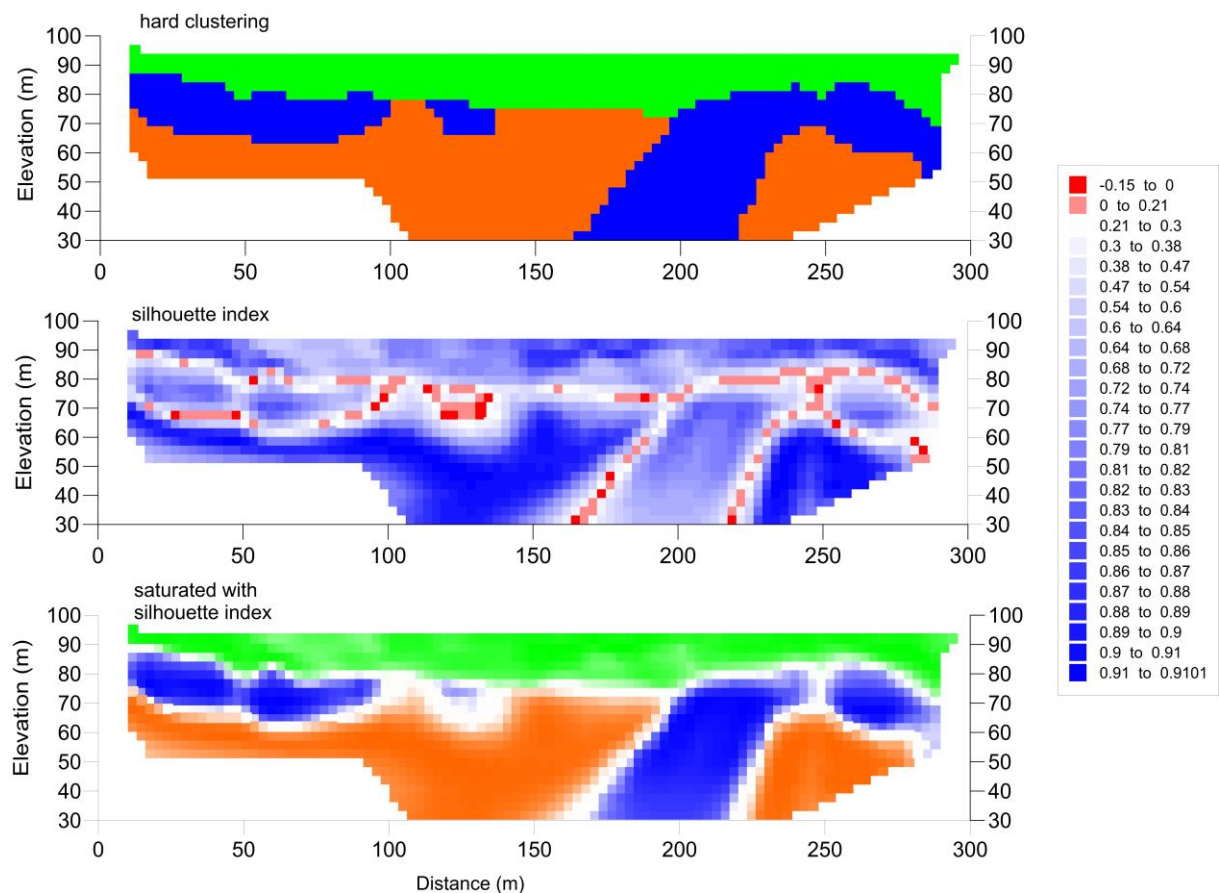
Figure 89 - Three different representation of the integration method on Profile "2": at the top the hard clustering version obtained with k-means algorithm, in the middle the tomography of Silhouette index where the red color is representative of values lower than 0.21 and at the bottom the integrated tomography obtained with k-means algorithm saturated with Silhouette index.

As the previous analyses, the integrated imaging of input ERT and SRT are showed in 3 different manners: 1) the tomography is showed by assigning a different color to each cluster, 2) the SI with a color scale ranging from the minimum to the maxim value and 3) the version saturated with SI. Low values of SI, ranging from -0.15 to 0.3 are indicative of the main uncertainties in the clustered observations, where the colors red and white represent the interfaces of geophysical facies. The version saturated with SI take into consideration the uncertainty in clustering analysis: the comparison with the hard

clustering version shows few differences into the cluster 3 (i.e. in blue) at a distance of about 120 m and an elevation of about 80 m, where observation assigned at the cluster have a high uncertainty.
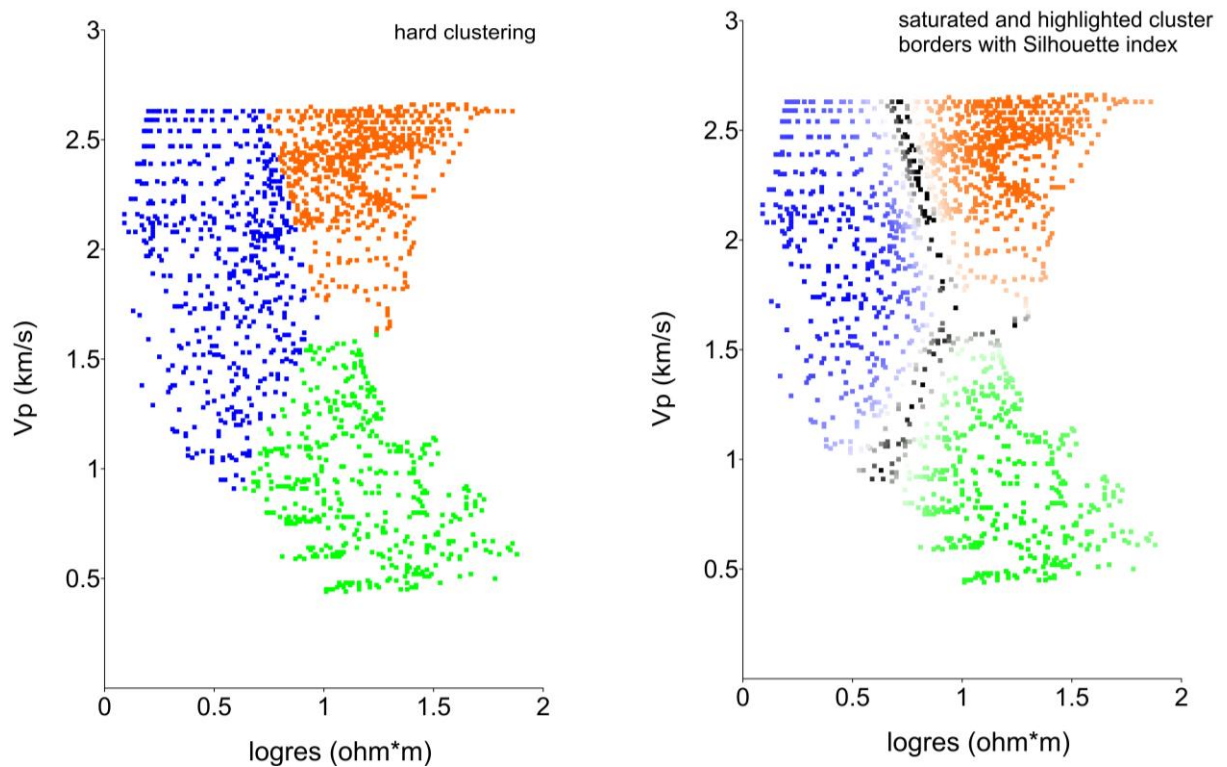
Figure 90 - Clustered dataset in the joint parameter space with clusters represented with different colors. In the right graph the dataset has colors saturated with Silhouette index and observations with values than 0.21 are highlighted in black.

In both profiles the cluster 1 (i.e. green) is characterized by low values of resistivity and velocity and in the 2D tomographic space starts as a horizontal layer from the topographic surface to a maximum depth of about 20 m. Clusters 2 (i.e. orange) and 3 (i.e. blue) are representative of two higher velocity geophysical facies, and respectively with high and low resistivity. These clusters are disposed in the 2D tomographic space at depths greater than 10 m up to the bottom of the integrated images (Figure 86 and Figure 89).

6.3.2   Integrated seismic attributes

As described in the § 4.2, the first step in order to generate a neural network based on the dataset distribution, is estimate the two net dimensions (i.e. the number of the neurons along the X and Y directions). Through the principal component analysis one can estimate the ratio between X and Y as the ratio between the variances of first two dimensions. As suggested by Vesanto and Alhoniemi (2000), the number of the neurons should be bigger than the clusters into the dataset (i.e. expected homogeneous zones within the seismic image).

In order to find out the best integrated images, were carried out several SOM analyses with different grid dimensions on the same dataset. As a tool to assess the statistical goodness of observations classified within neurons was calculated a coefficient of intra-neuron homogeneity, based on coefficient of variation of attributes. Thus, a visual inspection allowed to choose the most meaningful images, based on the subsurface features expected.

In the first attempt was used a dataset composed by dip angle, energy, entropy and similarity after the data sphering. The variables contribution and their quality were evaluated in the correlation circle (Abdi and Williams, 2010): in Figure 91, a, are represented the three variables with their cos2 (i.e. square cosine, squared coordinates) where energy and similarity showed a better representation on the principal component compared with dip angle. Also the contribution in principal component is lower for dip angle. Anyway the values calculated have not so much difference to exclude dip angle from the analysis. The variances for the first two dimensions are respectively 35.2% and 26.2 and their ratio is equal to 1.34.
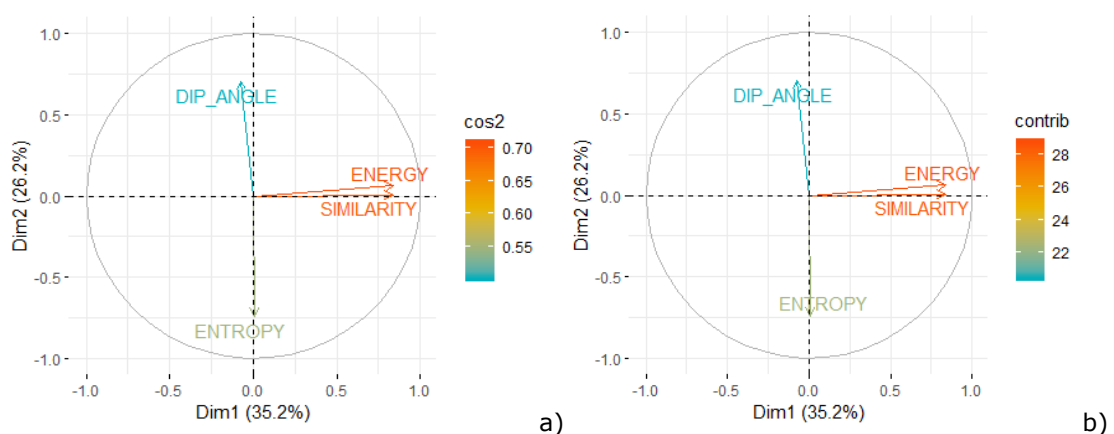
Figure 91 - Principal components analysis for the dataset composed by DIP ANGLE, ENERGY, ENTROPY and SIMILARITY seismic attributes: in a) are represented the attributes in the correlation circle with square cosine highlighted with the color scale, in b) are represented the attributes in the correlation circle with contribution highlighted with the color scale. In both analyses ENERGY and SIMILARITY showed the best representation in first two dimensions.

The multivariate dataset composed by dip angle, energy, entropy and similarity was employed as input to the SOM analysis after the data sphering process. The neural grid tested were 7x5, 10x7, 12x9, and 14x10 with the side ratio close to 1.34. The neural grid was obtained in the hexagonal disposition with size 7x5 (Figure 93) and the net was trained for 9000 iterations.

Analyzed dataset can be explained by comparing the RGB color map, the heatmaps and the Umatrix, in order to highlight the main features. The observations classified into the neurons 5, 6, 7, 13 and 14 (Figure 93) are representative of zones with energy and similarity medium to maximum, sub horizontal dip and medium to low entropy (Figure 92). Moreover the Umatrix showed a high distance among neurons, consistent with the very different average values mainly of energy and entropy (Figure 92). On the integrated section are imaged as sub-horizontal zones from the depth of about 50 m up to 300 m. The neurons 21, 28 and 35 show high dip angle, low values of similarity and energy, and medium to high entropy. Are well imaged as thin sub-vertical path from the depth of 100 m up to 350 m, mainly into the central part of the section, cutting the sub-horizontal portions highlighted by the neurons 5, 6, 7, 13 and 14. Also the neuron 20 has similar values, with lower dip and a slightly higher similarity. Neurons 2 and 9 contain observation with medium to low energy, high dip (toward left) and entropy and medium similarity and are mainly present on the left side of the section from the depth of about 50 m up to 300 m, creating an asymmetrical zone on the section. In general observations colored with light green on the integrated section (Figure 93) show very low energy, variable dip, medium similarity and entropy and are spatially distributed at the border of above described observations.
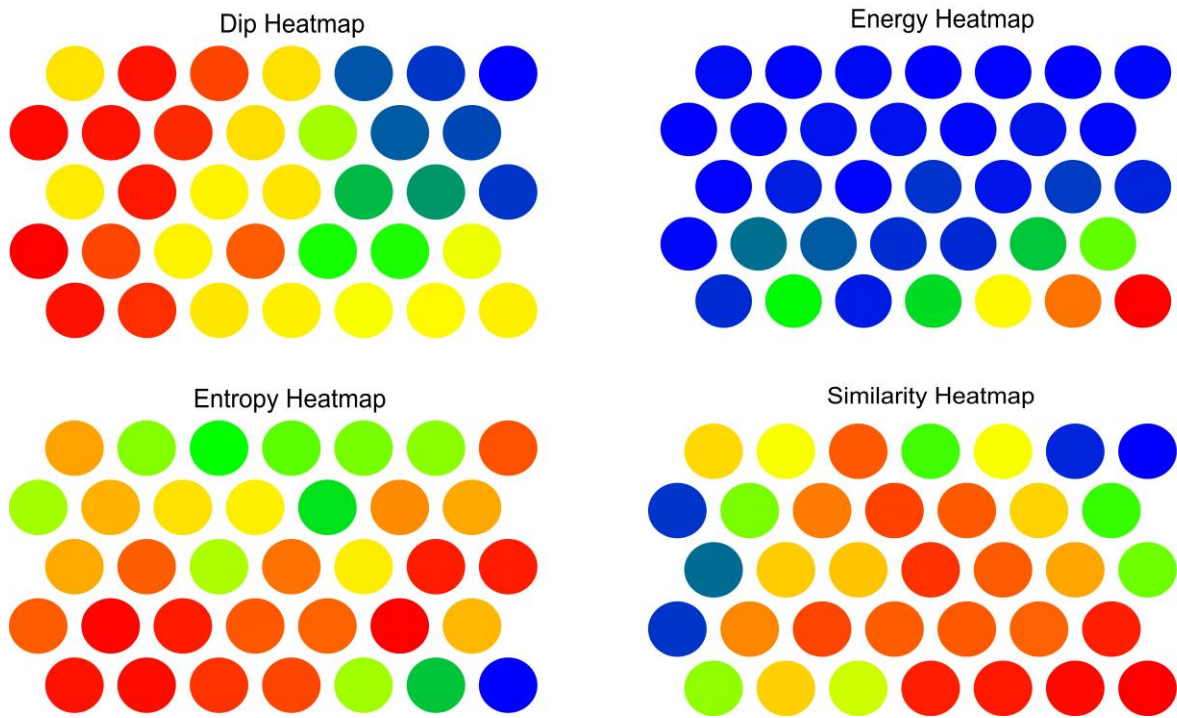
Figure 92 - Distribution of average attributes values on 7x5 trained SOM for the dataset composed by DIP ANGLE, ENERGY, ENTROPY and SIMILARITY: cold colors represent low values and warm colors represent high values. In the left side of the map are classified observations with high dip (toward right), low to medium similarity, low energy and medium to low entropy. Maximum values of energy and similarity with low entropy and sub-horizontal reflectors are classified in the lower right part of the map.
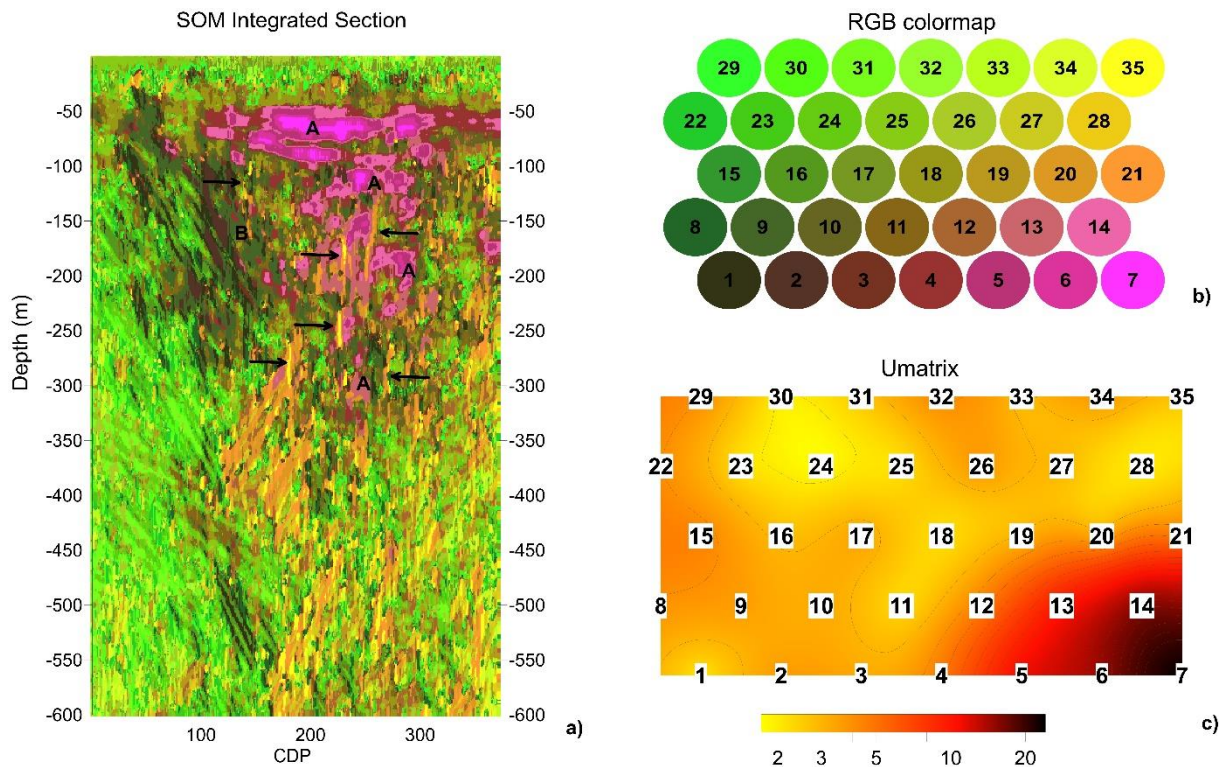
Figure 93 - In the left side a) the integrated section with similar zones individuated with colors. In the upper right side b) the RGB color map based on the position of neurons and c) the U-matrix. With "A" are highlighted sub-horizontal high energy and similarity zones interrupted by sub-vertical orange colored zones with high entropy and highlighted with black arrows. With the letter B the asymmetrical left border of the crater.

127

A second analysis with a 12x9 neural grid was carried out, in order to evaluate differences on the 2D distribution of observations across the net, especially to have a more gradual distribution of attribute energy. In general high values of similarity are well correlated with energy and classified into the upper right corner of the net (Figure 94). These neurons have low values of entropy and about 0 dip, as revealed by the analysis with net size 7x5. Neurons with dip high in absolute value are well correlated with low values of similarity and medium to high values of entropy and constitute sub-vertical path in the central portion of the integrated section. The same became less dipping in the lower part of the section. As expected, neurons with medium to high energy are imaged with more gradational colors, because of the higher number of neurons. The Umatrix in the upper right corner (Figure 95), in agreement with this distribution, showed lower distances (i. e. lower gradient of energy) among neurons with respect to the previous analysis (i.e. net size 7x5). Moreover, the higher grid size led to a major discrimination of observations with very low energy, low to medium similarity and energy and variable dip.
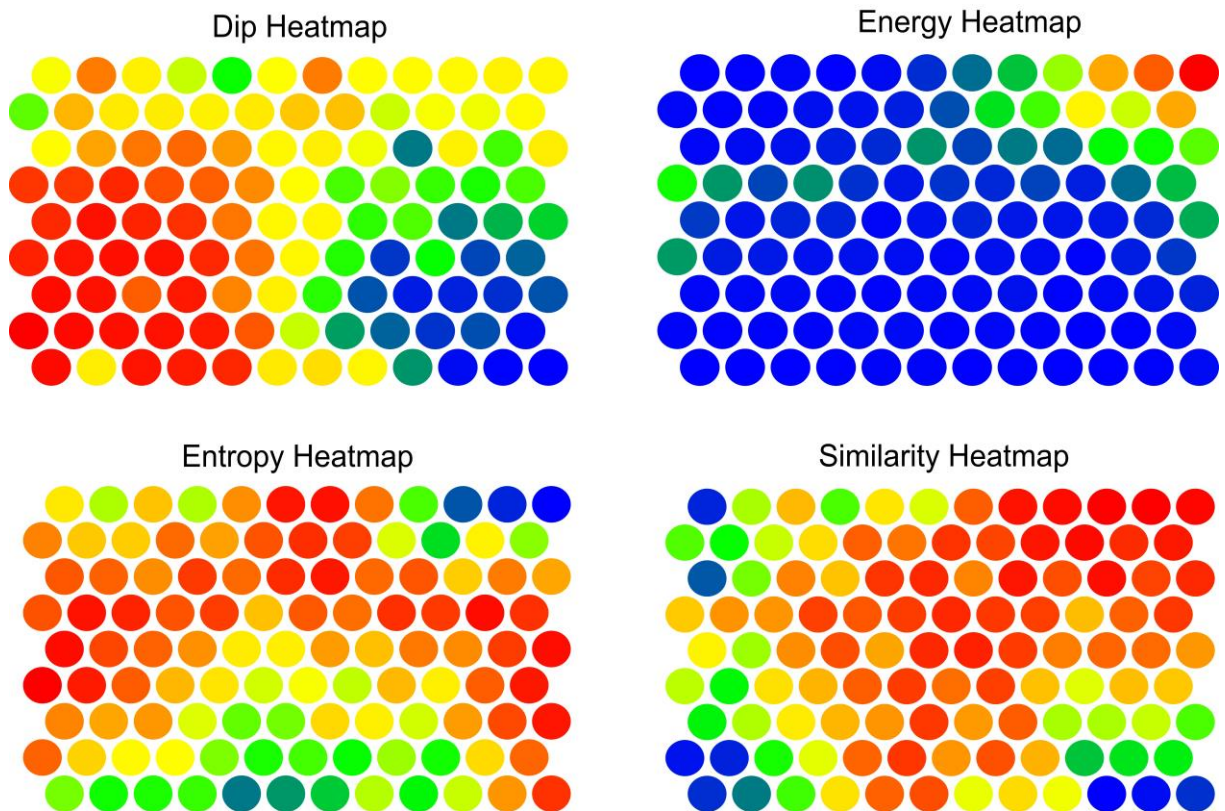
Figure 94 - Distribution of average attributes values on 12x9 trained SOM for the dataset composed by DIP ANGLE, ENERGY, ENTROPY and SIMILARITY: cold colors represent low values and warm colors represent high values. In the left side of the map are classified observations with high dip (toward right reflectors), low to medium similarity, low energy and medium to low entropy. Maximum values of energy and similarity with low entropy and sub-horizontal reflectors are classified in the upper right part of the map.
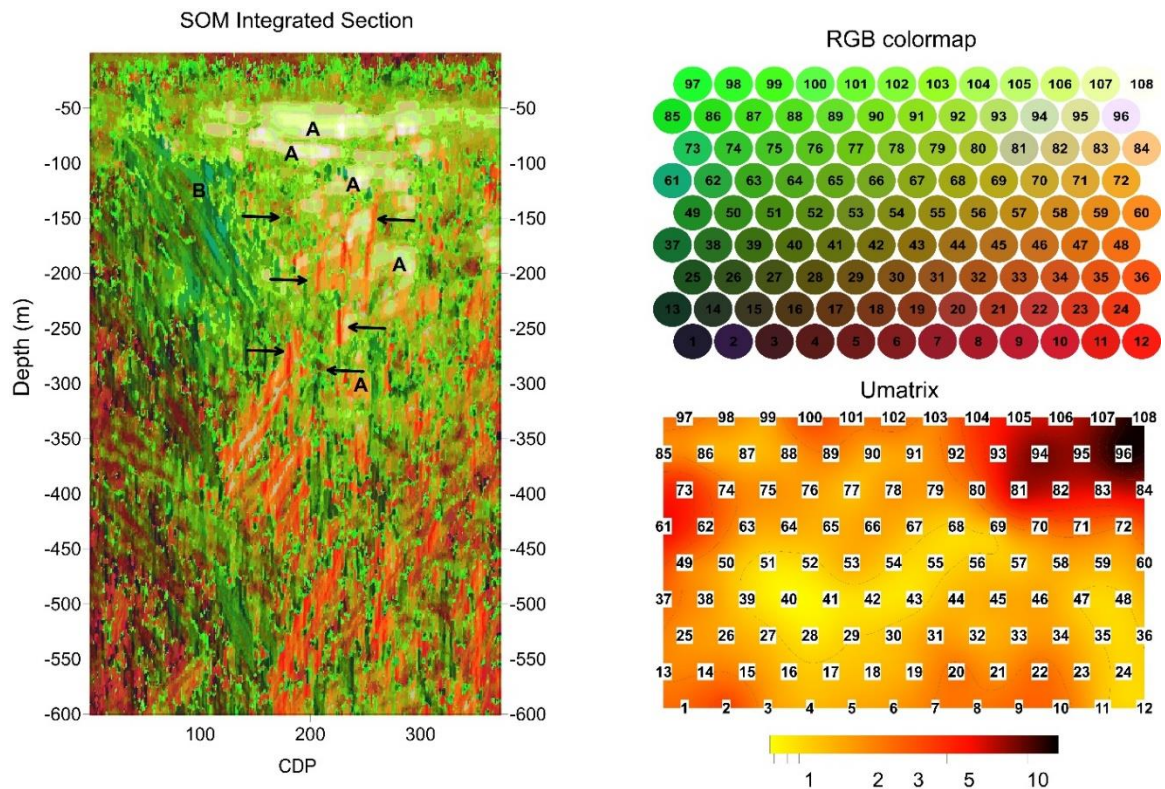
Figure 95 - In the left side the integrated section with similar zones individuated with colors. In the upper right side the RGB color map based on the position of neurons and the U-matrix. With "A" are highlighted sub-horizontal high energy and similarity zones interrupted by sub-vertical and thin zones and highlighted with black arrows, representing high entropy observations. With the letter B the asymmetrical left border of the crater.

In order to find out the presence of other pattern not revealed by the previous analyses, was analyzed a dataset composed by dip angle, similarity, energy, instantaneous frequency and several complementary and redundant texture attributes. The PC analysis into the correlation circle resulted in a very low contribution for instantaneous frequency and also a low quality of representation in first two principal components.
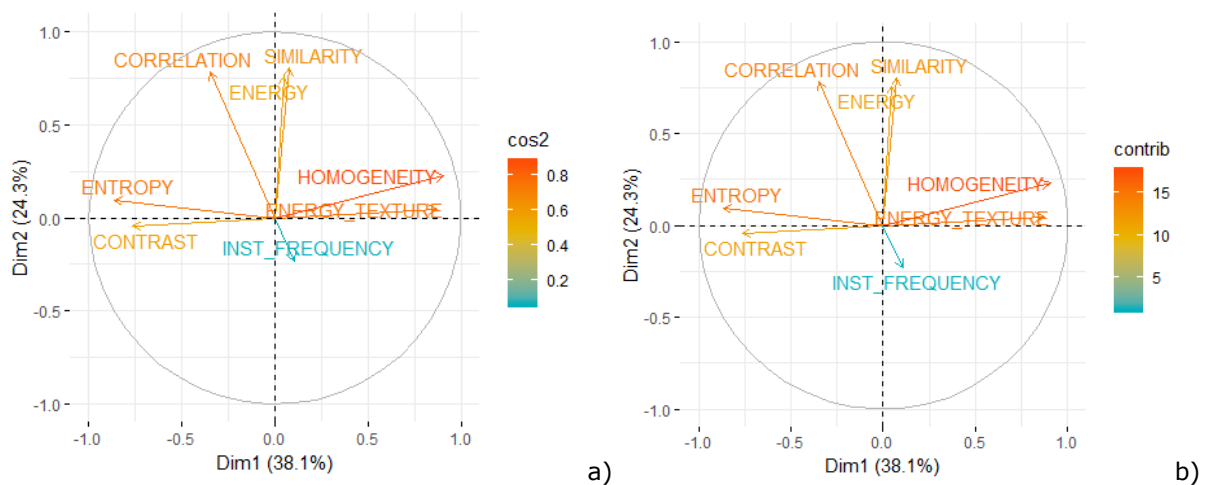
Figure 96 - Principal components analysis for the dataset composed by geometric, physics and texture seismic attributes: in a) are represented the attributes in the correlation circle, where the square cosine is highlighted with the color scale, in b) are represented the attributes in the correlation circle with contribution highlighted through the color scale. In both analyses instantaneous frequency showed the worst representation in first two dimensions, with values much lower than other variables.

129

Due to the low contribution of instantaneous frequency, the attribute was removed from the dataset because of possible source of noise in the SOM analysis. The PC analysis carried out on the dataset, after removing the instantaneous frequency, gave more uniform results among the different variables, either in terms of quality of representation (i.e. cos2) or in contribution. The variances for the first two principal components are 43.5% and 27.4 and their ratio is equal to 1.58.
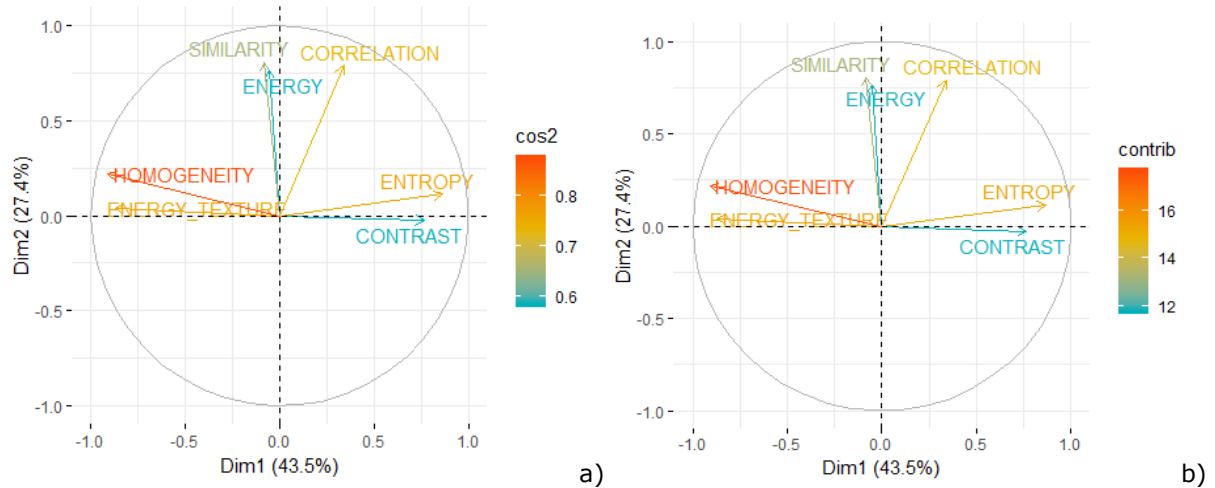
Figure 97 - Principal components analysis for the dataset composed by geometric, physics and texture seismic attributes: in a) are represented the attributes in the correlation circle, where the square cosine is highlighted with the color scale, in b) are represented the attributes in the correlation circle with contribution highlighted through the color scale. In both analyses Homogeneity showed the best representation and Energy the worst, in first two dimensions.

The SOM analysis on the multivariate dataset, composed by dip angle, energy, similarity and a combination of texture attribute, was performed on a grids 6x4, 8x5, 9x6, 12x8 and 16x10 with a sides ratio close to 1.58 (i.e. ratio between variances of first two principal component). At a visual inspection the neural net 12x8 was the most consistent with the geological setting of Solfatara and then chosen as best result.

As seen in Figure 97, energy (texture) results well correlated with homogeneity and has a similar 2D distribution across the neural network (Figure 98). Entropy, being a measure of disorder showed an opposite trend, with higher values where energy (texture) and homogeneity have low values. High values of contrast are classified within neurons 60, 72, 83, 84, 95, 96. Energy, similarity and correlation exhibited a similar distribution across the 2D net, with maximum values on the neurons 10, 11, 12, 23, 24 and 36. Neurons 6, 7 and 20 have high values of texture attributes as energy, homogeneity and correlation, medium to high similarity and medium to low energy. In the integrated section (Figure 99) neurons 6, 7, 10, 11, 12, 20, 23, 24 and 36 are disposed as sub-horizontal layers truncated laterally by observations classified within the neurons 60, 72, 83, 84, 95 and 96. In particular, observations within the neuron 84 are mainly disposed on the left side of the integrated image (Figure 99) dipping toward the center of the section. The others are mapped on the integrated section as thick vertical path. These features were

depicted from the depth of 50 m up to the maximum depth of 350 m; the remaining zones of the integrated section showed no particular features.
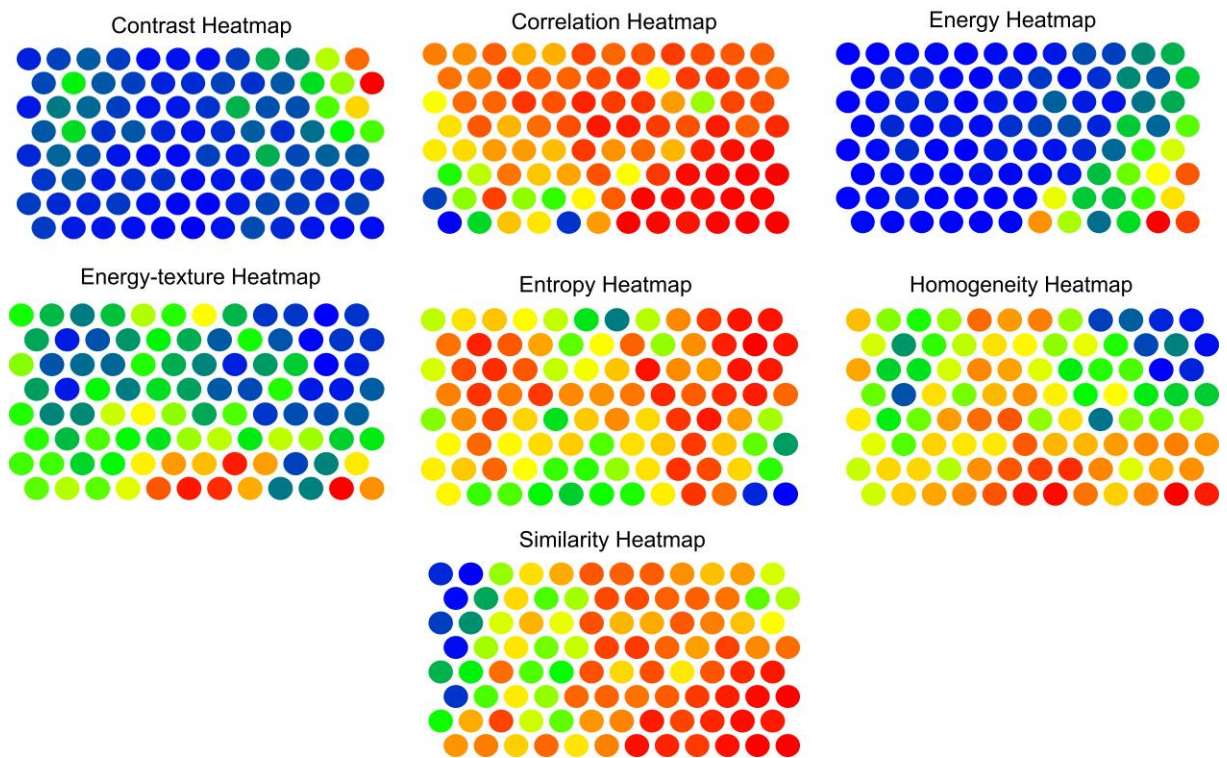
Figure 98 - Distribution of average attributes values on 12x8 trained SOM for the dataset composed by geometric, physics and texture: cold colors represent low values and warm colors represent high values. The distribution of attributes across the 2D map appears complex but consistent among different attributes.
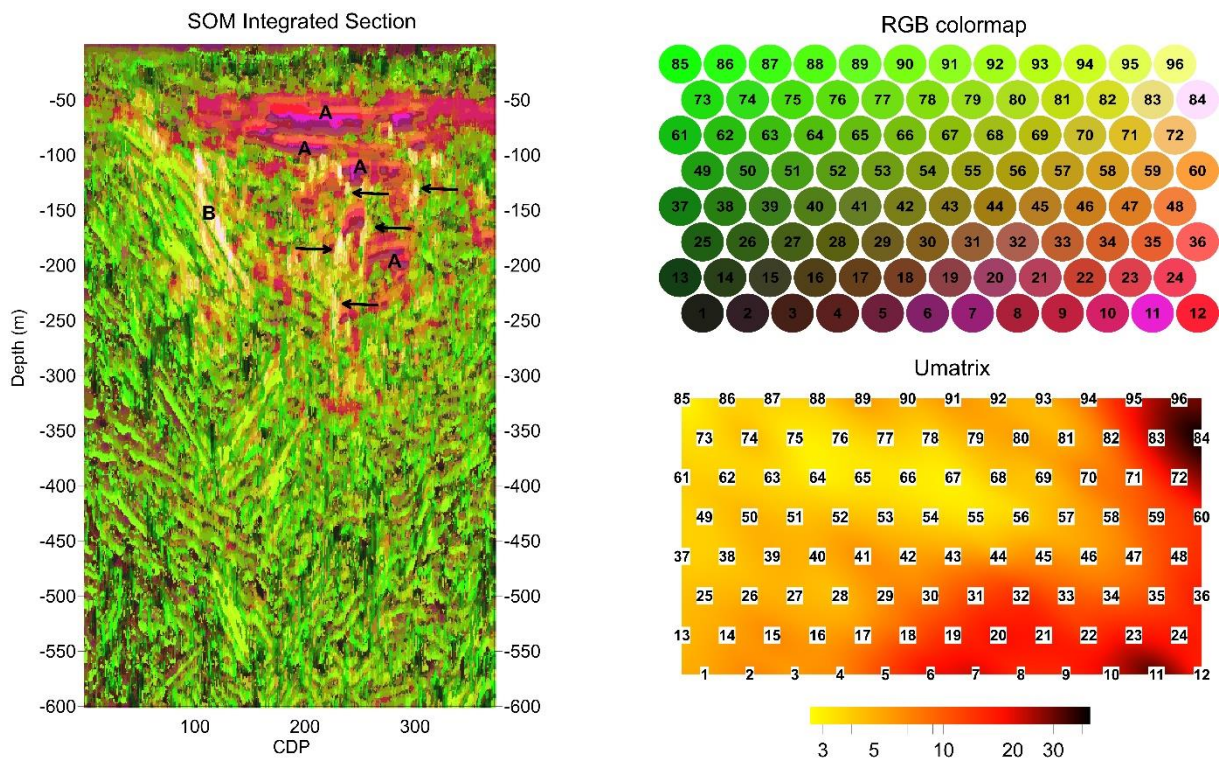
Figure 99 - In the left side the integrated section with similar zones individuated with colors. In the upper right side the RGB color map based on the position of neurons and the U-matrix. With"A" are highlighted sub-horizontal high energy, similarity and correlation, and low contrast zones, interrupted by sub-vertical and thin zones and highlighted with black arrows, representing high entropy observations. Whit the letter "B" the asymmetrical left border of the crater.

131

7 Discussion

The groundwater resource is nowadays requiring a huge attention: the changing in climate and the major research of water increased hydrogeophysical studies. The improved geophysical techniques of subsoil exploration, allow to characterize spatially principal hydrogeological information as: stratigraphy, system boundaries, permeability and saturation and the connectivity between aquifers and aquitards.

The approach I proposed in this thesis, for the production of a subsurface conceptual model for hydrogeological purposes, is based on the geophysical characterization of the subsurface by integrating high resolution datasets obtained through two of the more diffused surface geophysical techniques, the electrical resistivity tomography and the seismic refraction/reflection survey.

As seen in the §3, geophysical parameters as electrical resistivity and P waves velocity are influenced by several factors as porosity, mineralogy, texture, fluids saturation (i.e. gas or liquid). This make hard to distinguish different homogeneous zone (i.e. different lithology) in the case of univariate geophysical techniques, i.e., low resistivity values can be obtained for clayey soils or for high fractured saturated rocks or clayey rocks (Figure 31). In the tomographic integration, the co-located knowledge of electrical resistivity and seismic velocity reduced this uncertainty and increased the characterization level of the subsoil. Without a priori geological informations, integrated tomography allowed to recognize different geophysical facies, giving a subsurface discretization meaningful in the exploration phase. Ranges of bivariate parameters (i.e. electrical resistivity and seismic velocity) individuated by clustering the datasets, also without stratigraphic information, can be used in a qualitative way in order to estimate main lithologies and their changes (i.e. deposits from rocks, saturated from unsaturated). In presence of stratigraphic log information, this method allowed to infer this information to the geophysical facies individuated, and distinguish different lithologies settings (i.e. characterize potential aquifers).

Seismic reflection can resolve system boundaries, changes in stratigraphy and fractures, but not directly measures hydraulic parameters. An aquifer is generally controlled by locating geological or topographic boundaries. Geological boundaries can take a range of different forms but are often associated with fault planes and the discontinuity of lithology. High resolution seismic reflection is powerful in determining the location of these boundaries. Seismic attributes are a quantitative measure of seismic features, depending directly from the subsurface setting. In this optic, seismic attributes were used as a quantitative parameters to characterize main features of an aquifer system. Furthermore, once aquifer properties can be connected to the character of the seismic images by drill hole informations, it may be possible to infer that property more generally with the seismic data in a particular setting. The lack of stratigraphic informations reduces

the degree of knowledge of subsurface, but changes in lithology and fractured zones are still well recognized.

The two developed methods of integration, the first based upon k-means algorithm and silhouette index on tomographic data and the second based on Self-Organized maps and seismic attributes, allowed studying the subsoil in terms of similar and homogeneous geophysical facies at different scales. In connection with geophysical parameters applied, as electrical resistivity, P-waves seismic velocity and seismic attributes, were interpreted homogeneous zones in a geological/hydrogeological point of view. More in specific, the integration of tomographic dataset, according with the average values of electrical resistivity and P-waves seismic velocity, was employed to distinguish porous saturated/non saturated zones from low permeability media. The integration of seismic attributes, was applied in order to reveal principal structural pattern, as faulted/fractured zones, main paths for fluids (i.e. gas and liquids) and major changes in lithology revealed by impedance contrast, changes in frequency and textural features of the seismic images.

7.1   Shallow aquifer characterization: Laterina

In the case study of Laterina, a shallow multi-layer aquifer was explored, where the shallower is represented by an unconfined aquifer in alluvial deposits and the second by a confined aquifer in fractured rocks. With the aim to verify the results a borehole was performed, placed at a distance of about 50 m to the east part from the survey line (for logistic reasons) (Figure 54). A high resolution characterization of the lithologies drilled was carried out also through the use of geophysical loggings, in particular were executed caliper, electrical resistivity, gamma ray, temperature and conductivity of the liquid. In Figure 53 geophysical logs are in agreement with the stratigraphic log, as follows: from the surface up the depth of 9 m the subsoil consists of brown silt and sand (bna), from the depth of 9 up to 14 m a layer of polygenic pebbles with sand (bna), then a level of clay of 6 m (VIL); from the depth of 20 m to the bottom of the borehole (42 m) alternated levels of sandstone and marls (Monte Cervarola unit) with a different degree of fracturation. In the depth range from 32 to 39 m the Monte Cervarola unit revealed a very high fractured level, with the fracturation degree increasing from the upper part to the lower. This state of fracturation can be associated to the high angle direct fault presents at the east border of the valley (Figure 45). As seen in Figure 45 the water depth from surface has a variable value across the valley and a measured depth of 5.9 m along the survey line.

As explained in § 4.1 in general the geophysical models from ERT and SRT have a different spatial discretization (Figure 32), thus was necessary an interpolation process for the co-location of the bivariate dataset. With the aim to evaluate possible effects of the grid spacing in the co-location process, two integration processes on 2x7 and 2x2 m spaced datasets were carried out. In both cases the Silhouette index exhibited a maximum for 6 clusters with similar distribution of the bivariate dataset in the joint parameter space. A

comparison of the two integrated tomographies (Figure 64) shows similar results, with more details in the image provided by the 2x2 m spaced dataset. In order to build the conceptual model of the subsurface was taken into consideration the more resoluted image.

In agreement with the geological setting of the area and stratigraphic log from the borehole "Laterina 1", the integration method showed the optimal clustering of the dataset with 6 geophysical facies (Figure 66, Table 8). The facies represented by the cluster 2, with average values of resistivity (i.e. $\log_{10}$ resistivity) and p-wave velocity respectively of 1.78 ohm*m and 0.87 km/s, can be associated with silty and sandy (bna) unsaturated deposits. The clusters 3 and 6, respectively with average resistivity of 2.28 and 1.68 ohm*m and 1.4 and 1.76 km/s describe the unconfined and saturated aquifer, consistent with the groundwater depth level of 5.9 m observed in the survey area (Figure 45). The cluster 3, with a higher resistivity value, is mainly composed of pebbles with sand and the cluster 6, with a lower resistivity and slightly higher velocity values, composed of silty sand. The cluster 1 has the lowest resistivity and an average velocity of 1.97 km/s, associable with clayey Villafranchian deposits. The clusters 4 and 5 have average velocity values higher than 2.4 km/s, which can be associated with sandstones and marls of the Monte Cervarola unit.

The Silhouette index allowed the determination of a value for each observation of the dataset: this fact permitted evaluating the degree of membership in a cluster, i.e. how the observation differs from the centroid values resistivity and velocity of the cluster. As seen in Figure 68, observations with a silhouette value lower than 0.3, highlighted in black, are generally placed at the borders of the clusters in the joint parameter space. An improvement of the classical hard clustering obtained by the k-means, was obtained through the use of Silhouette index. In Figure 67 the Silhouette index highlighted main discontinuities (in red) between different lithologies and heterogeneities intra-cluster (low saturated colors zones).

A direct comparison of the stratigraphic log projected over the integrated tomography can show a high coherency of the results (Figure 100). The first layer individuated reaches the depth of 10 m instead of 9 m from the stratigraphic log, the second layer with resistivity and velocity values characteristics of a saturated sand with pebbles is present until the depth of 19 m. The stratigraphy showed the presence of a level with pebbles for a thickness of 6 m and then other 6 m of clay, which is missing in the integrated tomography. The lack of the clayey level, also in the electrical resistivity tomography, can be probably due to the local absence of the same layer along the survey line, located about 40 m on the left of the borehole (Figure 54). The shallow depth and the thickness not negligible, with the high resolution of the electrical resistivity survey make improbable a "loss" of the clayey layer during the acquisition.

Moreover, the integrated tomography allowed to constrain the interpretation of the main lateral variation at the distance of about 200 m in the elevation range between 170 and 140 m, where the resistivity (Figure 56) imaged a low lateral variation, but with a high velocity decreasing from the left to the right of the seismic refraction tomography. The low resistivity variation, resulted ambiguous in the univariate interpretation, coupled with the P-waves seismic velocity allowed to highlight a sub vertical interface, interpreted as the contact between the Mount Cervarola unit (i.e. sandstone/marl flysch) and the fluvial-lacustrine Villafranchian deposits (i.e. clayey deposits). This contact is consistent with the direct fault seen during the geological fieldworks. The third layer individuated in the integrated tomography is consistent with the sandstone/marl level from the depth of 20 m to 39 m. Although the cluster 5 is at the bottom of the tomograms (i.e. low sensitivity and low resoluted zone), is consistent with the borehole loggings and can be associated with a very low fractured sandstone portion.

As seen above, the use of Silhouette index can be considered as an improvement with respect to the classical hard clustering version of the integration method. In particular, the Silhouette values are used as saturation value of the colors in the integrated tomography (Figure 67). The main difference with respect to the hard clustering imaging is the possibility to highlight local dissimilarity from the average values of the geophysical facies individuated. The orange layer (cluster 4) in Figure 101 showed a low saturated color zone in the upper part, with respect to the lower (high saturated color), indicating a local difference of resistivity and velocity values compared to the average values of the geophysical facies. This difference is consistent with the low fractured/ high fractured zone individuated from the stratigraphy and geophysical logs.
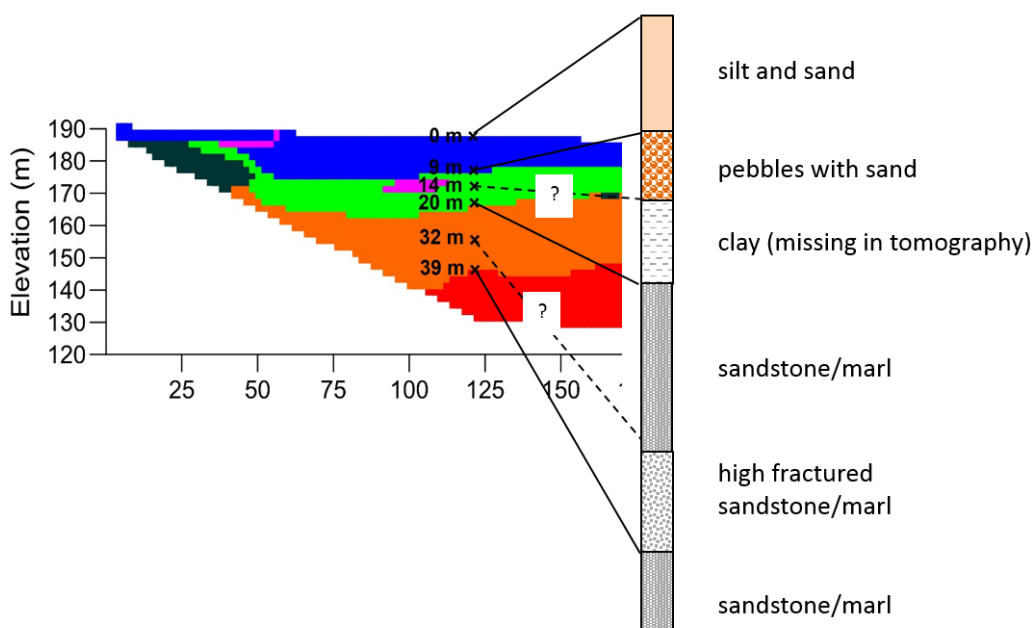
Figure 100 – Comparison of geophysical facies individuated through the hard clustering version, with the projected stratigraphic log.
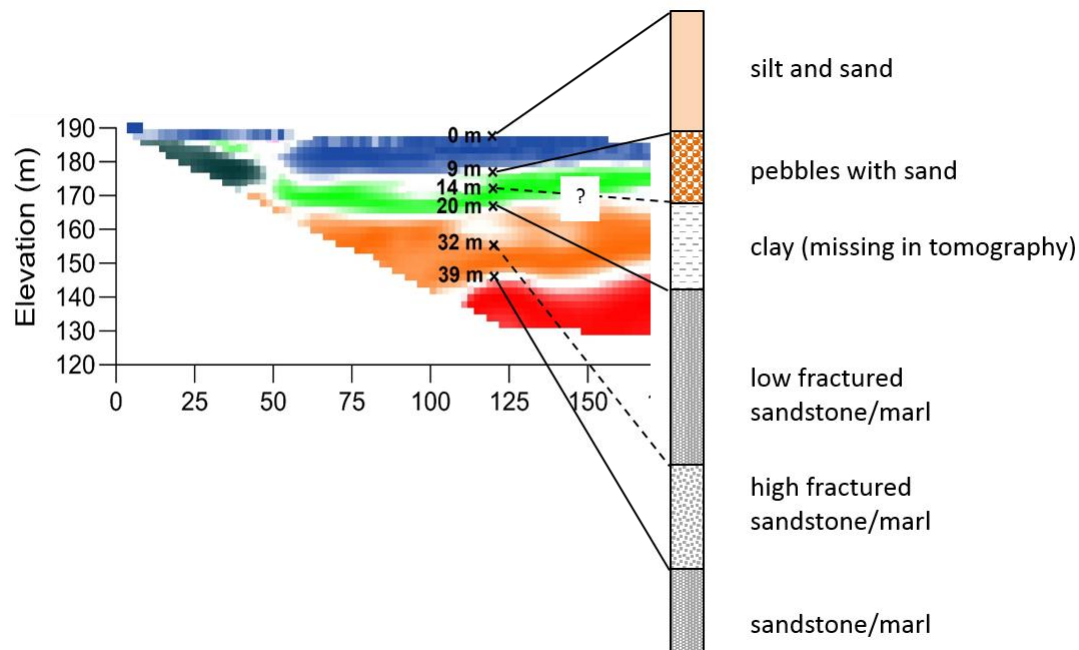
Figure 101 - Comparison of geophysical facies individuated through the Silhouette saturated version, with the projected stratigraphic log.

The tomographic integration method characterized the subsurface in two main potential aquifers and two aquitard:

- Geophysical facies individuated by clusters 2, 3 and 6 (Table 8, Figure 68) constitute the unconfined aquifer in recent deposits, with a different degree of saturation;
- Geophysical facies individuated by cluster 4 represents the confined aquifer in fractured rocks;
- Geophysical facies individuated by cluster 1 and 5 represent the main aquitards.

The resulting conceptual model is characterized of four different zones:

- unsaturated alluvial aquifer;
- saturated alluvial aquifer;
- saturated rock aquifer;
- aquiclude.
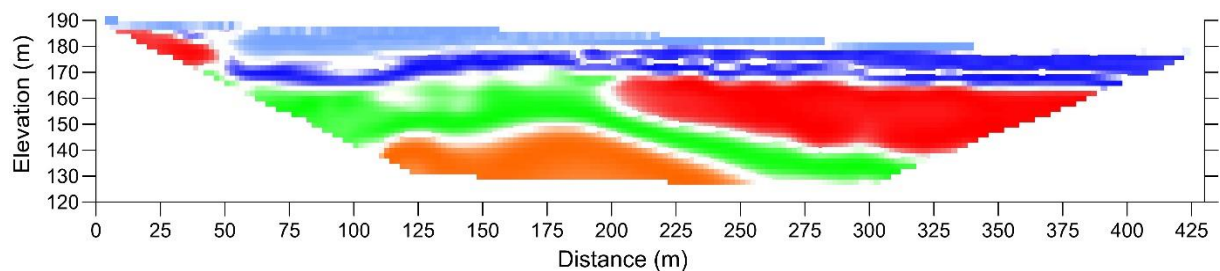
Figure 102 – Conceptual model from the integrated analysis of tomographic dataset: in light blue unsaturated alluvial aquifer; in blue saturated alluvial aquifer; in green saturated rock aquifer; in orange low fractured sandstone (aquitard); in red clayey deposits (aquitard). The white portions are representative of borders between different lithologies and intra-unit dissimilarities.

136

Self-Organized maps (or Kohonen maps) applied to the seismic multivariate dataset, composed by the attributes carried out from the high-resolution shallow seismic section, had the purpose to mainly distinguish unconsolidated sediments (i.e. alluvial recent deposits and Villafranchian deposits) from the arenaceous-marly substrate and its degree of fracturation (i.e. Monte Cervarola unit). The best integrated image was obtained for the 7x5 neural network which allowed to distinguish main geological characters of the subsurface.

The Self-Organized maps allowed to image at a large scale the geological setting of sediments and the substrate with its alteration degree. Anyway, mainly from about CDP 200, a complex seismic section characterized by a low impedance contrast led to a difficult discrimination between the Villafranchian clayey deposits and the top of Monte Cervarola unit. As seen in the tomographic integration (Figure 67), a main distinction can be made differentiating the geological setting before and after the distance of about 200 m, where an abrupt lateral variation of geophysical parameters allowed to interpret the Monte Cervarola unit in contact with the clayey Villafranchian deposits. This interpretation is in agreement with the geological evidences of a direct fault in the Southeast part of the Laterina basin, where outcrop cataclasites and fault breccias (Figure 45 and Figure 48).

In Figure 71 a, the zone marked with the letter "B" with medium to high values of similarity, low energy and instantaneous frequency and low to medium entropy, is interpreted as top of Mount Cervarola unit. The depth of the top is in agreement with the stratigraphic log provided by the borehole "Laterina 1" and its shape in agreement with the setting highlighted by the integrated tomography. The low energy values with medium to high similarity and low to medium entropy can be due to its degree of fracturation, resulting with a non-high impedance contrast. Moreover, the change of seismic facies is well imaged by the lowest frequency values. On the top of this zone a non-reflective facies is individuated with the letter "A", and represents the shallow unconfined aquifer on unconsolidated recent sediments. At the letter "B", in agreement with the stratigraphic log, can be associated the aquifer in fractured rocks, on the low-weathered portion of the top of Monte Cervarola unit. With the letters "C1" and "C2" were highlighted zones having medium to high energy, low entropy and high similarity values. The "C1" can be interpreted as a lower fractured portion of the Mount Cervarola unit because of the higher value of energy of the zone "B". This interpretation is consistent with the spatial disposition of sandstone/marly individuated by the integrated tomography (cluster 4). The letter "C2" includes portions with higher energy values than "C1" and, differently from "C1", matches with the Villafranchian clayey deposits interpreted with the cluster 1 in Figure 67 and Figure 68.

The sub-horizontal interface well imaged in the tomographic integrated image (Figure 67), between the clusters 1 and 4 at distances greater than 220 m and depths deeper than 40 m, interpreted respectively as Villafranchian clayey deposits and the top of

Monte Cervarola unit, is not recognized in the seismic integrated section (Figure 71). Probably, the low impedance contrast between the Villafranchian deposits and the sandstone/marly rocks led to a hard distinction of the two lithologies based only on the seismic attributes. Anyway, in Figure 71 the portion individuated with the letter "D" highlighted high entropy values, in correspondence with the above mentioned zone, consistent with a very high weathered/fractured state of the sandstone/marly deposits.

7.2 Complex hydrothermal aquifer characterization: Solfatara volcano

The case study of Solfatara had the aim to characterize the complex hydrothermal environment consisting in sets of sub-vertical paths for liquid and gas into the caldera, local variations of permeability and fluids saturation.

The lack of intra-caldera stratigraphic log informations led to a qualitative interpretation based on previous works in the Solfatara area.

The integration of tomographic dataset was performed on Profile 1 and Profile 2 as in Figure 74. As detailed in § 4.1, in both datasets, was conducted the preliminary test in order to evaluate effects on the integration method, due to the process of co-location. With the aim to preserve informations from the more resoluted seismic models, a co-located 3x3 m grid spacing and the larger 10x6 m were applied, due to the large spacing of the ERT model. In Profile 1 the clustering analysis showed 3 and 6 geophysical facies, respectively for 3x3m and 10x6m grid spacing. The different result for the more and lower resoluted datasets can be probably due to two main reasons. As seen in Figure 32 the resolution of ERT coarses with the depth: 1) using the minimum resolution of geoelectrical model led to lose informations from both univariate datasets; 2) the density of the bivariate dataset influences the k-means algorithm in finding optimal centroids. Moreover, due to the small size of co-located portion of seismic refraction and electrical resistivity tomograms, the lower resoluted grid analyzed resulted in a very low number of observations, compared with the more resoluted. Nevertheless, Figure 82 revealed similar pattern in the joint parameter space (i.e. resistivity – velocity space), where the higher number of clusters can be reconduced to the 3 clusters of the higher resoluted dataset. Clusters 5 and 6 constitute the cluster 2, 4 and 1 the cluster 3 and 2 the cluster 1. Cluster 3 of the 10x6m grid spaced dataset results the unique not directly ascribable as a part of clusters individuated in the 3x3m grid spaced dataset analysis. In terms of integrated images, the lower resoluted grid resulted in a more discretized way (Figure 83), with an intermediate layer provided by the observations within the cluster 3. As on the Profile 1, the same preliminary test on Profile 2 was performed, with grid spacing 3x3 m and 10x6 m. The integration method provided similar results for Profile 2, in terms of geophysical facies with substantially the same Silhouette index (Table 11 and Figure 84). Due to this, the more resoluted dataset as integrated tomography was taken into consideration. Moreover, considering the result obtained for the Profile 2 (i.e. 3 geophysical facies in the

similar environment of Profile 1), the loss of informations due to reduced amount of observations and the mainly nested partitioning of clusters between 3x3m and 10x6 m for Profile 1, led to interpret also for Profile 1 the results with 3 geophysical facies.

In both profiles (Figure 86 and Figure 89), the first layer (i.e. geophysical facies) individuated by the cluster 1, with low velocity and high resistivity, can be associated with shallow poorly consolidated tephra sediments and affected by CO2 degassing (Chiodini et al., 2001; Bruno et al., 2007; Bruno et al., 2017; Gresse et al., 2017). The cluster 2, characterized by average values of the log-resistivity 1.29 and 1.17 ohm*m, and P-wave velocity 2.26 and 2.37 km/s, respectively for Profile 1 and Profile 2. These homogeneous zones, comparing P-wave velocities, are consistent with high pressure gas-dominated porous media individuated in the work of Bruno et al., (2007) and De Landro et al., (2017) and with the work of Byrdina et al., (2014), Troiano et al., (2014) and Gresse et al., (2017), comparing electrical resistivity values. The clusters 3, with low resistivity and high velocity are consistent with liquid-dominated zones. More in specific, the sub vertical disposition of cluster 3 on Profile 2 at the distance of about 200 m (Figure 89) matches with the condensate flow individuated in the work of Gresse et al., (2017). Locally suspended zones individuated by cluster 3 can be associated with liquid saturated portions, which in the topographic surface are represented by the Fangaia mud-pool. A minimal lack of consistency in depth can be due to the dynamic setting of the caldera and changements occurred between the field surveys.

The homogeneous zones individuated by the tomographic integration method can be summarized as follows:

- Geophysical facies individuated by cluster 1 represent the unsaturated - poorly consolidated sediments;
- Geophysical facies individuated by cluster 2 represent the high pressure gas dominated porous media;
- Geophysical facies individuated by cluster 3 represent the liquid saturated media (i.e., mud pools and condensate paths).

The resulting conceptual model for the shallow hydrothermal system is characterized by three different zones:

- unsaturated sediments affected by CO_2 degassing;
- gas dominated porous media;
- liquid dominated porous media.
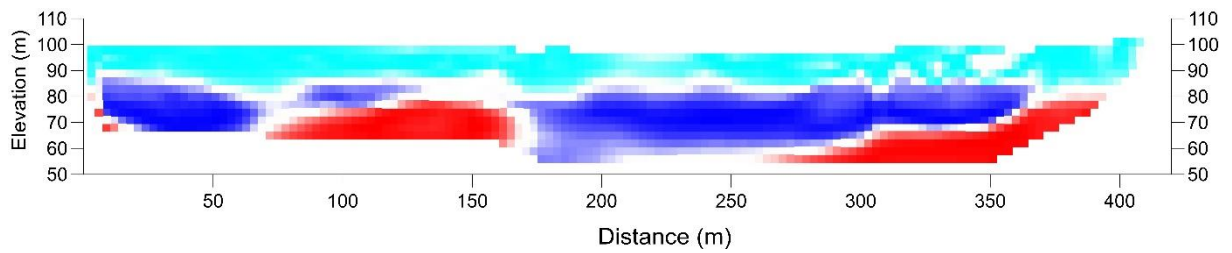
Figure 103 – Conceptual model from the integrated analysis of tomographic dataset: in light blue, unsaturated poorly consolidated sediments affected by CO$_2$ degassing; in blue, liquid dominated porous media; in red, gas dominated porous media. The white portions are representative of borders between different hydrogeological zones and intra-zone dissimilarities.
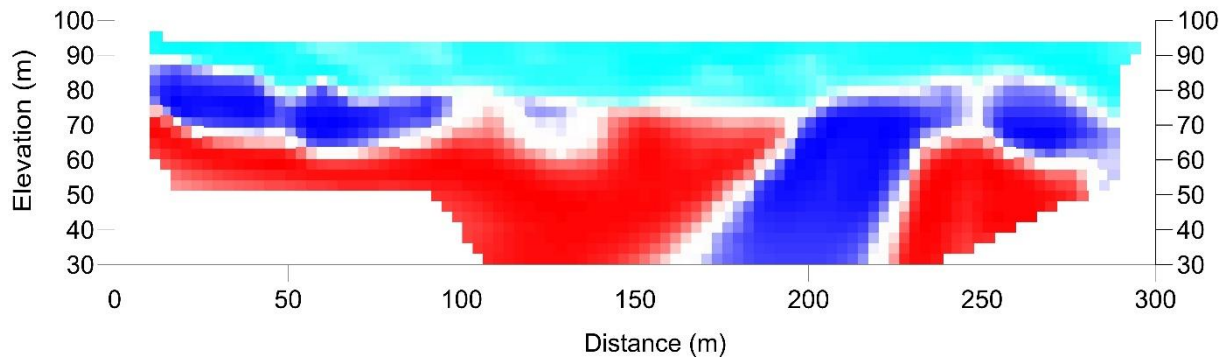
Figure 104 - Conceptual model from the integrated analysis of tomographic dataset: in light blue, unsaturated poorly consolidated sediments affected by CO$_2$ degassing; in blue, liquid dominated porous media; in red, gas dominated porous media. The white portions are representative of borders between different hydrogeological zones and intra-zone dissimilarities.

The integrated analysis of seismic attributes based upon Self-Organized Maps, had the aim to produce a unique image with zones having similar characteristics in terms of seismic response. In particular attributes as dip, energy, similarity and texture were used to reveal principal faults and changes in acoustic impedance, where the latter can be interpreted as a vapor-saturated porous media (Bruno et al., 2017). The Self-Organizing Maps allowed to classify observation of the multivariate datasets, revealing particular and different attribute signature for the different zones of the seismic image.

As seen in many works based on Self-Organizing maps (e.g.; de Matos et al., 2006; Roden et al., 2015), one of the most limit is the absence of a quantitative criteria to assess the validity of the results. Anyway the quantitative use of heatmaps helps in understanding the spatial similarities and correlations individuated by the algorithm, respectively among the observations and attributes.

Before to perform SOM analysis, a principal component analysis was carried out to define the net dimensions (Kohonen, 2014) and evaluate the contribution of variables (i.e. attributes) in the reduced 2D space. The first multivariate dataset composed by dip, energy, entropy and similarity, was integrated by using several grids, where the neural network 7x5 and 12x8 showed the most meaningful integrated images.

The most representative patterns in Figure 93 (i.e. som analysis with grid 7x5) were individuated with an asymmetrical shape up to a depth of 400 m. The neurons 5, 6, 7, 13

140

and 14 (Figure 83), representative of zones with energy and similarity medium to maximum, sub horizontal dip and medium to low entropy (Figure 82) can be produced by high seismic impedance contrast. Probably, the alternance of low-permeability and high-permeability horizons create stratigraphic traps for the uprising hydrothermal fluids (Bruno et al., 2017). These areas are well imaged from the depth of about 50 m up to 350 m, repeated at different depths, where the largest is at the top of the section (Figure 93). This fact can be explained as the presence of a continuous low permeability layer at the top of the caldera. Starting from depths greater than 150 m this sub-horizontal high energy areas are truncated by sub-vertical path, having very low energy, high entropy and dip, and low similarity. This kind of signature can be interpreted as fractured/faulted zones. In a hydrogeological point of view, these are the most important paths for uprising fluids

Moreover, the Umatrix showed a high distance among neurons, consistent with the very different average values mainly of energy and entropy (Figure 82). On the integrated section are imaged as sub-horizontal zones from the depth of about 50 m up to 300 m. The neurons 21, 28 and 35 show high dip angle, low values of similarity and energy, and medium to high entropy. Are well imaged as thin sub-vertical path from the depth of 100 m up to 350 m, mainly into the central part of the section, cutting the sub-horizontal portions highlighted by the neurons 5, 6, 7, 13 and 14. The neuron 20 also has similar values, with lower dip and a slightly higher similarity. Neurons 2 and 9 contain observation with medium to low energy, high dip (toward left) and entropy and medium similarity and are mainly present on the left side of the section from the depth of about 50 m up to 300 m, creating an asymmetrical zone on the section. In general observations colored with light green on the integrated section (Figure 83) show very low energy, variable dip, medium similarity and entropy and are spatially distributed at the border of above described observations. These are representative of non-reflective zones and at a depth greater than about 400 m can be interpreted as pre-eruptive basement showing features consistent with the emplacement of shallow intrusive bodies (Bruno et al., 2017).

The second analysis carried out with a 12x9 neural grid, in order to evaluate differences on the 2D distribution of observations across the dense net, showed well 2D variations of attributes among neurons (Figure 94). The higher discretization introduced by the 12x9 net, allowed to distinguish observations in a more gradational way, the most for the observations classified into the neurons 105,106,107,94,95,96,83 and 84. In general, the 12x9 neural network analysis imaged features consistent with the 7x5 grid analysis. Compared with the latter, the 12x9 neural net analysis, imaged in a better way intra crater sub-horizontal pattern, as discussed above, interpreted as traps for hydrothermal fluids. Furthermore, several sets of high angle fractures/faults can be associated with neurons 12, 24, 36, 48 and 13, 25, 37, respectively for sub vertical pattern dipping toward left and right. (Figure 84). Due to the high angle and their intra-crateric position, as seen in Bruno et al., (2017) and Isaia et al., (2017), these structures can be

associated to collapses and explosions of the center of the crater. The asymmetrical dipping structure at the left border of the crater resulted better imaged by the SOM analysis 7x5 (Figure 93), probably due to the higher discretization introduced by the 12x8 grid (Figure 95). In general the increasing of the net dimensions led to a lower intra neurons dispersion of variables, with the drawback to over-discretize similar zones.

The dataset composed by dip angle, similarity, energy, instantaneous frequency and several complementary and redundant texture attributes, revealed a low contribution of instantaneous frequency. Due to this the latter was removed from the dataset, in order to use only meaningful informations and reduce possible sources of noise. The 2D distribution of attributes among the 12x8 neural net revealed complex patterns, still consistent with the previous dataset analyzed. The textural contrast attribute showed its main contribution in highlighting the lateral variations of seismic texture, as at the contact on the left border and the volcanoclastic filled crater, as the intra crater discontinuities interpreted as faults. The use of redundant attributes, as energy- texture, homogeneity and entropy, although they showed a high contribution in principal component analysis, not improved the integrated image with respect to the previous dataset analyzed. Moreover, at a visual inspection, the integrated image resulted more chaotic than the previous analyses.

The correlations revealed by the Self-Organizing maps among the most meaningful attributes on the Solfatara datasets can be summarized in four main categories:

1. high energy, high similarity, low entropy and sub-horizontal dip consistent with intra crater bright spot;
2. high entropy, lowest energy, medium to low similarity and high absolute values of dip;
3. high dip angle, medium energy and similarity and low to medium entropy and high contrast;
4. low energy, medium to low similarity, medium entropy and low contrast.

The four categories above enunciated and obtained through the Self-Organizing maps resulted as seismic signatures, powerful to distinguish main features connected with the hydrothermal system, as intra-crater faults/fractures and deep reservoirs. This allowed to characterize the Solfatara volcano at depths greater than 400 m.

7.3 Improvement from Silhouette index

As seen in the work of Di Giuseppe et al., (2014), the tomographic integration method based on the k-means algorithm and the bivariate dataset composed by electrical resistivity and P-waves seismic velocity proposed in this thesis, highlighted several homogeneous zones by clustering the dataset in the joint parameter space. Conversely from Di Giuseppe et al., (2014), in order to individuate the optimal clustering was used the Silhouette index instead of RS index. The Silhouette index allowed to calculate each observation how well is assigned within the cluster, compared to the samples into the other

clusters. In general at the border of clusters, observation have values whose differ from values of the centroids individuated with the k-means, a low degree of membership within the cluster, and can reveal local dissimilarities of the geophysical facies. Thus, the use of Silhouette index as color saturation value, improved the graphical imaging of integrated tomography. In this manner, the integrated tomograms saturated with Silhouette index resulted more meaningful than the hard clustering version (Figure 67, Figure 86 and Figure 89). In particular, in the case of Laterina, the hard clustering version of integrated tomography highlighted an initial lack of consistency with the stratigraphic log (Figure 100). Conversely, the image saturated with the membership degree provided by Silhouette index, revealed a local dissimilarity of the orange zone, more consistent with the different degree of fracturation (Figure 101).

8   Conclusions

The aim of this thesis was the geophysical characterization of aquifers, through dual integrated approaches based on two of the most common geophysical surveys, electrical resistivity tomography and seismic refraction/reflection methods. More in specific, were proposed two different approaches, one using tomographic data and one seismic attributes. The first allowed to characterize aquifers and aquitards thanks to the different sensitivity of electrical resistivity and P-waves velocity to parameters as mineralogy, lithology, permeability, fluids saturation, texture and grain size. The second had the aim to reveal principal structural and stratigraphic variations, in order to study at a larger scale the subsurface geological setting. Final result was a unique conceptual model of subsoil, differentiated in geophysical homogeneous zones and strictly connected with lithology variations and water content.

The main advantages brought from this thesis in groundwater research, can be pointed out as follows: i) uncertainty reduction of geophysical model (compared with the univariate case) helpful in the well location phase, ii) building of a reliable conceptual model at the base of hydrogeological modeling and iii) high-resolution large-scale geological structural characterization.

As seen in both cases study, the tomographic integration distinguished several zones characterized by average values of electrical resistivity and P-waves velocity, connected with different hydrogeological settings. The Silhouette index permitted to individuate the optimal number of homogeneous zones and local dissimilarities of observations from the values of the centroid within the clusters. The use of uncertainty was also adopted to take into consideration well known limits of k-means in partitioning complex dataset, where not well natural separated clusters within the dataset, can lead to a not clear assignment at the borders. In this manner, an eventual misclassification of observations has a low impact in the integrated image.

The Self-Organized maps revealed as a useful method to recognize and classify portions of subsurface with similar seismic response. In this work, the 3D color map, based on the neural network and the Umatrix, provided a crucial instrument to image similar (or dissimilar) zones. Correlations individuated among attributes for similar portions of subsurface were used as signature to interpret the different features imaged in the integrated sections. One of the limits of this approach appeared when a high number of neurons is present, with subtle differences among neighbor neurons in terms of colors (i.e. similar classified observation). A lack of a sharp changement of colors made hard in some cases reveal not well marked differences among observations. The use of more complementary and redundant texture attributes, as in the case of Solfatara, increased the complexity of the integrated image, without any meaningful other contribution. Thus, the attributes must be limited to a minimal significant number individuated by their

144

theoretical contribution in imaging the desired target and trials to evaluate the real contribution in discriminating different features of the subsurface.

9 Further developments

In both cases (i.e. for tomographic data and seismic attributes), the applied integration methods are post processing and depend on the spatial distribution of initial geophysical parameters. In order to evaluate the reliability of the method, in the tomographic integration further analysis will be applied varying the inversion process of univariate models. As known one of the limits of tomographic inversions is their spatial smooth variation of the geophysical parameter. In particular will be tested inversion methods less smoothed with respect to the applied in this thesis, in order to highlight sharpen variations of electrical resistivity and P-waves velocity. Theoretically, by using tomographic models with a sharp increase (or decrease) of the variable, one can obtain datasets with clusters more separated with respect to the datasets obtained with a smoothed variation of the geophysical parameter.

In the tomographic integration a significant improvement can be the use of more geophysical parameters, in specific the coupled use of S-waves with P-waves velocity can be powerful for groundwater purposes, because of the higher sensitivity to the water content of the ratio Vp/Vs or its manipulations. One of the limits about the use of S-waves is due to the source, generally less energetic and reliable of P-waves sources, with the consequence of a reduced range of offset. This fact is the major limit connected with the tomographic inversion of first arrivals. Anyway, the increased interest in the use of S-waves for shallow surveys led to the production of vibrational sources capable to carry out high quality datasets and with a range of offset comparable with the P-waves sources for shallow surveys.

In the Self-Organizing maps analysis, one of the major drawback with the increasing of the number of the neurons was the difficult in interpreting the results. In futures analyses will be tested an automated interpretation of results based on the clustering of codebook vectors. In this manner similar neurons can be grouped within similar clusters in a reduced number of final similar zones. Moreover, the lack of a consolidated quantitative method to assess the validity of results appears as a limit for choosing the optimal result among the analyzed neural network. Thus, will be evaluated methods based on the distances among observations intra and among neurons.

One of the more promising use of neural network will be based on their capability to recognize similar features on a new dataset, based on previous learning of the trained net on a dataset, as the human brain does. In presence of known stratigraphic informations, and based on a trained net, the neural network algorithms can classify similar observations of the new dataset in a supervised way. Thanks to this capability, sparse stratigraphic informations can be spreaded on the new dataset based on the recognition capabilities of the previously trained neural network.

10 References

- Abdi, Hervé, and Lynne J. Williams. 2010. "Principal Component Analysis." John Wiley and Sons, Inc. WIREs Comp Stat 2: 433–59. http://staff.ustc.edu.cn/~zwp/teach/MVA/abdi-awPCA2010.pdf.
- Abdulkadir, Y. A., & Eritro, T. H. (2017). 2D resistivity imaging and magnetic survey for characterization of thermal springs: A case study of Gergedi thermal springs in the northwest of Wonji, Main Ethiopian Rift, Ethiopia. Journal of African Earth Sciences, 133, 95-103.
- Ackermann, H. D., Pankratz, L. W., & Dansereau, D. (1986). Resolution of ambiguities of seismic refraction traveltime curves. Geophysics, 51(2), 223-235.
- Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. Transactions of the AIME, 146(01), 54-62.
- Azaria, A., Zelt, C., & Levander, A. (2003, April). High-resolution seismic mapping at a groundwater contamination site: 3-D traveltime tomography of refraction data. In EGS-AGU-EUG Joint Assembly.
- Banton, O., Cimon, M. A., & Seguin, M. K. (1997). Mapping field-scale physical properties of soil with electrical resistivity. Soil Science Society of America Journal, 61(4), 1010-1017.
- Bauer, K., Schulze, A., Ryberg, T., Sobolev, S. V., & Weber, M. H. (2003). Classification of lithology from seismic tomography: A case study from the Messum igneous complex, Namibia. Journal of Geophysical Research: Solid Earth, 108(B3).
- Bauer, K., Munoz, G., Moeck, I., 2012. Pattern recognition and lithological interpretation of collocated seismic and magnetotelluric models using self-organizing maps. Geophys. J. Int. 189, 984–998. http://dx.doi.org/10.1111/j.1365-246X.2012.05402.x.
- Bedrosian, P.A.,Maercklin, N.,Weckmann, U., Bartov, Y., Ryberg, T., Ritter, O., 2007. Lithologyderived structure classification from the joint interpretation of magnetotelluric and seismic models. Geophys. J. Int. 170, 737–748. http://dx.doi.org/10.1111/j.1365-246X.2007.03440.x.
- Bernardinetti, S., Pieruccioni, D., Mugnaioli, E., Talarico, F. M., Trotta, M., Harroud, A., & Tufarolo, E. (2018). A pilot study to test the reliability of the ERT method in the identification of mixed sulphides bearing dykes: The example of Sidi Flah mine (Anti-Atlas, Morocco). Ore Geology Reviews, 101, 819-838.
- Berrino G., Corrado G., Luongo G., Toro B. (1984). Ground deformation and gravity changes accompanying the 1982 Pozzuoli uplift. Bulletin of Volcanology, 47(2), 187-200. https://doi.org/10.1007/BF01961548
- Bianco F., Del Pezzo E., Saccorotti G., Ventura G. (2004). The role of hydrothermal fluids in triggering the July-August 2000 seismic swarm at Campi Flegrei, Italy: evidence from seismological and mesostructural data. Journal of Volcanology and Geothermal Research, 133, 229–246. https://doi.org/10.1016/S0377-0273(03)00400-1
- Binley, A., Hubbard, S. S., Huisman, J. A., Revil, A., Robinson, D. A., Singha, K., & Slater, L. D. (2015). The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. Water resources research, 51(6), 3837-3866.
- Biot, M. A. (1956). Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range. The Journal of the acoustical Society of america, 28(2), 179-191.
- Biot, M. A. (1962). Mechanics of deformation and acoustic propagation in porous media. Journal of applied physics, 33(4), 1482-1498.

- Blainey, J. B., Ferré, T. P., & Cordova, J. T. (2007). Assessing the likely value of gravity and drawdown measurements to constrain estimates of hydraulic conductivity and specific yield during unconfined aquifer testing. Water resources research, 43(12).
- Blake, S., Henry, T., Muller, M. R., Jones, A. G., Moore, J. P., Murray, J., ... & Rath, V. (2016). Understanding hydrothermal circulation patterns at a low-enthalpy thermal spring using audio-magnetotelluric data: A case study from Ireland. Journal of Applied Geophysics, 132, 1-16.
- Block, G. I., & Harris, J. G. (2006). Conductivity dependence of seismoelectric wave phenomena in fluid-saturated sediments. Journal of Geophysical Research: Solid Earth, 111(B1).
- Boadu, F. K., Gyamfi, J., & Owusu, E. (2005). Determining subsurface fracture characteristics from azimuthal resistivity surveys: A case study at Nsawam, Ghana. Geophysics, 70(5), B35-B42.
- Boccaletti, M., Calamita, F., Deiana, G., Gelati, R., Massari, F., Moratti, G., Ricci Lucchi, F. (1990) – Migrating foredeep thrust belt system in the Northern Apennines and Southern Alps. Palaeogeogr., Palaeoclimatol., Palaeoecol., 77: 3-14
- Boccaletti, M., Bonini, M., Moratti, G., Sani, F. (1996) Tectono-sedimantary Neogene evolution of the post-nappe basins in the Northern Appennines. Notes et Mém Serv. Géol. Maroc., n. 387: 117-193.
- Bosch, M. (1999). Lithologic tomography: From plural geophysical data to lithology estimation. Journal of Geophysical Research: Solid Earth, 104(B1), 749-766.
- Bosch, M., Zamora, M., & Utama, W. (2002). Lithology discrimination from physical rock properties. Geophysics, 67(2), 573-581.
- Bossio A., Cerri R., Costantini A., Gandin A., Lazzarotto A., Magi M., Mazzanti R., Mazzei R., Sagri M., Salvatorini G.(1992) – I bacini distensivi neogenici e quaternari della Toscana. 76 Adun. Estiva Soc. Geol. It., Guida alle escursioni, in stampa, Firenze.
- G. Brock, V. Pihur, S. Datta, and S. Datta. clValid: An R package for cluster validation. Journal of Statistical Software, 25(4), March 2008. URL http://www.jstatsoft.org/v25/i04.
- Brock, G., Pihur, V., Datta, S., & Datta, S. (2011). clValid, an R package for cluster validation. Journal of Statistical Software
- Bruno P.P., Ricciardi G.P., Petrillo Z., Di Fiore V., Troiano A., Chiodini G. (2007). Geophysical and hydrogeological experiments from a shallow hydrothermal system at Solfatara Volcano, Campi Flegrei, Italy: response to caldera unrest. Journal of Geophysical Research: Solid Earth, 112, 1–17. https://doi.org/10.1029/2006JB004383
- Bruno, P. P., Castiello, A., & Improta, L. (2010). Ultrashallow seismic imaging of the causative fault of the 1980, M6. 9, southern Italy earthquake by pre-stack depth migration of dense wide-aperture data. Geophysical Research Letters, 37(19).
- Bruno, P. P., Castiello, A., Villani, F., & Improta, L. (2013). High-Resolution Densely Spaced Wide-Aperture Seismic Profiling as a Tool to Aid Seismic Hazard Assessment of Fault-Bounded intramontane Basins: Application to Vallo di Diano, Southern Italy. Bulletin of the Seismological Society of America, 103(3), 1969-1980.
- Bruno, P. P. G., Maraio, S., & Festa, G. (2017). The shallow structure of Solfatara Volcano, Italy, revealed by dense, wide-aperture seismic profiling. Scientific reports, 7(1), 17386.
- Burger H. R. (1992). Exploration Geophysics of the Shallow Subsurface. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 489 pp.
- Byrdina S., Vandemeulebrouck J., Cardellini C., Legaz A., Camerlynck C., Chiodini G., Lebourg T., Gresse M., Bascou P., Motos G., Carrier A., Caliro S. (2014). Relations between electrical resistivity, carbon dioxide flux, and self-potential in the shallow hydrothermal system of Solfatara (Phlegrean Fields, Italy). Journal of Volcanology and

Geothermal Research, 283, 172–182. https://doi.org/10.1016/j.jvolgeores.2014.07.010

- Caliro S., Chiodini G.,Moretti R., Avino R., Granieri D., Russo M., Fiebig J. (2007). The origin of the fumaroles of La Solfatara (Campi Flegrei, South Italy). Geochimica et Cosmochimica Acta, 71, 3040–3055. https://doi.org/10.1016/j.gca.2007.04.007
- Campbell, R. B., Bower, C. A., & Richards, L. A. (1949). Change of Electrical Conductivity With Temperature and the Relation of Osmotic Pressure to Electrical Conductivity and Ion Concentration for Soil Extracts 1. Soil Science Society of America Journal, 13(C), 66-69.
- Cardellini C, Chio G, Frondini F., Avino R., Bagnato E., Caliro S., Lelli M., Rosiello A. (2017). Monitoring diffuse volcanic degassing during volcanic unrests: the case of Campi Flegrei (Italy). Scientific Reports, 7, 1–15. https://doi.org/10.1038/s41598-017-06941-2
- Castagna, J. P., Batzle, M. L., & Eastwood, R. L. (1985). Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. Geophysics, 50(4), 571-581.
- Catchings, R. D., Goldman, M. R., Lee, W. H. K., Rymer, M. J., & Ponti, D. J. (1998). Faulting apparently related to the 1994 Northridge, California, earthquake and possible co-seismic origin of surface cracks in Potrero Canyon, Los Angeles County, California. Bulletin of the Seismological Society of America, 88(6), 1379-1391.
- Chambers, J., Ogilvy, R., Kuras, O., Cripps, J., & Meldrum, P. (2002). 3D electrical imaging of known targets at a controlled environmental test site. Environmental Geology, 41(6), 690-704.
- Chiodini, G., Frondini F., Cardellini C., Granieri D., Marini L., Ventura G. (2001). CO2 degassing and energy release at Solfatara volcano, Phlegraean Fields, Italy. Journal of Geophysical Research: Solid Earth, 106, 16213–16221. https://doi.org/10.1029/2001JB000246
- Chiodini G., Granieri D., Avino R., Caliro S., Costa A., Werner C. (2005). Carbon dioxide diffuse degassing and estimation of heat release from volcanic and hydrothermal systems. Journal of Geophysical Research: Solid Earth, 110, B08204. https://doi.org/10.1029/2004JB003542
- Chiodini G. (2009). CO2/CH4 ratio in fumaroles a powerful tool to detect magma degassing episodes at quiescent volcanoes. Geophysical Research Letters, 36, L02302, https://doi.org/10.1029/2008Gl036347
- Chiodini G, Avino R, Caliro S, Minopoli C (2011). Temperature and pressure gas geoindicators at the Solfatara fumaroles (Campi Flegrei). Annales of Geophysics, 54, 151–160. https://doi.org/10.4401/ag-5002
- Chiodini G., Vandemeulebrouck J., Caliro S., D'Auria L., De Martino P., Mangiacapra A., Petrillo Z. (2015). Evidence of thermal-driven processes triggering the 2005–2014 unrest at Campi Flegrei caldera. Earth and Planetary Science Letters, 414,58–67. https://doi.org/10.1016/j.epsl.2015.01.012
- Christiansen, L., Binning, P. J., Rosbjerg, D., Andersen, O. B., & Bauer-Gottwein, P. (2011). Using time-lapse gravity for groundwater model calibration: An application to alluvial aquifer storage. Water Resources Research, 47(6).
- Chopra, S. and Alexeev, V., 2005. Application of texture attribute analysis to 3D seismic data. CSEG Recorder, Sep. 2005 pp 29-32.
- Chopra, S., & Marfurt, K. J. (2005). 75th Anniversary Seismic attributes—A historical perspective. GEOPHYSICS, 70(5).
- Coléou, T., Poupon, M., & Azbel, K. (2003). Unsupervised seismic facies classification: A review and comparison of techniques and implementation. The Leading Edge, 22(10), 942-953.
- Cox, T. F., & Cox, M. A. (2000). Multidimensional scaling. CRC press.

- Cusano P., Petrosino P., Saccorotti G. (2008). Hydrothermal origin for sustained Long-Period (LP) activity at Phlegraean Fields volcanic complex, Italy. Journal of Volcanology and Geothermal Research, 177, 1035–1044. https://doi.org/10.1016/j.jvolgeores.2008.07.019
- Dahlin, T., & Zhou, B. (2004). A numerical comparison of 2D resistivity imaging with 10 electrode arrays. Geophysical prospecting, 52(5), 379-398.
- Dahlin, T., & Zhou, B. (2006). Multiple-gradient array measurements for multichannel 2D resistivity imaging. Near Surface Geophysics, 4(2), 113-123.
- Daily, W., Ramirez, A., LaBrecque, D., & Nitao, J. (1992). Electrical resistivity tomography of vadose water movement. Water Resources Research, 28(5), 1429-1442.
- De Franco, R., Biella, G., Tosi, L., Teatini, P., Lozej, A., Chiozzotto, B. & Bassan, V. (2009). Monitoring the saltwater intrusion by time lapse electrical resistivity tomography: The Chioggia test site (Venice Lagoon, Italy). Journal of Applied Geophysics, 69(3-4), 117-130.
- de Groot-Hedlin, C. and Constable, S. (1990). Occam's inversion to generate smooth, two dimensional models from magnetotelluric data. Geophysics, 55, 1613-1624
- De Landro G., Serlenga V., Russo G., Amoroso O., Festa G., Bruno P.P., Gresse M., Vandemeulebrouck J., Zollo A. (2017). 3D ultra-high resolution seismic imaging of shallow Solfatara crater in Campi Flegrei (Italy): New insights on deep hydrothermal fluid circulation processes. Scientific Reports, 7, 1-10. https://doi.org/10.1038/s41598-017-03604-0
- de Matos, M. C., Osorio, P. L., & Johann, P. R. (2007). Unsupervised seismic facies analysis using wavelet transform and self-organizing maps. Geophysics, 72(1), P9-P21.
- Matos, M. C. D., Marfurt, K. J., & Johann, P. R. (2010). Seismic interpretation of self-organizing maps using 2D color displays. Revista Brasileira de Geofísica, 28(4), 631-642.
- De Natale G., Pingue F., Allard P., Zollo A. (1991). Geophysical and geochemical modelling of the 1982–1984 unrest phenomena at Phlegraean Fields caldera (southern Italy). Journal of Volcanology and Geothermal Research, 48, 199–222. https://doi.org/10.1016/0377-0273(91)90043-Y
- De Vita S., Orsi G., Civetta L. et al (1999) The Agnano-Monte Spina eruption (4100 years BP) in the restless Campi Flegrei caldera (Italy). Journal of Volcanology and Geothermal Research, 91, 269–301. https://doi.org/10.1016/S0377-0273(99)00039-6
- De Vivo B., Orsi G., Civetta L., Carandente A., D'Antonio M., Deino A., Cesare T., Di Vito M.A., Fisher R.V., Isaia R., Marotta E., Necco A., Ort M., Pappalardo L., Piochi M., Southon J. (2001). New constraints on the pyroclastic eruptive history of the Campanian volcanic Plain (Italy). Mineralogy and Petrology, 73, 47–65. https://doi.org/10.1016/S0377-0273(99)00039-6
- Dell'Aversana, P., Bernasconi, G., Chiappa, F. (2016). A global integration platform for optimizing cooperative modeling and simultaneous joint inversion of multi-domain geophysical data. AIMS Geosciences, 2(1), 1-31
- Di Giuseppe, M. G., Troiano, A., Troise, C., & De Natale, G. (2014). k-Means clustering as tool for multivariate geophysical data analysis. An application to shallow fault zone imaging. Journal of Applied Geophysics, 101, 108-115.
- Di Giuseppe MG, Troiano A, Fedele A., Caputo T., Patella D., Troise C., De Natale G. (2015). Electrical resistivity tomography imaging of the near-surface structure of the Solfatara crater, Campi Flegrei (Naples, Italy). Bullettin of Volcanology, 77, 27. https://doi.org/10.1007/s00445-015-0910-6

- Di Giuseppe, M. G., Troiano, A., Patella, D., Piochi, M., & Carlino, S. (2017). A geophysical k-means cluster analysis of the Solfatara-Pisciarelli volcano-geothermal system, Campi Flegrei (Naples, Italy). Journal of Applied Geophysics.
- Di Vito M., Lirer L., Mastrolorenzo G., Rolandi V. (1987). The 1538 Monte Nuovo eruption (Campi Flegrei, Italy). Bullettin of Volcanology, 49, 608–615. https://doi.org/10.1007/BF01079966
- Dietrich, P., Fechner, T., Whittaker, J., & Teutsch, G. (1998). An integrated hydrogeophysical approach to subsurface characterization. IAHS Publication(International Association of Hydrological Sciences), (250), 513-519.
- dos Reis Jr, J. A., de Castro, D. L., de Jesus, T. E. S., & Lima Filho, F. P. (2014). Characterization of collapsed paleocave systems using GPR attributes. Journal of Applied Geophysics, 103, 43-56.
- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern classification. 2nd. Edition. New York, 55.
- Dumay, J., & Fournier, F. (1988). Multivariate statistical analyses applied to seismic facies recognition. Geophysics, 53(9), 1151-1159.
- Dupuis, J. C., and K. E.Butler (2006), Vertical seismoelectric profiling in a boreholepenetrating glaciofluvial sediments, Geophys. Res. Lett., 33,L16301, doi:10.1029/2006GL026385
- Edwards, L. S. (1977). A modified pseudosection for resistivity and IP. Geophysics, 42(5), 1020-1036.
- Fayyad, M.U., Piatesky-Shapiro, G., Smuth P., Uthurusamy, R. (1996). Advances in Knowledge Discovery and Data Mining. AAAI Press.
- Fechner, T., & Dietrich, P. (1997, August). Lithologic Inversion of Tomographic Data. In 3rd EEGS Meeting.
- Forte, E., Pipan, M., Casabianca, D., Di Cuia, R., & Riva, A. (2012). Imaging and characterization of a carbonate hydrocarbon reservoir analogue using GPR attributes. Journal of Applied Geophysics, 81, 76-87.
- Foti, S., Lai, C. G., & Lancellotta, R. (2002). Porosity of fluid-saturated porous media from measured seismic wave velocities. Géotechnique, 52(5), 359-373.
- Francese, R., Giudici, M., Schmitt, D. R., & Zaja, A. (2005). Mapping the geometry of an aquifer system with a high-resolution reflection seismic profile. Geophysical Prospecting, 53(6), 817-828.
- Fukue, M., Minato, T., Horibe, H., & Taya, N. (1999). The micro-structures of clay given by resistivity measurements. Engineering geology, 54(1-2), 43-53.
- Gallardo, L., Meju, M.A., (2004). Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints. J. Geophys. Res. 109.
- Gao, D. (2003). Volume texture extraction for 3D seismic visualization and interpretation. Geophysics, 68(4), 1294-1302.
- Gehman, C. L., Harry, D. L., Sanford, W. E., Stednick, J. D., & Beckman, N. A. (2009). Estimating specific yield and storage change in an unconfined aquifer using temporal gravity surveys. Water Resources Research, 45(4).
- Gersho, A. (1982). On the structure of vector quantizers. IEEE Transactions on Information Theory, 28(2), 157-166
- Gilbert P. (1972). Iterative methods for the three-dimensional reconstruction of an object from projections. J. Theor Biol 1972;36:105±17.
- Glamoclija M., Garrel L., Berthon J. (2004). Biosignatures and bacterial diversity in hydrothermal deposits of Solfatara Crater, Italy. Geomicrobiology Journal, 21, 529–541. https://doi.org/10.1080/01490450490888235
- Gray, R. (1984). Vector quantization. IEEE Assp Magazine, 1(2), 4-29
- Gresse M., Vandemeulebrouck J., Byrdina S., Chiodini G., Revil A., Johnson T.C., Ricci T., Vilardo G., Mangiacapra A., Lebourg T., Grangeon J., Bascou P., Metral L. (2017).

Three-dimensional electrical resistivity tomography of the Solfatara Crater (Italy): Implication for the multiphase flow structure of the shallow hydrothermal system. Journal of Geophysical Research: Solid Earth, 122, 8749–8768. https://doi.org/10.1002/2017JB014389

- Griffiths, D. H., & Turnbull, J. (1985). A multi-electrode array for resistivity surveying. First break, 3(7), 16-20.
- Griffiths, D. H., Turnbull, J., & Olayinka, A. I. (1990). Two-dimensional resistivity mapping with a computer-controlled array. First break, 8(4), 121-129.
- Gruber, W. I. L. F. R. I. E. D., & Rieger, R. O. B. E. R. T. (2003). High resolution seismic reflection-constraints and pitfalls in groundwater exploration. MATERIALS AND GEOENVIRONMENT, 50, 133-136.
- Güler, C., Thyne, G. D., McCray, J. E., & Turner, K. A. (2002). Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology journal, 10(4), 455-474.
- Haberland, C., Rietbrock, A., Schurr, B., & Brasse, H. (2003). Coincident anomalies of seismic attenuation and electrical resistivity beneath the southern Bolivian Altiplano plateau. Geophysical Research Letters, 30(18).
- Haeni, F. P., 1986, Application of seismic refraction methods in groundwater modeling studies in New England: Geophysics, 51, no. 02, 236-249.
- Halkidi M., Batistakis Y., Vazirgiannis M. (2001) On Clustering Validation Techniques, Journal of Intelligent Information Systems, 17:2/3, 107–145.
- Hammah, R. E., & Curran, J. H. (1998). Fuzzy cluster algorithm for the automatic identification of joint sets. International Journal of Rock Mechanics and Mining Sciences, 35(7), 889-905.
- Han, D. H., Nur, A., & Morgan, D. (1986). Effects of porosity and clay content on wave velocities in sandstones. Geophysics, 51(11), 2093-2107.
- Hasselström, B. (1969). Water prospecting and rock-investigation by the seismic refraction method. Geoexploration, 7(2), 113-132.
- Improta, L., Zollo, A., Bruno, P. P., Herrero, A., & Villani, F. (2003). High-resolution seismic tomography across the 1980 (Ms 6.9) Southern Italy earthquake fault scarp. Geophysical research letters, 30(10).
- Irving, J., Knight, R., & Holliger, K. (2009). Estimation of the lateral correlation structure of subsurface water content from surface-based ground-penetrating radar reflection images. Water Resources Research, 45(12).
- Isaia R., Marianelli P., Sbrana A. (2009). Caldera unrest prior to intense volcanism in Campi Flegrei (Italy) at 4.0 ka B.P.: implications for caldera dynamics and future eruptive scenarios. Geophysical Research Letters, 36, 1–6. https://doi.org/10.1029/2009GL040513
- Isaia R. ,Vitale S., Di Giuseppe M.G., Iannuzzi E., D'Assisi Tramparulo F., Troiano A. (2015). Stratigraphy, structure, and volcano-tectonic evolution of Solfatara maar-diatreme (Campi Flegrei, Italy). Geological Society of American Bullettin, 127, 1–20. https://doi.org/10.1130/B31183.1
- Iyer, H. M., & Hirahara, K. (Eds.). (1993). Seismic tomography: Theory and practice. Springer Science & Business Media.
- Jackson, J. E. A user's guide to principal components. 1991. John Willey &Sons, New York.
- Jackson, P. D., Smith, D. T., & Stanford, P. N. (1978). Resistivity-porosity-particle shape relationships for marine sands. Geophysics, 43(6), 1250-1268.
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. ACM computing surveys (CSUR), 31(3), 264-323.

- Jin, S., & Feng, G. (2013). Large-scale variations of global groundwater from satellite gravimetry and hydrological models, 2002–2012. Global and Planetary Change, 106, 20-30.
- Kalinski, R. J., & Kelly, W. E. (1993). Estimating water content of soils from electrical resistivity. Geotechnical Testing Journal, 16(3), 323-329.
- Kalteh, A. M., Hjorth, P., & Berndtsson, R. (2008). Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. Environmental Modelling & Software, 23(7), 835-845.
- Kassambara, A. (2017). Practical Guide To Principal Component Methods in R: PCA, M (CA), FAMD, MFA, HCPC, factoextra (Vol. 2). STHDA.
- Kaufman L., and P. J. Rousseeuw (1990). Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken, NJ: John Wiley & Sons, Inc.
- Kearey, P., Brooks, M., & Hill, I. (2013). An introduction to geophysical exploration. John Wiley & Sons.
- Keller, G. V., & Frischknecht, F. C. (1966). Electrical methods in geophysical prospecting.
- Klemela, J., (2009). Smoothing of Multivariate Data. Density Estimation and Visualization. Wiley and Sons, New Jersey.
- Knackstedt, M. A., Arns, C. H., & Pinczewski, W. V. (2005). Velocity–porosity relationships: Predictive velocity model for cemented sands composed of multiple mineral phases. Geophysical Prospecting, 53(3), 349-372.
- Kohonen, T. (2014). MATLAB Implementations and Applications of the Self-organizing Map. Unigrafia Oy, Helsinki, Finland, 11-23.
- Kowallis, B. J., Jones, L. E. A., & Wang, H. F. (1984). Velocity-porosity-clay content systematics of poorly consolidated sandstones. Journal of Geophysical Research: Solid Earth, 89(B12), 10355-10364.
- Kraaijveld, M. A. (1992, January). A non-linear projection method based on Kohonen's topology preserving maps. In Pattern Recognition, 1992. Vol. II. Conference B: Pattern Recognition Methodology and Systems, Proceedings., 11th IAPR International Conference on (pp. 41-45). IEEE.
- Haeni, F. P. (1986). Application of seismic refraction methods in groundwater modeling studies in New England. Geophysics, 51(2), 236-249.
- Han, D. H., Nur, A., & Morgan, D. (1986). Effects of porosity and clay content on wave velocities in sandstones. Geophysics, 51(11), 2093-2107.
- Hayashi, K., & Takahashi, T. (2001). High resolution seismic refraction method using surface and borehole data for site characterization of rocks. International Journal of Rock Mechanics and Mining Sciences, 38(6), 807-813.
- Lees, J.M., Van Decar, J.C., (1991). Seismic tomography constrained by Bouguer gravity anomalies: applications in western Washington. Pure Appl. Geophys. 135, 31–52.
- Legchenko, A., Baltassat, J. M., Beauce, A., & Bernard, J. (2002). Nuclear magnetic resonance as a geophysical tool for hydrogeologists. Journal of Applied Geophysics, 50(1-2), 21-46.
- Lloyd, S. (1982). Least squares quantization in PCM. IEEE transactions on information theory, 28(2), 129-137.
- Linari, V., Santiago, M., Pastore, C., Azbel, K. and Poupon, M. (2003) Seismic facies analysis based on 3D multi attribute volume classification, La Palm Field, Maraibo, Venezuela. The Leading Edge, 22(1), 32-36.
- Liner, C., Li, C. F., Gersztenkorn, A., & Smythe, J. (2004, January). SPICE: A new general seismic attribute. In 2004 SEG Annual Meeting. Society of Exploration Geophysicists, Expanded Abstracts, 433–436.

- Lines, L., Schultz, A., Treitel, S., 1988. Cooperative inversion of geophysical data. Geophysics 53, 8–20.
- Lloyd, S., (1982). Least squares quantization in PCM. IEEE Trans. Inf. Theory 28, 129–137 (Originally as an unpublished Bell laboratories Technical Note (1957)).
- Lochbühler, T., Doetsch, J., Brauchler, R., & Linde, N. (2013). Structure-coupled joint inversion of geophysical and hydrological data. Geophysics, 78(3), ID1-ID14.
- Love, P. L., & Simaan, M. (1984). Segmentation of stacked seismic data by the classification of image texture. In SEG Technical Program Expanded Abstracts 1984 (pp. 480-482). Society of Exploration Geophysicists.
- Loke, M. H., & Barker, R. D. (1996). Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method 1. Geophysical prospecting, 44(1), 131-152.
- Loke, M. H. (2002). RES2DINV ver. 3.50. Rapid 2-D resistivity and IP inversion using the least square method.
- Loke, M. H. (2004). Tutorial: 2-D and 3-D electrical imaging surveys.
- MacQueen, J.B. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of 5th Berkley Symposium on Mathematical Statistics and Probability, Volume I: Statistics, pp. 281–297.
- Maillet, G. M., Rizzo, E., Revil, A., & Vella, C. (2005). High resolution electrical resistivity tomography (ERT) in a transition zone environment: application for detailed internal architecture and infilling processes study of a Rhône River paleo-channel. Marine Geophysical Researches, 26(2-4), 317-328.
- Manzi, M. S., Durrheim, R. J., Hein, K. A., & King, N. (2012). 3D edge detection seismic attributes used to map potential conduits for water and methane in deep gold mines in the Witwatersrand basin, South Africa. Geophysics, 77(5), WC133-WC147.
- Maraio, S., Bruno, P. P. G., Picotti, V., Mair, V., & Brardinoni, F. (2018). High-resolution seismic imaging of debris-flow fans, alluvial valley fills and hosting bedrock geometry in Vinschgau/Val Venosta, Eastern Italian Alps. Journal of Applied Geophysics.
- Marion, D., Nur, A., Yin, H., & Han, D. H. (1992). Compressional velocity and porosity in sand-clay mixtures. Geophysics, 57(4), 554-563.
- Martin, R., Harris, B., & Martin, R. (2013). High resolution seismic reflection for hydrogeology-where is the value?. ASEG Extended Abstracts, 2013(1), 1-4
- Martinez W. L., Martinez A. R., (2002). Computational Statistics Handbook with MATLAB. CHAPMAN &HALL/CRC.
- Martinez W. L., Martinez A. R., (2005). Exploratory Data Analysis with MATLAB. CHAPMAN &HALL/CRC.
- Martini I.P., Sagri M., (1993). Tectono sedimentary characteristics of late Miocene-Quaternary extensional basins of the Northern Apennines, Italy. Earth Science Reviews, 34, 197-133.
- Maercklin N., Bedrosian P.A., Haberland C., Ritter O., Ryberg T.& Weber, M., 2005. Characterizing a large shear-zone with seismic and magnetotelluric methods: the case of the Dead Sea Transform, Geophys. Res. Lett., 32, L15303, doi: 10.1029/2005GL022724.
- Mayer K., Scheu B., Montanaro C., Yilmaz T.I., Isaia R., Aßbichler D.,Dingwell D.B. (2016). Hydrothermal alteration of surficial rocks at Solfatara (Campi Flegrei): petrophysical properties and implications for phreatic eruption processes. Journal of Volcanology and Geothermal Research, 320, 128–143. https://doi.org/10.1016/j.jvolgeores.2016.04.020
- McNeill, J. D. (1980). Electrical conductivity of soils and rocks. Geonics Limited.
- Michot, D., Benderitter, Y., Dorigny, A., Nicoullaud, B., King, D., & Tabbagh, A. (2003). Spatial and temporal monitoring of soil water content with an irrigated corn crop cover using surface electrical resistivity tomography. Water Resources Research, 39(5).

- Moon, W. M., Kim, J. S., Lodha, G., Serzu, M., & Soonawala, N. (1993). Application of high resolution seismic reflection techniques for fracture mapping in crystalline rocks. Engineering geology, 34(3-4), 261-280.
- Moretti R, Arienzo I, Civetta L., Orsi G., Papale P. (2013). Multiple magma degassing sources at an explosive volcano. Earth and Planetary Science Letters, 367, 95–104. https://doi.org/10.1016/j.epsl.2013.02.013
- Moser, T. J. (1991). Shortest path calculation of seismic rays. Geophysics, 56(1), 59-67.
- Nicholson, C., & Simpson, D. W. (1985). Changes in Vp/Vs with depth: Implications for appropriate velocity models, improved earthquake locations, and material properties of the upper crust. Bulletin of the Seismological Society of America, 75(4), 1105-1123.
- Orsi G., De Vita S., Di Vito M.A. (1996). The restless, resurgent Phlegraean Fields nested caldera (Italy): constraints on its evolution and configuration. Journal of Volcanology and Geothermal Research,74, 179–214. https://doi.org/10.1016/S0377-0273(96)00063-7
- Orsi G., Civetta L., Del Gaudio C., De Vita S., Di Vito M.A., Isaia R., Petrazzuoli S.M., Ricciardi G.P., Ricco C. (1999). Short-term ground deformations and seismicity in the resurgent Campi Flegrei Caldera (Italy): An example of active block-resurgence in a densely populated area. Journal of Volcanology and Geothermal Research, 91, 415–451. https://doi.org/10.1016/S0377-0273(99)00050-5
- Paasche, H., Tronicke, J., Holliger, K., Green, A. G., & Maurer, H. (2006). Integration of diverse physical-property models: Subsurface zonation and petrophysical parameter estimation based on fuzzy c-means cluster analyses. Geophysics, 71(3), H33-H44.
- Paasche, H., Tronicke, J., & Dietrich, P. (2010). Automated integration of partially colocated models: Subsurface zonation using a modified fuzzy c-means cluster analysis algorithm. Geophysics, 75(3), P11-P22.
- Palacky, G.J., 1987. Clay mapping using electromagnetic methods. First Break 5, 295–306.
- Parra, J. O., Hackert, C., Bennett, M., & Collier, H. A. (2003). Permeability and porosity images based on NMR, sonic, and seismic reflectivity: Application to a carbonate aquifer. The Leading Edge, 22(11), 1102-1108.
- Pazdirek, O., & Blaha, V. (1996, June). Examples of resistivity imaging using ME-100 resistivity field acquisition system. In 58th EAGE Conference and Exhibition.
- Petrosino S, Damiano N, Cusano P., Di Vito M.A., De Vita S., Del Pezzo E. (2012) .Subsurface structure of the Solfatara volcano (Campi Flegrei caldera, Italy) as deduced from joint seismic-noise array, volcanological and morphostructural analysis. Geochemistry, Geophysics, Geosystems, 13, 1–25. https://doi.org/10.1029/2011GC004030
- Piochi M., Mormone A., Balassone G., StraussH., Troise C., De NataleG. (2015). Native sulfur, solfate and sulfides from the active Campi Flegrei volcano (southern Italy): genetic environments and degassing dynamics revealed by mineralogy and isotope geochemistry. Journal of Volcanology and Geothermal Research, 304, 180–193. https://doi.org/10.1016/j.jvolgeores.2015.08.017
- Prasad, M., Fabricius, I. L., & Olsen, C. (2005). Rock physics and statistical well log analyses in marly chalk. The Leading Edge, 24(5), 491-495.
- Pussak, M., Bauer, K., Stiller, M., & Bujakowski, W. (2014). Improved 3D seismic attribute mapping by CRS stacking instead of NMO stacking: Application to a geothermal reservoir in the Polish Basin. Journal of Applied Geophysics, 103, 186-198.
- Rawashdeh, M., & Ralescu, A. (2012, August). Crisp and fuzzy cluster validity: Generalized intra-inter silhouette index. In Fuzzy Information Processing Society (NAFIPS), 2012 Annual Meeting of the North American (pp. 1-6). IEEE.

- Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana (2004). Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain fullwaveform tomography: application to a thrust belt, Geophys. J. Int. 159, no. 3, 1032–1056, doi: 10.1111/j.1365-246X.2004.02442.x.
- Raymer, L. L., Hunt, E. R., & Gardner, J. S. (1980, January). An improved sonic transit time-to-porosity transform. In SPWLA 21st annual logging symposium. Society of Petrophysicists and Well-Log Analysts.
- Revil A., Naudet V, Nouzaret J, and Pessel M (2003) Principles of electrography applied to self-potential electrokinetic sources and hydrogeological applications. Water Resources Research 39(5), 1114.
- Revil A and Linde N (2006) Chemico-electromechanical coupling in microporous media. Journal of Colloid and Interface Sciences 302: 682–694.
- Revil, A., Barnier, G., Karaoulis, M., Sava, P., Jardani, A., & Kulessa, B. (2013). Seismoelectric coupling in unsaturated porous media: theory, petrophysics, and saturation front localization using an electroacoustic approach. Geophysical Journal International, 196(2), 867-884.
- Roden, R., Smith, T., & Sacrey, D. (2015). Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps. Interpretation, 3(4), SAE59-SAE83.
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53-65.
- Rubin, Y., & Hubbard, S. S. (Eds.). (2006). Hydrogeophysics (Vol. 50). Springer Science & Business Media.
- Ruggeri, P., Irving, J., Gloaguen, E., & Holliger, K. (2013). Regional–scale integration of multiresolution hydrological and geophysical data using a two-step Bayesian sequential simulation approach. Geophysical Journal International, 194(1), 289-303.
- Rye R.O. (2005). A review of the stable-isotope geochemistry of sulfate minerals in selected igneous environments and related hydrothermal systems selected igneous environments and related hydrothermal systems. Chemical Geology, 215, 5–36. https://doi.org/10.1016/j.chemgeo.2004.06.034
- Saccorotti, G., Petrosino S., Bianco F., Castellano M., Galluzzo D., La Rocca M., Del Pezzo E., Zaccarelli L., Cusano P. (2007). Seismicity associated with the 2004–2006 renewed ground uplift at Phlegraean Fields Caldera, Italy. Physics of the Earth and Planetary Interiors, 165, 14–24. https://doi.org/10.1016/j.pepi.2007.07.006
- Sagri, M., Martini, I.P., Pascucci, V. (2004) – Sedimentary and tectonic evolution of selected Neogene – Quaternary basins of the Apennines (Italy). In: Guerrieri L., Rischia I. & Serval L. (2004): Field Trip Guidebooks, 32nd IGC, Florence, 20-28 August 2004, Mem. Descr. Carta Geol. It., 63 (1), from PR01 to PR15, APAT, Roma, 56 pp.
- Samouëlian, A., Cousin, I., Tabbagh, A., Bruand, A., & Richard, G. (2005). Electrical resistivity survey in soil science: a review. Soil and Tillage research, 83(2), 173-193.
- Sander, J. E. (1978). The Blind Zone in Seismic Ground-Water Exploration. Groundwater, 16(6), 394-397.
- Sasaki, Y. (1992). Resolution of resistivity tomography inferred from numerical simulations. Geophysical prospecting, 40, 453-464.
- Sava, P., Revil, A., & Karaoulis, M. (2014). High definition cross-well electrical resistivity imaging using seismoelectric focusing and image-guided inversion. Geophysical Journal International, 198(2), 880-894.
- Scotellaro, C. (2006). Influenza delle argille sulle proprietà elastiche ed elettriche di mezzi eterogenei e porosi ed applicazione del Modello Elettrosismico (Doctoral dissertation, Università degli Studi di Napoli Federico II).
- Sharma, S.C. (1996). Applied Multivariate Techniques. John Wiley and Sons.

- Sheets, K. R., & Hendrickx, J. M. (1995). Noninvasive soil water content measurement using electromagnetic induction. Water resources research, 31(10), 2401-2409.
- Singha, K., & Gorelick, S. M. (2005). Saline tracer visualized with three-dimensional electrical resistivity tomography: Field-scale spatial moment analysis. Water Resources Research, 41(5).
- Stümpel, H., Kähler, S., Meissner, R., & Milkereit, B. (1984). The use of seismic shear waves and compressional waves for lithological problems of shallow sediments. Geophysical Prospecting, 32(4), 662-675.
- Šumanovac, F., & Weisser, M. (2001). Evaluation of resistivity and seismic methods for hydrogeological mapping in karst terrains. Journal of Applied Geophysics, 47(1), 13-28.
- Šumanovac, F., & Orešković, J. (2018). Exploration of buried carbonate aquifers by the inverse and forward modelling of the Controlled Source Audio-Magnetotelluric data. Journal of Applied Geophysics, 153, 47-63.
- Taner, M. T., Walls, J. D., Smith, M., Taylor, G., Carr, M. B., & Dumas, D. (2001, January). Reservoir Characterization By Calibration Of Selforganized Map Clusters. In 2001 SEG Annual Meeting. Society of Exploration Geophysicists.
- Taylor, R. W., & Fleming, A. H. (1988). Characterizing jointed systems by azimuthal resistivity surveys. Groundwater, 26(4), 464-474.
- Todesco M., Rinaldi A.P., Bonafede M. (2010). Modeling of unrest signals in heterogeneous hydrothermal systems. Journal of Geophysical Research: Solid Earth, 115, B09213. https://doi.org/10.1029/2010JB007474
- Topp, G. C., Davis, J. L., & Annan, A. P. (1980). Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. Water resources research, 16(3), 574-582.
- Tosaya, C., & Nur, A. (1982). Effects of diagenesis and clays on compressional velocities in rocks. Geophysical Research Letters, 9(1), 5-8.
- Troiano, A., Di Giuseppe, M. G., Patella, D., Troise, C., & De Natale, G. (2014). Electromagnetic outline of the Solfatara–Pisciarelli hydrothermal system, Campi Flegrei (southern Italy). Journal of Volcanology and Geothermal Research, 277, 9-21.
- Tronicke, J., Holliger, K., Barrash, W., Knoll, M.D., 2004. Multivariate analysis of cross-hole georadar velocity and attenuation tomograms for aquifer zonation. Water Resour. Res. 40, W01519. http://dx.doi.org/10.1029/2003WR002031.
- Ultsch, A. (1990). Kohonen's self organizing feature maps for exploratory data analysis. Proc. INNC90, 305-308.
- Valentino G.M., Stanzione D. (2004). Geochemical monitoring of the thermal waters of the Phlegraean Fields. Journal of Volcanology and Geothermal Research, 133, 261–289. https://doi.org/10.1016/S0377-0273(03)00402-5
- Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. IEEE Transactions on neural networks, 11(3), 586-600.
- Vereecken, H., Binley, A., Cassiani, G., Revil, A., & Titov, K. (2006). Applied hydrogeophysics. In Applied Hydrogeophysics (pp. 1-8). Springer, Dordrecht.
- Vitale S., Isaia R. (2014). Fractures and faults in volcanic rocks (Campi Flegrei, southern Italy): insight into volcano-tectonic processes. International Journal of Earth Sciences, 103, 801–819. https://doi.org/10.1007/s00531-013-0979-0
- Wehrens, R., & Buydens, L. M. (2007). Self-and super-organizing maps in R: the Kohonen package. Journal of Statistical Software, 21(5), 1-19.
- West, B. P., May, S. R., Eastwood, J. E., & Rossen, C. (2002). Interactive seismic facies classification using textural attributes and neural networks. The Leading Edge, 21(10), 1042-1049.
- White, R. E. (1991). Properties of instantaneous seismic attributes. The Leading Edge, 10(7), 26-32.

- Yang, X., Lassen, R. N., Jensen, K. H., & Looms, M. C. (2015). Monitoring CO2 migration in a shallow sand aquifer using 3D crosshole electrical resistivity tomography. International Journal of Greenhouse Gas Control, 42, 534-544.
- Yaramanci, U., Lange, G., & Hertrich, M. (2002). Aquifer characterisation using Surface NMR jointly with other geophysical techniques at the Nauen/Berlin test site. Journal of Applied Geophysics, 50(1-2), 47-65.
- Yechieli, Y., Kafri, U., Goldman, M., & Voss, C. (2001). Factors controlling the configuration of the fresh–saline water interface in the Dead Sea coastal aquifers: synthesis of TDEM surveys and numerical groundwater modeling. Hydrogeology Journal, 9(4), 367-377.
- YOSHIMI, Y., Tanaka, K., & TOKIMATSU, K. (1989). Liquefaction resistance of a partially saturated sand. Soils and foundations, 29(3), 157-162.
- Zhao, W., Forte, E., Pipan, M., & Tian, G. (2013). Ground penetrating radar (GPR) attribute analysis for archaeological prospection. Journal of Applied Geophysics, 97, 107-117.
- Zarroca, M., Bach, J., Linares, R., & Pellicer, X. M. (2011). Electrical methods (VES and ERT) for identifying, mapping and monitoring different saline domains in a coastal plain region (Alt Empordà, Northern Spain). Journal of Hydrology, 409(1-2), 407-422.
- Zelt, C. A., Azaria, A., & Levander, A. (2006). 3D seismic refraction traveltime tomography at a groundwater contamination site. Geophysics, 71(5), H67-H78.
- Zhou, Q. Y., Shimada, J., & Sato, A. (2001). Three-dimensional spatial and temporal monitoring of soil water content using electrical resistivity tomography. Water Resources Research, 37(2), 273-285.
- Zhou, B., & Dahlin, T. (2003). Properties and effects of measurement errors on 2D resistivity imaging surveying. Near Surface Geophysics, 1(3), 105-117.
- Zohdy, A. A., Eaton, G. P., & Mabey, D. R. (1974). Application of surface geophysics to ground-water investigations.
- Wallace, D. C. (1970). Thermoelastic theory of stressed crystals and higher-order elastic constants. In Solid state physics (Vol. 25, pp. 301-404). Academic Press.
- Watson, D. B., Doll, W. E., Jeffrey Gamey, T., Sheehan, J. R., & Jardine, P. M. (2005). Plume and lithologic profiling with surface resistivity and seismic tomography. Groundwater, 43(2), 169-177.
- Wehrens, R., & Buydens, L. M. (2007). Self-and super-organizing maps in R: the Kohonen package. Journal of Statistical Software, 21(5), 1-19.
- Whiteley, R. J., Hunter, J. A., Pullan, S. E., & Nutalaya, P. (1998). "Optimum offset" seismic reflection mapping of shallow aquifers near Bangkok, Thailand. Geophysics, 63(4), 1385-1394.
- Wyllie, M. R. J., Gregory, A. R., & Gardner, L. W. (1956). Elastic wave velocities in heterogeneous and porous media. Geophysics, 21(1), 41-70.